СПРАВОЧНИК по теории вероятностей и математической статистике

СПРАВОЧНИК ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

ИЗДАНИЕ ВТОРОЕ
ПЕРЕРАБОТАННОЕ И ДОПОЛНЕННОЕ

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1985

22.17
С74
УДК 519/(03)

В. С. КОРОЛЮК, Н. И. ПОРТЕНКО,
А. В. СКОРОХОД, А. Ф. ТУРБИН

Справочник по теории вероятностей и математической статистике/В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин.— М.: Наука. Главная редакция физико-математической литературы, 1985. — 640 с.

Справочник представляет собой расширенное и переработанное издание книги «Справочник по теории вероятностей и математической статистике» под редакцией В. С. Королюка, вышедшей в 1978 г. в издательстве «Наукова думка».

По широте охвата основных идей, методов и конкретных результатов современной теории вероятностей, теории случайных процессов и отчасти математической статистики «Справочник» является единственным изданием подобного рода.

Для научных работников и инженеров.

***Владимир Семенович* КОРОЛЮК, *Николай Иванович* ПОРТЕНКО, *Анатолий Владимирович* СКОРОХОД, *Анатолий Федорович* ТУРБИН**

СПРАВОЧНИК ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Редактор *Н. И. Воронина*
Художественный редактор *Т. Н. Кольченко*
Технический редактор *В. Н. Кондакова*
Корректор *Е. В. Сидоркина*

ИБ № 12662 314467

Сдано в набор 26.12.84. Подписано к печати 23.10.85. Формат 84×108 1/32. Бумага тип. № 2. Гарнитура литературная. Печать высокая. Усл. печ. л. 33,6. Усл. кр.-от. 33,6. Уч.-изд. л. 43,88. Тираж 70 000 экз. Заказ № 458. Цена 2 р. 50 к.

Ордена Трудового Красного Знамени издательство «Наука»
Главная редакция физико-математической литературы
117071 Москва В-71, Ленинский проспект, 15

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 198052 г. Ленинград, Л-52, Измайловский проспект, 29

$С \frac{1702060000-161}{053(02)-85}$ 51—85

ОГЛАВЛЕНИЕ

Предисловие ко второму изданию 8

Часть первая
ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Глава 1. **Вероятностное пространство** 11
1.1. Случайный эксперимент 11
1.2. Аксиомы и основные свойства вероятности 13
1.3. Определение вероятностного пространства 16
1.4. Случайные величины 18
1.5. Группы случайных величин 21
1.6. Математическое ожидание 27
1.7. Условные вероятности и математические ожидания . . 32

Глава 2. **Последовательности независимых событий и величин** 38
2.1. Закон нуля и единицы 38
2.2. Схема Бернулли 40
2.3. Предельные теоремы для схемы Бернулли 41
2.4. Последовательности независимых случайных величин. Закон больших чисел 43
2.5. Неравенство Колмогорова. Усиленный закон больших чисел 47
2.6. Закон повторного логарифма 50
2.7. Ряды из независимых случайных величин 52
2.8. Сходимость функций распределения 53

Глава 3. **Аналитический аппарат** 55
3.1. Производящие функции 55
3.2. Преобразование Лапласа 60
3.3. Характеристические функции 64

Глава 4. **Центральная предельная теорема** 70
4.1. Центральная предельная теорема для последовательностей независимых случайных величин 71
4.2. Центральная предельная теорема для независимых случайных векторов 77
4.3. Локальные предельные теоремы 79
4.4. Уточнение центральной предельной теоремы и асимптотические разложения 82
4.5. Большие уклонения 87

Глава 5. **Безгранично делимые распределения** 89
5.1. Суммы независимых случайных величин и их распределения 89

5.2. Определение и основные свойства безгранично делимых распределений 91
5.3. Предельные теоремы для схемы серий 96
5.4. Предельные теоремы для нарастающих сумм в **R** . . 99

Глава 6. **Основные вероятностные распределения** 105
6.1. Характеристики случайных величин 105
6.2. Дискретные распределения 108
6.3. Непрерывные распределения 116
6.4. Распределения Пирсона 133
6.5. Многомерные распределения 135
6.6. Устойчивые распределения 141
6.7. Сингулярные распределения 143

Глава 7. **Случайные блуждания** 145
7.1. Процессы восстановления 145
7.2. Классификация случайных блужданий на прямой . . 150
7.3. Функционалы на случайном блуждании 152
7.4. «Задача о разорении» для полунепрерывных случайных блужданий 155
7.5. Факторизационные тождества 156

Глава 8. **Цепи Маркова** 163
8.1. Определение цепи Маркова 163
8.2. Однородные цепи Маркова 172
8.3. Цепи Маркова с дискретным множеством состояний 192

Часть вторая

ТЕОРИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

Глава 9. **Основные понятия теории случайных процессов** . . 205
9.1. Определение случайного процесса 205
9.2. Измеримость и интегрируемость случайных процессов 212
9.3. Сепарабельность. Свойства выборочных функций . . . 215
9.4. Потоки σ-алгебр и согласованные с ними процессы . . 219
9.5. Абсолютная непрерывность мер, соответствующих случайным процессам 224

Глава 10. **$\mathscr{L}_2$-теория** 229
10.1. Пространство гильбертовых случайных величин $\mathscr{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$ 229
10.2. Стохастические меры и интегралы 233
10.3. Линейный прогноз и фильтрация гильбертовых случайных функций 238

Глава 11. **Стационарные процессы** 240
11.1. Стационарные в широком смысле случайные процессы 240
11.2. Спектральное представление корреляционных функций 245
11.3. Спектральное представление стационарных процессов 248
11.4. Аналитические свойства стационарных процессов и их траекторий 251
11.5 Эргодическая теорема 253
11.6. Линейные преобразования (фильтры) 255

11.7. Процессы с дробно-рациональными спектральными плотностями 260
11.8. Прогнозирование, интерполирование и фильтрация стационарных процессов 263
11.9. Разложение стационарного процесса 269
11.10. Решение задач линейного прогнозирования, интерполирования и фильтрации 271
11.11. Стационарные в узком смысле случайные процессы 277

Глава 12. **Случайные поля** 285

12.1. Основные определения 285
12.2. Свойства выборочных функций 288
12.3. Однородные случайные поля 291
12.4. Изотропные случайные поля 301
12.5. Обобщенные случайные поля 303

Глава 13. **Мартингалы** 311

13.1. Определения. Общие свойства 311
13.2. Мартингалы с дискретным параметром 314
13.3. Мартингалы с непрерывным временем 320
13.4. Семимартингалы и стохастические интегралы 329

Глава 14. **Марковские процессы** 339

14.1. Марковские случайные функции 339
14.2. Марковские процессы. Определение и основные свойства 344
14.3. Мультипликативные функционалы от марковских процессов 353

Глава 15. **Однородные марковские процессы** 358

15.1. Определение и основные свойства 358
15.2. Полугруппы операторов, связанные с однородными марковскими процессами 361
15.3. Характеристические операторы строго марковских процессов 366
15.4. Процессы со счетным множеством состояний 370
15.5. Функционалы от марковских процессов 378
15.6. Преобразования марковских процессов 384
15.7. Однородные диффузионные процессы в евклидовых пространствах 391
15.8. Непрерывные процессы на прямой 396
15.9. Предельное поведение вероятностей перехода эргодических марковских процессов 400

Глава 16. **Процессы с независимыми приращениями** 403

16.1. Определение и общие свойства 403
16.2. Стохастически непрерывные процессы с независимыми приращениями 406
16.3. Однородные процессы. Асимптотические свойства . . 409
16.4. Функционалы от процессов с независимыми приращениями 414
16.5. Процесс Пуассона 420
16.6. Винеровский процесс 422

Глава 17. **Ветвящиеся процессы** 427
17.1. Ветвящиеся процессы с одним типом частиц (дискретное время) 427
17.2. Ветвящиеся процессы с одим типом частиц (непрерывное время) 432
17.3. Ветвящиеся процессы с конечным числом типов частиц (дискретное время) 438
17.4. Ветвящиеся процессы с конечным числом типов частиц (непрерывное время) 444
17.5. Общие марковские ветвящиеся процессы 448

Глава 18. **Предельные теоремы для случайных процессов** . . 451
18.1. Слабая сходимость мер в метрических пространствах 451
18.2. Слабая сходимость мер в гильбертовом пространстве 454
18.3. Предельные теоремы для непрерывных случайных процессов 458
18.4. Предельные теоремы для процессов без разрывов второго рода 464

Глава 19. **Стохастические дифференциальные уравнения** . . 468
19.1. Диффузионные процессы 468
19.2. Стохастические интегралы по винеровскому процессу 472
19.3. Стохастические дифференциальные уравнения для непрерывных процессов 483
19.4. Стохастические дифференциальные уравнения для процессов с разрывами 511

Часть третья

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Глава 20. **Основные понятия и задачи математической статистики** 525
20.1. Статистическая структура 525
20.2. Статистики 527
20.3. Основные задачи математической статистики 531
20.4. Распределение выборки 534
20.5. Процедуры проверки гипотез 542

Глава 21. **Теория оценивания параметров** 550
21.1. Задача оценивания и свойства оценок 550
21.2. Методы построения оценок 556
21.3. Доверительные области 559

Глава 22. **Оценки параметров некоторых распределений** . . . 562
22.1. Оценки параметров нормального распределения . . . 562
22.2. Оценки параметров биномиального и пуассоновского распределений 565
22.3. Оценки параметров равномерного распределения и Г-распределения 567

Глава 23. **Метод наименьших квадратов** 570
23.1. Линейные модели регрессии 570
23.2. Свойства МНК-оценок 572
23.3. Оценка параметров линейной регрессии 577

Глава 24. **Статистика случайных процессов** 581
24.1. Различение гипотез 581
24.2. Различение гипотез для процессов с независимыми приращениями 584
24.3. Различение гипотез для диффузионных процессов . . 591
24.4. Различение гипотез о среднем значении гауссовского процесса 594
24.5. Различение гипотез о корреляционной функции гауссовского процесса 598
24.6. Оценки параметров распределений для случайных процессов 605

Глава 25. **Статистика стационарных в широком смысле случайных процессов** 610
25.1. Свойства статистических оценок характеристик стационарных процессов 610
25.2. Оценки неизвестного среднего 611
25.3. Оценки параметров регрессии 617
25.4. Оценки спектральной плотности и спектральной функции стационарных последовательностей 621
25.5. Оценки параметров спектральной плотности 626

Список литературы 628

Предметный указатель 633

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Настоящее второе издание справочника в основном содержит тот же материал, что и предыдущее, однако некоторые главы существенно переработаны с учетом современного состояния теории вероятностей.

Очевидно, что ни для какой активно развивающейся науки невозможно дать полное ее изложение даже в справочной форме, поэтому справочник содержит систематическое изложение основных положений теории вероятностей, теории случайных процессов и математической статистики, а также описание методов и идей, уже четко оформившихся и получивших широкое применение как в теоретических, так и прикладных исследованиях.

Среди многочисленной и разнообразной литературы, посвященной теории вероятностей и ее приложениям, справочная литература выполняет свою функцию, представляя в сжатой форме специалистам по теории вероятностей материал из смежных областей, а специалистам, работающим в области приложений, в обозримом виде классификацию основных направлений и методов теории вероятностей, а также возможность ознакомиться с основными фактами избранного направления без доказательств и обоснований. Для более глубокого изучения интересующего вопроса нужно использовать приведенную в справочнике литературу.

Теория вероятностей использует довольно широко аппарат других разделов математики, информацию о которых следует искать в справочной литературе по соответствующим разделам. Что касается теоретико-вероятностных сведений, то они все содержатся в справочнике, при этом отдельные главы справочника написаны так, чтобы ими можно было пользоваться практически независимо (хотя в отдельных главах могут быть ссылки на любую из остальных).

Справочник содержит большой фактический материал и может быть полезен как читателям, желающим ознакомиться с фактами с целью их применения, но не интересующимся математическими доказательствами, так и специалистам в области теории вероятностей для справок в их научной работе.

Справочник не содержит материала по приложениям (теория массового обслуживания, теория надежности, теория информации, стохастическая аппроксимация, управляемые процессы), для освещения этих вопросов нужен специальный справочник. Тем не менее, в качестве иллюстраций основных методов теории вероятностей прикладные задачи широко используются.

Справочник состоит из трех частей.

Часть первая — теория вероятностей (гл. 1—8) — содержит основные определения вероятностного пространства, случайной величины, математического ожидания, условных вероятностей и математических ожиданий, независимости. Здесь рассматриваются последовательности независимых событий и величин, а также связанные с ними законы больших чисел, центральная предельная теорема, безгранично делимые распределения и предельные теоремы для схемы серий. Описаны аналитические методы, используемые при исследовании сумм независимых случайных величин (гл. 3). В гл. 6 приведены сведения обо всех основных распределениях, используемых в теории вероятностей и ее приложениях. Глава 8 содержит большой материал (включая новейшие результаты), широко используемый в приложениях, по цепям Маркова. Здесь, в частности, приводятся самые общие результаты по эргодической проблеме и центральной предельной теореме.

Первые две главы этой части составляют необходимую основу для понимания всех разделов теории вероятностей.

Остальной материал этой части составляют специальные вопросы теории и ее приложений (центральная предельная теорема с уточнениями, случайные блуждания, цепи Маркова).

Наибольшая вторая часть справочника посвящена теории случайных процессов. В главах 9, 10, 12, 18 изложены общие понятия, методы и факты теории случайных процессов (и функций), остальные главы посвящены конкретным классам случайных процессов: стационарные процессы, мартингалы, марковские процессы, однородные марковские процессы, процессы с независимыми приращениями, ветвящиеся процессы, стохастические дифференциальные уравнения.

Во второй части содержится весь фактический материал, который обычно излагается в книгах по теории случайных процессов. Наибольший интерес для приложений традиционно представляют стационарные процессы, однородные марковские процессы, процессы с независимыми приращениями и ветвящиеся процессы. Материал по общей теории случайных процессов специалистам по приложениям может быть полезен для большего понимания конкретных фактов. Активно развивающиеся в настоящее время мартингальные методы, стохастические дифференциальные уравнения и предельные теоремы несомненно имеют большие перспективы в приложениях. Главы, посвященные этим разделам, во втором издании существенно расширены.

В третьей части излагаются основные понятия математической статистики, а также приведены важнейшие факты из теории проверки статистических гипотез, теории оценивания, изложен метод наименьших квадратов. Статистика случайных процессов в настоящее время находится в процессе интенсивного развития и трудно выделить прочно установившуюся структуру. Исключением является линейная теория статистики стационарных процессов, которая довольно полно отражена в справочнике. Сведения же по общим вопросам статистики случайных процессов не претендуют на полноту, а имеют цель иллюстрировать основные постановки задач и методы их решения.

Главы разбиты на пункты, нумерация которых проведена по главам, пункты — на подпункты, все они имеют названия. Это позволит читателям непосредственно знакомиться с интересующим их вопросом. Для удобства читателей справочник снабжен предметным указателем. Читатели, интересующиеся более подробным изложением вопроса или доказательствами приведенных утверждений, могут воспользоваться литературой, указанной в конце глав.

Работу над справочником авторы разделили следующим образом: В. С. Королюк — гл. 3, 7, 10, 18, 20; Н. И. Портенко — гл. 2, 8, 14, 15, 19, 21, 22; А. В. Скороход — гл. 1, 5, 9, 12, 13, 16, 18, 24; А. Ф. Турбин — гл. 4, 6, 11, 17, 20, 23, 25.

Авторы допускают, что изложение как по форме, так и по содержанию не лишено недостатков. Все замечания и предложения будут приняты с благодарностью.

При подготовке рукописи к печати большую помощь оказали Н. Ф. Рябова и Л. В. Лобанова, которым авторы выражают глубокую признательность.

Авторы

Часть первая

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Глава 1. ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

1.1. Случайный эксперимент

1.1.1. Определение. Одним из наиболее важных понятий в теории вероятностей является понятие *эксперимента*. Эксперимент состоит в том, что производится испытание при выполнении некоторого комплекса условий, которые либо создаются искусственно, либо осуществляются независимо от воли экспериментатора. Эксперимент задан, если определены его условия и указаны события, наступление или ненаступление которых следует наблюдать.

Эксперименты можно разделить на два класса. В одном из них результаты экспериментов заранее предсказуемы на основании естественнонаучных законов. Эти эксперименты носят название *детерминированных*. В другом классе экспериментов при одних и тех же условиях возможно наступление исключающих друг друга событий. Теоретическое изучение таких экспериментов и составляет предмет теории вероятностей; они носят название *случайных* (*стохастических*) или *вероятностных* экспериментов.

Примеры случайных экспериментов.

1. Изделия выпускаются партиями по n штук. Проверка качества изделий приводит к их разрушению. Поэтому для проверки партии на качество отбирают m изделий ($m < n$). Эксперимент заключается в выборе m изделий из партии и их проверке. Результат эксперимента — число обнаруженных дефектных изделий.

2. Розыгрыш лотереи можно рассматривать как случайный эксперимент, результатом которого является выпадение выигрышей на определенные лотерейные билеты.

3. В биологическом опыте самоопыления растение, полученное перекрестным опылением двух сортов, наследует по каждому признаку гены обоих родителей. Нельзя сказать заранее, как эти гены скомбинированы в том или ином семени, полученном в результате самоопыления: по каждому гену (если он различен в материнском и отцовском сорте) возможны три комбинации: доминантный — доминантный, доминантный — рецессивный, рецессивный — рецессивный. Следовательно, такой опыт можно рассматривать как случайный эксперимент.

4. Взвешенная в жидкости частица движется в результате столкновений с молекулами жидкости, находящимися в хаотическом тепловом движении. Опыт, в котором наблюдается движение такой частицы, можно рассматривать как случайный эксперимент, результатом которого является траектория движения броуновской частицы.

1.1.2. Алгебра событий. Рассмотрим множество $\mathfrak{M}$ событий, которые можно наблюдать в некотором стохастическом эксперименте.

Выделим прежде всего два специальных события — *достоверное событие* U, которое обязательно происходит в эксперименте, и *невозможное событие* V, которое не может произойти в эксперименте.

Для каждого события A из $\mathfrak{M}$ введем *противоположное событие* $\bar{A}$, которое состоит в том, что A не произошло. Событие $A+B$ $(A\cup B)$, заключающееся в том, что из двух событий A и B происходит по крайней мере одно, называется *суммой* (или *объединением*) *событий* A и B. Событие AB $(A\cap B)$, заключающееся в том, что происходят одновременно A и B, называется *произведением* (или *пересечением*) *событий* A и B. Событие $A\setminus B$ называется *разностью событий* A и B; оно состоит в том, что происходит A и не происходит B.

Два события A и B *несовместимы*, если $A\cap B$ — невозможное событие.

События $E_1, E_2, \ldots, E_n$ образуют *полную группу событий*, если они попарно несовместимы и $E_1\cup E_2\cup\ldots\cup E_n = \bigcup\limits_k E_k = U$, т. е. из этих событий происходит одно и только одно.

Непустое множество событий $\mathfrak{A}$, которое удовлетворяет условиям:

1) если $A\in\mathfrak{A}$, то $\bar{A}\in\mathfrak{A}$;

2) если $A, B\in\mathfrak{A}$, то $A\cup B\in\mathfrak{A}$,

называется *алгеброй событий*.

1.1.3. Элементарные события. Говорят, что *событие A влечет событие B* $(A\subset B)$, если событие B наступает всегда, когда наступает A. Событие E называется *элементарным*, если для всякого события A случайного эксперимента оно влечет либо A, либо $\bar{A}$. Случайный эксперимент называется *конечным*, если имеется полная группа элементарных событий. В теории вероятностей рассматриваются лишь такие случайные эксперименты, в которых каждое событие является суммой всех элементарных событий, влекущих это событие. Такой случайный эксперимент описывается множеством элементарных событий Ω (его элементы обозначают буквой ω с различными индексами, например, ω', ω'', ω^1, ω_1 и т. п.) и некоторым классом его подмножеств $\mathfrak{A}$, называемых *событиями*. Этот класс подмножеств должен удовлетворять следующим условиям:

1) $\Omega\in\mathfrak{A}$ (Ω — достоверное событие; оно происходит, какое бы элементарное событие ни произошло);

2) $\mathfrak{A}$ содержит пустое подмножество $\varnothing$, которое интерпретируется как невозможное событие;

3) если $A\in\mathfrak{A}$, то $\bar{A}\in\mathfrak{A}$, где $\bar{A}=\Omega\setminus A$;

4) если $A\in\mathfrak{A}$ и $B\in\mathfrak{A}$, то $A\cup B\in\mathfrak{A}$ и $A\cap B\in\mathfrak{A}$.

Класс подмножеств $\mathfrak{A}$, удовлетворяющий условиям 1) — 4), называется *алгеброй множеств*.

В том случае, когда Ω конечно, $\mathfrak{A}$ совпадает с классом всех подмножеств Ω.

Важным примером случайного эксперимента является эксперимент, в котором измеряется некоторая величина ξ. В качестве элементарных событий здесь можно взять события вида $(\xi = x)$, где x — некоторое фиксированное значение. Поэтому множество элементарных событий естественно отождествить с множеством точек на прямой. Если априори известно, что ξ может принимать лишь значения из некоторого множества M, то это множество и следует рассматривать как множество элементарных событий. В процессе измерения естественно предполагать возможность наблюдения события

$\{a \leqslant \xi < b\}$, где $a < b$ — произвольные числа. Всевозможные конечные суммы таких полуинтервалов (a и b могут принимать и бесконечные значения) можно рассматривать как алгебру событий, связанных с экспериментом.

1.2. Аксиомы и основные свойства вероятности

1.2.1. Частоты событий. Одной из существенных особенностей случайных экспериментов является возможность повторять их сколь угодно большое (в принципе неограниченное) число раз. Если Ω — множество элементарных событий эксперимента, то осуществление эксперимента означает выбор некоторой точки $\omega \in \Omega$, а повторение этого эксперимента n раз означает выбор последовательности точек $\omega_1, \ldots, \omega_n$ в Ω. Пусть $\mathfrak{A}$ — алгебра событий, наблюдаемых в эксперименте, $A \in \mathfrak{A}$. Обозначим число появлений события A в n экспериментах через $k_n(A)$ (если $\omega_i \in A$, то A произошло в i-м эксперименте). Величина $\nu_n(A) = \frac{1}{n} k_n(A)$ называется *частотой* появления события A в n экспериментах. Она до некоторой степени характеризует объективную связь между условиями эксперимента и событием A, указывая, как часто эти условия вызывают событие A. Отметим, что $\nu_n(A)$ меняется как с n, так и с изменением серии экспериментов.

Отметим основные свойства частот:

1) если U — достоверное событие, то $\nu_n(U) = 1$;
2) если V — невозможное событие, то $\nu_n(V) = 0$;
3) для всякого $A \in \mathfrak{A}$ имеет место неравенство $0 \leqslant \nu_n(A) \leqslant 1$;
4) если $A \subset B$, то $\nu_n(A) \leqslant \nu_n(B)$;
5) если A и B несовместимы, то $\nu_n(A+B) = \nu_n(A) + \nu_n(B)$;
6) если $A_1, A_2, \ldots, A_k$ попарно несовместимы, то

$$\nu_n\left(\sum_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \nu_n(A_i);$$

7) для всех $A \in \mathfrak{A}$ справедливо равенство

$$\nu_n(\bar{A}) = 1 - \nu_n(A).$$

1.2.2. Аксиомы вероятности. Важным, экспериментально установленным фактом является свойство устойчивости частот. При увеличении числа экспериментов частоты событий колеблются около некоторых чисел, не зависящих ни от числа, ни от серии экспериментов, причем частоты неограниченно приближаются к этим числам, когда $n \to \infty$. Эти числа естественно связать с каждым событием, происходящим в случайном эксперименте. Они называются *вероятностями* и определяются чисто аксиоматически. В теории вероятностей существование вероятностей постулируется и свойства их определяются *аксиомами вероятности*, которые приводятся ниже.

1) Каждому событию $A \in \mathfrak{A}$ отвечает число $\mathsf{P}(A)$, принимающее значение из $[0, 1]$ и называемое вероятностью A.

2) Если A и B — несовместимые события, то $\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$.

3) $\mathsf{P}(U) = 1$, где U — достоверное событие.

Если понимать вероятность как предел частоты, то естественность перечисленных аксиом следует из свойств 1), 3), 5) частот.

Из этих аксиом вытекают следующие свойства.

4) Если V — невозможное событие, то $\mathsf{P}(V)=0$.

5) $\mathsf{P}(\overline{A})=1-\mathsf{P}(A)$.

6) При $A\subset B$ справедливо неравенство $\mathsf{P}(A)\leqslant\mathsf{P}(B)$.

7) Если $A_1, A_2, \ldots, A_k$ — попарно непересекающиеся события, то

$$\mathsf{P}\left(\sum_{i=1}^{k}A_i\right)=\sum_{i=1}^{k}\mathsf{P}(A_i).$$

8) Для любых двух событий A и B

$$\mathsf{P}(A+B)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(A\cap B).$$

9) Для любых событий $A_1, A_2, \ldots, A_k$

$$\mathsf{P}\left(\sum_{i=1}^{k}A_i\right)\leqslant\sum_{i=1}^{k}\mathsf{P}(A_i).$$

10) Пусть $A_1, A_2, \ldots, A_n$ — некоторые события и $A_{i_1i_2\ldots i_k}=A_{i_1}A_{i_2}\ldots A_{i_k}$. Тогда

$$\mathsf{P}(A_1+A_2+\ldots+A_n)=\sum_{i=1}^{n}\mathsf{P}(A_i)-\sum_{i<j}\mathsf{P}(A_{ij})+\ldots$$
$$\ldots+(-1)^{k-1}\sum_{i_1<\ldots<i_k}\mathsf{P}\left(A_{i_1\ldots i_k}\right)+\ldots+(-1)^{n-1}\mathsf{P}\left(A_{1\ldots n}\right).$$

1.2.3. Классическое определение вероятности. Предположим, что в эксперименте имеется полная группа элементарных событий: $E_1, E_2, \ldots, E_n$. Тогда каждое событие из $\mathfrak{A}$ имеет вид

$$A=\sum_{k=1}^{m}E_{i_k}, \tag{2.1}$$

где $(i_1, i_2, \ldots, i_m)$ — некоторое подмножество множества $(1, 2, \ldots, n)$. Следовательно, по свойству 7)

$$\mathsf{P}(A)=\sum_{k=1}^{m}\mathsf{P}\left(E_{i_k}\right)=\sum_{E_i\subset A}\mathsf{P}(E_i).$$

Таким образом, в случае конечного эксперимента вероятность любого события определяется вероятностями элементарных событий (исходов).

Для многих конечных экспериментов из соображений симметрии можно априори установить, что элементарные события имеют одинаковую вероятность. Тогда вероятность каждого элементарного события равна $1/n$ (n — число исходов), а вероятность события A вида (2.1) равна m/n. Если элементарные события имеют одинаковую вероятность, то они называются *равновозможными исходами*, а те из них, которые влекут A, — *благоприятствующими исходами*. Следовательно, в этом случае $\mathsf{P}(A)$ равно отношению числа благоприятствующих исходов к числу всех равновозможных исходов.

Приведенное определение вероятности называется *классическим.*

При решении задач на классическое определение вероятности следует вычислить число всех равновозможных исходов в эксперименте, а затем — число благоприятствующих исходов. Обычно это можно сделать комбинаторными методами.

Пример. m шаров раскладываются по n ящикам $(m > n)$. Все расположения равновероятны. Какова вероятность того, что не найдется ни одного пустого ящика?

Занумеруем ящики, и пусть m_i — число шаров в ящике с номером i. В качестве множества элементарных событий возьмем группы по n чисел $(m_1, m_2, \ldots, m_n)$, где $m_i \geqslant 0$ и $\sum m_i = m$. Число элементарных событий можем определить так: поставим каждому событию в соответствие последовательность из нулей и единиц по такому правилу:

$$\underbrace{0\ldots0}_{m_1}\,1\,\underbrace{0\ldots0}_{m_2}\,1\ldots\,1\,\underbrace{0\ldots0}_{m_n}.$$

В этой последовательности $n-1$ единиц и m нулей. Каждой такой последовательности из нулей и единиц соответствует элементарное событие $(m_1, m_2, \ldots, m_n)$, где m_1 — число нулей до первой единицы, m_2 — число нулей между первой и второй единицами и т. д. Количество указанных последовательностей, очевидно, равно C_{m+n-1}^{n-1}. Чтобы определить число благоприятствующих исходов, нужно сосчитать число последовательностей, для которых $m_i \geqslant 1$. Но оно совпадает с числом последовательностей $(m_1', m_2', \ldots, m_n')$, для которых $m_i' \geqslant 0$ и $\sum m_i' = m - n$ $(m_i' = m_i - 1)$. Значит, число благоприятствующих исходов равно C_{m-1}^{n-1}. Искомая вероятность имеет вид

$$p = \frac{C_{m-1}^{n-1}}{C_{m+n-1}^{n-1}} = \frac{(m-1)!\,(n-1)!\,m!}{(n-1)!\,(m-n)!\,(m+n-1)!} = \frac{m!\,(m-1)!}{(m-n)!\,(m+n-1)!}.$$

1.2.4. Геометрические вероятности. Это вероятности в экспериментах с бесконечным числом исходов, которые интерпретируются как выбор наудачу точки из некоторого множества в $\mathbf{R}^m$. Предполагается, что множество имеет некоторую геометрическую форму. Событием назовем следующее: выбранная точка принадлежит заданной части фигуры. Вероятность такого события определим как отношение евклидова объема (площади, длины) части фигуры к объему (площади, длине) всей фигуры.

Пример (задача о встрече). Два лица договариваются о встрече на заданном промежутке времени l. Лицо, пришедшее первым, ожидает в течение времени $a < l$, затем уходит. Какова вероятность встречи?

В качестве множества элементарных событий рассмотрим квадрат, состоящий из точек (x, y), $0 \leqslant x \leqslant l$, $0 \leqslant y \leqslant l$, где x и y — время прихода первого и второго лица. Благоприятствующие исходы образуют точки, для которых $|x-y|<a$, т. е. точки квадрата между прямыми $y=x-a$, $y=x+a$. Легко вычислить, что площадь этой фигуры равна $l^2-(l-a)^2$, площадь квадрата составляет l^2, искомая вероятность равна

$$p=1-(l-a)^2/l^2.$$

1.3. Определение вероятностного пространства

1.3.1. σ-алгебра событий. В тех случайных экспериментах, в которых алгебра событий содержит бесконечное множество событий, приходится рассматривать и бесконечные последовательности событий, и операции над ними. Простейшими среди этих операций являются объединение и пересечение бесконечной последовательности событий. Если алгебра событий такова, что с каждой бесконечной последовательностью событий A_k она содержит и события $\bigcap\limits_k A_k$, $\bigcup\limits_k A_k$, то такая алгебра называется *σ-алгеброй*. Событие $\bigcap\limits_k A_k$ состоит в том, что происходят все события A_k одновременно, а событие $\bigcup\limits_k A_k$ — в том, что из последовательности событий A_k происходит по крайней мере одно.

Последовательность A_k называется *монотонно убывающей*, если $A_k\supset A_{k+1}$ для всех k, и *монотонно возрастающей*, если $A_k\subset A_{k+1}$. Событие $\bigcap\limits_k A_k$ называется *пределом убывающей последовательности*, а событие $\bigcup\limits_k A_k$ — *пределом возрастающей последовательности событий*. Предел монотонной (т. е. монотонно возрастающей или монотонно убывающей) последовательности A_k обозначается $\lim A_k$.

Алгебра событий $\mathfrak{A}$ будет σ-алгеброй, если со всякой монотонной последовательностью она содержит и ее предел.

При построении σ-алгебр множеств широко используется операция σ-замыкания алгебры множеств. σ-замыканием алгебры $\mathfrak{A}_0$ называется *наименьшая* σ-алгебра $\mathfrak{A}$, содержащая $\mathfrak{A}_0$; она обозначается $\sigma(\mathfrak{A}_0)$. Ее еще называют *σ-алгеброй, порожденной алгеброй* $\mathfrak{A}_0$.

Класс множеств μ называется *монотонным*, если с каждой монотонной последовательностью множеств он содержит и ее предел. При рассмотрении σ-алгебр часто полезна следующая теорема о монотонном классе.

Теорема. $\sigma(\mathfrak{A}_0)$ *совпадает с наименьшим монотонным классом, содержащим* $\mathfrak{A}_0$.

Если вероятность определена на σ-алгебре, то предполагается, что она удовлетворяет еще одной аксиоме, обобщающей аксиому вероятности 2) (п. 1.2.2) и называемой *расширенной аксиомой сложения*:

2′) Если A_k — последовательность попарно несовместимых событий, то

$$\mathsf{P}\left(\bigcup_k A_k\right)=\sum_k \mathsf{P}(A_k).$$

Эта аксиома эквивалентна следующей *аксиоме непрерывности*.

3) Для всякой монотонной последовательности A_k

$$\mathsf{P}(\lim A_k) = \lim \mathsf{P}(A_k).$$

1.3.2. Вероятностное пространство. *Вероятностным пространством* (*полем вероятностей*) $\{\Omega, \mathfrak{A}, \mathsf{P}\}$ называется совокупность трех объектов — пространства элементарных событий Ω, σ-алгебры $\mathfrak{A}$ подмножеств пространства Ω — σ-алгебры событий, вероятностной меры $\mathsf{P}(A)$, определенной для $A \in \mathfrak{A}$, для которой $\mathsf{P}(\Omega) = 1$.

Мерой на σ-алгебре подмножеств $\mathfrak{A}$ называется неотрицательная счетно-аддитивная функция $\mathsf{P}(A)$ множества, т. е. такая функция, для которой

$$\mathsf{P}\Big(\bigcup_k A_k\Big) = \sum_k \mathsf{P}(A_k)$$

для всякой последовательности попарно не пересекающихся множеств A_k из $\mathfrak{A}$.

Если $\mathsf{P}(\Omega) = 1$, то мера называется *нормированной.*

Измеримым пространством $\{\Omega, \mathfrak{A}\}$ называется пара объектов — некоторое множество Ω и некоторая σ-алгебра его подмножеств $\mathfrak{A}$. Таким образом, вероятностное пространство — это измеримое пространство с нормированной мерой на нем.

Если Ω содержит не более счетного числа элементов и $\mathfrak{A}$ есть множество всех подмножеств Ω, то вероятность полностью определяется своими значениями на элементарных событиях. Пусть $\Omega = \{\omega_1, \omega_2, \ldots\}$, $\mathsf{P}(\{\omega_k\}) = p_k$, $\{\omega_k\}$ — одноточечное множество, содержащее ω_k. Тогда

$$\mathsf{P}(A) = \sum_k p_k \chi_A(\omega_k),$$

где $\chi_A(\omega) = 1$ при $\omega \in A$, $\chi_A(\omega) = 0$ при $\omega \notin A$. Функция $\chi_A(\omega)$ называется *индикатором* события A. Вероятностные пространства описанного вида называются *дискретными.*

Рассмотрим вероятностное пространство, для которого Ω совпадает с m-мерным евклидовым пространством R^m. Такое пространство исходов естественно рассматривать в тех экспериментах, в которых наблюдаются значения m вещественных величин. Будем обозначать координаты точки $\bar{x} \in \mathrm{R}^m$ через $(x^1, x^2, \ldots, x^m)$. В качестве $\mathfrak{A}$ возьмем σ-алгебру, содержащую множества точек вида

$$\{\bar{x}\colon a_1 \leqslant x^1 < b_1, \ldots, a_m \leqslant x^m < b_m\}, \tag{3.1}$$

где $-\infty \leqslant a_i < b_i \leqslant +\infty$ — вещественные числа. Такие множества называются *полуоткрытыми справа параллелепипедами.* Конечные суммы полуоткрытых справа параллелепипедов образуют алгебру $\mathfrak{A}_0$ в R^m. Наименьшая σ-алгебра $\mathfrak{A}$, содержащая алгебру $\mathfrak{A}_0$, совпадает с наименьшей σ-алгеброй множеств, содержащих все открытые и замкнутые множества из R^m. Эта σ-алгебра называется *борелевской* σ-алгеброй, а множества из $\mathfrak{A}$ — *борелевскими.*

Всякое множество из $\mathfrak{A}$ получается с помощью операции предельного перехода, примененного не более счетного числа раз к множествам из $\mathfrak{A}_0$. Поэтому для задания вероятности на $\mathfrak{A}$ (учитывая аксиому непрерывности) достаточно задать ее на $\mathfrak{A}_0$. Поскольку множества из $\mathfrak{A}_0$ представимы в виде суммы попарно непересекающихся полуоткрытых параллелепипедов, то достаточно определить меру на множествах вида (3.1).

1.3.3. Функции распределения. Пусть

$$G(b_1, \ldots, b_m) = \mathsf{P}\{\bar{x}: -\infty < x^1 < b_1, \ldots, -\infty < x^m < b_m\}. \quad (3.2)$$

Обозначим через $\Delta^{(k)}_{[a, b)} G(x^1, \ldots, x^m) = G(x^1, \ldots, x^{k-1}, b, x^{k+1}, \ldots \ldots, x^m) - G(x^1, \ldots, x^{k-1}, a, x^{k+1}, \ldots, x^m)$ приращение функции $G(x^1, \ldots, x^m)$ по k-му аргументу на полуинтервале $[a, b)$. Тогда справедлива формула

$$\mathsf{P}(\{\bar{x}: a_1 \leqslant x^1 < b_1, \ldots, a_m \leqslant x^m < b_m\}) = \\ = \Delta^{(1)}_{[a_1, b_1)} \ldots \Delta^{(m)}_{[a_m, b_m)} G(x^1, \ldots, x^m). \quad (3.3)$$

Таким образом, всякая мера на измеримом пространстве $\{\mathbf{R}^m, \mathfrak{A}\}$ однозначно определяется функцией $G(x^1, \ldots, x^m)$ вида (3.2). Чтобы соответствующая мера была нормированной, необходимо и достаточно, чтобы выполнялось условие

1) $\lim\limits_{x^1 \to \infty, \ldots, x^m \to \infty} G(x^1, \ldots, x^m) = 1.$

Укажем еще некоторые условия, которым необходимо удовлетворяет G. Из аксиомы непрерывности следует, что

2) $\lim\limits_{x^k \to -\infty} G(x^1, \ldots, x^m) = 0$ для всех $k = 1, \ldots, m.$

3) $\lim\limits_{x^1 \uparrow b_1, \ldots, x^m \uparrow b_m} G(x^1, \ldots, x^m) = G(b_1, \ldots, b_m)$, каковы бы ни были $b_1, \ldots, b_m$, т. е. функция $G(x^1, \ldots, x^m)$ непрерывна по совокупности аргументов слева.

Из (3.3) следует

4) $\Delta^{(1)}_{[a_1, b_1)} \ldots \Delta^{(m)}_{[a_m, b_m)} G(x^1, \ldots, x^m) \geqslant 0.$

Функция $G(x^1, \ldots, x^m)$, удовлетворяющая условиям 1) — 4), называется *m-мерной функцией распределения. Одномерной функцией* распределения ($m = 1$) называется неубывающая непрерывная слева функция $F(x)$, определенная на $\mathbf{R}$ и удовлетворяющая условиям

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1.$$

Всякой m-мерной функции распределения отвечает единственная вероятностная мера на $\{\mathbf{R}^m, \mathfrak{A}\}$.

1.4. Случайные величины

1.4.1. Определение случайной величины. Случайные величины — это величины, измеряемые в случайных экспериментах. Случайная величина полностью определена, если известен исход эксперимента ω. Таким образом, случайная величина ξ на вероятностном пространстве $\{\Omega, \mathfrak{A}, \mathsf{P}\}$, описывающем данный случайный эксперимент, есть некоторая функция элементарного события $\xi(\omega)$. Тот факт, что мы можем измерять эту величину в нашем эксперименте, означает, что возможно наблюдать событие: значение величины ξ принадлежит данному интервалу Δ, каков бы ни был этот интервал. Значит,

$$\{\omega: \xi(\omega) \in \Delta\} \in \mathfrak{A}. \tag{4.1}$$

Функции $\xi(\omega)$, удовлетворяющие для всех интервалов Δ условию (4.1), называются *измеримыми* относительно σ-алгебры $\mathfrak{A}$ или $\mathfrak{A}$-*измеримыми*. Для измеримой относительно $\mathfrak{A}$ функции $\xi(\omega)$ соотношение (4.1) выполнено для всякого борелевского множества $\Delta \subset \mathbf{R}$.

Случайной величиной на вероятностном пространстве $\{\Omega, \mathfrak{A}, \mathsf{P}\}$ называется всякая $\mathfrak{A}$-измеримая функция $\xi(\omega)$, определенная на Ω. Случайные величины часто обозначают ξ вместо $\xi(\omega)$, не указывая на зависимость от элементарного события.

Простейшим примером случайной величины является величина $\chi_A(\omega)$ — индикатор события A: $\chi_A(\omega) = 1$, если $\omega \in A$; $\chi_A(\omega) = 0$, если $\omega \notin A$.

Другим примером случайной величины служит дискретная случайная величина, принимающая не более чем счетное множество различных значений $\{x_1, x_2, \ldots\}$. Очевидно, что события $\{\xi(\omega) = x_i\} = A_i$ попарно несовместимы и $\bigcup_i A_i = \Omega$. Пусть

$$\mathsf{P}(A_i) = \mathsf{P}\{\xi(\omega) = x_i\} = \mathsf{P}\{\xi = x_i\} = p_i.$$

Набор вероятностей $\{p_i\}$ и чисел $\{x_i\}$ называется *распределением дискретной величины* ξ. Оно определяет вероятность попадания величины ξ в любое множество Λ на прямой:

$$\mathsf{P}\{\xi \in \Lambda\} = \sum_{x_i \in \Lambda} p_i.$$

1.4.2. Распределение случайной величины. *Распределением произвольной* (*случайной*) *величины* ξ называется мера

$$\mathsf{P}_\xi(\Lambda) = \mathsf{P}(\{\omega: \xi(\omega) \in \Lambda\}), \tag{4.2}$$

заданная на σ-алгебре борелевских множеств из $\mathbf{R}$. Из (4.1) следует, что для всех борелевских множеств $\{\omega: \xi(\omega) \in \Lambda\} \in \mathfrak{A}$, и поэтому правая часть (4.2) определена.

Как следует из результатов § 1.3, для задания распределения величины ξ достаточно задать функцию

$$F_\xi(x) = \mathsf{P}_\xi((-\infty, x)) = \mathsf{P}\{\xi < x\},$$

которая называется *функцией распределения* величины ξ и является *одномерной функцией распределения*.

Если ξ — дискретная величина, для которой $\mathsf{P}\{\xi = x_i\} = p_i$, то

$$F_\xi(x) = \sum_{x_i < x} p_i = \sum_i p_i \varepsilon(x - x_i),$$

где $\varepsilon(x) = 1$, если $x > 0$; $\varepsilon(x) = 0$, если $x \leqslant 0$. Обозначим через $F_\xi(x+0)$ предел справа $F_\xi(x)$ в точке x. Величина скачка функции распределения $F_\xi(x+0) - F_\xi(x)$ совпадает с вероятностью

$\mathsf{P}\{\xi = x\}$; если $\mathsf{P}\{\xi = x\} > 0$, то x называют *атомом* распределения случайной величины ξ. Говорят, что ξ имеет *непрерывное распределение*, если $F_\xi(x)$ — непрерывная функция. В этом случае любое фиксированное значение ξ может принимать лишь с вероятностью 0. Величина ξ имеет *абсолютно непрерывное распределение*, если существует такая функция $f_\xi(x)$, что

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)\,dt. \tag{4.3}$$

Функция $f_\xi(x)$, удовлетворяющая соотношению (4.3), называется *плотностью распределения* величины ξ. Если ξ имеет плотность распределения, то ее распределение выражается формулой

$$\mathsf{P}_\xi(\Lambda) = \int_{\Lambda} f_\xi(t)\,dt \tag{4.4}$$

(интеграл в (4.4) понимается как интеграл Лебега). В частности,

$$\mathsf{P}_\xi((a, b)) = \int_a^b f_\xi(t)\,dt = F_\xi(b) - F_\xi(a).$$

Плотность распределения удовлетворяет следующим двум очевидным условиям:

а) $f_\xi(t) \geqslant 0$ для почти всех t;

б) $$\int f_\xi(t)\,dt = \int_{-\infty}^{\infty} f_\xi(t)\,dt = 1.$$

Любая измеримая по Лебегу функция $f_\xi(t)$, удовлетворяющая этим двум условиям, может выступать в качестве плотности некоторой случайной величины.

Примеры плотностей распределения.

1. Плотность величины ξ, равномерно распределенной на $[a, b]$:

$$f_\xi(x) = \begin{cases} 0, & x < a, \\ 1/(b-a), & a \leqslant x \leqslant b, \\ 0, & x > b. \end{cases}$$

2. *Показательная плотность* распределения:

$$f_\xi(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geqslant 0. \end{cases}$$

3. *Нормальная плотность* распределения:

$$f_\xi(x) = \frac{1}{\sqrt{2\pi b}}\, e^{-x^2/(2b)}.$$

Величина ξ имеет *решетчатое распределение*, если она дискретна и все ее возможные значения имеют вид $a + kh$, $k = 0, \pm 1, \ldots$

Величина h называется *шагом распределения*. Максимальное $\bar{h}$, для которого при некотором a

$$\sum_k \mathsf{P}\{\xi = a + k\bar{h}\} = 1,$$

существует, если только ξ принимает с положительной вероятностью хотя бы два значения. Это $\bar{h}$ называется *максимальным шагом распределения*. Если возможные значения ξ равны $a + kh$, то $\bar{h} = dh$, где d — наибольший общий делитель таких разностей $k_1 - k_2$, для которых $\mathsf{P}\{\xi = a + k_1 h\} > 0$ и $\mathsf{P}\{\xi = a + k_2 h\} > 0$. Среди решетчатых величин выделяется важный класс *целочисленных величин*, для которых

$$\sum_{k=-\infty}^{\infty} \mathsf{P}\{\xi = k\} = 1.$$

Примем обозначение $p_k = \mathsf{P}\{\xi = k\}$.

4. Величина ξ имеет *биномиальное распределение*, если при некоторых $0 < a < 1$, $n > 0$

$$p_k = 0, \quad k < 0; \quad p_k = C_n^k a^k (1-a)^{n-k}, \quad k \leqslant n; \quad p_k = 0, \quad k > n.$$

5. Величина ξ имеет *геометрическое распределение*, если при некотором $0 < a < 1$

$$p_k = 0, \quad k < 0; \quad p_k = a^k (1-a), \quad k \geqslant 0.$$

6. Величина ξ имеет *распределение Пуассона*, если при некотором $a > 0$

$$p_k = 0, \quad k < 0; \quad p_k = \frac{a^k}{k!} e^{-a}, \quad k \geqslant 0.$$

1.5. Группы случайных величин

1.5.1. Совместное распределение случайных величин. Пусть на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$ заданы m случайных величин $\xi_1(\omega), \ldots, \xi_m(\omega)$. Тогда для всех $a_1 < b_1, \ldots, a_m < b_m$

$$\{\omega\colon a_1 \leqslant \xi_1(\omega) < b_1, \ldots, a_m \leqslant \xi_m(\omega) < b_m\} = \\ = \bigcap_{k=1}^{m} \{\omega\colon a_k \leqslant \xi_k(\omega) < b_k\} \in \mathfrak{A}. \qquad (5.1)$$

Обозначим через $(\xi_1(\omega), \ldots, \xi_m(\omega))$ точку в $\mathbf{R}^m$, а через Δ полуоткрытый параллелепипед:

$$\Delta = \{\bar{x}\colon a_1 \leqslant x^1 < b_1, \ldots, a_m \leqslant x^m < b_m\}.$$

Соотношение (5.1) можно переписать так:

$$\{\omega\colon (\xi_1(\omega), \ldots, \xi_m(\omega)) \in \Delta\} \in \mathfrak{A}. \qquad (5.2)$$

Используя (5.2) и равенства

$$\bigcap_k \{\omega: (\xi_1(\omega), \ldots, \xi_m(\omega)) \in B_k\} = \left\{\omega: (\xi_1(\omega), \ldots, \xi_m(\omega)) \in \bigcap_k B_k\right\},$$

$$\bigcup_k \{\omega: (\xi_1(\omega), \ldots, \xi_m(\omega)) \in B_k\} = \left\{\omega: (\xi_1(\omega), \ldots, \xi_m(\omega)) \in \bigcup_k B_k\right\},$$

справедливые для всякой последовательности множеств из $\mathbf{R}^m$, устанавливаем, что (5.2) справедливо, если Δ — произвольное борелевское множество из $\mathbf{R}^m$.

Мера $\mu_{\xi_1, \ldots, \xi_m}$, определенная на борелевских множествах соотношением

$$\mu_{\xi_1, \ldots, \xi_m}(B) = \mathbf{P}\left(\{\omega: (\xi_1(\omega), \ldots, \xi_m(\omega)) \in B\}\right), \tag{5.3}$$

называется *совместным распределением случайных величин* $\xi_1, \ldots, \xi_m$ или *распределением случайного вектора* $\xi = (\xi_1(\omega), \ldots, \xi_m(\omega))$ в $\mathbf{R}^m$. Как указывалось в § 1.3, для определения меры $\mu_{\xi_1, \ldots, \xi_m}$ достаточно задать функцию

$$F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \mathbf{P}\left(\{\omega: \xi_1(\omega) < x_1, \ldots, \xi_m(\omega) < x_m\}\right) = \\ = \mathbf{P}\{\xi_1 < x_1, \ldots, \xi_m < x_m\}, \tag{5.4}$$

которая называется *совместной функцией распределения* величин $\xi_1, \ldots, \xi_m$. Эта функция является m-мерной функцией распределения и, следовательно, удовлетворяет условиям 1) — 4). Зная совместную функцию распределения величин $\xi_1, \ldots, \xi_m$, можно определить и совместную функцию распределения величин $\xi_{i_1}, \ldots, \xi_{i_k}$, где $0 < i_1 < \ldots < i_k \leqslant m$:

$$F_{\xi_{i_1}, \ldots, \xi_{i_k}}\left(x_{i_1}, \ldots, x_{i_k}\right) = F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)\Big|_{j \neq i_1, \ldots, i_k}^{x_j = +\infty} \tag{5.5}$$

(под $F(+\infty)$ понимается $\lim\limits_{x \to +\infty} F(x)$). Соотношение (5.5) следует непосредственно из (5.4), если только $\{\xi_j < +\infty\}$ — достоверное событие. Совместные функции распределения подмножества случайных величин, получаемые из функции распределения всех величин, называются *маргинальными* (частными) функциями распределения (формула (5.5) определяет *k-мерные маргинальные* распределения). В частности, зная $F_{\xi_1, \ldots, \xi_m}$, определяем и функции распределений величин ξ_k:

$$F_{\xi_k}(x) = F_{\xi_1, \ldots, \xi_m}(\overbrace{+\infty, \ldots, +\infty, x}^{k}, +\infty, \ldots, +\infty).$$

1.5.2. Дискретные и непрерывные распределения. Если каждая из величин ξ_k имеет дискретное распределение, то говорят, что случайный вектор $(\xi_1, \ldots, \xi_m)$ также имеет дискретное распределение (или что совместное распределение величин $\xi_1, \ldots, \xi_m$ дискретно).

Пусть ξ_k принимает значения $\{y_1^k, y_2^k, \ldots\}$. Тогда совместное распределение величин $\xi_1, \ldots, \xi_m$ определяется вероятностями

$$p_{i_1, \ldots, i_m} = \mathsf{P}\left\{\xi_1 = y_{i_1}^1,\ \xi_2 = y_{i_2}^2, \ldots, \xi_m = y_{i_m}^m\right\}.$$

Мера, задающая совместное распределение $\xi_1, \ldots, \xi_m$, задается в этом случае равенством

$$\mu_{\xi_1, \ldots, \xi_m}(B) = \sum p_{i_1, \ldots, i_m} \chi_B\left(y_{i_1}^1, \ldots, y_{i_m}^m\right),$$

где $\chi_B(y^1, \ldots, y^m) = 1$, если $(y^1, \ldots, y^m) \in B$; $\chi_B(y^1, \ldots, y^m) = 0$, если $(y^1, \ldots, y^m) \notin B$; $(y^1, \ldots, y^m)$ — точка в $\mathbf{R}^m$. Совместная функция распределения задается формулой

$$F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \sum_{y_{i_1}^1 < x_1, \ldots, y_{i_m}^m < x_m} p_{i_1, \ldots, i_m}.$$

Примером m-мерного дискретного распределения является m-мерное *полиномиальное* распределение. Величины ξ_k целочисленны, причем при некоторых $p_i \geqslant 0$, $i = 1, \ldots, m$, $\sum_{i=1}^{m} p_i = 1$ имеет место равенство

$$p\{\xi_1 = i_1, \ldots, \xi_m = i_m\} = \begin{cases} \dfrac{n!}{i_1! \ldots i_m!} p_1^{i_1} \ldots p_m^{i_m}, & \text{если } i_k \geqslant 0;\ k = 1, \ldots, m; \\ & i_1 + \ldots + i_m = n; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Величины $\xi_1, \ldots, \xi_m$ имеют *совместное абсолютно непрерывное* распределение, если существует такая измеримая по Лебегу функция $f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)$, что совместное распределение величин $\xi_1, \ldots, \xi_m$ определяется формулой

$$\mu_{\xi_1, \ldots, \xi_m}(B) = \int \ldots \int_B f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m$$

(справа записан m-кратный интеграл Лебега). Тогда функция $f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)$ называется *совместной плотностью* распределения величин $\xi_1, \ldots, \xi_m$. Совместная функция распределения выражается через совместную плотность по формуле

$$F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_m} f_{\xi_1, \ldots, \xi_m}(y_1, \ldots, y_m)\, dy_1 \ldots dy_m.$$

Из этого соотношения следует формула для плотности:

$$f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \frac{\partial^m}{\partial x_1 \ldots \partial x_m} F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m).$$

Отметим, что существование производной, стоящей в правой части последнего равенства, для почти всех $x_1, \ldots, x_m$ еще не обеспечивает существования плотности. Чтобы последняя существовала, необходимо и достаточно выполнения условия

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \frac{\partial^m}{\partial x_1 \ldots \partial x_m} F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m = 1.$$

Свойства совместной плотности:

а) $f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) \geqslant 0$ для почти всех $x_1, \ldots, x_m$;

б) $\displaystyle\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m = 1.$

Всякая измеримая по Лебегу функция $g(x_1, \ldots, x_m)$, удовлетворяющая этим двум условиям, может выступать в качестве совместной плотности некоторых m случайных величин и называется *m-мерной плотностью*.

Проинтегрировав плотность по аргументам x_j, $j \neq i_1, \ldots, i_k$, от $-\infty$ до $+\infty$, получим совместную плотность распределения величин $\xi_{i_1}, \ldots, \xi_{i_k}$. В частности,

$$f_{\xi_k}(x) = \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{m-1} f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_{k-1}, x, x_{k+1}, \ldots, x_m) \times$$
$$\times dx_1 \ldots dx_{k-1}\, dx_{k+1} \ldots dx_m.$$

Примеры m-мерных плотностей.

1. Случайный вектор $(\xi_1, \ldots, \xi_m)$ равномерно распределен в ограниченном измеримом множестве $G \subset \mathbf{R}^m$, если совместная плотность величин $\xi_1, \ldots, \xi_m$ имеет вид

$$f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \begin{cases} \dfrac{1}{\operatorname{mes} G}, & (x_1, \ldots, x_m) \in G; \\ 0, & (x_1, \ldots, x_m) \notin G, \end{cases}$$

где $\operatorname{mes} G$ — лебегова мера G в $\mathbf{R}^m$.

2. Случайный вектор $(\xi_1, \ldots, \xi_m)$ имеет невырожденное нормальное распределение, если существуют такие числа $a_1, \ldots, a_m$ и невырожденная симметрическая положительно определенная матрица $C = \| C_{ij} \|$, что совместная плотность распределения величин $\xi_1, \ldots, \xi_m$ имеет вид

$$f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = (2\pi)^{-m/2} [\det C]^{1/2} \times$$
$$\times \exp\left\{ \sum_{i, j=1}^{m} c_{ij}(x_i - a_i)(x_j - a_j) \right\}.$$

1.5.3. Функции случайных величин. Зная совместное распределение величин $\xi_1, \ldots, \xi_m$, можем определить функцию распределения некоторой функции $g(\xi_1, \ldots, \xi_m)$ от этих случайных величин, где $g(x_1, \ldots, x_m)$ — некоторая борелевская функция, определенная на $\mathbf{R}^m$, т. е. функция, измеримая относительно σ-алгебры борелевских множеств в $\mathbf{R}^m$. Пусть $\eta = g(\xi_1, \ldots, \xi_m)$, тогда

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \int\limits_{g(x_1, \ldots, x_m) < x} \cdots \int \mu_{\xi_1, \ldots, \xi_m}(d\bar{x})$$

(здесь $\bar{x} = (x_1, \ldots, x_m)$ — переменная интегрирования в $\mathbf{R}^m$; интеграл — интеграл Лебега по мере $\mu_{\xi_1, \ldots, \xi_m}$ — берется по области $\{\bar{x}\colon g(x_1, \ldots, x_m) < x\}$, являющейся борелевским подмножеством в $\mathbf{R}^m$).

Пусть $g(x_1, \ldots, x_m)$ — дифференцируемая функция и

$$|\operatorname{grad} g(x_1, \ldots, x_m)| = \sqrt{\sum_{k=1}^{m} \left(\frac{\partial g}{\partial x_k}(x_1, \ldots, x_m)\right)^2} > 0.$$

Если существует совместная плотность величин $\xi_1, \ldots, \xi_m$, то величина η также имеет плотность, определяемую формулой

$$f_\eta(x) = \overbrace{\int\limits_{g(x_1, \ldots, x_m) = x} \cdots \int}^{m-1} f(x_1, \ldots, x_m) \frac{dS_{m-1}}{|\operatorname{grad} g(x_1, \ldots, x_m)|},$$

где интеграл справа — поверхностный интеграл по $(m-1)$-мерной поверхности в $\mathbf{R}^m$, задаваемой уравнением $g(x_1, \ldots, x_m) = x$.

Пусть $g_1(x_1, \ldots, x_m)$, $g_2(x_1, \ldots, x_m)$, $\ldots$, $g_k(x_1, \ldots, x_m)$ — борелевские функции, заданные на $\mathbf{R}^m$. Положим $\eta_i = g_i(\xi_1, \ldots, \xi_m)$. Тогда совместное распределение величин $\eta_1, \ldots, \eta_k$ определяется формулой

$$\mu_{\eta_1, \ldots, \eta_k}(C) = \int\limits_{(g_1(x_1, \ldots, x_m), \ldots, g_k(x_1, \ldots, x_m)) \in C} \cdots \int \mu_{\xi_1, \ldots, \xi_m}(d\bar{x})$$

для борелевского множества $C \subset \mathbf{R}^k$; интеграл справа берется по множеству $\{\bar{x}\colon \bar{g}(\bar{x}) \in C\}$, где $\bar{g} = (g_1, \ldots, g_k)$ — точка в $\mathbf{R}^k$.

Предположим, что функции $g_1, \ldots, g_k$ $(k < m)$ дифференцируемы и существует совместная плотность распределения величин $\xi_1, \ldots, \xi_m$, тогда совместная плотность распределения величин $\eta_1, \ldots, \eta_k$ выражается формулой

$$f_{\eta_1, \ldots, \eta_k}(y_1, \ldots, y_k) = \overbrace{\int\limits_{\substack{(g_1(x_1, \ldots, x_m) \ldots \\ \ldots g_k(x_1, \ldots, x_m)) \in C}} \cdots \int}^{m-k} f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) \times$$

$$\times \frac{dS_{m-k}}{\left(\sum\limits_{i_1 < \ldots < i_k} \left[\dfrac{D(g_1, \ldots, g_k)}{D(x_{i_1}, \ldots, x_{i_k})}\right]^2\right)^{1/2}}.$$

Интеграл справа — поверхностный интеграл по поверхности размерности $m-k$, определяемой системой уравнений $g_1(x_1, \ldots, x_m) = y_1; \ldots; g_k(x_1, \ldots, x_m) = y_k$, а

$$\frac{D(g_1, \ldots, g_k)}{D(x_{i_1}, \ldots, x_{i_k})} = \begin{vmatrix} \frac{\partial g_1}{\partial x_{i_1}} & \ldots & \frac{\partial g_1}{\partial x_{i_k}} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \frac{\partial g_k}{\partial x_{i_1}} & \ldots & \frac{\partial g_k}{\partial x_{i_k}} \end{vmatrix}$$

есть якобиан функций $g_1, \ldots, g_k$ по переменным $x_{i_1}, \ldots, x_{i_k}$.

Рассмотрим распределения простейших функций от пары случайных величин.

Распределение суммы (разности) двух величин. Функция распределения суммы (разности) задается формулой

$$F_{\xi_1 \pm \xi_2}(x) = \iint\limits_{x^1 \pm x^2 < x} \mu_{\xi_1, \xi_2}(dx^1, dx^2).$$

Если существует f_{ξ_1, ξ_2}, то

$$f_{\xi_1 \pm \xi_2}(x) = \int f_{\xi_1, \xi_2}(x \mp y, y)\, dy.$$

Если ξ_1 и ξ_2 — дискретные целочисленные величины: $p_{kl} = \mathsf{P}\{\xi_1 = k, \xi_2 = l\}$, то

$$\mathsf{P}\{\xi_1 \pm \xi_2 = l\} = \sum_{k=-\infty}^{\infty} p_{l \mp k\, k}.$$

Распределение произведения двух величин. Функция распределения произведения задается формулой

$$F_{\xi_1\xi_2}(x) = \iint\limits_{x^1 x^2 < x} \mu_{\xi_1, \xi_2}(dx^1, dx^2).$$

Если существует f_{ξ_1, ξ_2}, то

$$f_{\xi_1\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1, \xi_2}\left(t, \frac{x}{t}\right) \frac{dt}{|t|}.$$

Распределение отношения двух величин. Функция распределения отношения задается формулой

$$F_{\xi_1/\xi_2}(x) = \int\limits_{x^1/x^2 < x} \mu_{\xi_1, \xi_2}(dx^1, dx^2).$$

Если существует f_{ξ_1/ξ_2}, то

$$f_{\xi_1/\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1, \xi_2}\left(t, \frac{t}{x}\right) \frac{|t|\, dt}{x^2}.$$

1.6. Математическое ожидание

1.6.1. Математическое ожидание дискретной величины. Пусть в случайном эксперименте наблюдается некоторая случайная величина ξ, которая может принимать конечное число значений $a_1, \ldots, a_N$ с вероятностями $p_1, \ldots, p_N$. Если $x_1, \ldots, x_n$ — наблюдения нашей величины в n последовательных осуществлениях эксперимента, то среднее значение реализаций можно представить в виде

$$\frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \sum_{k=1}^{N} a_k \nu_n(A_k),$$

где A_k — события $\{\xi = a_k\}$, ν_n — частота события. Заменяя частоты на вероятности, получим выражение

$$\mathsf{M}\xi = \sum_{k=1}^{N} a_k p_k,$$

которое называется *средним* или *математическим ожиданием* случайной величины ξ.

Если ξ — произвольная дискретная случайная величина, принимающая значения a_k $(k = 1, 2, \ldots)$ с вероятностями p_k, то $\mathsf{M}\xi = \sum_{k=1}^{\infty} a_k p_k$, если только ряд справа сходится абсолютно. Отметим некоторые свойства математического ожидания дискретной величины.

1) $\mathsf{M}(\xi_1 + \xi_2) = \mathsf{M}\xi_1 + \mathsf{M}\xi_2$;
2) $\mathsf{M}(\lambda\xi) = \lambda\mathsf{M}\xi$ для всех λ;
3) если $\mathsf{P}\{\xi_1 = \xi_2\} = 1$, то $\mathsf{M}\xi_1 = \mathsf{M}\xi_2$;
4) если $\xi \geqslant 0$, то $\mathsf{M}\xi \geqslant 0$;
5) если $\mathsf{P}\{\xi = c\} = 1$, то $\mathsf{M}\xi = c$.

1.6.2. Математическое ожидание произвольной величины. Для определения математического ожидания произвольной случайной величины ξ введем последовательность дискретных случайных величин ξ_n, определяемых равенством $\xi_n = k/n$, если $k/n \leqslant \xi < (k+1)/n$, $k = 0, \pm 1, \pm 2, \ldots$; $n = 1, 2, \ldots$ Очевидно, что $|\xi_n - \xi| \leqslant 1/n$. Если $\mathsf{M}\xi_n$ существует при некотором n, то оно существует для всех n и существует предел

$$\lim_{n\to\infty} \mathsf{M}\xi_n = \lim_{n\to\infty} \sum_{k=-\infty}^{\infty} \frac{k}{n} \mathsf{P}\left\{\frac{k}{n} \leqslant \xi < \frac{k+1}{n}\right\}.$$

Этот предел называется *математическим ожиданием* величины ξ и обозначается $\mathsf{M}\xi$. Таким образом определенное математическое ожидание также удовлетворяет свойствам 1) — 5). Если ξ — неотрицательная случайная величина, то считаем $\mathsf{M}\xi$ всегда определенным и равным $+\infty$ в том случае, когда ряд $\sum_{k=0}^{\infty} \frac{k}{n} \mathsf{P}\left\{\frac{k}{n} \leqslant \xi < \frac{k+1}{n}\right\}$ расходится.

1.6.3. Формулы для вычисления математического ожидания. Если $F_\xi(x)$ — функция распределения величины ξ, то

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x\, dF_\xi(x) \quad \text{при} \quad \int_{-\infty}^{\infty} |x|\, dF_\xi(x) < \infty$$

(интегралы справа являются интегралами Стилтьеса и вычисляются как пределы интегральных сумм). Если существует плотность $f_\xi(x)$ величны ξ, то

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x f_\xi(x)\, dx \quad \text{при} \quad \int_{-\infty}^{\infty} |x|\, f_\xi(x)\, dx < \infty.$$

Если величина $\xi = \xi(\omega)$ задана на вероятностном пространстве $\{\Omega, \mathfrak{A}, \mathsf{P}\}$, то ее математическое ожидание может быть вычислено с помощью интеграла Лебега по мере P:

$$\mathsf{M}\xi(\omega) = \int \xi(\omega)\, \mathsf{P}(d\omega),$$

при условии, что интеграл справа существует.

Пусть $\xi_1, \ldots, \xi_m$ — случайные величины, $F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)$ — их совместная функция распределения, $g(x_1, \ldots, x_m)$ — некоторая борелевская функция. Тогда

$$\mathsf{M}g(\xi_1, \ldots, \xi_m) = \int \ldots \int g(x_1, \ldots, x_m)\, dF_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m),$$

если только интеграл справа абсолютно сходится (он понимается как m-кратный интеграл Лебега — Стилтьеса); если g — непрерывная функция, то его можно вычислять как интеграл Римана — Стилтьеса. В том случае, когда существует совместная плотность величин $\xi_1, \ldots, \xi_m$, предыдущая формула принимает вид

$$\mathsf{M}g(\xi_1, \ldots, \xi_m) = \\ = \int \ldots \int g(x_1, \ldots, x_m)\, f_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m,$$

если только m-кратный интеграл Лебега в правой части абсолютно сходится.

1.6.4. Моменты случайных величин. Величина

$$\mathsf{M}\xi^k = \int x^k\, dF_\xi(x), \qquad k = 1, 2, \ldots,$$

называется *k-м моментом* величины ξ (если указанное математическое ожидание существует); k-й момент величины $(\xi - \mathsf{M}\xi)$ называется *k-м центральным моментом*. Он вычисляется по формуле

$$\mathsf{M}(\xi - \mathsf{M}\xi)^k = \int (x - \mathsf{M}\xi)^k\, dF_\xi(x).$$

k-й момент случайной величины $|\xi|$ называется *абсолютным k-м моментом* величины ξ.

Особую роль играет второй центральный момент, который называется *дисперсией* величины и обозначается $\mathsf{D}\xi$:

$$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2 =$$
$$= \int (x - \mathsf{M}\xi)^2 \, dF_\xi(x) = \int x^2 \, dF_\xi(x) - \left(\int x \, dF_\xi(x)\right)^2.$$

Для абсолютно непрерывных величин дисперсия вычисляется по формуле

$$\mathsf{D}\xi = \int (x - \mathsf{M}\xi)^2 f_\xi(x) \, dx = \int x^2 f_\xi(x) \, dx - \left(\int x f_\xi(x) \, dx\right)^2.$$

Для дискретной величины ξ, принимающей значения a_k с вероятностями p_k,

$$\mathsf{D}\xi = \sum_k a_k^2 p_k - \left(\sum_k a_k p_k\right)^2.$$

Заметим, что $\mathsf{D}\xi$ всегда определена, если определено $\mathsf{M}\xi$, но может принимать значения $+\infty$.

Величина $\sigma = \sqrt{\mathsf{D}\xi}$ называется *среднеквадратическим отклонением* величины ξ.

Отметим одно важное свойство величины $\mathsf{D}\xi$: если $\mathsf{D}\xi = 0$, то $\mathsf{P}\{\xi = \mathsf{M}\xi\} = 1$, т. е. в этом случае величина ξ с вероятностью 1 постоянна.

Пусть $\xi_1, \ldots, \xi_m$ — случайные величины с совместной функцией распределения $F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m)$. Величины

$$m_{\xi_1, \ldots, \xi_m}(k_1, \ldots, k_m) =$$
$$= \int \ldots \int x_1^{k_1} \ldots x_m^{k_m} \, dF_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) = \mathsf{M}\xi_1^{k_1} \ldots \xi_m^{k_m},$$

где $k_1, \ldots, k_m \geqslant 0$, $k_1 + \ldots + k_m = k$, называются *смешанными моментами* величин $\xi_1, \ldots, \xi_m$ порядка k. Среди смешанных моментов особую роль играют смешанные моменты второго порядка. Величина

$$\mathsf{M}(\xi - \mathsf{M}\xi)(\eta - \mathsf{M}\eta) = \operatorname{cov}(\xi, \eta)$$

называется *ковариацией величин* ξ и η. Величина

$$r_{\xi, \eta} = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{\mathsf{D}\xi \, \mathsf{D}\eta}}$$

называется *коэффициентом корреляции* величин ξ и η. Отметим некоторые свойства коэффициента корреляции:

1) $-1 \leqslant r_{\xi, \eta} \leqslant 1$;

2) если $|r_{\xi, \eta}| = 1$, то с вероятностью 1 выполняется соотношение

$$\eta = r_{\xi, \eta} \sqrt{\frac{\mathsf{D}\eta}{\mathsf{D}\xi}} (\xi - \mathsf{M}\xi) + \mathsf{M}\eta$$

(т. е. в этом случае ξ и η связаны линейным соотношением). Поэтому $r_{\xi, \eta}$ можно рассматривать как меру линейной зависимости величин ξ и η. Если величины ξ и η таковы, что $r_{\xi,\eta}=0$, то они называются *некоррелированными*.

Пусть $\xi_1, \ldots, \xi_m$ попарно некоррелированы. Тогда

$$\mathsf{D}\sum_{k=1}^{m}\xi_k=\sum_{k=1}^{m}\mathsf{D}\xi_k.$$

Если $\xi_1, \ldots, \xi_m$ — случайные величины, для которых $\mathsf{M}\xi_i^2<\infty$, $i=1, \ldots, m$, то матрица $R=\|b_{ij}\|_{i,\,j=1,\ldots,m}$, где $b_{ij}=\operatorname{cov}(\xi_i, \xi_j)$ называется *ковариационной* (*корреляционной*) матрицей для величин $\xi_1, \ldots, \xi_m$ (вектора $(\xi_1, \ldots, \xi_m)$). Корреляционная матрица положительно определена, т. е. для любых комплексных $\alpha_1, \ldots, \alpha_m$ выполнено неравенство

$$\sum_{i,\,j=1}^{m} b_{ij}\alpha_i\bar{\alpha}_j\geqslant 0,$$

где $\bar{\alpha}$ — комплексно сопряженное число к α.

Вычислим математические ожидания и дисперсии некоторых случайных величин с дискретными и абсолютно непрерывными распределениями:

а) ξ — целочисленная величина, равномерно распределенная на $[0, N]$: $\mathsf{P}\{\xi=k\}=\dfrac{1}{N+1}$, $k=0, 1, \ldots, N$,

$$\mathsf{M}\xi=\sum_{k=0}^{N}k\frac{1}{N+1}=\frac{N}{2},$$

$$\mathsf{D}\xi=\sum_{k=0}^{N}k^2\frac{1}{N+1}-\left(\frac{N}{2}\right)^2=\frac{N^2}{12}+\frac{N}{6};$$

б) ξ имеет биномиальное распределение: $\mathsf{P}\{\xi=k\}=C_N^k a^k(1-a)^{N-k}$, $k=0, 1, \ldots, N$; $0<a<1$,

$$\mathsf{M}\xi=\sum_{k=0}^{N}kC_N^k a^k(1-a)^{N-k}=Na,$$

$$\mathsf{D}\xi=\sum_{k=0}^{N}k^2C_N^k a^k(1-a)^{N-k}-N^2a^2=Na(1-a);$$

в) ξ имеет геометрическое распределение:

$\mathsf{P}\{\xi=k\}=(1-a)a^k$, $k=0, 1, \ldots$; $0<a<1$,

$$\mathsf{M}\xi=\sum_{k=0}^{\infty}k(1-a)a^k=\frac{a}{1-a},$$

$$\mathsf{D}\xi=\sum_{k=0}^{\infty}k^2(1-a)a^k-\left(\frac{a}{1-a}\right)^2=\frac{a+a^2}{(1-a)^2}-\frac{a^2}{(1-a)^2}=\frac{a}{(1-a)^2};$$

г) ξ имеет распределение Пуассона: $\mathsf{P}\{\xi = k\} = \frac{a^k}{k!} e^{-a}$, $k = 0, 1, 2, \ldots$; $a > 0$,

$$\mathsf{M}\xi = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} = a,$$

$$\mathsf{D}\xi = \sum_{k=0}^{\infty} k^2 \frac{a^k}{k!} e^{-a} - a^2 = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} + \\ + \sum_{k=0}^{\infty} k(k-1) \frac{a^k}{k!} e^{-a} - a^2 = a;$$

д) ξ имеет равномерное распределение на $[a, b]$:

$$\mathsf{M}\xi = \frac{1}{b-a} \int_a^b x\,dx = \frac{b+a}{2},$$

$$\mathsf{D}\xi = \frac{1}{b-a} \int_a^b x^2\,dx - \left(\frac{b+a}{2}\right)^2 = \left(\frac{b-a}{12}\right)^2;$$

е) ξ имеет показательное распределение: $f_\xi(x) = \lambda e^{-\lambda x}$, $x \geqslant 0$; $f_\xi(x) = 0$, $x \leqslant 0$,

$$\mathsf{M}\xi = \lambda \int_0^\infty x e^{-\lambda x}\,dx = \frac{1}{\lambda},$$

$$\mathsf{D}\xi = \lambda \int_0^\infty x^2 e^{-\lambda x}\,dx - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2};$$

ж) ξ имеет нормальное распределение: $f_\xi(x) = \frac{1}{\sqrt{2\pi b}} \times \exp\left\{-\frac{(x-a)^2}{2b}\right\}$,

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi b}}\, x \exp\left\{-\frac{(x-a)^2}{2b}\right\} dx = a;$$

$$\mathsf{D}\xi = \frac{1}{\sqrt{2\pi b}} \int_{-\infty}^{\infty} (x-a)^2 \exp\left\{-\frac{(x-a)^2}{2b}\right\} dx = b.$$

1.7. Условные вероятности и математические ожидания

1.7.1. Условная вероятность относительно события. *Условной вероятностью* события A относительно события B, для которого $\mathsf{P}(B) > 0$, называется величина

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)}.$$

Отсюда вытекает *формула умножения вероятностей*

$$\mathsf{P}(A \cap B) = \mathsf{P}(A \mid B)\,\mathsf{P}(B) = \mathsf{P}(B \mid A)\,\mathsf{P}(A),$$

позволяющая выразить $\mathsf{P}(A \mid B)$ через $\mathsf{P}(B \mid A)$:

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(B \mid A)\,\mathsf{P}(A)}{\mathsf{P}(B)}.$$

Приведем также общую формулу умножения вероятностей:

$$\mathsf{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathsf{P}(A_1)\,\mathsf{P}(A_2 \mid A_1) \ldots \mathsf{P}\left(A_n \,\middle|\, \bigcap_{k=1}^{n-1} A_k\right).$$

Формула полной вероятности: пусть $E_1, \ldots, E_n$ — полная группа событий (предполагаем, что $\mathsf{P}(E_i) > 0$, $i = 1, \ldots, n$). Тогда для всякого события A

$$\mathsf{P}(A) = \sum_{k=1}^{n} \mathsf{P}(A \mid E_k)\,\mathsf{P}(E_k).$$

Формула Байеса дает выражение для условных вероятностей одного из событий E_k полной группы $E_1, \ldots, E_n$ при условии, что произошло событие A:

$$\mathsf{P}(E_k \mid A) = \frac{\mathsf{P}(A \mid E_k)\,\mathsf{P}(E_k)}{\sum_{i=1}^{n} \mathsf{P}(A \mid E_i)\,\mathsf{P}(E_i)}.$$

Эта формула называется также *формулой для вероятности гипотезы после испытания*. Предположим, что событие A может происходить при гипотезе H_i, заключающейся в том, что произошло событие E_i, с вероятностью $\mathsf{P}(A \mid E_i)$, а $\mathsf{P}(E_i)$ — вероятность гипотезы H_i. Формула Байеса позволяет вычислить условную вероятность гипотезы H_k при условии, что произошло событие A, через вероятности гипотез и вероятности события A при различных гипотезах.

Пример. Имеется n урн, содержащих черные и белые шары. Вероятность события, заключающегося в том, что извлечен черный шар из k-й урны, равна p_k. Выбирается наудачу (с вероятностью $1/n$) одна урна, после чего из нее извлекается шар. Какова вероятность того, что мы выбрали k-ю урну, если шар оказался черным? Если событие E_k заключается в том, что выбранная урна была k-й, а событие A — в том, что шар оказался черным, то

$$\mathsf{P}(E_k \mid A) = \frac{p_k}{p_1 + \ldots + p_n}.$$

1.7.2. Условное распределение случайной величины. Рассмотрим некоторую величину ξ. Выражение

$$F_\xi(x \mid A) = \frac{\mathsf{P}(\{\xi < x\} \cap A)}{\mathsf{P}(A)}$$

называется *условной функцией распределения* величины ξ относительно события A. Она определена, если $\mathsf{P}(A) > 0$. Если $F_\xi(x \mid A)$ абсолютно непрерывна и

$$F_\xi(x \mid A) = \int_{-\infty}^{x} f_\xi(t \mid A)\, dt,$$

то $f_\xi(t \mid A)$ называется *условной плотностью распределения* величины ξ относительно события A. Как условная функция распределения, так и условная плотность распределения обладают свойствами функции распределения и плотности распределения соответственно. Моменты, вычисленные по условной функции распределения, называются *условными моментами* величины. В частности, выражение

$$\mathsf{M}(\xi \mid A) = \int x\, dF_\xi(x \mid A),$$

если интеграл справа сходится абсолютно, называется *условным математическим ожиданием* величины ξ относительно события A. Если ξ задана на вероятностном пространстве $(\Omega, \mathfrak{A}, \mathsf{P})$, то для условного математического ожидания можно привести другое выражение:

$$\mathsf{M}(\xi \mid A) = \frac{1}{\mathsf{P}(A)} \int_A \xi(\omega)\, \mathsf{P}(d\omega).$$

Пусть $E_1, \ldots, E_n$ — полная группа событий, $\mathsf{P}(E_i) > 0$ $(i = 1, \ldots, n)$. Справедлива следующая *формула полного математического ожидания*:

$$\mathsf{M}\xi = \sum_{k=1}^{n} \mathsf{M}(\xi \mid E_k)\, \mathsf{P}(E_k).$$

Можно привести и некоторое обобщение этой формулы. Если C имеет вид $C = \bigcup E_{i_k}$, то

$$\int_C \xi(\omega)\, \mathsf{P}(d\omega) = \sum_{E_k \subset C} \mathsf{M}(\xi \mid E_k)\, \mathsf{P}(E_k). \tag{7.1}$$

Предположим, что событие A заключается в том, что $\{a \leqslant \xi < b\}$. Тогда условная функция распределения

$$F_\xi(x \mid a \leqslant \xi < b) = \begin{cases} 0, & x < a, \\ \dfrac{F_\xi(x) - F_\xi(a)}{F_\xi(b) - F_\xi(a)}, & a \leqslant x < b, \\ 1, & x > b, \end{cases}$$

есть распределение *урезанной* величины ξ или *урезанное распределение*. Запишем математическое ожидание и дисперсию для урезанного распределения:

$$\mathsf{M}(\xi \mid \{a \leqslant \xi < b\}) = \frac{1}{F_\xi(b) - F_\xi(a)} \int_a^b x\, dF_\xi(x),$$

$$\mathsf{D}(\xi \mid \{a \leqslant \xi < b\}) =$$

$$= \frac{1}{F_\xi(b) - F_\xi(a)} \int_a^b x^2\, dF_\xi(x) - \left(\frac{1}{F_\xi(b) - F_\xi(a)} \int_a^b x\, dF_\xi(x)\right)^2.$$

1.7.3. Условная вероятность и условное математическое ожидание относительно σ-алгебры. Если $E_1, \ldots, E_n$ — полная группа событий, то всевозможные объединения $\bigcup E_{i_k}$ и пустое множество $\varnothing$ образуют наименьшую σ-алгебру, содержащую множества E_k. Обозначим эту σ-алгебру через $\mathfrak{A}_0$. Предположим, что $\mathsf{M}\xi$ существует. Случайная величина $\mathsf{M}(\xi \mid \mathfrak{A}_0)$, на E_k равная $\mathsf{M}(\xi \mid E_k)$, называется *условным математическим ожиданием величины ξ относительно σ-алгебры* $\mathfrak{A}_0$. Она удовлетворяет следующим условиям:

1) $\mathsf{M}(\xi \mid \mathfrak{A}_0)$ измерима относительно σ-алгебры $\mathfrak{A}_0$;

2) для всех $C \in \mathfrak{A}_0$

$$\int_C \xi(\omega)\, \mathsf{P}(d\omega) = \int_C \mathsf{M}(\xi \mid \mathfrak{A}_0)\, \mathsf{P}(d\omega).$$

Первое условие означает, что $\mathsf{M}(\xi \mid \mathfrak{A}_0)$ постоянно на множествах E_k. Второе условие следует из (7.1).

Рассмотрим общий случай. Величина η *$\mathfrak{A}_0$-измерима*, где $\mathfrak{A}_0$ — некоторая σ-алгебра событий из $\mathfrak{A}$, если для всякого интервала Δ

$$\{\omega\colon \eta(\omega) \in \Delta\} \in \mathfrak{A}_0.$$

Случайная величина η называется *условным математическим ожиданием величины ξ относительно σ-алгебры* $\mathfrak{A}_0 \subset \mathfrak{A}$, если: а) величина η $\mathfrak{A}_0$-измерима; б) для всех $C \in \mathfrak{A}_0$

$$\int_C \xi(\omega)\, \mathsf{P}(d\omega) = \int_C \eta(\omega)\, \mathsf{P}(d\omega).$$

Заметим, что эти два условия однозначно с вероятностью 1 определяют величину $\eta(\omega)$. Если $\eta_1(\omega)$ также удовлетворяет этим условиям, то

$$\int |\eta(\omega) - \eta_1(\omega)|\, \mathsf{P}(d\omega) =$$

$$= \int_{\{\eta_1 - \eta \geqslant 0\}} (\eta_1(\omega) - \eta(\omega))\, \mathsf{P}(d\omega) + \int_{\{\eta - \eta_1 > 0\}} (\eta(\omega) - \eta_1(\omega))\, \mathsf{P}(d\omega) =$$

$$= \int_{\{\eta_1 - \eta \geqslant 0\}} (\xi(\omega) - \xi(\omega))\, \mathsf{P}(d\omega) + \int_{\{\eta - \eta_1 > 0\}} (\xi(\omega) - \xi(\omega))\, \mathsf{P}(d\omega) = 0,$$

так как $\{\eta_1 - \eta \geqslant 0\}$, $\{\eta - \eta_1 > 0\} \in \mathfrak{A}_0$ в силу $\mathfrak{A}_0$-измеримости $\eta - \eta_1$. Условное математическое ожидание величины ξ относительно σ-алгебры $\mathfrak{A}_0$ обозначается $\mathsf{M}(\xi \mid \mathfrak{A}_0)$. Оно всегда существует, если $\mathsf{M}|\xi| < \infty$.

Выражение $\mathsf{P}(A \mid \mathfrak{A}_0) = \mathsf{M}(\chi_A(\omega) \mid \mathfrak{A}_0)$, где $\chi_A(\omega)$ — индикатор множества A, называется *условной вероятностью A относительно σ-алгебры* $\mathfrak{A}_0$.

Будем говорить, что σ-алгебра $\mathfrak{A}$ *счетно порождена*, если существует такая алгебра $\mathfrak{A}'$, содержащая не более чем счетное число множеств, что $\mathfrak{A}$ является наименьшей σ-алгеброй, содержащей $\mathfrak{A}'$. Так как условная вероятность определяется лишь с вероятностью 1, то $\mathsf{P}(A/\mathfrak{A}_0)$ можно менять на множестве P-меры 0. Если $\mathfrak{A}$ счетно порождена, то выражение для $\mathsf{P}(A \mid \mathfrak{A}_0)$ можно определить так, чтобы почти при всех ω оно было вероятностной мерой по A. При этом условное математическое ожидание величины вычисляется по формуле

$$\mathsf{M}(\xi \mid \mathfrak{A}_0) = \int \xi(\omega)\, \mathsf{P}(d\omega \mid \mathfrak{A}_0).$$

Выражение $F_\xi(x \mid \mathfrak{A}_0) = \mathsf{P}(\{\omega: \xi(\omega) < x\} \mid \mathfrak{A}_0)$ называется *условной функцией распределения величины ξ относительно σ-алгебры* $\mathfrak{A}_0$; если же

$$F_\xi(x \mid \mathfrak{A}_0) = \int_{-\infty}^{x} f_\xi(t \mid \mathfrak{A}_0)\, dt, \quad \text{то} \quad f_\xi(x \mid \mathfrak{A}_0)$$

называется *условной плотностью распределения величины ξ относительно σ-алгебры* $\mathfrak{A}$.

1.7.4. Свойства условных вероятностей и условных математических ожиданий.

1) Формула полного математического ожидания:

$$\mathsf{M}\xi = \mathsf{M}[\mathsf{M}(\xi \mid \mathfrak{A}_0)];$$

2) формула полной вероятности:

$$\mathsf{P}(A) = \mathsf{M}[\mathsf{P}(A \mid \mathfrak{A}_0)];$$

3) $\mathfrak{A}_0$-измеримый множитель η может быть вынесен из-под знака условного математического ожидания:

$$\mathsf{M}(\xi\eta \mid \mathfrak{A}_0) = \eta\mathsf{M}(\xi \mid \mathfrak{A}_0);$$

4) формула повторных математических ожиданий: если $\mathfrak{A}_0 \subset \mathfrak{A}_1 \subset \mathfrak{A}$, $\mathfrak{A}_0$ и $\mathfrak{A}_1$ — σ-алгебры, то с вероятностью 1

$$\mathsf{M}(\xi \mid \mathfrak{A}_0) = \mathsf{M}(\mathsf{M}(\xi \mid \mathfrak{A}_1) \mid \mathfrak{A}_0);$$

5) предельный переход по условию в условном математическом ожидании: пусть $\mathfrak{A}_k$ — σ-алгебры ($\mathfrak{A}_k \subset \mathfrak{A}_{k+1} \subset \mathfrak{A}$) и $\mathfrak{A}_\infty$ — наименьшая σ-алгебра, содержащая $\bigcup_k \mathfrak{A}_k$. Тогда с вероятностью 1

$$\mathsf{M}(\xi \mid \mathfrak{A}_\infty) = \lim_{k\to\infty} \mathsf{M}(\xi \mid \mathfrak{A}_k).$$

1.7.5. Способы вычисления условных распределений. Пусть ξ и η — две случайные величины. Через $\mathfrak{A}_\eta$ обозначим σ-алгебру множеств вида $\{\omega\colon \eta \in B\}$, где B — всевозможные борелевские множества на прямой; $\mathfrak{A}_\eta$ называется *σ-алгеброй, порожденной величиной* η. Функция $F_\xi(x \mid \mathfrak{A}_\eta)$ называется *условной функцией распределения величины* ξ *при заданном* η. Эта случайная величина $\mathfrak{A}_\eta$-измерима и потому является борелевской функцией от η. Будем обозначать ее значение при $\eta = y$ как $F_\xi(x \mid \eta = y)$. Пусть η имеет плотность распределения $f_\eta(y)$. Тогда

$$F_\xi(x \mid \eta = y) = \frac{1}{f_\eta(y)} \frac{\partial}{\partial y} F_{\xi,\eta}(x, y),$$

где $F_{\xi,\eta}(x, y)$ — совместная функция распределения величин ξ и η. Если существует совместная плотность $f_{\xi,\eta}(x, y)$ величин ξ и η, то выражение

$$f_\xi(x \mid \eta = y) = \frac{1}{f_\eta(y)} f_{\xi,\eta}(x, y)$$

определяет *условную плотность распределения величины* ξ *при заданном* η:

$$\mathsf{M}(\xi \mid \eta = y) = \frac{1}{f_\eta(y)} \int x\, d_x \frac{\partial}{\partial y} F_{\xi,\eta}(x, y)$$

будет *условным математическим ожиданием величины* ξ *при заданном* η.

Пусть $\xi_1, \ldots, \xi_m$ и $\eta_1, \ldots, \eta_n$ — некоторые случайные величины, $\mathfrak{A}_{\eta_1,\ldots,\eta_n}$ — σ-алгебра, порожденная величинами $\eta_1, \ldots, \eta_n$, т. е. σ-алгебра множеств вида

$$\{\omega\colon (\eta_1(\omega), \ldots, \eta_n(\omega)) \in B\},$$

где B — всевозможные борелевские множества в $\mathbf{R}^n$; $(\eta_1, \ldots, \eta_n)$ — точка в $\mathbf{R}^n$.

Функция

$$F_{\xi_1,\ldots,\xi_m}(x_1, \ldots, x_m \mid \mathfrak{A}_{\eta_1,\ldots,\eta_n}) = \\ = \mathsf{P}\left(\left\{\bigcap_{k=1}^{m} \{\omega\colon \xi_k(\omega) < x\}\right\} \middle| \mathfrak{A}_{\eta_1,\ldots,\eta_n}\right)$$

называется *совместной условной функцией распределения* величин $\xi_1, \ldots, \xi_m$ при заданных $\eta_1, \ldots, \eta_n$. Эта случайная величина является функцией величин $\eta_1, \ldots, \eta_n$. При $\eta_1 = y_1, \ldots, \eta_n = y_n$ будем обозначать ее

$$F_{\xi_1,\ldots,\xi_m \mid \eta_1,\ldots,\eta_n}(x_1, \ldots, x_m \mid y_1, \ldots, y_n).$$

Если $\eta_1, \ldots, \eta_n$ имеют совместную плотность распределения $f_{\eta_1, \ldots, \eta_n}(y_1, \ldots, y_m)$, то

$$F_{\xi_1, \ldots, \xi_m | \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m | y_1, \ldots, y_n) =$$
$$= \frac{1}{f_{\eta_1, \ldots, \eta_n}(y_1, \ldots, y_n)} \frac{\partial^n}{\partial y_1 \ldots \partial y_n} F_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m, y_1, \ldots, y_n),$$

где $F_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ — совместная функция распределения величин $\xi_1, \ldots, \xi_m$ и $\eta_1, \ldots, \eta_n$. Если существует совместная плотность $f_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ величин $\xi_1, \ldots, \xi_m$ и $\eta_1, \ldots, \eta_n$, то

$$f_{\xi_1, \ldots, \xi_m | \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m | y_1, \ldots, y_n) =$$
$$= \left(1/f_{\eta_1, \ldots, \eta_n}(y_1, \ldots, y_n)\right) f_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m, y_1, \ldots, y_n)$$

является условной совместной плотностью величин $\xi_1, \ldots, \xi_m$ при условии, что $\eta_1 = y_1, \ldots, \eta_n = y_n$.

1.7.6. Независимость. События A и B называются *независимыми*, если $\mathsf{P}(A \cap B) = \mathsf{P}(A)\,\mathsf{P}(B)$. Для независимых событий

$$\mathsf{P}(A | B) = \mathsf{P}(A), \quad \mathsf{P}(B | A) = \mathsf{P}(B).$$

Величина ξ не зависит от события A, если

$$F_\xi(x | A) = F_\xi(x).$$

Событие A не зависит от σ-алгебры $\mathfrak{A}_0$, если с вероятностью 1

$$\mathsf{P}(A | \mathfrak{A}_0) = \mathsf{P}(A).$$

Чтобы A не зависело от $\mathfrak{A}_0$, необходимо, чтобы A не зависело от любого события $B \subset \mathfrak{A}_0$, и достаточно, чтобы A не зависело от событий из некоторой алгебры $\widehat{\mathfrak{A}}_0$, такой, что $\mathfrak{A}_0$ — наименьшая σ-алгебра, содержащая $\widehat{\mathfrak{A}}_0$. Если $\mathsf{P}(A \cap A_1) = \mathsf{P}(A)\,\mathsf{P}(A_1)$ для всех $A \in \mathfrak{A}_0$, $A_1 \subset \mathfrak{A}_1$, то σ-алгебры $\mathfrak{A}_0$ и $\mathfrak{A}_1$ *независимы*. Величина ξ не зависит от $\mathfrak{A}_0$, если $\mathfrak{A}_\xi$ (σ-алгебра, порожденная величиной ξ) и $\mathfrak{A}_0$ независимы. Для этого необходимо и достаточно, чтобы с вероятностью 1

$$F_\xi(x | \mathfrak{A}_0) = F_\xi(x).$$

Величины ξ и η *независимы*, если независимы σ-алгебры $\mathfrak{A}_\xi$ и $\mathfrak{A}_\eta$. Для независимых величин ξ и η

$$F_{\xi, \eta}(x, u) = F_\xi(x)\,F_\eta(u),$$

где $F_{\xi,\eta}(x, u)$ — совместная функция распределения ξ и η; $F_\xi(x)$ и $F_\eta(u)$ — функции распределения ξ и η соответственно.

Величины $\xi_1, \ldots, \xi_n$ *независимы в совокупности*, если

$$F_{\xi_1, \ldots, \xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n).$$

Две *группы величин* $(\xi_1, \ldots, \xi_m)$ *и* $(\eta_1, \ldots, \eta_n)$ *независимы*, если

$$F_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m, y_1, \ldots, y_n) = F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m) F_{\eta_1, \ldots, \eta_n}(y_1, \ldots, y_n).$$

Здесь $F_{\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n}$, $F_{\xi_1, \ldots, \xi_m}$, $F_{\eta_1, \ldots, \eta_n}$ — совместные функции распределения величин $(\xi_1, \ldots, \eta_n)$, $(\xi_1, \ldots, \xi_m)$, $(\eta_1, \ldots, \eta_n)$ соответственно. Если группы величин $(\xi_1, \ldots, \xi_m)$ и $(\eta_1, \ldots, \eta_n)$ независимы, то

$$F_{\xi_1, \ldots, \xi_m | \eta_1, \ldots, \eta_n}(x_1, \ldots, x_m | y_1, \ldots, y_n) = F_{\xi_1, \ldots, \xi_m}(x_1, \ldots, x_m).$$

Обозначим через $\mathfrak{A}(A_1, \ldots, A_n)$ наименьшую алгебру, которая содержит множества $A_1, \ldots, A_n$. События $A_1, \ldots, A_n$ называются *независимыми* (в совокупности), если для всех $k = 1, 2, \ldots, n$ событие A_k не зависит от алгебры $\mathfrak{A}(A_1, \ldots, A_{k-1}, A_{k+1}, \ldots, A_n)$. Если события A_k независимы, то

$$\mathsf{P}\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} \mathsf{P}(A_k).$$

События $A_1, \ldots, A_n$ независимы тогда и только тогда, когда для всех $1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n$, $m \leqslant n$,

$$\mathsf{P}\left(\bigcap_{k=1}^{m} A_{i_k}\right) = \prod_{k=1}^{m} \mathsf{P}(A_{i_k}).$$

Пример. Пусть эксперимент состоит в выборе наудачу одного из четырех шаров. Предположим, что три из них занумерованы цифрами 1, 2, 3, а на четвертом шаре имеются все эти три цифры. Обозначим через A_i $(i = 1, 2, 3)$ событие, состоящее в том, что на выбранном шаре имеется цифра i. Очевидно, что

$$\mathsf{P}(A_1) = \mathsf{P}(A_2) = \mathsf{P}(A_3) = 1/2,$$
$$\mathsf{P}(A_1 \cap A_2) = \mathsf{P}(A_1 \cap A_3) = \mathsf{P}(A_2 \cap A_3) = 1/4,$$
$$\mathsf{P}(A_1 \cap A_2 \cap A_3) = 1/4,$$

так что события A_1, A_2, A_3 попарно независимы, но не являются таковыми в совокупности.

Литература: [19, 23, 44, 68, 89].

Глава 2. ПОСЛЕДОВАТЕЛЬНОСТИ НЕЗАВИСИМЫХ СОБЫТИЙ И ВЕЛИЧИН

2.1. Закон нуля и единицы

2.1.1. Теорема Бореля — Кантелли. Пусть $\{A_n, n \geqslant 1\}$ — последовательность событий (предполагается фиксированным вероятностное пространство $\{\Omega, \mathfrak{A}, \mathsf{P}\}$ и $A_n \in \mathfrak{A}$). Событие, состоящее в том, что происходит бесконечное число событий A_k, называется *верхним пределом* последовательности $\{A_n, n \geqslant 1\}$ и обозначается $\overline{\lim} A_n$.

Нижним пределом $\underline{\lim}\, A_n$ последовательности $\{A_n,\ n \geqslant 1\}$ называется событие, состоящее в том, что не происходит лишь конечное число событий A_k. Имеют место равенства

$$\overline{\lim}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \quad \underline{\lim}\, A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Последовательность событий $\{A_n,\ n \geqslant 1\}$ называется *последовательностью независимых событий*, если для каждого n события $A_1, A_2, \ldots, A_n$ независимы, т. е.

$$\mathsf{P}\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} \mathsf{P}\left(A_{i_k}\right)$$

при любых $n = 1, 2, \ldots;\ 1 \leqslant i_1 < i_2 < \ldots < i_n < \infty$.

Теорема Бореля — Кантелли. 1. *Если последовательность событий* $\{A_n,\ n \geqslant 1\}$ *такова, что* $\sum\limits_n \mathsf{P}(A_n) < \infty$, *то* $\mathsf{P}(\overline{\lim}\, A_n) = 0$.

2. *Если для последовательности независимых событий* $\{A_n,\ n \geqslant 1\}$ $\sum\limits_n \mathsf{P}(A_n) = +\infty$, *то* $\mathsf{P}(\overline{\lim}\, A_n) = 1$.

Из этой теоремы следует, что для последовательности независимых событий $\{A_n,\ n \geqslant 1\}$ событие $\overline{\lim}\, A_n$ происходит с вероятностью либо 0, либо 1.

2.1.2. Закон нуля и единицы Колмогорова. Пусть $\{\mathfrak{A}_n,\ n \geqslant 1\}$ — последовательность σ-алгебр событий ($\mathfrak{A}_n \subset \mathfrak{A}$). Обозначим через $\bigcup\limits_{n=m}^{\infty} \mathfrak{A}_n$ минимальную σ-алгебру, содержащую все σ-алгебры $\mathfrak{A}_n$, $n \geqslant m$, а через $\overline{\lim}\, \mathfrak{A}_n$ — σ-алгебру $\bigcap\limits_{m=1}^{\infty} \bigcup\limits_{n=m}^{\infty} \mathfrak{A}_n$, т. е. σ-алгебру событий, входящих в σ-алгебры $\bigcup\limits_{n=m}^{\infty} \mathfrak{A}_n$ при всех $m = 1, 2, \ldots$.

Последовательность σ-алгебр $\{\mathfrak{A}_n,\ n \geqslant 1\}$ называется *последовательностью независимых σ-алгебр*, если всякая последовательность событий $\{A_n,\ n \geqslant 1\}$ такая, что $A_n \in \mathfrak{A}_n$, является последовательностью независимых событий.

Теорема. *Если* $\{\mathfrak{A}_n,\ n \geqslant 1\}$ — *последовательность независимых σ-алгебр, то всякое событие из σ-алгебры* $\overline{\lim}\, \mathfrak{A}_n$ *имеет вероятность либо* 0, *либо* 1.

Последовательность случайных величин $\{\xi_k,\ k \geqslant 1\}$ называется *последовательностью независимых случайных величин*, если для всех действительных чисел x_k ($k = 1, 2, \ldots$) последовательность событий $\{\xi_k < x_k\}$ ($k = 1, 2, \ldots$) является последовательностью независимых событий. Если $\{\xi_n,\ n \geqslant 1\}$ — последовательность независимых случайных величин, а $f(x_1, \ldots, x_n, \ldots)$ — такая функция бесконечного числа переменных x_k, что величина $\zeta = f(\xi_1, \ldots, \xi_n, \ldots)$ измерима относительно $\overline{\lim}\, \mathfrak{A}_n$, то из закона нуля и единицы следует, что

ζ с вероятностью 1 принимает одно значение. В частности, для независимых случайных величин $\{\xi_k,\ k \geqslant 1\}$ величины $\varliminf \xi_n$, $\varlimsup \xi_n$, $\varliminf \frac{1}{n} \sum_{k=1}^{n} \xi_k$, $\varlimsup \frac{1}{n} \sum_{k=1}^{n} \xi_k$ с вероятностью 1 постоянны. Для последовательности независимых случайных величин $\{\xi_k,\ k \geqslant 1\}$ ряд $\sum_{k=1}^{\infty} \xi_k$ либо с вероятностью 1 сходится, либо с вероятностью 1 расходится.

2.2. Схема Бернулли

2.2.1. Биномиальное распределение (распределение Бернулли). Пусть $A_1, A_2, \ldots, A_n$ — независимые события, каждое из которых имеет вероятность p. Положим $q = 1 - p$, и пусть ν — число тех событий A_k из совокупности $\{A_1, A_2, \ldots, A_n\}$, которые произошли. Тогда ν — биномиально распределенная случайная величина (иначе говоря, величина ν имеет распределение Бернулли), т. е.

$$B_p(n,\ m) = \mathsf{P}\{\nu = m\} = C_n^m p^m q^{n-m}, \qquad m = 0,\ 1,\ 2,\ \ldots,\ n. \tag{2.1}$$

На практике очень часто используется *схема Бернулли*: производится n независимых (в теоретико-вероятностном смысле) испытаний (экспериментов), в каждом из которых с вероятностью p может произойти некоторое фиксированное событие A. Тогда вероятность того, что в серии из n испытаний событие A произойдет ровно m раз, равна $B_p(n,\ m)$ $(m = 0, 1, 2, \ldots, n)$.

Пример. Предположим, что каждое изделие, выпускаемое на некотором предприятии, может оказаться бракованным с вероятностью p. Тогда вероятность того, что в партии из n изделий окажется не более чем m бракованных, равна $\sum_{k=0}^{m} B_p(n,\ k)$.

Среднее число появлений события A в серии из n испытаний равно $\sum_{m=0}^{n} m B_p(n,\ m) = np$.

Дисперсия числа появлений события A в n испытаниях равна $\sum_{m=0}^{n} m^2 B_p(n,\ m) - n^2p^2 = npq$.

Если m изменяется от 0 до n, то вероятности $B_p(n,\ m)$ сначала возрастают, а затем убывают, достигая наибольшего значения при $m = [np + p]$, если число $np + p$ нецелое; если же $np + p$ целое, то имеются две максимальные вероятности: $B_p(n,\ np + p)$ и $B_p(n,\ np - q)$.

2.2.2. Полиномиальное распределение. Предположим, что в результате каждого из n независимых испытаний может произойти одно из m событий $A_1, A_2, \ldots, A_m$ с вероятностями $p_1, p_2, \ldots, p_m$ соответственно $(p_1 + p_2 + \ldots + p_m = 1,\ p_i \geqslant 0)$. Обозначим через ν_i

число тех испытаний, в которых произошло событие A_i, $i = 1, 2, \ldots$ $\ldots, m$. Тогда

$$\mathsf{P}\{\nu_1 = i_1, \nu_2 = i_2, \ldots, \nu_m = i_m\} = \frac{n!}{i_1! i_2! \ldots i_m!} p_1^{i_1} p_2^{i_2} \ldots p_m^{i_m}, \qquad (2.2)$$

где $i_k \geqslant 0$ и $i_1 + i_2 + \ldots + i_m = n$. Это распределение называется *полиномиальным*. При $m = 2$ оно превращается в биномиальное.

2.2.3. Биномиальное приближение для гипергеометрического распределения. Предположим, что из совокупности n предметов, из которых n_1 предметов 1-го вида и n_2 предметов 2-го вида ($n_1 + n_2 =$ $= n$), производится выборка без возвращения группы k предметов ($k \leqslant n$). Обозначим через ν число предметов 1-го вида в выборке. Распределение величины ν называется *гипергеометрическим* и вычисляется по формуле

$$P_{n_1 n}(k, m) = \mathsf{P}\{\nu = m\} = \frac{C_{n_1}^m C_{n_2}^{k-m}}{C_n^k}, \qquad m = 0, 1, 2, \ldots, \min(k, n_1). \qquad (2.3)$$

При $n \to \infty$ и $n_1/n \to p$ справедливо соотношение

$$P_{n_1 n}(k, m) \to B_p(k, m),$$

так что схему Бернулли можно рассматривать как выбор без возвращения из бесконечной совокупности предметов.

Аналогично пусть n — совокупность предметов, из которых имеется n_i предметов i-го вида ($i = 1, 2, \ldots, r$), и из этой совокупности производится выбор без возвращения группы k предметов. Обозначив через ν_i число предметов i-го вида в выборке, будем иметь

$$\mathsf{P}\{\nu_1 = m_1, \ldots, \nu_r = m_r\} = \frac{C_{n_1}^{m_1} C_{n_2}^{m_2} \ldots C_{n_r}^{m_r}}{C_n^k}.$$

Здесь $m_1 + m_2 + \ldots + m_r = k$, $0 \leqslant m_i \leqslant n_i$, $n_1 + n_2 + \ldots + n_r = n$. При $n \to \infty$ и $n_1/n \to p_1, \ldots, n_r/n \to p_r$

$$\mathsf{P}\{\nu_1 = m_1, \ldots, \nu_r = m_r\} \to \frac{k!}{m_1! m_2! \ldots m_r!} p_1^{m_1} p_2^{m_2} \ldots p_r^{m_r},$$

т. е. в пределе получаются полиномиальные вероятности.

2.3. Предельные теоремы для схемы Бернулли

2.3.1. Закон больших чисел. Обозначим через ν_n число появлений события A в серии из n независимых испытаний. Пусть p — вероятность появления события A в одном испытании. Тогда при любом $\varepsilon > 0$

$$\lim_{n \to \infty} \mathsf{P}\left\{\left|\frac{\nu_n}{n} - p\right| > \varepsilon\right\} = 0. \qquad (3.1)$$

Говорят, что последовательность случайных величин $\{\xi_n, n \geqslant 1\}$ *сходится по вероятности* к случайной величине ξ, если для каждого $\varepsilon > 0$

$$\lim_{n \to \infty} \mathsf{P}\{|\xi_n - \xi| > \varepsilon\} = 0.$$

Таким образом, предыдущее утверждение означает, что частота ν_n/n появления события A в серии из n независимых испытаний сходится по вероятности к вероятности p появления события A в одном испытании.

2.3.2. Нормальное приближение для биномиального распределения. При больших n подсчет вероятностей $B_p(n, m)$ может оказаться весьма затруднительным. Поэтому очень важной задачей является отыскание асимптотических формул для величин $B_p(n, m)$ при $n \to \infty$.

Теорема Муавра — Лапласа. *Положим $np = a_n$, $npq = B_n^2$, $x_{n,m} = (m - a_n)/B_n$. Если $B_n \to \infty$ при $n \to \infty$ и $x_{n,m}$ ограничено, то*

$$\lim_{n\to\infty} \frac{B_n B_p(n, m)}{\varphi(x_{n,m})} = 1, \tag{3.2}$$

где $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Утверждение теоремы справедливо также и в случае, когда $x_{n,m}$ ограничено, а p и q зависят от n, но так, что $B_n \to \infty$.

Интегральная предельная теорема Муавра — Лапласа. *В условиях предыдущей теоремы для произвольных $x_1 < x_2$ справедлива асимптотическая формула*

$$\mathbf{P}\left\{x_1 < \frac{\nu_n - a_n}{B_n} < x_2\right\} \sim \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-t^2/2}\, dt, \quad n \to \infty, \tag{3.3}$$

где ν_n — число появлений события A в серии из n независимых испытаний, если в одном испытании событие A происходит с вероятностью p.

Другими словами, распределение величины $(\nu_n - np)/\sqrt{npq}$ асимптотически нормально со средним 0 и дисперсией 1.

2.3.3. Пуассоновское приближение биномиального распределения. Пусть производится серия из n независимых испытаний и вероятность p осуществления события A в одном испытании зависит от n так, что последовательность $b_n = np_nq_n$ ограничена.

Если $b_n \to 0$, то либо $p_n \to 0$, либо $q_n \to 0$ при $n \to \infty$. В первом случае

$$B_{p_n}(n, 0) = q_n^n \sim \left(1 - \frac{b_n}{n}\right)^n \sim e^{-b_n} \sim 1.$$

Поэтому

$$\sum_{m=1}^{n} B_{p_n}(n, m) = 1 - B_{p_n}(n, 0) \to 0 \text{ при } n \to \infty.$$

Во втором случае $B_{p_n}(n, n) \sim 1$ и $\sum\limits_{m=0}^{n-1} B_{p_n}(n, m) \to 0$.

Рассмотрим теперь случай, когда существуют такие постоянные C_1 и C_2, что $0 < C_1 < C_2 < \infty$ и $C_1 < b_n < C_2$ при всех n. Тогда либо $p_n \to 0$, либо $q_n \to 0$. Рассмотрим первый случай, ибо второй сводится к первому в силу формулы

$$B_p(n, m) = B_q(n, n-m).$$

Так как $p_n \to 0$, то $q_n \to 1$ и, стало быть, $b_n \sim a_n = np_n$.

Теорема Пуассона. *Если при некоторых постоянных C_1 и C_2 $0 < C_1 < a_n < C_2 < \infty$, то для всех $m = 0, 1, \ldots$*

$$B_{p_n}(n, m) \sim \frac{a_n^m}{m!} e^{-a_n}. \tag{3.4}$$

В частности, если $a_n \to a$ при $n \to \infty$, то величина ν_n — число появлений события A в серии из n независимых испытаний — при указанных предположениях имеет асимптотически пуассоновское распределение с параметром a.

Формула (3.4) выполнена и в том случае, когда $np_n^2 \to 0$ $m^2/n \to 0$ при $n \to \infty$.

2.3.4. Асимптотическое поведение полиномиальных вероятностей. Положим

$$p_n(m_1, m_2, \ldots, m_r) = \frac{n!}{m_1! m_2! \ldots m_r!} p_1^{m_1} p_2^{m_2} \ldots p_r^{m_r},$$

где $m_1 + m_2 + \ldots + m_r = n$, $m_i \geqslant 0$; $p_1 + p_2 + \ldots + p_r = 1$, $p_i \geqslant 0$; r — некоторое фиксированное целое число ($r \geqslant 2$). Если величины $np_1 = a_1$, $np_2 = a_2$, ..., $np_{r-1} = a_{r-1}$ ограничены, то при $n \to \infty$

$$p_n(m_1, m_2, \ldots, m_r) \sim$$
$$\sim \frac{1}{m_1! m_2! \ldots m_{r-1}!} a_1^{m_1} a_2^{m_2} \ldots a_{r-1}^{m_{r-1}} e^{-(a_1 + \ldots + a_{r-1})}. \tag{3.5}$$

Здесь $m_1, m_2, \ldots, m_{r-1}$ — произвольные неотрицательные целые числа, $m_r = n - \sum\limits_{i=1}^{r-1} m_i$.

Если r также может изменяться вместе с n, $n \to \infty$, и все величины $a_1, a_2, \ldots, a_{r-1}, a_r = np_r$ ограничены, то для произвольных целых неотрицательных $m_1, m_2, \ldots, m_r$ при $n \to \infty$

$$p_n(m_1, m_2, \ldots, m_r) \sim \frac{\sqrt{2\pi n}\, e^{-n}}{m_1! \ldots m_r!} a_1^{m_1} \ldots a_r^{m_r}. \tag{3.6}$$

2.4. Последовательности независимых случайных величин. Закон больших чисел

2.4.1. Критерий независимости последовательности случайных величин. В п. 2.1.2 введено понятие последовательности независимых случайных величин. Пусть $\{\xi_n, n \geqslant 1\}$ — такая последовательность. Обозначим через $\mathfrak{A}_k$ σ-алгебру событий вида $\{\xi_k \in A\}$, где A — произвольное борелевское множество на прямой. Из определения последовательности независимых случайных величин следует, что последовательность σ-алгебр $\{\mathfrak{A}_n, n \geqslant 1\}$ является последовательностью независимых σ-алгебр. Приведем один критерий независимости случайных величин.

Теорема 1. *Для того чтобы величины $\xi_1, \xi_2, \ldots, \xi_n$ были независимы, необходимо, чтобы для всех ограниченных борелевских функций*

$$\mathsf{M}(g_1(\xi_1) \ldots g_n(\xi_n)) = \mathsf{M}g_1(\xi_1) \ldots \mathsf{M}g_n(\xi_n),$$

и достаточно, чтобы это равенство выполнялось для всех непрерывных ограниченных функций $g_1, \ldots, g_n$.

В частности, если $\xi_1, \xi_2, \ldots, \xi_n$ независимы и $\mathsf{M}\xi_k$ $(k = 1, \ldots, n)$ существует, то

$$\mathsf{M} \prod_{i=1}^{n} \xi_i = \prod_{i=1}^{n} \mathsf{M}\xi_i.$$

Для независимых случайных величин $\xi_1, \ldots, \xi_n$

$$\mathsf{D} \sum_{k=1}^{n} \xi_k = \sum_{k=1}^{n} \mathsf{D}\xi_k,$$

если только существуют $\mathsf{D}\xi_k$ $(k = 1, 2, \ldots, n)$.

Если независимы две группы случайных величин $(\xi_1, \ldots, \xi_n)$ и $(\zeta_1, \ldots, \zeta_m)$, то для произвольных борелевских функций $f(x_1, \ldots, x_n)$ и $g(y_1, \ldots, y_m)$ случайные величины $\xi = f(\xi_1, \ldots, \xi_n)$ и $\zeta = g(\zeta_1, \ldots, \zeta_m)$ независимы.

2.4.2. Неравенство Чебышева. Пусть ξ — произвольная случайная величина и $g(x)$ — неотрицательная четная и неубывающая на $[0, +\infty)$ функция. Тогда для всех $a \geqslant 0$

$$\frac{\mathsf{M}g(\xi) - g(a)}{\text{п. н. } \sup g(\xi)} \leqslant \mathsf{P}\{|\xi| \geqslant a\} \leqslant \frac{\mathsf{M}g(\xi)}{g(a)}. \tag{4.1}$$

Величина, стоящая в знаменателе левой части неравенства, называется *почти наверное верхней гранью случайной величины* $g(\xi)$ и определяется так:

$$\text{п. н. } \sup g(\xi) = \inf\{C: C \geqslant 0 \text{ и } \mathsf{P}\{g(\xi) > C\} = 0\}.$$

Полагая в неравенстве (4.1) $g(x) = |x|^r$ $(r > 0)$, получаем

$$\frac{\mathsf{M}|\xi|^r - a^r}{\text{п. н. } \sup |\xi|^r} \leqslant \mathsf{P}\{|\xi| \geqslant a\} \leqslant \frac{\mathsf{M}|\xi|^r}{a^r}. \tag{4.2}$$

Применив это неравенство к величине $\xi - \mathsf{M}\xi$, получим

$$\frac{\mathsf{M}|\xi - \mathsf{M}\xi|^r - a^r}{\text{п. н. } \sup |\xi - \mathsf{M}\xi|^r} \leqslant \mathsf{P}\{|\xi - \mathsf{M}\xi| \geqslant a\} \leqslant \frac{\mathsf{M}|\xi - \mathsf{M}\xi|^r}{a^r}. \tag{4.3}$$

При $r = 2$ отсюда следует *неравенство Чебышева*:

$$\mathsf{P}\{|\xi - \mathsf{M}\xi| \geqslant a\} \leqslant \frac{\mathsf{D}\xi}{a^2}. \tag{4.4}$$

2.4.3. Сходимость в среднем. Последовательность случайных величин $\{\xi_n, n \geqslant 1\}$ *сходится в среднем порядка* r $(r > 0)$ к случайной величине ξ, если

$$\lim_{n\to\infty} \mathsf{M}\,|\xi_n - \xi|^r = 0.$$

Сходимость в среднем порядка 2 называется *сходимостью в среднеквадратическом.* Если $\xi_n \to \xi$ в среднем порядка r, то $\xi_n \to \xi$ в среднем порядка r' при всех $r' \leqslant r$ и $\mathsf{M}\,|\xi_n|^{r'} \to \mathsf{M}\,|\xi|^{r'}$. Применяя правое неравенство в (4.2) к величине $\xi_n - \xi$, получаем, что из сходимости в среднем порядка r $(r > 0)$ следует *сходимость по вероятности.* Левое неравенство в (4.2) дает обратное утверждение: если последовательность $\{\xi_n, n \geqslant 1\}$ сходится по вероятности и п. н. равномерно ограничена, то она сходится и в среднем порядка r при любом $r > 0$.

Последовательность случайных величин $\{\xi_n, n \geqslant 1\}$ называется *равномерно интегрируемой*, если

$$\lim_{a\to\infty} \sup_n \int\limits_{\{|\xi_n|>a\}} |\xi_n|\, d\mathsf{P}(\omega) = 0.$$

Для того чтобы $\xi_n \to \xi$ в среднем порядка r, необходимо и достаточно, чтобы $\xi_n \to \xi$ по вероятности и последовательность $|\xi_n|^r$ была равномерно интегрируемой.

Пусть $g(x)$ — произвольная борелевская функция, заданная на прямой, и пусть A — множество точек разрыва $g(x)$. Если последовательность $\{\xi_n, n \geqslant 1\}$ сходится по вероятности к случайной величине ξ и $\mathsf{P}\{\xi \in A\} = 0$, то $g(\xi_n)$ сходится по вероятности к величине $g(\xi)$. Если, кроме того, $\sup\limits_n \mathsf{M}\,|g(\xi_n)|^{r_0} = C < \infty$ при некотором $r_0 > 0$, то $g(\xi_n) \to g(\xi)$ в среднем порядка r при любом $r < r_0$. В частности, если $\xi_n \to \xi$ по вероятности, то $\lim\limits_{n\to\infty} \mathsf{P}\{\xi_n < x\} = \mathsf{P}\{\xi < x\}$ для любого такого x, что $\mathsf{P}\{\xi = x\} = 0$.

2.4.4. Закон больших чисел. Так называют теоремы, дающие условия, при которых

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k \to 0$$

по вероятности. В п. 2.3.1 приведено утверждение, относящееся к схеме Бернулли. Заметим, что если ν_n — число появлений события A в серии из n независимых испытаний, то $\nu_n = \sum\limits_{i=1}^{n} \xi_i$, где ξ_i — случайная величина, равная 1, если в i-м испытании событие A произошло, и равная 0 в противном случае. Тогда

$$\frac{\nu_n}{n} = \frac{1}{n}\sum_{k=1}^{n} \xi_k, \quad p = \frac{1}{n}\sum_{k=1}^{n} \mathsf{M}\xi_k.$$

Рассмотрим более общую задачу. Пусть заданы последовательность независимых случайных величин $\{\xi_n,\ n=1,\ 2,\ \ldots\}$ и числовая последовательность $\{\beta_n,\ n=1,\ 2,\ \ldots\}$ такая, что $\beta_n\to\infty$ при $n\to\infty$. При каких условиях существует такая числовая последовательность $\{\alpha_n,\ n=1,\ 2,\ \ldots\}$, что при $n\to\infty$

$$\frac{1}{\beta_n}\sum_{k=1}^{n}\xi_k-\alpha_n\to 0$$

по вероятности? Ответ на этот вопрос дает следующая теорема.

Теорема 2. *Пусть $\{\xi_n,\ n\geqslant 1\}$ — последовательность независимых случайных величин $F_n(x)=\mathsf{P}\{\xi_n<x\}$. Обозначим через m_n медиану случайной величины ξ_n, т. е. любое из чисел, удовлетворяющих неравенствам $\mathsf{P}\{\xi_n\geqslant m_n\}\geqslant 1/2$ и $\mathsf{P}\{\xi_n\leqslant m_n\}\geqslant 1/2$. Для того чтобы существовала последовательность постоянных $\{\alpha_n,\ n\geqslant 1\}$ такая, что при $n\to\infty$*

$$\frac{1}{\beta_n}\sum_{k=1}^{n}\xi_k-\alpha_n\to 0$$

по вероятности, необходимо и достаточно выполнения условия

$$\sum_{k=1}^{n}\int_{-\infty}^{\infty}\frac{(x-m_k)^2}{\beta_k^2+(x-m_k)^2}\,dF_k(x)\to 0. \tag{4.5}$$

Если это условие выполнено, то

$$\alpha_n=\frac{1}{\beta_n}\sum_{k=1}^{n}\left(m_k+\int\limits_{|x-m_k|<\tau\beta_n}(x-m_k)\,dF_k(x)\right)+o(1),$$

где τ — произвольная положительная постоянная.

Следующая теорема дает условия применимости к данной последовательности независимых случайных величин закона больших чисел в классической форме.

Теорема 3. *Пусть $\{\xi_n,\ n\geqslant 1\}$ — последовательность независимых случайных величин, $F_k(x)=\mathsf{P}\{\xi_k<x\}$. Предположим, что $\mathsf{M}|\xi_k|<\infty$ $(k=1,\ 2,\ \ldots)$. Для того чтобы при $n\to\infty$*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k-\frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\xi_k\to 0$$

по вероятности, необходимо и достаточно выполнения следующих условий:

1) $$\sum_{k=1}^{n}\int\limits_{|x-\mathsf{M}\xi_k|\geqslant n}dF_k(x)\to 0;$$

$$2)\ \frac{1}{n}\sum_{k=1}^{n}\int\limits_{|x-\mathsf{M}\xi_k|<n}(x-\mathsf{M}\xi_k)\,dF_k(x)\to 0;$$

$$3)\ \frac{1}{n^2}\sum_{k=1}^{n}\left\{\int\limits_{|x-\mathsf{M}\xi_k|<n}(x-\mathsf{M}\xi_k)^2\,dF_k(x)-\right.$$
$$\left.-\left(\int\limits_{|x-\mathsf{M}\xi_k|<n}(x-\mathsf{M}\xi_k)\,dF_k(x)\right)^2\right\}\to 0.$$

Простое условие применимости закона больших чисел содержится в следующей теореме.

Теорема 4. *Если последовательность независимых случайных величин* $\{\xi_n,\ n\geqslant 1\}$ *такова, что* $\mathsf{D}\xi_n$ *существует и* $\frac{\mathsf{D}\xi_n}{n}\to 0$ *при* $n\to\infty$, *то*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k-\frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\xi_k\to 0$$

по вероятности.

Если величины ξ_n $(n=1,2,\ldots)$ имеют одну и ту же функцию распределения $F(x)=\mathsf{P}\{\xi_n<x\}$, то они называются *одинаково распределенными.*

Теорема 5. *Если* $\{\xi_n,\ n\geqslant 1\}$ — *последовательность независимых, одинаково распределенных случайных величин и если существует математическое ожидание* $\mathsf{M}\xi_n=a$, *то при* $n\to\infty$

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k\to a$$

по вероятности.

2.5. Неравенство Колмогорова. Усиленный закон больших чисел

2.5.1. Неравенства типа неравенства Колмогорова.

Теорема 1 (Колмогорова). *Если случайные величины* $\xi_1,\xi_2,\ldots,\xi_n$ *независимы и* $\mathsf{D}\xi_k<\infty$, *то при любом* $a>0$

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}(\xi_i-\mathsf{M}\xi_i)\right|\geqslant a\right\}\leqslant\frac{1}{a^2}\sum_{k=1}^{n}\mathsf{D}\xi_k.\tag{5.1}$$

Если, кроме того, $|\xi_k|\leqslant C<\infty$ *для всех* $k=1,2,\ldots$, *то*

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}(\xi_i-\mathsf{M}\xi_i)\right|\geqslant a\right\}\geqslant 1-(a+2C)^2\Big/\sum_{k=1}^{n}\mathsf{D}\xi_k.\tag{5.2}$$

Распределение величины ξ называется *симметричным*, если величины ξ и $-\xi$ одинаково распределены. Величина ξ с симметричным распределением называется *симметричной.* Если $F(x)$ — функция распределения симметричной случайной величины, то $F(x) = 1 - F(-x+0)$.

Теорема 2. *Если величины* $\xi_1, \xi_2, \ldots, \xi_n$ *независимы и симметричны, то*

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}\xi_i\right|\geqslant a\right\}\leqslant 2\mathsf{P}\left\{\left|\sum_{i=1}^{n}\xi_i\right|\geqslant a\right\}. \tag{5.3}$$

Теорема 3. *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n$ *независимы и для некоторых* $a>0$ *и* $\alpha<1$

$$\mathsf{P}\left\{\left|\sum_{i=k+1}^{n}\xi_i\right|\geqslant a\right\}\leqslant\alpha,\qquad k=1, 2, \ldots, n-1,$$

то при всех $C>0$

$$\mathsf{P}\left\{\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}\xi_i\right|\geqslant a+C\right\}\leqslant\frac{1}{1-\alpha}\mathsf{P}\left\{\left|\sum_{k=1}^{n}\xi_k\right|\geqslant C\right\}. \tag{5.4}$$

2.5.2. Сходимость почти наверное. Последовательность случайных величин $\{\xi_n, n\geqslant 1\}$ *сходится почти наверное* (или с вероятностью 1) к случайной величине ξ ($\xi_n\to\xi$ п. н. или $\xi_n\to\xi$ (mod P)), если

$$\mathsf{P}\{\omega: |\xi_n(\omega)-\xi(\omega)|\to 0\}=1.$$

Последовательность $\xi_n\to\xi$ п. н. (почти наверное) тогда и только тогда, когда для всякого $\varepsilon>0$

$$\mathsf{P}\left\{\bigcup_{m=1}^{\infty}\{|\xi_{n+m}-\xi|\geqslant\varepsilon\}\right\}\to 0$$

при $n\to\infty$.

Из сходимости п. н. следует сходимость по вероятности. Обратное, вообще говоря, не верно. Однако всякая последовательность ξ_n, сходящаяся к ξ по вероятности, содержит подпоследовательность, сходящуюся п. н. Более того, $\xi_n\to\xi$ по вероятности тогда и только тогда, когда всякая подпоследовательность последовательности ξ_n содержит подпоследовательность, сходящуюся к ξ п. н. Из теоремы Бореля — Кантелли следует, что достаточным условием сходимости ξ_k к ξ с вероятностью 1 является сходимость для каждого $\varepsilon>0$ ряда

$$\sum_{n=1}^{\infty}\mathsf{P}\{|\xi_n-\xi|>\varepsilon\}.$$

2.5.3. Усиленный закон больших чисел. Это одна из форм закона больших чисел, в которой вместо сходимости по вероятности утверждается сходимость п. н. (с вероятностью 1).

Наиболее общо задача может быть сформулирована так: дана последовательность независимых случайных величин $\{\xi_n, n=1, 2, \ldots\}$ и числовая последовательность β_n такая, что $\beta_n\to\infty$

при $n \to \infty$. При каких условиях существует такая числовая последовательность $\{\alpha_n, n = 1, 2, \ldots\}$, что при $n \to \infty$

$$\frac{1}{\beta_n} \sum_{k=1}^{n} \xi_k - \alpha_n \to 0$$

с вероятностью 1?

Теорема 4. *Пусть* $\{\xi_n, n = 1, 2, \ldots\}$ — *последовательность независимых случайных величин. Предположим, что* $\beta_n \to \infty$ *при* $n \to \infty$, *и пусть существуют такая подпоследовательность* β_{n_k} *и такие числа* C_1 *и* C_2, *что для всех достаточно больших* k

$$1 < C_1 \leqslant \frac{\beta_{n_{k+1}}}{\beta_{n_k}} \leqslant C_2 < \infty.$$

Положим $S_n = \sum_{k=1}^{n} \xi_k$, $T_k = \dfrac{S_{n_k} - S_{n_{k-1}}}{\beta_{n_k}}$, $k = 1, 2, \ldots, S_{n_0} = 0$.

Для того чтобы при $n \to \infty$

$$\frac{1}{\beta_n}(S_n - mS_n) \to 0$$

почти наверное, необходимо и достаточно, чтобы выполнялось одно из условий:

1) $T_k - mT_k \to 0$ *п. н. при* $k \to \infty$;

2) $\sum_{k=1}^{\infty} \mathsf{P}\{|T_k - mT_k| \geqslant \varepsilon\} < \infty$ *для любого* $\varepsilon > 0$.

Здесь mS_k и mT_k обозначают медианы случайных величин S_k и T_k соответственно.

Удобное для проверки достаточное условие применимости усиленного закона больших чисел к данной последовательности независимых случайных величин содержится в следующей теореме.

Теорема 5. *Если* $\{\xi_n, n \geqslant 1\}$ — *последовательность независимых случайных величин, для которой ряд* $\sum_{k=1}^{\infty} \dfrac{\mathsf{D}\xi_k}{k^2}$ *сходится, то с вероятностью* 1 *при* $n \to \infty$

$$\frac{1}{n} \sum_{k=1}^{n} \xi_k - \frac{1}{n} \sum_{k=1}^{n} \mathsf{M}\xi_k \to 0.$$

Следующая теорема относится к одинаково распределенным слагаемым.

Теорема 6. *Если* $\{\xi_n, n \geqslant 1\}$ — *последовательность независимых, одинаково распределенных случайных величин, для которых*

$\mathsf{M}\xi_n$ конечно, то с вероятностью 1

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \to \mathsf{M}\xi_1.$$

Если же величины ξ_n не имеют конечного математического ожидания, то последовательность $\frac{1}{n}\sum_{k=1}^{n}\xi_k$ с вероятностью 1 не ограничена.

Следствие. *Если ν_n — число появлений события A в серии из n независимых испытаний, то $\nu_n/n \to p$ с вероятностью 1 при $n\to\infty$, где p — вероятность появления события A в одном испытании.*

2.6. Закон повторного логарифма

Пусть задана последовательность независимых случайных величин $\{\xi_n, n\geqslant 1\}$ с нулевыми математическими ожиданиями: $\mathsf{M}\xi_n=0$ $(n=1, 2, \ldots)$. Приведенные в § 2.5 теоремы типа усиленного закона больших чисел дают информацию о порядке роста сумм $S_n=\sum_{k=1}^{n}\xi_k$ при $n\to\infty$. Так, если выполнены условия теоремы 5 (п. 2.5.3), то $S_n/n\to 0$ при $n\to\infty$ с вероятностью 1. Оказывается, что при некоторых дополнительных предположениях о случайных величинах ξ_k можно получить более точную информацию об асимптотическом поведении таких сумм.

Говорят, что числовая последовательность $\{\alpha_n, n\geqslant 1\}$ принадлежит *нижнему классу*, если P {найдется бесконечно много значений $n\geqslant 1$, для которых $S_n>\alpha_n\}=1$. Если же эта вероятность равна нулю, то скажем, что последовательность $\{\alpha_n, n\geqslant 1\}$ принадлежит *верхнему классу*.

Ниже будут приведены теоремы, дающие условия, при которых с вероятностью 1 $\limsup c_n^{-1}S_n=1$, где $\{c_n, n\geqslant 1\}$ — некоторая неограниченно возрастающая числовая последовательность. Это будет означать, что при любом $\varepsilon>0$ числовая последовательность $\{(1-\varepsilon)c_n, n\geqslant 1\}$ принадлежит нижнему классу, а последовательность $\{(1+\varepsilon)c_n, n\geqslant 1\}$ — верхнему классу. Тем самым последовательность $\{c_n, n\geqslant 1\}$ в некотором смысле характеризует порядок роста сумм S_n при $n\to\infty$.

Теорема 1. *Пусть $\{\xi_n, n\geqslant 1\}$ — последовательность независимых случайных величин, для которых*

$$\mathsf{M}\xi_k=0,\quad \mathsf{M}\xi_k^2<\infty.$$

Обозначим через $F_k(x)$ функцию распределения величин ξ_k и положим

$$L_k(x)=\int\limits_{|u|\geqslant x}u^2\,dF_k(x),\quad x\geqslant 0;\qquad B_n=\sum_{k=1}^{n}\mathsf{M}\xi_k^2.$$

Предположим, что $B_n \to \infty$, и пусть выполнены условия

$$B_n^{-1} \sum_{k=k_0}^{n} L_k\left(\varepsilon\left(\frac{B_k}{\ln\ln B_k}\right)^{1/2}\right) \to 0, \tag{6.1}$$

$$\sum_{k=k_0}^{\infty} (B_k \ln\ln B_k)^{-1} L_k\left(\varepsilon\left(\frac{B_k}{\ln\ln B_k}\right)^{1/2}\right) < \infty \tag{6.2}$$

для любого $\varepsilon > 0$ и какого-нибудь k_0, при котором $\ln\ln B_{k_0} > 0$.

Тогда с вероятностью 1

$$\limsup_{n\to\infty} \frac{S_n}{(2B_n \ln\ln B_n)^{1/2}} = 1. \tag{6.3}$$

Если для последовательности независимых величин $\{\xi_k,\ k \geqslant 1\}$ выполнено с вероятностью 1 соотношение (6.3), то говорят, что эта последовательность удовлетворяет закону повторного логарифма. Таким образом, условия (6.1), (6.2) являются достаточными для того, чтобы последовательность $\{\xi_k,\ k \geqslant 1\}$ удовлетворяла закону повторного логарифма. Достаточным условием для выполнимости соотношений (6.1) и (6.2) является следующее условие:

$$\sum_k B_k^{-1} L_k\left(\varepsilon\left(\frac{B_k}{\ln\ln B_k}\right)^{1/2}\right) < \infty \tag{6.4}$$

при всяком $\varepsilon > 0$.

Следующая теорема также содержит достаточное условие для того, чтобы последовательность независимых случайных величин удовлетворяла закону повторного логарифма.

Теорема 2. *Пусть задана последовательность независимых случайных величин $\{\xi_k,\ k \geqslant 1\}$, для которой* $\mathsf{M}\xi_k = 0$, $\mathsf{M}\xi_k^2 < \infty$, $F_k(x) = \mathsf{P}\{\xi_k < x\}$, $B_n = \sum_{k=1}^{n} \mathsf{M}\xi_k^2$. *Пусть $B_n \to \infty$ при $n \to \infty$ и*

$$\ln B_n (\ln\ln\ln B_n)^{1+\delta} B_n^{-1} \sum_{k=k_0}^{n} L_k\left(\varepsilon\left(\frac{B_n}{\ln\ln B_n}\right)^{1/2}\right) \to 0$$

при некотором $\delta > 0$ и любом фиксированном $\varepsilon > 0$, где k_0 и L_k таковы, как и в теореме 1. Тогда последовательность $\{\xi_n,\ n \geqslant 1\}$ удовлетворяет закону повторного логарифма.

Теоремы 1, 2 применимы и к величинам $-\xi_k$, и потому в условиях этих теорем

$$\liminf_{n\to\infty} \frac{S_n}{(2B_n \ln\ln B_n)^{1/2}} = -1$$

с вероятностью 1, а следовательно, и

$$\limsup_{n\to\infty} \frac{|S_n|}{(2B_n \ln\ln B_n)^{1/2}} = 1$$

с вероятностью 1.

Следующая теорема относится к одинаково распределенным случайным величинам.

Теорема 3. *Пусть* $\{\xi_k,\ k \geqslant 1\}$ — *последовательность независимых одинаково распределенных случайных величин, для которых* $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = 1$. *Тогда с вероятностью* 1 *при* $n \to \infty$

$$\limsup \frac{S_n}{\sqrt{2n \ln \ln n}} = 1. \tag{6.5}$$

Теорема 3 допускает следующее обращение.

Теорема 4. *Пусть* $\{\xi_n,\ n \geqslant 1\}$ — *последовательность независимых одинаково распределенных случайных величин, для которых с вероятностью* 1 *выполнено условие* (6.5). *Тогда*

$$\mathsf{M}\xi_k = 0, \quad \mathsf{M}\xi_k^2 = 1.$$

2.7. Ряды из независимых случайных величин

Пусть $\{\xi_n,\ n \geqslant 1\}$ — некоторая последовательность случайных величин. Ряд $\sum\limits_{k=1}^{\infty} \xi_k$ называется *сходящимся с вероятностью* 1, если частичные суммы $\sum\limits_{k=1}^{n} \xi_k$ сходятся с вероятностью 1 при $n \to \infty$.

Теорема 1. *Пусть* $\{\xi_n,\ n \geqslant 1\}$ — *последовательность независимых случайных величин. Если сходятся ряды* $\sum\limits_{k=1}^{\infty} \mathsf{M}\xi_k$ *и* $\sum\limits_{k=1}^{\infty} \mathsf{D}\xi_k$, *то ряд* $\sum\limits_{k=1}^{\infty} \xi_k$ *сходится с вероятностью* 1. *Обратно, если величины* ξ_n *с вероятностью* 1 *ограничены* (*при некотором* $C > 0$ *для всех* n $\mathsf{P}\{|\xi_n| > C\} = 0$) *и ряд* $\sum\limits_{k=1}^{\infty} \xi_k$ *сходится с вероятностью* 1, *то сходятся ряды* $\sum\limits_{k=1}^{\infty} \mathsf{M}\xi_k$ *и* $\sum\limits_{k=1}^{\infty} \mathsf{D}\xi_k$.

Теорема 2 (о трех рядах). *Для сходимости с вероятностью* 1 *ряда из независимых случайных величин* $\{\xi_n,\ n \geqslant 1\}$ *необходимо, чтобы для каждого* $C > 0$, *и достаточно, чтобы при некотором* $C > 0$ *сходились ряды*

$$\sum_{k=1}^{\infty} \mathsf{P}\{|\xi_k| > C\}, \quad \sum_{k=1}^{\infty} \mathsf{M}\xi_k(C), \quad \sum_{k=1}^{\infty} \mathsf{D}\xi_k(C),$$

где

$$\xi_k(C) = \begin{cases} \xi_k, & \text{если } |\xi_k| \leqslant C, \\ 0, & \text{если } |\xi_k| > C. \end{cases}$$

Теорема 3. *Если* $\{\xi_n,\ n \geqslant 1\}$ — *последовательность неотрицательных случайных величин, то для сходимости с вероятностью* 1 *ряда* $\sum_{k=1}^{\infty} \xi_k$ *достаточно, чтобы при некотором* $C > 0$, *и необходимо, чтобы при всех* $C > 0$ *сходились ряды*

$$\sum_{k=1}^{\infty} \mathsf{P}\{\xi_k > C\}, \quad \sum_{k=1}^{\infty} \mathsf{M}\xi_k(C),$$

где

$$\xi_k(C) = \begin{cases} \xi_k, & \text{если } \xi_k \leqslant C, \\ 0, & \text{если } \xi_k > C. \end{cases}$$

Ряд, составленный из величин ξ_k, называется *сходящимся по вероятности*, если сходятся по вероятности его частичные суммы. Для рядов из независимых случайных величин сходимость по вероятности влечет сходимость с вероятностью 1.

Последовательность случайных величин $\{\eta_n, n \geqslant 1\}$ называется *ограниченной по вероятности*, если

$$\lim_{A \to +\infty} \sup_n \mathsf{P}\{|\eta_n| > A\} = 0.$$

Если $\{\xi_n, n \geqslant 1\}$ — последовательность независимых симметричных случайных величин и $\sum_{k=1}^{n} \xi_k$ ограничены по вероятности, то ряд $\sum_{k=1}^{\infty} \xi_k$ сходится с вероятностью 1.

Аналогичные результаты справедливы и в том случае, когда рассматриваются суммы независимых случайных векторов. Аналогом теоремы о трех рядах в этом случае будет следующая теорема.

Теорема 4. *Пусть* $\sum_{k=1}^{\infty} \xi_k$ — *ряд, составленный из независимых случайных векторов. Для сходимости этого ряда почти наверное необходимо, чтобы для всех* $C > 0$ *сходились ряды*

$$\sum_{k=1}^{\infty} \mathsf{M}\xi_k(C), \quad \sum_{k=1}^{\infty} \mathsf{M}|\xi_k(C) - \mathsf{M}\xi_k(C)|^2, \quad \sum_{k=1}^{\infty} \mathsf{P}\{|\xi_k| > C\},$$

и достаточно, чтобы эти ряды сходились при некотором $C > 0$.

Здесь $\xi_k(C)$ — вектор, определяемый формулой

$$\xi_k(C) = \begin{cases} \xi_k, & \text{если } |\xi_k| \leqslant C, \\ 0, & \text{если } |\xi_k| > C. \end{cases}$$

2.8. Сходимость функций распределения

Излагаемые здесь определения и факты естественным образом переносятся на общий случай сходимости вероятностных распределений в метрических пространствах (см. § 18.1).

2.8.1. Определения. Последовательность функций распределения (ф. р.) $P_n(x)$ *сходится в основном* к ф. р. $P(x)$ ($P_n \Rightarrow P$), если при $n \to \infty$

$$P_n(x) \to P(x), \quad x \in C(P), \tag{8.1}$$

где $C(P)$ — множество точек непрерывности предельной ф. р. $P(x)$. Для сходимости ф. р. в основном достаточно сходимости ф. р. на счетном всюду плотном множестве точек числовой прямой $\mathbf{R}$.

Последовательность ф. р. $P_n(x)$ *слабо сходится* к ф. р. $P(x)$ $(P_n \to P)$, если при $n \to \infty$

$$\int_{-\infty}^{\infty} f(x)\, dP_n(x) \to \int_{-\infty}^{\infty} f(x)\, dP(x), \quad f(x) \in C(\mathbf{R}), \tag{8.2}$$

где $C(\mathbf{R})$ — множество непрерывных и ограниченных функций на числовой прямой $\mathbf{R}$.

Слабая сходимость и сходимость в основном ф. р. эквивалентны и обозначаются чаще всего так: $P_n \Rightarrow P$.

Пусть α_n — последовательность случайных величин с ф. р. $P_n(x) = \mathsf{P}\{\alpha_n < x\}$ и α — случайная величина с ф. р. $P(x) = \mathsf{P}\{\alpha < x\}$. Слабая сходимость ф. р. $P_n \Rightarrow P$ означает, что

$$\lim_{n \to \infty} \mathsf{M} f(\alpha_n) = \mathsf{M} f(\alpha), \quad f(x) \in C(\mathbf{R}). \tag{8.3}$$

Если имеет место слабая сходимость ф. р. (8.2), то говорят, что последовательность случайных величин α_n *сходится по распределению к* α $(\alpha_n \xrightarrow{D} \alpha)$.

Из сходимости последовательности случайных величин α_n *по вероятности* к случайной величине α следует сходимость *по распределению*. Обратное утверждение неверно, за исключением случая, когда имеет место сходимость к постоянной.

Слабая сходимость (8.2) может быть обобщена на более широкий класс *равномерно интегрируемых* функций $g(x)$:

$$\lim_{a \to \infty} \sup_n \int_{|x| > a} |g(x)|\, dP_n(x) = 0. \tag{8.4}$$

Теорема о сходимости. *Если последовательность ф. р. $P_n(x)$ слабо сходится к ф. р. $P(x)$ и функция $g(x)$ равномерно интегрируема и непрерывна, то*

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} g(x)\, dP_n(x) = \int_{-\infty}^{\infty} g(x)\, dP(x). \tag{8.5}$$

2.8.2. Компактность и плотность семейства функций распределений. Семейство ф. р. $\{P_\theta,\ \theta \in \Theta\}$ *слабо компактно*, если любая последовательность ф. р. семейства содержит слабо сходящуюся к ф. р. последовательность. Предельная ф. р. не обязательно принадлежит данному семейству.

В определении слабой компактности существенно, что пределом последовательности служит функция распределения, удовлетворяющая условиям нормировки: $P(-\infty) = 0$, $P(+\infty) = 1$.

Если расширить класс ф. р., не предполагая выполненными условиями нормировки, то слабая компактность будет иметь место для любого семейства таких обобщенных функций распределения.

Теорема 1 (Хелли). *Всякое семейство $\{\mathfrak{S}_\theta : \theta \in \Theta\}$ неубывающих, непрерывных слева функций $\mathfrak{S}_\theta(x)$, удовлетворяющих* *условию* $0 \leqslant \mathfrak{S}_\theta(-\infty)$, $\mathfrak{S}_\theta(+\infty) \leqslant 1$, *компактно (секвенциально), т. е. из любой последовательности функций семейства можно выделить подпоследовательность, сходящуюся в основном к функции* $\mathfrak{S}(x)$ *того же типа: неубывающей непрерывной слева и удовлетворяющей условию* $0 \leqslant \mathfrak{S}(-\infty)$, $\mathfrak{S}(+\infty) \leqslant 1$.

Семейство ф. р. $\{P_\theta\colon \theta \in \Theta\}$ *плотное*, если для каждого $\varepsilon > 0$ можно указать такое число $K_\varepsilon > 0$, что

$$\sup_{\theta \in \Theta} [1 - P_\theta(K_\varepsilon) + P_\theta(-K_\varepsilon)] \leqslant \varepsilon. \tag{8.6}$$

Фундаментальную роль в вопросах слабой сходимости ф. р. играет следующая теорема.

Теорема 2 (Прохорова). *Семейство ф. р.* $\{P_\theta\colon \theta \in \Theta\}$ *слабо компактно тогда и только тогда, когда оно плотное.*

2.8.3. Расстояние между функциями распределения. Выделение различных типов сходимости ф. р. эквивалентно, по существу, введению в пространстве всех ф. р. структуры определенного топологического пространства. При этом топология, как правило, задается метрикой (расстоянием) в пространстве ф. р., отвечающей тому или иному типу сходимости.

Наиболее употребительными метриками являются расстояние по вариации и расстояние Леви.

Расстояние по вариации между двумя ф. р. P и Q определяется формулой

$$V(P, Q) = \int_{-\infty}^{\infty} |dP(x) - dQ(x)|. \tag{8.7}$$

Расстояние Леви между двумя ф. р. P и Q определяется соотношением

$$L(P, Q) = \inf\{h\colon P(x-h) - h \leqslant Q(x) \leqslant P(x+h) + h \ \ \forall x \in \mathbf{R}\}. \tag{8.8}$$

Слабая сходимость ф. р. эквивалентна сходимости к нулю расстояния Леви, т. е.

$$P_n \Rightarrow P \text{ эквивалентно } L(P_n, P) \to 0,\ n \to \infty. \tag{8.9}$$

Заметим, что в случае, когда предельная ф. р. P непрерывна, слабая сходимость *равномерна*:

$$\sup_{-\infty < x < +\infty} |P_n(x) - P(x)| \to 0, \quad n \to \infty. \tag{8.10}$$

Литература: [44, 60, 63, 64, 81, 83].

Глава 3. АНАЛИТИЧЕСКИЙ АППАРАТ

3.1. Производящие функции

3.1.1. Определения, свойства. Пусть ν — целочисленная неотрицательная случайная величина с распределением вероятностей

$$\mathbf{P}\{\nu = k\} = p_k, \quad k = 0, 1, 2, \ldots \tag{1.1}$$

Производящей функцией распределения (1.1) случайной величины ν называется ряд

$$p(z) = \mathsf{M}z^{\nu} = \sum_{k=0}^{\infty} z^k p_k, \quad |z| \leqslant 1. \tag{1.2}$$

Для группы n целочисленных неотрицательных случайных величин $\nu_1, \nu_2, \ldots, \nu_n$ *совместная производящая функция* определяется рядом

$$p(z_1, z_2, \ldots, z_n) = \mathsf{M}z_1^{\nu_1} z_2^{\nu_2} \ldots z_n^{\nu_n} =$$

$$= \sum_{k_1, k_2, \ldots, k_n \geqslant 0} z_1^{k_1} z_2^{k_2} \ldots z_n^{k_n} p_{k_1 k_2 \ldots k_n}, \tag{1.3}$$

где $p_{k_1 k_2 \ldots k_n} = \mathsf{P}(\nu_1 = k_1, \nu_2 = k_2, \ldots, \nu_n = k_n)$.

Производящая функция аналитична внутри единичного круга $|z| < 1$. Распределение вероятностей (1.1) однозначно определяется своей производящей функцией:

$$p_k = \frac{1}{k!} p^{(k)}(0), \quad p^{(k)}(0) = \frac{d^k}{dz^k} p(z)\Big|_{z=0}, \quad k \geqslant 0. \tag{1.4}$$

Часто оказывается полезным (особенно в асимптотическом анализе) представление распределения интегралом Коши:

$$p_k = \frac{1}{2\pi i} \int\limits_{|z|=\rho} \frac{p(z)}{z^{k+1}} dz, \quad 0 < \rho \leqslant 1. \tag{1.5}$$

Определим хвост распределения:

$$\mathsf{P}(\nu > k) = q_k = \sum_{r=1}^{\infty} p_{k+r}, \quad k \geqslant 0. \tag{1.6}$$

Производящая функция $Q(z) = \sum_{k=0}^{\infty} z^k q_k$ последовательности $\{q_k, k \geqslant 0\}$ связана с производящей функцией $p(z)$ распределения $\{p_k, k \geqslant 0\}$ соотношением

$$Q(z) = \frac{1 - p(z)}{1 - z}. \tag{1.7}$$

В частности, математическое ожидание $\mathsf{M}\nu$ выражается формулой

$$\mathsf{M}\nu = Q(1) = \sum_{k=0}^{\infty} q_k. \tag{1.8}$$

Факториальные моменты случайной величины $\mathsf{M}\nu^{[m]} = \mathsf{M}[\nu \times \times (\nu - 1) \ldots (\nu - m + 1)]$ вычисляются по формуле

$$\mathsf{M}\nu^{[m]} = \frac{d^m}{dz^m} p(z)\Big|_{z=1}, \quad m \geqslant 1. \tag{1.9}$$

В частности, математическое ожидание $\mathsf{M}\nu$ и дисперсия $\mathsf{D}\nu$ определяются соответственно формулами

$$\begin{aligned} \mathsf{M}\nu &= p'(1); \\ \mathsf{D}\nu &= p''(1) + p'(1) - [p'(1)]^2. \end{aligned} \tag{1.10}$$

Для вычисления факториальных моментов можно также использовать следующее разложение производящей функции:

$$p(z+1) = \sum_{m=0}^{\infty} z^m B_m, \quad B_m = \frac{1}{m!} \mathsf{M}\nu^{[m]}. \tag{1.11}$$

Если производящая функция $p(z)$ случайной величины ν определена при всех $|z| < z_0$ (для некоторого $z_0 > 1$), то все моменты $m_k = \mathsf{M}\nu^k$ при $k \geqslant 1$ существуют и определяются производящей функцией

$$P(s) = p(e^s) = \sum_{k=0}^{\infty} \frac{s^k}{k!} m_k. \tag{1.12}$$

Производящая функция на интервале $(0, 1)$ имеет вероятностный смысл:

$$p(s) = \mathsf{P}\{\nu \leqslant \tau\}, \quad 0 < s < 1, \tag{1.13}$$

где τ — случайная величина, не зависящая от ν и имеющая геометрическое распределение с параметром s:

$$\mathsf{P}\{\tau = k\} = s^k (1-s), \quad k \geqslant 0, \tag{1.14}$$

с производящей функцией

$$p_\tau(z) = \mathsf{M}z^\tau = \frac{1-s}{1-sz}. \tag{1.15}$$

Равенство (1.13) можно интерпретировать следующим образом: $p(s) = \mathsf{M}s^\nu$ есть вероятность того, что число успехов ν в некоторой серии испытаний не превосходит числа τ наступивших подряд успехов в испытаниях Бернулли с вероятностью успеха в отдельном испытании, равной s.

Композиция (*свертка*) двух целочисленных распределений $\{p_k, k \geqslant 0\}$ и $\{q_k, k \geqslant 0\}$ задается формулой

$$c_k = \sum_{r=0}^{k} p_r q_{k-r} = \sum_{r=0}^{k} p_{k-r} q_r, \quad k \geqslant 0, \tag{1.16}$$

и обозначается через

$$\{c_k, k \geqslant 0\} = \{p_k, k \geqslant 0\} * \{q_k, k \geqslant 0\}. \tag{1.17}$$

Распределение суммы $\gamma = \mu + \nu$ двух независимых случайных величин с распределениями вероятностей слагаемых $\{p_k, k \geqslant 0\}$ и $\{q_k, k \geqslant 0\}$ определяется композицией (1.16):

$$\mathsf{P}\{\mu + \nu = k\} = c_k = \sum_{r=0}^{k} p_r q_{k-r}, \quad k \geqslant 0. \tag{1.18}$$

Производящая функция $p_\nu(z)$ *суммы* $\nu=\nu_1+\nu_2+\ldots+\nu_n$ независимых случайных величин равна произведению производящих функций слагаемых:

$$p_\nu(z)=p_{\nu_1}(z)\,p_{\nu_2}(z)\ldots p_{\nu_n}(z). \tag{1.19}$$

Пусть ν — неотрицательная целочисленная случайная величина с производящей функцией $\varphi(z)=\mathsf{M}z^\nu$. Производящая функция $p_\gamma(z)$ суммы $\gamma=\sum_{k=1}^{\nu}\mu_k$ независимых между собой и не зависящих от ν одинаково распределенных случайных величин μ_k с производящей функцией $p(z)=\mathsf{M}z^{\mu_k}$ равна *суперпозиции* производящих функций $\varphi(z)$ и $p(z)$:

$$p_\gamma(z)=\varphi[p(z)]. \tag{1.20}$$

3.1.2. Примеры. 1. *Биномиальное распределение* $B_p(n,k)=C_n^kp^kq^{n-k}$ $(q=1-p)$ имеет производящую функцию

$$b_p(n,z)=(q+pz)^n. \tag{1.21}$$

Случайная величина ν с биномиальным распределением представима в виде суммы $\nu=\sum_{k=1}^{n}\mu_k$ независимых, одинаково распределенных *бернуллиевых* случайных величин с производящей функцией $p_{\mu_k}(z)=q+pz$, принимающих два значения: 0 с вероятностью q и 1 с вероятностью p.

2. *Распределение Пуассона* $p_k(a)=\frac{a^k}{k!}e^{-a}$ $(k\geqslant 0,\ a>0)$ имеет производящую функцию

$$p(z,a)=e^{-a(1-z)}. \tag{1.22}$$

Сумма $\mu+\nu$ независимых случайных величин, имеющих пуассоновское распределение с параметрами a и b, имеет пуассоновское распределение с параметром $a+b$.

3. *Сложное распределение Пуассона* задается производящей функцией

$$\Pi(z)=e^{-a[1-p(z)]},\quad p(z)=\mathsf{M}z^\mu, \tag{1.23}$$

где $p(z)$ — производящая функция некоторых неотрицательных случайных величин μ.

Случайная величина γ со сложным распределением Пуассона (1.23) представима в виде суммы $\gamma=\sum_{k=1}^{\nu}\mu_k$, в которой слагаемые μ_k — независимые одинаково распределенные величины с производящей функцией $p(z)=\mathsf{M}z^{\mu_k}$, а число слагаемых ν, не зависящее от μ_k $(k\geqslant 1)$, имеет распределение Пуассона с производящей функцией $\mathsf{M}z^\nu=e^{-a(1-z)}$.

3.1.3. Теорема непрерывности и тауберова теорема.

Теорема непрерывности. *Пусть* $\{p_k^{(n)}, k \geqslant 0\}$ $(n = 1, 2, \ldots)$ — *последовательность распределений вероятностей целочисленных случайных величин* ν_n *с производящими функциями* $p_n(z)$. *Для сходимости распределений при каждом конечном* $k \geqslant 0$

$$\lim_{n\to\infty} p_k^{(n)} = p_k \tag{1.24}$$

необходимо и достаточно, чтобы при любом s *из полуинтервала* $0 \leqslant s < 1$

$$\lim_{n\to\infty} p_n(s) = p(s) = \sum_{k=0}^{\infty} s^k p_k. \tag{1.25}$$

Заданная на полуоси $(0, +\infty)$ функция $L(t)$ называется *медленно меняющейся* при $t \to \infty$, если при любом $x > 0$

$$\lim_{t\to\infty} L(tx)/L(t) = 1.$$

Медленно меняющаяся функция $L(t)$ представима в виде

$$L(t) = a(t) \exp\left\{\int_1^t \frac{\varepsilon(y)}{y}\,dy\right\}, \tag{1.26}$$

где $\varepsilon(t) \to 0$ и $a(t) \to a < \infty$ при $t \to \infty$, $a \neq 0$.

Тауберова теорема. *Пусть ряд* $a(s) = \sum_{n=0}^{\infty} a_n s^n$ *сходится при* $0 \leqslant s < 1$ *и* $a_n \geqslant 0$. *Тогда при* $0 \leqslant c < \infty$ *следующие два соотношения эквивалентны:*

$$(1-s)^c a(s) \sim L\left(\frac{1}{1-s}\right), \qquad s \to 1-0, \tag{1.27}$$

$$\frac{1}{n^c}\sum_{k=0}^{n} a_k \sim \frac{1}{\Gamma(c+1)} L(n), \quad n \to \infty. \tag{1.28}$$

Если последовательность a_n *монотонна и* $0 < c < \infty$, *то соотношение* (1.27) *равносильно соотношению*

$$\frac{a_n}{n^{c-1}} \sim \frac{1}{\Gamma(c)} L(n), \quad n \to \infty, \tag{1.29}$$

где $\Gamma(c) = \int_0^{\infty} x^{c-1} e^{-x}\,dx$ — *гамма-функция Эйлера.*

В частности, при $L(t) \equiv A$ соотношения (1.27)—(1.29) имеют соответственно следующий вид:

$$\lim_{s\to 1-0} (1-s)^c a(s) = A; \tag{1.30}$$

$$\lim_{n\to\infty} \frac{1}{n^c} \sum_{k=0}^{n} a_k = \frac{A}{\Gamma(c+1)}; \tag{1.31}$$

$$\lim_{n\to\infty} \frac{a_n}{n^{c-1}} = \frac{A}{\Gamma(c)}. \tag{1.32}$$

3.2. Преобразование Лапласа

3.2.1. Определение. Формулы обращения. Пусть ξ — неотрицательная случайная величина с функцией распределения вероятностей $P(x) = \mathsf{P}\{\xi \leqslant x\}$.

Преобразованием Лапласа $p(\lambda)$ функции распределения $P(x)$ (или случайной величины ξ) называется функция

$$p(\lambda) = \mathsf{M}e^{-\lambda\xi} = \int_0^\infty e^{-\lambda x}\, dP(x), \tag{2.1}$$

определенная для $\operatorname{Re}\lambda \geqslant 0$ и аналитическая при $\operatorname{Re}\lambda > 0$.

Предполагается, что точка 0 включена в область интегрирования.

Теорема обращения. *Функция распределения $P(x)$ однозначно определяется своим преобразованием Лапласа $p(\lambda)$: в каждой точке непрерывности функции распределения*

$$P(x) = \lim_{\lambda\to\infty} \sum_{k\leqslant\lambda x} \frac{(-\lambda)^k}{k!} p^{(k)}(\lambda). \tag{2.2}$$

Здесь $p^{(k)}(\lambda)$ — производная k-го порядка: $p^{(k)}(\lambda) = (-1)^k \times \int_0^\infty e^{-\lambda x} x^k\, dP(x)$.

Преобразование Лапласа $p(\lambda)$ можно представить в виде ряда

$$p(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} m_k, \quad m_k = \mathsf{M}\xi^k = \int_0^\infty x^k\, dP(x), \quad k \geqslant 1, \tag{2.3}$$

в любом интервале $0 \leqslant \lambda < \lambda_0$, в котором ряд сходится. Если такой интервал сходимости ряда (2.3) существует, то последовательность моментов $\{m_k, k \geqslant 0\}$ однозначно определяет функцию распределения $P(x)$.

Преобразование Лапласа $p(\lambda)$ при вещественных $\lambda > 0$ имеет вероятностный смысл:

$$p(\lambda) = \mathsf{P}\{\xi \leqslant \tau\}, \quad \lambda > 0, \tag{2.4}$$

где τ — случайная величина, не зависящая от ξ и имеющая показательное распределение с параметром λ:

$$\mathsf{P}\{\tau > t\} = e^{-\lambda t} \tag{2.5}$$

с преобразованием Лапласа $\mathsf{M}e^{-s\tau} = \lambda/(\lambda + s)$.

Равенство (2.4) в приложениях интерпретируется следующим образом: $p(\lambda) = \mathsf{M}e^{-\lambda\xi}$ при $\lambda > 0$ есть вероятность того, что момент успеха (восстановления, вызова, отказа и т. д.) ξ наступит до момента прекращения наблюдений τ, имеющего показательное распределение.

Если случайная величина ξ имеет плотность $u(x) = P'(x)$, то преобразование Лапласа записывается в виде

$$p(\lambda) = \mathsf{M}e^{-\lambda\xi} = \int_0^\infty e^{-\lambda x} u(x)\, dx. \tag{2.6}$$

Формула обращения в этом случае имеет вид

$$u(x) = \lim_{n\to\infty} \frac{(-1)^{n-1}}{(n-1)!} \left(\frac{n}{x}\right)^n p^{(n-1)}\left(\frac{n}{x}\right), \tag{2.7}$$

или для почти всех $x \geqslant 0$

$$u(x) = \frac{d}{dx} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} p(\lambda)\, e^{\lambda x} \frac{d\lambda}{\lambda}, \tag{2.8}$$

где $c > 0$ и интеграл берется вдоль любой прямой $\operatorname{Re}\lambda = c > 0$ и понимается в смысле *главного значения*, т. е. как предел интеграла вдоль отрезка $[c - iA, c + iA]$ при $A \to \infty$.

Если плотность распределения $u(x)$ непрерывна, то

$$u(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} p(\lambda)\, e^{\lambda x}\, d\lambda.$$

Функция распределения $F(x)$ суммы $\zeta = \xi + \eta$ двух независимых неотрицательных случайных величин ξ и η с функциями распределения $P(x)$ и $Q(x)$ определяется *сверткой*

$$F(x) = \int_0^x Q(x-y)\, dP(y) = \int_0^x P(x-y)\, dQ(y) \tag{2.9}$$

и обозначается $F = P*Q$.

Преобразование Лапласа $p_{\xi+\eta}(\lambda)$ *суммы* $\xi + \eta$ двух независимых случайных величин равно произведению преобразований Лапласа слагаемых:

$$p_{\xi+\eta}(\lambda) = p_\xi(\lambda)\, p_\eta(\lambda). \tag{2.10}$$

Равенство (2.10) равносильно следующему:

$$\mathsf{M}e^{-\lambda(\xi+\eta)} = \mathsf{M}e^{-\lambda\xi}\mathsf{M}e^{-\lambda\eta}. \qquad (2.11)$$

Пусть ν — целочисленная неотрицательная случайная величина с производящей функцией $\varphi(z) = \mathsf{M}z^{\nu}$. Преобразование Лапласа $p_{\eta}(\lambda)$ суммы $\eta = \sum_{k=1}^{\nu} \xi_k$ независимых между собой и не зависящих от ν, одинаково распределенных случайных величин ξ_k с преобразованием Лапласа $p(\lambda) = \mathsf{M}e^{-\lambda\xi_k}$ определяется суперпозицией функций $\varphi(z)$ и $p(\lambda)$:

$$p_{\eta}(\lambda) = \mathsf{M}\exp\left\{-s\sum_{k=1}^{\nu}\xi_k\right\} = \varphi[p(\lambda)]. \qquad (2.12)$$

3.2.2. Предельная теорема. Вполне монотонные функции.

Теорема непрерывности. *Если последовательность $P_n(x)$ функций распределения сходится к функции распределения $P(x)$, то их преобразования Лапласа $p_n(\lambda)$ сходятся к $p(\lambda)$ — преобразованию Лапласа предельной функции распределения $P(x)$ — в каждой точке $\lambda > 0$. Обратно, если последовательность преобразований Лапласа $p_n(\lambda)$ сходится при каждом $\lambda > 0$ к пределу $p(\lambda)$, то $p(\lambda)$ есть преобразование Лапласа функции распределения $P(x)$ и $P_n(x)$ сходятся к $P(x)$. Предельная функция распределения $P(x)$ — собственная ($P(+\infty) = 1$) тогда и только тогда, когда $p(\lambda) \to 1$ при $\lambda \to 0$*

Заданная на $(0, \infty)$ функция $p(\lambda)$ называется *вполне монотонной*, если она имеет производные $p^{(n)}(\lambda)$ при любом $n > 0$, и

$$(-1)^n p^{(n)}(\lambda) \geqslant 0, \quad \lambda > 0. \qquad (2.13)$$

Теорема о представлении. *Функция $p(\lambda)$ на $(0, \infty)$ является преобразованием Лапласа некоторого распределения вероятностей тогда и только тогда, когда она вполне монотонна и $p(0) = 1$.*

Критерий вполне монотонности. 1. *Произведение вполне монотонных функций — вполне монотонная функция.*

2. *Суперпозиция $\varphi(p(\lambda))$ вполне монотонной функции $\varphi(\lambda)$ с положительной функцией $p(\lambda)$, производная которой вполне монотонна, также вполне монотонна.*

3. *Предел последовательности вполне монотонных функций есть вполне монотонная функция.*

Для описания меры U, сосредоточенной на полуоси $(0, +\infty)$, используется неубывающая функция распределения $U(x) = \mathsf{U}((0, x))$.

Принцип сдвига. *Пусть U — мера на $(0, \infty)$ с преобразованием Лапласа $u(\lambda) = \int_0^{\infty} e^{-\lambda x}\mathsf{U}(dx)$, сходящимся при $\lambda \geqslant 0$. Тогда функция*

$$p(\lambda) = \frac{u(\lambda + a)}{u(a)} = \int_0^{\infty} e^{-\lambda x}\, dP(x) \qquad (2.14)$$

является преобразованием Лапласа распределения вероятностей

$$P(x) = \frac{1}{u(a)} \int_0^x e^{-ay} \mathsf{U}(dy). \tag{2.15}$$

Принцип сдвига позволяет все утверждения относительно преобразований Лапласа распределений вероятностей перенести на преобразования Лапласа мер, сосредоточенных на полуоси $(0, \infty)$.

Тауберова теорема. *Пусть* $u(\lambda) = \int_0^\infty e^{-\lambda x} \mathsf{U}(dx)$ — *преобразование Лапласа меры* U, $U(x) = \mathsf{U}((0, x))$. *При* $0 \leqslant c < \infty$ *соотношения*

$$u(\lambda) \backsim \lambda^{-c} L(1/\lambda), \quad \lambda \to 0. \tag{2.16}$$

$$U(x) \backsim \frac{x}{\Gamma(c+1)} L(x), \quad x \to \infty, \tag{2.17}$$

равносильны. Если существует монотонная плотность $U'(x)$, *то из* (2.17) *при* $0 < c < \infty$ *следует, что*

$$U'(x) \backsim \frac{x^{c-1}}{\Gamma(c)} L(x), \quad x \to \infty. \tag{2.18}$$

Здесь $L(x)$ — *медленно меняющаяся функция*: $\lim\limits_{x\to\infty} (L(xt)/L(x)) = 1$ *при любом* $t > 0$.

Примеры. 1. *Экспоненциальное распределение* $P(x) = 1 - e^{-ax}$ $(a > 0)$ имеет преобразование Лапласа

$$p(\lambda) = a/(a+\lambda). \tag{2.19}$$

2. *Гамма-распределение* с плотностью вероятностей

$$u(x;\ \rho,\ a) = \frac{1}{\Gamma(\rho)} a^\rho x^{\rho-1} e^{-ax}, \quad x > 0, \quad a > 0, \quad \rho > 0, \tag{2.20}$$

имеет преобразование Лапласа

$$p(\lambda;\ \rho,\ a) = \left(\frac{a}{a+\lambda}\right)^\rho. \tag{2.21}$$

3. *Сложное пуассоновское распределение* задается преобразованием Лапласа

$$p(\lambda) = \exp\left\{ a \int_0^\infty (e^{-\lambda x} - 1)\, dF(x) \right\}, \quad a > 0. \tag{2.22}$$

Соответствующая функция распределения представима в виде

$$P(x) = e^{-a} \sum_{n=0}^\infty \frac{a^n}{n!} F^{n*}(x). \tag{2.23}$$

Здесь $F^{n*}(x)$ — n-кратная свертка функции распределения $F(x)$ сама с собой.

Случайная величина ζ с распределением (2.23) представима в виде $\zeta = \sum_{k=0}^{\nu} \xi_k$, где ξ_k — независимые одинаково распределенные неотрицательные случайные величины с функцией распределения $F(x)$, ν — не зависящая от ξ_k случайная величина с пуассоновским распределением с параметром a.

Понятие преобразования Лапласа распространяется естественным образом на многомерные распределения. Определение (2.1) сохраняется в случае $P(x) = P(x_1, x_2, \ldots, x_n)$, если только понимать x и λ как векторы: $x = (x_1, x_2, \ldots, x_n)$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, а их произведение λx — как скалярное произведение векторов: $\lambda x = \sum_{k=1}^{n} \lambda_k x_k$.

3.3. Характеристические функции

3.3.1. Определение. Основные свойства. *Характеристической функцией* (*х. ф.*) случайной величины ξ с функцией распределения $F(x) = \mathsf{P}(\xi < x)$ называется комплекснозначная функция

$$f(t) = \mathsf{M}e^{it\xi} = \int_{-\infty}^{\infty} e^{itx}\, dF(x). \tag{3.1}$$

В частности, если существует плотность распределения вероятностей $p(x) = F'(x)$, то характеристическая функция есть преобразование Фурье плотности распределения:

$$f(t) = \int_{-\infty}^{\infty} e^{itx} p(x)\, dx. \tag{3.2}$$

Для дискретной случайной величины, принимающей значения x_k с вероятностями p_k, х. ф. представима рядом

$$f(t) = \sum_k e^{itx_k} p_k. \tag{3.3}$$

Характеристическая функция определена при вещественных t для любой случайной величины. Приведем основные свойства характеристических функций:

1) $f(0) = 1$, $|f(t)| \leqslant 1$, $-\infty < t < +\infty$;

2) $f(t)$ равномерно непрерывна на числовой оси;

3) при каждом целом $n > 0$ для любых комплексных чисел $z_1, z_2, \ldots, z_n$ и любых вещественных чисел $t_1, t_2, \ldots, t_n$

$$\sum_{k,\, r=1}^{n} f(t_k - t_r) z_k \bar{z}_r \geqslant 0 \tag{3.4}$$

(положительная определенность х. ф.);

4) $\overline{f(-t)} = f(t)$ — эрмитовость;

5) х. ф. суммы независимых случайных величин равна произведению х. ф. слагаемых:

$$f_{\xi_1+\xi_2}(t) = f_{\xi_1}(t)\, f_{\xi_2}(t); \tag{3.4$'$}$$

6) при $\eta = a\xi + b$, где a и b — постоянные,

$$f_\eta(t) = f_\xi(at)\, e^{ibt}. \tag{3.4$''$}$$

Основные свойства х. ф. 1) — 4) являются определяющими.

Теорема (Бохнера — Хинчина). *Для того чтобы непрерывная функция $f(t)$, заданная на вещественной оси и удовлетворяющая условию $f(0) = 1$, была характеристической функцией, необходимо и достаточно, чтобы она была положительно определенной.*

3.3.2. Примеры. 1. *Нормальное распределение* с плотностью $\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}$ имеет х. ф. $f(t) = \exp\left\{iat - \frac{1}{2}\sigma^2 t^2\right\}$.

2. *Равномерное распределение* на интервале $|x| < a$ имеет х. ф. $f(t) = \sin at/(at)$.

3. *Пуассоновское распределение* $p_k = \frac{a^k}{k!}e^{-a}$ $(k \geqslant 0)$ имеет х. ф. $f(t) = \exp\{a(e^{it} - 1)\}$.

4. *Распределение Бернулли* (биномиальное распределение) $B_p(n, k) = C_n^k p^k q^{n-k}$ $(0 \leqslant k \leqslant n)$ имеет х. ф. $f(t) = (q + pe^{it})^n$ $(q = 1 - p)$.

5. *Гамма-распределение* с плотностью $\frac{1}{\Gamma(\rho)} x^{\rho-1}e^{-x}$ $(x > 0,$ $\rho > 0)$ имеет х. ф. $f(t) = (1 - it)^{-\rho}$.

3.3.3. Взаимная однозначность и непрерывность соответствия между х. ф. и распределениями вероятностей.

Теорема обращения. *Функция распределения $F(x)$ однозначно определяется своей х. ф. $f(t)$. Если x, y — точки непрерывности $F(x)$, то имеет место формула обращения*

$$F(x) - F(y) = \frac{1}{2\pi} \lim_{c\to\infty} \int_{-c}^{c} \frac{e^{-itx} - e^{-ity}}{it} f(t)\, dt. \tag{3.5}$$

В частности, если $|f(t)/t|$ интегрируема на бесконечности, то

$$F(x) - F(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itx} - e^{-ity}}{it} f(t)\, dt. \tag{3.6}$$

Если же х. ф. $f(t)$ суммируема на вещественной оси, то функция распределения $F(x)$ имеет ограниченную непрерывную плотность $p(x) = F'(x)$, которая определяется формулой

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t)\, dt. \tag{3.7}$$

Для решетчатого распределения

$$p_k = \mathsf{P}\{\xi = a + kh\} = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{-it(a+kh)} f(t)\, dt. \tag{3.8}$$

Различные формулы обращения можно получить, используя *равенство Парсеваля.* Пусть $f(t) = \mathsf{M}e^{it\xi}$ и $\varphi(t) = \mathsf{M}e^{it\eta}$ — х. ф. независимых случайных величин ξ и η с функциями распределения $F(x) = \mathsf{P}\{\xi < x\}$ и $\Phi(x) = \mathsf{P}\{\eta < x\}$. Равенство Парсеваля задается формулой

$$\int_{-\infty}^{\infty} \varphi(t)\, dF(t) = \int_{-\infty}^{\infty} f(t)\, d\Phi(t). \tag{3.9}$$

Правая и левая части формулы (3.9) являются различными формами записи выражения $\mathsf{M}e^{i\xi\eta}$.

Одним из вариантов записи равенства Парсеваля является *формула обращения* (*со сглаживанием*)

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x-t)^2}{2\sigma^2}\right\} dF(t) = $$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{-itx - \frac{\sigma^2 t^2}{2}\right\} f(t)\, dt. \tag{3.10}$$

Величина скачков функции распределения определяется соотношением

$$F(x+0) - F(x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx} f(t)\, dt, \tag{3.11}$$

так что в точках непрерывности x функции распределения $F(x)$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx} f(t)\, dt = 0. \tag{3.12}$$

Теорема непрерывности. *Последовательность функций распределения $F_n(x)$ слабо сходится к функции распределения $F(x)$ тогда и только тогда, когда последовательность их х. ф. $f_n(t)$ сходится к непрерывной предельной функции $f(t)$. При этом $f(t)$ есть х. ф. предельной функции распределения $F(x)$ и сходимость $f_n(t)$ к $f(t)$ равномерная в каждом конечном интервале.*

Если последовательность интегрируемых х. ф. $f_n(t)$ сходится в среднем к предельной х. ф. $f(t)$, т. е. $\int_{-\infty}^{\infty} |f_n(t) - f(t)|\, dt \to 0$ при $n \to \infty$, то последовательность соответствующих плотностей распре-

деления $p_n(x)$ сходится равномерно к предельной плотности распределения $p(x)$.

3.3.4. Свойства регулярности. 1) Если функция распределения $F(x)$ *абсолютно непрерывна*, то $\lim\limits_{|t|\to\infty} |f(t)| = 0$.

2) Если функция распределения $F(x)$ имеет *абсолютно непрерывную компоненту*, то $\lim\limits_{|t|\to\infty} \sup |f(t)| < 1$.

3) Для *решетчатого распределения* $p_k = \mathsf{P}\{\xi = a + kh\}$ х. ф. представима в виде

$$f(t) = e^{iat} \sum_{k=-\infty}^{\infty} e^{ikht} p_k, \tag{3.13}$$

так что $|f(2\pi/h)| = 1$. Обратно, если при некотором $t_0 \neq 0$ справедливо равенство $|f(t_0)| = 1$, то соответствующее распределение решетчатое.

Максимальный шаг распределения равен h тогда и только тогда, когда модуль х. ф. меньше единицы при $0 < |t| < 2\pi/h$ и равен единице при $t = 2\pi/h$.

4) Для произвольной х. ф. $f(t)$ существует

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2 dt = \sum_k p_k^2, \tag{3.14}$$

где p_k — величины скачков функции распределения, и суммирование ведется по всем скачкам.

5) Если функция распределения $F(x)$ удовлетворяет *условию Липшица* с показателем $\gamma < 1$, то при $T \to \infty$

$$\frac{1}{T} \int_{-T}^{T} |f(t)|^2 dt = O\left(T^{-\gamma}\right). \tag{3.15}$$

Из условия

$$\int_{1}^{\infty} t^{\gamma-1} |f(t)| dt < \infty \tag{3.16}$$

следует, что функция распределения $F(x)$ удовлетворяет условию Липшица с показателем γ.

6) Пусть $f(t)$ — непрерывная неотрицательная четная функция, выпуклая в области $t > 0$ и удовлетворяющая условиям $f(0) = 1$, $\lim\limits_{t\to\infty} f(t) = 0$. Тогда $f(t)$ есть характеристическая функция. Отсюда следует, что существуют х. ф., совпадающие на конечном или бесконечном интервале, но не тождественно равные.

7) Если х. ф. имеет вид $f(t) = \exp\{P_k(t)\}$, где $P_k(t)$ — полином степени k, то $k \leqslant 2$, т. е. в этом случае $f(t) = \exp\left\{iat - \frac{\sigma^2 t^2}{2}\right\}$, где a и σ — вещественные параметры.

8) Характеристическая функция неотрицательной случайной величины не может обращаться в нуль на конечном интервале.

9) Характеристическая функция $f(x)$ вещественна ($\overline{f(t)} = f(t)$) тогда и только тогда, когда соответствующее распределение симметрично: $1 - F(-x+0) = F(x)$.

3.3.5. Моменты и семиинварианты. Моменты случайной величины ξ определяются значениями соответствующих производных характеристической функции:

$$m_k = \mathbf{M}\xi^k = (-i)^k f^{(k)}(0), \quad k \geqslant 1. \tag{3.17}$$

Если существует абсолютный момент $a_N = \mathbf{M}|\xi|^N < \infty$, то имеет место разложение

$$f(t) = 1 + \sum_{k=1}^{N} \frac{(it)^k}{k!} m_k + o(t^N). \tag{3.18}$$

Для достаточно малых значений t главная ветвь $\log f(t)$, которая стремится к нулю вместе с t, представима в виде

$$\log f(t) = \sum_{k=1}^{N} \frac{(it)^k}{k!} \gamma_k + o(t^N), \tag{3.19}$$

где *семиинварианты* (*кумулянты*) γ_k определяются формулой

$$\gamma_k = \frac{1}{i^k}\left[\frac{d^k}{dt^k} \log f(t)\right]_{t=0}. \tag{3.20}$$

Связь семиинвариантов γ_k и моментов $m_k = \mathbf{M}\xi^k$ выражается формулой

$$\gamma_k = k! \sum_{r=1}^{k} (-1)^{n_1+\ldots+n_{r-1}} (n_1 + \ldots + n_{r-1})! \prod_{l=1}^{r} \frac{1}{n_l!}\left(\frac{m_l}{l!}\right)^{n_l}. \tag{3.21}$$

Суммирование производится по всем целым неотрицательным решениям уравнения $n_1 + 2n_2 + \ldots + kn_k = k$.

Для того чтобы существовал абсолютный момент четного порядка $a_{2n} < \infty$, необходимо и достаточно, чтобы $f(t) = P_{2n}(t) + o(t^{2n})$ при $t \to 0$, где $P_{2n}(t)$ — полином степени $2n$.

Для существования производной $f^{(k)}(0)$ необходимо и достаточно, чтобы были выполнены условия

$$\lim_{x\to+\infty} x^k(1 - F(x) + F(-x)) = 0; \quad \lim_{c\to\infty} \int_{-c}^{c} x^k\, dF(x) = m_k. \tag{3.22}$$

При этом $f^{(k)}(0) = i^k m_k$.

3.3.6. Неравенства. Для оценок х. ф. используется следующее неравенство:

$$\left| e^{it} - \sum_{k=0}^{n} \frac{(it)^k}{k!} \right| \leqslant \frac{t^{n+1}}{(n+1)!} \tag{3.23}$$

для любых $n \geqslant 1$ и $t > 0$.

1. Пусть $\tau > 0$ и $x > 0$ таковы, что $\tau x > 1$. Тогда

$$\left(1 - \frac{1}{\tau x}\right) \mathsf{P}\{|\xi| \leqslant x\} \geqslant \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} f(t)\,dt \right| - \frac{1}{\tau x}. \tag{3.24}$$

Из этого неравенства, в частности, следует, что равностепенная непрерывность семейства х. ф. в нуле равносильна слабой компактности соответствующего семейства распределений.

2. Для всех вещественных t имеет место неравенство

$$0 \leqslant 1 - \operatorname{Re} f(2t) \leqslant 4(1 - \operatorname{Re} f(t)). \tag{3.25}$$

3. Если $|f(t)| \leqslant c < 1$ для $|t| \geqslant \varepsilon > 0$, то при $|t| < \varepsilon$

$$|f(t)| \leqslant 1 - \frac{1 - c^2}{8\varepsilon^2} t^2. \tag{3.26}$$

4. Для ограниченной случайной величины ξ с $|\xi| < c$ и дисперсией σ^2 х. ф. удовлетворяет неравенствам

$$e^{-\sigma^2 t^2} \leqslant |f(t)| \leqslant e^{-\sigma^2 t^2/3} \quad \text{для} \quad |t| \leqslant 1/(4c). \tag{3.27}$$

5. Характеристическая функция $f(t)$ случайной величины с ограниченной плотностью $p(x) \leqslant c$ и конечной дисперсией σ^2 удовлетворяет неравенствам

$$|f(t)| \leqslant \exp\left\{-\frac{A}{c^2\sigma^2}\right\} \quad \text{для} \quad |t| \geqslant \frac{\pi}{\sigma}; \tag{3.28}$$

$$|f(t)| \leqslant \exp\left\{-\frac{t^2}{96c^2(2\sigma^2|t| + \pi)^2}\right\} \quad \text{для всех} \quad t. \tag{3.29}$$

Здесь A — абсолютная постоянная.

3.3.7. Характеристические функции многомерных распределений. Пусть $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ — случайный вектор со значениями в евклидовом пространстве $\mathbf{R}^n$ с ф. р. $F(x) = \mathsf{P}\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n\}$, $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$.

Х. ф. случайного вектора ξ определяется равенством

$$f(t) = \mathsf{M} e^{i(t, \xi)} = \int_{\mathbf{R}^n} e^{i(t, x)}\,dF(x), \tag{3.30}$$

где $t = (t_1, t_2, \ldots, t_n)$, $(t, x) = \sum_{k=1}^{n} t_k x_k$ — скалярное произведение векторов t и x.

Свойства х. ф. многомерных распределений аналогичны свойствам х. ф. случайных величин. Укажем некоторые отличия. *Моментами* (*смешанными моментами*) случайного вектора $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ называются числа

$$m_{k_1 k_2 \ldots k_n} = \mathsf{M}\left(\xi_1^{k_1} \xi_2^{k_2} \ldots \xi_n^{k_n}\right). \tag{3.31}$$

Число $k = k_1 + k_2 + \ldots + k_n$ называется *порядком момента*. Моменты с целыми индексами можно определить дифференцированием х. ф.:

$$m_{k_1 k_2 \ldots k_n} = \left. \frac{(-i)^k \partial^k f(t)}{\partial t_1^{k_1} \partial t_2^{k_2} \ldots \partial t_n^{k_n}} \right|_{t=0}. \tag{3.32}$$

Примеры. 1. Двумерное нормальное распределение задается плотностью вероятностей

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-a)^2}{\sigma_1^2} - \right.\right.$$
$$\left.\left. - 2\rho \frac{(x-a)(x-b)}{\sigma_1\sigma_2} + \frac{(x-b)^2}{\sigma_2^2} \right] \right\}$$

с х. ф.

$$f(t_1, t_2) = \exp\left\{ iat_1 + ibt_2 - \frac{1}{2}\left(\sigma_1^2 t_1^2 + 2\sigma_1\sigma_2\rho t_1 t_2 + \sigma_2^2 t_2^2\right) \right\}.$$

Параметры распределения имеют следующий смысл:

$$a = \mathsf{M}\xi_1, \quad b = \mathsf{M}\xi_2; \quad \sigma_1^2 = \mathsf{M}\xi_1^2, \quad \sigma_2^2 = \mathsf{M}\xi_2^2; \quad \rho = \frac{\mathsf{M}(\xi_1\xi_2)}{\sigma_1\sigma_2}.$$

2. Многомерное нормальное распределение задается плотностью

$$p(x_1, \ldots, x_n) = \frac{\sqrt{D}}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2} Q(x_1, x_2, \ldots, x_n) \right\},$$

где $Q(x_1, x_2, \ldots, x_n) = \sum_{k, r=1}^{n} b_{kr}(x_k - a_k)(x_r - a_r)$ — положительно определенная форма, D — определитель матрицы $B = \{b_{kr}, 1 \leqslant k, r \leqslant n\}$. Соответствующая х. ф. имеет вид

$$f(t_1, t_2, \ldots, t_n) = \exp\left\{ iat - \frac{1}{2} t\sigma t' \right\},$$

где матрица вторых моментов $\sigma = \{\sigma_{ij}, 1 \leqslant i \leqslant j \leqslant n\}$ определяется соотношениями $\sigma = B^{-1}$; $\sigma_{ij} = \mathsf{M}\xi_i\xi_j$, $t\sigma t' = \sum_{i, j} \sigma_{ij} t_i t_j$.

Литература: [12, 23, 64, 89, 98].

Глава 4. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Термин *центральная предельная теорема* в теории вероятностей означает любое утверждение о том, что при выполнении определенных условий функция распределения суммы индивидуально малых случайных величин с ростом числа слагаемых сходится к нормальной функции распределения. Исключительная важность центральной предельной теоремы объясняется тем, что она дает теоретическое объяснение следующему многократно подтвержденному практикой наблюдению: если исход случайного эксперимента определяется большим числом случайных факторов, влияние каждого из которых пренебрежимо мало, то такой эксперимент хорошо аппроксимируется нормальным распределением с соответствующим образом подобранными математическим ожиданием и дисперсией.

4.1. Центральная предельная теорема для последовательностей независимых случайных величин

4.1.1. Центральная предельная теорема при наличии конечных дисперсий. Пусть $\{\xi_k,\ k \geqslant 1\}$ — последовательность взаимно независимых случайных величин с функциями распределения $G_k(x) = \mathsf{P}\{\xi_k < x\}$, имеющих конечные математические ожидания $\mathsf{M}\xi_k = a_k$ и дисперсии $\mathsf{D}\xi_k = \sigma_k^2$, причем $B_n^2 = \sum_{k=1}^{n} \sigma_k^2 > 0$ для $n \geqslant 1$.

Нормированной суммой случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ называется случайная величина

$$\eta_n = B_n^{-1} \sum_{k=1}^{n} (\xi_k - a_k),$$

которая характеризуется тем, что $\mathsf{M}\eta_n = 0$, $\mathsf{D}\eta_n = 1$ для любого $n \geqslant 1$.

Пусть $F_n(x)$ — функция распределения нормированной суммы η_n и $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\,dz$ — нормальная (0, 1) функция распределения. При наличии конечных дисперсий центральная предельная теорема устанавливает условия, при которых имеет место соотношение

$$\lim_{n\to\infty} F_n(x) = \Phi(x) \tag{1.1}$$

равномерно относительно $x \in (-\infty, \infty)$.

Одна из наиболее простых и в то же время наиболее часто применяемых (особенно в статистических приложениях) форм центральной предельной теоремы связана с последовательностью одинаково распределенных случайных величин.

Теорема Леви — Линдеберга. *Если $\{\xi_k,\ k \geqslant 1\}$ — последовательность взаимно независимых одинаково распределенных случайных величин, то для функции распределения $F_n(x)$ нормированной суммы $\eta_n = \frac{1}{\sigma\sqrt{n}}\left(\sum \xi_k - na\right)$ имеет место соотношение* (1.1), $a = \mathsf{M}\xi_k$, $\sigma^2 = \mathsf{D}\xi_k$.

Важный частный случай теоремы Леви — Линдеберга, формулируемый для случайных величин ξ_k, имеющих распределение Бернулли, представляет следующая теорема.

Центральная предельная теорема Муавра — Лапласа (интегральная теорема Муавра — Лапласа). *Если ν_n есть число наступлений некоторого события в серии из n независимых испытаний, в каждом из которых вероятность наступления этого* *события равна p, причем $0 < p < 1$, то для функции распределения $F_n(x)$ нормированного отклонения от среднего числа наступления события $\eta_n = \dfrac{\nu_n - np}{\sqrt{np(1-p)}}$ имеет место соотношение* (1.1).

Пример 1. Требуется оценить вероятность отклонения частоты ν_n/n появления события в схеме испытаний Бернулли от вероятности p, $0 < p < 1$, не более чем на ε, где ε — произвольное положительное число:

$$\mathsf{P}\left\{\left|\frac{\nu_n}{n} - p\right| < \varepsilon\right\} \simeq \Phi\left(\varepsilon\sqrt{\frac{n}{p(1-p)}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{p(1-p)}}\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\varepsilon\sqrt{n/(p(1-p))}}^{\varepsilon\sqrt{n/(p(1-p))}} e^{-z^2/2}\,dz.$$

Для конкретных значений n, p, ε правая часть этого равенства находится из таблиц нормальной функции распределения.

В случае разнораспределенных случайных величин одна из основных причин, в силу которой функция распределения $F_n(x)$ нормированной суммы η_n может не сходиться к нормальной функции распределения, связана с неравноценностью слагаемых в сумме η_n. Одним из условий, обеспечивающих равномерность вкладов слагаемых $(\xi_k - a_k)/B_n$ в η_n, является *условие равномерной малости*

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant n} \frac{\sigma_k}{B_n} = 0. \tag{1.2}$$

Это условие не достаточно для выполнения центральной предельной теоремы. Это показывает следующий пример.

Пример 2. Пусть $\mathsf{P}\{\xi_k = 0\} = 1 - 1/k^2$, $\mathsf{P}\{\xi_k = \pm k\} = 1/(2k^2)$. Тогда $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = 1$, $\eta_n = \dfrac{1}{\sqrt{n}} \sum\limits_{k=1}^{n} \xi_k$. Для величины η_n соотношение (1.1) не выполняется, так как $\eta_n \to 0$ с вероятностью 1, поскольку ряд $\sum\limits_{k=1}^{\infty} \xi_k$ сходится с вероятностью 1. Последнее следует из того, что

$$\sum_{k=1}^{\infty} \mathsf{P}\{\xi_k \neq 0\} = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty,$$

и значит, в силу теоремы Бореля — Кантелли, среди величин ξ_k с вероятностью 1 лишь конечное число отлично от нуля.

Наиболее удобно проверяемые достаточные условия дает следующая теорема.

Теорема Ляпунова. *Если для последовательности взаимно независимых случайных величин $\{\xi_k,\ k \geqslant 1\}$ существует $\delta > 0$ такое, что*

$$\lim_{n\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathsf{M}\,|\,\xi_k - a_k\,|^{2+\delta} = 0, \qquad (1.3)$$

то для функции распределения $F_n(x)$ нормированной суммы $\eta_n = \frac{1}{B_n} \sum_{k=1}^{n} (\xi_k - a_k)$ имеет место (1.1).

Условие (1.3) называется *условием Ляпунова.*

Наряду с (1.1) условие Ляпунова достаточно для выполнения соотношения

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} |\,x\,|^{2+\delta}\, dF_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |\,x\,|^{2+\delta} e^{-x^2/2}\, dx = \int_{-\infty}^{\infty} |\,x\,|^{2+\delta}\, d\Phi(x). \qquad (1.4)$$

Если имеет место сходимость к нормальному распределению (1.1) и выполнено соотношение (1.4), то условие Ляпунова необходимо для того, чтобы имела место равномерная малость в смысле (1.2).

Пример 3. Пусть случайные величины ξ_k взаимно независимы и имеют распределение

$$\mathsf{P}\{\xi_k = \pm k\} = 1/2.$$

В этом случае $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = \sigma_k^2 = k^2$,

$$B_n^2 = \sum_{1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6},$$

$$\mathsf{M}\,|\,\xi_k\,|^3 = k^3, \quad \sum_{k=1}^{n} \mathsf{M}\,|\,\xi_k\,|^3 = \sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}.$$

Следовательно,

$$\lim_{n\to\infty} \frac{1}{B_n^3} \sum_{k=1}^{n} \mathsf{M}\,|\,\xi_n\,|^3 = \lim_{n\to\infty} \frac{3\sqrt{3}}{4} \cdot \frac{n^4}{n^{4+1/2}} = 0,$$

и, таким образом, условие Ляпунова выполнено.

Наиболее общим условием, обеспечивающим выполнимость центральной предельной теоремы для последовательности случайных величин $\{\xi_k,\ k \geqslant 1\}$, имеющих конечные дисперсии, является *условие Линдеберга*: для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\varepsilon B_n} (x-a_k)^2\, dG_k(x) = 0,$$

где $G_k(x)$ — функции распределения случайных величин ξ_k.

Теорема Линдеберга — Феллера. *Пусть* $\{\xi_k,\ k \geqslant 1\}$ — *последовательность взаимно независимых случайных величин. Для того чтобы функция распределения* $F_n(x)$ *нормированной суммы* $\eta_n = \frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k)$ *удовлетворяла соотношению* (1.1) *и условию равномерной малости* (1.2), *необходимо и достаточно, чтобы выполнялось условие Линдеберга.*

Условие Ляпунова является достаточным для выполнения условия Линдеберга в силу неравенства

$$\frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\varepsilon B_n} (x-a_k)^2\, dG_k(x) \leqslant \frac{1}{\varepsilon^\delta} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathsf{M}\,|\xi_k - a_k|^{2+\delta}.$$

4.1.2. Общие условия сходимости к нормальному распределению для последовательностей независимых случайных величин. Пусть $\{\xi_k,\ k \geqslant 1\}$ — последовательность взаимно независимых случайных величин с функциями распределения $G_k(x) = \mathsf{P}\{\xi_k < x\}$ и $\eta_n = \frac{1}{\beta_n}\sum_{k=1}^{n}\xi_k - \alpha_n$, где $\{\alpha_n\}$ и $\{\beta_n > 0\}$ — некоторые последовательности постоянных.

В отсутствие предположения о конечности моментов может оказаться тем не менее, что существуют последовательности постоянных $\{\alpha_n\}$ и $\{\beta_n > 0\}$ такие, что для функции распределения $F_n(x) = \mathsf{P}\{\eta_n < x\}$ имеет место соотношение (1.1).

Теорема 1. *Пусть случайные величины* ξ_k *одинаково распределены,* $G(x) = \mathsf{P}\{\xi_k < x\}$. *Для того чтобы существовали последовательности постоянных* $\{\alpha_n\}$ *и* $\{\beta_n > 0\}$ *такие, что для функции распределения* $F_n(x) = \mathsf{P}\{\eta_n < x\}$ *имело бы место соотношение* (1.1), *необходимо и достаточно, чтобы*

$$\lim_{x\to\infty} x^2 \mathsf{P}\{|\xi_1| \geqslant x\} \Big/ \int\limits_{|z|<x} z^2\, dG(z) = 0. \tag{1.5}$$

Условие (1.5) эквивалентно следующему: функция $d(x) = \int\limits_{|z|<x} z^2\, dG(z)$ *медленно меняющаяся*, т. е. для любого $c > 0$

$$\lim_{x\to\infty} \frac{d(cx)}{d(x)} = 1. \tag{1.6}$$

Принято говорить, что если существуют последовательности постоянных $\{\alpha_n\}$ и $\{\beta_n > 0\}$ такие, что функция распределения $F_n(x)$ случайной величины $\eta_n = \frac{1}{\beta_n}\sum \xi_k - \alpha_n$, где ξ_k имеют общее распределение $G(x)$, сходится к функции распределения $F(x)$, то $G(x)$ притягивается к $F(x)$ или $G(x)$ принадлежит области притяжения распределения $F(x)$; таким образом, условие (1.5) является необходимым и достаточным условием притяжения распределения $G(x)$ к нормальному распределению. Общие вопросы, связанные с описанием всех возможных предельных распределений для η_n, рассмотрены в гл. 5.

Пример. Пусть случайные величины ξ_k имеют общую плотность распределения

$$g(x) = \begin{cases} \frac{2}{|x|^3}\ln|x|, & |x| \geqslant 1, \\ 0, & |x| < 1. \end{cases}$$

Дисперсия случайных величин с такой плотностью бесконечна. Тем не менее урезанный (по уровню $x \geqslant 1$) второй момент

$$d(x) = \int\limits_{|z|<x} z^2 g(z)\,dz = 4\int\limits_1^x \frac{\ln z}{z}\,dz = 2\ln^2 x$$

является медленно меняющейся функцией. Следовательно,

$$\mathsf{P}\left\{\sum_{k=1}^{n} \xi_k < x\sqrt{2\pi \ln n}\right\} \xrightarrow[n\to\infty]{} \Phi(x).$$

Теорема 2. *Пусть случайные величины* ξ_k $(k \geqslant 1)$ *взаимно независимы и* $G_k(x)$ — *их функции распределения. Для того чтобы существовали последовательности постоянных* $\{\alpha_n\}$ *и* $\{\beta_n > 0\}$ *такие, что выполнено условие равномерной малости* $\lim\max \mathsf{P}\{|\xi_k| > \varepsilon\beta_n\} = 0$ *для любого* $\varepsilon > 0$ *и для функции распределения* $F_n(x)$ *случайной величины* $\eta_n = \frac{1}{\beta_n}\sum_{k=1}^{n} \xi_k - \alpha_n$ *имеет место соотношение* (1.1), *необходимо и достаточно, чтобы существовала последовательность постоянных* γ_n $(\gamma_n \to \infty$ *при* $n \to \infty)$ *такая, что*

$$\sum_{k=1}^{n} \int\limits_{|x|>\gamma_n} dG_k(x) \to 0,$$

$$\frac{1}{\gamma_n^2}\sum_{k=1}^{n}\left\{\int\limits_{|x|<\gamma_n} x^2\,dG_k(x) - \left(\int\limits_{|x|<\gamma_n} x\,dG_k(x)\right)^2\right\} \to 1.$$

Если такая последовательность γ_n существует, в качестве α_n и β_n^2 можно взять суммы урезанных математических ожиданий и дисперсий:

$$\beta_n^2 = \sum_{k=1}^{n} \left\{ \int_{|x|<\gamma_n} x^2\, dG_k(x) - \left(\int_{|x|<\gamma_n} x\, dG_k(x) \right)^2 \right\},$$

$$\alpha_n = \frac{1}{\beta_n} \sum_{k=1}^{n} \int_{|x|<\gamma_n} x\, dG_k(x).$$

4.1.3. Центральная предельная теорема в схеме серий. *Схемой серий* называется двойная последовательность случайных величин $\{\xi_{nk},\ 1 \leqslant k \leqslant k_n,\ k_n \to \infty,\ n \geqslant 1\}$, в которой случайные величины $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n}$, образующие n-ю серию, взаимно независимы для любого n. Схема суммирования последовательностей есть частный случай схемы серий. Так, в случае конечных дисперсий n-я серия имеет вид $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nn}$, где

$$\xi_{ni} = (\xi_i - \mathsf{M}\xi_i) \left(\sum_{k=1}^{n} \mathsf{D}\xi_k \right)^{-1/2}.$$

Общая форма центральной предельной теоремы в схеме серий. *Пусть $\{\xi_{nk},\ 1 \leqslant k \leqslant k_n,\ n \geqslant 1\}$ — схема серий, а $F_{nk}(x)$ и $F_n(x)$ — функции распределения случайных величин ξ_{nk} и $\xi_n = \sum_{k=1}^{k_n} \xi_{nk}$ соответственно.*

Для того чтобы $\lim\limits_{n\to\infty} F_n(x) = \Phi(x)$ равномерно по $x \in (-\infty, \infty)$ и выполнялось условие равномерной малости

$$\lim_{n\to\infty} \max_{1 \leqslant i \leqslant k_n} \mathsf{P}\{|\xi_{ni}| \geqslant \varepsilon\} = 0 \tag{1.7}$$

для любого фиксированного $\varepsilon > 0$, необходимо и достаточно, чтобы были выполнены условия

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} \mathsf{P}\{|\xi_{nk}| \geqslant \varepsilon\} = 0,$$

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} \int_{|x|<\varepsilon} x\, dF_{nk}(x) = 0, \tag{1.8}$$

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} \left\{ \int_{|x|<\varepsilon} x^2\, dF_{nk}(x) - \left(\int_{|x|<\varepsilon} x\, dF_{nk}(x) \right)^2 \right\} = 1.$$

4.2. Центральная предельная теорема для независимых случайных векторов

4.2.1. Многомерный аналог интегральной теоремы Муавра — Лапласа. Рассмотрим схему независимых испытаний, в каждом из которых может осуществляться одно из m событий $A_1, A_2, \ldots, A_m$ с вероятностями $p_1, p_2, \ldots, p_m$, причем $0 < p_i < 1$, $q = 1 - \sum_{i=1}^{m} p_i > 0$.

Пусть $\nu_n(i)$ — число наступлений события A_i в серии из n испытаний, $\eta_n(i) = (\nu_n(i) - np_i)/\sqrt{np_i(1-p_i)}$ — нормированное отклонение от среднего числа появлений события A_i в серии из n испытаний, $\eta_n = (\eta_n(1), \eta_n(2), \ldots, \eta_n(m))$ — вектор нормированных отклонений, компоненты которого — зависимые случайные величины,

$$F_n(x) = F_n(x_1, \ldots, x_m) = \mathsf{P}\{\eta_n(1) < x_1, \ldots, \eta_n(m) < x_m\}.$$

Теорема.

$$\lim_{n\to\infty} F_n(x_1, \ldots, x_m) = \frac{1}{\sqrt{(2\pi)^m \det C}} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} \exp\left\{-\frac{1}{2} \sum_{i,j=1}^{m} c_{ij}^{(-1)} z_i z_j\right\} \times dz_1 \ldots dz_m, \quad (2.1)$$

где $C = \{c_{ij},\ i, j = 1, \ldots, m\}$ — *матрица ковариаций вектора* η_n;

$$c_{ij} = \mathsf{M}\eta_n(i)\,\eta_n(j) = \begin{cases} 1, & i = j, \\ -\sqrt{\dfrac{p_i p_j}{(1-p_i)(1-p_j)}}, & i \neq j; \end{cases}$$

$$\det C = \frac{q}{(1-p_1)(1-p_2)\ldots(1-p_m)} \neq 0;$$

$c_{ij}^{(-1)}$ — (i, j)-*й элемент матрицы* C^{-1}, *причем*

$$c_{ij}^{(-1)} = \begin{cases} \dfrac{(1-p_i)(p_i+q)}{q}, & i = j, \\ \dfrac{\sqrt{p_i p_j (1-p_i)(1-p_j)}}{q}, & i \neq j. \end{cases}$$

4.2.2. Многомерные аналоги теорем Леви — Линдеберга и Линдеберга — Феллера. Пусть $\left\{\xi_n = \left(\xi_1^n, \xi_2^n, \ldots, \xi_k^n\right),\ n = 1, 2, \ldots\right\}$ — последовательность взаимно независимых случайных векторов с математическими ожиданиями $\mathsf{M}\xi_n = a_n$ и ковариационными матрицами $C_n = \operatorname{cov}\xi_n = \mathsf{M}[\xi_n - a_n][\xi_n - a_n]^T$ (здесь ^{T} означает транспонирование вектор-столбца $[\xi_n - a_n]$). Пусть $B_n = \sum_{i=1}^{n} C_i$ — ковариационная матрица суммы $\sum_{l=1}^{n} \xi_l$ и $B_n^{1/2}$ — ее квадратный корень. Если матрица B_n положительно определена, то вектор $\eta_n = B_n^{-1/2} \sum_{l=1}^{n} (\xi_l - a_l)$ называется *нормированной суммой* случайных векторов $\xi_1, \ldots, \xi_n$. Вектор η_n характеризуется тем, что $\mathsf{M}\eta_n = 0$ (нулевой вектор) и $\operatorname{cov} \eta_n = \mathsf{M}\eta_n\eta_n^T = I$ (единичная $k \times k$-матрица).

Обозначим через $F_n(x) = F_n(x_1, x_2, \ldots, x_k)$ функцию распределения нормированной суммы η_n:

$$F_n(x_1, x_2, \ldots, x_k) = \mathsf{P}\left\{\eta_1^n < x_1, \eta_2^n < x_2, \ldots, \eta_k^n < x_k\right\},$$

где η_l^n — l-я компонента вектора η_n, и пусть $\Phi_{0, I}(x)$ — k-мерная нормальная функция распределения с нулевым средним и единичной ковариационной матрицей.

Теорема 1. *Если $\{\xi_n, n \geqslant 1\}$ — последовательность одинаково распределенных взаимно независимых случайных векторов, $\mathsf{M}\xi_k = a$ и матрица $C = \operatorname{cov} \xi_n$ положительно определена, то для функции распределения $F_n(x) = F_n(x_1, x_2, \ldots, x_k)$ нормированной суммы $\eta_n = \frac{1}{\sqrt{n}} C^{-1/2} \sum_{l=1}^{n} (\xi_l - a_l)$ имеет место соотношение*

$$\lim_{n \to \infty} F_n(x) = \Phi_{0, I}(x) \tag{2.2}$$

равномерно относительно $x \in \mathbb{R}^k$.

В случае разнораспределенных случайных векторов ξ_n имеет место следующее утверждение.

Теорема 2. *Пусть $G_l(x) = G_l(x_1, \ldots, x_k)$ — функции распределения случайных векторов ξ_l, $\mathsf{M}\xi_l = a_l$, и ковариационная матрица B_n суммы $\sum_{l=1}^{n} \xi_l$ положительно определена при любом n. Для того чтобы для функции распределения $F_n(x)$ нормированной суммы $\eta_n = B_n^{-1/2} \sum_{l=1}^{n} (\xi_l - a_l)$ имело место соотношение (2.2) и для любого $\varepsilon > 0$*

$$\lim_{n \to \infty} \max_{1 \leqslant l \leqslant n} \mathsf{P}\left\{\| \xi_l - a_l \| > \varepsilon \sqrt{\operatorname{Sp} B_n}\right\} = 0,$$

где $\operatorname{Sp} B_n$ — след матрицы B_n, необходимо и достаточно, чтобы выполнялись условия: для любого $\varepsilon > 0$

$$\lim_{n \to \infty} \frac{1}{\operatorname{Sp} B_n} \sum_{l=1}^{n} \int\limits_{\| x - a_l \| > \varepsilon \sqrt{\operatorname{Sp} B_n}} \| x - a_l \|^2 \, dG_l(x) = 0 \tag{2.3}$$

(*многомерный аналог условия Линдеберга*) *и*

$$\varliminf_{n\to\infty} \inf_{x\in \mathrm{R}^k} \frac{(B_n x, x)}{\|x\|^2} > 0. \tag{2.4}$$

4.2.3. Центральная предельная теорема для случайных векторов в схеме серий. Пусть $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n}$ $(n = 1, 2, \ldots)$ — последовательность серий независимых и одинаково распределенных в каждой серии случайных векторов со значениями в евклидовом пространстве R^k, и пусть $G_n(x)$ — функции распределения векторов ξ_{nl} $(l = 1, \ldots, k_n)$.

Теорема 3. *Если существуют вектор $a \in \mathrm{R}^k$ и неотрицательно определенная симметричная матрица B, для которых при всех $z \in \mathrm{R}^k$ и $\varepsilon > 0$ выполняются равенства*

$$\lim_{n\to\infty} k_n \int\limits_{\|x\|\leqslant\varepsilon} (z, x)\, G_n(dx) = (z, a),$$

$$\lim_{n\to\infty} k_n \left[\int\limits_{\|x\|\leqslant\varepsilon} (z, x)^2\, G_n(dx) - \left(\int\limits_{\|x\|<\varepsilon} (z, x)\, G_n(dx)\right)^2\right] = (Bz, z), \tag{2.5}$$

$$\lim_{n\to\infty} k_n \int\limits_{\|x\|>\varepsilon} G_n(dx) = 0,$$

то распределение случайного вектора $\eta_n = \sum_{l=1}^{k_n} \xi_{nl}$ *при* $n\to\infty$ *слабо сходится к нормальному распределению с характеристической функцией*

$$\varphi(z) = \exp\left\{ i(a, z) - \frac{1}{2}(Bz, z)\right\}. \tag{2.6}$$

4.3. Локальные предельные теоремы

4.3.1. Локальные предельные теоремы для плотностей. Пусть $\{\xi_k,\ k \geqslant 1\}$ — последовательность взаимно независимых случайных величин с функциями распределения $G_k(x) = \mathsf{P}\{\xi_k < x\}$ таких, что сумма $\sum_{k=1}^{n} \xi_k$ для $n \geqslant n_0$ имеет плотность распределения. Не уменьшая общности, можно считать, что $n_0 = 1$. Локальные предельные теоремы для плотностей выясняют условия, при которых плотности $f_n(x)$ распределений нормированных сумм либо (в общем случае) сумм вида $\eta_n = \frac{1}{\beta_n}\sum_{k=1}^{n} \xi_k - \alpha_n$ с соответствующим образом подобранными последовательностями постоянных $\{\alpha_n\}$ и $\{\beta_n > 0\}$, удовлетворяют соотношению

$$\lim_{n\to\infty} f_n(x) = \varphi(x) \tag{3.1}$$

равномерно относительно $x \in (-\infty, \infty)$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ — плотность стандартного нормального распределения.

I. Случай конечных дисперсий.

Теорема 1. *Если случайные величины в последовательности $\{\xi_k, k \geqslant 1\}$ одинаково распределены, имеют конечное математическое ожидание $\mathsf{M}\xi_k = a$ и дисперсию $\mathsf{D}\xi_k = \sigma^2$, и $f_n(x)$ — плотность распределения нормированной суммы* $\eta_n = \frac{1}{\sigma\sqrt{n}}\left(\sum \xi_k - na\right)$, *то для того, чтобы имело место соотношение* (3.1), *необходимо и достаточно, чтобы существовало такое N, что*

$$\sup_x f_N(x) < \infty.$$

В случае разнораспределенных случайных величин ξ_k определим класс M_r последовательностей случайных величин $\{\xi_k, k \geqslant 1\}$ с конечными математическими ожиданиями $\mathsf{M}\xi_k = a_k$ и дисперсиями $\sigma_k^2 = \mathsf{D}\xi_k$, причем $B_n^2 = \sum \sigma_k^2 > 0$, $n \geqslant 1$, характеризующийся тем, что среди распределений $G_k(x)$ случайных величин ξ_k не более r различных.

Пусть n_k $(k = 1, \ldots, r)$ есть число распределений k-го типа первых n членов последовательности $\{\xi_k, k \geqslant 1\}$ из M_r.

Теорема 2. *Если последовательность случайных величин $\{\xi_k, k \geqslant 1\}$ принадлежит классу M_r, причем $B_n \to \infty$ при $n \to \infty$, существует плотность $f_n(x)$ нормированной суммы* $\eta_n = \frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k)$ *и выполнено условие*

$$\lim_{n\to\infty} \min_{1 \leqslant k \leqslant r} \frac{n_k}{\ln B_n} = \infty, \tag{3.2}$$

то для того, чтобы имело место соотношение (3.1), *необходимо и достаточно, чтобы существовало такое N, что*

$$\sup_x f_N(x) < \infty.$$

II. Общие условия сходимости к плотности нормального распределения. Обозначим через $f_n(x)$ плотность распределения случайной величины

$$\eta_n = \frac{1}{\beta_n}\sum_{k=1}^{n} \xi_k - \alpha_n,$$

где случайные величины ξ_k независимы и имеют одинаковое распределение $G(x)$, а $\{\alpha_n\}$ и $\{\beta_n > 0\}$ — некоторые последовательности постоянных.

Теорема 3. *Для того чтобы существовали последовательности постоянных $\{\alpha_n\}$ и $\{\beta_n > 0\}$ такие, что для $f_n(x)$ имеет место* (3.1), *необходимо и достаточно, чтобы выполнялись следующие условия:*

1) *функция распределения $G(x)$ принадлежит области притяжения нормального распределения* (п. 4.1.2);

2) *существует такое N, что* $\sup_x f_N(x) < \infty$.

4.3.2. Локальные теоремы для решетчатых распределений. Пусть $\{\xi_n, n \geqslant 1\}$ — последовательность взаимно независимых случайных величин, имеющих одинаковое решетчатое распределение $G(x)$, т. е. (см. п. 1.4.2) случайные величины ξ_n принимают значения из некоторой арифметической прогрессии $\{m+kh\}$, $h>0$, $k=0, \pm 1, \pm 2, \ldots$ Пусть случайные величины ξ_n имеют конечное математическое ожидание $\mathsf{M}\xi_k = a$ и дисперсию $\mathsf{D}\xi_k = \sigma^2$, и пусть $P_n(r) = \mathsf{P}\left\{\sum_{l=1}^{n} \xi_l = nm + rh\right\}$.

Теорема Гнеденко. *Для того чтобы равномерно относительно r $(-\infty < r < \infty)$ имело место соотношение*

$$\lim_{n\to\infty}\left|\frac{\sigma\sqrt{n}}{h} P_n(r) - \varphi(x_{nr})\right| = 0, \tag{3.3}$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ *— плотность нормального (0, 1) распределения, а* $x_{nr} = (n(m-a)+rh)/(\sigma\sqrt{n})$, *необходимо и достаточно, чтобы шаг h распределения $G(x)$ был максимальным.*

В частности, если случайные величины ξ_k имеют распределение Бернулли, то теорема Гнеденко переходит в локальную теорему Муавра — Лапласа.

Локальная теорема Муавра — Лапласа. *Если вероятность p наступления события отлична от нуля и единицы, то вероятность $P_n(r)$ того, что в серии из n независимых испытаний событие наступит ровно r раз, удовлетворяет соотношению* (3.1) *с*

$$x_{nr} = \frac{r-np}{\sqrt{np(1-p)}} \quad \text{и} \quad \sigma = \sqrt{p(1-p)}.$$

В отсутствие предположения о конечности моментов имеет место следующая теорема.

Теорема 4. *Для того чтобы для некоторых последовательностей постоянных $\{\alpha_n\}$ и $\{\beta_n > 0\}$ равномерно по r $(-\infty < r < \infty)$ имело место соотношение*

$$\lim_{h\to\infty}\left|\frac{\beta_n}{h} P_n(r) - \varphi(x_{nr})\right| = 0, \tag{3.4}$$

где $\varphi(x)$ — плотность нормального (0, 1) *распределения и* $x_{nr} = (nm - \alpha_n\beta_n + rh)/\beta_n$, *необходимо и достаточно, чтобы были выполнены следующие условия:* 1) *общая функция распределения $G(x)$ случайных величин ξ_k притягивается к нормальному закону* (п. 4.1.2); 2) *шаг h распределения $G(x)$ максимален.*

4.4. Уточнение центральной предельной теоремы и асимптотические разложения

4.4.1. Неравенства Эссеена и Берри — Эссеена. Пусть $\{\xi_k, k \geqslant 1\}$ — последовательность взаимно независимых случайных величин, имеющих конечные математические ожидания $\mathsf{M}\xi_k = a_k$ и дисперсии $\mathsf{D}\xi_k = \sigma_k^2$ и $B_n^2 = \sum_{k=1}^{n} \sigma_k^2 > 0$. Предположение о существовании моментов порядка выше чем 2 позволяет установить не только факт слабой сходимости функции распределения $F_n(x)$ нормированной суммы $\eta_n = \frac{1}{B_n} \sum_{k=1}^{n} (\xi_k - a_k)$ к нормальной (0, 1) функции распределения $\Phi(x)$, но и выяснить, каким образом это происходит.

Теорема 1. *Если для некоторого положительного $\delta \leqslant 1$ существуют $\mathsf{M} |\xi_k - a_k|^{2+\delta}$, то*

$$\sup_x | F_n(x) - \Phi(x) | \leqslant A \frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathsf{M} | \xi_k - a_k |^{2+\delta}, \tag{4.1}$$

где A — абсолютная константа.

Неравенство (4.1) в случае $\delta = 1$ называют обычно *неравенством Эссеена*. В частности, если случайные величины $\xi_1, \ldots, \xi_n$ имеют одинаковое распределение и $\delta = 1$, неравенство (4.1) переходит в неравенство

$$\sup_x | F_n(x) - \Phi(x) | \leqslant A \frac{\mathsf{M} | \xi_1 - a |^3}{\sigma^3 \sqrt{n}}. \tag{4.2}$$

Неравенство (4.2) называется *неравенством Берри — Эссеена.*

Абсолютные константы в (4.1) и (4.2) не могут быть меньше величины $1/\sqrt{2\pi}$. Наименьшее значение константы A в неравенстве Берри — Эссеена равно

$$\sup \sqrt{n} \frac{\sigma^3}{\mathsf{M} | \xi_1 - a |^3} \sup_x | F_n(x) - \Phi(x) |,$$

где первый sup берется по всем n и всем функциям распределения $F_n(x)$, имеющим конечный третий момент и нулевое среднее. Точное значение этой константы неизвестно. Известно, однако, что

$$\varlimsup_{n \to \infty} \sup_F \sup_x \sqrt{n} \frac{\sigma^3}{\mathsf{M} | \xi_1 - a |^3} | F_n(x) - \Phi(x) | = \frac{\sqrt{10} + 3}{6\sqrt{2\pi}}.$$

По современным оценкам значение абсолютной константы A в неравенстве (4.1) не превышает число 0,7915, а в неравенстве (4.2) — число 0,7655. Неравенство Берри — Эссеена допускает следующие усиления и модификации:

1) $| F_n(x) - \Phi(x) | < A \frac{\mathsf{M} | \xi_1 - a |^3}{\sigma^3 \sqrt{n} (1 + | x |^3)}$;

2) если $\rho^{(p)}(F_n, \Phi)$ — расстояние между F_n и Φ в метрическом пространстве L_p $(p \geqslant 1)$, т. е.

$$\rho^{(p)}(F_n, \Phi) = \left\{ \int_{-\infty}^{\infty} |F_n(x) - \Phi(x)|^p \, dx \right\}^{1/p},$$

то

$$\rho^{(p)}(F_n, \Phi) \leqslant A \frac{\mathsf{M}\,|\xi_1 - a|^3}{\sigma^3 \sqrt{n}}.$$

Порядок оценок в (4.1) и (4.2) нельзя улучшить, не вводя дополнительных предположений.

4.4.2. Уточнения центральной предельной теоремы в многомерном случае. Пусть $\{\xi_k, k \geqslant 1\}$ — последовательность взаимно независимых одинаково распределенных случайных векторов со значениями в $\mathbf{R}^k$ с вектором математических ожиданий $\mathsf{M}\xi_k = a$ и положительно определенной ковариационной матрицей $B = \operatorname{cov} \xi_k$.

Неравенство Берри — Эссеена в многомерном случае имеет вид: если $G_n(x)$ — функция распределения вектора $\frac{1}{\sqrt{n}} \sum_{l=1}^{n} \xi_l$, то

$$\sup_{x \in \mathbf{R}^k} |G_n(x) - \Phi_{0,B}(x)| \leqslant A(k) \left(\sum_{l=1}^{k} \frac{\Lambda_{ll}}{\Lambda} \rho_l \right) \frac{1}{\sqrt{n}}, \tag{4.3}$$

где $\Phi_{0,B}(x)$ — нормальная функция распределения с вектором математических ожиданий 0 и ковариационной матрицей B; $A(k)$ — абсолютная константа, зависящая только от размерности k; $\rho_l = \mathsf{M}\,|\xi_l^n|^3 / \left(\mathsf{M}(\xi_l^n)^2\right)^{3/2}$; ξ_l^n — l-я компонента вектора ξ_n; $\Lambda = \det B$; Λ_{ll} — l-й главный минор ковариационной матрицы B.

В частности, при $k = 2$ неравенство (4.3) принимает вид

$$\sup_{x \in \mathbf{R}^2} |G_n(x) - \Phi_{0,B}(x)| \leqslant A(2) \frac{\rho_1 + \rho_2}{1 - \lambda^2} \cdot \frac{1}{\sqrt{n}}. \tag{4.4}$$

Оценка (4.4) имеет смысл лишь для значений λ, которые не очень близки к ± 1, т. е. когда распределение вектора ξ_1 не очень близко к вырожденному.

Оценка скорости сходимости, верная при любых предположениях о характере зависимости компонент вектора ξ_1, имеет вид

$$\sup_{x \in \mathbf{R}^k} |G_n(x) - \Phi_{0,B}(x)| \leqslant B(k) \sqrt{\frac{1}{\sqrt{n}} \sum_{l=1}^{k} \rho_l}, \tag{4.5}$$

где $B(k)$ — постоянная, зависящая только от размерности k.

4.4.3. Асимптотические разложения для сумм случайных величин. Асимптотические разложения в центральной предельной теореме основаны на разложении функций по *полиномам Чебышева — Эрмита* $H_m(x)$, определяемым любым из равенств

$$H_m(x) = (-1)^m e^{x^2/2} \frac{d^m}{dx^m} e^{-x^2/2},$$
$$H_m(x) = m! \sum_{k=0}^{[m/2]} \frac{(-1)^k x^{m-2k}}{k!\,(m-2k)!\,2^k}, \tag{4.6}$$

где $[m/2]$ означает целую часть числа $m/2$.

Несколько первых полиномов Чебышева — Эрмита имеют вид

$$H_0(x) = 1;\quad H_1(x) = x;\quad H_2(x) = x^2 - 1;\quad H_3(x) = x^3 - 3x;$$
$$H_4(x) = x^4 - 6x^2 + 3;\quad H_5(x) = x^5 - 10x^3 + 15x; \ldots$$

Обозначим j-й семиинвариант случайной величины ξ_1 через γ_j, и пусть

$$Q_m(x) = -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sum H_{m+2s-1}(x) \prod_{l=1}^{m} \frac{1}{k_l!}\left(\frac{\gamma_{l+2}}{(l+2)!\,\sigma^{l+2}}\right)^{k_l}, \tag{4.7}$$

$$q_m(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sum H_{m+2s}(x) \prod_{l=1}^{m} \frac{1}{k_l!}\left(\frac{\gamma_{l+2}}{(l+2)!\,\sigma^{l+2}}\right)^{k_l}, \tag{4.8}$$

где суммирование ведется по всем неотрицательным целочисленным решениям уравнения $k_1 + 2k_2 + \ldots + mk_m = m$, а $s = k_1 + k_2 + \ldots + k_m$.

Теорема 2. *Если взаимно независимые и одинаково распределенные случайные величины ξ_n $(n \geqslant 1)$ имеют конечный абсолютный момент порядка $r \geqslant 3$, $\mathsf{M}\xi_n = a$, $\mathsf{D}\xi_n = \sigma^2$ и для них выполнено условие Крамера* (С)

$$\overline{\lim_{z\to\infty}} |g(z)| < 1, \tag{4.9}$$

где $g(z)$ — характеристическая функция распределения $G(x) = \mathsf{P}\{\xi_n < x\}$, то равномерно относительно $x \in (-\infty, \infty)$ для функции распределения $F_n(x)$ нормированной суммы $\eta_n = \frac{1}{\sigma\sqrt{n}} \times \left(\sum_{l=1}^{n} \xi_l - na\right)$ имеет место асимптотическое разложение

$$F_n(x) = \Phi(x) + \sum_{m=1}^{r-2} \frac{Q_m(x)}{\sqrt{n^m}} + o\left(n^{-(r-2)/2}\right), \tag{4.10}$$

где $\Phi(x)$ — *нормальная* (0, 1) *функция распределения, а* $Q_m(x)$ *определяются равенством* (4.7).

В частности, если $r = 3$ и $\mu_3 = \int\limits_{-\infty}^{\infty} (x-a)^3 G(dx)$, то

$$F_n(x) = \Phi(x) + \frac{\mu_3}{6\sigma^3\sqrt{n}}(1-x^2)\frac{1}{\sqrt{2\pi}}e^{-x^2/2} + o\left(\frac{1}{\sqrt{n}}\right).$$

Разложение (4.10) имеет различные модификации и усиления, а именно в условиях сформулированной выше теоремы

$$(1+|x|^r)\left|F_n(x) - \Phi(x) - \sum_{m=1}^{r-2}\frac{Q_m(x)}{\sqrt{n^m}}\right| = o\left(n^{-(r-2)/2}\right). \quad (4.11)$$

Для любого $p > 1/r$

$$\int\limits_{-\infty}^{\infty}\left|F_n(x) - \Phi(x) - \sum_{l=1}^{r-2}\frac{Q_l(x)}{\sqrt{n^l}}\right|^p dx = o\left(n^{-(r-2)p/2}\right). \quad (4.12)$$

Для любого $p \geqslant 1$

$$\int\limits_{-\infty}^{\infty}|F_n(x) - \Phi(x)|^p\,dx = \int\limits_{-\infty}^{\infty}\left|\sum_{l=1}^{r-2}\frac{Q_l(x)}{\sqrt{n^l}}\right|^p dx + o\left(n^{-(r+p-3)/2}\right). \quad (4.13)$$

Для любого $p \geqslant 1$

$$\|F_n(x) - \Phi(x)\|_p = \left\|\sum_{l=1}^{r-2}\frac{Q_l(x)}{\sqrt{n^l}}\right\|_p + o\left(n^{-(r-2)/2}\right), \quad (4.14)$$

где $\|f(x)\|_p = \left(\int\limits_{-\infty}^{\infty}|f(x)|^p\,dx\right)^{1/p}$, если функция $f(x)$ удовлетворяет условию $\int\limits_{-\infty}^{\infty}|f(x)|^p\,dx < \infty$.

Теорема 3. *Если взаимно независимые и одинаково распределенные случайные величины* ξ_n $(n \geqslant 1)$ *имеют решетчатое распределение со значениями в прогрессии* $\{m + hk\}$, $h > 0$, $k = 0, \pm1, \pm2, \ldots$, *шаг* h *максимален,* $\mathsf{M}\xi_n = a$, $\mathsf{D}\xi_n = \sigma^2 > 0$ *и существует конечный абсолютный момент порядка* $r \geqslant 3$, *то равномерно относительно* $x \in (-\infty, \infty)$ *для функции распределения* $F_n(x)$ *нормированной суммы* $\eta_n = \frac{1}{\sigma\sqrt{n}}\left(\sum\limits_{l=1}^{n}\xi_l - na\right)$ *имеет место асимптоти-*

ческое разложение

$$F_n(x) = \Phi_{nr}(x) + \sum_{l=1}^{r-2} d_l \left(\frac{h}{\sigma\sqrt{n}}\right)^l S_l \left\{\frac{\sigma x\sqrt{n}}{h} - \right.$$
$$\left. - \left(\frac{nm}{h} - \left[\frac{nm}{h}\right]\right)\right\} \frac{d^l}{dx^l}\Phi_{nr}(x) + o\left(n^{-(r-2)/2}\right), \qquad (4.15)$$

где

$$\Phi_{nr}(x) = \Phi(x) + \sum_{l=1}^{r-2} \frac{Q_l(x)}{\sqrt{n^l}},$$

$$d_l = \begin{cases} 1, & \text{если } l \text{ имеет вид } l = 4k+1 \text{ или } l = 4k+2, \\ -1, & \text{если } l \text{ имеет вид } l = 4k+3 \text{ или } l = 4k; \end{cases}$$

$$S_{2l}(x) = \sum_{j=1}^{\infty} \frac{\cos 2\pi jx}{2^{2l-1}(\pi j)^{2l}}, \qquad S_{2l+1}(x) = \sum_{j=1}^{\infty} \frac{\sin 2\pi jx}{2^{2l}(\pi j)^{2l+1}}.$$

Теорема 4. *Если взаимно независимые и одинаково распределенные случайные величины ξ_n $(n \geqslant 1)$ имеют конечный абсолютный момент порядка $r \geqslant 3$, $\mathsf{M}\xi_n = a$, $\mathsf{D}\xi_n = \sigma^2 > 0$ и плотность $f_n(x)$ распределения нормированной суммы $\eta_n = \frac{1}{\sigma\sqrt{n}}\left(\sum_{l=1}^{n}\xi_l - na\right)$ ограничена для некоторого $n = N$, то равномерно относительно $x \in (-\infty, \infty)$ имеет место асимптотическое разложение*

$$f_n(x) = \varphi(x) + \sum_{l=1}^{r-2} \frac{q_l(x)}{\sqrt{n^l}} + o\left(n^{-(r-2)/2}\right), \qquad (4.16)$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ *— плотность нормального* (0, 1) *распределения;* $q_l(x)$ *определяются формулой* (4.8).

Теорема 5. *Если взаимно независимые и одинаково распределенные случайные величины ξ_n $(n \geqslant 1)$ принимают лишь целочисленные значения, максимальный шаг распределения равен 1 и существует конечный абсолютный момент порядка $r \geqslant 3$, то равномерно относительно $k \in (-\infty, \infty)$*

$$\sigma\sqrt{n}\,P_n(k) = \varphi(x_{nk}) + \sum_{l=1}^{r-2} \frac{q_l(x_{nk})}{\sqrt{n^l}} + o\left(n^{-(r-2)/2}\right),$$

где $P_n(k) = \mathsf{P}\left\{\sum_{l=1}^{n}\xi_l = k\right\}$; $x_{nk} = \frac{k - na}{\sigma\sqrt{n}}$; $\varphi(x)$ *— плотность нормального* (0, 1) *распределения, а функции* $q_l(x)$ *определяются формулой* (4.8).

4.5. Большие уклонения

4.5.1. Зоны нормальной сходимости. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — последовательность одинаково распределенных взаимно независимых случайных величин.

Если случайные величины ξ_k удовлетворяют условиям (интегральной) центральной предельной теоремы, то из равномерной сходимости функции распределения $F_n(x)$ нормированной суммы η_n к нормальной (0, 1) функции распределения $\Phi(x)$ следует, что равномерно по x из каждого конечного интервала имеют место соотношения

$$\frac{1-F_n(x)}{1-\Phi(x)} \to 1, \quad n \to \infty, \qquad \frac{F_n(-x)}{\Phi(-x)} \to 1, \quad n \to \infty. \tag{5.1}$$

Аналогично, если случайные величины ξ_k удовлетворяют условиям локальной предельной теоремы для плотностей и $f_n(x)$ есть плотность распределения нормированной суммы η_n, то равномерно по x из каждого конечного интервала имеет место соотношение

$$f_n(x) \Big/ \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \to 1, \quad n \to \infty. \tag{5.2}$$

Соотношения (5.1) и (5.2) могут иметь место равномерно по x, меняющихся в отрезках $[0, \Lambda(n)]$ или $[-\Lambda(n), 0]$, где $\Lambda(n)$ — неубывающая функция, неограниченно возрастающая вместе с n. Такие интервалы называются *зонами* (*интегральной* в случае (5.1) и *локальной* в случае (5.2)) *нормальной сходимости*. Представление об интегральной зоне нормальной сходимости дает следующий пример.

Пример. Пусть случайные величины $\xi_1, \xi_2, \ldots$ описывают схему независимых испытаний Бернулли. Из того, что

$$1-F_n(x) = \mathrm{P}\left\{\sum_{k=1}^{n} \xi_k > x\sqrt{np(1-p)} + np\right\},$$

следует, что для любых $x > \sqrt{n(1-p)/p}$ имеет место равенство $\dfrac{1-F_n(x)}{1-\Phi(x)} = 0$. Таким образом, в интервале $[0, O(\sqrt{n})]$ соотношение (5.1) может не иметь места.

Для зон нормальной сходимости $\Lambda(n) = o(\sqrt{n})$. В частности, если $\Lambda(n) = o(\sqrt[6]{n})$, то соответствующие зоны называются *узкими*, если же $\Lambda(n) = n^{\alpha}$, где $0 < \alpha < 1/2$ — заданное число, то соответствующие зоны называются *мономиальными*.

4.5.2. Индивидуальные асимптотические разложения в схеме больших уклонений. Пусть случайные величины ξ_k, определенные выше, удовлетворяют *условию Крамера*:

$$\exists h > 0 \quad \text{такое, что} \quad \mathrm{M}\exp\{h|\xi_k|\} < \infty, \tag{5.3}$$

обеспечивающему существование всех моментов ξ_k. В этом случае зоной интегральной и локальной (если существует ограниченная плотность вероятности случайных величин ξ_k) нормальной сходимости является узкая зона.

Обозначим через $f(z)$ характеристическую функцию случайных величин ξ_k, и пусть $\psi(z) = \ln f(z)$. При выполнении условия Крамера (5.3) $\psi(z)$ является аналитической функцией в окрестности нуля. Для достаточно малых z равенство $\psi'(s) = z\sqrt{\mathsf{D}\xi_1}$ определяет s как аналитическую функцию переменного z.

Степенной ряд $\lambda(z) = \lambda_1 + z\lambda_2 + z^2\lambda_3 + \ldots$, определяемый соотношением

$$z^3\lambda(z) = \psi(s) - s\psi'(s) + \frac{1}{2}\psi''(s),$$

называется *рядом Крамера*. Если $\mathsf{M}\xi_n = 0$, то

$$\lambda_1 = \frac{\gamma_3}{3!\sigma^3}, \quad \lambda_2 = \frac{\gamma_4\sigma^2 - 3\gamma_3^2}{4!\sigma^6}, \quad \lambda_3 = \frac{\gamma_5\sigma^2 - 10\gamma_4\gamma_3\sigma + 15\gamma_3^3}{5!\sigma^9}, \ldots,$$

где $\sigma^2 = \mathsf{D}\xi_k$ и γ_i — i-й семиинвариант случайной величины ξ_n.

При выполнении условия Крамера (5.3) для $x \geqslant 0$ и $x = o(\sqrt{n})$ имеют место соотношения

$$\begin{gathered}\frac{1 - F_n(x)}{1 - \Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{x+1}{\sqrt{n}}\right)\right),\\ \frac{F_n(-x)}{\Phi(-x)} = \exp\left\{-\frac{x^3}{\sqrt{n}}\lambda\left(-\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{x+1}{\sqrt{n}}\right)\right).\end{gathered} \tag{5.4}$$

Если случайные величины ξ_n, кроме того, имеют непрерывную и ограниченную на всей оси плотность вероятности $f(x)$, то для $x \geqslant 1$ и $x = o(\sqrt{n})$ имеют место соотношения

$$\begin{gathered}\frac{f_n(x)}{\varphi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right),\\ \frac{f_n(-x)}{\varphi(x)} = \exp\left\{-\frac{x^3}{\sqrt{n}}\lambda\left(-\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right).\end{gathered} \tag{5.5}$$

Соотношения (5.4) и (5.5) носят индивидуальный характер по той причине, что ряд Крамера $\lambda(z)$ определяется всеми семиинвариантами и потому однозначно определяется соответствующей случайной величиной.

4.5.3. Зоны нормальной сходимости и асимптотические разложения. Пусть $\rho(n)$ — как угодно медленно растущая к бесконечности положительная функция и $0 < \alpha < 1/2$. Для того чтобы зоны $[0, n^\alpha\rho(n)]$ и $[-n^\alpha\rho(n), 0]$ были зонами нормальной сходимости, необходимо и достаточно, чтобы

$$\mathsf{M}\exp\left\{|\xi_k|^{4\alpha/(2\alpha+1)}\right\} < \infty. \tag{5.6}$$

При $\alpha < 1/6$ условие (5.6) является необходимым для того, чтобы зоны $[0, n^\alpha\rho(n)]$ и $[-n^\alpha\rho(n), 0]$ были зонами локальной нормальной сходимости, и достаточным, чтобы таковыми были зоны $[0, n^\alpha/\rho(n)]$ и $[-n^\alpha/\rho(n), 0]$.

Пусть $1/6 \leqslant \alpha < 1/2$. Рассмотрим ряд чисел

$$\frac{1}{6}, \frac{1}{4}, \frac{3}{10}, \ldots, \frac{1}{2} \cdot \frac{n+1}{n+3}, \ldots \to \frac{1}{2}, \tag{5.7}$$

и пусть s таково, что

$$\frac{1}{2} \frac{s+1}{s+3} \leqslant \alpha < \frac{1}{2} \frac{s+2}{s+4}. \tag{5.8}$$

Для того чтобы зоны $[0, n^{\alpha}\rho(n)]$ и $[-n^{\alpha}\rho(n), 0]$ были зонами (интегральной) нормальной сходимости, необходимо, чтобы было выполнено условие (5.6) и чтобы все моменты ξ_k вплоть до $(s+3)$-го совпадали с моментами нормального (0, 1) распределения. Эти два условия достаточны, чтобы зоны $[0, n^{\alpha}/\rho(n)]$ и $[-n^{\alpha}/\rho(n), 0]$ были зонами (интегральной) нормальной сходимости.

При выполнении условия (5.6) в зоне $[0, n^{\alpha}/\rho(n)]$ равномерно по x имеют место соотношения

$$\begin{gathered} 1 - F_n(x) \sim [1 - \Phi(x)] \exp\left\{ \frac{x^3}{\sqrt{n}} \lambda^{[s]} \left(\frac{x}{\sqrt{n}} \right) \right\}, \\ F_n(-x) \sim \Phi(-x) \exp\left\{ -\frac{x^3}{\sqrt{n}} \lambda^{[s]} \left(\frac{x}{\sqrt{n}} \right) \right\}, \end{gathered} \tag{5.9}$$

где $\lambda^{[s]}(z)$ — отрезок ряда Крамера длиной s, а s определяется условием (5.8). В отличие от разложений (5.4) соотношения (5.9) имеют собирательный характер, ибо они верны для классов случайных величин, удовлетворяющих условию (5.6) и имеющих одинаковые отрезки ряда Крамера длины s, т. е. одинаковые моменты до $(s+3)$-го порядка включительно.

Литература: [5, 24, 34, 60, 64, 74].

Глава 5. БЕЗГРАНИЧНО ДЕЛИМЫЕ РАСПРЕДЕЛЕНИЯ

5.1. Суммы независимых случайных величин и их распределения

5.1.1. Свертки распределений. Пусть ξ_1 и ξ_2 — две независимые случайные величины со значениями в $\mathbf{R}^m$, а μ_1 и μ_2 — соответственно их распределения, т. е. меры, определенные на борелевских множествах A из $\mathbf{R}^m$ соотношениями $\mu_i(A) = \mathsf{P}\{\xi_i \in A\}$. Тогда распределение суммы $\xi_1 + \xi_2$, которая, очевидно, также является случайной величиной в $\mathbf{R}^m$, задается мерой

$$\mu(A) = \int \mu_1(A - x)\, \mu_2(dx),$$

где $A - x = \{y\colon y + x \in A\}$. Мера μ называется *сверткой мер* μ_1 и μ_2, она может быть еще представлена так:

$$\mu(A) = \iint\limits_{x+y \in A} \mu_1(dx)\, \mu_2(dy). \tag{1.1}$$

Свертка мер μ_1 и μ_2 обозначается $\mu_1 * \mu_2$. Из (1.1) видно, что операция свертки коммутативна: $\mu_1 * \mu_2 = \mu_2 * \mu_1$. Пусть $F_i(x)$ $(x \in \mathbf{R}^m)$ — функция распределения величины ξ_i. Функция распределения величины $\xi_1 + \xi_2$ определяется равенством

$$F(x) = \int F_1(x-y)\, dF_2(y).$$

Функция $F(x)$ называется *сверткой функций распределения* F_1 и F_2 и обозначается $F = F_1 * F_2$. Если существует плотность распределения $f_i(x)$ величин ξ_i, то будет существовать и плотность $f(x)$ их суммы, при этом

$$f(x) = \int f_1(x-y)\, f_2(y)\, dy;$$

f также называется *сверткой* f_1 и f_2.

Заметим, что для существования плотности распределения суммы $\xi_1 + \xi_2$ достаточно, чтобы лишь одно слагаемое имело плотность. Если, например, существует $f_1(x)$, то

$$f(x) = \int f_1(x-y)\, \mu_2(dy).$$

Если $\xi_1, \xi_2, \ldots, \xi_k$ — независимые случайные величины из $\mathbf{R}^m$, то распределение μ их суммы $\xi = \xi_1 + \xi_2 + \ldots + \xi_k$ задается сверткой распределений μ_i отдельных слагаемых:

$$\mu = \mu_1 * \mu_2 * \ldots * \mu_k,$$

где $\mu_1 * \mu_2 * \ldots * \mu_k = (\mu_1 * \ldots * \mu_{k-1}) * \mu_k$ определяется индуктивно. Используя коммутативность и ассоциативность сложения случайных величин, легко убедиться, что и операция свертки обладает этими свойствами.

5.1.2. Характеристическая функция суммы независимых случайных величин. Распределение случайной величины определяется ее характеристической функцией, т. е. преобразованием Фурье соответствующей меры. Оказывается, что при сложении независимых случайных величин характеристическая функция суммы весьма просто выражается через характеристические функции отдельных слагаемых.

Теорема. *Пусть $\xi_1, \xi_2, \ldots, \xi_k$ — независимые случайные величины со значениями в $\mathbf{R}^m$, $f_j(z) = \mathsf{M}e^{i(z, \xi_j)}$ $(z \in \mathbf{R}^m)$ — характеристическая функция величины ξ_j, $\xi = \xi_1 + \ldots + \xi_k$, $f(z) = \mathsf{M}e^{i(z, \xi)}$. Тогда $f(z) = f_1(z) \ldots f_k(z)$.*

Доказательство этого утверждения следует из того, что математическое ожидание произведения независимых случайных величин равно произведению математических ожиданий, и из независимости сомножителей в правой части равенства

$$e^{i(z, \xi)} = \prod_{j=1}^{k} e^{(z, \xi_j)}.$$

Можем установить аналогичное соотношение для преобразований Лапласа: если ξ_j $(j = 1, \ldots, k)$ — независимые неотрицательные случайные величины и $\xi=\xi_1+\ldots+\xi_k$, $\varphi_j(\lambda)=\mathsf{M}e^{-\lambda\xi_j}$, $\varphi(\lambda)=\mathsf{M}e^{-\lambda\xi}$, то $\varphi(\lambda)=\prod_{j=1}^{k}\varphi_j(\lambda)$.

Предположим, что величины ξ_j принимают значения лишь из целочисленной решетки в $\mathbf{R}^m$ (т. е. ξ_j имеют с вероятностью 1 целочисленные координаты). Тогда вместо характеристических функций удобно рассматривать производящие функции

$$h_j(z)=\mathsf{M}z^{\xi_j},$$

где $z=(z^1, \ldots, z^m)$ — точка из $\mathbf{C}^m$ — комплексного m-мерного пространства, $|z^j|=1$, а $z^x=\prod_{k=1}^{m}(z^k)^{x^k}$ $(x\in\mathbf{R}^m)$, $x=(x^1, \ldots, x^m)$ (см. § 3.1).

Обозначим через $h(z)$ производящую функцию величины $\xi=\xi_1+\ldots+\xi_k$. Тогда $h(z)=h_1(z)\ldots h_k(z)$.

5.1.3. **Примеры.** 1. Пусть ξ_1 и ξ_2 — независимые гауссовские величины в $\mathbf{R}^m$, $\mathsf{M}\xi_k=a_k$, B_k — корреляционная матрица величины ξ_k. Тогда характеристическая функция величины ξ_k равна

$$f_k(z)=\exp\left\{i(z, a_k)-\frac{1}{2}(B_kz, z)\right\}.$$

Если $f(z)$ — характеристическая функция величины $\xi_1+\xi_2$, то

$$f(z)=\exp\left\{i(z, a_1+a_2)-\frac{1}{2}((B_1+B_2)z, z)\right\}.$$

Таким образом, $\xi_1+\xi_2$ также имеет гауссовское распределение со средним a_1+a_2 и корреляционной матрицей B_1+B_2.

2. Пусть ξ_1 и ξ_2 — независимые пуассоновские случайные величины с параметрами a_1 и a_2 соответственно. Их характеристические функции имеют вид

$$f_k(t)=\exp\{a_k(e^{it}-1)\}.$$

Характеристическая функция суммы записывается в виде

$$f(t)=\exp\{(a_1+a_2)(e^{it}-1)\}.$$

Сумма также имеет пуассоновское распределение с параметром a_1+a_2.

5.2. Определение и основные свойства безгранично делимых распределений

5.2.1. **Определение.** Распределение вероятностей μ в $\mathbf{R}^m$ называется *безгранично делимым*, если для всякого n можно указать такое распределение μ_n, что μ представимо в виде n-кратной свертки распределения μ_n самого с собой:

$$\mu=\underbrace{\mu_n*\mu_n*\ldots*\mu_n}_{n\text{ раз}}.$$

Таким образом, величина ξ имеет безгранично делимое распределение, если для всякого n существуют независимые одинаково распределенные величины $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nn}$ такие, что

$$\xi = \xi_{n1} + \ldots + \xi_{nn}.$$

Можно дать определение безгранично делимого распределения и в терминах характеристических функций. Пусть $\varphi(z)$ $(z \in R^m)$ — характеристическая функция распределения

$$\varphi(z) = \int e^{i(z, x)} \mu(dx);$$

если μ безгранично делимо, то для всякого n существует характеристическая функция $\varphi_n(z)$ такая, что

$$\varphi(z) = \varphi_n(z)^n.$$

Характеристические функции безгранично делимых распределений называют *безгранично делимыми характеристическими функциями.* Перечислим некоторые их существенные свойства:

1) безгранично делимая характеристическая функция не обращается в нуль;

2) $\arg \varphi(z)$ всегда можно считать непрерывной функцией;

3) если определить для $t > 0$

$$\varphi(z)^t = |\varphi(z)|^t \exp\{it \arg \varphi(z)\},$$

где $\arg \varphi(z)$ — непрерывная функция, то $\varphi(z)^t$ для всех $t > 0$ является характеристической функцией, притом безгранично делимой.

5.2.2. Общий вид безгранично делимой характеристической функции. Для каждой безгранично делимой характеристической функции $\varphi(z)$ в R^m можно указать: 1) $a \in R^m$; 2) линейный неотрицательный оператор B, действующий в R^m; 3) конечную меру $\Pi(\cdot)$ на борелевских множествах R^m, для которой $\Pi(\{0\}) = 0$ ($\{0\}$ — множество, состоящее из одной точки 0) такие, что имеет место формула

$$\varphi(z) = \exp\left\{ i(a, z) - \frac{1}{2}(Bz, z) + \right.$$
$$\left. + \int \left(e^{i(z, x)} - 1 - \frac{i(z, x)}{1 + |x|^2} \right) \frac{1 + |x|^2}{|x|^2} \Pi(dx) \right\}, \quad (2.1)$$

где $|x| = \sqrt{(x, x)}$. Эта формула называется *формулой Леви.* В том случае, когда $\Pi = 0$, $\varphi(z)$ будет характеристической функцией гауссовского распределения.

Формула (2.1) дает каноническое представление безгранично делимой характеристической функции. Три элемента — a, B, Π — определяются характеристической функцией однозначно. Приведем теорему о сходимости безгранично делимых характеристических функций.

Теорема. *Последовательность безгранично делимых характеристических функций может сходиться только к безгранично делимой характеристической функции. Если $\varphi_n(z)$ определяются формулой* (2.1), *в которую вместо a, B и Π подставлены соответственно*

a_n, B_n, Π_n, то $\varphi_n(z)$ сходится к $\varphi(z)$, определяемой формулой (2.1), *тогда и только тогда, когда выполняются условия:*

а) *для всякой непрерывной ограниченной функции $g(x)$ на $\mathbf{R}^m$, для которой $g(0) = 0$,*

$$\lim_{n\to\infty} \int g(x)\, \Pi_n(dx) = \int g(x)\, \Pi(dx);$$

б) $$\lim_{n\to\infty} \left\{ (B_n z, z) + \int \frac{(z, x)^2}{|x|^2}\, \Pi_n(dx) \right\} = (Bz, z) + \int \frac{(z, x)^2}{|x|^2}\, \Pi(dx);$$

в) $\lim a_n = a$.

Для характеристических функций в $\mathbf{R}$ можно уменьшить число определяющих элементов до двух. Для всякой безгранично делимой функции $\varphi(z)$ в $\mathbf{R}$ существуют $\gamma \in \mathbf{R}$ и неубывающая непрерывная справа ограниченная функция $G(x)$, для которой $G(-\infty) = 0$ такие, что

$$\varphi(z) = \exp\left\{ i\gamma z + \int_{-\infty}^{+\infty} \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2}\, dG(x) \right\}. \tag{2.2}$$

При этом подынтегральная функция $\left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) \frac{1+x^2}{x^2}$ при $x = 0$ доопределена по непрерывности равной $-z^2/2$. Формулу (2.2) называют *каноническим представлением Леви — Хинчина*; его элементы γ и G однозначно определяются характеристической функцией. Из теоремы следует, что для сходимости последовательности характеристических функций $\varphi_n(z)$, представимых по формуле (2.2), если в ней γ и G заменить на γ_n и G_n, к функции $\varphi(z)$, задаваемой формулой (2.2), необходимо и достаточно, чтобы выполнялись условия: а) $\gamma_n \to \gamma$; б) $G_n(x) \to G(x)$ почти для всех x и $G_n(+\infty) \to G(+\infty)$.

Приведем еще некоторые формулы для безгранично делимой характеристической функции в одномерном случае.

Каноническое представление Леви:

$$\varphi(z) = \exp\left\{ i\gamma z - \frac{bz^2}{2} + \int_{-\infty}^{0} \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) dN(x) + \int_{0}^{+\infty} \left(e^{izx} - 1 - \frac{izx}{1+x^2} \right) dM(x) \right\}, \tag{2.3}$$

где $\gamma \in \mathbf{R}$, $b > 0$, $N(x)$ и $M(x)$ не убывают соответственно на $(-\infty, 0)$ и $(0, \infty)$, $N(-\infty) = 0$, $M(+\infty) = 0$ и

$$\int_{-1}^{0} x^2\, dN(x) + \int_{0}^{1} x^2\, dM(x) < \infty.$$

Для безгранично делимых распределений в $\mathbf{R}$ с конечной дисперсией характеристическая функция может быть представлена *по формуле Колмогорова*:

$$\varphi(z) = \exp\left\{i\gamma z + \int (e^{izx} - 1 - izx)\frac{1}{x^2}\,dK(x)\right\}, \tag{2.4}$$

где $K(x)$ —неубывающая непрерывная справа ограниченная функция, для которой $K(-\infty) = 0$; подынтегральная функция при $x=0$ доопределяется по непрерывности равной $-z^2/2$.

Если величина ξ неотрицательна и имеет безгранично делимое распределение, то ее характеристическая функция имеет вид

$$\varphi(z) = \exp\left\{i\gamma z + \int_0^\infty (e^{izx} - 1)\,dM(x)\right\}, \tag{2.5}$$

где $\gamma \geqslant 0$, а $M(x)$ — неубывающая функция, для которой $M(+\infty) = 0$, $\int_0^1 x\,dM(x) < \infty$.

Если величина ξ имеет безгранично делимое арифметическое распределение с шагом h, то ее характеристическая функция имеет вид

$$\varphi(z) = \exp\left\{i\gamma z + \sum_{k=-\infty}^{\infty} (e^{izkh} - 1)\,C_k\right\}, \tag{2.6}$$

где γ и k — целые, $C_k \geqslant 0$, $\sum C_k < \infty$.

5.2.3. Примеры безгранично делимых распределений и характеристических функций.

1. *Распределение Пуассона.* Величина ξ имеет арифметическое распределение с шагом 1 и $\mathsf{P}(\xi = k) = a^k e^{-a}/k!$ $(k \geqslant 0)$. Характеристическая функция имеет вид

$$\varphi(z) = \exp\{a(e^{iz} - 1)\}, \tag{2.7}$$

т. е. представима по формуле (2.6) с $h = 1$, $\gamma = 0$, $C_1 = a$, $C_k = 0$, $k \neq 1$.

2. *Обобщенное (сложное) распределение Пуассона.* Пусть ξ_1, ξ_2, ... — последовательность независимых одинаково распределенных величин в $\mathbf{R}^m$ с распределением μ, а ν — не зависящая от них случайная величина, принимающая целые неотрицательные значения. Положим $s_0 = 0$, $s_n = \sum_{k=1}^{n} \xi_k$. Тогда величина $\xi = s_\nu$ имеет обобщенное распределение Пуассона. Если μ^{n*} обозначает n-кратную свертку меры μ, μ^{0*} — распределение величины, равной 0 с вероятностью 1, то

$$\mathsf{P}\{\xi \in A\} = \sum_{n=0}^{\infty} \frac{a^n e^{-a}}{n!}\,\mu^{n*}(A), \tag{2.8}$$

где a — параметр пуассоновского распределения, входящий в формулу (2.7). Характеристическая функция величины ξ имеет вид

$$\varphi(z) = \exp\left\{a \int (e^{izx} - 1)\,\mu(dx)\right\}, \tag{2.9}$$

т. е. может быть представлена по формуле (2.1) с мерой Π, определяемой равенством $\Pi(A) = \int\limits_A \frac{|x|^2}{1+|x|^2}\,\mu(dx)$, $B = 0$ и $a \in \mathbf{R}^m$, для которого $(a, z) = \int \frac{(z, x)}{1+|x|^2}\,\mu(dx)$.

3. *Нормальное распределение.* Характеристическая функция такого распределения в $\mathbf{R}^m$ получается, если в формуле (2.1) положить $\Pi = 0$, а в одномерном случае — если в формуле (2.2) взять $G(x) = 0$ для $x < 0$, $G(x) = G(0)$ для $x > 0$.

4. Γ-*распределение.* Такое распределение в $\mathbf{R}$ задается плотностью:

$$p_\alpha(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-x}, \quad x \geqslant 0, \quad p_\alpha(x) = 0, \quad x < 0, \quad \alpha > 0.$$

Характеристическая функция такого распределения имеет вид

$$\varphi_\alpha(z) = \frac{1}{(1-iz)^\alpha} = \exp\left\{\alpha \int\limits_0^\infty \frac{e^{izx}-1}{x} e^{-x}\,dx\right\} \tag{2.10}$$

и представима по формуле (2.3) с $\gamma = \alpha \int\limits_0^\infty \frac{1}{1+x^2} e^{-x}\,dx$, $b = 0$, $N(x) = 0$, $M(x) = -\int\limits_x^\infty \frac{e^{-y}}{y}\,dy$.

5. *Устойчивые распределения.* Так называется семейство распределений в $\mathbf{R}$, для которых характеристические функции задаются равенством

$$\varphi(z) = \exp\{i\gamma z - C|z|^\alpha (1 + i\beta\omega(z, \alpha))\}, \tag{2.11}$$

где $\gamma \in \mathbf{R}$, $C > 0$, $|\beta| \leqslant 1$, $0 < \alpha \leqslant 2$, $\omega(z, \alpha) = \operatorname{sign} z \operatorname{tg}\frac{\pi}{2}\alpha$, $\alpha \neq 1$, $\omega(z, 1) = \frac{2}{\pi}\ln(z)$.

При $\alpha = 2$ устойчивое распределение является нормальным. Характеристическая функция устойчивого закона при $\alpha < 2$ может быть записана по формуле (2.3), если положить в этой формуле $b = 0$, $N(x) = C_1/|x|^{1+\alpha}$, $M(x) = -C_2/x^{1+\alpha}$, где $C_1 > 0$, $C_2 > 0$ — некоторые постоянные. Для плотностей устойчивых распределений явные выражения отсутствуют, за исключением случаев: 1) $\alpha = 1/2$,

$\beta = \pm 1$; 2) $\alpha = 1$, $\beta = 0$; 3) $\alpha = 2$. Для $\alpha = 1/2$, $\beta = 1$, $\gamma = 0$ плотность имеет вид

$$p(x) = \frac{C}{\sqrt{2\pi}} x^{-3/2} e^{-C^2/2x};$$

при $\alpha = 1$, $\beta = 0$, $\gamma = 0$ получаем распределение Коши с плотностью

$$p(x) = \frac{C}{\pi (x^2 + C^2)}.$$

Для всех устойчивых распределений плотность существует, и ее можно вычислить по формуле обращения, так как $\varphi(z)$ абсолютно интегрируема.

Отметим характерную особенность устойчивых функций распределения: F будет устойчивой функцией распределения, если для любых $a_1 > 0$, $a_2 > 0$, b_1 и b_2 существуют $a > 0$ и b такие, что

$$F(a_1 x + b_1) * F(a_2 x + b_2) = F(ax + b)$$

(другими словами, свертки однотипных распределений приводят к распределению того же типа). С помощью этого свойства определяют иногда класс устойчивых распределений, а затем выводят формулу (2.11) для характеристической функции.

5.3. Предельные теоремы для схемы серий

5.3.1. Общие теоремы. Рассмотрим последовательность серий случайных величин $\xi_{n1}, \ldots, \xi_{nk_n}$ (первый индекс указывает номер серии, второй — номер величины в серии), принимающих значения из $\mathbf{R}^m$ и независимых в каждой серии. Эти величины называются *предельно пренебрегаемыми* (*или бесконечно малыми*), если для всякого $\varepsilon > 0$ выполняется условие

$$\lim_{n \to \infty} \sup_{i \leqslant k_n} \mathsf{P}\{|\xi_{ni}| > \varepsilon\} = 0.$$

В этом пункте формулируются условия, при которых суммы $\zeta_n = \sum_{i=1}^{k_n} \xi_{ni}$ предельно пренебрегаемых величин имеют предельное распределение. Первый важный факт, который здесь установлен, можно сформулировать так: *если предельное распределение величин ζ_n существует, то оно обязательно безгранично делимо.*

Используя этот факт, мы сведем общую задачу о предельных распределениях для сумм независимых случайных величин к следующей: найти условия, которые нужно наложить на распределения отдельных слагаемых, чтобы суммы ζ_n имели в качестве предельного данное безгранично делимое распределение.

Обозначим через μ_{ni} распределение величины ξ_{ni} в $\mathbf{R}^m$ и определим далее такое $a_{ni} \in \mathbf{R}^m$, чтобы для всех $z \in \mathbf{R}^m$ выполнялось равенство

$$(a_{ni}, z) = \int \frac{(x, z)}{1 + (x, x)} \mu_{ni}(dx).$$

Введем на $\mathbf{R}^m$ меру Π_n так, чтобы для всякой ограниченной непрерывной функции $g(x)$

$$\int g(x)\,\Pi_n(dx) = \sum_{j=1}^{k_n} \int g(x-a_{nj}) \frac{|x-a_{nj}|^2}{1+|x-a_{nj}|^2}\,\mu_{nj}(dx).$$

Теорема 1. *Для того чтобы последовательность ν_n распределений величин ζ_n слабо сходилась к безгранично делимому распределению с характеристической функцией $\varphi(z)$, определяемой равенством* (2.1), *необходимо и достаточно выполнения следующих условий*:

а) $\lim\limits_{n\to\infty} \sum\limits_{j=1}^{k_n} a_{nj} = a$;

б) $\lim\limits_{\varepsilon\downarrow 0} \overline{\lim\limits_{n\to\infty}} \left| \sum\limits_{j=1}^{k_n} \int\limits_{|x|\leqslant\varepsilon} (x-a_{nj}, z)^2\,\mu_{nj}(dx) - (Bz, z) \right| = 0$;

в) *для всякой непрерывной ограниченной функции* $g(x)$

$$\lim_{n\to\infty} \int [g(x)-g(0)]\,\Pi_n(dx) = \int [g(x)-g(0)]\,\Pi(dx).$$

Замечание. Если выполняются только два последних условия теоремы, то распределение $\zeta_n - a_n$, где $a_n = \sum\limits_{j=1}^{k_n} a_{nj}$, сходится к безгранично делимому распределению, характеристическая функция которого $\varphi(z)$ задается формулой (2.1), если в ней положить $a=0$. Наоборот, если при некотором выборе векторов $a_n' \in \mathbf{R}^m$ величина $\zeta_n - a_n'$ имеет предельное распределение с характеристической функцией (2.1), то выполняются два последних условия теоремы и

$$\lim_{n\to\infty} \left(\sum_{j=1}^{k_n} a_{nj} - a_n' \right) = a.$$

Для случайных величин, принимающих значения в $\mathbf{R}$, условия сходимости более простые. Пусть $F_{ni}(x)$ — функция распределения величины ξ_{ni},

$$a_{ni} = \int \frac{x}{1+x^2}\,dF_{ni}(x), \quad G_n(x) = \sum_{i=1}^{k_n} \int\limits_{-\infty}^{x} \frac{y^2}{1+y^2}\,dF_{ni}(y+a_{ni}).$$

Теорема 2. *Для того чтобы последовательность $F_n(x)$ функций распределения величин ζ_n слабо сходилась к некоторой предельной функции распределения, необходимо и достаточно выполнения условий*: а) *существует* $\lim\limits_{n\to\infty} \sum\limits_{j=1}^{k_n} a_{nj} = \gamma$; б) *последовательность* *функций $G_n(x)$ слабо сходится к некоторой неубывающей функции $G(x)$. Если эти условия выполнены, то характеристическая функция предельного распределения задается формулой* (2.2).

Замечание. Если выполнено второе условие, то $\zeta_n - \sum_{j=1}^{k_n} a_{nj}$ имеет предельное распределение, характеристическая функция которого определяется формулой (2.2) с $\gamma = 0$.

5.3.2. Применение общих теорем. Используем эти общие результаты для формулировки сходимости к конкретным безгранично делимым распределениям.

1. Условия сходимости к вырожденному распределению. *Пусть задана последовательность серий случайных величин $\xi_{n1}, \ldots, \xi_{nk_n}$ со значениями в $\mathbf{R}^m$, независимых в каждой серии. Для того чтобы существовала такая последовательность векторов $a_n \in \mathbf{R}^n$, что для всякого $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\{|\zeta_n - a_n| > \varepsilon\} = 0,$$

где $\zeta_n = \sum_1^{k_n} \xi_{nl}$ (т. е. $\zeta_n - a_n \to 0$ по вероятности), необходимо и достаточно выполнения условий:

а) $\sum_{j=1}^{k_n} a_{nj} - a_n \to 0;$

б) $\lim_{n\to\infty} \sum_{j=1}^{k_n} \int \frac{|x - a_{nj}|^2}{1 + |x - a_{nj}|^2} \mu_{nj}(dx) = 0$

(обозначения такие же, как в теореме 1).

2. Условия сходимости к нормальному распределению. *Для того чтобы случайная величина ζ_n имела предельное нормальное распределение в $\mathbf{R}^m$ с характеристической функцией*

$$\varphi(z) = \exp\left\{ i(a, z) - \frac{1}{2}(Bz, z) \right\},$$

необходимо и достаточно, чтобы выполнялись условия:

а) $\lim_{n\to\infty} \sum_{j=1}^{k_n} a_{nj} = a;$

б) *для всех $\varepsilon > 0$*

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \mathsf{P}\{|\xi_{nj}| > \varepsilon\} = 0;$$

в) *при некотором $\varepsilon > 0$*

$$\lim_{n\to\infty} \left[(Bz, z) - \sum_{j=1}^{k_n} \int_{|x| \leqslant \varepsilon} (x - a_{nj}, z)^2 \mu_{nj}(dx) \right] = 0.$$

3. Условия сходимости к обобщенному распределению Пуассона. *Последовательность ζ_n имеет предельное обобщенное распределение Пуассона, если выполняются условия:*

а) *существует* $\lim\limits_{n\to\infty} \sum\limits_{j=1}^{k_n} \mathsf{P}\{\xi_{nj} \neq 0\}$;

б) *существует мера* $\nu(dx)$ *на* $\mathbf{R}^m$ *такая, что для всякой ограниченной непрерывной функции* $g(x)$ *на* $\mathbf{R}^m$

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \int [g(x) - g(0)]\, \mu_{nj}(dx) = \int [g(x) - g(0)]\, \nu(dx).$$

Если эти условия выполнены, то характеристическая функция предельного распределения задается формулой

$$\varphi(z) = \exp\left\{ \int \left(e^{i(z,x)} - 1\right) \nu(dx) \right\}.$$

4. **Условия сходимости к арифметическим распределениям.** *Пусть величины ξ_{nj} принимают только целые значения. Для того чтобы ζ_n имели предельное распределение, необходимо и достаточно, чтобы выполнялись условия:*

а) *существует* $\lim\limits_{n\to\infty} \sum\limits_{j=1}^{k_n} \mathsf{P}\{\xi_{nj} \neq 0\} = C$;

б) *для всякого целого m существует*

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \mathsf{P}\{\xi_{nj} = m\} = C_m$$

и $C = \sum\limits_m C_m$. *Если эти условия выполнены, то характеристическая функция предельного распределения имеет вид*

$$\varphi(z) = \exp\left\{ -C + \sum_m C_m e^{izm} \right\} = \exp\left\{ \sum_m C_m \left(e^{izm} - 1\right) \right\}.$$

Замечание. Если положить во втором условии $C_m = 0$ для $m \neq 1$, получим условия сходимости к распределению Пуассона.

5.4. Предельные теоремы для нарастающих сумм в R

5.4.1. Теорема для нарастающих сумм. Будем рассматривать последовательность независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ в $\mathbf{R}$. Нас будут интересовать следующие вопросы: когда существуют такие постоянные A_n и B_n, что величины

$$\zeta_n = \frac{1}{B_n}\left(\sum_{k=1}^{n} \xi_k - A_n\right) \tag{4.1}$$

имеют предельное распределение; как выбрать постоянные A_n и B_n; каким будет это предельное распределение? Эту задачу можно свести к задаче суммирования величин в схеме серий, если положить

$$\xi_{nk} = \frac{1}{B_n}(\xi_k - a_{nk}), \tag{4.2}$$

где a_{nk} такие, что $\sum\limits_{k=1}^{n} a_{nk} = A_n$. Если величины (4.2) будут предельно пренебрегаемыми при некотором выборе a_{nk}, то они будут предельно пренебрегаемыми, если положить $a_{nk} = m_k$, где m_k — медиана величины ξ_n, т. е. такое число, что $\mathsf{P}\{\xi_k \geqslant m_k\} \geqslant 1/2$, $\mathsf{P}\{\xi_k \leqslant m_k\} \geqslant 1/2$. Постоянные B_n следует выбрать так, чтобы существовал отличный от нуля конечный предел

$$\lim_{n\to\infty} \sum_{k=1}^{n} \inf_{\alpha} \int \frac{(x-\alpha)^2}{B_n^2 + (x^2-\alpha)^2}\, dF_k(x+m_k), \tag{4.3}$$

где $F_k(x)$ — функция распределения величины ξ_k. После того как выбраны B_n, можно взять

$$A_n = \sum_{k=1}^{n} \left[m_k + B_n \int \frac{x}{B_n^2 + x^2}\, dF_k(x+m_k) \right]. \tag{4.4}$$

Теорема 1. *Если для данной последовательности ξ_k можно выбрать B_n так, чтобы выполнялось условие* (4.3), *и при этом для всякого $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\{|\xi_k - m_k| > \varepsilon B_n\} = 0,$$

то для существования у ζ_n предельного распределения необходимо и достаточно, чтобы выполнялись следующие условия: существует неубывающая функция $G(x)$ такая, что $G(-\infty) = 0$, $G(+\infty) < \infty$ и почти для всех y

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int_{-\infty}^{yB_n} \frac{x^2}{B_n^2 + x^2}\, dF_k(x + m_k + \alpha_{nk}) = G(y),$$

а

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int_{-\infty}^{\infty} \frac{x^2}{B_n^2 + x^2}\, dF_k(x + m_k + \alpha_{nk}) = G(+\infty),$$

где

$$\alpha_{nk} = B_n \int_{-\infty}^{\infty} \frac{x^2}{B_n^2 + x^2}\, dF_k(x + m_k).$$

Если это условие выполнено, то характеристическая функция предельного закона задается формулой (2.2) *с $\gamma = 0$.*

В случае, когда выполнены условия теоремы 1, в качестве предельного распределения для величины ζ_n могут выступать не все безгранично делимые распределения, а лишь некоторый класс таких распределений, называемый *классом* L.

Распределения класса L характеризуются следующим свойством: для них функция G в формуле (2.2) обязательно в каждой точке $x \neq 0$ имеет левую и правую производные и функция $\frac{1+x^2}{x} G'(x)$ не возрастает при $x < 0$ и не убывает при $x > 0$ (при этом $G'(x)$ может обозначать любую производную — левую или правую, возможно, не одну и ту же в различных точках). Если использовать представление (2.3) для характеристических функций безгранично делимых распределений, то класс L совпадает с множеством тех распределений, для которых функции N и M в формуле (2.3) логарифмически выпуклы, т. е. $N(-e^{-x})$ и $M(e^x)$ выпуклы (первая — вниз, вторая — вверх). Примерами распределений класса L могут служить устойчивые (в том числе нормальные) распределения.

5.4.2. Применение общей теоремы.

1. *Сходимость к вырожденному распределению.* Выбором достаточно быстро растущих постоянных B_n можно добиться того, чтобы ζ_n сходилась к нулю по вероятности. Этот факт не представляет интереса, если $A_n/B_n \to 0$. Если же последнее условие не выполняется, то можно указать такие постоянные C_n, что $\frac{1}{C_n}\sum_{k=1}^{n}\xi_k - 1$ сходится по вероятности к нулю. В этом случае постоянные C_n в определенном смысле характеризуют рост случайных сумм $\sum_{k=1}^{n}\xi_k$, а сами суммы называются *относительно устойчивыми.*

Пусть постоянные A_n и B_n выбраны в соответствии с формулами (4.3) и (4.4). Если $\lim\limits_{n\to\infty}\frac{A_n}{B_n} = \infty$, то для всякого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{1}{A_n}\sum_{k=1}^{n}\xi_k - 1\right| > \varepsilon\right\} = 0. \tag{4.5}$$

Возьмем другое условие относительной устойчивости. Пусть существуют такие постоянные C_n, что выполнены условия:

1) $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}\mathsf{P}\{|\xi_k| > C_n\} = 0$;

2) если $F_k(x)$ — функция распределения величины ξ_k, то

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{C_n}\int_{-C_n}^{C_n} x\,dF_k(x) = +\infty.$$

Тогда соотношение (4.5) выполняется, если положить

$$A_n = \sum_{k=1}^{n} \int_{-C_n}^{C_n} x\, dF_k(x).$$

Предположим теперь, что величины ξ_k одинаково распределены и неотрицательны, $\mathsf{P}\{\xi_k > 0\} > 0$. Для относительной устойчивости сумм ζ_n необходимо и достаточно выполнения следующего условия:

$$\lim_{t\to\infty} \frac{1}{t[1-F(t)]} \int_0^t x\, dF(x) = +\infty$$

(в случае, когда $1 - F(t_0) = 0$ при некотором t_0, считаем, что $C/0 = +\infty$ при $C > 0$, так что это условие выполняется). Постоянные A_n такие, что выполняется соотношение (4.5), можно взять равными $A_n = n \int_0^{C_n} x\, dF(x)$, если C_n таковы, что $A_n/C_n \to +\infty$ и $\lim_{n\to\infty} n(1 - F(C_n)) = 0$. Здесь $F(x)$ — общая функция распределения случайных величин ξ_k.

Пример. Пусть величины ξ_k принимают значения 2^n ($n = 0, 1, 2, \ldots$) с вероятностью $1/2^{n+1}$. Тогда для $2^n \leqslant t < 2^{n+1}$

$$\frac{1}{t[1-F(t)]} \int_0^t x\, dF(x) = \frac{2^n}{t} n \to +\infty.$$

Если взять $A_n = n \ln n$, то будет выполнено (4.5); C_n здесь можно взять, например, равным $n\sqrt{\ln n}$.

2. *Сходимость к нормальному распределению.* Пусть ξ_k — последовательность независимых случайных величин с функциями распределения $F_k(x)$. Не ограничивая общности, будем считать, что их медианы $m_k = 0$ (иначе можно было бы рассматривать величины $\xi_k - m_k$).

Для того чтобы существовали такие постоянные A_n и B_n, что случайная величина

$$\zeta_n = \frac{1}{B_n}\left(\sum_1^n \xi_k - A_n\right)$$

при $n \to \infty$ имеет предельное нормальное распределение и для всех $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\{|\xi_k| > \varepsilon B_n\} = 0,$$

необходимо и достаточно, чтобы существовали такие постоянные C_n, что

1) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \mathsf{P}\{|\xi_k| > C_n\} = 0,$

2) $\lim\limits_{n\to\infty} \frac{1}{C_n^2} \sum\limits_{k=1}^{n} \left\{ \int\limits_{|x|<C_n} x^2\, dF_k(x) - \left(\int\limits_{|x|<C_n} x\, dF_k(x) \right)^2 \right\} = +\infty.$

При этом можно положить

$$B_n = \sum_{k=1}^{n} \left\{ \int\limits_{|x|<C_n} x^2\, dF_k(x) - \left(\int\limits_{|x|<C_n} x\, dF_k(x) \right)^2 \right\},$$

$$A_n = \sum_{k=1}^{n} \int\limits_{|x|<C_n} x\, dF_k(x).$$

Характеристическая функция предельного нормального закона будет иметь вид $\varphi(z) = e^{-z^2/2}$.

5.4.3. Предельные теоремы для одинаково распределенных слагаемых. Пусть $\xi_1, \xi_2, \ldots$ независимы и одинаково распределены. Укажем условия, при которых существуют такие постоянные A_n и B_n, что величина

$$\zeta_n = \frac{1}{B_n}\left(\sum_{1}^{n} \xi_k - A_n \right)$$

при $n \to \infty$ имеет предельное распределение. В этом пункте нас будут интересовать предельные распределения, отличные от вырожденного и нормального, так как сходимость к последним уже рассматривалась ранее. Кроме того, ограничимся лишь случаем величин со значениями в R.

Теорема 2. *Если величины ζ_n имеют предельное распределение, то оно обязательно устойчиво.*

Функция $h(t)$, определенная для $t > 0$ (либо для всех целых $t > 0$), называется *регулярно меняющейся*, если для всех $k > 0$ существует

$$\lim \frac{h(t_1)}{h(t_2)}$$

при $t_1 \to \infty$, $t_1/t_2 \to k$. Этот предел (он, естественно, зависит лишь от k) обязательно имеет вид k^α $(-\infty < \alpha < +\infty)$. Показатель α называется *степенью регулярно меняющейся функции*; при $\alpha = 0$ функция называется *медленно меняющейся*. Регулярно меняющаяся функция степени α представима в виде $h(t) = t^\alpha h_0(t)$, где $h_0(t)$ — медленно меняющаяся функция.

Примерами медленно меняющихся функций служат:

а) функция $h(t)$, для которой $\lim\limits_{t\to\infty} h(t)$ существует и не равен 0;

б) $h(t) = (\log t)^\beta$, каков бы ни был показатель β;

в) $h(t) = [\log\log(t+1)]^\beta$ для всех β.

Общий вид медленно меняющейся функции дает формула

$$h(t) = C(t) \exp\left\{\int_{t_0}^{t} \frac{\alpha(z)}{z}\, dz\right\},$$

где существуют $\lim\limits_{t\to+\infty} C(t) \neq 0$ и $\lim\limits_{z\to+\infty} \alpha(z) = 0$.

Условия сходимости к устойчивым законам с показателем α дает следующая теорема.

Теорема 3. *Для того чтобы существовали постоянные A_n и B_n такие, что величины ζ_n имели предельное устойчивое распределение с показателем α, необходимо и достаточно, чтобы*

а) *функция $h(t) = 1 - F(t) + F(-t)$ $(t > 0)$ была регулярно меняющейся степени α $(0 < \alpha < 2)$;*

б) *существовал предел*

$$\lim_{t\to\infty} \frac{1 - F(t)}{h(t)} = \lambda.$$

Если эти условия выполнены, то характеристическая функция предельного устойчивого закона будет иметь вид

$$\varphi(z) = \exp\left\{i\gamma z - C|z|^\alpha [1 + i\beta\omega(z, \alpha)]\right\} \tag{4.6}$$

(см. формулу (2.11)), *где γ и C — некоторые постоянные, зависящие от выбора постоянных A_n и B_n; α — то, что упомянуто в условии* а); *$\beta = 2\lambda - 1$, где λ взято из условия* б).

Выбор постоянных B_n можно осуществить способом, не зависящим от значений α и λ, именно, B_n можно выбрать так, чтобы

$$\lim_{n\to\infty} nh(B_n) = 1 \tag{4.7}$$

(из регулярности функции h следует, что это всегда возможно).

При выборе A_n следует рассмотреть три случая.

1) $\alpha < 1$, $A_n = 0$. Если B_n выбрано в соответствии с (4.7), то в формуле (4.6) $\gamma = 0$, $C = \dfrac{\Gamma(1-\alpha)}{\alpha} \cos\dfrac{\pi}{2}\alpha$ (Γ — гамма-функция Эйлера).

2) $1 < \alpha < 2$. В этом случае существует $\mathsf{M}\xi_k = a$ и $A_n = na$. При таком выборе A_n и B_n будем иметь в (4.6) $\gamma = 0$, $C = -\dfrac{\Gamma(2-\alpha)}{\alpha(\alpha-1)} \cos\dfrac{\pi}{2}\alpha$ (Γ — гамма-функция Эйлера).

3) $\alpha = 1$. В этом случае можно взять $A_n = nB_n \displaystyle\int \frac{x}{x^2 + B_n^2} \times$ $\times\, dF(x)$ (B_n определяются из (4.7)). При таком выборе постоянных A_n и B_n будем иметь

$$C = \frac{\pi}{2}, \quad \gamma = \int_0^\infty \left[\frac{\sin v}{v^2} - \frac{1}{v(1+v^2)}\right] dv.$$

Литература: [23, 24, 33, 34, 51, 60, 64, 68, 89, 94].

Глава 6. ОСНОВНЫЕ ВЕРОЯТНОСТНЫЕ РАСПРЕДЕЛЕНИЯ

В этой главе систематизированы результаты о функциях, определяемых случайными величинами и случайными векторами, и приводятся основные сведения о наиболее важных вероятностных распределениях.

6.1. Характеристики случайных величин

6.1.1. Функции распределения и плотность распределения. Пусть ξ — случайная величина. *Функцией распределения* случайной величины ξ называется функция

$$F_\xi(x) = \mathsf{P}\{\xi < x\}.$$

Имеют место следующие свойства:

а) $0 \leqslant F_\xi(x) \leqslant 1$;

б) $F_\xi(x)$ не убывает, непрерывна слева *) и удовлетворяет условиям

$$F_\xi(-\infty) = \lim_{x\to\infty} F_\xi(-x) = 0, \quad F_\xi(+\infty) = \lim_{x\to\infty} F_\xi(x) \leqslant 1$$

(в случае $F_\xi(\infty) = 1$ случайную величину ξ называют *собственной*, в случае $F_\xi(\infty) \leqslant 1$ — *несобственной*, при этом $1 - F_\xi(\infty) = \mathsf{P}\{\xi = \infty\}$);

в) $F_\xi(x)$ имеет не более чем счетное число разрывов (скачков);

г) $F_\xi(x)$ может быть представлена в виде

$$F_\xi(x) = \alpha F_\xi^{(d)}(x) + \beta F_\xi^{(ac)}(x) + \gamma F_\xi^{(s)}(x) \tag{1.1}$$

(разложение Лебега), где $\alpha \geqslant 0$, $\beta \geqslant 0$, $\gamma \geqslant 0$, $\alpha + \beta + \gamma = 1$, $F_\xi^{(d)}(x)$ — ступенчатая функция распределения, $F_\xi^{(ac)}(x)$ — абсолютно непрерывная функция распределения, $F_\xi^{(s)}(x)$ — сингулярная функция распределения (т. е. непрерывная функция распределения такая, что $\frac{d}{dx} F_\xi^{(s)}(x) = 0$ почти всюду по мере Лебега).

Распределения, порождаемые функциями распределения $F_\xi^{(d)}(x)$ $F_\xi^{(ac)}(x)$, $F_\xi^{(s)}(x)$, называются соответственно *дискретными*, *абсолютно непрерывными* (относительно меры Лебега) и *сингулярными*. Если $F_\xi(x) = F_\xi^{(ac)}(x)$, то существует неотрицательная функция $f_\xi(x)$ такая, что

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(u)\,du,$$

*) Если функция распределения определена равенством $F_\xi(x) = \mathsf{P}\{\xi \leqslant x\}$, то $F_\xi(x)$ непрерывна справа.

которая называется *плотностью распределения* (порождаемого $F_\xi(z)$). Таким образом, любая функция распределения является смесью дискретной, непрерывной и сингулярной функций распределения.

6.1.2. Моменты и семиинварианты. Обозначим через m_k и μ_k соответственно k-й начальный и k-й центральный моменты случайной величины ξ (см. п. 1.6.4):

$$m_k = \mathrm{M}\xi^k = \int_{-\infty}^{+\infty} x^k \, dF_\xi(x),$$

$$\mu_k = \mathrm{M}(\xi - \mathrm{M}\xi)^k = \int_{-\infty}^{\infty} (x - \mathrm{M}\xi)^k \, dF_\xi(x).$$

Начальные и центральные моменты связаны очевидными соотношениями ($m_0 = \mu_0 = 1$)

$$\mu_k = \sum_{j=0}^{k} C_k^j (-m_1)^j \mu_{k-j}, \quad m_k = \sum_{j=0}^{k} C_k^j m_1^j \mu_{k-j}.$$

Пусть $\nu_k = \mathrm{M}|\xi|^k$ — k-й абсолютный момент. Имеет место неравенство

$$\nu_{k_2}^{k_3-k_1} \leqslant \nu_{k_1}^{k_3-k_2} \nu_{k_3}^{k_2-k_1}$$

для любых $0 \leqslant k_1 \leqslant k_2 \leqslant k_3$ (*неравенство Ляпунова*), $\nu_1 \leqslant \sqrt{\nu_2} \leqslant \sqrt[3]{\nu_3} \leqslant \ldots \leqslant \sqrt[r]{\nu_r}$.

В связи с моментами возникают следующие два вопроса (так называемая *проблема моментов*).

Имеются константы $a_0, a_1, a_2, \ldots$.

1. При каких дополнительных условиях эти константы являются моментами некоторой случайной величины ξ?

2. Если эти константы являются моментами некоторой случайной величины ξ, то может ли существовать другая случайная величина η с теми же моментами? Если да, то при каких дополнительных условиях моменты определяют случайную величину ξ единственным образом?

Ответ на поставленные вопросы дает следующая теорема.

Теорема. *Последовательность чисел $a_0, a_1, a_2, \ldots$ с $a_0 = 1$ является последовательностью моментов некоторой случайной величины ξ, принимающей значения в* $[0, 1]$ *тогда и только тогда, когда*

$$(-1)^r \Delta^r a_k \geqslant 0, \quad k \geqslant 1, \quad r \geqslant 1,$$

где $\Delta^r a_k$ — r-я разность, определяемая рекуррентно:

$$\Delta^1 a_k = a_{k+1} - a_k,$$

$$\Delta^2 a_k = \Delta(\Delta a_k) = a_{k+2} - 2a_{k+1} + a_k,$$

$$\cdots\cdots\cdots$$

$$\Delta^r a_k = \Delta(\Delta^{r-1} a_k).$$

В общем случае ответ на первую половину второго вопроса утвердительный; например, у семейства плотностей

$$f_\varepsilon(x)=\begin{cases} c\exp\{-\alpha x^\lambda\}\{1+\varepsilon\sin\beta x^\lambda\}, & x\geqslant 0,\\ 0, \quad x<0,\end{cases}$$

где c — нормирующая константа,

$$\lambda\in(0,\ 1/2),\quad \alpha>0,\quad \beta=\alpha\operatorname{tg}\lambda\pi,\quad |\varepsilon|<1,$$

моменты совпадают для всех ε.

Для абсолютно непрерывных распределений моменты однозначно определяют распределение, если

1) в случае $x\in(-\infty,\ 0)$

$$f_\xi(x)<M\,|\,x\,|^{\beta-1}\exp\{-\alpha\,|\,x\,|^\lambda\}$$

для M, α, $\beta>0$, $\lambda\geqslant 1$ и всех $|x|>x_0$ для некоторого $x_0>0$;

2) в случае $x\in[0,\ \infty)$

$$f_\xi(x)<Mx^{\beta-1}\exp\{-\alpha x^\lambda\}$$

для M, α, $\beta>0$, $\lambda\geqslant 1/2$ и всех $x>x_0$, начиная с некоторого $x_0>0$ (критерий Стилтьеса).

Для произвольных распределений, сосредоточенных на $(-\infty,\infty)$, моменты однозначно определяют распределение, если, например,

а) $\limsup\limits_{n\to\infty}\dfrac{\sqrt[n]{\nu_n}}{n}<\infty$, либо

б) $\limsup\limits_{n\to\infty}\dfrac{\sqrt[2n]{\mu_{2n}}}{2n}<\infty$, либо

в) $\sum\limits_{n=0}^{\infty}\dfrac{1}{\sqrt[2n]{\mu_{2n}}}=\infty$.

6.1.3. Характеристики формы и расположения. Если случайная величина ξ абсолютно непрерывна, то значения x, в которых плотность $f_\xi(x)$ достигает своего максимального значения, называются *модами*. Если мода единственна, то распределение случайной величины называют *унимодальным*, в противном случае — *мультимодальным*.

Если ξ — дискретная случайная величина и $\mathsf{P}\{\xi=x_i\}=p_i$, то ее модами называются те значения x_j, для которых

$$\mathsf{P}\{\xi=x_j\}=\max_i p_i.$$

Медианой случайной величины ξ называется значение x', для которого

$$F_\xi(x')\leqslant 1/2,\quad F_\xi(x'+0)\geqslant 1/2.$$

Для случайной величины ξ с абсолютно непрерывным распределением медиана определяется как значение x', для которого

$$\int_{-\infty}^{x'} f(x)\,dx = \int_{x'}^{\infty} f(x)\,dx = 1/2.$$

Квантиль порядка α, $\alpha \in (0, 1)$, есть значение x_α, для которого

$$F_\xi(x_\alpha) \leqslant \alpha, \quad F_\xi(x_\alpha + 0) \geqslant \alpha.$$

Если ξ — случайная величина с абсолютно непрерывным распределением, то квантиль x_α порядка α определяется равенством

$$F_\xi(x_\alpha) = \alpha.$$

Медиана является квантилью порядка 1/2.

Если случайная величина ξ имеет конечные моменты до четвертого включительно, то величина

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mathsf{M}(\xi - \mathsf{M}\xi)^3}{\sqrt{(\mathsf{D}\xi)^3}}$$

называется *коэффициентом асимметрии*, а

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mathsf{M}(\xi - \mathsf{M}\xi)^4}{(\mathsf{D}\xi)^2} - 3$$

есть *коэффициент эксцесса* ее распределения. Эти величины характеризуют степень отличия функции распределения $F_\xi(x)$ от функции распределения $\Phi(x)$ стандартного нормального распределения, для которого коэффициенты асимметрии и эксцесса равны нулю.

6.2. Дискретные распределения

6.2.1. Вырожденное распределение. 1. Случайная величина ξ имеет вырожденное распределение, сосредоточенное в a, если

$$\mathsf{P}\{\xi = a\} = 1.$$

Функция распределения:

$$F(x) = \begin{cases} 0, & x \leqslant a, \\ 1, & x > a. \end{cases}$$

2. Характеристическая функция:

$$\varphi(t) = e^{ita}.$$

Моменты:

$$\mathsf{M}\xi^k = a^k; \quad \mathsf{D}\xi = 0.$$

3. Вырожденное распределение описывает неслучайные величины. Верно обратное утверждение: если случайная величина ξ имеет конечное математическое ожидание и нулевую дисперсию, то

$$\mathsf{P}\{\xi = \mathsf{M}\xi\} = 1.$$

6.2.2. Распределение Бернулли. 1. Случайная величина ξ имеет распределение Бернулли с параметром p $(0 < p < 1)$, если

$$\mathsf{P}\{\xi = 1\} = p, \quad \mathsf{P}\{\xi = 0\} = 1 - p.$$

Функция распределения:

$$F(x) = \begin{cases} 0, & x \leqslant 0, \\ 1 - p, & 0 < x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

2. Характеристическая функция:

$$\varphi(t) = 1 + p(e^{it} - 1).$$

Моменты:

$$\mathsf{M}\xi^k = p; \quad \mathsf{D}\xi = p(1 - p).$$

3. Распределение Бернулли играет фундаментальную роль в теории вероятностей и математической статистике, являясь моделью любого случайного эксперимента, исходы которого принадлежат двум взаимно исключающим классам.

6.2.3. Биномиальное распределение. 1. Случайная величина ξ имеет биномиальное распределение с параметрами (n, p) $(0<p<1$, $n \geqslant 1)$, если

$$\mathsf{P}\left\{\xi = k\right\} = C_n^k p^k (1 - p)^{n-k}, \quad k = 0, \ldots, n.$$

Функция распределения:

$$F(x) = \begin{cases} \sum\limits_{k=1}^{l} C_n^k p^k (1 - p)^{n-k}, & l < x \leqslant l + 1, \\ 1, \quad x > n, \\ 0, \quad x \leqslant 0. \end{cases}$$

2. Характеристическая функция:

$$\varphi(t) = [1 + p(e^{it} - 1)]^n.$$

Моменты:

$$\mathsf{M}\xi = np, \quad \mathsf{M}\xi^2 = np + n(n-1)p^2, \quad \mathsf{M}\xi^3 = np(1-p)(1-2p),$$
$$\mathsf{M}\xi^4 = 3n^2p^2(1 - p^2) + np(1 - p)(1 - 6p(1 - p)); \quad \mathsf{D}\xi = np(1 - p).$$

Мода: $m = p(n + 1) - 1 \leqslant x \leqslant p(n + 1)$.

Коэффициент асимметрии: $\gamma_1 = (q - p)/\sqrt{np(1 - p)}$; коэффициент эксцесса: $\gamma_2 = \dfrac{1 - 6p(1 - p)}{np(1 - p)}$.

Центральные моменты $\mu_k = \mathsf{M}(\xi - \mathsf{M}\xi)^k$ могут быть вычислены по формуле

$$\mu_{k+1} = p(1 - p)\left[n^k \mu_{k-1} + \frac{d\mu_k}{dp}\right].$$

3. Биномиальное распределение является моделью случайных экспериментов, состоящих из n независимых однородных испытаний

Бернулли: если ξ_k $(k=1, \ldots, n)$ независимы и имеют распределение Бернулли с параметром p, то случайная величина $\xi = \sum_{k=1}^{n} \xi_k$ имеет биномиальное распределение с параметрами (n, p).

4. Если p таково, что $np(1-p) > 9$ и $1/(n+1) < p < n/(n+1)$, то можно пользоваться следующими приближенными формулами:

$$B_p(n, k) = \mathsf{P}\{\xi = k\} \approx \frac{1}{\sqrt{np(1-p)}} \varphi\left(\frac{x-np}{\sqrt{np(1-p)}}\right),$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ — плотность стандартного нормального распределения, либо

$$B_p(n, k) \approx \Phi\left(\frac{x+0{,}5-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{x-0{,}5-np}{\sqrt{np(1-p)}}\right),$$

где $\Phi(x) = \int\limits_{-\infty}^{x} \varphi(u)\, du = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-u^2/2}\, du$ — функция стандартного нормального распределения.

При тех же значениях p для функции распределения $F(x)$ можно использовать приближение

$$F(x) \approx \Phi\left(\frac{x+0{,}5-np}{\sqrt{np(1-p)}}\right).$$

Если $np^{3/2} > 1{,}07$, то ошибка при использовании нормальной функции распределения вместо биномиальной не превосходит 0,05 при всех x.

Если p имеет одинаковый с $1/n$ порядок при больших n либо $p < 0{,}1$, можно использовать приближение распределением Пуассона

$$B_p(n, k) \approx \frac{(np)^k}{k!} e^{-np}, \quad F(k) \approx \sum_{l=0}^{k} \frac{(np)^l}{l!} e^{-np}.$$

Пусть $F_{\alpha\beta}(x)$ — функция распределения бета-распределения с параметрами α и β. Тогда

$$\mathsf{P}\{\xi \leqslant k\} = F_{n-k,\, k+1}(1-p).$$

Если η_{m_1, m_2} — случайная величина, имеющая F-распределение с (m_1, m_2) степенями свободы, то

$$\mathsf{P}\{\xi \leqslant k\} = \mathsf{P}\left\{\eta_{2(n-k),\, 2(k+1)} \leqslant \frac{(k+1)(1-p)}{(n-k)p}\right\}.$$

6.2.4. Отрицательное биномиальное распределение (распределение Паскаля). 1. Случайная величина ξ имеет отрицательное биномиальное распределение (распределение Паскаля) с параметрами (r, p), если *)

$$\mathsf{P}\{\xi=k\}=C_{r+k-1}^{k}p^{r}(1-p)^{k}, \quad k=0,\ 1,\ 2,\ \ldots$$

2. Характеристическая функция:

$$\varphi(t)=\left[\frac{p}{1-(1-p)e^{it}}\right]^{r}.$$

Моменты:

$$\mathsf{M}\xi=r(1-p)/p;\quad \mathsf{D}\xi=r(1-p)/p^{2},$$

$$\mathsf{M}(\xi-\mathsf{M}\xi)^{3}=\frac{r(1-p)(2-p)}{p^{3}},$$

$$\mathsf{M}(\xi-\mathsf{M}\xi)^{4}=\frac{r(1-p)}{p^{4}}(3r(1-p)+6(1-p)+p^{2}).$$

Коэффициент асимметрии: $\gamma_1=\dfrac{2-p}{\sqrt{r(1-p)}}$; коэффициент эксцесса: $\gamma_2=\dfrac{6}{r}+\dfrac{p^2}{r(1-p)}$.

3. При натуральном r отрицательное биномиальное распределение описывает число испытаний в схеме Бернулли, необходимых для того, чтобы получить значение 1 ровно r раз.

Если случайные величины ξ_k $(k=0,\ 1,\ 2,\ \ldots)$ независимы и имеют логарифмическое распределение, то случайная величина $\xi=\sum\limits_{k=0}^{\nu}\xi_k$, где ν не зависит от ξ_k и распределена по закону Пуассона с параметром λ, имеет отрицательное биномиальное распределение с параметром $r=-\lambda/\ln p$.

Еще одна характеризация отрицательного биномиального распределения состоит в следующем.

Пусть μ — случайная величина, имеющая Γ-распределение с плотностью $f(x)=\dfrac{1}{\Gamma(\alpha)}\lambda^{\alpha}x^{\alpha-1}e^{-\lambda x}$ $(x\geqslant 0)$. Пусть η — случайная величина, при фиксированном значении $\mu=m$ имеющая распределение Пуассона с параметром m: $\mathsf{P}(\eta=k\mid\mu=m)=e^{-m}\dfrac{m^k}{k!}$ $(k=0,\ 1,\ \ldots)$. Если положить $\lambda=p/(1-p)$, $\alpha=r$, то безусловное распределение случайной величины η будет отрицательным биномиальным:

$$\mathsf{P}(\eta=k)=C_{r+k-1}^{k}p^{r}(1-p)^{k}.$$

---

*) Для нецелых r величина C_{r+k-1}^{k} определяется так: $C_{r+k-1}^{k}=\dfrac{(r+k-1)(r+k-2)\ldots r}{k!}$.

В такой интерпретации отрицательное биномиальное распределение имеет приложения к статистике несчастных случаев и заболеваний, к задачам, связанным с количеством особей данного вида в выборках из биологических популяций, и т. д.

4. При фиксированном p отрицательное биномиальное распределение безгранично делимо.

6.2.5. Геометрическое распределение. 1. Случайная величина ξ имеет геометрическое распределение с параметром p $(0 < p \leqslant 1)$, если

$$\mathsf{P}\{\xi = k\} = p(1-p)^k, \quad k = 0, 1, 2, \ldots$$

2. Характеристическая функция:

$$\varphi(t) = \frac{p}{1-(1-p)e^{it}}.$$

Моменты:

$$\mathsf{M}\xi = (1-p)/p; \quad \mathsf{D}\xi = (1-p)/p^2,$$

$$\mathsf{M}(\xi - \mathsf{M}\xi)^3 = \frac{(1-p)(2-p)}{p^3}, \quad \mathsf{M}(\xi - \mathsf{M}\xi)^4 = \frac{9(1-p)^2}{p^4} + \frac{1-p}{p^2}.$$

Коэффициент асимметрии: $\gamma_1 = \dfrac{2-p}{\sqrt{1-p}}$; коэффициент эксцесса: $\gamma_2 = 6 + \dfrac{p^2}{1-p}$.

3. Является частным случаем отрицательного биномиального распределения с $r = 1$. Описывает число испытаний в схеме Бернулли, необходимых для того, чтобы получить значение 1 ровно один раз.

4. Важность геометрического распределения объясняется свойством, называемым *отсутствием последействия*: для любых $m, n \geqslant 0$

$$\mathsf{P}\{\xi \geqslant m+n \mid \xi \geqslant m\} = \mathsf{P}\{\xi \geqslant n\}.$$

6.2.6. Гипергеометрическое распределение. 1. Случайная величина ξ имеет гипергеометрическое распределение с параметрами (N, p, n) $(0 < p < 1)$, если

$$\mathsf{P}\{\xi = k\} = \frac{C_{Np}^k C_{N(1-p)}^{n-k}}{C_N^n}, \quad k = 0, \ldots, n.$$

2. Характеристическая функция:

$$\varphi(t) = \frac{[N(1-p)]^{[n]}}{N^{[n]}} \sum_{l=0}^{n} \frac{[Np]^{[l]} n^{[l]} e^{ilt}}{[N(1-p)-n+l]^{[l]} l!},$$

где $C^{[l]} = C(C-1)(C-2)\ldots(C-l+1)$; $\varphi(t)$ является решением дифференциального уравнения

$$(1-e^{it})\left\{\frac{d^2\varphi}{dt^2} - (n+Np)\frac{d\varphi}{dt} + Npn\varphi\right\} - Npn\varphi + N\frac{d\varphi}{dt} = 0.$$

Моменты:

$$\mathsf{M}\xi = np; \quad \mathsf{D}\xi = \frac{N-n}{N-1}\, np\,(1-p),$$

$$\mathsf{M}\,(\xi - \mathsf{M}\xi)^3 = \frac{(N-n)\,(N-2n)}{(N-1)\,(N-2)}\, np\,(1-p)\,(1-2p);$$

$$\mathsf{M}\,(\xi - \mathsf{M}\xi)^4 = \frac{np\,(1-p)\,(N-n)}{(N-1)\,(N-2)\,(N-3)}\,\{N\,(N+1) + 6n\,(N-n) + 3p\,(1-p)\,[n\,(N-n)\,(N+6) - 2N^2]\}.$$

Коэффициент асимметрии: $\gamma_1 = \dfrac{(1-2p)\,(N-2n)\,\sqrt{N-1}}{\sqrt{np\,(1-p)\,(N-n)\,(N-2)}}$;

коэффициент эксцесса:

$$\gamma_2 = \frac{N^2\,(N-1)}{n\,(N-2)\,(N-3)\,(N-n)} \times \left[\frac{6n - 6N + 1}{Np\,(1-p)} + \frac{3n\,(N-n)\,(10N-12)}{N^2\,(-N-1)} - 6\right].$$

3. Типичная схема, в которой появляется гипергеометрическое распределение: проверяется партия готовой продукции, которая содержит Np годных и $N(1-p)$ негодных изделий. Случайным образом выбирают n изделий. Число годных изделий среди выбранных описывается гипергеометрическим распределением.

4. Если n мало́ по сравнению с N (практически, когда $n < 0{,}1N$), то

$$\frac{C_{Np}^{k}\, C_{N(1-p)}^{n-k}}{C_N^n} \approx C_n^k p^k (1-p)^{n-k}.$$

6.2.7. Распределение Пойа. 1. Случайная величина ξ имеет распределение Пойа с параметрами $(N,\ p,\ n,\ s)$, если

$$\mathsf{P}\{\xi = k\} = C_n^k\, \frac{b\,(b+s)\,\ldots\,[b+(k-1)\,s]\; c\,(c+s)\,\ldots\,[c+(n-k-1)\,s]}{N\,(N+s)\,\ldots\,[N+(n-1)\,s]},$$

где $b = Np$, $c = N(1-p)$.

2. Моменты:

$$\mathsf{M}\xi = np, \quad \mathsf{M}\xi^2 = np\,\frac{n + p + 1 + s/N}{1 + s/N};$$

$$\mathsf{D}\xi = np\,(1-p)\,\frac{1 + ns/N}{1 + s/N}.$$

3. Распределение Пойа появляется как модель следующего эксперимента: имеется урна, в которой находится Np белых и $N(1-p)$ черных шаров. Наугад вынимают шар и после определения цвета его возвращают в урну вместе с s новыми шарами того же цвета. Тогда ξ означает число выниманий белого шара в серии из n выниманий. Этот эксперимент широко используется при моделировании эпидемий заразных заболеваний.

4. Если n мало по сравнению с N, то

$$\mathsf{P}\{\xi=k\}\approx C_n^k p^k(1-p)^{n-k}.$$

6.2.8. Распределение Пуассона. 1. Случайная величина ξ имеет распределение Пуассона с параметром λ ($\lambda>0$), если

$$\mathsf{P}\{\xi=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,\,1,\,2,\,\ldots$$

2. Характеристическая функция:

$$\varphi(t)=e^{\lambda\left(e^{it}-1\right)}.$$

Моменты:

$$\mathsf{M}\xi=\lambda,\ \mathsf{M}\xi^2=\lambda^2+\lambda;\ \mathsf{D}\xi=\lambda.$$

Центральные моменты $\mu_k=\mathsf{M}(\xi-\mathsf{M}\xi)^k$ могут быть найдены из соотношений:

$$\mu_k=\lambda\sum_{l=0}^{k-2}C_{k-1}^l\mu_l\ \text{ либо }\ \mu_{k+1}=k\lambda\mu_{k-1}+\lambda\frac{d\mu_k}{d\lambda},$$

начальные моменты: для $k\geqslant 1$ $\mathsf{M}\xi^k=\lambda\sum\limits_{l=0}^{k-1}C_{k-1}^l\mathsf{M}\xi^l.$

Коэффициент асимметрии: $\gamma_1=1/\sqrt{\lambda}$; коэффициент эксцесса: $\gamma_2=1/\lambda$.

3. Распределение Пуассона является приемлемой моделью для описания случайного числа появления определенных событий в фиксированном промежутке времени, в фиксированной области пространства.

4. Если $np\to\lambda$, то $C_n^k p^k(1-p)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}$. Для больших λ имеет место приближение

$$\mathsf{P}\{\xi\leqslant k\}\approx\Phi\left(\frac{k+0{,}5-\lambda}{\sqrt{\lambda}}\right),$$

где $\Phi(x)$ — нормальная (0, 1) функция распределения.

Если $F_m(x)$ — функция распределения χ^2-распределения с m степенями свободы, то

$$\mathsf{P}\{\xi\leqslant k\}=1-F_{2(k+1)}(2\lambda).$$

Если ξ имеет распределение Пуассона с параметром λ, то для больших λ случайная величина $\sqrt{\xi}$ имеет распределение, близкое к нормальному с параметрами $(\sqrt{\lambda},\,1/4)$.

5. Распределение Пуассона безгранично делимо. Если сумма независимых случайных величин распределена по закону Пуассона, то и каждое слагаемое распределено по этому закону.

Пусть $\xi_1,\ \xi_2,\ \ldots$ — последовательность независимых одинаково распределенных случайных величин и ν — случайная величина, имеющая распределение Пуассона с параметром λ. Распределение случайной величины $\xi = \sum_{k=1}^{\nu} \xi_k$ называется *сложным пуассоновским распределением.*

Характеристическая функция случайной величины ξ:

$$\varphi(t) = e^{-\lambda + \lambda\psi(t)},$$

где $\psi(t)$ — характеристическая функция случайной величины ξ_k.

6.2.9. Обобщенное биномиальное распределение. 1. Случайная величина ξ имеет обобщенное биномиальное распределение с параметрами $(n, p_1, p_2, \ldots, p_n)$ $(0 < p_i < 1,\ i = 1, \ldots, n)$, если

$$\mathsf{P}\{\xi = k\} = \begin{cases} \prod_{l=1}^{n} (1 - p_l), & k = 0, \\ \sum_{l_k=1}^{n} \cdots \sum_{l_1=1}^{n} \prod_{\substack{i \neq l_1 \\ \cdots \\ i \neq l_k}} (1 - p_i) \prod_{j=1}^{k} p_{l_j}, & k = 1, \ldots, n-1, \\ \prod_{l=1}^{n} p_l, & k = n. \end{cases}$$

2. Характеристическая функция:

$$\varphi(t) = \prod_{k=1}^{n} [1 + p_k (e^{it} - 1)].$$

Моменты:

$$\mathsf{M}\xi = \sum_{k=1}^{n} p_k, \quad \mathsf{M}\xi^2 = \sum_{k=1}^{n} p_k + \sum_{k \neq l} p_k p_l; \quad \mathsf{D}\xi = \sum_{k=1}^{n} p_k (1 - p_k).$$

3. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины, имеющие распределение Бернулли с параметрами $p_1, p_2, \ldots, p_n$ соответственно. Тогда случайная величина $\xi = \sum_{k=1}^{n} \xi_k$ имеет обобщенное биномиальное распределение.

Если $\bar{p} = \frac{1}{n} \sum_{k=1}^{n} p_k$ и $\sigma_{\bar{p}} = \frac{1}{n} \sum_{k=1}^{n} p_k^2 - \bar{p}^2$, то $\mathsf{D}\xi = n\bar{p}(1 - \bar{p}) - n\sigma_{\bar{p}}\bar{p}$. Поэтому дисперсия биномиального распределения с параметрами $(n, \bar{p})$ на n дисперсий случайной величины ζ с распределением $\mathsf{P}\{\zeta = p_k\} = 1/n$ больше, чем распределение случайной величины ξ с обобщенным биномиальным распределением.

6.2.10. Логарифмическое распределение. 1. Случайная величина ξ имеет логарифмическое распределение с параметром p $(0 < p < 1)$, если

$$\mathsf{P}\{\xi = k\} = -\frac{(1-p)^k}{k \ln p}, \quad k = 1, 2, \ldots$$

2. Характеристическая функция:

$$\varphi(t)=\frac{1}{\ln p}\ln\left[1-(1-p)e^{it}\right]=1-\frac{1}{\ln p}\ln\left[1-\frac{(1-p)t}{p\cdot 1!}-\frac{(1-p)t^2}{p\cdot 2!}-\ldots\right].$$

Моменты:

$$\mathsf{M}\xi=\frac{1-p}{p\ln p},\quad \mathsf{M}\xi^2=-\frac{1-p}{p^2\ln p},\quad \mathsf{M}\xi^3=\frac{(1-p)(2-p)}{p^3\ln p};$$

$$\mathsf{D}\xi=-\frac{1-p}{p^2\ln p}\left[1+\frac{1-p}{\ln p}\right].$$

3. Логарифмическое распределение является предельным для отрицательного биномиального распределения в следующем смысле. Если η_r — случайная величина, имеющая отрицательное биномиальное распределение с параметром (r, p), то

$$\lim_{r\to 0}\mathsf{P}\{\eta_r=k\,|\,\eta_r>0\}=-\frac{(1-p)^k}{k\ln p}.$$

4. Логарифмическим называют также распределение случайной величины ξ, у которой

$$\mathsf{P}\{\xi=k\}=\log_m(k+1)-\log_m k,\qquad k=1,\ldots,m-1.$$

6.2.11. Распределение Бореля — Таннера. 1. Случайная величина ξ имеет распределение Бореля — Таннера с параметрами (r, α) $(0<\alpha<1)$, если

$$\mathsf{P}\{\xi=k\}=\frac{r}{(k-r)!}k^{k-r-1}e^{-\alpha k}\alpha^{k-r},\qquad k=r, r+1, \ldots$$

2. Моменты: $\mathsf{M}\xi=r/(r-\alpha)$; $\mathsf{D}\xi=\alpha r/(1-\alpha)^3$.

3. В теории массового обслуживания распределение Бореля — Таннера определяется как распределение числа требований, обслуженных на периоде занятости в системе массового обслуживания с пуассоновским входящим потоком с параметром α и постоянным временем обслуживания в том случае, когда в начальный момент длина очереди равна r.

6.3. Непрерывные распределения

6.3.1. Равномерное распределение. 1. Случайная величина ξ имеет равномерное распределение на отрезке $[a, b]$ $(a<b)$, если

$$f(x)=\begin{cases}1/(b-a), & x\in[a, b],\\ 0, & x\notin[a, b].\end{cases}$$

Здесь и далее через $f(x)$ обозначена плотность вероятности соответствующей случайной величины (рис. 1).

2. Характеристическая функция:

$$\varphi(t)=(e^{itb}-e^{ita})/(b-a)\,it.$$

Моменты:

$$\mathrm{M}\xi^k = \frac{b^{k+1} - a^{k+1}}{(b-a)(k+1)} \quad (k = 1, 2, \ldots); \quad \mathrm{D}\xi = \frac{(b-a)^2}{12}.$$

Коэффициент асимметрии: $\gamma_1 = 0$; коэффициент эксцесса: $\gamma_2 = 9/5$.

3. С помощью линейного преобразования $\eta = (\xi - a)/(b - a)$ приводится к равномерному распределению на отрезке [0, 1]. Равномерное распределение является непрерывным аналогом распределений классической теории вероятностей, описывающих случайные эксперименты с равновероятными исходами.

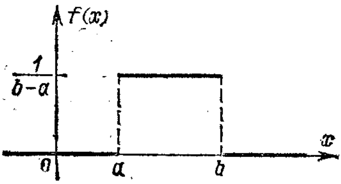

Рис. 1. Плотность равномерного распределения

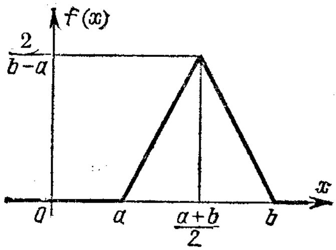

Рис. 2. Плотность треугольного распределения

4. Погрешность, происходящая от округления числа, удовлетворительно описывается равномерным распределением на отрезке [—1/2, 1/2].

Если случайная величина ζ имеет непрерывную функцию распределения $F_\zeta(x)$, то случайная величина $\xi = F_\zeta(\zeta)$ имеет равномерное распределение на отрезке [0, 1]. Этим объясняется широкое использование равномерного распределения в статистическом моделировании (методы Монте-Карло).

6.3.2. Треугольное распределение (распределение Симпсона). 1. Случайная величина ξ имеет треугольное распределение (распределение Симпсона) на отрезке $[a, b]$, если

$$f(x) = \begin{cases} \dfrac{2}{b-a} - \dfrac{2}{(b-a)^2}\,|a + b - 2x|, & x \in [a, b], \\ 0, \quad x \notin [a, b] \end{cases}$$

(рис. 2).

2. Характеристическая функция:

$$\varphi(t) = \left[\frac{2\left(e^{itb/2} - e^{ita/2}\right)}{(b-a)\,it}\right]^2.$$

Моменты:

$$\mathrm{M}\xi^k = \frac{4}{(b-a)^2(k+1)(k+2)}\left[a^{k+2} + b^{k+2} - 2\left(\frac{a+b}{2}\right)^{k+2}\right];$$

$$\mathrm{D}\xi = \frac{(b-a)^2}{24}.$$

Коэффициент асимметрии: $\gamma_1 = 0$; коэффициент эксцесса: $\gamma_2 = -3/5$.

3. Если ξ_1 и ξ_2 — независимые случайные величины, равномерно распределенные на отрезке $[a/2, b/2]$, то случайная величина $\xi = \xi_1 + \xi_2$ имеет треугольное распределение.

6.3.3. Показательное (экспоненциальное) распределение. 1. Случайная величина ξ имеет показательное (экспоненциальное) распределение с параметром $\lambda > 0$, если

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0 \end{cases}$$

(рис. 3).

2. Характеристическая функция:

$$\varphi(t) = \lambda/(\lambda - it).$$

Моменты:

$$\mathsf{M}\xi^k = k!/\lambda^k; \quad \mathsf{D}\xi = 1/\lambda^2.$$

Медиана: $m = \ln 2/\lambda$; коэффициент асимметрии: $\gamma_1 = 2$; коэффициент эксцесса: $\gamma_2 = 6$.

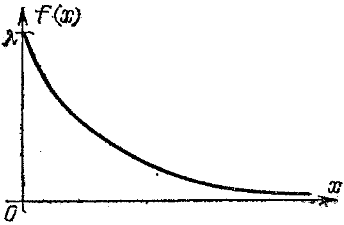

Рис. 3. Плотность показательного распределения

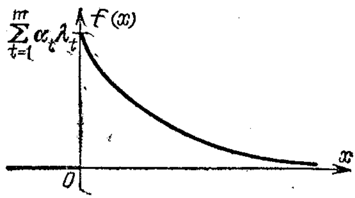

Рис. 4. Плотность гиперэкспоненциального распределения

3. Непрерывный аналог геометрического распределения. Обладает свойством *отсутствия последействия:*

$$\mathsf{P}\{\xi > t + s \mid \xi > s\} = \mathsf{P}\{\xi > t\},$$

в связи с чем является основным в теории скачкообразных марковских процессов.

6.3.4. Гиперэкспоненциальное распределение. 1. Случайная величина ξ имеет гиперэкспоненциальное распределение с параметрами $(m;\ \alpha_1, \alpha_2, \ldots, \alpha_m;\ \lambda_1, \lambda_2, \ldots, \lambda_m)$, $\alpha_k\lambda_k \geqslant 0$, $\sum_{k=1}^{m} \alpha_k = 1$ (смесь экспоненциальных распределений), если

$$f(x) = \begin{cases} \sum_{k=1}^{m} \alpha_k \lambda_k e^{-\lambda_k x}, & x \geqslant 0, \\ 0, & x < 0 \end{cases}$$

(рис. 4).

2. Характеристическая функция:

$$\varphi(t)=\sum_{k=1}^{m}\frac{\alpha_k\lambda_k}{\lambda_k-it}.$$

Моменты:

$$\mathsf{M}\xi^k=\sum_{l=1}^{m}\alpha_l\frac{k!}{\lambda_l^k};\quad \mathsf{D}\xi=2\sum_{l=1}^{m}\frac{\alpha_l}{\lambda_l^2}-\left(\sum_{l=1}^{m}\frac{\alpha_l}{\lambda_l}\right)^2.$$

3. Введем случайные величины I_k ($k=1,\ldots,m$), принимающие значения 0 или 1, причем $\mathsf{P}\{I_k=1\}=\alpha_k$, $\mathsf{P}\left\{\sum_{k=1}^{m}I_k=1\right\}=1$. Если ξ_k ($k=1,\ldots,m$) — независимые случайные величины, имеющие экспоненциальное распределение с параметром λ_k соответственно и не зависящие от I_k, то случайная величина $\xi=\sum_{k=1}^{m}I_k\xi_k$ имеет гиперэкспоненциальное распределение с параметрами (m; $\alpha_1,\ldots,\alpha_m$; $\lambda_1,\ldots,\lambda_m$).

6.3.5. Нормальное распределение. 1. Случайная величина ξ имеет нормальное распределение с параметрами (m, σ^2), если

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-(x-m)^2/2\sigma^2},\quad x\in(-\infty,\infty)$$

(рис. 5).

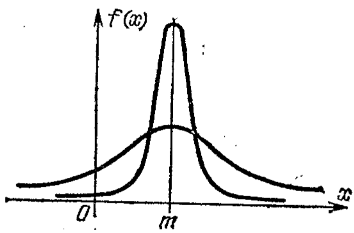

Рис. 5. Плотность нормального распределения

Наряду с приведенным выше представлением плотности нормального распределения используется следующее:

$$f(x)=\frac{\rho}{\sqrt{\pi}E}\exp\left\{\frac{\rho^2(x-m)^2}{E^2}\right\},$$

где $\rho=0{,}4769\ldots$ является решением уравнения

$$\int_0^{\rho}e^{-x^2/2}dx=\sqrt{\pi},$$

а $E=\rho\sqrt{2}\,\sigma$ определяется из соотношения

$$\mathsf{P}\{|\xi-m|\leqslant E\}=\mathsf{P}\{|\xi-m|>E\}=0{,}5$$

и называется *срединным* (или *вероятностным*) *отклонением*.

2. Характеристическая функция:

$$\varphi(t)=\exp\left\{imt-\frac{t^2\sigma^2}{2}\right\}.$$

Моменты:

$$\mu_{2k+1} = 0, \quad \mu_{2k} = 1 \cdot 3 \cdot \ldots \cdot (2k-1)\,\sigma^{2k},$$

где $\mu_k = \mathsf{M}(\xi - \mathsf{M}\xi)^k$, $\mathsf{M}\xi = m$; $\mathsf{D}\xi = \sigma^2$ для $k \geqslant 2$, $\mathsf{M}\xi^k = m\mathsf{M}\xi^{k-1} + (k-1)\,\sigma^2\mathsf{M}\xi^{k-2}$.

Коэффициенты асимметрии и эксцесса равны нулю, мода и медиана совпадают с m.

3. Фундаментальная роль, которую играет нормальное распределение, объясняется тем, что при широких предположениях суммы случайных величин с ростом числа слагаемых ведут себя асимптотически нормально. Соответствующие условия составляют содержание *центральной предельной теоремы.*

С помощью линейного преобразования $\eta = (\xi - m)/\sigma$ нормальное распределение с произвольными параметрами (m, σ) приводится к стандартному нормальному закону с параметрами $(0, 1)$ и с функцией распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\,dz.$$

Табулируется, как правило, функция $\Phi_0(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{0}^{x} e^{-z^2/2}\,dz$, связанная с $\Phi(x)$ соотношением $\Phi(x) = 1/2 + \Phi_0(x)$.

Для вычисления $\Phi(x)$ при малых x можно использовать

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left(x - \frac{x^3}{2 \cdot 3} + \frac{x^5}{2!\,2^2 \cdot 5} + \ldots (-1)^n \frac{x^{2n+1}}{n!\,(2n+1)\,2^n} + \ldots \right)$$

либо *отношение Миллса* $R(x) = \sqrt{2\pi}\,\dfrac{1 - \Phi(x)}{e^{-x^2/2}} = \displaystyle\int_{x}^{\infty} e^{(x^2 - y^2)/2}\,dy$,

которое раскладывается в непрерывную дробь:

$$R(x) = \frac{1}{x+} \cdot \frac{1}{x+} \cdot \frac{2}{x+} \cdot \frac{3}{x+} \cdot \ldots \cdot \frac{n}{x+} \cdot \ldots$$

Нормально распределенная случайная величина с большой вероятностью принимает значения, близкие к своему математическому ожиданию, что выражается *правилом сигм*:

$$\mathsf{P}\{|\xi - m| \geqslant k\sigma\} = \begin{cases} 0{,}3173\ \ldots, & k = 1, \\ 0{,}0455\ \ldots, & k = 2, \\ 0{,}0027\ \ldots, & k = 3. \end{cases}$$

Чаще всего используется *правило трех сигм.*

4. Нормальное распределение безгранично делимо. Если сумма двух независимых случайных величин имеет нормальное распределение, то и каждая случайная величина имеет нормальное распределение.

Нормальное распределение может встретиться под названиями: второй закон Лапласа, лапласовское распределение, гауссовское распределение, распределение Лапласа — Гаусса, распределение Гаусса — Лапласа.

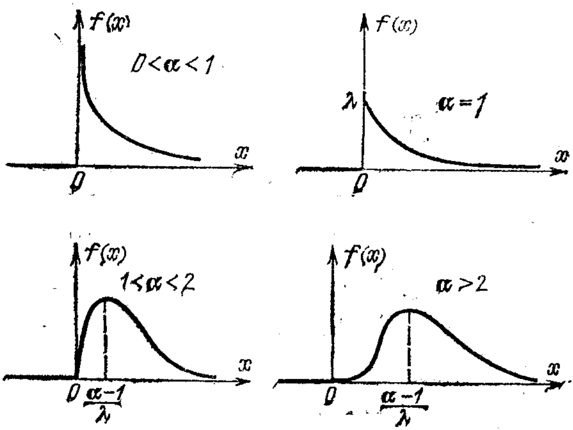

Рис. 6. Плотность гамма-распределения

6.3.6. Гамма-распределение. 1. Случайная величина ξ имеет гамма-распределение с параметрами (α, λ) $(\alpha > 0,\ \lambda > 0)$, если

$$f(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

(рис. 6).

2. Характеристическая функция: $\varphi(t) = (1 - it/\lambda)^{-\alpha}$.

Моменты:

$$\mathsf{M}\xi^k = \frac{\alpha(\alpha+1)\ \ldots\ (\alpha+k-1)}{\lambda^k}; \quad \mathsf{D}\xi = \frac{\alpha}{\lambda^2}.$$

Мода: $\eta = (\alpha - 1)/\lambda$, $\alpha \geqslant 1$; коэффициент асимметрии: $\gamma_1 = 2/\sqrt{\alpha}$; коэффициент эксцесса: $\gamma_2 = 6/\alpha$.

3. Гамма-распределение является непрерывным аналогом отрицательного биномиального распределения. При $\alpha = 1$ гамма-распределение совпадает с показательным, а при $\alpha = n/2$, $\lambda = 1/2$ — с χ^2-распределением с n степенями свободы. При $\lambda = n\mu$ и $\alpha = n$ гамма-распределение называется *эрланговским распределением* с параметрами (n, μ) и описывает распределение длительности интервала времени до появления n событий процесса Пуассона с пара-

метром μ, используемым в теории массового обслуживания и теории надежности.

При $n \to \infty$ эрланговское распределение стремится к вырожденному.

При $\alpha = m + 1$ и $\lambda = 1$ гамма-распределение называют *показательно-степенным* распределением с параметром m, функция распределения которого имеет вид

$$F(x) = 1 - e^{-x} \sum_{l=0}^{m} \frac{x^l}{l!}.$$

При фиксированном λ гамма-распределение безгранично делимо.

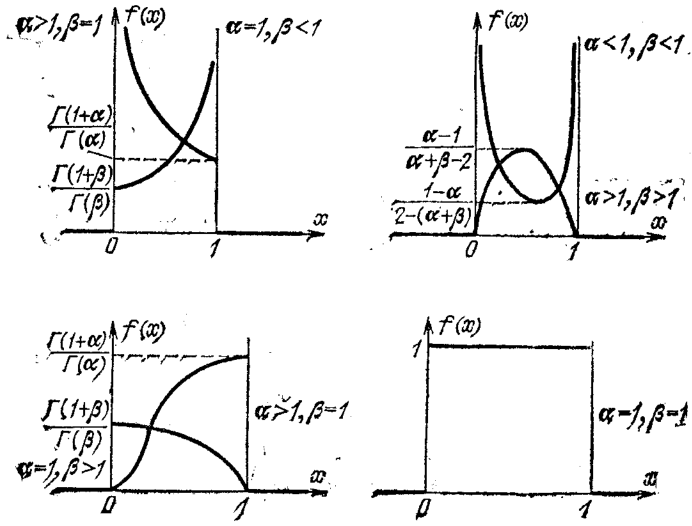

Рис. 7. Плотность бета-распределения

6.3.7. Бета-распределение. 1. Случайная величина ξ имеет бета-распределение с параметрами (α, β) $(\alpha > 0, \beta > 0)$, если

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, & x \in [0, 1]; \\ 0, & x \notin [0, 1] \end{cases}$$

(рис. 7).

2. Характеристическая функция:

$$\varphi(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{(it)^k \Gamma(\alpha+k)}{k!\Gamma(\alpha+\beta+k)}.$$

Моменты:

$$\mathsf{M}\xi^k = \frac{\alpha(\alpha+1)\ldots(\alpha+k-1)}{(\alpha+\beta)(\alpha+\beta+1)\ldots(\alpha+\beta+k-1)} = \frac{\Gamma(\alpha+k)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\alpha+\beta+k)};$$

$$\mathsf{D}\xi = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Коэффициент асимметрии: $\gamma_1 = \dfrac{2(\beta-\alpha)\sqrt{1+\alpha+\beta}}{(2+\alpha+\beta)\sqrt{\alpha\beta}}$; коэффициент эксцесса: $\gamma_2 = \dfrac{6[(\alpha-\beta)^2(\alpha+\beta+1)-\alpha\beta(\alpha+\beta+2)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}$.

3. Бета-распределение появляется, например, как распределение порядковых статистик.

Если $\xi_1, \xi_2, \ldots, \xi_n$ независимы и равномерно распределены на отрезке $[0, 1]$ и $\xi_{(1)}, \xi_{(2)}, \ldots, \xi_{(n)}$ — упорядоченные по возрастанию величины ξ_k $(k = 1, \ldots, n)$ ($\xi_{(k)}$ называется *k-й порядковой статистикой*), то плотность вероятности $f_{(k)}(x)$ k-й порядковой статистики $\xi_{(k)}$ имеет бета-распределение с $\alpha = k$, $\beta = n - k + 1$.

При $\alpha > 1$, $\beta > 1$ бета-распределение унимодально с модой в точке $x = (\alpha-1)/(\alpha+\beta-2)$. Если $\alpha = \beta = 1$, то бета-распределение совпадает с равномерным на отрезке $[0, 1]$ распределением. При $\beta = \alpha + 1$ бета-распределение называется *обобщенным распределением арксинуса*, при $\alpha = \beta = 1/2$ — *распределением арксинуса*.

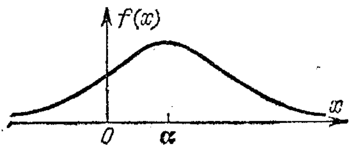

Рис. 8. Плотность распределения Коши

6.3.8. Распределение Коши. 1. Случайная величина ξ имеет распределение Коши с параметрами (α, λ) $(\lambda > 0)$, если

$$f(x) = \frac{1}{\pi} \cdot \frac{\lambda}{\lambda^2 + (x-\alpha)}$$

(рис. 8).

2. Характеристическая функция:

$$\varphi(t) = \exp\{i\alpha t - \lambda|t|\}.$$

Математическое ожидание случайной величины, имеющей распределение Коши, не существует, хотя интеграл в смысле главного значения равен

$$\lim_{A\to\infty} \frac{\lambda}{\pi} \int_{-A}^{A} \frac{x\,dx}{\lambda^2 + (x-\alpha)^2} = \frac{\alpha}{\lambda}.$$

Второй момент бесконечен.

3. Параметр α является модой и медианой. Распределение Коши безгранично делимо.

6.3.9. Распределение Лапласа (двойное экспоненциальное распределение). 1. Случайная величина ξ имеет распределение Лапласа (двойное экспоненциальное распределение) с параметрами (α, λ) $(\lambda > 0)$, если

$$f(x) = \frac{\lambda}{2} e^{-\lambda |x-\alpha|}, \quad x \in (-\infty, \infty)$$

(рис. 9).

2. Характеристическая функция:

$$\varphi(t) = \lambda^2 e^{it\alpha}/(t^2 + \lambda^2).$$

Моменты:

$$\mathsf{M}\xi^{2k+1} = \alpha^{2k+1}(2k+1)!,$$

$$\mathsf{M}\xi^{2k} = \left[\frac{\alpha^{2k}}{(2k)!} + \frac{\alpha^{2(k-1)}}{2(k-1)!\,\lambda^2} + \ldots + \frac{1}{\lambda^{2k}}\right](2k)!; \quad \mathsf{D}\xi = \frac{2}{\lambda^2}.$$

Мода и медиана совпадают с α; коэффициент асимметрии: $\gamma_1 = 0$; коэффициент эксцесса: $\gamma_2 = 6\lambda - 3$.

3. Пусть ξ_1, ξ_2 — независимые случайные величины, имеющие экспоненциальное распределение с параметром λ. Случайная величина $\xi = \xi_1 - \xi_2 + \alpha$ имеет распределение Лапласа с параметрами (α, λ). Появляется в качестве предельного распределения в схемах суммирования случайного числа случайных слагаемых.

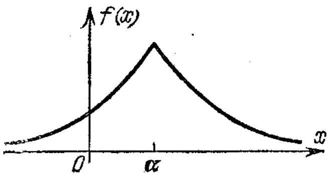

Рис. 9. Плотность двойного экспоненциального распределения

6.3.10. χ^2-распределение. 1. Случайная величина ξ имеет χ^2-распределение с α степенями свободы, если

$$f(x) = \begin{cases} \dfrac{1}{2^{\alpha/2}\Gamma(\alpha/2)} x^{\alpha/2-1} e^{-x/2}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

(рис. 10).

2. Характеристическая функция:

$$\varphi(t) = (1 - 2it)^{-\alpha/2}.$$

Моменты:

$$\mathsf{M}\xi^k = \alpha(\alpha+2)\ldots[\alpha + 2(k-1)]; \quad \mathsf{D}\xi = 2\alpha,$$

$$\mu_3 = 8\alpha, \quad \mu_4 = 48\alpha + 12\alpha^2, \ldots$$

Коэффициент асимметрии: $\gamma_1 = \sqrt{8/\alpha}$; коэффициент эксцесса: $\gamma_2 = 12/\alpha$.

3. Многочисленные применения χ^2-распределения в теории вероятностей и математической статистике основаны на следующей его интерпретации.

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины, имеющие стандартное нормальное распределение. Случайная величина $\xi = \sum_{k=1}^{n} \xi_k^2$ имеет χ^2-распределение с n степенями свободы.

Если $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$ — n-мерный нормальный вектор с математическим ожиданием $m = (m_1, m_2, \ldots, m_n)$ и невырожденной ковариационной матрицей $C = \{c_{ij},\ i, j = 1, \ldots, n\}$, то случайная величина

$$\xi = (\eta - m)^T C^{-1} (\eta - m) = \sum_{i,\ j=1}^{n} c_{ij}^{-1} (\eta_i - m_i)(\eta_j - m_j),$$

где ^{T} означает операцию транспонирования и c_{ij}^{-1} — элементы матрицы C^{-1}, имеет χ^2-распределение с n степенями свободы.

Если ξ — случайная величина, имеющая χ^2-распределение с n степенями свободы, то случайная величина $\sqrt{2\xi} - \sqrt{2n-1}$ имеет приближенно стандартное нормальное распределение.

4. Один из критериев проверки согласия эмпирических данных с гипотетической функцией распределения $F(x)$ основан на изучении χ^2-статистики Пирсона

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i - np_i}{np_i},$$

где $p_i = F(x_i) - F(x_{i-1})$, $x_0 = -\infty < x_1 < \ldots < x_k = \infty$, — произвольное разбиение интервала $(-\infty, \infty)$, n_i — число наблюдений, попавших в полуинтервал $[x_{i-1}, x_i)$. В предположении истинности гипотезы χ^2-статистика Пирсона имеет в пределе при $n = \sum_i n_i \to \infty$ χ^2-распределение с $(k-1)$-степенями свободы и не зависит от $F(x)$.

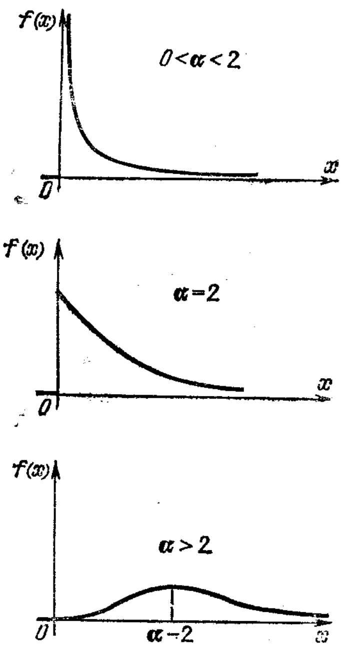

Рис. 10. Плотность χ^2-распределения

5. Если $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины, имеющие нормальные распределения с параметрами (m_1, σ^2), (m_2, σ^2), ..., (m_n, σ^2) соответственно, то распределение случайной величины $\xi = \frac{1}{\sigma^2} \sum_{i=1}^{n} \xi_i^2$ называется *нецентральным χ^2-распределением* с n степенями свободы и параметром нецентральности $m = \frac{1}{\sigma^2} \sum_{i=1}^{n} m_i^2$.

Плотность распределения вероятностей:

$$f(x) = \frac{e^{(x+m)/2}}{2^{n/2}} x^{n/2-1} \sum_{j=0}^{\infty} \frac{(mx/4)^j}{j!\,\Gamma(j+n/2)}, \quad x \geqslant 0.$$

Характеристическая функция:

$$\varphi(t)=\exp\left\{\frac{imt}{1-2it}\right\}(1-2it)^{-n/2}.$$

Моменты:

$$\mathsf{M}\xi=m+n;\quad \mathsf{D}\xi=4m+2n.$$

Для больших значений параметра m нецентральное χ^2-распределение с n степенями свободы приближенно совпадает с нормальным $\left(n+m,\ \sqrt{2(n+2m)}\right)$ распределением.

χ^2-статистика Пирсона в предположении, что теоретические вероятности равны $\left(p_1', p_2', \ldots, p_k'\right)$, имеет асимптотически нецентральное χ^2-распределение с $k-1$ степенями свободы и параметром нецентральности

$$m=\sum_{i=1}^{k}\frac{\left(p_i'-p_i\right)^2}{p_i}.$$

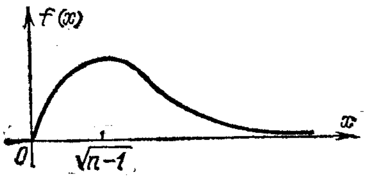

Рис. 11. Плотность χ-распределения

6.3.11. χ-распределение. 1. Случайная величина ξ имеет χ-распределение с α степенями свободы ($\alpha>0$), если

$$f(x)=\begin{cases}\dfrac{1}{2^{\alpha/2-1}\Gamma(\alpha/2)}x^{\alpha-1}e^{-x^2/2}, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

(рис. 11).

2. Характеристическая функция:

$$\varphi(t)=\frac{1}{\Gamma(\alpha/2)}\sum_{k=0}^{\infty}\frac{(i\sqrt{2}t)^k}{k!}\Gamma\left(\frac{\alpha+k}{2}\right).$$

Моменты:

$$\mathsf{M}\xi^k=\frac{2^{k/2}\Gamma((\alpha+k)/2)}{\Gamma(\alpha/2)};\quad \mathsf{D}\xi=\alpha+\left(2-\sqrt{2}\right)\left[\frac{\Gamma((\alpha+1)/2)}{\Gamma(\alpha/2)}\right]^2.$$

3. Если $\xi_1, \xi_2, \ldots, \xi_n$ — независимые случайные величины, имеющие стандартное нормальное распределение, то $\xi=\sqrt{\sum_{i=1}^{n}\xi_i^2}$ имеет χ-распределение с n степенями свободы.

При $n=2$ χ-распределение называется иногда *распределением Рэлея* (*Рэлея — Райса*).

При $n=3$ χ-распределение называется *распределением Максвелла* и описывает распределение скоростей молекул разреженного газа.

6.3.12. Распределение Стьюдента (t-распределение). 1. Случайная величина ξ имеет распределение Стьюдента (t-распределение) с α степенями свободы ($\alpha > 0$), если

$$f(x) = \frac{\Gamma((\alpha+1)/2)}{\sqrt{\alpha\pi}\,\Gamma(\alpha/2)}\left(1+\frac{x^2}{\alpha}\right)^{-(\alpha+1)/2}, \quad x \in (-\infty, \infty)$$

(рис. 12).

2. Характеристическая функция:

$$\varphi(t) = \frac{\sqrt{\pi}\,\Gamma((\alpha+1)/2)}{\Gamma(\alpha/2)} \times$$

$$\times \frac{\exp\{-\sqrt{\alpha}\,|t|\}}{2^{2(n-1)}(n-1)!}\sum_{k=0}^{n-1}(2k)!\,C_{n-1+k}^{2k}(2\sqrt{\alpha}\,|t|)^{n-1-k},$$

если $n = (\alpha+1)/2$ — целое число.

Моменты:

$$\mathsf{M}\xi^{2k-1} = 0, \quad \mathsf{M}\xi^{2k} = \frac{\alpha^k\Gamma(\alpha/2-k)\,\Gamma(k+1/2)}{\sqrt{\pi}\,\Gamma(\alpha/2)}, \quad 2k < \alpha;$$

$$\mathsf{D}\xi = \begin{cases} \alpha/(\alpha-2), & \text{если } \alpha > 2, \\ \infty, & \text{если } \alpha \leqslant 2. \end{cases}$$

3. Если η и ζ — независимые случайные величины, причем η имеет стандартное нормальное распределение, а ζ — χ^2-распределение с n степенями свободы, то $\xi = \eta\sqrt{n/\zeta}$ имеет распределение Стьюдента с n степенями свободы.

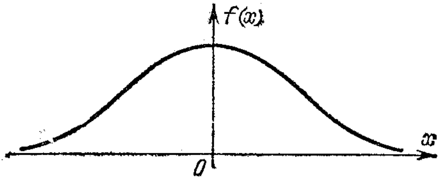

Рис. 12. Плотность распределения Стьюдента

Во многих статистических приложениях параметр α — натуральное число. При $\alpha = 1$ распределение Стьюдента совпадает с распределением Коши. Распределение Стьюдента появляется при проверке гипотезы о среднем нормально распределенной генеральной совокупности при неизвестной дисперсии.

При больших значениях α распределение Стьюдента асимптотически сближается с стандартным нормальным распределением. t-распределение безгранично делимо.

6.3.13. F-распределение (распределение Фишера — Снедекора). 1. Случайная величина ξ имеет F-распределение (распределение

Фишера — Снедекора) с (m_1, m_2) степенями свободы, если

$$f(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{m_1+m_2}{2}\right)m_1^{m_1/2}m_2^{m_2/2}}{\Gamma\left(\dfrac{m_1}{2}\right)\Gamma\left(\dfrac{m_2}{2}\right)}x^{m_1/2-1}(m_2+m_1x)^{-(m_1+m_2)/2}, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

(рис. 13).

2. Моменты:

$$\mathsf{M}\xi^k=\frac{\Gamma(m_1/2+k)\,\Gamma(m_2/2-k)\,m_2^k}{\Gamma(m_1/2)\,m_1^k\,\Gamma(m_2/2)},$$

если

$$2k<m_2,\quad \mathsf{M}\xi=\frac{m_2}{m_2-2}\quad (m_2>2);$$

$$\mathsf{D}\xi=\frac{2m_2^2(m_1+m_2-2)}{m_1(m_2-2)^2(m_2-4)}\quad (m_2>4).$$

Мода: $\dfrac{m_2(m_1-2)}{m_1(m_2+2)}$, $m_1>2$.

Коэффициент асимметрии: $\gamma_1=\dfrac{(2m_1+m_2-2)\sqrt{8(m_2-4)}}{(m_2-6)\sqrt{m_1+m_2-2}}$, $m_2>6$

3. Если ξ_1 и ξ_2 — независимые случайные величины, имеющие χ^2-распределение соответственно с m_1 и m_2 степенями свободы, то случайная величина $\xi=\dfrac{\xi_1/m_1}{\xi_2/m_2}$ имеет F-распределение с (m_1, m_2)

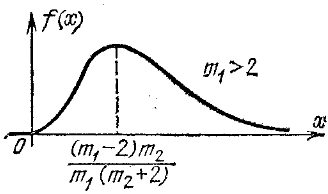

Рис. 13. Плотность F-распределения

степенями свободы, в связи с чем m_1 называют *числом степеней свободы числителя*, а m_2 —*числом степеней свободы знаменателя*. В частности, если $x_1, x_2, \ldots, x_{m_1}$ — выборка из нормальной (a_1, σ^2) генеральной совокупности, а $y_1, y_2, \ldots, y_{m_2}$ — выборка из нормальной (a_2, σ^2) генеральной совокупности, то статистика

$$\frac{1}{m_1-1}\sum_{i=1}^{m_1}(x_i-\bar{x})^2\Bigg/\frac{1}{m_2-1}\sum_{i=1}^{m_2}(y_i-\bar{y})^2,$$

где $\bar{x} = \frac{1}{m_1} \sum_{i=1}^{m_1} x_i$, $\bar{y} = \frac{1}{m_2} \sum_{i=1}^{m_2} y_i$, имеет F-распределение с $(m_1 - 1, m_2 - 1)$ степенями свободы.

4. Если ξ_1 и ξ_2 — независимые случайные величины, причем ξ_2 имеет χ^2-распределение с m_2 степенями свободы, а ξ_1 имеет нецентральное χ^2-распределение с m_1 степенями свободы и параметром нецентральности m, то $\xi = \frac{\xi_1/m_1}{\xi_2/m_2}$ имеет *нецентральное F-распределение* с (m_1, m_2) степенями свободы и параметром нецентральности m.

6.3.14. Логарифмически нормальное (логнормальное) распределение. 1. Случайная величина ξ имеет логарифмически нормальное (логнормальное) распределение с параметрами (m, σ^2), если

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln x - m)^2}{2\sigma^2}\right\}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

(рис. 14).

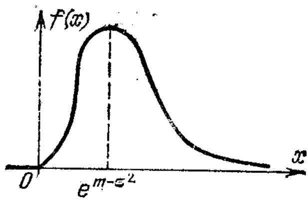

Рис. 14. Плотность логарифмически нормального распределения

2. Моменты:

$$\mathsf{M}\xi^k = \exp\left\{\frac{1}{2} k^2\sigma^2 + km\right\};$$

$$\mathsf{D}\xi = e^{\sigma^2 + 2m}\left[e^{\sigma^2} - 1\right].$$

Коэффициент асимметрии: $\gamma_1 = \left(e^{\sigma^2} + 2\right)\sqrt{e^{\sigma^2} - 1}$; коэффициент эксцесса: $\gamma_2 = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$.

3. Если η имеет нормальное $(0, 1)$ распределение, то $\xi = \exp\{\sigma\eta + m\}$ имеет логнормальное распределение с параметрами (m, σ^2).

Широко используется в статистической физике, статистической геологии, экономической статистике, биологии и т. д.

4. Логнормальное распределение можно получить как частный случай так называемого *распределения Кептейна*, имеющего плотность вероятности

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{[G(x) - m]^2}{2\sigma^2}\right\} \left|\frac{dG(x)}{dx}\right|,$$

где $G(x)$ — монотонная дифференцируемая функция. Если ξ имеет распределение Кептейна $(m, \sigma^2, G(x))$, то $\eta = G(\xi)$ имеет нормальное $(0, 1)$ распределение. Распределение Кептейна появляется как результат применения центральной предельной теоремы к схеме суммирования вида $x_{i+1} = x_i + z_{i+1}g(x_i)$, где z_i — независимые случайные величины, $G(x) = \int\limits_{x_0}^{x} \frac{du}{g(u)}$.

6.3.15. Логистическое распределение. 1. Случайная величина ξ имеет логистическое распределение с параметрами (m, σ^2), если

$$f(x)=\frac{\pi\exp\left\{-\frac{\pi}{\sqrt{3}}\left(\frac{x-m}{\sigma}\right)\right\}}{\sigma\sqrt{3}\left(1+\exp\left\{-\frac{\pi(x-m)}{\sigma\sqrt{3}}\right\}\right)^2},\quad x\in(-\infty,\infty)$$

(рис. 15).

2. Характеристическая функция:

$$\varphi(t)=e^{itm}\Gamma\left(1-i\frac{\sigma\sqrt{3}}{\pi}t\right)\Gamma\left(1+i\frac{\sigma\sqrt{3}}{\pi}t\right).$$

Моменты: $\mathrm{M}\xi=m$; $\mathrm{D}\xi=\sigma^2$.

Мода и медиана равны m; коэффициент асимметрии: $\gamma_1=0$; коэффициент эксцесса: $\gamma_2=1{,}2$.

3. Логистическая функция распределения по форме мало отличается от нормальной функции распределения и наряду с последней используется, например, в медико-биологических исследованиях для анализа эффекта различных лекарств, ядов и т. д.

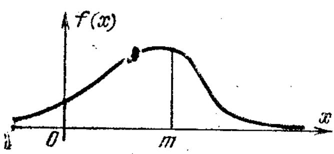

Рис. 15. Плотность логистического распределения

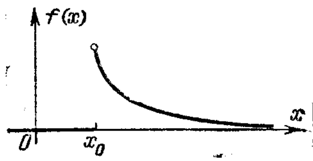

Рис. 16. Плотность распределения Парето

6.3.16. Распределение Парето. 1. Случайная величина ξ имеет распределение Парето с параметрами (x_0, α) $(\alpha>0,\ x_0>0)$, если

$$f(x)=\begin{cases}\frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1}, & x>x_0,\\ 0, & x\leqslant x_0\end{cases}$$

(рис. 16).

2. Моменты:

$$\mathrm{M}\xi^k=\frac{\alpha}{\alpha-k}x_0^k,\quad k<\alpha;\qquad \mathrm{D}\xi=\begin{cases}\frac{\alpha}{(\alpha-1)(\alpha-2)}x_0^2, & \alpha>2,\\ \infty, & \alpha\leqslant 2.\end{cases}$$

3. Распределение Парето есть усечение на интервале (x_0, ∞) степенного распределения с параметром α, имеющего плотность вероятности

$$f(x)=\begin{cases}\alpha x^{-(\alpha+1)}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

Встречается в задачах экономической статистики.

6.3.17. Распределение Шермана. 1. Случайная величина ξ имеет распределение Шермана с n степенями свободы (параметром n), если

$$f(x)=\begin{cases}\sum\limits_{m=1}^{n} mb_m x^{m-1}, & x\in\left[0,\ \frac{n}{n+1}\right],\\ 0, & x\notin\left[0,\ \frac{n}{n+1}\right],\end{cases}$$

где $b_m=\sum\limits_{j=0}^{r}(-1)^{(m+j+1)}C_{n+1}^{j+1}C_{m+j}^{j}\left[C_{n-j}^{n+1}\right]^{n-m}$, r — целое неотрицательное число, удовлетворяющее неравенствам

$$\frac{n-r-1}{n+1}\leqslant x<\frac{n-r}{n+1}$$

(рис. 17).

2. Моменты:

$$\mathsf{M}\xi=\left(1-\frac{1}{n+1}\right)^{n+1};$$

$$\mathsf{D}\xi=\frac{2n^{n+2}+n(n-1)^{n+2}}{(n+2)(n+1)^{n+2}}-\left(1-\frac{1}{n+1}\right)^{2(n+1)}.$$

3. Пусть на отрезке [0, 1] с равномерным распределением выбраны n точек. Упорядочим точки на отрезке, и пусть ξ_i — длина i-го интервала слева из $n+1$ возможных интервалов.

Случайная величина $\xi=\frac{1}{2}\sum\limits_{j=1}^{n+1}\left|\xi_j-\frac{1}{n+1}\right|$ имеет распределение Шермана с n степенями свободы.

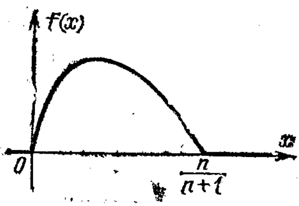

Рис. 17. Плотность распределения Шермана

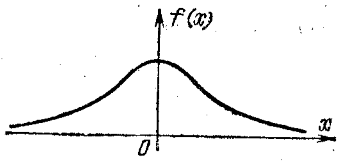

Рис. 18. Плотность распределения дисперсионного отношения Фишера

6.3.18. z-распределение (распределение дисперсионного отношения Фишера). 1. Случайная величина ξ имеет z-распределение (распределение дисперсионного отношения Фишера) с (m_1, m_2) степенями свободы, если

$$f(x)=\frac{2m_1^{m_1/2}m_2^{m_2/2}\Gamma\left(\frac{m_1+m_2}{2}\right)e^{m_1x}}{\Gamma\left(\frac{m_1}{2}\right)\Gamma\left(\frac{m_2}{2}\right)\left(m_2+m_1e^{2x}\right)^{(m_1+m_2)/2}},\quad x\in(-\infty,\infty)$$

(рис. 18).

2. Характеристическая функция:

$$\varphi(t) = \left(\frac{m_2}{m_1}\right)^{it/2} \frac{\Gamma\left(\frac{m_1+it}{2}\right)\Gamma\left(\frac{m_2-it}{2}\right)}{\Gamma\left(\frac{m_1}{2}\right)\Gamma\left(\frac{m_2}{2}\right)}.$$

Моменты:

$$\mathsf{M}\xi = 0; \quad \mathsf{D}\xi = \frac{1}{2}\,\frac{m_1+m_2}{m_1 m_2}.$$

3. Если η — случайная величина, имеющая F-распределение с (m_1, m_2) степенями свободы, то $\xi = \frac{1}{2}\ln\eta$ имеет z-распределение с (m_1, m_2) степенями свободы.

Широко используется в дисперсионном анализе.

6.3.19. Распределение Вейбулла — Гнеденко. 1. Случайная величина ξ имеет распределение Вейбулла — Гнеденко с параметрами (α, λ) $(\alpha, \lambda > 0)$, если

$$f(x) = \begin{cases} \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

(рис. 19).

2. Моменты:

$$\mathsf{M}\xi^k = \lambda^{-k/\alpha}\Gamma\left(\frac{k}{\alpha}+1\right);$$

$$\mathsf{D}\xi = \lambda^{-2/\alpha}\left\{\frac{2}{\alpha}\Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha^2}\left[\Gamma\left(\frac{1}{\alpha}\right)\right]^2\right\}.$$

3. Появляется как предельное распределение максимума. Пусть случайные величины ξ_k взаимно независимы и одинаково распределены и $\mathsf{P}\{\xi_k \leqslant x\} < 1$ для $x < \infty$. Положим

$$\xi_n^* = \max\left[\xi_1, \xi_2, \ldots, \xi_n\right].$$

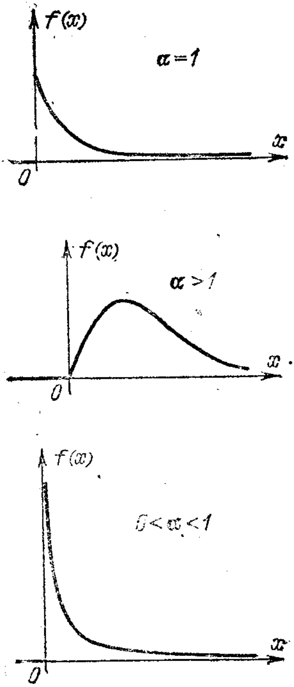

Рис. 19. Плотность распределения Вейбулла — Гнеденко

Для того чтобы при некотором выборе констант a_n распределения $F_n(x)$ случайных величин $\frac{1}{a_n}\xi_n^*$ сходились к распределению $F(x)$, не сосредоточенному в 0, необходимо и достаточно, чтобы функция $1 - F(x)$ была правильно меняющейся *) с показателем α $(\alpha < 0)$. В этом случае $F(x) = e^{-\lambda x^{\alpha}}$.

*) Функция $1 - F(x)$ называется *правильно меняющейся с показателем* α, если $1 - F(x) = x^{\alpha}L(x)$, где $L(x)$ — медленно меняющаяся функция, т. е. такая, что $L(tx)/L(t) \to 1$ при $t \to \infty$ и любом $x > 0$.

Распределение Вейбулла — Гнеденко часто используется в теории надежности для описания времени безотказной работы приборов.

6.4. Распределения Пирсона

6.4.1. Определение. *Распределениями Пирсона* называются непрерывные распределения, плотности вероятности которых являются решениями дифференциального уравнения

$$\frac{df(x)}{dx}=\frac{a_1x+a_0}{b_0+2b_1x+b_2x^2}f(x), \tag{4.1}$$

где a_0, a_1, b_0, b_1, b_2 — параметры распределения. Распределения Пирсона полностью определяются первыми четырьмя моментами.

Пусть μ_k — k-й центральный момент случайной величины, имеющей распределение Пирсона. Тогда, если $a_1=1$, то

$$\begin{aligned} a_0&=\frac{\mu_3\left(\mu_4+3\mu_2^2\right)}{A}, & b_1&=-\frac{\mu_3\left(\mu_4+3\mu_2^2\right)}{2A},\\ b_0&=-\frac{\mu_2\left(4\mu_2\mu_4-3\mu_3^2\right)}{A}, & b_2&=-\frac{2\mu_2\mu_4-3\mu_3^2-6\mu_2^3}{A}, \end{aligned} \tag{4.2}$$

где $A=10\mu_4\mu_2-18\mu_2^3-12\mu_3^2$.

В соответствии с распределением корней квадратного трехчлена $b_0+b_1x+b_2x^2$ различают 12 типов распределений Пирсона.

6.4.2. Тип I *). 1. $D<0$, $\lambda<0$, $b_0+2b_1x+b_2x^2=(\alpha+x)(-\beta+x)b_2$ $(\alpha,\ \beta>0)$.

2. Плотность вероятности:

$$f(x)=\begin{cases}\dfrac{\alpha^{2m}\beta^{2n}}{(\alpha+\beta)^{m+n+1}B(m+1,\ n+1)}(\alpha+x)^m(\beta-x)^n, & x\in[-\alpha,\ \beta],\\ 0, & x\notin[-\alpha,\ \beta],\end{cases}$$

где $B(m+1,\ n+1)=\dfrac{\Gamma(m+1)\Gamma(n+1)}{\Gamma(m+n+2)}$, $m>-1$, $n>-1$.

3. Если $\alpha=\beta$ (следовательно, $\lambda=0$) и $m=n$, то соответствующие распределения имеют тип II. Распределения типа II имеют вертикальную ось симметрии.

Если $m=-n$, то соответствующее распределение имеет **тип XII.**

4. Распределениями Пирсона типа I являются бета-распределения.

6.4.3. Тип III. 1. $D<0$, $\lambda=\infty$, $b_0+2b_1x+b_2x^2=2(\alpha+x)b_1$.

2. Плотность вероятности:

$$f(x)=\begin{cases}\dfrac{k^{m+1}}{\Gamma(m+1)}(x+\alpha)^m e^{-k(\alpha+x)}, & x>-\alpha,\ k>0,\\ 0, & x\leqslant-\alpha.\end{cases}$$

*) Через D обозначен дискриминант квадратного трехчлена $D=b_0b_2-b_1^2$, $\lambda=b_1^2/\left(b_0b_2\right)$.

3. Распределениями Пирсона типа III являются гамма-распределения.

6.4.4. Тип IV. 1. $D > 0$, $0 < \lambda < 1$, $b_0 + 2b_1x + b_2x^2 = (x^2 + \alpha^2)b_2$.

2. Плотность вероятности:

$$f(x) = c(\alpha^2 + x^2)^{-m} \exp\left\{-\nu \operatorname{arctg}\frac{x}{\alpha}\right\}, \quad x \in (-\infty, \infty), \quad m \geqslant 1/2,$$

где

$$c^{-1} = \int_{-\infty}^{\infty} (\alpha^2 + x^2)^{-m} \exp\left\{-\nu \operatorname{arctg}\frac{x}{\alpha}\right\} dx.$$

6.4.5. Тип VII. 1. $D > 0$, $\lambda = 0$, $b_0 + 2b_1x + b_2x^2 = (\alpha^2 + x^2)b_2$.

2. Плотность вероятности:

$$f(x) = \frac{\alpha}{B(m - 1/2,\ 1/2)} (\alpha^2 + x^2)^{-m}, \quad x \in (-\infty, \infty), \quad m \geqslant \frac{1}{2}.$$

3. Распределением Пирсона типа VII является распределение Стьюдента.

6.4.6. Тип V. 1. $D = 0$, $\lambda = 1$, $b_0 + 2b_1x + b_2x^2 = (x + \alpha)^2 b_2$.

2. Плотность вероятности:

$$f(x) = \begin{cases} \dfrac{\gamma^{m-1}}{\Gamma(m-1)} x^{-m} e^{-\gamma/x}, & \gamma > 0, \quad m > 1, \quad x > 0, \\ 0, \quad x \leqslant 0. \end{cases}$$

6.4.7. Тип VI. 1. $D < 0$, $\lambda > 1$, $b_0 + 2b_1x + b_2x^2 = (x + \alpha) \times (x - \beta)b_2$.

2. Плотность вероятности:

$$f(x) = \begin{cases} \dfrac{(\alpha + \beta)^{-(m+n+1)}}{B(-m-n-1,\ n+1)} (x + \alpha)^m (x - \beta)^n, & x > \beta, \quad m - 1 > 0, \\ & n > -1, \\ 0, \quad x \leqslant \beta. \end{cases}$$

6.4.8. Тип VIII. 1. $D < 0$, $\lambda < 0$, $b_0 + 2b_1x + b_2x^2 = x(\alpha + x)b_2$.

2. Плотность вероятности:

$$f(x) = \begin{cases} \dfrac{m+1}{\alpha^{m+1}} (x + \alpha)^m, & x \in [-\alpha, 0], \quad -1 < m < 0, \\ 0, \quad x \notin [-\alpha, 0]. \end{cases}$$

6.4.9. Тип IX. 1. $D < 0$, $\lambda < 0$, $b_0 + 2b_1x + b_2x^2 = x(\alpha + x)b_2$.

2. Плотность вероятности:

$$f(x) = \begin{cases} \dfrac{m+1}{\alpha^{m+1}} (x + \alpha)^m, & x \in [-\alpha, 0], \quad m < -1, \\ 0, \quad x \notin [-\alpha, 0]. \end{cases}$$

6.4.10. Тип X. 1. $D=0$, $\lambda=0$, $b_0+2b_1x+b_2x^2=b_0$, $a_1=0$.

2. Плотность вероятности:

$$f(x)=\begin{cases}\gamma e^{-\gamma x}, & x>0, \quad \gamma>0,\\ 0, & x\leqslant 0.\end{cases}$$

3. Распределением Пирсона типа X является показательное распределение.

6.4.11. Тип XI. 1. $D=0$, λ неопределенно, $b_0+2b_1x+b_2x^2=b_0$, $a_1\neq 0$.

2. Плотность вероятности:

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\},\quad x\in(-\infty,\infty).$$

3. Распределением Пирсона типа XI является нормальное распределение.

4. Распределения Пирсона широко используются в математической статистике при сглаживании распределений эмпирических данных. Для определения распределения Пирсона, аппроксимирующего наблюденные данные, вычисляют первые четыре момента и из уравнений (3.2) находят оценки параметров.

6.5. Многомерные распределения

6.5.1. Полиномиальное распределение. 1. k-мерный случайный вектор $\xi=(\xi_1, \xi_2, \ldots, \xi_k)$ имеет полиномиальное распределение с параметрами $(n;\ p_1, p_2, \ldots, p_k)$ $\left(0<p_i<1,\ \sum p_i=1\right)$, если

$$\mathsf{P}\{\xi=m\}=\mathsf{P}\{\xi_1=m_1, \xi_2=m_2, \ldots, \xi_k=m_k\}=$$
$$=\frac{n!}{m_1!m_2!\ldots m_k!}p_1^{m_1}p_2^{m_2}\ldots p_k^{m_k}$$

для $m=(m_1, m_2, \ldots, m_k)$, $\sum\limits_{i=1}^{k} m_i=n$.

2. Характеристическая функция:

$$\varphi(t)=\varphi(t_1, t_2, \ldots, t_k)=\left(\sum_{l=1}^{k}p_le^{it_l}\right)^n.$$

Моменты:

$$\mathsf{M}\xi=np=n(p_1, p_2, \ldots, p_k),$$
$$\operatorname{cov}(\xi_i, \xi_j)=\mathsf{M}(\xi_i-\mathsf{M}\xi_i)(\xi_j-\mathsf{M}\xi_j)=-np_ip_j,\quad i\neq j;$$
$$\mathsf{D}\xi_i=np_i(1-p_i).$$

3. Полиномиальное распределение есть многомерный аналог биномиального распределения. Маргинальное распределение каждой из компонент вектора ξ есть биномиальное распределение. Если

$\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(r)}$ — независимые k-мерные случайные величины с параметрами (n_1, p), (n_2, p), ..., (n_r, p) соответственно, то вектор $\xi = \sum_{l=1}^{r} \xi^{(l)}$ имеет полиномиальное распределение с параметрами $\left(\sum_{l=1}^{r} n_l, p\right)$ $(p = (p_1, p_2, \ldots, p_k))$.

Полиномиальное распределение является моделью случайного эксперимента, представляющего n независимых испытаний, исходом каждого из которых является событие одного из k непересекающихся классов; p_i $(0 < p_i < 1)$ означает вероятность того, что результат любого испытания принадлежит i-му классу, причем $\sum_{i=1}^{k} p_i = 1$.

6.5.2. Равномерное распределение. 1. Пусть S — ограниченное борелевское множество в $\mathbf{R}^k$. Обозначим через $\operatorname{mes} S$ его k-мерную меру Лебега.

Случайный вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_k)$ имеет равномерное распределение в S, если $\operatorname{mes} S > 0$, и его плотность вероятности $f(x) = f(x_1, x_2, \ldots, x_k)$ равна

$$f(x) = \begin{cases} 1/\operatorname{mes} S, & x \in S, \\ 0, & x \notin S. \end{cases}$$

В том важном частном случае, когда S — k-мерный параллелепипед $S = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_k, b_k]$,

$$f(x) = \begin{cases} \left(\prod_{l=1}^{k} (b_l - a_l)\right)^{-1}, & x \in S, \\ 0, \quad x \notin S. \end{cases}$$

2. Характеристическая функция в случае прямоугольной области имеет вид

$$\varphi(t) = \varphi(t_1, t_2, \ldots, t_k) = \frac{\prod_{l=1}^{k} \left(e^{it_l b_l} - e^{it_l a_l}\right)}{\prod_{l=1}^{k} (b_l - a_l)\, i^k \prod_{l=1}^{k} t_l}.$$

6.5.3. Двумерное нормальное распределение. 1. Двумерный случайный вектор $\xi = (\xi_1, \xi_2)$ имеет нормальное распределение, если его характеристическая функция $\varphi(t) = \varphi(t_1, t_2)$ равна

$$\varphi(t) = \exp\left\{i\left(m_1 t_1 + m_2 t_2\right) - \frac{1}{2}\left(c_{11} t_1^2 + 2c_{12} t_1 t_2 + c_{22} t_2^2\right)\right\},$$

причем квадратичная форма $Q(t) = Q(t_1, t_2) = \sum_{i,j=1}^{2} c_{ij} t_i t_j$, $c_{12} = c_{21}$, неотрицательна, т. е. $Q(t) \geqslant 0$ для любых действительных t_1, t_2. Если ранг квадратичной формы $Q(t)$ равен 2, т. е. детерминант матрицы $C = \{c_{ij}, i, j = 1, 2\}$ отличен от нуля, то двумерное нормальное распределение называется *невырожденным* или *собственным.* Если ранг квадратичной формы $Q(t)$ равен 0 или 1, то двумерное нормальное распределение называется *вырожденным* или *несобственным.*

2. Если случайный вектор ξ имеет невырожденное нормальное распределение, то его плотность $f(x) = f(x_1, x_2)$ равна

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$
$$\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right]\right\},$$

где

$$\det C = c_{11}c_{22} - c_{12}^2 = \sigma_1^2\sigma_2^2(1-\rho^2)$$

(рис. 20).

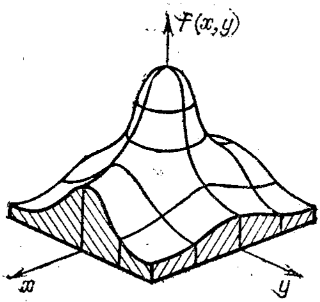

Рис. 20. Плотность невырожденного двумерного нормального распределения

Смысл величин m_1, m_2, σ_1, σ_2 и ρ следующий: $m_i = \mathsf{M}\xi_i$, $\sigma_i^2 = \mathsf{D}\xi_i$, $i=1, 2$; $\mathsf{M}\xi_1\xi_2 = \mathsf{M}\xi_2\xi_1 = \sigma_1\sigma_2\rho + m_1m_2$; $\rho = \dfrac{\mathsf{M}(\xi_1-m_1)(\xi_2-m_2)}{\sqrt{\mathsf{D}\xi_1\mathsf{D}\xi_2}}$ — коэффициент корреляции компонент ξ_1 и ξ_2.

Плотность невырожденного двумерного нормального распределения удобно записывать в виде

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}Q^{-1}(x_1-m_1, x_2-m_2)\right\} =$$
$$= \frac{1}{2\pi\sqrt{\det C}} \exp\left\{-\frac{1}{2}\left[c_{11}^{-1}(x_1-m_1)^2 + 2c_{12}^{-1}(x_1-m_1)(x_2-m_2) + c_{22}^{-1}(x_2-m_2)^2\right]\right\},$$

где $Q^{-1}(t_1, t_2) = \sum\limits_{i,j=1}^{2} c_{ij}^{-1}t_it_j$, c_{ij}^{-1} — элементы матрицы C^{-1}.

Маргинальные плотности нормального распределения:

$$f_{\xi_i}(x_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i}\exp\left\{-\frac{(x_i-m_i)^2}{2\sigma_i^2}\right\}, \quad i = 1, 2.$$

Условная плотность распределения компоненты ξ_1 при условии, что $\xi_2 = a$, равна

$$f(x_1 \mid \xi_2 = a) =$$
$$= \frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left[x_1 - m_1 - \frac{\rho\sigma_1}{\sigma_2}(a-m_2)\right]^2\right\},$$

$$\mathsf{M}\left\{\xi_1 \mid \xi_2 = a\right\} = m_1 + \frac{\rho\sigma_1}{\sigma_2}\left(a - m_2\right), \quad \mathsf{M}\left\{\left(\xi_1 - m_1\right)^2 \mid \xi_2 = a\right\} = \sigma_1^2\left(1 - \rho^2\right).$$

Плотность невырожденного нормального распределения сохраняет постоянное значение на эллипсах

$$\frac{1}{2\left(1-\rho^2\right)}\left[\frac{\left(x_1^2 - m_1\right)^2}{\sigma_1^2} - \frac{2\rho\left(x_1 - m_1\right)\left(x_2 - m_2\right)}{\sigma_1\sigma_2} + \frac{\left(x_2 - m_2\right)^2}{\sigma_2^2}\right] = \lambda^2,$$

называемых *эллипсами равных вероятностей*, причем вероятность попадания вектора $\xi = (\xi_1, \xi_2)$ внутрь такого эллипса равна $1 - e^{-\lambda^2}$.

3. Если двумерное нормальное распределение вырождено, то в случае, когда ранг квадратичной формы $Q(t_1, t_2)$ равен нулю, оно сосредоточено в точке $m = (m_1, m_2)$, т. е.

$$\mathsf{P}\{\xi = m\} = 1.$$

Если ранг квадратичной формы $Q(t_1, t_2)$ равен единице, то нормальное распределение сосредоточено на прямой, определяемой собственным вектором матрицы C, который соответствует ненулевому собственному значению.

6.5.4. Многомерное нормальное распределение. 1. Случайный вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_k)$ имеет многомерное нормальное распределение, если его характеристическая функция $\varphi(t) = \varphi(t_1, t_2, \ldots, t_k)$ имеет вид

$$\varphi(t) = \exp\left\{it^T m - \frac{1}{2}t^T C t\right\},$$

где C — неотрицательно определенная симметрическая $k \times k$-матрица, t^T — транспонированный вектор-столбец $t \in \mathbf{R}^k$.

Если ранг матрицы C равен k, т. е. $\det C \neq 0$, то распределение называется *невырожденным* или *собственным*. Смысл вектора $m = (m_1, \ldots, m_k)$ и матрицы $C = \{c_{ij},\ i, j = 1, \ldots, k\}$ следующий:

$$m = (m_1, \ldots, m_k) = (\mathsf{M}\xi_1, \ldots, \mathsf{M}\xi_k) = \mathsf{M}\xi,$$

$$C = \left\{c_{ij},\ i, j = 1, \ldots, k\right\} = \left\{\mathsf{M}\left(\xi_i - m_i\right)\left(\xi_j - m_j\right),\ i, j = 1, \ldots, k\right\} = \mathsf{M}\,(\xi - m)(\xi - m)^T$$ — ковариационная матрица.

Если $\det C \neq 0$, то C^{-1} называют *матрицей точности*.

2. Если случайный вектор ξ имеет невырожденное нормальное распределение, то его плотность распределения $f(x) = f(x_1, \ldots, x_k)$ равна

$$f(x) = \frac{1}{\sqrt{(2\pi)^k \det C}} \exp\left\{-\frac{1}{2}(x - m)^T C^{-1}(x - m)\right\}.$$

Пусть $\xi^{(1)} = (\xi_1, \ldots, \xi_r)$, $\xi^{(2)} = (\xi_{r+1}, \ldots, \xi_k)$. Соответственно такому разбиению случайного вектора ξ вектор математических ожиданий m представим в виде $m^{(1)} = (m_1, \ldots, m_r)$, $m^{(2)} = (m_{r+1}, \ldots, m_k)$, причем $m^{(i)} = \mathsf{M}\xi^{(i)}$ $(i = 1, 2)$, и ковариационная матрица C представима в блочном виде

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

где C_{11}, C_{12}, C_{21}, C_{22} являются соответственно $r\times r$-, $r\times(k-r)$-, $(k-r)\times r$-, $(k-r)\times(k-r)$-матрицами, причем

$$C_{ij}=\mathsf{M}\,(\xi^{(i)}-m^{(i)})(\xi^{(j)}-m^{(j)})^T.$$

Маргинальное распределение вектора $\xi^{(j)}$ является нормальным с вектором математических ожиданий $m^{(j)}$ и ковариационной матрицей C_{jj}, и соответствующая плотность распределения, при условии что $\det C_{jj}\neq 0$, имеет вид

$$f^{(j)}(x^{(j)})=\frac{1}{\sqrt{(2\pi)^{r_j}\det C_{jj}}}\exp\left\{-\frac{1}{2}(x^{(j)}-m^{(j)})^T C_{jj}^{-1}(x^{(j)}-m^{(j)})\right\},$$

$$r_1=r,\quad r_2=k-r.$$

Если случайный вектор $\xi=(\xi_1,\ \ldots,\ \xi_k)$ имеет невырожденное нормальное распределение, то условная плотность распределения $f^{(1)}(x^{(1)}\,|\,\xi^{(2)}=a^{(2)})$ вектора $\xi^{(1)}$, при условии, что $\xi^{(2)}=a^{(2)}$, равна

$$f^{(1)}(x^{(1)}\,|\,\xi^{(2)}=a^{(2)})=\frac{1}{\sqrt{(2\pi)^r}}\sqrt{\frac{\det C_{22}}{\det C}}\times$$

$$\times\exp\left\{-\frac{1}{2}\left(x^{(1)}-b^{(1)}\right)^T\left(C_{11}-C_{12}C_{22}^{-1}C_{21}\right)^{-1}\left(x^{(1)}-b^{(1)}\right)\right\},$$

где $b^{(1)}=m^{(1)}+C_{12}C_{22}^{-1}\left(a^{(2)}-m^{(2)}\right)$, причем

$$\mathsf{M}\left\{\xi^{(1)}\,|\,\xi^{(2)}=a^{(2)}\right\}=m^{(1)}+C_{12}C_{22}^{-1}\left(a^{(2)}-m^{(2)}\right);$$

$$\mathsf{M}\left\{\left(\xi^{(1)}-m^{(1)}\right)\left(\xi^{(1)}-m^{(1)}\right)^T\,|\,\xi^{(2)}=a^{(2)}\right\}=C_{11}-C_{12}C_{22}^{-1}C_{21}.$$

3. Если случайный вектор $\xi=(\xi_1,\ \ldots,\ \xi_k)$ имеет вырожденное нормальное распределение и ковариационная матрица C имеет ранг r $(0\leqslant r<k)$, то существует прямоугольная $k\times r$-матрица B такая, что $C=BB^T$, и r-мерный случайный вектор ξ_0, имеющий нулевое математическое ожидание и единичную ковариационную матрицу такие, что $\xi=m+B\xi_0$.

4. Многомерное нормальное распределение является наряду с одномерным одним из основных распределений теории вероятностей и математической статистики, что, прежде всего, связано с центральной предельной теоремой.

6.5.5. Распределение Дирихле. 1. Случайный вектор $\xi=(\xi_1,\ \ldots,\ \xi_k)$ имеет распределение Дирихле с векторным параметром $\alpha=(\alpha_1,\ \ldots,\ \alpha_k)$ $(\alpha_i>0,\ i=1,\ \ldots,\ k)$, если

$$f(x)=f(x_1,\ \ldots,\ x_k)=\begin{cases}\dfrac{\Gamma(\alpha_1+\alpha_2+\ldots+\alpha_k)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\ldots\Gamma(\alpha_k)}\,x_1^{\alpha_1}\ldots x_k^{\alpha_k}, & x\in S,\\ 0,\quad x\notin S, & \end{cases}$$

где S — $(k-1)$-мерный симплекс: $S=\left\{x\in\mathbf{R}^k\colon\ \sum\limits_{i=1}^{k}x_i=1,\ x_i\geqslant 0,\right.$ $\left.i=1,\ \ldots,\ k\right\}.$

2. Моменты:

$$\mathsf{M}\xi = \frac{1}{\alpha_0}\alpha, \quad \alpha_0 = \sum_{i=1}^{k}\alpha_i; \quad \operatorname{cov}\xi_i\xi_j = -\frac{\alpha_i\alpha_j}{\alpha_0^2(1+\alpha_0)}; \quad \mathsf{D}\xi_i = \frac{\alpha_i}{\alpha_0}.$$

3. Распределение Дирихле является многомерным аналогом бета-распределения.

6.5.6. Многомерное распределение Стьюдента (t-распределение). 1. Случайный вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_k)$ имеет *k-мерное распределение Стьюдента* (t-распределение) с n степенями свободы, вектором сдвига m и матрицей точности T, если

$$f(x) = \frac{\Gamma\left(\frac{k+n}{2}\right)\sqrt{\det T}}{\Gamma\left(\frac{n}{2}\right)\sqrt{(2\pi)^k}}\left[1 + \frac{1}{n}(x-m)^T T (x-m)\right]^{-(k+n)/2},$$

где T — симметрическая положительно определенная матрица.

2. Моменты:

$$\mathsf{M}\xi = m; \quad \operatorname{cov}\xi = \mathsf{M}(\xi - m)(\xi - m)^T = \frac{n}{n-2}T^{-1}, \quad n > 2.$$

3. Если η имеет нормальное распределение с нулевым вектором математических ожиданий и невырожденной ковариационной матрицей $C = T^{-1}$, а ζ имеет χ^2-распределение с n степенями свободы, то вектор $\xi = (\xi_1, \ldots, \xi_k)$, где $\xi_i = \frac{\sqrt{n}}{\sqrt{\zeta}}\eta_i + m_i$, имеет k-мерное распределение Стьюдента с n степенями свободы, вектором сдвига m и матрицей точности T.

Если случайный вектор ξ имеет k-мерное распределение Стьюдента с n степенями свободы, вектором сдвига m и матрицей точности T, то случайная величина

$$\zeta = \frac{1}{k}(\xi - m)^T T (\xi - m)$$

имеет F-распределение с k и n степенями свободы.

6.5.7. Распределение Уишарта. 1. Пусть ξ — симметрическая $k \times k$-матрица либо, что эквивалентно, $k(k+1)/2$-мерный вектор.

Обозначим через X симметрическую положительно определенную $k \times k$-матрицу. Случайная матрица ξ имеет невырожденное распределение Уишарта с n степенями свободы и матрицей точности T, если $\det T \neq 0$ и

$$f(X) = \frac{(\det T)^{n/2}(\det X)^{(n-k-1)/2}}{2^{nk/2}\pi^{k(k-1)/4}\prod_{j=1}^{k}\Gamma\left(\frac{n+1-j}{2}\right)}\exp\left\{-\frac{1}{2}\operatorname{Sp}TX\right\},$$

где $\operatorname{Sp}TX$ обозначает след матрицы TX.

2. Характеристическая функция: пусть t — симметрическая матрица вида

$$t = \begin{bmatrix} 2t_{11} & t_{12} & \ldots & t_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ t_{1k} & t_{2k} & \ldots & 2t_{kk} \end{bmatrix}.$$

Характеристическая функция распределения Уишарта определяется на множестве симметрических матриц вида (4.1):

$$\varphi(t) = \left[\frac{\det T}{\det(T - it)}\right]^{n/2}.$$

3. Пусть $\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(n)}$ — независимые нормально распределенные векторы с нулевыми математическими ожиданиями и невырожденной ковариационной матрицей C. Случайная матрица $\xi = \sum_{l=1}^{n} \xi^{(l)}\xi^{(l)^T}$ имеет распределение Уишарта с n степенями свободы и матрицей точности $T = C^{-1}$. Распределение Уишарта является многомерным аналогом χ^2-распределения.

6.6. Устойчивые распределения

6.6.1. Определение. Функция распределения $F(x)$ называется *устойчивой*, если для любых действительных $a_1 > 0$, $a_2 > 0$, b_1, b_2 найдутся $a > 0$ и b такие, что имеет место равенство

$$F(a_1x + b_1) * F(a_2x + b_2) = F(ax + b), \tag{6.1}$$

где $*$ — операция свертки.

Если $\varphi(t)$ — характеристическая функция устойчивого распределения, то для любых $a_1 > 0$ и $a_2 > 0$ найдутся b и $a > 0$ такие, что

$$\varphi\left(\frac{t}{a_1}\right)\varphi\left(\frac{t}{a_2}\right) = \varphi\left(\frac{t}{a}\right)e^{-itb}.$$

Важная роль устойчивых распределений связана со следующим результатом.

Теорема 1. *Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые одинаково распределенные случайные величины и*

$$\eta_n = \frac{1}{\beta_n}\sum_{k=1}^{n} \xi_k - \alpha_n, \tag{6.2}$$

где $\beta_n > 0$ и α_n — некоторые нормирующие и центрирующие константы соответственно.

Если $F_n(x)$ — функция распределения случайных величин η_n, то предельными распределениями для $F_n(x)$ при $n \to \infty$ могут быть лишь устойчивые распределения.

Обратно, для любого устойчивого распределения $F(x)$ существует последовательность случайных величин вида (6.2) такая, что $F_n(x)$ сходится к $F(x)$ при $n \to \infty$.

Отсюда, в частности, следует, что устойчивые распределения безгранично делимы.

6.6.2. Характеризация устойчивых распределений.

Теорема 2. *Для того чтобы функция распределения $\varphi(t)$ была устойчивой, необходимо и достаточно, чтобы логарифм ее характеристической функции $\varphi(t)$ имел вид*

$$\ln\varphi(t) = i\gamma t - c|t|^{\alpha}\left\{1 + i\beta\frac{t}{|t|}\omega(t, \alpha)\right\}, \tag{6.3}$$

где α, β, γ, c — *постоянные, причем* $-1 \leqslant \beta \leqslant 1$, $0 \leqslant \alpha \leqslant 2$, $c \geqslant 0$ *и*

$$\omega(t, \alpha) = \begin{cases} \operatorname{tg} \dfrac{\pi}{2} \alpha, & \text{если } \alpha \neq 1, \\ \dfrac{2}{\pi} \ln |t|, & \text{если } \alpha = 1. \end{cases} \tag{6.4}$$

Параметр α называется *параметром устойчивости* или *характеристическим показателем*.

Все устойчивые распределения с характеристическим показателем $\alpha > 0$ непрерывны, и их плотности в каждой точке имеют производные всех порядков.

Из (6.3) следует, что плотность вероятности устойчивого распределения имеет вид

$$f(x; \alpha, \beta, \gamma, c) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \exp\left\{i\gamma t - c|t|^{\alpha} \times \right.$$
$$\left. \times \left(1 + i\beta \frac{t}{|t|} \omega(t, \alpha)\right)\right\} dt. \tag{6.5}$$

Плотности устойчивых распределений не выражаются через элементарные функции, за исключением следующих случаев:

1) нормальное распределение устойчиво с характеристическим показателем $\alpha = 2$;

2) распределение Коши устойчиво с характеристическим показателем $\alpha = 1$;

3) распределение с плотностью

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi x^3}} e^{-1/2x}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

устойчиво с характеристическим показателем $\alpha = 1/2$;

4) вырожденное распределение устойчиво с характеристическим показателем $\alpha = 0$.

Для плотности $f(x; \alpha, \beta, \gamma, c)$ устойчивого распределения с $\alpha > 0$ имеют место равенства

$$f(x; \alpha, \beta, \gamma, c) = \begin{cases} c^{-1/\alpha} f\left((x-\gamma) c^{-1/2}; \alpha, \beta, 0, 1\right), & \alpha \neq 1, \\ c^{-1} f\left(\left(\dfrac{x-\gamma}{c}\right) - \beta \dfrac{2}{\pi} \ln c; 1, \beta, 0, 1\right), & \alpha = 1. \end{cases} \tag{6.6}$$

Следовательно, не уменьшая общности, можно считать, что $\gamma = 0$, $c = 1$. Если $f(x; \alpha, \beta) = f(x; \alpha, \beta, 0, 1)$, то

$$f(x; \alpha, \beta) = f(-x; \alpha, -\beta). \tag{6.7}$$

6.6.3. Асимптотические разложения. Если $F(x)$ — устойчивое распределение с характеристическим показателем α, то существует константа $c_1 > 0$ такая, что

$$\lim_{x \to \infty} x^{\alpha} \{1 - F(x) + F(-x)\} = c_1. \tag{6.8}$$

Каждый устойчивый закон с характеристическим показателем α $(0 < \alpha < 2)$ имеет конечные абсолютные моменты всех порядков p $(0 < p < \alpha)$.

Пусть $f(x;\ \alpha,\ \beta)$ — плотность устойчивого распределения. Если $0 < \alpha < 1$ и $x > 0$, то

$$f(x;\ \alpha,\ \beta) = \frac{1}{\pi x}\sum_{k=1}^{\infty}(-1)^{k+1}\sin\left[\frac{k\pi}{2}\alpha(\beta+1)\right]\frac{\Gamma(k\alpha+1)}{k!}x^{-k\alpha}. \tag{6.9}$$

В случае $\alpha > 1$ ряд справа в (6.9) расходится.

Если $1 < \alpha < 2$, $x > 0$, то

$$f(x;\ \alpha,\ \beta) = \frac{1}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{\Gamma((k+1)/\alpha)}{\alpha k!}x^k\times$$
$$\times\cos\left[\frac{k\pi}{2}\left(1+\left(1+\frac{1}{k}\right)\frac{2-\alpha}{\alpha}\beta\right)\right]. \tag{6.10}$$

В случае $\alpha < 1$ ряд справа в (6.10) расходится.

Если $\alpha = 1$, $x > 0$, то для любого N

$$f\left(x+\frac{2\beta}{\pi}\ln x,\ 1,\ \beta\right) = \frac{1}{\pi x}\sum_{k=0}^{N}\frac{b_k}{k!}x^{-k}+O\left(x^{-N-2}\right),$$

где $b_k = \operatorname{Im}\int_0^{\infty}e^{-t}t^k\left(i+i\beta-\frac{2\beta}{\pi}\ln t\right)^k dt.$

6.7. Сингулярные распределения

6.7.1. Распределения канторовского типа. Пусть p и q — целые положительные числа, $L(q) = \{i_1,\ i_2,\ \ldots,\ i_q\}$, где $1 \leqslant i_1 < i_2 < \ldots$ $\ldots < i_q \leqslant p$. Положим $C_0 = [0,\ 1]$ и построим последовательность $C_1,\ C_2,\ \ldots,\ C_n,\ \ldots$ подмножеств отрезка $[0,\ 1]$ с помощью следующего алгоритма: разобьем C_0 на p равных частей, обозначим через $\Delta_j^{(1)}$ j-й по счету открытый интервал длины $1/p$, полученный в результате разбиения, и положим $C_1 = C_0 \diagdown \bigcup_{j\in L(q)}\Delta_j^{(1)}$; с каждым из $p-q$ отрезков, составляющих C_1, поступим аналогично. Описанный алгоритм приводит к искомой последовательности множеств C_0, $C_1,\ C_2,\ \ldots,\ C_n,\ \ldots$

Примем обозначения

$$f_n(x) = \begin{cases}\gamma_n, & x\in C_n,\\ 0, & x\notin C_n,\end{cases}\qquad F_n(x) = \int_0^x f_n(u)\,du,$$

где γ_n определяется условием

$$\int_0^1 f_n(x)\,dx = 1, \quad \text{т. е.} \quad \gamma_n = \left(\frac{p}{p-q}\right)^n.$$

Предел $F(x) = \lim\limits_{n\to\infty} F_n(x)$ назовем *функцией распределения канторовского типа* $C\{p, L(q)\}$. Поскольку $\lambda(C_n) = \left(\frac{p-q}{p}\right)^n \to 0$, $n \to \infty$, где $\lambda(\cdot)$ — мера Лебега, то распределение, порождаемое $F(x)$, сингулярно: $\frac{d}{dx} F(x) = 0$ для $x \in [0, 1] \setminus C$. Совершенное множество C является фрактальным в том смысле, что его размерность Хаусдорфа — Безиковича равна

$$d = \frac{\ln(p-q)}{\ln p}.$$

6.7.2. Квазиравномерные функции распределения канторовского типа. Пусть $p = 2m + 1$ (п. 6.7.1) — нечетное положительное число, $C_0, C_1, C_2, \ldots, C_n, \ldots$ — последовательность замкнутых множеств, определенная выше. Определим новую последовательность C_0^k, C_1^k, $C_2^k, \ldots, C_n^k, \ldots$ следующим образом: $C_0^k = C_0$, $C_1^k = C_1$ и, начиная с $n = 2$, через C_n^k будем обозначать множество, получаемое из C_n так, чтобы соседние отрезки, образующие C_n (для $n \leqslant k$ k фиксировано), были равноудалены друг от друга в C_n^k (для $n > k$ C_n^k строятся так же, как в п. 6.7.1). Обозначим

$$f_n(x) = \begin{cases} \gamma_n, & x \in C_n^k, \\ 0, & x \notin C_n^k, \end{cases}$$

$$F_n(x) = \int_0^x f_n(u)\,du,$$

где $\gamma_n = \left(\frac{p}{p-q}\right)^n = \left(\frac{2m+1}{m+1}\right)^n$. Предел $F^{(k)}(x) = \lim\limits_{n\to\infty} F_n^{(k)}(x)$ назовем *квазиравномерной функцией распределения канторовского типа* $C\{p, k\}$, поскольку для любого $x \in [0, 1]$, $\varepsilon > 0$ найдется k такое, что

$$\lim_{n\to\infty} \left| F_n^{(k)}(x) - x \right| < \varepsilon.$$

6.7.3. Сингулярные распределения Салема. Пусть $\alpha \in (0, 1)$ и $\alpha \neq 1/2$. Определим $S(\alpha)$ — преобразование отрезка прямой, соединяющей точки $\{x, y\}$ и $\{x + \Delta x, y + \Delta y\}$ плоскости ($\Delta x > 0$, $\Delta y > 0$) в двухзвенную ломаную, соединяющую точки $\{x, y\}$, $\{x + \Delta x/2, y + \alpha\Delta y\}$ и $\{x + \Delta x/2, y + \alpha\Delta y\}$, $\{x + \Delta x, y + \Delta y\}$. Положим $F_0(x) = x$ ($x \in [0, 1]$). К каждому из отрезков, соединяющих $\{0, 0\}$ и $\{1/2, 1/2\}$, $\{1/2, 1/2\}$ и $\{1, 1\}$, применим преобразование $S(\alpha)$ и обозначим результат через $F_1(x)$. Далее поступаем аналогичным образом: делим полученные отрезки ломаных пополам и к каждому применяем преобразование $S(\alpha)$. Пусть $F_n(x)$ — функция, полученная после n-го шага описанного алгоритма.

Функцию $F(x) = \lim\limits_{n\to\infty} F_n(x)$ назовем *функцией распределения Салема* типа $S(\alpha)$; она определяет сингулярное распределение, поскольку $\frac{d}{dx} F(x) = 0$ для всех $x \in [0, 1]$, за исключением счетного числа точек вида $k/2^m$, $k = 1, \ldots, 2^m - 1$, $m = 1, 2, \ldots$

Литература: [8, 41, 49, 72].

Глава 7. СЛУЧАЙНЫЕ БЛУЖДАНИЯ

7.1. Процессы восстановления

7.1.1. Определение. Основные теоремы. Последовательные суммы

$$\zeta_n = \sum_{k=0}^{n} \xi_k, \quad n \geqslant 0, \quad \zeta_0 = \xi_0 = 0, \tag{1.1}$$

независимых неотрицательных одинаково распределенных случайных величин с функцией распределения $F(x) = \mathsf{P}\{\xi_k \leqslant x\}$ образуют *моменты* (*точки*) *восстановления* (переключения, наступления некоторого события и т. п.).

Процессом восстановления называется величина

$$\nu(t) = \max\{n\colon \zeta_n \leqslant t\}, \tag{1.2}$$

определенная при каждом $t \geqslant 0$. Так что

$$\nu(t) = n, \quad \zeta_n \leqslant t < \zeta_{n+1}, \ n \geqslant 0, \tag{1.3}$$

или иначе

$$\nu(t) = \sum_{n=1}^{\infty} \chi(\zeta_n \leqslant t). \tag{1.4}$$

Траектории процесса восстановления ступенчаты, непрерывны справа и имеют скачки, равные единице, в точках восстановления ζ_n $(n \geqslant 0)$.

Процесс восстановления $\nu(\tau)$, остановленный в показательно распределенный момент времени τ с параметром s $(0 < s < 1)$, имеет *геометрическое распределение* с параметром $f(s) = \mathsf{M}e^{-s\xi_k}$:

$$\mathsf{P}\{\nu(\tau) = n\} = f^n(s)\,(1 - f(s)). \tag{1.5}$$

Производящая функция

$$\mathsf{M}z^{\nu(\tau)} = s\int_0^{\infty} e^{-st}\mathsf{M}z^{\nu(t)}\,dt \tag{1.6}$$

представима формулой

$$\mathsf{M}z^{\nu(\tau)} = \frac{1 - f(s)}{1 - zf(s)}. \tag{1.7}$$

Для факториальных моментов

$$\mathsf{M}\nu(\tau)^{[n]} = s\int_0^\infty e^{-st}\mathsf{M}\nu(t)^{[n]}\,dt \tag{1.8}$$

имеет место формула

$$\mathsf{M}\nu(\tau)^{[n]} = n!\left[\frac{f(s)}{1 - f(s)}\right]^n, \quad n \geqslant 1. \tag{1.9}$$

В частности,

$$\mathsf{M}\nu(\tau) = s\int_0^\infty e^{-st}\mathsf{M}\nu(t)\,dt = \frac{f(s)}{1 - f(s)}. \tag{1.10}$$

Функция $N(t) = \mathsf{M}\nu(t) + 1$ называется *функцией восстановления.* Функция восстановления $N(t)$ конечна при всех $t > 0$ и представима через распределения сумм в виде (см. (1.4))

$$N(t) = \sum_{n=0}^\infty \mathsf{P}\{\zeta_n \leqslant t\}. \tag{1.11}$$

Функция восстановления удовлетворяет *интегральному уравнению восстановления*

$$N(t) = 1 + \int_0^t N(t - x)\,dF(x), \quad t > 0. \tag{1.12}$$

Асимптотическое поведение функции восстановления определяется соотношением при $\mathsf{M}\xi \leqslant \infty$

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mathsf{M}\xi}. \tag{1.13}$$

Теорема 1. *Уравнение восстановления*

$$U(t) = g(t) + \int_0^t U(t - x)\,dF(x), \quad t > 0, \tag{1.14}$$

в котором $g(t)$ ограничена в каждом конечном интервале, имеет единственное решение $U(t)$ на полуоси $[0, +\infty)$, представимое в виде

$$U(t) = \int_0^t g(t - x)\,dN(x). \tag{1.15}$$

Теорема восстановления. *Если распределение величин ξ_k неарифметическое, то для всех $h > 0$*

$$\lim_{t\to\infty} [N(t + h) - N(t)] = h/\mathsf{M}\xi_1. \tag{1.16}$$

Если же величины ξ_k имеют арифметическое распределение, то (1.16) *имеет место для h, кратных шагу распределения ξ_k.*

Функция $g(t)$, заданная на полуоси $[0, +\infty)$, *непосредственно интегрируема по Риману*, если

$$\int_0^\infty g(t)\,dt = \lim_{\varepsilon\to 0} \varepsilon \sum_{n=1}^{\infty} \underline{g}_n^\varepsilon = \lim_{\varepsilon\to 0} \varepsilon \sum_{n=1}^{\infty} \bar{g}_n^\varepsilon, \tag{1.17}$$

где нижние и верхние римановы суммы в (1.17) определяются экстремумами функции g

$$\underline{g}_n^\varepsilon = \min_{(n-1)\varepsilon\leqslant t\leqslant n\varepsilon} g(t), \quad \bar{g}_n^\varepsilon = \max_{(n-1)\varepsilon\leqslant t\leqslant n\varepsilon} g(t), \quad n = 1, 2, \ldots$$

Теорема 2 (восстановления узловая). *Для функции $g(t)$, непосредственно интегрируемой по Риману на полуоси $[0, +\infty)$, в случае неарифметического распределения ξ_k выполняется соотношение*

$$\lim_{t\to\infty} \int_0^t g(t-x)\,dN(x) = \frac{1}{\mathsf{M}\xi_1}\int_0^\infty g(t)\,dt; \tag{1.18}$$

в случае арифметического распределения ξ_k с шагом d выполняется соотношение

$$\lim_{n\to\infty} \int_0^{t+nd} g(t+nd-x)\,dN(x) = \frac{d}{\mathsf{M}\xi_1}\sum_{k=0}^{\infty} g(t+kd). \tag{1.19}$$

Процесс восстановления с запаздыванием задается начальным распределением $F_0(t) = \mathsf{P}(\xi_0 \leqslant t)$, отличным от распределения $F(t)$ остальных интервалов между моментами восстановления.

Процесс восстановления с постоянной скоростью восстановления, когда функция восстановления $N(t)$ линейна по t: $N(t) = at + 1$, называется *стационарным* процессом восстановления.

Примеры. 1. Процесс восстановления с показательным распределением интервалов между скачками $\mathsf{P}(\xi_k \leqslant t) = 1 - e^{-at}$ $(t \geqslant 0)$ является пуассоновским процессом с функцией восстановления $N(t) = at + 1$ $(t \geqslant 0)$. Таким образом, пуассоновский процесс восстановления стационарный.

2. Процесс восстановления с начальным распределением

$$F_0(t) = \mathsf{P}(\xi_0 \leqslant t) = \frac{1}{\mathsf{M}\xi_1}\int_0^t [1 - F(x)]\,dx \tag{1.20}$$

также стационарный. Его функция восстановления $N(t) = t/(\mathsf{M}\xi_1) + 1$. Распределение (1.20) называют *стационарным* распределением запаздывания.

Распределение (1.20) служит предельным для линейчатых процессов (см. п. 7.1.3, следствие 1).

7.1.2. Дополнения и уточнения. 1. Если интервалы между моментами восстановления ξ_k имеют равномерно непрерывную плотность $p(t)$ и функция $q(t) = \sup_{x\geqslant t} p(x)$ интегрируема на полуоси

$[0, +\infty)$, то функция восстановления $N(t)$ имеет ограниченную равномерно непрерывную производную $N'(t)$, для которой

$$\lim_{t\to\infty} N'(t) = 1/\mathsf{M}\xi_1. \tag{1.21}$$

2. Для неарифметического распределения $F(t)$ с конечной дисперсией $\mathsf{D}\xi_k = \sigma^2 < \infty$ выполняется соотношение

$$\lim_{t\to\infty}\left[N(t) - \frac{t}{\mathsf{M}\xi_1}\right] = \left[\frac{\sigma}{\mathsf{M}\xi_1}\right]^2 + \frac{1}{2}. \tag{1.22}$$

3. Для процесса восстановления имеет место усиленный закон больших чисел:

$$\mathsf{P}\left(\lim_{t\to\infty}\frac{1}{t}\nu(t) = \frac{1}{\mathsf{M}\xi_1}\right) = 1. \tag{1.23}$$

4. Если величины ξ_k имеют *устойчивое распределение* с параметром $\alpha > 1$ и $\mathsf{M}e^{-\lambda\xi_k} = e^{-c\lambda^\alpha}$ $(c > 0)$, то

$$\lim_{t\to\infty}\frac{N(t)}{t^\alpha} = \frac{1}{c}. \tag{1.24}$$

5. Если величины ξ_k имеют интегрируемую характеристическую функцию $f(s) = \mathsf{M}e^{is\xi_k}$ и $f(s) - 1 \sim cs\log s$ при $s \to 0$, то

$$N(t) \sim \frac{t}{c\log t}, \quad t \to \infty. \tag{1.25}$$

6. Пусть η_k $(k \geqslant 1)$ — последовательность независимых одинаково распределенных случайных величин с $\mathsf{M}e^{i\lambda\eta_k} = \varphi(\lambda)$. *Общий процесс восстановления* задается соотношением

$$\nu_0(t) = \sum_{k=1}^{\nu(t)} \eta_k, \tag{1.26}$$

где $\nu(t)$ — процесс восстановления, определяемый моментами восстановления $\zeta_n = \sum\limits_{k=0}^{n} \xi_k$ $(n > 0)$ с $\mathsf{M}e^{-s\xi_k} = f(s)$.

Характеристическая функция общего процесса восстановления, остановленного в показательно распределенный момент τ с параметром s $(0 < s < 1)$, имеет вид

$$\mathsf{M}e^{i\lambda\nu_0(\tau)} = s\int_0^\infty e^{-st}\mathsf{M}e^{i\lambda\nu_0(t)}\,dt \tag{1.27}$$

и определяется соотношением

$$\mathsf{M}e^{i\lambda\nu_0(\tau)} = \frac{1 - f(s)}{1 - \varphi(\lambda)f(s)}. \tag{1.28}$$

В частности, при $\mathsf{M}|\eta_k| < \infty$

$$\mathsf{M}\nu_0(t) = \mathsf{M}\eta_k\mathsf{M}\nu(t); \tag{1.29}$$

при $M\xi_k < \infty$

$$P\left(\lim_{t\to\infty} \nu_0(t)/\nu(t) = M\eta_k\right) = 1. \tag{1.30}$$

7.1.3. Линейчатые процессы восстановления. Этими процессами определяются следующие соотношения: *перескок*, или *остающееся время ожидания* момента восстановления,

$$\gamma_t^+ = \zeta_{\nu(t)+1} - t, \quad t \geqslant 0; \tag{1.31}$$

недоскок, или *время, прошедшее с момента восстановления*,

$$\gamma_t^- = t - \zeta_{\nu(t)}, \quad t \geqslant 0. \tag{1.32}$$

Процессы γ_t^+ и γ_t^- являются кусочно-линейными. Между двумя соседними точками восстановления ζ_n и ζ_{n+1} процесс γ_t^+ убывает линейно от ξ_{n+1} до нуля, а процесс γ_t^- возрастает линейно от нуля до ξ_{n+1}.

Сумма $\gamma_t = \gamma_t^+ + \gamma_t^- = \zeta_{\nu(t)+1}$ есть длина интервала между соседними моментами восстановления, накрывающего точку t. Заметим, что распределение $\zeta_{\nu(t)+1}$ не совпадает с распределением ξ_k при фиксированном k.

В качестве исходных соотношений для изучения линейчатых процессов восстановления можно принять следующие стохастические соотношения, в которых равенство $\doteq$ понимается в смысле совпадения распределений случайных величин, стоящих по обеим сторонам равенств:

$$\gamma_t^+ \doteq \gamma_{t-\xi_1}^+, \quad t \geqslant 0; \quad \gamma_t = -t, \quad t < 0; \tag{1.33}$$

$$\gamma_t^- \doteq \chi(\xi_1 \leqslant t)\,\gamma_{t-\xi_1}^- + \chi(\xi_1 > t)\,t, \quad t \geqslant 0. \tag{1.34}$$

Совместная производящая функция для γ_t^+ и γ_t^- задается формулой

$$M\exp\left\{-\lambda\gamma_\tau^+ - \mu\gamma_\tau^-\right\} = \frac{s}{s+\mu-\lambda}\cdot\frac{M\exp\{-\lambda\xi_1\} - M\exp\{-(\mu+s)\xi_1\}}{1 - M\exp\{-s\xi_1\}}. \tag{1.35}$$

Производящие функции линейчатых процессов восстановления определяются формулами

$$M\exp\left\{-\lambda\gamma_\tau^+\right\} = s\int_0^\infty \exp\{-st\}\, M\exp\left\{-\lambda\gamma_t^+\right\} dt = \frac{s}{s-\lambda}\cdot\frac{M\exp\{-\lambda\xi_1\} - M\exp\{-s\xi_1\}}{1 - M\exp\{-s\xi_1\}} \tag{1.36}$$

и

$$M\exp\left\{-\lambda\gamma_\tau^-\right\} = s\int_0^\infty \exp\{-st\}\, M\exp\left\{-\lambda\gamma_t^-\right\} dt = \frac{s}{s+\lambda}\cdot\frac{1 - M\exp\{-(s+\lambda)\xi_1\}}{1 - M\exp\{-s\xi_1\}}. \tag{1.37}$$

Теорема 3. *При* $\mathsf{M}\xi_k < \infty$ *в случае неарифметического распределения величин* ξ_k *существует*

$$\lim_{t\to\infty} \mathsf{P}\left(\gamma_t^+ > x,\ \gamma_t^- > y\right) = \frac{1}{\mathsf{M}\xi_1} \int_{x+y}^{\infty} [1 - F(u)]\, du; \tag{1.38}$$

в случае целочисленных величин ξ_k *существует*

$$\lim_{t\to\infty} \mathsf{P}\left(\gamma_t^+ = k,\ \gamma_t^- = r\right) = \frac{1}{\mathsf{M}\xi_1} \mathsf{P}(\xi_1 \geqslant k + r), \tag{1.39}$$

когда t *пробегает натуральный ряд чисел.*

Следствие 1. *Предельные распределения* γ_t^+ *и* γ_t^- *при* $t\to\infty$ *совпадают с распределением стационарного запаздывания:*

$$\lim_{t\to\infty} \mathsf{P}\left(\gamma_t^+ \leqslant x\right) = \lim_{t\to\infty} \mathsf{P}\left(\gamma_t^- \leqslant x\right) = \frac{1}{\mathsf{M}\xi_1} \int_0^x (1 - F(y))\, dy. \tag{1.40}$$

Следствие 2.

$$\lim_{t\to\infty} \mathsf{P}\left(\gamma_t^+ + \gamma_t^- \leqslant x\right) = \frac{1}{\mathsf{M}\xi_1} \int_0^x y\, dF(y). \tag{1.41}$$

В частности,

$$\lim_{t\to\infty} \mathsf{M}\left[\gamma_t^+ + \gamma_t^-\right] = 2\mathsf{M}\xi_1. \tag{1.42}$$

Примеры. 1. Если $\mathsf{P}(\xi_k \leqslant x) = 1 - e^{-ax}$, то и

$$\mathsf{P}\left(\gamma_t^+ \leqslant x\right) = 1 - e^{-ax}. \tag{1.43}$$

2. Если $\mathsf{P}(\xi_0 \leqslant x) = \dfrac{1}{\mathsf{M}\xi_1} \displaystyle\int_0^x [1 - F(y)]\, dy$, то

$$\mathsf{P}\left(\gamma_t^+ \leqslant x\right) = \frac{1}{\mathsf{M}\xi_1} \int_0^x [1 - F(y)]\, dy, \tag{1.44}$$

что поясняет стационарность процесса восстановления со стационарным запаздыванием ξ_0.

7.2. Классификация случайных блужданий на прямой

7.2.1. Критерий возвратности. Последовательность сумм

$$\zeta_n = \sum_{k=1}^{n} \xi_n, \quad n \geqslant 0, \quad \zeta_0 = 0, \tag{2.1}$$

независимых одинаково распределенных случайных величин ξ_k с функцией распределения $F(x)$ $(0 < F(0) < 1)$ определяет случайное блуждание на вещественной прямой.

Величины ξ_k называют *шагами* (*скачками*) блуждания, суммы ζ_n определяют *положение блуждания в момент n* (*после n шагов*).

Функцией восстановления случайного блуждания ζ_n $(n \geqslant 0)$ называется функция интервалов

$$N(A) = \sum_{n=0}^{\infty} \mathsf{P}(\zeta_n \in A), \tag{2.2}$$

где A — интервал на вещественной прямой.

Если F — арифметическое распределение с шагом d, то всегда предполагается, что интервал A содержит по крайней мере одну точку вида nd.

Теорема 1. *Имеет место альтернатива: либо $N(A) < \infty$ для всех конечных интервалов, либо $N(A) = \infty$ для всех интервалов.*

Случайное блуждание ζ_n $(n \geqslant 0)$ называется *невозвратным*, если $N(A) < \infty$ для всех конечных интервалов, и *возвратным*, если $N(A) = \infty$ для всех интервалов.

Введем случайную величину ν_A — число попаданий случайного блуждания ζ_n $(n \geqslant 0)$ в интервал A:

$$\nu_A = \sum_{n=0}^{\infty} \chi(\zeta_n \in A). \tag{2.3}$$

Теорема 2. *Для невозвратного случайного блуждания число попаданий в каждый конечный интервал конечно с вероятностью единица, а математическое ожидание числа попаданий в интервал A равно $\mathsf{M}\nu_A = N(A)$.*

Для возвратного случайного блуждания число попаданий в каждый конечный интервал равно бесконечности с вероятностью единица.

Теорема 3 (общий критерий возвратности случайного блуждания). *Случайное блуждание со скачками ξ_k невозвратно тогда и только тогда, когда*

$$\lim_{s \uparrow 1} \int_C \operatorname{Re} \frac{1}{1 - s\mathsf{M}\exp\{i\lambda\xi_k\}}\, d\lambda < \infty. \tag{2.4}$$

Критерий (2.4) сохраняется и для случайных блужданий в конечномерном евклидовом пространстве, т. е. когда $\xi_k = (\xi_{k1}, \ldots, \xi_{km})$ — m-мерные случайные векторы. При этом $\lambda\xi_k = \sum_{r=1}^{m} \lambda_r \xi_{kr}$ — скалярное произведение векторов λ и ξ_k, C — контур, содержащий начало координат.

В частности, *двумерное случайное блуждание возвратно*, если $\mathsf{M}\xi_k = 0$ и $\mathsf{M}\xi_k^2 < \infty$.

Случайное блуждание в трехмерном (а также в m-мерном при $m > 3$) евклидовом пространстве невозвратно.

Теорема 4. *Одномерное случайное блуждание со скачками ξ_k и $\mathsf{M}|\xi_k| < \infty$ возвратно тогда и только тогда, когда $\mathsf{M}\xi_k = 0$.*

7.2.2. Типы случайных блужданий. Введем случайные величины

$$\zeta^+ = \sup_{n \geqslant 0} \zeta_n, \quad \zeta^- = \inf_{n \geqslant 0} \zeta_n. \tag{2.5}$$

Теорема 5. *Существует только три типа случайных блужданий:*

1) *осциллирующий:* $\mathsf{P}\{\zeta^+ = \infty\} = \mathsf{P}\{\zeta^- = -\infty\} = 1$;

2) *уходящий в* $-\infty$: $\mathsf{P}(\zeta^- = -\infty) = 1$, $\mathsf{P}(\zeta^+ < \infty) = 1$;

3) *уходящий в* $+\infty$: $\mathsf{P}(\zeta^+ = \infty) = 1$, $\mathsf{P}(\zeta^- > -\infty) = 1$.

Случайные блуждания, уходящие в $-\infty$ или $+\infty$, очевидным образом невозвратны.

Среди осциллирующих случайных блужданий имеются как возвратные, так и невозвратные. Например, случайное блуждание с распределением Коши скачков — возвратное и осциллирующее; случайное блуждание с симметричным устойчивым распределением скачков с параметром $\alpha < 1$ — невозвратное и осциллирующее.

Факторизационные тождества (§ 7.5) дают следующий критерий уходящих в $-\infty$ случайных блужданий.

Теорема 6. *Для того чтобы* $\mathsf{P}(\sup\limits_{n \geqslant 0} \zeta_n < \infty) = 1$, *необходимо и достаточно, чтобы выполнялось условие*

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n > 0) < \infty. \tag{2.6}$$

При этом

$$\mathsf{P}(\sup_{n \geqslant 0} \zeta_n = 0) = \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n > 0)\right\}. \tag{2.7}$$

Аналогичный критерий имеет место для случайных блужданий, уходящих в $+\infty$.

Заметим, что всегда

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n = 0) < \infty. \tag{2.8}$$

Поэтому условие (2.6) эквивалентно также одному из следующих условий:

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n \geqslant 0) < \infty, \quad \mathsf{P}(\sup_{n \geqslant 1} \zeta_n < 0) > 0. \tag{2.9}$$

7.3. Функционалы на случайном блуждании

7.3.1. Лестничные величины. Строгие верхние лестничные точки случайного блуждания ζ_n $(n \geqslant 0)$ определяются соотношениями

$$\tau_+^{(1)} = \min\{n \geqslant 1 \colon \zeta_n > 0\}, \tag{3.1}$$

$$\gamma_+^{(1)} = \zeta_{\tau_+^{(1)}}. \tag{3.1'}$$

Случайная величина $\tau_+^{(1)}$ есть *момент первого вхождения* случайного блуждания ζ_n $(n \geqslant 0)$ на полуось $(0, +\infty)$, а $\gamma_+^{(1)}$ — положение случайного блуждания в момент первого вхождения на полуось $(0, +\infty)$, или, иначе, *величина первого перескока* нулевого уровня.

Случайное блуждание $\zeta_n^{(1)} = \zeta_{\tau_+^{(1)}+n} - \zeta_{\tau_+^{(1)}}$ $(n \geqslant 0)$ *стохастически эквивалентно* случайному блужданию ζ_n $(n \geqslant 0)$. Строгие верхние лестничные точки $\tau_+^{(2)} = \min\left\{n \geqslant 1\colon\ \zeta_n^{(1)} > 0\right\}$ и $\gamma_+^{(2)} = \zeta^{(1)}_{\tau_+^{(2)}}$ блуждания $\zeta_n^{(1)}$ $(n \geqslant 0)$ являются вторыми лестничными точками блуждания ζ_n $(n \geqslant 0)$. При этом пары случайных величин $\left(\tau_+^{(1)}, \gamma_+^{(1)}\right)$ и $\left(\tau_+^{(2)}, \gamma_+^{(2)}\right)$ независимы и одинаково распределены. Аналогично могут быть определены $\left(\tau_+^{(k)}, \gamma_+^{(k)}\right)$ при каждом целом $k \geqslant 1$.

Последовательности $\left\{\tau_+^{(k)}, k \geqslant 1\right\}$ и $\left\{\gamma_+^{(k)}, k \geqslant 1\right\}$ определяют процессы *восстановления, вложенные в случайное блуждание* ζ_n $(n \geqslant 0)$.

Вложенные процессы восстановления обрываются для случайных блужданий, уходящих в $-\infty$. Вероятность обрыва процесса на конечном шаге равна дефекту величины τ_+:

$$\mathsf{P}(\tau_+ = \infty) = \mathsf{P}\left(\sup_{n \geqslant 0} \zeta_n = 0\right) = \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n > 0)\right\}. \tag{3.2}$$

Если $\mathsf{P}\left(\sup\limits_{n \geqslant 0} \zeta_n = \infty\right) = 1$, то τ_+ — собственная случайная величина и

$$\mathsf{M}\tau_+ = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n \leqslant 0)\right\}. \tag{3.3}$$

Аналогичные соотношения имеют место для *строгих нижних лестничных моментов*

$$\tau_- = \min\{n \geqslant 1\colon\ \zeta_n < 0\}, \tag{3.4}$$

а именно:

$$\mathsf{P}(\tau_- = \infty) = \mathsf{P}\left(\inf_{n \geqslant 0} \zeta_n = 0\right) = \exp\left\{-\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n < 0)\right\} \tag{3.5}$$

и

$$\mathsf{M}\tau_- = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n \geqslant 0)\right\}. \tag{3.6}$$

Таким образом, имеют место соотношения

$$\begin{aligned} \mathsf{M}\tau_+ &= \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n = 0)\right\} \Big/ \mathsf{P}(\tau_- = \infty), \\ \mathsf{M}\tau_- &= \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n} \mathsf{P}(\zeta_n = 0)\right\} \Big/ \mathsf{P}(\tau_+ = \infty). \end{aligned} \tag{3.7}$$

Имеет место альтернатива: один из лестничных моментов τ_+ или τ_- — несобственная случайная величина, блуждание уходит в $-\infty$ или $+\infty$, и тогда $\mathsf{M}\tau_- < \infty$ или $\mathsf{M}\tau_+ < \infty$; или $\mathsf{M}\tau_+ = \mathsf{M}\tau_- = \infty$ (для осциллирующих блужданий).

Для лестничных величин τ_+ и γ_+ при $\mathsf{M}\tau_+ < \infty$ имеет место *тождество Вальда*:

$$\mathsf{M}\,\tau_+ = \mathsf{M}\gamma_+\mathsf{M}\xi_1, \tag{3.8}$$

при $0 \leqslant \mathsf{M}\xi_1 < \infty$.

Пример 1. *Случайное блуждание в схеме Бернулли* определяется скачками $\xi_k = \pm 1$ с вероятностями p и $q = 1 - p$ соответственно. При $p > q$ бернуллиевское случайное блуждание уходит в $+\infty$ и $\mathsf{M}\tau_+ = 1/(p-q)$, $\mathsf{P}(\tau_- = \infty) = 1 - q/p$; при $p < q$ блуждание уходит в $-\infty$ и $\mathsf{M}\tau_- = 1/(q-p)$, $\mathsf{P}(\tau_+ = \infty) = 1 - p/q$; при $p = q = 1/2$ случайное блуждание в схеме Бернулли — осциллирующее и возвратное с $\mathsf{M}\tau_+ = \mathsf{M}\tau_- = \infty$.

7.3.2. Верхние функционалы. Верхние функционалы от случайного блуждания ζ_n $(n \geqslant 0)$ определяются соотношениями

$$\bar{\zeta}_n = \max_{0 \leqslant k \leqslant n} \zeta_k, \quad n \geqslant 0, \quad \bar{\zeta}_0 = 0, \tag{3.9}$$

$$\theta_n = \min\{k\colon \zeta_k = \bar{\zeta}_n\}, \quad n \geqslant 0, \quad \theta_0 = 0, \tag{3.10}$$

$$\nu_n = \sum_{k=1}^{n} \chi(\zeta_k > 0), \quad n \geqslant 0, \quad \nu_0 = 0. \tag{3.11}$$

Теорема 1. *Имеют место следующие соотношения*:

$$\sum_{n=1}^{\infty} z^n \mathsf{P}(\bar{\zeta}_n = 0) = \exp\left\{\sum_{n=1}^{\infty} \frac{z^n}{n} \mathsf{P}(\zeta_n \leqslant 0)\right\}, \tag{3.12}$$

$$\sum_{n=1}^{\infty} z^n \mathsf{P}(\nu_n = n) = \exp\left\{\sum_{n=1}^{\infty} \frac{z^n}{n} \mathsf{P}(\zeta_n > 0)\right\}, \tag{3.13}$$

$$\mathsf{P}(\theta_n = k) = \mathsf{P}(\theta_k = k)\,\mathsf{P}(\bar{\zeta}_{n-k} = 0). \tag{3.14}$$

Кроме того, случайные величины θ_n (*номер первого максимума*) и ν_n (*число положительных членов последовательности*) ζ_k $(1 \leqslant k \leqslant n)$ одинаково распределены.

Для симметричного и непрерывного распределения скачков ξ_k при всех n: $\mathsf{P}(\zeta_n > 0) = \mathsf{P}(\zeta_n \geqslant 0) = 1/2$, распределения величин τ_+, $\bar{\zeta}_n$ и θ_n не зависят от распределения ξ_k:

$$\mathsf{M}z^{\tau_+} = 1 - \sqrt{1-z}, \quad \mathsf{P}(\tau_+ = n) = \frac{\mathsf{P}(\bar{\zeta}_n = 0)}{2n-1}, \tag{3.15}$$

$$\mathsf{P}(\theta_n = n) = \mathsf{P}(\bar{\zeta}_n = 0) = \frac{(2n)!}{2^{2n}(n!)^2}. \tag{3.16}$$

С помощью формулы Стирлинга определяется *асимптотическое поведение вероятностей* (3.15) и (3.16):

$$\mathsf{P}(\theta_n = n) = \mathsf{P}(\bar{\zeta}_n = 0) \sim 1/\sqrt{\pi n} \tag{3.17}$$

и

$$\mathsf{P}(\theta_n = k) \sim 1/\left(\pi\sqrt{k(n-k)}\right). \tag{3.18}$$

Последнее соотношение называется *локальным законом арксинуса*. В общем случае имеет место следующий закон арксинуса.

Теорема 2. *При условии сходимости ряда*

$$\left|\sum_{n=1}^{\infty}\frac{1}{n}\left[\frac{1}{2}-\mathsf{P}(\zeta_n>0)\right]\right|<\infty \tag{3.19}$$

$$\lim_{n\to\infty}\mathsf{P}(\theta_n < xn) = \lim_{n\to\infty}\mathsf{P}(\nu_n < xn) = \frac{2}{\pi}\arcsin\sqrt{x}. \tag{3.20}$$

Условие (3.19) выполняется очевидным образом для симметричных случайных величин ξ_k, а также при $\mathsf{M}\xi_k = 0$ и $\mathsf{D}\xi_k < \infty$.

Пример 2. В теории систем обслуживания важную роль играет *процесс ожидания* W_n $(n \geqslant 0)$, который задается соотношением

$$W_{n+1} = \max(0,\ W_n + \xi_n), \qquad n \geqslant 0,$$

при заданном значении (или распределении) W_0 и заданной последовательности случайных величин ξ_n, $n \geqslant 0$.

Если случайные величины последовательности ξ_n $(n \geqslant 0)$ независимы и одинаково распределены и $W_0 = \xi_0 = 0$, то случайные величины W_n и $\bar{\zeta}_n = \max \sum_{k=1}^{n} \xi_k$ имеют одинаковое распределение.

7.4. «Задача о разорении» для полунепрерывных случайных блужданий

Различные практические ситуации приводят к «задачам о разорении», которые в терминах случайных блужданий ζ_n $(n \geqslant 0)$ формулируются следующим образом. Имеется два поглощающих экрана: верхний в точке $x > 0$ и нижний в точке $-y$ $(y > 0)$; $T = x + y$.

Определим *моменты выхода* случайного блуждания ζ_n $(n \geqslant 0)$, $\zeta_0 = 0$ из отрезка $[-y,\ x]$ (*моменты разорения*) через нижний и верхний уровни:

$$\tau_y = \min\{n\colon \zeta_n \leqslant -y\}, \quad \tau^x = \min\{n\colon \zeta_n \geqslant x\}. \tag{4.1}$$

Вероятности выхода случайного блуждания ζ_n $(n \geqslant 0)$ (вероятности разорения) через нижний и верхний уровни определяются следующим образом:

$$\begin{aligned} Q_T(x) &= \mathsf{P}\{\zeta_n < x,\ 0 \leqslant n \leqslant \tau_{T-x}\}, \\ Q^T(x) &= \mathsf{P}\{\zeta_n > x - T,\ 0 \leqslant n \leqslant \tau^x\}. \end{aligned} \tag{4.2}$$

Производящие функции моментов разорения определяются соотношениями

$$\begin{aligned} B_T(x, z) &= \mathsf{M}\left[z^{\tau_{T-x}}\chi(\zeta_n < x,\ 0 \leqslant n \leqslant \tau_{T-x})\right], \\ B^T(x, z) &= \mathsf{M}\left[z^{\tau^x}\chi(\zeta_n > x - T,\ 0 \leqslant n \leqslant \tau^x)\right]. \end{aligned} \tag{4.3}$$

Пусть распределение величины скачков ξ решетчатое и *полунепрерывное снизу*, т. е. $\mathsf{P}(\xi \geqslant -1) = 1$ и $\mathsf{M}z^{\xi} = p(z)$.

Потенциал $\{R_k, k \geqslant 0\}$ на полуоси $k \geqslant 0$ случайного блуждания ζ_n $(n \geqslant 0)$ с производящей функцией $p(z)$ скачков задается соотношением

$$r(z) = \sum_{k=0}^{\infty} z^k R_k = [p(z) - 1]^{-1}. \tag{4.4}$$

Резольвента $\{R_k(\lambda), k \geqslant 0\}$ на полуоси $k \geqslant 0$ случайного блуждания ζ_n $(n \geqslant 0)$ с производящей функцией скачков $p(z)$ задается соотношением

$$r_\lambda(z) = \sum_{k=0}^{\infty} z^k R_k(\lambda) = [\lambda p(z) - 1]^{-1}. \tag{4.5}$$

Теорема. *Для вероятностей разорения* (4.2) *имеет место формула*

$$Q_T(x) = 1 - Q^T(x) = R_x/R_T, \quad 0 \leqslant x \leqslant T. \tag{4.6}$$

Для производящих функций моментов разорения (4.3) *имеют место формулы*

$$B_T(x, \lambda) = R_x(\lambda)/R_T(\lambda), \tag{4.7}$$

$$B^T(x, \lambda) = \left[1 - (\lambda - 1)\sum_{k=1}^{T} R_k(\lambda)\right] R_x(\lambda)/R_T(\lambda) + (\lambda - 1)\sum_{k=1}^{x} R_k(\lambda). \tag{4.8}$$

Заметим, что при $\mathsf{M}\xi < 0$ потенциал может быть задан распределением максимума $\zeta = \max\limits_{n \geqslant 0} \zeta_n$:

$$R_k = \mathsf{P}(\max_{n \geqslant 0} \zeta_n < k)/(\mathsf{P}(\max_{n \geqslant 0} \zeta_n = 0)\,\mathsf{P}(\xi = -1)). \tag{4.9}$$

7.5. Факторизационные тождества

7.5.1. Основные факторизационные тождества. Пусть задана последовательность ξ_k $(k \geqslant 1)$ независимых одинаково распределенных величин с характеристической функцией $\varphi(\lambda) = \mathsf{M}\exp\{i\lambda\xi_k\}$.

Последовательность сумм ζ_n $(n \geqslant 0)$, $\zeta_0 = 0$, определяет случайное блуждание на числовой оси.

В *теории случайных блужданий* важную роль играют факторизационные тождества для функции $1 - z\varphi(\lambda)$ вида

$$1 - z\varphi(\lambda) = \psi_+(z, \lambda)\,\psi_-(z, \lambda), \quad \operatorname{Im}\lambda = 0, \tag{5.1}$$

в котором множители факторизации $\psi_\pm(z, \lambda)$ аналитичны в областях $\operatorname{Im}\lambda > 0$ и $\operatorname{Im}\lambda < 0$, непрерывны и ограничены в замкнутых полуплоскостях $\operatorname{Im}\lambda \geqslant 0$ и $\operatorname{Im}\lambda \leqslant 0$ соответственно.

Функция $\psi_+(z, \lambda)$ $(\psi_-(z, \lambda))$ называется *положительной* (*отрицательной*) *компонентой факторизации*.

Задача факторизации, т. е. представления х. ф. в виде (5.1), является одним из вариантов проблемы Коши — Римана в теории граничных задач для аналитических функций.

Основное факторизационное тождество легко следует из разложения

$$1 - z\varphi(\lambda) = \exp\{\ln(1 - z\varphi(\lambda))\} = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k}\varphi^k(\lambda)\right\}. \qquad (5.2)$$

Теорема 1. *При* $|z| < 1$, $\operatorname{Im}\lambda = 0$ *функция* $1 - z\varphi(\lambda)$ *представима в виде*

$$1 - z\varphi(\lambda) = \varphi_+(z, \lambda)\,\varphi_-(z, \lambda)\,\varphi_0(z), \qquad (5.3)$$

где

$$\varphi_+(z, \lambda) = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k} \mathsf{M}\left[e^{i\lambda\zeta_k}\chi(\zeta_k > 0)\right]\right\},$$

$$\varphi_-(z, \lambda) = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k} \mathsf{M}\left[e^{i\lambda\zeta_k}\chi(\zeta_k < 0)\right]\right\}, \qquad (5.4)$$

$$\varphi_0(z) = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k} \mathsf{P}(\zeta_k = 0)\right\}.$$

Функции $\varphi_+(z, \lambda)$ и $\varphi_-(z, \lambda)$ являются соответственно *положительной* и *отрицательной компонентами факторизации* и удовлетворяют дополнительно следующим условиям:

$$\inf_{\operatorname{Im}\lambda \gtrless 0} |\varphi_\pm(z, \lambda)| > 0, \quad \varphi_\pm(z, \pm i\infty) = 1. \qquad (5.5)$$

Компоненты факторизации $\varphi_\pm(z, \lambda)$ и $\varphi_0(z)$ имеют вероятностную интерпретацию в терминах *граничных функционалов* от случайного блуждания ζ_n $(n \geqslant 0)$.

Строгие лестничные моменты вводятся следующим образом:

$$\tau_+ = \min\{k \geqslant 1: \zeta_k > 0\}, \quad \tau_- = \min\{k \geqslant 1: \zeta_k < 0\}. \qquad (5.6)$$

Лестничные величины определяются соотношениями

$$\gamma_+ = \zeta_{\tau_+}, \quad \gamma_- = \zeta_{\tau_-}. \qquad (5.7)$$

Величина τ_+ называется *временем первого достижения* (*вхождения*) *положительной полуоси* $(0, +\infty)$.

Аналогично τ_- есть *время первого достижения отрицательной полуоси* $(-\infty, 0)$.

Лестничная величина γ_+ (γ_-) называется *первой положительной* (*отрицательной*) *суммой* или *точкой вхождения* на полуось $(0, +\infty)$ $((-\infty, 0))$.

Заметим, что случайная величина τ_+ определена только на тех последовательностях сумм ζ_n $(n \geqslant 1)$, для которых $\bar{\zeta} = \sup_{n \geqslant 1} \zeta_n > 0$. В противном случае, когда $\bar{\zeta} \leqslant 0$, полагаем $\tau_+ = \infty$.

Аналогично τ_- — несобственная случайная величина, если $\mathsf{P}(\underline{\zeta} = \inf_{n \geqslant 1} \zeta_n \geqslant 0) > 0$.

Теорема 2. *При* $|z| \leqslant 1$, $\operatorname{Im}\lambda = 0$ *справедливо факторизационное тождество*

$$1 - z\varphi(\lambda) =$$
$$= \left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_+} z^{\tau_+} \chi(\tau_+ < \infty)\right]\right\} \left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_-} z^{\tau_-} \chi(\tau_- < \infty)\right]\right\} \varphi_0(z). \tag{5.8}$$

Функция $\varphi_0(z)$ *определена в* (5.4).

Определим лестничные величины, порожденные случайным блужданием ζ_n, $n \geqslant 0$, соотношениями

$$\tau_+^0 = \min\{k \geqslant 1 \colon \zeta_k \geqslant 0\},$$
$$\tau_-^0 = \min\{k \geqslant 1 \colon \zeta_k \leqslant 0\}, \tag{5.9}$$

$$\gamma_\pm^0 = \zeta_{\tau_\pm^0}. \tag{5.10}$$

Теорема 3. *При* $|z| \leqslant 1$ *имеют место равенства*

$$1 - \mathsf{M}\left[e^{i\lambda\gamma_+^0} z^{\tau_+^0} \chi\left(\tau_+^0 < \infty\right)\right] = \varphi_0(z) \left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_+} z^{\tau_+} \chi(\tau_+ < \infty)\right]\right\}, \tag{5.11}$$

$$1 - \mathsf{M}\left[e^{i\lambda\gamma_-^0} z^{\tau_-^0} \chi\left(\tau_-^0 < \infty\right)\right] = \varphi_0(z) \left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_-} z^{\tau_-} \chi(\tau_- < \infty)\right]\right\} \tag{5.12}$$

При этом

$$\varphi_0(z) = 1 - \mathsf{M}\left[z^{\tau_+^0} \chi\left(\gamma_+^0 = 0\right)\right] = 1 - \mathsf{M}\left[z^{\tau_-^0} \chi\left(\gamma_-^0 = 0\right)\right]. \tag{5.13}$$

Теоремы 1—3 позволяют получить целый ряд соотношений для граничных функционалов от последовательности сумм ζ_n ($n \geqslant 0$).

Следствие 1. *При* $|z| \leqslant 1$

$$1 - \mathsf{M}\left[e^{i\lambda\gamma_+} z^{\tau_+} \chi(\tau_+ < \infty)\right] = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k} \mathsf{M}\left[e^{i\lambda\zeta_k} \chi(\zeta_k > 0)\right]\right\}, \tag{5.14}$$

$$1 - \mathsf{M}\left[e^{i\lambda\gamma_-} z^{\tau_-} \chi(\tau_- < \infty)\right] = \exp\left\{-\sum_{k=1}^{\infty} \frac{z^k}{k} \mathsf{M}\left[e^{i\lambda\zeta_k} \chi(\zeta_k < 0)\right]\right\}. \tag{5.15}$$

Следствие 2. *При* $\operatorname{Im}\lambda = 0$

$$1 - \varphi(\lambda) = \left[1 - \mathsf{M}\left(e^{i\lambda\gamma_+} \chi(\tau_+ < \infty)\right)\right] \left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_-^0} \chi\left(\tau_-^0 < \infty\right)\right]\right\}. \tag{5.16}$$

Распределение максимума $\zeta = \max\limits_{n \geqslant 0} \zeta_n$ последовательности сумм ζ_n ($n \geqslant 0$) также определяется компонентами факторизации.

Теорема 4. *Если* $p = \mathsf{P}(\tau_+ < \infty) < 1$, *то при* $\operatorname{Im}\lambda \geqslant 0$

$$\mathsf{M}e^{i\lambda\zeta} = (1-p)\left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_+}\chi(\tau_+ < \infty)\right]\right\}^{-1}, \tag{5.17}$$

или

$$\mathsf{M}e^{i\lambda\zeta} = (1-p_0)\left\{1 - \mathsf{M}\left[e^{i\lambda\gamma_+^0}\chi\left(\tau_+^0 < \infty\right)\right]\right\}^{-1}. \tag{5.18}$$

Производящая функция распределения максимумов $\overline{\zeta}_n = \max(0, \zeta_1, \zeta_2, \ldots, \zeta_n)$ определяется следующим образом.

Тождество Поллачека — Спицера. При $|z| < 1$, $\operatorname{Im}\lambda \geqslant 0$

$$\sum_{n=0}^{\infty} z^n \mathsf{M}e^{i\lambda\overline{\zeta}_n} = \frac{1}{1-z}\exp\left\{\sum_{k=1}^{\infty}\frac{z^k}{k}\mathsf{M}\left[\left(e^{i\lambda\zeta_n}-1\right)\chi(\zeta_n > 0)\right]\right\} =$$

$$= \exp\left\{\sum_{k=1}^{\infty}\frac{z^k}{k}\mathsf{M}\exp\{i\lambda\max(0, \zeta_n)\}\right\}. \tag{5.19}$$

7.5.2. Примеры явных формул. Существует класс распределений, для которых компоненты факторизационных тождеств имеют явные аналитические выражения.

1. *Распределения лестничных величин* γ_+ *и* γ_-^0 *и максимума* $\zeta = \max\limits_{n\geqslant 0}\zeta_n$ имеют простые явные аналитические выражения в одном важном для приложений случае (в теории систем обслуживания, задачах о разорении и др.) — случае *полунепрерывных распределений* величины скачков, когда, например, правый хвост распределения скачков показательный:

$$1 - \Phi(x) = \mathsf{P}(\xi \geqslant x) = Ce^{-ax} \quad (C > 0,\ a > 0), \quad x > 0. \tag{5.20}$$

При $x < 0$ распределение $\mathsf{P}(\xi < x) = \Phi(x)$ произвольно. В этом случае величины γ_+ и τ_+ независимы и

$$\mathsf{P}(\gamma_+ < x) = p\left(1 - e^{-ax}\right), \quad p = \mathsf{P}(\tau_+ < \infty). \tag{5.21}$$

Кроме того,

$$\mathsf{P}(\zeta < x) = 1 - pe^{-a(1-p)x}, \quad 1 - p = \mathsf{P}(\zeta = 0). \tag{5.22}$$

Если существует $\mathsf{M}\xi = 0$, то $p = \mathsf{P}(\tau_+ < \infty) = 1$, но $\mathsf{P}(\tau_- < \infty) = 1 - a\mathsf{M}\xi < 1$, и тогда

$$P(x) = \mathsf{P}\left(\gamma_-^0 \leqslant x\right) = \Phi(x) + a\int_{-\infty}^{x}\Phi(y)\,dy, \quad x \leqslant 0, \tag{5.23}$$

и имеет место *формула Хинчина — Поллачека*

$$\mathsf{P}\left(\min_{n\geqslant 0}\zeta_n \leqslant x\right) = (1 - a\mathsf{M}\xi)\sum_{n=0}^{\infty}P^{n*}(x), \quad x \leqslant 0. \tag{5.24}$$

Если $\mathsf{M}\xi < 0$, то $p = \mathsf{P}(\tau_+ < \infty) < 1$ (но $\mathsf{P}(\tau_- < \infty) = 1$). Константа p в этом случае определяется равенством

$$p = 1 - \mu_0/a, \tag{5.25}$$

где μ_0 — единственный положительный корень уравнения $\varphi(-i\mu_0) = 1$.

Распределение γ_-^0 определяется формулой (5.23).

Приведенные в примере формулы также имеют место, когда величина скачков представима в виде разности двух независимых неотрицательных случайных величин $\xi = \xi_+ - \xi_-$ с $\mathsf{P}(\xi_+ \geqslant x) = e^{-ax}$ $(x \geqslant 0)$ и $\mathsf{P}(\xi_- < x) = \Phi_-(x)$ $(x \geqslant 0)$. В этом случае при $x > 0$

$$\Phi(x) = \mathsf{P}(\xi < x) = 1 - Ce^{-ax}, \quad C = \int\limits_0^\infty e^{-ax}\, d\Phi_-(x), \tag{5.26}$$

а при $x \leqslant 0$

$$\Phi(x) = \mathsf{P}(\xi < x) = 1 - \Phi_-(-x) - e^{-ax}\int\limits_{-x}^\infty e^{-ay}\, d\Phi_-(y). \tag{5.27}$$

При этом лестничная величина γ_-^0 имеет плотность

$$\frac{d}{dx}\mathsf{P}\left(\gamma_-^0 < x\right) = a\Phi_-(x).$$

Аналогичные результаты имеют место для решетчатых распределений с геометрическим распределением положительных скачков:

$$\mathsf{P}(\xi \geqslant k) = Cq^{k-1} \quad (C > 0,\ q \geqslant 0), \qquad k = 1, 2, \ldots \tag{5.28}$$

(при $q = 0$ полагаем $q^0 = 1$).

2. *Показательное случайное блуждание* задается двусторонним показательным распределением величины скачков:

$$\varphi(\lambda) = \mathsf{M}e^{i\lambda\xi} = \frac{b}{b + i\lambda}\,\frac{a}{a - i\lambda}. \tag{5.29}$$

В этом случае явные выражения для компонент факторизации $\varphi_\pm(z, \lambda)$ (см. теорему 1) определяются формулами (при $b \leqslant a$)

$$\varphi_+(z, \lambda) = 1 - \frac{u(z)}{a - i\lambda}, \quad \varphi_-(z, \lambda) = 1 - \frac{u(z)}{b + i\lambda}, \tag{5.30}$$

$$2u(z) = a + b - \sqrt{(a+b)^2 - 4abz}. \tag{5.31}$$

Из теоремы 2 следуют выражения для производящих функций лестничных моментов

$$\begin{aligned} &\mathsf{M}\left[z^{\tau_+}\left[\chi\left(\tau_+^0 < \infty\right)\right]\right] = u(z)/a, \\ &\mathsf{M}\left[z^{\tau_-}\left[\chi\left(\tau_-^0 < \infty\right)\right]\right] = u(z)/b. \end{aligned} \tag{5.32}$$

При $b < a$ дефект случайной величины τ_+ равен $1 - b/a = \mathsf{P}(\tau_+ = \infty)$.

Распределения лестничных величин $\gamma_\pm$ являются показательными с плотностями ae^{-ax} и be^{-bx} соответственно.

Распределение максимума (при $b < a$) также показательное:

$$\mathsf{P}\left(\max_{n\geqslant 0}\zeta_n \leqslant x\right) = 1 - \frac{b}{a}e^{-(a-b)x}. \tag{5.33}$$

3. *Биномиальное случайное блуждание* определяется биномиальным распределением величины скачков: $\mathsf{P}(\xi_k = +1) = p$, $\mathsf{P}(\xi_k = -1) = q$, $p + q = 1$, $\varphi(\lambda) = \mathsf{M}e^{i\lambda\xi_k} = pe^{i\lambda} + qe^{-i\lambda}$.

Компоненты факторизации определяются формулами

$$\varphi_+(z, \lambda) = 1 - e^{i\lambda}u_+(z), \quad \varphi_-(z, \lambda) = 1 - e^{i\lambda}u_-(z), \tag{5.34}$$

$$u_+(z) = \left(1 - \sqrt{1 - 4pqz^2}\right)/2qz, \quad u_-(z) = \left(1 - \sqrt{1 - 4qpz^2}\right)/2pz, \tag{5.35}$$

$$\varphi_0(z) = \left(1 + \sqrt{1 - 4pqz^2}\right)/2. \tag{5.36}$$

7.5.3. Граничные функционалы. *Верхние граничные функционалы* на случайном блуждании ζ_n ($n \geqslant 0$), связанные с достижением положительного уровня, определяются следующим образом.

Момент первого достижения положительного уровня $x > 0$:

$$\tau_x = \min\{k \geqslant 1:\ \zeta_k \geqslant x\}. \tag{5.37}$$

Величина первого перескока положительного уровня $x > 0$:

$$\gamma_x = \zeta_{\tau_x} - x. \tag{5.38}$$

Теорема 5. *При* $|z| < 1$, $\operatorname{Im}\lambda \geqslant 0$, $\operatorname{Im}\mu > 0$

$$1 - \frac{\lambda - \mu}{\lambda}\int_0^\infty e^{i\lambda x}d_x\mathsf{M}\left[z^{\tau_x}e^{i\mu\gamma_x}\chi(\tau_x < \infty)\right] = \frac{\varphi_+(z, \mu)}{\varphi_+(z, \lambda)}. \tag{5.39}$$

При этом $\tau_{+0} = \tau_+$, $\gamma_{+0} = \gamma_+$.

Если характеристическая функция $\varphi(\lambda) = \mathsf{M}e^{i\lambda\xi_k}$ величины скачков аналитична в некоторой полосе $-\lambda_0 < \operatorname{Im}\lambda < 0$ и $\varphi(-i\lambda_0) > 1$, то имеет место *тождество Вальда*: при $x > 0$

$$\mathsf{M}\left[z^{\tau_x}e^{\lambda(z)\gamma_x}\chi(\tau_x < \infty)\right] = e^{-\lambda(z)x}, \tag{5.40}$$

где $\lambda(z)$ — наибольший корень уравнения

$$1 - zf(-i\lambda(z)) = 0. \tag{5.41}$$

Тождество Вальда имеет место в более общей ситуации для момента выхода случайного блуждания ζ_n ($n \geqslant 0$) из конечного интервала $(-a, b)$: $\tau_b^a = \min\{n:\ \zeta_n \notin (-a, b)\}$ и положения точки в момент выхода $\zeta_{\tau_b^a}$ в следующей форме:

$$\mathsf{M}\left\{[f(\lambda)]^{-\tau_b^a}\exp\left\{-\lambda\zeta_{\tau_b^a}\right\}\right\} = 1.$$

Это тождество используется в последовательном анализе для оценки распределения τ_b^a.

Для полунепрерывного распределения величины скачков, имеющего показательный или геометрический хвост $\mathsf{P}(\xi \geqslant x) = Ce^{-ax}$ ($C > 0$, $a > 0$, $x > 0$) или (в решетчатом случае) $\mathsf{P}(\xi \geqslant k) = Cq^{k-1}$ ($C > 0$, $q \geqslant 0$, $k = 1, 2, \ldots$), случайные величины τ_x и γ_x независимы и

$$\mathsf{P}(\gamma_x \geqslant t \mid \tau_x < \infty) = e^{-at} \tag{5.42}$$

или (в решетчатом случае)

$$\mathsf{P}(\gamma_x \geqslant k \mid \tau_x < \infty) = q^k. \tag{5.43}$$

Из тождества Вальда (5.40) следует

$$\mathsf{M}\left[z^{\tau_x}\chi(\tau_x < \infty)\right] = \frac{a - \lambda(z)}{a} e^{-\lambda(z)x} \tag{5.44}$$

или (в решетчатом случае)

$$\mathsf{M}\left[z^{\tau_x}\chi(\tau_x < \infty)\right] = \frac{1 - qe^{\lambda(z)}}{1 - q} e^{-\lambda(z)x}. \tag{5.45}$$

Из формул (5.44), (5.45) следует

$$\mathsf{P}(\tau_x = n) = \frac{x}{n} p_n(x) + \frac{1}{a}\frac{d}{dx}\left[\frac{x}{n} p_n(x)\right], \tag{5.46}$$

или (в решетчатом случае)

$$\mathsf{P}(\tau_x = n) = \\ = \frac{x}{n}\mathsf{P}(\zeta_n = x) + \frac{q}{1-q}\left[\frac{x}{n}\mathsf{P}(\zeta_n = x) - \frac{x-1}{n}\mathsf{P}(\zeta_n = x - 1)\right]. \tag{5.47}$$

Здесь $p_n(x) = \frac{d}{dx}\mathsf{P}(\zeta_n < x)$ — плотность распределения суммы ζ_n при $x > 0$.

Для полунепрерывного решетчатого случайного блуждания с $q = 0$, т. е. с $\mathsf{P}(\xi > 1) = 0$,

$$\mathsf{P}(\tau_x = n) = \frac{x}{n}\mathsf{P}(\zeta_n = x). \tag{5.48}$$

Определим момент 1-го достижения максимума

$$\theta_n = \min\{k\colon \zeta_k = \bar{\zeta}_n\}$$

и случайную величину ν_p с геометрическим распределением

$$\mathsf{P}\{\nu_p = k\} = (1 - p)p^k, \quad k \geqslant 0, \quad 0 < p < 1.$$

Теорема 6. *При* $|z| < 1$, $\operatorname{Im}\lambda = \operatorname{Im}\mu = 0$

$$\mathsf{M}\left[\exp\left\{i\lambda(\zeta_{\nu_p} - \bar{\zeta}_{\nu_p}) + i\mu\bar{\zeta}_{\nu_p}\right\} z^{\theta_{\nu_p}}\right] = \\ = \varphi_+(zp, \mu)\,\mathsf{M}\left[\exp\left\{i\lambda(\zeta_{\nu_p} - \bar{\zeta}_{\nu_p})\right\}\right]\mathsf{M}z^{\theta_{\nu_p}},$$

причем $\mathsf{M}z^{\theta_{\nu_p}} = \lim\limits_{i\lambda \to -\infty} \varphi_+(p, \lambda)/\varphi_+(zp, \lambda)$.

7.5.4. Время пребывания над положительным уровнем

$$Q_x(n) = \sum_{k=1}^{n} \chi(\zeta_k > x) \quad (x > 0).$$

Теорема 7. *Пусть* $D_n(z, x) = \mathrm{M} z^{Q_x(n)}$, $D(u, z, x) = \sum_{n=1}^{\infty} u^n D_n(z, x)$. *Тогда*

$$\int_{+0}^{\infty} e^{i\lambda x} d_x D(u, z, x) = \frac{1}{1-u} \cdot \frac{\varphi_+(u, \lambda)}{\varphi_+(uz, \lambda)} - 1 - D(u, z, +0),$$

$$D(u, z, +0) = \frac{1}{1-u} \lim_{i\lambda \to -\infty} \frac{\varphi_+(u, \lambda)}{\varphi_+(uz, \lambda)} - 1.$$

Если выполнено условие (5.20), то $\varphi_+(u, \lambda) = \dfrac{p_+(u)(a - i\lambda)}{ap_+(u) - i\lambda}$; следовательно,

$$D(u, z, x) = \frac{u}{1-u} - \left(\frac{u}{1-u} - D(u, z, +0)\right) e^{-xp_+(u)a},$$

$$D(u, z, +0) = \frac{1}{1-u} \cdot \frac{p_+(u)}{p_+(uz)} - 1, \quad p_+(u) = \mathrm{P}\{\bar{\zeta}_{\nu_u} = 0\}.$$

Литература: [11, 12, 47, 84, 86, 87, 89].

Глава 8. ЦЕПИ МАРКОВА

8.1. Определение цепи Маркова

8.1.1. Определение цепи Маркова. Одним из наиболее важных обобщений понятия последовательности независимых случайных величин является понятие последовательности величин, связанных в цепь Маркова.

Пусть задано вероятностное пространство $(\Omega, \mathfrak{F}, \mathrm{P})$. Измеримое отображение ξ: $(\Omega, \mathfrak{F}) \to (X, \mathfrak{B})$, где $(X, \mathfrak{B})$ — некоторое измеримое пространство, называется *случайным элементом* в $(X, \mathfrak{B})$.

Последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ случайных элементов в измеримом пространстве $(X, \mathfrak{B})$ называется *цепью Маркова*, если для любых $\Gamma \in \mathfrak{B}$ и $n = 1, 2, \ldots$ с вероятностью 1

$$\mathrm{P}\{\xi_n \in \Gamma \mid \xi_0, \xi_1, \ldots, \xi_{n-1}\} = \mathrm{P}\{\xi_n \in \Gamma \mid \xi_{n-1}\}.$$

Пространство $(X, \mathfrak{B})$ называется *фазовым пространством* цепи.

Всякую последовательность (случайных или неслучайных) элементов $\{\xi_n, n = 0, 1, \ldots\}$ пространства $(X, \mathfrak{B})$ можно рассматривать как движение некоторой системы (точки, частицы) в фазовом пространстве: из начального состояния ξ_0 в момент времени 1 система переходит в состояние ξ_1, затем в момент времени 2—в состояние ξ_2 и т. д. Понятие цепи Маркова, таким образом, выделяет из совокупности всевозможных движущихся систем так называемые *системы* *без последействия*, или *системы с отсутствием памяти*. В детерминированном случае это те системы, для которых состояние в момент времени n однозначно определяется состоянием этой системы в момент времени $n-1$ независимо от того, каким было движение до этого момента. В отличие от детерминированных стохастические системы без последействия обладают тем свойством, что по состоянию системы в момент времени $n-1$ однозначно определяется не состояние системы в момент времени n, а лишь вероятность, с какой она в этот момент времени находится в том или ином множестве состояний.

Примеры. 1. Последовательность независимых случайных элементов $\{\xi_n, n = 0, 1, 2, \ldots\}$ образует цепь Маркова, так как

$$\mathsf{P}\{\xi_n \in \Gamma \mid \xi_0, \xi_1, \ldots, \xi_{n-1}\} = \mathsf{P}\{\xi_n \in \Gamma \mid \xi_{n-1}\} = \mathsf{P}\{\xi_n \in \Gamma\}.$$

2. *Случайные блуждания*. Пусть X — аддитивная коммутативная группа и $\mathfrak{B}$ — некоторая σ-алгебра подмножеств X, согласованная с операцией сложения в X, т. е. если $\Gamma \in \mathfrak{B}$, то $\Gamma + x = \{x + z, z \in \Gamma\} \in \mathfrak{B}$ при любом $x \in X$. Предположим, что задана последовательность независимых случайных элементов $\{\eta_n, n = 0, 1, 2, \ldots\}$ в $(X, \mathfrak{B})$. Тогда последовательность $\{\xi_n = \eta_0 + \eta_1 + \ldots + \eta_n, n = 0, 1, 2, \ldots\}$ является цепью Маркова в фазовом пространстве $(X, \mathfrak{B})$, так как для всех $\Gamma \in \mathfrak{B}$ и $n = 1, 2, \ldots$

$$\mathsf{P}\{\xi_n \in \Gamma \mid \xi_0, \xi_1, \ldots, \xi_{n-1}\} = \mathsf{P}\{\xi_{n-1} + \eta_n \in \Gamma \mid \xi_{n-1}\}.$$

Такая цепь называется *случайным блужданием* в X.

Пусть, например, X — совокупность всех векторов из евклидова пространства $\mathbb{R}^m$, координаты которых в некотором фиксированном базисе $e_1, e_2, \ldots, e_m$ целочисленны; $\mathfrak{B}$ — σ-алгебра всех подмножеств X. Если случайные векторы $\eta_1, \eta_2, \ldots$ со значениями в X независимы и одинаково распределены, то блуждание $\{\xi_n = \eta_0 + \eta_1 + \ldots + \eta_n, n = 0, 1, 2, \ldots\}$ называется *случайным блужданием по целочисленной решетке* в $\mathbb{R}^m$. Здесь η_0 — произвольный не зависящий от последовательности $\{\eta_n, n = 1, 2, \ldots\}$ случайный вектор, принимающий значения в X. В частности, если векторы η_k $(k = 1, 2, \ldots)$ принимают значения $\pm e_1, \pm e_2, \ldots, \pm e_m$ с вероятностями $1/(2m)$ каждый, то блуждание происходит таким образом, что частица (система) за единицу времени из данной точки переходит с равной вероятностью в одну из соседних точек (соседними к точке $x \in X$ называются точки вида $x \pm e_1, x \pm e_2, \ldots, x \pm e_m$). Такое блуждание называется *простейшим симметричным блужданием* по целочисленной решетке в $\mathbb{R}^m$.

Другой пример случайного блуждания получим, если положим $X = \mathbb{R}^m$, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств $\mathbb{R}^m$, а $\{\eta_k, k = 1, 2, \ldots\}$ — независимые одинаково распределенные случайные векторы в $\mathbb{R}^m$.

3. *Случайные блуждания с границами*. Пусть X — коммутативная аддитивная группа, а $\mathfrak{B}$ — σ-алгебра подмножеств X, согласованная с операцией сложения в X (см. пример 2). Для произвольного множества $D \in \mathfrak{B}$ обозначим через $\mathfrak{B}_D$ след σ-алгебры $\mathfrak{B}$ на множестве D, т. е. σ-алгебру множества вида $\Gamma \cap D$, $\Gamma \in \mathfrak{B}$. Предположим, что для некоторого фиксированного множества $D \in \mathfrak{B}$ задано измеримое отображение φ: $(X \setminus D, \mathfrak{B}_{X \setminus D}) \to (D, \mathfrak{B}_D)$, а для некоторого подмножества $D' \subset D$, $D' \in \mathfrak{B}$, — измеримое отображение

ψ: $(D', \mathfrak{B}_{D'}) \to (D, \mathfrak{B}_D)$. Пусть $\{\eta_n, n = 1, 2, \ldots\}$ — последовательность независимых одинаково распределенных случайных элементов в $(X, \mathfrak{B})$, а η_0 — не зависящий от последовательности $\{\eta_n, n = 1, 2, \ldots\}$ случайный элемент в $(D, \mathfrak{B}_D)$. Положим $\xi_0 = \eta_0$ и

$$\xi_{n+1} = \chi_{D'}(\xi_n)\,\psi(\xi_n) + \chi_{D\setminus D'}(\xi_n)\,[(\xi_n + \eta_{n+1})\,\chi_D(\xi_n + \eta_{n+1}) + \\ + \varphi(\xi_n + \eta_{n+1})\,\chi_{X\setminus D}(\xi_n + \eta_{n+1})], \quad n = 0, 1, 2, \ldots, \quad (1.1)$$

где $\chi_\Gamma(x)$ — индикатор множества $\Gamma \subset X$, т. е. функция, равная 1 при $x \in \Gamma$ и равная 0 при $x \notin \Gamma$. Тогда последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ образует цепь Маркова в $(D, \mathfrak{B}_D)$, называемую *случайным блужданием с отражением*.

Движение частицы при таком блуждании можно описать следующим образом. Если в некоторый момент времени частица попала в точку $x \in D'$, то в следующий момент времени она попадает в точку $\psi(x) \in D$. Если же в момент времени n $\xi_n = x \in D \setminus D'$, то в момент времени $n+1$ частица попадает в точку $x + \eta_{n+1}$ при условии, что $x + \eta_{n+1} \in D$; в противном случае она переходит в точку $\varphi(x + \eta_{n+1})$.

Рассмотрим некоторые частные случаи этой модели. а) Пусть X — целочисленная решетка на прямой, D — множество всех неотрицательных целых чисел, D' — пустое множество, $\varphi(x) = 0$ для всех $x \in X \setminus D$, а $\{\eta_n, n = 1, 2, \ldots\}$ — последовательность независимых одинаково распределенных случайных величин, принимающих значения $+1$ и -1 с вероятностями 1/2 каждая. Тогда, если η_0 — не зависящая от последовательности $\{\eta_n, n \geqslant 1\}$ неотрицательная целочисленная случайная величина, то последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$, где в соответствии с формулой (1.1) $\xi_0 = \eta_0$, $\xi_{n+1} = (\xi_n + \eta_{n+1})\chi_D(\xi_n + \eta_{n+1})$ $(n = 0, 1, 2, \ldots)$, образует цепь Маркова, называемую *простейшим блужданием с задерживающим отражающим экраном в нуле*. При таком блуждании во всех точках $x > 0$ блуждающая частица ведет себя точно так же, как и при простейшем симметричном блуждании, т. е. из x она попадает в одну из точек $x + 1$ и $x - 1$ с вероятностями 1/2. Попав в точку $x = 0$, частица находится в ней некоторое случайное время, имеющее геометрическое распределение, а затем уходит в точку $x = 1$.

б) Чтобы получить простейшее блуждание с отражением в нуле без задержки, нужно выбрать X, D, φ и $\{\eta_n, n = 0, 1, 2, \ldots\}$ такими же, как и в случае а). В качестве D' возьмем множество, состоящее из одной точки $x = 0$ $(D' = \{0\})$, и пусть $\psi(0) = 1$. Тогда $\xi_0 = \eta_0$, $\xi_{n+1} = \chi_{\{0\}}(\xi_n) + \chi_{D\setminus\{0\}}(\xi_n)(\xi_n + \eta_{n+1})$ $(n = 0, 1, 2, \ldots)$. Это означает, что, попав в точку $x = 0$, блуждающая частица следующим шагом выйдет оттуда и попадет в точку $x = 1$. Во всех остальных точках $(x > 0)$ частица ведет себя точно так же, как и в предыдущем примере.

в) Аналогично можно построить случайные блуждания и с двумя отражающими экранами. Пусть, например, a и b — два целых числа, $a < b$, и пусть D означает совокупность всех целых чисел на отрезке $[a, b]$, $D' = \{b\}$, $\psi(b) = b - 1$, $\varphi(x) = a$ для всех $x < a$ и $\varphi(x) = b$ для всех целых $x > b$; последовательность $\{\eta_n, n = 1, 2, \ldots\}$ такая же, как и в случае б); η_0 — не зависящая от последовательности $\{\eta_n, n \geqslant 1\}$ случайная величина, значениями которой служат целые числа отрезка $[a, b]$. В этом случае $\xi_0 = \eta_0$,

$\xi_{n+1} = \chi_{\{b\}}(\xi_n)(b-1) + \chi_{[a,\,b)}(\xi_n)[(\xi_n + \eta_{n+1})\chi_{[a,\,b]}(\xi_n + \eta_{n+1}) + a\chi_{\{a-1\}}(\xi_n + \eta_{n+1})]$ $(n=0, 1, 2, \ldots)$, т. е. последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ представляет собой простейшее случайное блуждание, для которого точка a является задерживающим отражающим экраном, а в точке b происходит отражение без задержки.

г) Приведем еще один пример многомерного случайного блуждания с отражением. Пусть $X = \mathbf{R}^m$; $\mathfrak{B}$ — σ-алгебра борелевских подмножеств $\mathbf{R}^m$; $e_1, e_2, \ldots, e_m$ — фиксированный базис в $\mathbf{R}^m$. Через x^i обозначим i-ю координату вектора $x \in \mathbf{R}^m$, так что $x = \sum_{i=1}^{m} x^i e_i$.

Положим $D = \{x\colon x \in \mathbf{R}^m,\ x^1 \geqslant 0\}$, $\varphi(x) = -x^1 e_1 + \sum_{i=2}^{m} x^i e_i$ для всех $x \in \mathbf{R}^m \setminus D$, $D' = \varnothing$ ($\varnothing$ — пустое множество). Предположим еще, что задана последовательность независимых одинаково распределенных случайных векторов в $\mathbf{R}^m$ $\{\eta_n, n = 1, 2, \ldots\}$ и не зависящий от этой последовательности случайный вектор $\eta_0 \in D$. Тогда, если

$$\xi_0 = \eta_0,\ \xi_{n+1} = (\xi_n + \eta_{n+1})\chi_D(\xi_n + \eta_{n+1}) + \varphi(\xi_n + \eta_{n+1}) \times \\ \times \chi_{\mathbf{R}^m \setminus D}(\xi_n + \eta_{n+1}), \qquad n = 0, 1, 2, \ldots$$

то последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ представляет собой цепь Маркова в полупространстве $(D, \mathfrak{B}_D)$. Это пример случайного блуждания, при котором отражение происходит по закону отражения светового луча на гиперплоскости $x^1 = 0$.

д) Пусть $(X, \mathfrak{B})$ — аддитивная коммутативная группа с σ-алгеброй, такая же, как и в примере 2. По заданному множеству $D_0 \in \mathfrak{B}$ и последовательности $\{\eta_n, n = 0, 1, 2, \ldots\}$ независимых случайных элементов в $(X, \mathfrak{B})$ построим новую последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ по формуле $\xi_0 = \eta_0$, $\xi_{n+1} = \xi_n\chi_{D_0}(\xi_n) + (\xi_n + \eta_{n+1})\chi_{X \setminus D_0}(\xi_n)$ $(n = 0, 1, 2, \ldots)$. Эта последовательность образует цепь Маркова в $(X, \mathfrak{B})$. Она называется *случайным блужданием с поглощением* во множестве D_0. При таком блуждании частица, оказавшаяся в момент времени n в некоторой точке $x \notin D_0$, в следующий момент времени смещается в точку $x + \eta_{n+1}$. Если же частица попала в какую-либо точку множества D_0, то там и остается навсегда.

4. *Простейшее блуждание с переменными вероятностями.* Пусть X — совокупность всех неотрицательных целых чисел, $\mathfrak{B}$ — σ-алгебра всех подмножеств X. Точки $x \in X$ будем интерпретировать как состояния некоторой системы. Предположим, что эта система в моменты времени $0, 1, 2, \ldots$ изменяет свои состояния таким образом, что, оказавшись в момент времени n в состоянии $x > 0$, в момент времени $n+1$ она попадает в одно из состояний $x-1$, x, $x+1$ с вероятностями q_x, r_x, p_x соответственно ($p_x + q_x + r_x = 1$). Если $x = 0$, то переход возможен лишь в состояния 0, 1 с вероятностями r_0 и p_0 соответственно ($r_0 + p_0 = 1$). Предположим также, что указанные переходы в момент времени $n+1$ совершаются независимо (в теоретико-вероятностном смысле) от движения системы до момента n. Это означает, что если ξ_n — состояние системы в момент

времени n, то для $\Gamma \in \mathfrak{B}$ и $n = 0, 1, 2, \ldots$

$$\mathsf{P}\{\xi_{n+1} \in \Gamma \mid \xi_0, \xi_1, \ldots, \xi_n\} = \mathsf{P}\{\xi_{n+1} \in \Gamma \mid \xi_n\}.$$

Иначе говоря, последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ образует цепь Маркова, которую назовем *простейшим блужданием с переменными вероятностями.*

Строго говоря, цепь Маркова в этом примере нельзя считать заданной, поскольку априори не известно, можно ли задать вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ и случайные величины $\xi_0, \xi_1, \xi_2, \ldots$, определенные на Ω, так, чтобы выполнялись все условия, наложенные на ξ_n. Ниже будет сформулирована общая теорема (см. теорему 2), показывающая, в частности, что последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$, определяющая простейшее блуждание с переменными вероятностями, может быть задана на некотором вероятностном пространстве.

5. *Системы массового обслуживания.* Рассмотрим некоторую систему, имеющую m мест для обслуживания. Предположим, что в случайные моменты времени (целочисленные) в систему поступают требования на обслуживание, которые при наличии свободных мест сразу же начинают удовлетворяться. Если свободных мест нет, то поступившие требования создают очередь. Каждое требование обслуживается некоторое случайное время, после чего сразу же покидает систему. Пусть выполнены следующие условия:

а) в каждый момент времени с вероятностью p может поступить лишь одно требование на обслуживание независимо от числа требований, поступивших до этого момента;

б) если некоторое требование в момент времени n обслуживается, то с вероятностью q его обслуживание может закончиться к моменту времени $n+1$ независимо от того, сколь долго оно обслуживалось до этого момента;

в) обслуживание на каждом из m мест не зависит от обслуживания на остальных местах и, кроме того, не зависит от входящего потока требований.

Обозначим через ξ_n число всех требований в рассматриваемой системе обслуживания в момент времени n (включая те, которые обслуживаются, и те, которые стоят в очереди). Тогда $\{\xi_n, n = 0, 1, 2, \ldots\}$ — цепь Маркова, для которой при всех $k \in [1, m]$

$$\mathsf{P}\{\xi_{n+1} = k - j \mid \xi_n = k\} = (1-p)\, C_k^j q^j (1-q)^{k-j} + \\ + pC_k^{j+1} q^{j+1} (1-q)^{k-j-1}, \quad j = 0, 1, 2, \ldots, k,$$

$$\mathsf{P}\{\xi_{n+1} = k+1 \mid \xi_n = k\} = p(1-q)^k.$$

Заметим, что $C_k^{k+1} = 0$, так что $\mathsf{P}\{\xi_{n+1} = 0 \mid \xi_n = k\} = (1-p)\, q^k$.

При $k = 0$

$$\mathsf{P}\{\xi_{n+1} = 0 \mid \xi_n = 0\} = 1 - p, \quad \mathsf{P}\{\xi_{n+1} = 1 \mid \xi_n = 0\} = p.$$

Наконец, при $k > m$

$$\mathsf{P}\{\xi_{n+1} = k - j \mid \xi_n = k\} = (1-p)\, C_m^j q^j (1-q)^{m-j} + \\ + pC_m^{j+1} q^{j+1} (1-q)^{m-j-1}, \quad j = 0, 1, 2, \ldots m,$$

$$\mathsf{P}\{\xi_{n+1} = k+1 \mid \xi_n = k\} = p(1-q)^m.$$

Во всех этих случаях, если $r > 1$, то

$$\mathsf{P}\{\xi_{n+1} = k + r \mid \xi_n = k\} = 0.$$

Если $m = 1$, т. е. в системе имеется лишь одно место для обслуживания, то цепь $\{\xi_n,\ n = 0, 1, 2, \ldots\}$ совпадает с цепью примера 4, для которой $r_0 = 1 - p$, $p_0 = p$, а при $x > 0$

$$p_x = p(1-q), \quad q_x = q(1-p), \quad r_x = (1-p)(1-q) + pq.$$

8.1.2. Критерий марковости. Пусть задана последовательность $\{\xi_n,\ n = 0, 1, 2, \ldots\}$ случайных элементов в измеримом пространстве $(X, \mathfrak{B})$ (вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ считается фиксированным). Обозначим через $\mathfrak{F}_n$ минимальную σ-алгебру событий, относительно которой измеримы случайные элементы $\xi_0, \xi_1, \ldots, \xi_n$, а через $\mathfrak{F}^n$ — минимальную σ-алгебру событий, относительно которой измеримы случайные элементы $\xi_n, \xi_{n+1}, \ldots$ Иначе говоря, $\mathfrak{F}_n$ — σ-алгебра всех событий, связанных с эволюцией последовательности до момента n включительно, а $\mathfrak{F}^n$—σ-алгебра всех событий, связанных с эволюцией последовательности после момента n, включая сам момент времени n. σ-алгебра $\mathfrak{F}_n$ порождается событиями вида $\{\xi_k \in \Gamma\}$ при всех $k = 0, 1, 2, \ldots, n$, $\Gamma \in \mathfrak{B}$. Аналогично σ-алгебра $\mathfrak{F}^n$ порождается событиями $\{\xi_k \in \Gamma\}$ при $k \geqslant n$, $\Gamma \in \mathfrak{B}$.

Определение цепи Маркова означает, таким образом, что для всех $n = 0, 1, 2, \ldots$ и $\Gamma \in \mathfrak{B}$ с вероятностью 1

$$\mathsf{P}\{\xi_{n+1} \in \Gamma \mid \mathfrak{F}_n\} = \mathsf{P}\{\xi_{n+1} \in \Gamma \mid \xi_n\}.$$

Теорема 1. *Пусть* $\{\xi_n,\ n = 0, 1, 2, \ldots\}$ — *последовательность случайных элементов в измеримом пространстве* $(X, \mathfrak{B})$. *Следующие утверждения эквивалентны:*

1) *последовательность* $\{\xi_n,\ n = 0, 1, \ldots\}$ *является цепью Маркова;*

2) *для всех* $n = 0, 1, 2, \ldots$, $m = 1, 2, \ldots$, $\Gamma \in \mathfrak{B}$ *с вероятностью* 1

$$\mathsf{P}\{\xi_{n+m} \in \Gamma \mid \mathfrak{F}_n\} = \mathsf{P}\{\xi_{n+m} \in \Gamma \mid \xi_n\};$$

3) *для любых* $A \in \mathfrak{F}_{n-1}$, $B \in \mathfrak{F}^{n+1}$ $(n = 1, 2, \ldots)$ *с вероятностью* 1

$$\mathsf{P}\{A \cap B \mid \xi_n\} = \mathsf{P}\{A \mid \xi_n\}\,\mathsf{P}\{B \mid \xi_n\};$$

4) *для любой ограниченной* $\mathfrak{F}^{n+1}$-*измеримой случайной величины* η *с вероятностью* 1

$$\mathsf{M}\{\eta \mid \mathfrak{F}_n\} = \mathsf{M}\{\eta \mid \xi_n\}.$$

Если считать момент времени n «настоящим», то тогда $\mathfrak{F}_{n-1}$ — это «прошлое», а $\mathfrak{F}^{n+1}$ — «будущее». Утверждение 3) означает, таким образом, что для цепи Маркова при известном «настоящем» «прошлое» и «будущее» условно независимы.

8.1.3. Уравнение Колмогорова — Чепмена. Пусть $\{\xi_n,\ n = 0, 1, 2, \ldots\}$ — цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$ и $0 \leqslant k < m < n$. Тогда в силу свойств условных вероятностей и марковского свойства имеем с вероятностью 1

$$\mathsf{P}\{\xi_n \in \Gamma \mid \xi_k\} = \mathsf{M}\{\mathsf{P}\{\xi_n \in \Gamma \mid \xi_m\} \mid \xi_k\}.$$

Это соотношение называется *уравнением Колмогорова — Чепмена* и является фактически следствием формулы полной вероятности и марковского свойства.

Рассмотрим условную вероятность $\mathsf{P}(\xi_n \in \Gamma | \xi_k)$ $(0 \leqslant k < n$, $\Gamma \in \mathfrak{B})$. При фиксированных k, n, Γ эта вероятность представляет собой $\mathfrak{B}$-измеримую функцию от ξ_k. Вообще говоря, нельзя утверждать, что при фиксированных k, n, ω функция $\mathsf{P}(\xi_n \in \Gamma | \xi_k)$ является мерой на $\mathfrak{B}$.

В самом деле, из свойств условных вероятностей следует, что для каждой последовательности $\{\Gamma_r,\ r = 1,\ 2,\ \ldots\}$ непересекающихся множеств из σ-алгебры $\mathfrak{B}$ с вероятностью 1 выполнено

$$\mathsf{P}\left\{\xi_n \in \bigcup_{r=1}^{\infty} \Gamma_r \,|\, \xi_k\right\} = \sum_{r=1}^{\infty} \mathsf{P}\{\xi_n \in \Gamma_r \,|\, \xi_k\}.$$

При этом множество тех ω, для которых это равенство не имеет места, зависит от последовательности $\{\Gamma_r,\ r = 1,\ 2,\ \ldots\}$. Для другой последовательности это исключительное множество будет другим, и поэтому нельзя утверждать, что для почти всех ω условная вероятность $\mathsf{P}\{\xi_n \in \Gamma | \xi_k\}$ является мерой на $\mathfrak{B}$.

Тем не менее во многих случаях такое утверждение справедливо. Именно если X — полное метрическое сепарабельное пространство, а $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X, то существует такая функция $P(k,\ x,\ n,\ \Gamma)$, $0 \leqslant k < n$, $x \in X$, $\Gamma \in \mathfrak{B}$, что при каждых k, n и Γ с вероятностью 1

$$\mathsf{P}\{\xi_n \in \Gamma \,|\, \xi_k\} = P(k,\ \xi_k,\ n,\ \Gamma),$$

и при этом $P(k,\ n,\ x,\ \Gamma)$ $\mathfrak{B}$-измерима при фиксированных k, n, Γ, а при фиксированных k, x, n — это вероятностная мера на $\mathfrak{B}$. Очевидно, при $k = n$ должно быть $P(n,\ x,\ n,\ \Gamma) = \chi_\Gamma(x)$, где $\chi_\Gamma(x)$ — индикатор множества Γ.

Если для данной цепи $\{\xi_n,\ n = 0,\ 1,\ 2,\ \ldots\}$ в пространстве $(X,\ \mathfrak{B})$ такая функция $P(k,\ x,\ n,\ \Gamma)$ существует, то она называется *вероятностью перехода*. В терминах вероятностей перехода уравнение Колмогорова — Чепмена может быть записано так:

$$P(k,\ \xi_k,\ n,\ \Gamma) = \int_X P(m,\ y,\ n,\ \Gamma)\, P(k,\ \xi_k,\ m,\ dy).$$

Это равенство выполняется с вероятностью 1. Во многих случаях выполняется более сильное равенство

$$P(k,\ x,\ n,\ \Gamma) = \int_X P(m,\ y,\ n,\ \Gamma)\, P(k,\ x,\ m,\ dy)$$

для всех $0 \leqslant k \leqslant m \leqslant n$, $x \in X$, $\Gamma \in \mathfrak{B}$, которое также называется *уравнением Колмогорова — Чепмена* для вероятностей перехода. Вероятность перехода $P(k,\ x,\ n,\ \Gamma)$ можно интерпретировать как условную вероятность $\mathsf{P}\{\xi_n \in \Gamma | \xi_k = x\}$.

Заметим, что условные вероятности вида $\mathsf{P}\{\xi_n \in \Gamma | \xi_k = x\}$ для данной случайной последовательности $\{\xi_n,\ n = 0,\ 1,\ 2,\ \ldots\}$ могут удовлетворять уравнению Колмогорова — Чепмена без того, чтобы эта последовательность была цепью Маркова.

Пример 6. В урне лежат четыре шара, занумерованных цифрами 1, 2, 3, 4. Из нее наугад вынимают шар, замечают его номер и возвращают обратно. Предположим, что эта процедура продолжается сколь угодно долго. Обозначим через η_n номер шара, вынутого на n-м шаге. Для $j = 1, 2, 3$ пусть $A_j^{(n)}$ обозначает событие, состоящее в том, что $\eta_n = j$ или $\eta_n = 4$. Положим $\xi_{3(m-1)+j}$ ($m = 1, 2, \ldots$) равным 1 или 0 в соответствии с тем, произошло или не произошло событие $A_j^{(m)}$. Тогда для x_1, x_2, x_3, каждое из которых равно либо 0, либо 1, имеем

$$\mathsf{P}\{\xi_n = x_1\} = \mathsf{P}\{\xi_n = x_2 \mid \xi_m = x_3\} = 1/2, \quad n > m.$$

Поэтому для $k < m < n$ имеем

$$1/2 = \mathsf{P}\{\xi_n = x_2 \mid \xi_k = x_1\} = \mathsf{P}\{\xi_n = x_2 \mid \xi_m = 0\}\,\mathsf{P}\{\xi_m = 0 \mid \xi_k = x_1\} + \\ + \mathsf{P}\{\xi_n = x_2 \mid \xi_m = 1\}\,\mathsf{P}\{\xi_m = 1 \mid \xi_k = x_1\},$$

т. е. в данном случае уравнение Колмогорова — Чепмена для условных вероятностей выполнено. Тем не менее

$$\mathsf{P}\{\xi_{3m+3} = 1 \mid \xi_{3m+2} = 1, \quad \xi_{3m+1} = 1\} = 1, \qquad m = 1, 2, \ldots,$$

и потому последовательность $\{\xi_n, n = 1, 2, \ldots\}$ не является цепью Маркова.

8.1.4. Построение цепи Маркова по вероятности перехода. Пусть $(X, \mathfrak{B})$ — некоторое измеримое пространство. Предположим, что для всех $x \in X$, $\Gamma \in \mathfrak{B}$ и целых k, n таких, что $0 \leqslant k < n$, задана числовая функция $P(k, x, n, \Gamma)$, удовлетворяющая условиям:

а) при фиксированных k, n, Γ она $\mathfrak{B}$-измерима;

б) при фиксированных k, x, n она является вероятностной мерой на $\mathfrak{B}$;

в) для всех $0 \leqslant k < m < n$, $x \in X$ и $\Gamma \in \mathfrak{B}$ выполнено соотношение

$$P(k, x, n, \Gamma) = \int_X P(k, x, m, dy)\, P(m, y, n, \Gamma).$$

Спрашивается, существует ли на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ цепь Маркова $\{\xi_n, n = 0, 1, 2, \ldots\}$, для которой $P(k, x, n, \Gamma)$ была бы вероятностью перехода, т. е. с вероятностью 1

$$\mathsf{P}\{\xi_n \in \Gamma \mid \xi_k\} = P(k, \xi_k, n, \Gamma).$$

Ответ на этот вопрос дает следующая теорема.

Теорема 2. *Если функция $P(k, x, n, \Gamma)$ удовлетворяет условиям а) — в), то существуют вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ и последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ случайных элементов из $(X, \mathfrak{B})$ такие, что последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ является цепью Маркова с вероятностью перехода $P(k, x, n, \Gamma)$.*

Указанное в теореме 2 вероятностное пространство может быть построено следующим образом. Положим $\Omega = X^\infty$, $\mathfrak{F} = \mathfrak{B}^\infty$. Это означает, что элементами множества Ω являются всевозможные последовательности вида $\omega = (x_0, x_1, x_2, \ldots)$, где $x_i \in X$, а $\mathfrak{F}$ — минимальная σ-алгебра подмножеств Ω, содержащая все множества вида

$$\{\omega:\ x_0 \in \Gamma_0,\quad x_1 \in \Gamma_1,\ \ldots,\ x_n \in \Gamma_n\} \tag{1.2}$$

при всевозможных $n = 0, 1, 2, \ldots, \Gamma_0, \Gamma_1, \ldots, \Gamma_n \in \mathfrak{B}$.

Далее, пусть μ — произвольная вероятностная мера на $\mathfrak{B}$. На множествах вида (1.2) зададим числовую функцию P формулой

$$\mathsf{P}\{\omega:\ x_0 \in \Gamma_0,\quad x_1 \in \Gamma_1,\ \ldots,\ x_n \in \Gamma_n\} = \\ = \int\limits_{\Gamma_0} \mu(dx_0) \int\limits_{\Gamma_1} P(0, x_0, 1, dx_1) \int\limits_{\Gamma_2} P(1, x_1, 2, dx_2) \ldots \\ \ldots \int\limits_{\Gamma_n} P(n-1, x_{n-1}, n, dx_n).$$

Эта функция продолжается до вероятностной меры P на измеримом пространстве $(\Omega, \mathfrak{F})$. Положим для $n = 0, 1, 2, \ldots$ $\xi_n = \xi_n(\omega) = x_n$, если $\omega = (x_0, x_1, x_2, \ldots)$. Тогда на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ случайных элементов в $(X, \mathfrak{B})$ образует цепь Маркова, для которой заданная функция $P(k, x, n, \Gamma)$ является вероятностью перехода. При этом начальное состояние ξ_0 имеет распределение μ, называемое *начальным распределением* цепи.

По функции $P(k, x, n, \Gamma)$ цепь Маркова может быть построена неоднозначно: имеется произвол в выборе вероятностного пространства и начального распределения. Однако, если для двух цепей в одном и том же фазовом пространстве: $\{\xi_n, n = 0, 1, 2, \ldots\}$, заданной на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$, и $\{\xi'_n, n = 0, 1, 2, \ldots\}$, заданной на вероятностном пространстве $(\Omega', \mathfrak{F}', \mathsf{P}')$, — совпадают вероятности перехода и начальные распределения, то такие цепи стохастически эквивалентны в том смысле, что для любых $n = 0, 1, 2, \ldots$ и произвольного набора $\{\Gamma_i, i = 0, 1, 2, \ldots, n\}$ измеримых множеств из фазового пространства выполнено

$$\mathsf{P}\left\{\xi_0 \in \Gamma_0, \xi_1 \in \Gamma_1, \ldots, \xi_n \in \Gamma_n\right\} = \mathsf{P}'\left\{\xi'_0 \in \Gamma_0, \\ \xi'_1 \in \Gamma_1, \ldots, \xi'_n \in \Gamma_n\right\}.$$

Это означает, что цепь Маркова в указанном смысле однозначно определяется своей вероятностью перехода и начальным распределением.

8.1.5. Другое определение цепи Маркова. Пусть заданы вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ и определенная на нем цепь Маркова $\{\xi_n, n = 0, 1, 2, \ldots\}$ в фазовом пространстве $(X, \mathfrak{B})$. Предположим, что вероятность перехода этой цепи удовлетворяет условиям а) — в) п. 8.1.4. Тогда при любых $k \geqslant 0$ и $x \in X$ на σ алгебре $\mathfrak{F}^k$, порожденной случайными элементами $\xi_k, \xi_{k+1}, \ldots$, определены вероятностные меры P_{kx} такие, что с вероятностью 1 для $A \in \mathfrak{F}^k$ выполнено $\mathsf{P}_{k\xi_k}(A) = \mathsf{P}\{A|\mathfrak{F}_k\}$ (напомним, что $\mathfrak{F}_k$ — σ-алгебра, порожденная случайными элементами $\xi_0, \xi_1, \ldots, \xi_k$). Таким образом, $\mathsf{P}_{kx}(A)$, $A \in \mathfrak{F}^k$, задает условную вероятность события A при условии $\xi_k = x$. В частности, вероятность перехода цепи определяется через меры P_{kx} по формуле

$$P(k, x, n, \Gamma) = \mathsf{P}_{kx}\{\xi_n \in \Gamma\}, \quad 0 \leqslant k < n, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

Иногда под цепью Маркова понимают следующую совокупность объектов:

1) измеримое пространство $(\Omega, \mathfrak{F})$;

2) последовательность $\{\xi_n = \xi_n(\omega),\ n = 0, 1, 2, \ldots\}$ измеримых при каждом n отображений пространства $(\Omega, \mathfrak{F})$ в измеримое пространство $(X, \mathfrak{B})$;

3) семейство вероятностных мер P_{kx} (k — целые неотрицательные числа, $x \in X$), заданных на σ-алгебрах $\mathfrak{F}^k \subset \mathfrak{F}$, если выполнены следующие условия:

а) при фиксированных k, n и Γ, $0 \leqslant k < n$, $\Gamma \in \mathfrak{B}$, функция $P(k, x, n, \Gamma) = \mathsf{P}_{kx}\{\xi_n \in \Gamma\}$ $\mathfrak{B}$-измерима;

б) $\mathsf{P}_{kx}\{\xi_k \in \Gamma\} = \chi_\Gamma(x)$;

в) для всех $x \in X$, $0 \leqslant k < n$ и $\Gamma \in \mathfrak{B}$

$$\mathsf{P}_{kx}\{\xi_{n+1} \in \Gamma \mid \mathfrak{F}_n\} = \mathsf{P}_{n\xi_n}\{\xi_{n+1} \in \Gamma\}$$

с P_{kx}-вероятностью 1.

Если задана цепь Маркова в смысле такого определения, то, положив для любой вероятностной меры μ, заданной на $(X, \mathfrak{B})$,

$$\mathsf{P}_\mu^{(k)}(A) = \int_X \mathsf{P}_{kx}(A)\,\mu(dx), \quad A \in \mathfrak{F}^k, \qquad k = 0, 1, 2, \ldots,$$

получим последовательность $\xi_k, \xi_{k+1}, \xi_{k+2}, \ldots$ случайных элементов в $(X, \mathfrak{B})$, заданную на вероятностном пространстве $\left(\Omega, \mathfrak{F}^k, \mathsf{P}_\mu^{(k)}\right)$, которая образует цепь Маркова в смысле определения, данного в начале § 8.1.

Таким образом, марковская цепь в смысле последнего определения представляет собой целое семейство цепей Маркова (в смысле первого определения), «начинающихся» в момент времени k в точке x.

Функция $P(k, x, n, \Gamma)$, определенная в условии а), называется *вероятностью перехода цепи.*

Две цепи Маркова (в смысле последнего определения), заданные в одном фазовом пространстве, называются *эквивалентными*, если у них совпадают вероятности перехода. Если по эквивалентным цепям построить цепи Маркова в смысле первого определения с одними и теми же начальным распределением и начальным моментом, то они будут стохастически эквивалентными.

Заметим, что по функции $P(k, x, n, \Gamma)$, удовлетворяющей условиям теоремы 2, всегда можно построить цепь Маркова в смысле последнего определения.

8.2. Однородные цепи Маркова

8.2.1. Определение однородной цепи Маркова. Пусть задано вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$. Цепь Маркова $\{\xi_n,\ n = 0, 1, 2, \ldots\}$ в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода $P(k, x, n, \Gamma)$ называется *однородной*, если $P(k, x, n, \Gamma)$ представляет собой функцию от $x \in X$, $\Gamma \in \mathfrak{B}$ и $n - k$, $0 \leqslant k < n$. Обозначим через $P(n, x, \Gamma)$, $n > 0$, $x \in X$, $\Gamma \in \mathfrak{B}$, функцию, для которой

$$P(k, x, n, \Gamma) = P(n - k, x, \Gamma).$$

При $n=0$ естественно положить $P(0, x, \Gamma)=\chi_\Gamma(x)$. Функция $P(n, x, \Gamma)$ называется *вероятностью перехода однородной цепи.* В соответствии с п. 8.1.4 она удовлетворяет условиям:

а) при фиксированных n и Γ, $n=0, 1, 2, \ldots, \Gamma\in\mathfrak{B}$, функция $P(n, x, \Gamma)$ $\mathfrak{B}$-измерима;

б) при фиксированных n и x, $n=1, 2, \ldots, x\in X$, функция $P(n, x, \Gamma)$ является вероятностной мерой на $\mathfrak{B}$;

в′) при всех $0\leqslant k<m<n$ и $\Gamma\in\mathfrak{B}$ с вероятностью 1 выполнено соотношение

$$P(n-k, \xi_k, \Gamma)=\int_X P(n-m, y, \Gamma)\,P(m-k, \xi_k, dy).$$

Ниже всюду будет предполагаться, что вероятность перехода однородной цепи Маркова удовлетворяет условиям а), б) и следующему, несколько более сильному, чем в′), условию:

в) при всех $m>0$, $n>0$, $x\in X$ и $\Gamma\in\mathfrak{B}$ выполнено соотношение

$$P(m+n, x, \Gamma)=\int_X P(m, x, dy)\,P(n, y, \Gamma),$$

называемое *уравнением Колмогорова — Чепмена.*

Положим $P(x,\Gamma)=P(1, x, \Gamma)$. Функция $P(x, \Gamma)$ называется *вероятностью перехода за один шаг.* Из уравнения Колмогорова — Чепмена следует, что вероятность перехода за n шагов, т. е. функция $P(n, x, \Gamma)$, выражается через $P(x, \Gamma)$ при помощи рекуррентных соотношений

$$P(n+1, x, \Gamma)=\int_X P(n, y, \Gamma)\,P(x, dy)=$$
$$=\int_X P(y, \Gamma)\,P(n, x, dy), \qquad n=1, 2, \ldots$$

Поэтому, зная начальное распределение μ однородной цепи (т. е. меру $\mu(\Gamma)=\mathsf{P}\{\xi_0\in\Gamma\}$, $\Gamma\in\mathfrak{B}$) и вероятность перехода за один шаг, можно в принципе определить вероятность произвольного события, связанного с эволюцией рассматриваемой цепи, т. е. произвольного события из σ-алгебры $\mathfrak{F}^0$, порожденной элементами ξ_0, ξ_1, $\xi_2, \ldots$ Именно для событий вида

$$A=\{\xi_0\in\Gamma_0, \xi_1\in\Gamma_1, \ldots, \xi_n\in\Gamma_n\},$$
$$n=0, 1, 2, \ldots, \Gamma_0, \Gamma_1, \ldots, \Gamma_n\in\mathfrak{B},$$

имеем

$$\mathsf{P}(A)=\int_{\Gamma_0}\mu(dx_0)\int_{\Gamma_1}P(x_0, dx_1)\ldots\int_{\Gamma_{n-1}}P(x_{n-2}, dx_{n-1})\,P(x_{n-1},\Gamma_n).$$

Так как события указанного вида образуют алгебру, а $\mathfrak{F}^0$ — минимальная σ-алгебра, порожденная этой алгеброй, то вероятность произвольного события из $\mathfrak{F}^0$ однозначно восстанавливается по вероятностям всевозможных событий типа события A.

Отсюда следует, что все однородные цепи Маркова в одном и том же фазовом пространстве (возможно, на разных вероятностных пространствах), у которых совпадают начальные распределения и вероятности перехода за один шаг, *стохастически эквивалентны*. Это означает, что вероятности событий типа события A для всех таких цепей одни и те же.

Вероятность перехода за один шаг $P(x, \Gamma)$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяет условиям:

А) при фиксированном $\Gamma \in \mathfrak{B}$ функция $P(x, \Gamma)$ $\mathfrak{B}$-измерима по x;

Б) при фиксированном $x \in X$ она является вероятностной мерой на $\mathfrak{B}$.

Если на некотором измеримом пространстве $(X, \mathfrak{B})$ задана функция $P(x, \Gamma)$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяющая условиям А) и Б), то можно построить однородную цепь Маркова, для которой эта функция была бы вероятностью перехода за один шаг. Разумеется, с заданной вероятностью перехода за один шаг существует не одна цепь. Однако все они отличаются друг от друга (с точностью до стохастической эквивалентности) лишь начальным распределением.

8.2.2. Другое определение однородной цепи Маркова. При изучении однородных цепей Маркова удобно не фиксировать начальное распределение, а рассматривать целое семейство однородных цепей, «начинающихся» в произвольной неслучайной точке фазового пространства. Пусть $\{\xi_n, n = 0, 1, 2, \ldots\}$ — однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$, заданная на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$. Для каждого $x \in X$ можно построить семейство мер P_x на σ-алгебре $\mathfrak{F}^0$, порожденной случайными элементами ξ_0, $\xi_1, \ldots$, задав их на событиях вида $A_n(\Gamma_0, \Gamma_1, \ldots, \Gamma_n) = \{\xi_0 \in \Gamma_0, \xi_1 \in \Gamma_1, \ldots, \xi_n \in \Gamma_n\}$ формулой

$$\mathsf{P}_x\{A_n(\Gamma_0, \Gamma_1, \ldots, \Gamma_n)\} = \\ = \chi_{\Gamma_0}(x) \int_{\Gamma_1} P(x, dx_1) \ldots \int_{\Gamma_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, \Gamma_n),$$

а затем продолжая P_x до меры на $\mathfrak{F}^0$. Если $A \in \mathfrak{F}^0$, то с вероятностью 1 $\mathsf{P}_{\xi_0}(A) = \mathsf{P}(A \mid \xi_0)$. Поэтому $\mathsf{P}_x(A)$, $A \in \mathfrak{F}^0$, $x \in X$, естественно интерпретировать как условную вероятность события A при условии $\xi_0 = x$. Если ξ_0 имеет распределение μ, то сужение меры P на σ-алгебру $\mathfrak{F}^0$ $(\mathfrak{F}^0 \subset \mathfrak{F})$ совпадает с мерой

$$\mathsf{P}_\mu(A) = \int_X \mu(dx) \mathsf{P}_x(A), \quad A \in \mathfrak{F}^0.$$

Построенное семейство мер $\{\mathsf{P}_x, x \in X\}$ на $\mathfrak{F}^0$ обладает следующими свойствами:

1) при фиксированных $\Gamma \in \mathfrak{B}$ и $n = 0, 1, 2, \ldots$ функция $P(n, x, \Gamma) = \mathsf{P}_x\{\xi_n \in \Gamma\}$ $\mathfrak{B}$-измерима;

2) $\mathsf{P}_x\{\xi_0 \in \Gamma\} = \chi_\Gamma(x)$, $x \in X$, $\Gamma \in \mathfrak{B}$;

3) при всех $x \in X$, $\Gamma \in \mathfrak{B}$ и $n, m = 0, 1, 2, \ldots$ с P_x-вероятностью 1 выполнено соотношение

$$\mathsf{P}_x\{\xi_{n+m} \in \Gamma \mid \mathfrak{F}_m\} = \mathsf{P}_{\xi_m}\{\xi_n \in \Gamma\}.$$

Здесь $\mathfrak{F}_m$ — минимальная σ-алгебра, относительно которой измеримы элементы $\xi_0, \xi_1, \ldots, \xi_m$.

Иногда под однородной цепью Маркова понимают совокупность объектов:

а) последовательность $\{\xi_n = \xi_n(\omega),\ n = 0, 1, 2, \ldots\}$ измеримых отображений измеримого пространства $(\Omega, \mathfrak{F})$ в измеримое пространство $(X, \mathfrak{B})$;

б) семейство вероятностных мер $\{\mathsf{P}_x,\ x \in X\}$, заданных на σ-алгебре $\mathfrak{F}^0$, порожденной случайными элементами $\xi_0, \xi_1, \xi_2, \ldots$, если только выполнены условия 1) — 3).

При таком определении однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$ будет обозначаться (ξ_n, P_x). Фактически это целое семейство однородных цепей Маркова в первоначальном определении. Для того чтобы получить однородную цепь Маркова с фиксированным начальным распределением μ, нужно рассмотреть последовательность $\{\xi_n, n=0, 1, 2, \ldots\}$ на вероятностном пространстве $(\Omega, \mathfrak{F}^0, \mathsf{P}_\mu)$, где

$$\mathsf{P}_\mu(A) = \int_X \mathsf{P}_x(A)\,\mu(dx), \quad A \in \mathfrak{F}^0.$$

Две однородные цепи Маркова (ξ_n, P_x) и (ξ'_n, P'_x) в одном и том же фазовом пространстве $(X, \mathfrak{B})$ (возможно на разных вероятностных пространствах) *эквивалентны*, если для всех $x \in X$, $\Gamma \in \mathfrak{B}$

$$\mathsf{P}_x\{\xi_1 \in \Gamma\} = \mathsf{P}'_x\{\xi'_1 \in \Gamma\}.$$

Если по эквивалентным цепям Маркова построить марковские цепи в смысле первоначального определения с одним и тем же начальным распределением, то они будут стохастически эквивалентны.

Заметим, что по функции $P(x, \Gamma)$, удовлетворяющей условиям А) и Б) п. 8.2.1, всегда можно построить цепь Маркова (ξ_n, P_x), для которой $\mathsf{P}_x\{\xi_1 \in \Gamma\} = P(x, \Gamma)$. Такая цепь единственна с точностью до эквивалентности.

8.2.3. Следствия марковского свойства. Пусть (ξ_n, P_x) — однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$ в смысле определения п. 8.2.2. На σ-алгебре $\mathfrak{F}^0$, порожденной случайными элементами $\xi_0, \xi_1, \ldots$, определим семейство операторов θ_k $(k = 0, 1, 2, \ldots)$, отображающих $\mathfrak{F}^0$ в $\mathfrak{F}^0$ следующим образом. Для событий вида $\{\xi_n \in \Gamma\}$ $(n = 0, 1, 2, \ldots, \Gamma \in \mathfrak{B})$ положим

$$\theta_k\{\xi_n \in \Gamma\} = \{\xi_{n+k} \in \Gamma\}.$$

Кроме того, потребуем, чтобы операторы θ_k сохраняли все теоретико-множественные операции, т. е. чтобы для всех $A_j \in \mathfrak{F}^0$ выполнялись соотношения

$$\theta_k\Big\{\bigcup_j A_j\Big\} = \bigcup_j \theta_k A_j, \quad \theta_k\Big\{\bigcap_j A_j\Big\} = \bigcap_j \theta_k A_j,$$
$$\theta_k(A_j \setminus A_i) = \theta_k A_j \setminus \theta_k A_i.$$

Тем самым операторы θ_k однозначно определены.

Если $A \in \mathfrak{F}^0$, то $\theta_k A \in \mathfrak{F}^k$, где $\mathfrak{F}^k$ — σ-алгебра, порожденная элементами $\xi_k, \xi_{k+1}, \ldots$ Из свойств марковости и однородности следует, что для всех $A \in \mathfrak{F}^0$, $x \in X$, $k = 0, 1, 2, \ldots$ с P_x-вероятностью 1 выполнено

$$\mathsf{P}_x\{\theta_k A \mid \mathfrak{F}_k\} = \mathsf{P}_{\xi_k}(A), \tag{2.1}$$

где $\mathfrak{F}_k$ — σ-алгебра, порожденная элементами $\xi_0, \xi_1, \ldots, \xi_k$. Более общим образом, если $B \in \mathfrak{F}_k$, $A \in \mathfrak{F}^0$, то при всех $x \in X$ выполнено

$$\mathsf{P}_x(B \cap \theta_k A) = \int_B \mathsf{P}_{\xi_k}(A)\, \mathsf{P}_x(d\omega). \tag{2.2}$$

Операторы θ_k можно применить и к случайным величинам. Для $\mathfrak{F}^0$-измеримой случайной величины ξ положим $\eta = \theta_k\xi$, если при всех вещественных a

$$\theta_k\{\xi = a\} = \{\eta = a\}.$$

Если ξ — $\mathfrak{F}^0$-измерима, то $\theta_k\xi$ — $\mathfrak{F}^k$-измерима.

Из равенств (2.1) и (2.2) следуют соотношения

$$\mathsf{M}_x\{\theta_k\xi \mid \mathfrak{F}_k\} = \mathsf{M}_{\xi_k}\xi, \qquad \mathsf{M}_x\{\eta\theta_k\xi\} = \mathsf{M}_x\{\eta \mathsf{M}_{\xi_k}\xi\},$$

где ξ — $\mathfrak{F}^0$-измеримая случайная величина, η — $\mathfrak{F}_k$-измеримая случайная величина, M_x — знак математического ожидания по мере P_x ($k = 0, 1, 2, \ldots$). Первое из этих равенств выполняется с P_x-вероятностью 1 при естественном требовании суммируемости по мере P_x величины ξ (т. е. интегрируемости функции $\xi(\omega)$ по мере P_x). Для выполнения второго равенства достаточно потребовать, чтобы величины ξ и $\eta\theta_k\xi$ были P_x-суммируемы.

Если ξ и η неотрицательны, то оба равенства справедливы без каких-либо дополнительных ограничений.

8.2.4. Строгая марковость. Пусть (ξ_n, P_x) — однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$. Если интерпретировать ξ_n как положение движущейся частицы в момент времени n, то формула (2.1) показывает, что в каждый фиксированный момент времени движение начинается как бы заново.

Оказывается, что таким свойством обладают и некоторые случайные моменты времени. Пусть $\mathfrak{F}_k$ ($k = 0, 1, \ldots$), как и прежде, обозначает минимальную σ-алгебру событий, порожденную элементами $\xi_0, \xi_1, \ldots, \xi_k$, а $\mathfrak{F}^k$ ($k = 0, 1, 2, \ldots$) — минимальную σ-алгебру событий, порожденную случайными элементами $\xi_k, \xi_{k+1}, \ldots$ Неотрицательная целочисленная случайная величина τ, вообще говоря, несобственная (т. е. для некоторых ω возможно $\tau(\omega) = +\infty$), называется *марковским моментом*, если при любом $n = 0, 1, 2, \ldots$ событие $\{\tau \leqslant n\}$ $\mathfrak{F}_n$-измеримо ($\{\tau \leqslant n\} \in \mathfrak{F}_n$). Это означает, что для того, чтобы узнать, произошло ли событие $\{\tau \leqslant n\}$, нужно «проследить» за эволюцией цепи лишь до момента времени n включительно. Заметим, что требование $\{\tau \leqslant n\} \in \mathfrak{F}_n$ при всех n равносильно требованию, чтобы $\{\tau > n\} \in \mathfrak{F}_n$ при всех n, или требованию, чтобы $\{\tau = n\} \in \mathfrak{F}_n$ при всех n.

Пусть τ — марковский момент для однородной цепи Маркова (ξ_n, P_x). Обозначим через $\mathfrak{F}_\tau$ совокупность всех событий $A \in \mathfrak{F}^0$, для которых $A \cap \{\tau \leqslant n\} \in \mathfrak{F}_n$ при всех $n = 0, 1, 2, \ldots$ Тогда $\mathfrak{F}_\tau$ — σ-алгебра событий. В σ-алгебру $\mathfrak{F}_\tau$ входят те события, для которых узнать о том, произошли они или нет, можно, проследив за эволюцией цепи лишь до случайного момента времени τ.

Простейшим примером марковского момента τ может служить фиксированный (неслучайный) момент времени n. Для него $\mathfrak{F}_\tau$ совпадает с $\mathfrak{F}_n$.

Другим примером являются моменты первого достижения некоторых множеств. Пусть Γ — некоторое измеримое множество из фазового пространства. Положим $\tau = \inf\{k: \ \xi_k \in \Gamma\}$, причем если для данного ω при всех $k = 0, 1, 2, \ldots$ $\xi_k(\omega) \notin \Gamma$, то полагаем $\tau(\omega) = +\infty$. Тогда τ является марковским моментом и называется *моментом первого достижения множества* Γ.

Заметим, что если τ — марковский момент для однородной цепи Маркова (ξ_n, P_x), то величина τ $\mathfrak{F}_\tau$-измерима. Если, кроме того, $\tau < +\infty$ почти наверное относительно P_x при всех $x \in X$, то случайный элемент ξ_τ также $\mathfrak{F}_\tau$-измерим. Здесь $\xi_\tau = \xi_{\tau(\omega)}(\omega) = \xi_n(\omega)$ для $\omega \in \{\tau = n\}$.

Всякая однородная цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ обладает следующим свойством *строгой марковости*: для любого марковского момента τ и любых целых $n \geqslant 0$, $x \in X$, $\Gamma \in \mathfrak{B}$ выполняется соотношение

$$\mathsf{P}_x\{\xi_{n+\tau} \in \Gamma \mid \mathfrak{F}_\tau\} = P(n, \xi_\tau, \Gamma)$$

почти наверное относительно меры P_x на множестве $\{\tau < +\infty\}$. Здесь $P(n, x, \Gamma)$ — вероятность перехода цепи за n шагов, т. е. $P(n, x, \Gamma) = \mathsf{P}_x\{\xi_n \in \Gamma\}$.

Свойство строгой марковости показывает, что при фиксированном состоянии ξ_τ в марковский момент τ последовательность $\{\xi_{n+\tau}, n = 0, 1, \ldots\}$ представляет собой однородную цепь Маркова с начальным состоянием ξ_τ, имеющую такие же вероятности перехода, как и исходная цепь, и не зависящую от σ-алгебры $\mathfrak{F}_\tau$. Иными словами, если τ — марковский момент для цепи (ξ_k, P_x) и $\tau < +\infty$, то частица, находящаяся в момент времени n в состоянии ξ_n, начинает движение заново в момент времени τ.

Пусть заданы однородная цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ и марковский момент τ. Положим для каждого $A \in \mathfrak{F}^0$

$$\theta_\tau A = \bigcup_{n=0}^{\infty} \{\tau = n\} \cap \theta_n A$$

и для всякой $\mathfrak{F}^0$-измеримой случайной величины ξ

$$\theta_\tau \xi = \theta_n \xi(\omega),$$

если $\tau(\omega) = n$. Тогда:

а) для любого $A \in \mathfrak{F}^0$ и $x \in X$

$$\mathsf{P}_x\{\theta_\tau A \mid \mathfrak{F}_\tau\} = \mathsf{P}_{\xi_\tau}(A)$$

почти наверное относительно P_x на множестве $\{\tau < +\infty\}$;

б) для любых $A \in \mathfrak{F}^0$, $B \in \mathfrak{F}_\tau$ и $x \in X$

$$\mathsf{P}_x\{B \cap \theta_\tau A\} = \int_B \mathsf{P}_{\xi_\tau}(A)\, \mathsf{P}_x(d\omega);$$

в) если η — $\mathfrak{F}^0$-измеримая P_x-суммируемая случайная величина, то

$$\mathsf{M}_x\{\theta_\tau \eta \mid \mathfrak{F}_\tau\} = \mathsf{M}_{\xi_\tau} \eta$$

почти наверное относительно P_x на множестве $\{\tau < +\infty\}$;

г) если η — $\mathfrak{F}^0$-измеримая случайная величина, а ζ — $\mathfrak{F}_\tau$-измеримая случайная величина, то при всех $x \in X$

$$\mathsf{M}_x\{\zeta\theta_\tau\eta\} = \mathsf{M}_x\{\zeta\mathsf{M}_{\xi_\tau}\eta\}$$

в предположении, что величины η и $\zeta\theta_\tau\eta$ суммируемы по мере P_x $(x \in X)$.

Если η и ζ неотрицательны, то в) и г) выполнены без каких-либо дополнительных ограничений.

8.2.5. Операторы, связанные с цепью Маркова. Пусть (ξ_n, P_x) — однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода за один шаг $P(x, \Gamma)$, $x \in X$, $\Gamma \in \mathfrak{B}$. Обозначим через $\mathfrak{A}$ банахово пространство всех счетно-аддитивных вещественных функций ограниченной вариации (зарядов), заданных на σ-алгебре $\mathfrak{B}$, с нормой, равной вариации заряда, а через $\mathfrak{M}$ — банахово пространство всех вещественных $\mathfrak{B}$-измеримых функций, заданных на X, с нормой, равной супремуму модуля функции. Ядро $P(x, \Gamma)$ порождает операторы в пространствах $\mathfrak{A}$ и $\mathfrak{M}$, действующие по формулам

$$\varphi P(\Gamma) = \int_X P(x, \Gamma)\,\varphi(dx), \quad \varphi \in \mathfrak{A}, \quad \Gamma \in \mathfrak{B},$$

$$Pf(x) = \int_X f(y)\,P(x, dy), \quad f \in \mathfrak{M}, \quad x \in X.$$

Оба эти оператора линейны, непрерывны и имеют нормы, не превосходящие 1. При этом они положительны в том смысле, что $\mu P \geqslant 0$ для $\mu \geqslant 0$ и $Pf \geqslant 0$ для $f \geqslant 0$. Для $\varphi \in \mathfrak{A}$ и $f \in \mathfrak{M}$ положим

$$\langle \varphi, f\rangle = \int_X f(x)\,\varphi(dx).$$

Тогда при всех $\varphi \in \mathfrak{A}$ и $f \in \mathfrak{M}$ $\langle \varphi P, f\rangle = \langle \varphi, Pf\rangle$.

Обозначим через P^n n-ю степень оператора P. Из уравнения Колмогорова — Чепмена следуют соотношения

$$P^n f(x) = \int_X f(y)\,P(n, x, dy) = \mathsf{M}_x f(\xi_n), \quad n = 1, 2, \ldots, \quad x \in X, \quad f \in \mathfrak{M},$$

$$\varphi P^n(\Gamma) = \int_X P(n, x, \Gamma)\,\varphi(dx) = $$

$$= \int_X \varphi(dx)\,\mathsf{P}_x\{\xi_n \in \Gamma\}, \quad n = 1, 2, \ldots, \quad \Gamma \in \mathfrak{B}, \quad \varphi \in \mathfrak{A},$$

$$\langle \varphi P^n, f\rangle = \langle \varphi, P^n f\rangle = \int_X \varphi(dx) \int_X P(n, x, dy)\,f(y) = \int_X \varphi(dx)\,\mathsf{M}_x f(\xi_n).$$

При $n = 0$ естественно считать $P^0 = I$, где I — тождественный оператор.

Разумеется, оператор P можно применять и к неограниченным $\mathfrak{B}$-измеримым функциям, а также к зарядам неограниченной вариации, лишь бы имели смысл интегралы, определяющие этот оператор.

$\mathfrak{B}$-измеримая вещественная функция $f(x)$ $(x \in X)$ называется *супергармонической* (*субгармонической*), если для всех $x \in X$ имеет место неравенство $Pf(x) \leqslant f(x)$ $(Pf(x) \geqslant f(x))$. Если для всех $x \in X$ имеет место равенство $f(x) = Pf(x)$, то $f(x)$ называется *гармонической*. Неотрицательная супергармоническая функция называется *эксцессивной*.

Заряд φ (вообще говоря, неограниченной вариации) называется *инвариантным*, если $\varphi P = \varphi$. Конечная инвариантная мера μ (т. е. неотрицательный инвариантный заряд ограниченной вариации) называется *стационарной*. Стационарную меру всегда можно нормировать и считать вероятностной. Если для данной цепи (ξ_n, P_x) существует стационарная мера μ, то, взяв ее в качестве начального распределения, т. е. положив $\mathsf{P}\{\xi_0 \in \Gamma\} = \mu(\Gamma)$, будем иметь

$$\mathsf{P}_\mu\{\xi_n \in \Gamma\} = \int\limits_X \mu(dx)\, \mathsf{P}_x\{\xi_n \in \Gamma\} = \mu P^n(\Gamma) = \\ = \mu(\Gamma), \quad n = 0, 1, 2, \ldots, \quad \Gamma \in \mathfrak{B},$$

Значит, распределение элемента ξ_n по мере P_μ не меняется во времени. Более того, для $0 \leqslant n_1 < n_2 < \ldots < n_k$, $\Gamma_1, \Gamma_2, \ldots, \Gamma_k \in \mathfrak{B}$, $r > 0$ имеем

$$\mathsf{P}_\mu\left\{\xi_{n_1+r} \in \Gamma_1, \xi_{n_2+r} \in \Gamma_2, \ldots, \xi_{n_k+r} \in \Gamma_k\right\} = \\ = \int\limits_X \mu(dx_0) \int\limits_{\Gamma_1} P(n_1 + r, x_0, dx_1) \int\limits_{\Gamma_2} P(n_2 - n_1, x_1, dx_2) \ldots \\ \ldots \int\limits_{\Gamma_k} P(n_k - n_{k-1}, x_{k-1}, dx_k) = \int\limits_X \mu(dx_0) \int\limits_{\Gamma_1} P(n_1, x_0, dx_1) \times \\ \times \int\limits_{\Gamma_2} P(n_2 - n_1, x_1, dx_2) \ldots \int\limits_{\Gamma_k} P(n_k - n_{k-1}, x_{k-1}, dx_k) = \\ = \mathsf{P}_\mu\left\{\xi_{n_1} \in \Gamma_1, \xi_{n_2} \in \Gamma_2, \ldots, \xi_{n_k} \in \Gamma_k\right\}, \quad k = 1, 2, \ldots$$

Это означает, что последовательность $\{\xi_n, n = 0, 1, 2, \ldots\}$ в фазовом пространстве $(X, \mathfrak{B})$, заданная на вероятностном пространстве $(\Omega, \mathfrak{F}^0, \mathsf{P}_\mu)$, является *стационарной последовательностью*.

Таким образом, если для данной цепи существует стационарная мера, то, взяв ее в качестве начального распределения, получаем цепь, называемую *стационарной цепью Маркова*.

Определим оператор $G = \sum\limits_{n=0}^{\infty} P^n$. Он называется *потенциалом* цепи. Этот оператор применим не ко всякой $\mathfrak{B}$-измеримой функции (например, для $f(x) \equiv 1$ $Gf(x) = \sum\limits_{n=0}^{\infty} P^n f(x) \equiv +\infty$). В частности, может оказаться, что область его определения состоит из единственной функции $f(x) \equiv 0$.

Для $x \in X$, $\Gamma \in \mathfrak{B}$ положим

$$G(x, \Gamma) = G\chi_\Gamma(x) = \sum_{n=0}^{\infty} P^n \chi_\Gamma(x) = \sum_{n=0}^{\infty} P(n, x, \Gamma).$$

При фиксированном $x \in X$ $G(x, \cdot)$ — мера на $\mathfrak{B}$, возможно, тождественно равная бесконечности. Функция $G(x, \Gamma)$ называется *ядром потенциала*. Если $G|f|(x) < \infty$ для некоторой $\mathfrak{B}$-измеримой функции $f(x)$, $x \in X$, то

$$Gf(x) = \int_X f(y)\, G(x, dy).$$

Ядро потенциала имеет простой вероятностный смысл. Поскольку $P(n, x, \Gamma) = \mathsf{M}_x \chi_\Gamma(\xi_n)$, то

$$G(x, \Gamma) = \mathsf{M}_x \sum_{n=0}^{\infty} \chi_\Gamma(\xi_n),$$

т. е. $G(x, \Gamma)$ есть среднее число моментов времени, в которые система находилась в состояниях множества Γ, при условии, что в начальный момент она находилась в состоянии x.

Пусть $f(x) \geqslant 0$ и $\varphi(x) = Gf(x) < +\infty$ $(x \in X)$. Тогда $(P - I)\varphi = -f$, где I — тождественный оператор. Последнее равенство означает, что оператор G в некотором смысле обратный к оператору $I - P$ и что потенциал неотрицательной функции эксцессивен. Наоборот, если $f(x)$ $(x \in X)$ — эксцессивная функция, то $f(x) = G\varphi(x) + h(x)$, где $\varphi(x) \geqslant 0$, а $h(x)$ — гармоническая функция, т. е. $h = Ph$. Это утверждение является аналогом известной теоремы Рисса из теории дифференциальных уравнений.

Пример 1. Пусть X — целочисленная решетка на прямой, $\mathfrak{B}$ — σ-алгебра всех подмножеств X. Предположим, что задана последовательность $\{\eta_n,\ n = 1, 2, \ldots\}$ независимых одинаково распределенных случайных величин со значениями в X. Если η_0 — не зависящая от последовательности $\{\eta_n,\ n \geqslant 1\}$ целочисленная случайная величина, то последовательность $\{\xi_n = \eta_0 + \eta_1 + \ldots + \eta_n,\ n = 0, 1, 2, \ldots\}$ образует однородную цепь Маркова (случайное блуждание по целочисленной решетке). Положим

$$\varphi(\theta) = \mathsf{M}_x \exp\{i\theta\eta_1\}, \quad \theta \in \mathbf{R}.$$

Для вероятности перехода за n шагов имеем формулу

$$P(n, x, y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\theta(y-x)} \varphi^n(\theta)\, d\theta, \quad n = 0, 1, 2, \ldots, \quad x, y \in X$$

Поскольку σ-алгебра $\mathfrak{B}$ состоит из всех подмножеств X, то достаточно знать $P(n, x, y)$ для всех $x, y \in X$, так как для $\Gamma \subset X$ имеем

$$P(n, x, \Gamma) = \sum_{y \in \Gamma} P(n, x, y).$$

Предположим далее, что $\varphi(\theta)$ обращается в единицу лишь в точках, кратных 2π, и пусть существует математическое ожидание величины η_k $(k = 1, 2, \ldots)$, причем $a = \mathsf{M}\eta_k \neq 0$. Тогда для ядра потенциала справедлива формула

$$G(x, y) = \sum_{n=0}^{\infty} P(n, x, y) = \frac{1}{2|a|} + \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Re}\frac{e^{i\theta(x-y)}}{1-\varphi(\theta)}\,d\theta =$$

$$= \frac{1}{2|a|} + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(1-\varphi_c(\theta))\cos\theta(x-y) - \varphi_s(\theta)\sin\theta(x-y)}{|1-\varphi(\theta)|^2}\,d\theta,$$

$$x,\ y \in X,$$

где $\varphi_c(\theta) = \operatorname{Re}\varphi(\theta)$, $\varphi_s(\theta) = \operatorname{Im}\varphi(\theta)$. В частности, если величины η_k $(k = 1, 2, \ldots)$ принимают лишь значения $+1$ и -1 с вероятностями p и q соответственно, $p + q = 1$, $p - q = a > 0$, то

$$G(x, y) = \begin{cases} \dfrac{1}{a} & \text{при } y \geqslant x, \\ \dfrac{1}{a}\left(\dfrac{1-a}{1+a}\right)^{x-y} & \text{при } y < x. \end{cases}$$

8.2.6. Вероятностное представление решения задачи Дирихле. Пусть $P(x, \Gamma)$ — вероятность перехода за один шаг некоторой однородной цепи Маркова в фазовом пространстве $(X, \mathfrak{B})$; $\mathfrak{B}$-измеримую вещественную функцию $f(x)$, заданную на X, назовем *гармонической на множестве* $\Gamma \in \mathfrak{B}$, если для всех $x \in \Gamma$ выполнено равенство $f(x) = Pf(x)$.

Следующая задача является аналогом задачи Дирихле из теории дифференциальных уравнений.

Пусть заданы множество $D \in \mathfrak{B}$ и определенная на множестве $X \setminus D$ вещественная $\mathfrak{B}$-измеримая функция $g(x)$. Требуется найти такую $\mathfrak{B}$-измеримую функцию $f(x)$, чтобы на множестве D она была гармонической, а вне D совпадала с заданной функцией $g(x)$.

Запишем решение такой задачи в вероятностных терминах. Пусть τ — момент первого попадания во множество $X \setminus D$ для цепи Маркова с вероятностью перехода за один шаг $P(x, \Gamma)$: $\tau = \inf\{n\colon n \geqslant 0,\ \xi_n \notin D\}$, причем, если для некоторого ω $\xi_n(\omega) \in D$ для всех $n = 0, 1, 2, \ldots$, то полагаем $\tau(\omega) = +\infty$. Для всех $x \in X$ положим $f(x) = \mathsf{M}_x g(\xi_\tau)$. При этом, если $\tau = +\infty$, то считаем $g(\xi_\tau) = 0$; для $x \notin D$ $f(x) = g(x)$. Нетрудно проверить также, что $f(x)$ гармонична на множестве D.

Решение такой задачи не единственно.

8.2.7. Функционалы от цепи Маркова. Пусть (ξ_n, P_x) — однородная цепь Маркова в фазовом пространстве $(X, \mathfrak{B})$. Для вещественной $\mathfrak{B}$-измеримой функции $v(x)$ $(x \in X)$ положим

$$\eta_n = \sum_{k=0}^{n} v(\xi_k).$$

Случайная величина η_n представляет собой функционал от цепи Маркова (ξ_n, P_x). Это означает, что η_n $\mathfrak{F}_n$-измерима.

Распределение величины η_n определено, если известна функция

$$u_\theta(x, \Gamma; \lambda) = \sum_{n=0}^{\infty} \mathsf{M}_x \left\{e^{i\lambda\eta_n} \mid \xi_n \in \Gamma\right\} P(n, x, \Gamma)\,\theta^n,$$

где $x \in X$, $\Gamma \in \mathfrak{B}$, θ и λ — вещественные числа, причем $0 \leqslant \theta < 1$.

Можно доказать, что функция u_θ является единственным решением двух интегральных уравнений:

$$u_\theta(x, \Gamma; \lambda) = P_\theta(x, \Gamma) + \int_X \left(1 - e^{-i\lambda v(y)}\right) u_\theta(y, \Gamma; \lambda)\, P_\theta(x, dy),$$

$$u_\theta(x, \Gamma; \lambda) = P_\theta(x, \Gamma) + \int_X \left(1 - e^{-i\lambda v(y)}\right) P_\theta(y, \Gamma)\, u_\theta(x, dy; \lambda),$$

где

$$P_\theta(x, \Gamma) = \sum_{n=0}^{\infty} \theta^n P(n, x, \Gamma), \quad x \in X, \quad \Gamma \in \mathfrak{B}, \quad 0 \leqslant \theta < 1.$$

Такие уравнения могут быть полезными при изучении предельного поведения величин η_n при $n \to \infty$. В частности, может оказаться, что

$$\eta = \sum_{n=0}^{\infty} v(\xi_n) < \infty.$$

Это будет так, если, например,

$$\int_X |v(y)|\, G(x, dy) < \infty.$$

Здесь $G(x, \Gamma)$ — ядро потенциала цепи. В этом случае, положив

$$u(x; \lambda) = \mathsf{M}_x e^{i\lambda\eta},$$

будем иметь интегральные уравнения для функции $u(x; \lambda)$:

$$u(x; \lambda) = e^{i\lambda v(y)} \int_X P(x, dy)\, u(y; \lambda),$$

$$u(x; \lambda) = 1 + \int_X \left(1 - e^{-i\lambda v(y)}\right) u(y; \lambda)\, G(x, dy), \tag{2.3}$$

где $P(x, \Gamma)$ — вероятность перехода за один шаг.

Пример 2. Пусть X счетно и $\mathfrak{B}$ — σ-алгебра всех подмножеств X. Положим для некоторого $y_0 \in X$

$$v(y) = \begin{cases} 1 & \text{при} \quad y = y_0, \\ 0 & \text{при} \quad y \neq y_0. \end{cases}$$

Тогда величина $\eta_{y_0} = \sum_{n=0}^{\infty} v(\xi_n)$ есть число тех моментов времени, в которые цепь находится в состоянии y_0 за время от 0 до $+\infty$.

В предположении, что $G(y_0, y_0) < \infty$, уравнение (2.3) примет вид
$$u(x;\lambda) = 1 + \left(1 - e^{-i\lambda}\right) u(y_0;\lambda)\, G(x, y_0).$$
Из этого уравнения находим
$$u(x;\lambda) = 1 - \frac{a}{d} + \frac{c}{d}\cdot\frac{1}{1 - d\left(1 - e^{-i\lambda}\right)},$$
где $c = G(x, y_0)$, $d = G(y_0, y_0)$ $(c \leqslant d)$. Отсюда
$$\mathsf{P}_x\{\eta_{y_0} = 0\} = 1 - c/d,$$
$$\mathsf{P}_x\{\eta_{y_0} = n\} = \frac{c}{d}\cdot\frac{(d-1)^{n-1}}{d^n}, \qquad n = 1, 2, \ldots, x \in X,$$
т. е. величина η_{y_0} распределена по геометрическому закону. При этом $\mathsf{M}_x\eta_{y_0} = G(x, y_0)$.

8.2.8. Предельные теоремы для цепей Маркова. Пусть задана однородная цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$.

Важной задачей является изучение предельного поведения вероятностей $P(n, x, \Gamma)$ при $n \to \infty$. В § 8.3 эта задача будет подробно рассмотрена для того случая, когда множество X счетно или конечно. Здесь приведены основные результаты для общего случая при выполнении *условия Деблина*, которое в дальнейшем будет обозначаться символом (Д). Сформулируем его.

Существуют конечная мера φ на σ-алгебре $\mathfrak{B}$ $(\varphi(X) > 0)$, целое число $k_0 \geqslant 1$ и положительное число ε такие, что для всех $x \in X$ и всякого $\Gamma \in \mathfrak{B}$, для которого $\varphi(\Gamma) \leqslant \varepsilon$, выполнено неравенство
$$P(k_0, x, \Gamma) \leqslant 1 - \varepsilon.$$

Приведем примеры марковских цепей, удовлетворяющих условию (Д).

3. Для цепи Маркова с конечным множеством состояний положим, что $\varphi(\Gamma)$ равна числу точек во множестве Γ. Тогда $\Gamma = \varnothing$, если только $\varphi(\Gamma) < 1$. Стало быть, при $\varphi(\Gamma) < 1$ $P(n, x, \Gamma) = 0$, и условие (Д) выполнено. Таким образом, условие (Д) не накладывает никаких ограничений на конечные цепи Маркова.

Если множество состояний счетно, то условие (Д) выполнено, например, в том случае, когда ряд $\sum\limits_{y \in X} P(x, y)$ сходится равномерно относительно $x \in X$. Здесь $P(x, y)$ — вероятность перехода за один шаг из состояния x в состояние y. Однако требование равномерной сходимости указанного ряда значительно сильнее условия (Д).

4. Пусть X — борелевское множество в $\mathbf{R}^m$, а $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X. Предположим, что существует такая борелевская функция двух переменных $p(x, y)$ $(x, y \in X)$, что
$$P(x, \Gamma) = \int_\Gamma p(x, y)\, dy, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$
Очевидно, должно быть $p(x, y) \geqslant 0$ и
$$\int_X p(x, y)\, dy = 1.$$

В этом случае условие (Д) выполнено, если, например, $\operatorname{mes} X < \infty$, а функция $p(x, y)$ ограничена или равномерно относительно x интегрируема по y. Однако и здесь оба эти условия значительно сильнее условия (Д), поскольку из них следует, что равномерно по x $P(x, \Gamma) \to 0$ при $\operatorname{mes} \Gamma \to 0$, в то время как условие (Д) требует лишь, чтобы при малых $\operatorname{mes} \Gamma$ функция $P(x, \Gamma)$ была меньше единицы равномерно по x.

Множество $\Gamma \in \mathfrak{B}$ называется *последующим за состоянием* x_0, если $P(n, x_0, \Gamma) = 1$ при всех $n = 1, 2, \ldots$ Из условия (Д) следует, что если Γ — множество, последующее за состоянием x_0, то $\varphi(\Gamma) > \varepsilon$. Если множество Γ является последующим за каждым входящим в него состоянием, то оно называется *инвариантным*. Инвариантное множество, не содержащее никаких инвариантных подмножеств меньшей φ-меры, называется *минимальным*. Всякое множество, последующее за некоторым состоянием, содержит минимальное инвариантное множество. Два минимальных инвариантных множества либо не пересекаются, либо отличаются друг от друга лишь на множество φ-меры 0.

Пусть $K^1, K^2, \ldots, K^N$ — такие минимальные инвариантные множества, что $K^i \cap K^j = \varnothing$ при $i \neq j$, $\varphi(K^j) > \varepsilon$ и произвольное минимальное инвариантное множество отличается от некоторого K^j лишь на множество φ-меры 0. Очевидно, что $1 \leqslant N \leqslant \varphi(X)/\varepsilon$. Если в некоторый момент времени система попадает во множество K^j, то она в нем остается навсегда. Более того, справедливо следующее утверждение.

Теорема 1. *Если выполнено условие* (Д), *то существуют постоянные C и ρ, $C > 0$, $0 < \rho < 1$ такие, что*

$$1 - P\left(n, x, \bigcup_{j=1}^{N} K^j\right) \leqslant C\rho^n, \qquad n = 1, 2, \ldots, \quad x \in X.$$

Из этой теоремы и теоремы Бореля — Кантелли следует, что, каково бы ни было начальное состояние системы, с вероятностью 1 через конечное число шагов система окажется в одном из множеств K^j. Обозначим через $F^j(x)$ вероятность того, что система, выйдя из состояния x, когда-либо достигнет множества K^j:

$$F^j(x) = \mathsf{P}_x\left(\bigcup_{n=1}^{\infty} \{\xi_n \in K^j\}\right).$$

Так как, попав в K^j, система навсегда там остается, то

$$F^j(x) = \lim_{n \to \infty} P(n, x, K^j).$$

Если $x \in K^j$, то $F^j(x) = 1$. Кроме того, для всех $x \in X$

$$\sum_{j=1}^{N} F^j(x) = 1.$$

Из множества K^j можно выбросить такое его подмножество $\bar{K}^j$ нулевой φ-меры (возможно, пустое), что $P(x, \bar{K}^j) = 0$ для всех $x \in K^j \setminus \bar{K}^j$, $\lim\limits_{n \to \infty} P(n, x, \bar{K}^j) = 0$ равномерно относительно $x \in X$ и при этом множество $K^j \setminus \bar{K}^j$ может быть разбито на d_j $(1 \leqslant d_j < \infty)$

непересекающихся подмножеств K_i^j $(i=0, 1, 2, \ldots, d_j-1)$, для которых $P(x, K_i^j)=1$ при $x \in K_{i-1}^j$ $(i=1, 2, \ldots, d_j)$ (под $K_{d_j}^j$ понимаем K_0^j). Будем считать, что $\tilde{K}^j$ из K^j уже выброшено, так что $\bigcup_{i=0}^{d_j-1} K_i^j = K^j$. Множества K^j $(j=1, 2, \ldots, N)$ называются *эргодическими классами*, а K_i^j $(i=0, 1, \ldots, d_j-1)$ — *циклическими подклассами* класса K^j. Если система на некотором шаге попадает в класс K^j, то при $d_j > 1$ во все последующие моменты времени она будет циклически двигаться по подклассам этого класса. При $d_j = 1$ класс K^j называется *апериодическим*.

Назовем множество $\Gamma \in \mathfrak{B}$ *множеством несущественных состояний*, если $\lim_{n\to\infty} P(n, x, \Gamma) = 0$ для всех $x \in X$.

Таким образом, множество состояний X марковской цепи, удовлетворяющей условию (Д), может быть разбито на некоторое число эргодических классов K^j $(j=1, 2, \ldots, N)$ и множество несущественных состояний $X \setminus \bigcup_{j=1}^{N} K^j$. При этом каждый эргодический класс можно разбить на некоторое число циклических подклассов.

Положим для $x \in X$, $j=1, 2, \ldots, N$, $i=0, 1, \ldots, d_j-1$

$$F_i^j(x) = \mathsf{P}_x\left(\bigcup_{n=1}^{\infty}\left\{\xi_{nd_j} \in K_i^j\right\}\right).$$

Если при некотором n $\xi_{nd_j} \in K_i^j$, то это же верно и для всех $k \geqslant n$. Отсюда

$$F_i^j(x) = \lim_{n\to\infty} P\left(nd_j, x, K_i^j\right).$$

Очевидно, $F_i^j(x) = 1$ для $x \in K_i^j$. Кроме того, для $x \in X$

$$\sum_{i=0}^{d_j-1} F_i^j(x) = F^j(x).$$

Сформулируем теперь основную теорему о предельном поведении вероятностей $P(n, x, \Gamma)$ при $n \to \infty$.

Теорема 2. *Пусть выполнено условие* (Д). *Тогда существует такая система вероятностных мер* π_i^j $(j=1, 2, \ldots, N,\ i=0, 1, \ldots, d_j-1)$, *заданных на* $\mathfrak{B}$, *что для всех* $x \in X$, $\Gamma \subset K^j$ (*разумеется*, $\Gamma \in \mathfrak{B}$)

$$\lim_{n\to\infty} P\left(nd_j + m, x, \Gamma\right) = \sum_{r=0}^{d_j-1} F_r^j(x)\, \pi_{r+m}^j(\Gamma),$$

где индекс $r+m$ *рассматривается по модулю* d_j. *При этом* $\pi_i^j\left(K_i^j\right) = 1$ *и* $\pi_i^j(\Gamma) > 0$ *при* $\varphi\left(\Gamma \cap K_i^j\right) > 0$. *Стремление к пределу в этом соотношении равномерно по* x *и* Γ.

В частности, если $x \in K_i^j$, то $\lim_{n\to\infty} P\left(nd_j+m,\ x,\ \Gamma\right) = \pi_r^j(\Gamma)$ при $r \equiv i+m\left(\mathrm{mod}\, d_j\right)$. Заметим, что $P\left(nd_j+m,\ x,\ \Gamma\right) = P\left(nd_j+m,\ x,\ \Gamma \cap K_r^j\right)$ при $x \in K_i^j$ и $r \equiv i+m\left(\mathrm{mod}\, d_j\right)$.

Из этой теоремы следует *сходимость средних по Чезаро*. Именно, равномерно относительно $x \in X$ и $\Gamma \in \mathfrak{B}$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P(k,\ x,\ \Gamma) = \sum_{j=1}^{N} F^j(x)\,\pi^j(\Gamma),$$

где

$$\pi^j(\Gamma) = \frac{1}{d_j}\sum_{i=0}^{d_j-1} \pi_i^j(\Gamma), \quad \Gamma \in \mathfrak{B},$$

так что π^j является вероятностной мерой на σ-алгебре $\mathfrak{B}$, причем $\pi^j(K^j) = 1$ и $\pi^j(\Gamma) > 0$, если $\varphi(\Gamma \cap K^j) > 0$.

Обозначим

$$\mu_x(\Gamma) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P(k,\ x,\ \Gamma). \tag{2.4}$$

При каждом $x \in X$ функция μ_x является вероятностной мерой на $\mathfrak{B}$, так как $F^j(x) \geqslant 0$ и $\sum_{j=1}^{N} F^j(x) = 1$. Если $x \in K^j$, то μ_x совпадает с мерой π^j. Если Γ — множество несущественных состояний, то $\mu_x(\Gamma) = 0$.

Теорема 3. *Если выполнено условие* (Д), *то при каждом $x \in X$ мера μ_x является стационарной* (см. п. 8.2.5). *Наоборот, всякое стационарное распределение может быть записано в виде $\sum_{j=1}^{N} \rho_j \pi^j$, где числа ρ_j неотрицательны и в сумме равны единице.*

Следствие. 1) *Предел в соотношении* (2.4) *тогда и только тогда не зависит от x, когда имеется только один эргодический класс.*

2) *Предел $P(n,\ x,\ \Gamma)$ при $n \to \infty$ существует тогда и только тогда, когда ни один из эргодических классов не содержит циклических подклассов, т. е. $d_j = 1$ при всех j.*

Предположим теперь, что на X задана вещественная $\mathfrak{B}$-измеримая функция $v(x)$. Следующая теорема является аналогом закона больших чисел для последовательности $\{v(\xi_n),\ n = 0,\ 1,\ 2,\ \ldots\}$.

Теорема 4. *Если выполнено условие* (Д) *и $v(x)$ $(x \in X)$ такова, что*

$$\int_{K^j} |v(y)|\,\pi^j(dy) < \infty, \quad j = 1,\ 2,\ \ldots,\ N,$$

то для любой вероятностной меры μ *на* $\mathfrak{B}$ *с* P_μ*-вероятностью* 1 *существует предел* $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=0}^{n} v(\xi_k)$, *и он равен* $\int\limits_{K^j} v(y)\,\pi^j(dy)$ *при* $\xi_0(\omega)\in K^j$, *где* $\mathsf{P}_\mu(\cdot)=\int\limits_X \mathsf{P}_x(\cdot)\,\mu(dx)$.

В частности, если имеется только один эргодический класс, т. е. $N=1$, то

$$\mathsf{P}_\mu\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n} v(\xi_k)=\int\limits_X v(y)\,\pi(dy)\right\}=1,$$

где π — единственное стационарное распределение. (Заметим, что $\int\limits_X v(y)\,\pi(dy)=\mathsf{M}_\pi v(\xi_k)$, где $\mathsf{M}_\pi(\cdot)$ есть среднее по мере P_π.)

Следующая теорема описывает флуктуации величины $\frac{1}{n}\sum\limits_{k=0}^{n} v(\xi_k)$ вокруг предельного значения.

Теорема 5. *Предположим, что для данной цепи Маркова выполняется условие* (Д), *существует только один эргодический класс и он апериодичен. Обозначим через* π *единственное стационарное распределение цепи. Пусть задана вещественная* $\mathfrak{B}$*-измеримая функция* $v(x)$ $(x\in X)$, *для которой при некотором* $\delta>0$

$$\int\limits_X |v(x)|^{2+\delta}\,\pi(dx)<\infty.$$

Тогда существует предел

$$\lim_{n\to\infty}\mathsf{M}_\pi\left\{\left[\frac{1}{\sqrt{n}}\sum_{k=0}^{n}\left(v(\xi_k)-\mathsf{M}_\pi v(\xi_k)\right)\right]^2\right\}=\sigma^2,$$

и если $\sigma^2>0$, *то для любого начального распределения* μ

$$\lim_{n\to\infty}\mathsf{P}_\mu\left\{\frac{1}{\sigma\sqrt{n}}\sum_{k=0}^{n}\left(v(\xi_k)-\mathsf{M}_\pi v(\xi_k)\right)<\alpha\right\}=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\alpha} e^{-\beta^2/2}\,d\beta$$

равномерно относительно $\alpha\in\mathbb{R}$.

8.2.9. Эргодичность. Возвратность по Харрису. Однородная цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ называется *эргодической*, если для всех $x\in X$ почти наверное относительно P_x существует неслучайный и не зависящий от x предел

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} v(\xi_k), \tag{2.5}$$

какова бы ни была ограниченная $\mathfrak{B}$-измеримая функция v.

Из теоремы 4 следует, что эргодической будет цепь Маркова, удовлетворяющая условию (Д) и имеющая лишь один эргодический класс (см. п. 8.2.8).

Теорема 6. *Пусть в фазовом пространстве* $(X, \mathfrak{B})$ *задана однородная цепь Маркова* (ξ_n, P_x). *Следующие утверждения эквивалентны:*

1) *цепь* $(\xi_n\ \mathsf{P}_x)$ *является эргодической;*

2) *для всякого* $x \in X$, $\Gamma \in \mathfrak{B}$ *существует не зависящий от* x *предел*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P(k, x, \Gamma);$$

значение этого предела равно $\pi(\Gamma)$, *где* π — *стационарная вероятностная мера для цепи* (ξ_n, P_x), *т. е. мера на* $(X, \mathfrak{B})$, *для которой* $\pi(X) = 1$ *и при всех* $\Gamma \in \mathfrak{B}$

$$\pi(\Gamma) = \int_X P(x, \Gamma)\, \pi(dx); \tag{2.6}$$

3) *для цепи* (ξ_n, P_x) *существует стационарная вероятностная мера и не существует ограниченных гармонических функций, отличных от постоянных.*

Для эргодической цепи (ξ_n, P_x) предел (2.5) (временно́е среднее) совпадает со средним от функции v по мере π (пространственное среднее):

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} v(\xi_k) = \int_X v(x)\, \pi(dx) \tag{2.7}$$

почти наверное относительно меры P_μ, каким бы ни было начальное распределение μ. Полагая здесь $v(x) = \chi_\Gamma(x)$, $x \in X$, $\Gamma \in \mathfrak{B}$, где $\chi_\Gamma(x)$ — индикатор множества Γ, получим

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \chi_\Gamma(\xi_k) = \pi(\Gamma),$$

откуда следует, что множество положительной π-меры цепь посещает бесконечно много раз почти наверное относительно меры P_x при любом $x \in X$. Это свойство эргодической цепи лежит в основе понятия возвратности по Харрису.

Однородная цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ называется *возвратной по Харрису*, если существует такая нетривиальная σ-конечная мера μ *на* $\mathfrak{B}$, что при любом $x \in X$ почти наверное относительно меры P_x

$$\sum_{n=0}^{\infty} \chi_\Gamma(\xi_k) = +\infty$$

для всякого такого множества $\Gamma \in \mathfrak{B}$, для которого $\mu(\Gamma) > 0$. Меру

μ в этом определении называют *мерой, доставляющей возвратность цепи*, а саму цепь называют также *μ-возвратной*. В соответствии с этим эргодическая цепь является π-возвратной.

Если цепь (ξ_n, P_x) возвратна по Харрису и для нее существует стационарная вероятностная мера, то такая цепь является эргодической.

При весьма широких предположениях о фазовом пространстве возвратные по Харрису цепи Маркова обладают инвариантной мерой, однако не обязательно конечной, а лишь σ-конечной. Иначе говоря, для возвратных по Харрису цепей уравнение (2.6) во многих случаях имеет неотрицательное решение $\pi(\Gamma)$, для которого, вообще говоря, $\pi(X) = \infty$, но существует счетный набор множеств $\Gamma_n \in \mathfrak{B}$, $n = 1, 2, \ldots$ таких, что $\Gamma_1 \subset \Gamma_2 \subset \ldots, \bigcup\limits_{n=1}^{\infty} \Gamma_n = X$ и $\pi(\Gamma_n) < \infty$. Точнее, справедлив следующий результат.

Теорема 7. *Предположим, что цепь (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ возвратна по Харрису, и пусть σ-алгебра $\mathfrak{B}$ порождается счетным набором своих элементов (в таком случае σ-алгебру $\mathfrak{B}$ называют счетно порожденной). Тогда для цепи (ξ_n, P_x) существует единственная с точностью до положительного множителя неотрицательная инвариантная σ-конечная мера π и цепь является π-возвратной. Если при этом некоторая σ-конечная мера μ доставляет возвратность цепи, то μ абсолютно непрерывна относительно π.*

Аналогом эргодической теоремы для возвратных по Харрису цепей (т. е. аналогом соотношения (2.7)) является следующее утверждение.

Теорема 8. *В условиях теоремы 7 при всех $x \in X$ P_x-почти наверное имеет место соотношение*

$$\lim_{n\to\infty}\left(\sum_{k=0}^{n} v(\xi_k)\Big/\sum_{k=0}^{n} u(\xi_k)\right) = \int\limits_X v(y)\,\pi(dy)\Big/\int\limits_X u(y)\,\pi(dy)$$

для любых $\mathfrak{B}$-измеримых и суммируемых по мере π функций u, v на X, причем $\int\limits_X u(y)\,\pi(dy) \neq 0$.

Циклическое разложение возвратной по Харрису цепи, а стало быть, и решение вопроса о предельном поведении вероятностей $P(n, x, \Gamma)$ при $n \to \infty$ можно получить из следующего утверждения. В формулировке этого утверждения фигурирует оператор θ_1, определение которого приведено в п. 8.2.3 (мы предполагаем, что пространство элементарных событий Ω замкнуто относительно сдвигов, т. е. для всякого $\omega \in \Omega$ существует $\omega' \in \Omega$ такое, что $\xi_n(\omega') = \xi_{n+1}(\omega)$ при всех $n = 0, 1, 2, \ldots$).

Комплексное число z называется *собственным числом* оператора θ_1, если существует такая комплекснозначная $\mathfrak{F}^0$-измеримая величина η, что $\theta_1\eta = z\eta$ P_π-почти наверное и $\mathsf{M}_\pi|\eta| > 0$. В этом случае величина η называется *собственной*. Заметим, что число $z = 1$ является, очевидно, собственным, и ему соответствует собственная величина $\eta(\omega) \equiv 1$.

Теорема 9. *В условиях теоремы 7 число d собственных чисел оператора θ_1 конечно и они совпадают с корнями степени d из 1. Собственному числу* $\exp\{2k\pi i/d\}$, $k = 0, 1, \ldots, d-1$, *отвечает единственная с точностью до мультипликативной константы и P_π-эквивалентности собственная величина* $\exp\{2k\pi i r(\xi_0)/d\}$, *где $r(x)$ — $\mathfrak{B}$-измеримая функция на X со значениями в конечном множестве* $\{0, 1, \ldots, d-1\}$.

(Здесь π под знаком экспоненты обозначает число, равное отношению длины окружности к ее диаметру, в отличие от π, обозначающего инвариантную меру.)

Из этой теоремы нетрудно вывести, что в фазовом пространстве X возвратной по Харрису цепи (со счетно порожденной σ-алгеброй $\mathfrak{B}$) можно выделить непересекающиеся подмножества $K_0, K_1, \ldots, K_{d-1} \in \mathfrak{B}$ такие, что $P(x, K_{j+1}) = 1$ для $x \in K_j$ ($j = 0, 1, \ldots, d-1$) (считаем $K_d = K_0$) и при этом для всех $x \in X$ с P_x-вероятностью 1 множество $X \setminus \bigcup_{j=0}^{d-1} K_j$ посещается цепью лишь конечное число раз. Другими словами, выходя из произвольного состояния $x \in X$, цепь рано или поздно попадет в одно из множеств K_j и дальше будет двигаться по этим множествам циклически.

Число d называется *периодом цепи*. Если $d = 1$, цепь называется апериодической. Множества K_j называются *циклическими подклассами* цепи.

Обозначим через τ момент первого попадания цепи во множество $\bigcup_{j=0}^{d-1} K_j$. Как уже отмечалось, этот момент конечен P_x-почти наверное при любом $x \in X$. Положим $F_j(x) = \mathsf{P}_x\{\xi_\tau \in K_j\}$, $x \in X$, $j = 0, 1, \ldots, d-1$. Тогда $\sum_{j=0}^{d-1} F_j(x) = 1$ при всех $x \in X$ и $F_j(x) = 1$ при $x \in K_j$.

Далее, возвратная по Харрису цепь называется *положительной*, если мера π, существование которой утверждается в теореме 7, конечна (предполагается, что σ-алгебра $\mathfrak{B}$ счетно порождена). Если же $\pi(X) = \infty$, то цепь называется *нулевой*.

Следующая теорема описывает поведение вероятностей $P(n, x, \Gamma)$ при $n \to \infty$ для возвратных по Харрису цепей.

Теорема 10. *Предположим, что возвратная по Харрису цепь Маркова* (ξ_n, P_x) *задана в фазовом пространстве* $(X, \mathfrak{B})$ *со счетно порожденной σ-алгеброй $\mathfrak{B}$. Тогда:*

1) *если цепь положительна со стационарной вероятностной мерой π и апериодична, то при всех $x \in X$ равномерно относительно* $\Gamma \in \mathfrak{B}$

$$\lim_{n\to\infty} P(n, x, \Gamma) = \pi(\Gamma);$$

2) *если цепь нулевая, то для всякой $\mathfrak{B}$-измеримой функции v, суммируемой по инвариантной мере,*

$$\lim_{n\to\infty} \int_X v(y)\, P(n, x, dy) = 0, \quad x \in X;$$

3) *если цепь положительна и имеет период $d > 1$, то при всех $x \in X$ равномерно относительно $\Gamma \in \mathfrak{B}$*

$$\lim_{n\to\infty} P(nd+m, x, \Gamma) = d\sum_{r=0}^{d-1} F_r(x)\,\pi(\Gamma\cap K_{r+m}),$$

где сумма $r+m$ рассматривается по модулю d; в частности, если $x \in K_j$, то

$$\lim_{n\to\infty} P(nd+m, x, \Gamma) = d\pi(\Gamma\cap K_r)$$

равномерно относительно Γ, где $j+m = r \pmod d$.

Сформулируем центральную предельную теорему для функционалов от цепей Маркова.

Теорема 11. *Пусть задана однородная эргодическая цепь Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ со счетно порожденной σ-алгеброй $\mathfrak{B}$, и пусть на пространстве $(X\times X, \mathfrak{B}\times\mathfrak{B})$ определена вещественная измеримая функция $v(x, y)$, для которой $\mathsf{M}_\pi v(\xi_0, \xi_1) = 0, \mathsf{M}_\pi v^2(\xi_0, \xi_1) < \infty$ (по-прежнему π означает стационарную меру цепи). Предположим далее, что найдется такое множество $D \in \mathfrak{B}$ положительной π-меры, для которого*

$$\begin{gathered}\sup_n \frac{1}{n}\int_D \pi(dx)\,\mathsf{M}_x\left[\sum_{k=1}^n v(\xi_{k-1}, \xi_k)\right]^2 < \infty,\\ \sup_n\left|\int_A \pi(dx)\,\mathsf{M}_x\sum_{k=1}^n v(\xi_{k-1}, \xi_k)\right| < \infty,\end{gathered} \tag{2.8}$$

каково бы ни было $A \subset D$ $(A \in \mathfrak{B})$.

Тогда независимо от начального распределения цепи случайные величины

$$\frac{1}{\sqrt{n}}\sum_{k=0}^n v(\xi_{k-1}, \xi_k)$$

имеют в пределе при $n\to\infty$ нормальное распределение со средним 0 и дисперсией $\sigma^2 \geqslant 0$. При этом $\sigma^2 = 0$ тогда и только тогда, когда найдется такая $\mathfrak{B}$-измеримая функция $a(x)$, что

$$v(\xi_0, \xi_1) = a(\xi_1) - a(\xi_0)$$

почти наверное относительно меры P_π.

Следующая теорема дает условия применимости центральной предельной теоремы в терминах резольвенты цепи.

Обозначим через $L_2 = L_2(X, \mathfrak{B}, \pi)$ гильбертово пространство всех вещественных $\mathfrak{B}$-измеримых функций с π-интегрируемым квадратом. Пусть P — оператор в L_2, действующий на функцию $f \in L_2$ по формуле

$$Pf(x) = \int_X P(x, dy)\,f(y),$$

а Π — проектор, проектирующий функцию $f \in L_2$ в постоянную функцию $\Pi f(x) = \int\limits_X \pi(dy) f(y)$.

Наконец, для $\lambda \in [0, 1)$ положим

$$R_\lambda = \sum_{n=0}^{\infty} \lambda^n P^n.$$

Теорема 12. *Пусть цепь Маркова* (ξ_n, P_x) *и функция* $v(x, y)$ *удовлетворяют всем условиям предыдущей теоремы, за исключением неравенств* (2.8). *Предположим, что*

$$\sup_{0<\lambda<1} \left\| R_\lambda - \frac{1}{1-\lambda} \Pi \right\|_{L_2} < \infty.$$

Тогда справедливо утверждение теоремы 11.

8.3. Цепи Маркова с дискретным множеством состояний

8.3.1. Матрицы вероятностей перехода. Рассмотрим однородные цепи Маркова (ξ_n, P_x) в фазовом пространстве $(X, \mathfrak{B})$ в предположении, что X счетно или конечно, а $\mathfrak{B}$ — σ-алгебра всех подмножеств X. Такие цепи определяются вероятностями перехода за один шаг в одноточечные множества $P(x, y) = \mathsf{P}_x\{\xi_1 = y\}$ $(x, y \in X)$, ибо для произвольного $\Gamma \subset X$ имеем $P(x, \Gamma) = \mathsf{P}_x\{\xi_1 \in \Gamma\} = \sum\limits_{y \in \Gamma} P(x, y)$. Числа $P(x, y)$ $(x, y \in X)$ образуют матрицу P, возможно бесконечную, в x-й строке которой расположены вероятности перехода за один шаг из состояния x во всевозможные состояния $y \in X$, а в y-м столбце — вероятности перехода за один шаг из всевозможных состояний $x \in X$ в состояние y. Элементы матрицы P неотрицательны, и сумма их по строке равна 1. Такие матрицы называются *стохастическими*. Каждая стохастическая матрица определяет единственную с точностью до эквивалентности однородную цепь Маркова, для которой вероятности перехода за один шаг совпадают с элементами этой матрицы.

Вероятности перехода за n шагов $P(n, x, y)$ также образуют стохастическую матрицу, и она равна n-й степени матрицы P, как это следует из уравнения Колмогорова — Чепмена:

$$P(n, x, y) = \sum_{z_1, \ldots, z_{n-1} \in X} P(x, z_1) P(z_1, z_2) \ldots P(z_{n-1}, y),$$

где $x, y \in X$, $n = 2, 3, \ldots$ При $n = 0$ естественно считать функцию $P(0, x, y) = 1$ для $x = y$ и $P(0, x, y) = 0$ для $x \neq y$, так что вероятности перехода за 0 шагов образуют единичную матрицу I.

В случае, когда X конечно, для вероятностей перехода за n шагов справедлива следующая формула:

$$P(n, x, y) = \sum_{k=1}^{r} \frac{1}{(m_k - 1)!} \frac{d^{m_k - 1}}{d\lambda^{m_k - 1}} \left[\frac{\lambda^n M_{xy}(\lambda)}{\psi_k(\lambda)} \right]_{\lambda = \lambda_k}, \tag{3.1}$$

где $n = 0, 1, 2, \ldots$; $x, y \in X$; $\lambda_1, \lambda_2, \ldots, \lambda_r$ — различные корни уравнения $\det(\lambda I - P) = 0$; $m_1, m_2, \ldots, m_r$ — их кратности соответственно; $M_{xy}(\lambda)$ — алгебраическое дополнение элемента y-й строки и x-го столбца матрицы $\lambda I - P$, $\psi_k(\lambda) = (\lambda - \lambda_k)^{-m_k} \det(\lambda I - P)$.

8.3.2. Классификация состояний. Состояние $y \in X$ достижимо из состояния $x \in X$, если при некотором $n = 0, 1, 2, \ldots$ $P(n, x, y) > 0$. Если y достижимо из x, а x достижимо из y, то состояния x и y называются *сообщающимися*. Тем самым во множестве X вводится «отношение», обладающее свойством симметрии, рефлексивности и транзитивности. Стало быть, множество X можно разбить на непересекающиеся классы сообщающихся между собой состояний. При этом никакие два состояния из разных классов не сообщаются между собой, однако для состояний данного класса могут быть достижимы состояния других классов. Иначе говоря, из данного состояния система может перейти с положительной вероятностью в любое другое состояние того же класса. Кроме того, возможен выход из данного класса состояний, но вернуться в исходный класс система не может.

Марковская цепь (ξ_n, P_x) называется *неприводимой*, если любая пара состояний $x, y \in X$ представляет собой сообщающиеся состояния. Другими словами, все состояния неприводимой цепи образуют один класс сообщающихся состояний.

Состояние x называется *существенным*, если любое состояние, достижимое из x, сообщается с x. В противном случае оно называется *несущественным*. Для несущественного состояния x существует по крайней мере одно состояние y, достижимое из x, такое, что x недостижимо из y.

Нетрудно убедиться в том, что из существенного состояния достижимы только существенные состояния. Отсюда следует, что в каждом классе сообщающихся состояний либо все состояния существенны, либо все несущественны.

На множестве классов сообщающихся состояний, на которые разбивается пространство всех возможных состояний данной цепи, можно ввести частичный порядок при помощи следующего отношения. Говорят, что класс X_β следует за классом X_α, если хотя бы для одного $x \in X_\alpha$ существует такое $y \in X_\beta$, что y достижимо из x. (Отсюда следует, что из любого состояния $x \in X_\alpha$ достижимо любое состояние $y \in X_\beta$, если только X_β следует из X_α.) Такое отношение между классами обладает свойством рефлексивности и транзитивности, однако не обладает симметрией, и потому порождает частичный порядок во множестве классов. Очевидно, что только классы существенных состояний (если только они есть) обладают тем свойством, что за ними не следует никакой другой класс. Если для данной цепи число классов сообщающихся состояний конечно, то существует хотя бы один класс существенных состояний. В общем случае может не найтись ни одного класса существенных состояний.

Если начальное состояние цепи находится в каком-либо классе существенных состояний, то система никогда не выйдет из этого класса. Поэтому, если все классы состоят из существенных состояний, то цепь фактически разлагается на несколько цепей, соответствующих каждому такому классу.

8.3.3. Периодичность. Обозначим через $d(x)$ $(x \in X)$ наибольший общий делитель тех чисел n, для которых $P(n, x, x) > 0$. Если

$P(n, x, x) = 0$ для всех $n = 1, 2, \ldots$, то будем считать $d(x) = \infty$. Можно показать, что в каждом классе сообщающихся состояний $d(x)$ постоянно. Общее значение $d = d(x)$ для x из данного класса называется *периодом* этого класса. Класс называется *апериодичным*, если его период $d = 1$. При $d > 1$ класс называется *периодичным*. В соответствии с этим неприводимая цепь называется *периодичной* или *апериодичной* в зависимости от того, будет ли ее период больше единицы или равен единице.

Следующая теорема показывает, как происходит движение в периодичном классе сообщающихся состояний.

Теорема 1. *Каждый класс K сообщающихся состояний периода d ($d < \infty$) можно разбить на d попарно не пересекающихся подмножеств $K_0, K_1, \ldots, K_{d-1}$ так, чтобы за один шаг из K_i ($i = 0, 1, \ldots, d-2$) система могла перейти лишь в состояния множества K_{i+1}, а из K_{d-1} — только в состояния множества K_0. При этом, если $x \in K_i$, $y \in K_r$, то существует такое $n_0 = n_0(x, y)$, что при всех $n > n_0$*

$$P(nd + r - i, x, y) > 0.$$

В частности, для всех $n > n_0 = n_0(x, x)$ $P(nd, x, x) > 0$.

Множества K_i ($i = 0, 1, 2, \ldots, d-1$) называются *подклассами* периодического класса сообщающихся состояний.

Пример. Пусть X — целочисленная решетка на прямой и система движется так, что из состояния $x \in X$ за один шаг возможны лишь переходы в состояние $x + 1$ с вероятностью p и в состояние $x - 1$ с вероятностью q, $p + q = 1$ $(p, q < 0)$. Такая цепь неприводима и периодична с периодом 2. Подклассы K_0 и K_1 — это соответственно совокупности четных и нечетных чисел.

8.3.4. Возвратность. Пусть τ_y — момент первого после нуля достижения состояния $y \in X$, т. е. $\tau_y = \inf\{n\colon n = 1, 2, \ldots, \xi_n = y\}$, причем в случае $\xi_n \neq y$ для всех $n = 1, 2, \ldots$ полагаем $\tau_y = +\infty$. Положим $f(0, x, y) = 0$ и $f(n, x, y) = \mathsf{P}_x\{\tau_y = n\}$ $(n = ., 2, \ldots$ $\ldots, x, y \in X)$. Обозначим

$$F(x, y) = \sum_{n=0}^{\infty} f(n, x, y) = \mathsf{P}_x\{\tau_y < \infty\}.$$

$F(x, y)$ есть вероятность того, что система, выйдя из состояния x, когда-либо попадает в состояние y; при $x = y$ $F(x, x)$ есть вероятность того, что система, выйдя из состояния x, когда-либо возвратится в это состояние.

Состояние x называется *возвратным*, если $F(x, x) = 1$, и *невозвратным*, если $F(x, x) < 1$.

Связь между вероятностями перехода и вероятностями первого достижения задается формулой

$$P(n, x, y) = \sum_{k=0}^{n} f(k, x, y)\, P(n-k, y, y), \quad n = 1, 2, \ldots, x, y \in X$$

Полагая для $\lambda \in [0, 1)$

$$P_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n P(n, x, y), \quad F_\lambda(x, y) = \sum_{n=0}^{\infty} \lambda^n f(n, x, y),$$

из предыдущей формулы получаем соотношения

$$P_\lambda(x, x) = \frac{1}{1 - F_\lambda(x, x)}, \quad P_\lambda(x, y) = F_\lambda(x, y)\, P_\lambda(y, y), \quad x \neq y.$$

Эти формулы приводят к следующему результату.

Теорема 2. *Состояние x возвратно тогда и только тогда, когда*

$$G(x, x) = \sum_{n=0}^{\infty} P(n, x, x) = \infty,$$

и невозвратно, когда

$$G(x, x) = \sum_{n=0}^{\infty} P(n, x, x) < \infty.$$

В невозвратном случае

$$G(x, x) = \frac{1}{1 - F(x, x)}.$$

Далее можно показать, что сообщающиеся состояния возвратны или невозвратны одновременно, так что свойство возвратности является свойством класса сообщающихся состояний. Поэтому можно говорить о возвратности или невозвратности неприводимой цепи Маркова.

Следующая теорема дает критерий возвратности неприводимой цепи Маркова в терминах эксцессивных функций.

Теорема 3. *Неприводимая цепь Маркова возвратна в том и только том случае, когда всякая эксцессивная функция постоянна.*

Более точно, если цепь неприводима и возвратна, то единственной эксцессивной функцией (с точностью до постоянного неотрицательного множителя) является функция, тождественно равная единице. Это означает, что система неравенств

$$\sum_{y \in X} P(x, y)\, \varphi(y) \leqslant \varphi(x), \quad x \in X,$$

не имеет неотрицательных решений, отличных от решений вида $\varphi(x) = \mathrm{const}$.

Наоборот, если цепь (не обязательно неприводимая) имеет хотя бы одно невозвратное состояние, то всегда существует непостоянная эксцессивная функция; например,

$$\varphi(y) = \begin{cases} 1, & \text{если} \quad y = x_0, \\ F(y, x_0), & \text{если} \quad y \neq x_0, \end{cases}$$

где x_0 — фиксированное невозвратное состояние.

8.3.5. Свойства возвратных состояний. Обозначим через $q(x, y)$ $(x, y \in X)$ вероятность того, что система, выйдя из состояния x, попадет в состояние y бесконечно много раз. Используя свойство строгой марковости цепи, можно показать, что $q(x, y) = F(x, y)$, если только состояние y возвратно. В частности, $q(y, y) = 1$ для всякого возвратного состояния y. Другими словами, вероятность того, что система побывает бесконечно много раз в возвратном состоянии y, выходя из x, равна вероятности того, что y будет когда-либо достигнуто из состояния x.

Более того, если состояние x возвратно и $F(x, y) > 0$, то $q(x, y) = 1$, т. е., выходя из возвратного состояния x, система обязана посетить достижимое из x состояние y бесконечно много раз. При этом $F(y, x) > 0$. В частности, если x возвратно и y достижимо из x, то y достижимо из x с вероятностью 1 ($F(x, y) = 1$).

Далее, если y невозвратно, то $q(x, y) = 0$ для всех $x \in X$ и, в частности, $q(y, y) = 0$. Это означает, что невозвратные состояния система посещает лишь конечное число раз (последнее, разумеется, следует также и из теоремы 2).

Таким образом, из возвратных состояний достижимы лишь возвратные состояния. Возвратные состояния существенны.

8.3.6. Предельное поведение вероятностей перехода.

Теорема 4. *Для $x, y \in X$ справедливо соотношение*

$$F(x, y) = \lim_{N \to \infty} \sum_{n=1}^{N} P(n, x, y) \Big/ \sum_{n=0}^{N} P(n, y, y).$$

Из этой формулы следует, что для всех x и y из одного возвратного класса $G(x, y) = \sum_{n=0}^{\infty} P(n, x, y) = +\infty$. Если же y невозвратно, то $G(x, y) < \infty$ при всех $x \in X$.

Более точные результаты могут быть получены применением теоремы восстановления. Обозначим через τ_y^1 момент первого после нуля достижения состояния y (см. п. 8.3.4). Далее, положим

$$\tau_y^{k+1} = \inf\{n\colon n > \tau_y^k,\ \xi_n = y\}, \qquad k = 1, 2, 3, \ldots,$$

причем полагаем $\tau_y^{k+1} = +\infty$, если $\xi_n \neq y$ для всех $n > \tau_y^k$. Из свойства строгой марковости цепи следует, что величины τ_y^1, $\tau_y^2 - \tau_y^1$, $\tau_y^3 - \tau_y^2$, ... независимы и одинаково распределены по мере P_y, если только y — возвратное состояние. Если начальным является состояние x, $F(x, y) = 1$ и y возвратно, то относительно меры P_x все эти величины независимы между собой и, за исключением первой, одинаково распределены. Это обстоятельство позволяет применить теорему восстановления для изучения предельного поведения вероятностей перехода $P(n, x, y)$ при $n \to \infty$.

Обозначим через m_y среднее число шагов до первого возвращения в состояние y, т. е.

$$m_y = \sum_{n=1}^{\infty} n f(n, y, y) = \mathsf{M}_y \tau_y^1 \leqslant +\infty.$$

Теорема 5. а) *Если состояние y невозвратно, то при всех $x \in X$*

$$\lim_{n \to \infty} P(n, x, y) = 0;$$

б) *если x и y лежат в различных классах существенных состояний то при всех n $P(n, x, y) = 0$;*

в) *если x и y принадлежат одному классу существенных состояний с периодом d, причем $x \in K_l$, $y \in K_r$, где K_l — подклассы, введенные в теореме* 1, *то*

$$\lim_{n\to\infty} P(nd+l, x, y) = d/m_y$$

при условии $l \equiv r - i \pmod d$ *и*

$$P(nd+l, x, y) = 0$$

при условии $l \not\equiv r - i \pmod d$; *в частности, если* $d = 1$, *т. е. класс апериодичен,* $\lim\limits_{n\to\infty} P(n, x, y) = 1/m_y$;

г) *если состояние x несущественно, а y принадлежит классу существенных состояний с периодом d, то для всех* $r = 0, 1, 2, \ldots, d-1$

$$\lim_{n\to\infty} P(nd+r, x, y) = F_r(x, y)\frac{d}{m_y},$$

где

$$F_r(x, y) = \sum_{\substack{n=1 \\ n\equiv r \pmod d}}^{\infty} f(n, x, y).$$

Заметим, что $F_r(x, y) \geqslant 0$ и $\sum\limits_{r=0}^{d-1} F_r(x, y) = F(x, y) \leqslant 1$. Кроме того, $F_r(x, y) = \mathsf{P}_x\{\xi_n = y$ для некоторого $n \equiv r \pmod d$, $n > 0\}$.

Из этой теоремы следует также несколько более грубый результат о сходимости средних по Чезаро.

Следствие. *Если считать, что для невозвратных состояний* y $m_y = \infty$, *то для всех* $x, y \in X$ *существует предел*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N} P(n, x, y) = \pi(x, y),$$

причем

$$\pi(x, y) = \frac{F(x, y)}{m_y}.$$

8.3.7. Положительные и нулевые классы. Возвратное состояние y называется *нулевым*, если $\lim\limits_{n\to\infty} P(nd(y), y, y) = 0$, и *положительным*, если $\lim\limits_{n\to\infty} P(nd(y), y, y) > 0$. Легко установить, что в возвратном классе состояний либо все состояния одновременно положительные, либо нулевые. Если y — положительное состояние, то $m_y < \infty$ и $\pi(y) = \pi(y, y) = 1/m_y$. Для нулевого состояния $\pi(y) = 0$.

Оказывается, свойство класса быть положительным или нулевым тесно связано с проблемой существования инвариантных мер.

Для марковских цепей со счетным или конечным множеством состояний естественно рассматривать меры (заряды) лишь на одноточечных множествах: $\mu(y) = \mu(\{y\})$, ибо для $\Gamma \subset X$ имеем

$$\mu(\Gamma) = \sum_{y\in\Gamma} \mu(y).$$

Заряд $\mu(y)$, заданный для $y \in K \subset X$, называется *инвариантным на множестве* K, если $\mu(y) = \sum_{x \in K} \mu(x) P(x, y)$ для всех $y \in K$.

Теорема 6. *Если K — класс существенных состояний, то всякий инвариантный на множестве K заряд μ, удовлетворяющий условию*

$$\sum_{y \in K} |\mu(y)| < \infty, \tag{3.2}$$

имеет вид $\mu(y) = c\pi(y)$, где $y \in K$, а c — произвольная постоянная.

Следствие. 1) *Если K — положительный возвратный класс, то единственным инвариантным на K зарядом, удовлетворяющим условию (3.2) и условию*

$$\sum_{y \in K} \mu(y) = 1, \tag{3.3}$$

является мера

$$\pi(y) = \lim_{n \to \infty} \frac{1}{N} \sum_{n=0}^{N} P(n, y, y), \quad y \in K.$$

При этом, если d — период класса K, то для любого подкласса K_j ($j = 0, 1, 2, \ldots, d-1$) имеем

$$\sum_{y \in K_j} \pi(y) = \frac{1}{d}.$$

2) *Если K — нулевой возвратный класс, то единственный инвариантный на K заряд, подчиненный условию (3.2), тривиален ($\mu(y) = 0$, $y \in K$).*

3) *Если марковская цепь произвольна, а $\mu(y)$ ($y \in X$) — абсолютно суммируемое решение системы уравнений*

$$\mu(y) = \sum_{x \in X} \mu(x) P(x, y), \quad y \in X, \tag{3.4}$$

то для всякого невозвратного состояния y_0 должно быть $\mu(y_0) = 0$.

4) *Для того чтобы неприводимая цепь Маркова была положительно возвратной, необходимо и достаточно, чтобы система уравнений (3.4) имела нетривиальное абсолютно суммируемое решение.*

5) *Неприводимая цепь Маркова имеет стационарное распределение тогда и только тогда, когда она положительно возвратна.*

6) *Если цепь неприводима, положительно возвратна и апериодична, то единственным решением системы (3.4), удовлетворяющим условиям (3.2) и (3.3), является мера*

$$\mu(y) = \lim_{n \to \infty} P(n, x, y).$$

Следующая теорема описывает все возможные стационарные распределения для данной цепи (если таковые имеются).

Теорема 7. *Пусть (ξ_n, P_x) — однородная цепь Маркова с дискретным множеством состояний. Обозначим через D_α ($\alpha \in A$) поло-*

жительные возвратные классы, причем $D_\alpha \neq D_\beta$ *при* $\alpha \neq \beta$, *где* A — *счетный набор индексов. Положим* $D = \bigcup_{\alpha \in A} D_\alpha$. *Мера* $\mu(x)$ $(x \in X)$ *стационарна тогда и только тогда, когда существует последовательность величин* $\{\lambda_\alpha,\ \alpha \in A\}$, $\lambda_\alpha \geqslant 0$ *и* $\sum_{\alpha \in A} \lambda_\alpha = 1$ *такая, что*

$$\mu(x) = \begin{cases} 0, & \text{если} \quad x \notin D, \\ \lambda_\alpha \pi(x), & \text{если} \quad x \in D_\alpha, \quad \alpha \in A. \end{cases}$$

8.3.8. Вероятности с запрещением. Пусть Γ — некоторое множество состояний, $\Gamma \subset X$. Положим для $x,\ y \in X$ и $n = 1, 2, \ldots$

$$_\Gamma P(n, x, y) = \sum_{z_1, z_2, \ldots, z_{n-1} \in X \setminus \Gamma} P(x, z_1) P(z_1, z_2) \ldots P(z_{n-1}, y).$$

Так определенная величина задает вероятность того, что система за n шагов перейдет из состояния x в состояние y, не побывав ни в один из моментов времени $1, 2, \ldots, n-1$ в состояниях множества Γ. Такие вероятности называются *вероятностями с запрещением* или *табу-вероятностями*. Если $\Gamma = \{z\}$, то будем писать $_zP(n, x, y)$ вместо $_{\{z\}}P(n, x, y)$. Очевидно, $f(n, x, y) = {_y}P(n, x, y)$.

Для всех $n = 1, 2, \ldots,$ $x, y, z \in X$, $\Gamma \subset X$, $z \notin \Gamma$, имеют место равенства

$$_\Gamma P(n, x, y) = {_{z,\Gamma}}P(n, x, y) + \sum_{k=1}^{n-1} {_{z,\Gamma}}P(k, x, z)\, {_\Gamma}P(n-k, z, y),$$

$$_\Gamma P(n, x, y) = {_{z,\Gamma}}P(n, x, y) + \sum_{k=1}^{n-1} {_\Gamma}P(k, x, z)\, {_{z,\Gamma}}P(n-k, z, y),$$

где $_{z,\Gamma}P(r, x, y) = {_{\{z\} \cup \Gamma}}P(r, x, y)$.

Положим $_\Gamma P(0, x, y) = \chi_{\{y\}}(x)$ при $x \notin \Gamma$ и $_\Gamma P(0, x, y) = 0$ при $x \in \Gamma$, где $\chi_B(x)$ — индикатор множества $B \subset X$. Далее, если $\Gamma = \{y, z\}$, то естественно обозначить $_zf(n, x, y) = {_{\{z, y\}}}P(n, x, y)$ для $n = 1, 2, \ldots,$ $x, y, z \in X$. Это есть вероятность того, что система, выйдя из состояния x, впервые на n-м шаге окажется в состоянии y, не заходя при этом в моменты времени $1, 2, \ldots, n-1$ в состояние z. Положим еще $f(0, x, y) = 0$. С учетом этих обозначений и соглашений из предыдущих равенств легко получить формулы

$$_zP(n, x, y) = \sum_{k=0}^{n} {_z}f(k, x, y)\, {_z}P(n-k, y, y) + \delta_{n0}\chi_{\{y\}}(x), \qquad z \neq y,$$

$$f(n, x, y) = \sum_{k=0}^{n} {_y}P(k, x, x)\, {_x}f(n-k, x, y), \quad x \neq y,$$

справедливые при $n = 0, 1, 2, \ldots,$ где $\chi_\Gamma(x)$ — индикатор множества Γ, а $\delta_{n0} = 1$ при $n = 0$ и $\delta_{n0} = 0$ при $n \neq 0$. Отсюда, полагая

$$_zG(x, y) = \sum_{n=0}^{\infty} {_z}P(n, x, y), \quad _zF(x, y) = \sum_{n=0}^{\infty} {_z}f(n, x, y),$$

находим

$$_zG(x, y) = \chi_{\{y\}}(x) + {_zF}(x, y)\,{_zG}(y, y), \quad z \neq y,$$
$$F(x, y) = {_yG}(x, x)\,{_xF}(x, y), \quad x \neq y.$$

Если x и y — сообщающиеся состояния, то очевидно, что $_xF(x, y) > 0$. Поэтому из второго соотношения имеем

$$0 < {_yG}(x, x) = \frac{F(x, y)}{_xF(x, y)} < \infty, \quad x \neq y,$$

если только x и y сообщаются. Теперь из первого соотношения для сообщающихся состояний y и z получаем

$$0 < {_zG}(z, y) = F(y, z)\,\frac{_zF(z, y)}{_yF(y, z)} < \infty, \quad y \neq z.$$

Далее, можно показать, что для x и y из одного возвратного класса справедливо соотношение

$$_xG(x, y) = \lim_{N \to \infty}\left[\sum_{n=0}^{N} P(n, x, y) \Big/ \sum_{n=0}^{N} P(n, x, x)\right].$$

Отсюда следует, что для возвратных состояний $_xG(x, x) = 1$. Кроме того, если x и y — состояния из одного положительного класса, то $_xG(x, y) = \pi(y)/\pi(x)$.

Обозначим $m_{xy} = \mathsf{M}_x\tau_y^1 = \sum\limits_{n=1}^{\infty} nf(n, x, y)$. Величина m_{xy} есть среднее время до первого достижения состояния y, если начальным было состояние x. При $x = y$ величина m_{yy} совпадает с введенной ранее величиной m_y.

Теорема 8. *Если* $F(x, y) = 1$, *то* $\sum\limits_{z \in X} {_yG}(x, z) = m_{xy}$.

Из этой теоремы следует, что ряд $\sum\limits_{y} {_xG}(x, y)$ сходится для положительно возвратного класса и расходится для нулевого класса.

Как следует из теоремы 6, в случае, когда K — нулевой возвратный класс, не существует инвариантного на K заряда ограниченной вариации, отличного от тривиального заряда. Следующая теорема показывает, что в этом случае существует инвариантная на K мера бесконечной полной массы.

Теорема 9. *Пусть K — возвратный класс. Единственным неотрицательным решением системы уравнений*

$$\mu(y) = \sum_{x \in K} \mu(x) P(x, y), \quad y \in K,$$

удовлетворяющим условию $\mu(z_0) = 1$ *для некоторого* $z_0 \in K$, *является мера* $_{z_0}G(z_0, y)$ $(y \in K)$.

8.3.9. Эргодическая теорема. Пусть (ξ_n, P_x) — однородная цепь Маркова, состояния которой образуют один возвратный класс X, Предположим, что на X задана вещественная функция $v(x)$, и рассмотрим функционал

$$\eta_n(v)=\sum_{k=0}^{n} v(\xi_k), \quad n=0, 1, 2, \ldots$$

Для $y \in X$ положим $\tau_y^1=\inf\{n: n=1, 2, \ldots, \xi_n=y\}$,

$$\tau_y^{k+1}=\inf\{n: n>\tau_y^k, \xi_n=y\}, \quad k=1, 2, \ldots$$

(см. п. 8.3.6). Так как все состояния рассматриваемой цепи образуют один возвратный класс, то при всех $x, y \in X$ и $k=1, 2, \ldots$, $\mathsf{P}_x\{\tau_y^k<\infty\}=1$. Величину $\eta_n(v)$ можно представить в виде

$$\eta_n(v)=\sum_{k=0}^{\tau_y^1-1} v(\xi_k)+\sum_{k=1}^{\nu_y(n)-1} \zeta_k(y, v)+\sum_{k=\tau_y^{\nu_y(n)}}^{n} v(\xi_k),$$

где

$$\zeta_k(y, v)=\sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} v(\xi_r), \quad k=1, 2, \ldots,$$

а $\nu_y(n)$ — случайная величина, для которой $\tau_y^{\nu_y(n)} \leqslant n$, $\tau_y^{\nu_y(n)+1}>n$ (иначе: $\nu_y(n)=\max\{k: \tau_y^k \leqslant n\}=\min\{k: \tau_y^{k+1}>n\}$). Заметим, что в силу строгой марковости цепи величины $\zeta_k(y, v)$ $(k=1, 2, \ldots)$ независимы и одинаково распределены, причем распределение величины $\zeta_k(y, v)$ по мере P_x не зависит от x. Действительно, для вещественных α имеем

$$\mathsf{P}_x\{\zeta_k(y, v)<\alpha\}=\mathsf{M}_x\mathsf{P}_x\{\theta_{\tau_y^k}\zeta_1(y, v)<\alpha \mid \mathfrak{F}_{\tau_y^k}\}=$$
$$=\mathsf{M}_x\mathsf{P}_{\xi_{\tau_y^k}}\{\zeta_1(y, v)<\alpha\}=\mathsf{P}_y\{\zeta_1(y, v)<\alpha\}.$$

Далее, можно показать, что из существования момента p-го порядка величины $\zeta_k(y, v)$ при некотором $y \in X$ следует существование его при любом $y \in X$. В частности, если $p=1$, то при условии

$$\sum_{z \in X} {}_yG(y, z)\,|v(z)|<\infty$$

имеем

$$\mathsf{M}_x\{\zeta_k(y, v)\}=\sum_{z \in X} {}_yG(y, z)\, v(z).$$

Обозначим

$$S_x(v)=\sum_{y \in X} {}_xG(x, y)\, v(y).$$

Оказывается, если $S_x(v)$ конечно (под этим будем понимать условие $\sum_{y \in X} {}_xG(x, y)|v(y)|=S_x(|v|)<\infty$) при некотором $x \in X$, то оно конечно и при всех x. Аналогично, если $S_x(v) \neq 0$ при некотором x, то это же справедливо и для всех x. Кроме того, если $S_x(u)$ и $S_x(v)$ конечны, то отношение $S_x(u)/S_x(v)$ не зависит от x. В случае положительно возвратной цепи ${}_xG(x, y) = \pi(y)/\pi(x)$, так что в этом случае

$$\pi(x)\, S_x(v) = \sum_{y \in X} \pi(y)\, v(y),$$

где $\pi(y)$ — стационарное распределение цепи.

Следующая теорема носит название *эргодической теоремы*. В ее доказательстве решающую роль играет представление величины $\eta_n(v)$ в виде суммы независимых одинаково распределенных случайных величин и некоторых добавок (см. выше).

Теорема 10. *Если цепь* (ξ_n, P_x) *неприводима и возвратна, а заданные на пространстве состояний функции* u *и* v *таковы, что величины* $S_x(u)$ *и* $S_x(v)$ *конечны и не равны нулю, то справедливо соотношение*

$$\lim_{N \to \infty} \left[\sum_{n=0}^{N} u(\xi_n) \Big/ \sum_{n=0}^{N} v(\xi_n) \right] = S_y(u)/S_y(v)$$

почти наверное относительно меры P_x *при любом* $x \in X$ (правая часть этого равенства, как отмечалось выше, не зависит от y).

Следствие. 1) *Если неприводимая цепь положительно возвратна, а функция* $v(x)$ $(x \in X)$ *удовлетворяет условию* $\sum_{y \in X} \pi(y)\, |v(y)| < \infty$, *то* P_x*-почти наверное (при всех* $x \in X$*)*:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N} v(\xi_n) = \sum_{y \in X} \pi(y)\, v(y).$$

Это означает, что среднее по времени для последовательности $\{v(\xi_n),\ n = 0, 1, 2, \ldots\}$ сходится к среднему от функции $v(x)$ по стационарному распределению.

2) *Если для неприводимой положительно возвратной цепи* $\sum_{y \in X} \pi(y)\, |v(y)| < \infty$, *то при всех* $x \in X$

$$\lim_{N \to \infty} \mathsf{M}_x \left\{ \left| \frac{1}{N} \sum_{n=0}^{N} v(\xi_n) - \sum_{y \in X} \pi(y)\, v(y) \right| \right\} = 0.$$

3) *Для неприводимой положительно возвратной цепи* P_x*-почти наверное (при всех* $x \in X$*)*

$$\lim_{n \to \infty} \frac{\nu_y(n)}{n} = \pi(y), \qquad \lim_{n \to \infty} \frac{\tau_y^{\nu_y(n)}}{n} = 1,$$

где $\nu_y(n)$ *и* τ_y^k *— величины, определенные выше.*

8.3.10. Центральная предельная теорема для цепей Маркова. Предположим, что цепь Маркова (ξ_n, P_x) неприводима и положительно возвратна, а $v(x)$ — вещественная функция, заданная на состояниях цепи. Выше отмечалось, что для существования среднего $\mathsf{M}_x\zeta_k(y, v)$ достаточно, чтобы

$$\sum_{y \in X} |v(y)| \pi(y) < \infty,$$

где $\pi(y)$ — стационарное распределение. При этом условии

$$\mathsf{M}_x\zeta_k(y, |v|) = \mathsf{M}_x \sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} |v(\xi_r)|.$$

Предположим теперь, что для функции v выполняется несколько менее жесткое условие, а именно

$$\mathsf{M}_x |\zeta_k(y, v)| = \mathsf{M}_x \left| \sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} v(\xi_r) \right| < \infty. \tag{3.5}$$

Положим для $y \in X$

$$\mu_y(v) = \mathsf{M}_x \sum_{r=\tau_y^k}^{\tau_y^{k+1}-1} v(\xi_r).$$

Величина $\mu_y(v)$ не зависит ни от $x \in X$, ни от $k = 1, 2, \ldots$

Теорема 11. *Если неприводимая положительно возвратная цепь Маркова и функция v таковы, что $\mu_y(v)$ существует* (т. е. *выполнено условие* (3.5)), *то предел по вероятности P_x $(x \in X)$ величины $\frac{1}{N}\sum_{n=0}^{N} v(\xi_n)$ существует и равен $\pi(y)\mu_y(v)$ $(y \in X)$.*

Отсюда, в частности, следует, что величина $\pi(y)\mu_y(v)$ не зависит от y.

Положим теперь для $y \in X$, $k = 1, 2, \ldots$

$$\delta_k(y, v) = \zeta_k(y, v) - \pi(y)\mu_y(v)\left(\tau_y^{k+1} - \tau_y^k\right).$$

Величины $\delta_k(y, v)$ $(k = 1, 2, \ldots)$ независимы и одинаково распределены по мере P_x $(x \in X)$, причем распределение $\mathsf{P}_x\{\delta_k(y, v) < \alpha\}$ не зависит от x. Очевидно, что $\mathsf{M}_x\delta_k(y, v) = 0$. Обозначим

$$\sigma_y^2(v) = \mathsf{M}_x[\delta_k(y, v)]^2.$$

Можно показать, что если $\sigma_y^2(v) < \infty$ для некоторого $y \in X$, то это справедливо и для всех $y \in X$.

Теорема 12. *Если для неприводимой положительно возвратной цепи Маркова и функции v выполнено $0 < \sigma_y^2(v) < \infty$, то справедливо соотношение*

$$\lim_{n\to\infty} \mathsf{P}_x\left\{\frac{\eta_n(v) - na}{\sqrt{bn}} < \alpha\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\beta^2/2}\, d\beta,$$

где $x \in X$, α — *произвольное вещественное число,* $a = \pi(y)\mu_y(v)$, $b = \pi(y)\sigma_y^2(v)$. *Величины* a *и* b *не зависят от выбора* $y \in X$.

Из теоремы 11 следует, что $\frac{1}{n}\eta_n(v) \to a$ по вероятности P_x. Теорема 12, таким образом, показывает, что флуктуации величины $\frac{1}{n}\eta_n(v)$ вокруг среднего значения a распределены асимптотически нормально, если только существует второй момент величины $\delta_k(y, v)$ и он отличен от нуля.

Замечание. Величину $\eta_n(v)$ можно представить в виде

$$\eta_n(v) = a\left(\tau_y^{\nu_y(n)} - \tau_y^1\right) + \sum_{k=1}^{\nu_y(n)-1} \delta_k(y, v) + V_n' + V_n'',$$

где V_n' и V_n'' совпадают соответственно с первым и третьим слагаемыми в представлении для величины $\eta_n(v)$, приведенном в п. 8.3.9. Обычный ход рассуждений при доказательстве предельных теорем для величины $\eta_n(v)$ состоит в доказательстве асимптотической малости величин V_n' и V_n'' (с определенным нормировочным множителем) и применении классических предельных теорем для сумм независимых случайных величин к сумме

$$\sum_{k=1}^{\nu_y(n)-1} \delta_k(y, v).$$

Предположение теоремы 12 означает, что распределение величины $\delta_k(y, v)$ принадлежит области притяжения нормального закона, и потому в пределе здесь получается нормальное распределение. Предполагая, что распределение величины $\delta_k(y, v)$ попадает в область притяжения других устойчивых законов, можно получить другие предельные теоремы для величин $\eta_n(v)$.

Литература: [19, 28, 31, 40, 63, 73, 78, 83, 89, 97, 100, 112, 117].

Часть вторая

ТЕОРИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

Глава 9. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

9.1. Определение случайного процесса

9.1.1. Конечномерные распределения. *Случайным процессом* на вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ называется семейство случайных величин $\xi(t, \omega)$, зависящих от вещественного параметра t, принимающего значения из некоторого множества T. Это множество называется *областью определения* процесса. Сами случайные величины $\xi(t, \omega)$ могут быть либо вещественными, либо комплексными, либо векторными. То пространство X, в котором $\xi(t, \omega)$ принимает свои значения, называется *фазовым пространством процесса.* В зависимости от фазового пространства процесса говорят о числовых, комплекснозначных или векторных процессах. Как и при обозначении случайных величин, для процессов аргумент ω часто опускают и пишут $\xi(t)$ вместо $\xi(t, \omega)$. Одной из основных характеристик случайного процесса являются его *конечномерные* (*частные*) *распределения* — набор функций, определенных для каждого натурального k соотношениями

$$F_{t_1, t_2, \ldots, t_k}(A_1, A_2, \ldots, A_k) = \mathsf{P}\left\{\bigcap_{j=1}^{k} \{\xi(t_j, \omega) \in A_j\}\right\},$$

где $t_1, t_2, \ldots, t_k \in T$; $A_1, A_2, \ldots, A_k$ — борелевские множества из области значений процесса.

Конечномерные распределения удовлетворяют следующим очевидным условиям:

1) при фиксированных $t_1, t_2, \ldots, t_k$ функция $F_{t_1, t_2, \ldots, t_k}(A_1, \ldots, A_k)$ является совместным распределением k случайных величин;

2) $F_{t_1, \ldots, t_k}(A_1, \ldots, A_k) = F_{t_{i_1}, \ldots, t_{i_k}}(A_{i_1}, \ldots, A_{i_k})$, какова бы ни была перестановка $i_1, \ldots, i_k$ чисел $1, 2, \ldots, k$;

3) если X — область значений процесса, то

$$F_{t_1, \ldots, t_{k-1}, t_k}(A_1, \ldots, A_{k-1}, X) = F_{t_1, \ldots, t_{k-1}}(A_1, \ldots, A_{k-1})$$

Конечномерные распределения $F_{t_1, \ldots, t_k}(A_1, \ldots, A_k)$ могут задаваться конечномерными плотностями распределения — такими функциями $f_{t_1, \ldots, t_k}(x_1, \ldots, x_k)$, что

$$F_{t_1, \ldots, t_k}(A_1, \ldots, A_k) = \int\limits_{A_1} \ldots \int\limits_{A_k} f_{t_1, \ldots, t_k}(x_1, \ldots, x_k) \times \\ \times dx_1 \ldots dx_k.$$

Ответ на вопрос, при каких условиях существует случайный процесс, для которого данные функции $F_{t_1, \ldots, t_k}(A_1, \ldots, A_k)$ являются конечномерными распределениями, дает следующая теорема.

Теорема 1 (Колмогорова). *Пусть функции* $F_{t_1, \ldots, t_k}(A_1, \ldots, A_k)$ *определены при* $t_1, \ldots, t_k \in T$, $A_1, \ldots, A_k \in \mathfrak{B}(X)$ (σ-*алгебра борелевских множеств в конечномерном евклидовом пространстве* X). *Тогда для того, чтобы существовал случайный процесс, для которого* $F_{t_1, \ldots, t_k}(A_1, \ldots, A_k)$ *являлись бы конечномерными распределениями, необходимо и достаточно, чтобы выполнялись условия* 1) — 3). *В качестве вероятностного пространства можно взять пространство* $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, *где* Ω — *множество всех функций* $\omega(t)$, *определенных на* T, *принимающих значения из* X; σ-*алгебра* $\mathfrak{S}$ — *это* σ-*алгебра, порожденная цилиндрическими множествами, т. е. множествами вида*

$$\{\omega\colon\ \omega(t_1) \in A_1, \ \ldots, \ \omega(t_k) \in A_k\} = C_{t_1, \ldots, t_k}(A_1, \ldots, A_k),$$

а мера P *определяется соотношением*

$$\mathsf{P}(C_{t_1, \ldots, t_k}(A_1, \ldots, A_k)) = F_{t_1, \ldots, t_k}(A_1, \ldots, A_k).$$

Искомый случайный процесс на этом вероятностном пространстве определяется равенством

$$\xi(t, \omega) = \omega(t).$$

Функции $\xi(t, \omega)$ при фиксированном ω называются *выборочными функциями* (или *траекториями*, или *реализациями*) случайного процесса.

Предложенная в теореме Колмогорова конструкция построения случайного процесса с заданными конечномерными распределениями приводит к слишком широкому пространству выборочных функций. Иногда желательно строить процесс, у которого выборочные функции обладают некоторыми свойствами регулярности (например, измеримы, непрерывны, дифференцируемы и т. д.).

Два случайных процесса называются *стохастически эквивалентными в широком смысле*, если совпадают их конечномерные распределения.

Теорема 2. *Для того чтобы для данного процесса существовал стохастически эквивалентный в широком смысле процесс, выборочные функции которого принадлежат множеству* $F \subset \Omega$, *необходимо и достаточно, чтобы* $\mathsf{P}^*(F) = 1$, *где* P^* — *внешняя мера, построенная по мере* P, *определенной в теореме Колмогорова*:

$$\mathsf{P}^*(G) = \inf_{G \subset \bigcup_k C_k} \sum_k \mathsf{P}(C_k),$$

где C_k — *цилиндрические множества и* G — *произвольное множество из* Ω.

При выполнении этого условия в качестве вероятностного пространства, на котором задан процесс, можно взять $\{F, \mathfrak{S}^*, \mathsf{P}^*\}$; здесь

$\mathfrak{S}^*$ — σ-алгебра подмножеств Ω вида $F \cap C$, где $C \in \mathfrak{S}$, а сам процесс по-прежнему определяется соотношением

$$\xi(t, \omega) = \omega(t).$$

Пусть $\xi(t, \omega)$ $(t \in T)$ — некоторый случайный процесс (вещественный, комплексный или векторный), определенный на произвольном вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$. Если F_T — пространство всех функций с теми же значениями, что и $\xi(t, \omega)$, а $\mathfrak{S}_T$ — наименьшая σ-алгебра, содержащая все цилиндрические множества из F_T, то отображение

$$\omega \to \xi(\cdot, \omega)$$

является измеримым отображением пространства $\{\Omega, \mathfrak{S}\}$ в $\{F_T, \mathfrak{S}_T\}$, т. е. для всякого $A \in \mathfrak{S}_T$ $\{\omega: \xi(\cdot, \omega) \in A\} \in \mathfrak{S}$. Это отображение переводит меру P в некоторую меру μ_ξ:

$$\mu_\xi(A) = \mathsf{P}(\{\omega: \xi(\cdot, \omega) \in A\}), \quad A \in \mathfrak{S}_T.$$

Мера $\mu_\xi(\cdot)$ называется мерой, *соответствующей случайному процессу* $\xi(t, \omega)$. Она совпадает с мерой, построенной по конечномерным распределениям в теореме Колмогорова.

Конечномерные распределения процесса $\xi(t, \omega)$ удобно задавать с помощью *характеристического функционала* процесса:

$$\chi(g) = \mathsf{M} \exp\left\{i \int (\xi(t, \omega), \quad dg(t))\right\},$$

определенного для всех ступенчатых функций g на T со значениями в X; (ξ, dg) — скалярное произведение в X; интеграл в показателе экспоненты — интеграл Стилтьеса.

9.1.2. Моментные функции. Пусть $\xi(t, \omega)$ — числовой случайный процесс, для которого $\mathsf{M}|\xi(t, \omega)|^m < \infty$. Тогда при $k \leqslant m$ определены функции

$$m_k(t_1, \ldots, t_k) = \mathsf{M}\xi(t_1, \omega) \ldots \xi(t_k, \omega).$$

Функция $m_k(t_1, \ldots, t_k)$ называется *k-й моментной функцией* процесса $\xi(t, \omega)$. Если $\mathsf{M}|\xi(t, \omega)|^k < \infty$, $t \in T$, для всех k, то для процесса определены моментные функции всех порядков.

Среди моментных функций наиболее употребительны функции первых двух порядков: $m_1(t)$ — среднее значение процесса; вместо $m_2(t_1, t_2)$ обычно рассматривают функцию $R(t_1, t_2) = m_2(t_1, t_2) - m_1(t_1)m_1(t_2)$, которая называется *корреляционной функцией*. Среднее значение может быть любой функцией, определенной на T. Корреляционная функция $R(t_1, t_2)$ положительно определена: для всех $t_1, t_2, \ldots, t_n$ из T и вещественных $x_1, x_2, \ldots, x_n$

$$\sum_{i, k} R(t_i, t_k) x_i x_k \geqslant 0.$$

Всякая положительно определенная функция $R(t_1, t_2)$ является корреляционной функцией некоторого процесса.

Пусть $\xi(t, \omega)$ — случайный процесс со значениями в конечномерном евклидовом пространстве X. Функция $a(t)$, определенная на T и принимающая значения из X, называется *средним значением процесса*, если для всех $z \in X$

$$\mathsf{M}\,(\xi\,(t,\ \omega),\ z) = (a\,(t),\ z).$$

Функция $B(t, s)$, определенная при $t,\ s \in T$, значениями которой служат линейные операторы в X, называется *операторной корреляционной функцией* векторного процесса, если при $z,\ u \in X$

$$\mathsf{M}\,(\xi\,(t,\ \omega),\ z)\,(\xi\,(s,\ \omega),\ u) = (B\,(t,\ s)\,z,\ u) + (a\,(t),\ z)\,(a\,(s),\ u).$$

Операторная функция $B(t, s)$ также положительно определена: если $z_1, \ldots, z_k \in X$, $t_1, \ldots, t_k \in T$, то

$$\sum_{i,\,j} (B\,(t_i,\ t_j)\,z_i,\ z_j) \geqslant 0.$$

Кроме того, она симметрична в следующем смысле: $B(t, s) = B^*(s, t)$, где B^* — оператор, сопряженный к B.

Операторная корреляционная функция может быть задана своей матрицей в некотором базисе; такая матрица называется *корреляционной матрицей* векторного процесса.

Пусть $\xi_1(t)$ и $\xi_2(t)$ — два случайных процесса на одном и том же вероятностном пространстве. Функция

$$b_{12}\,(t,\ s) = \mathsf{M}\xi_1\,(t)\,\xi_2\,(s) - \mathsf{M}\xi_1\,(t)\,\mathsf{M}\xi_2\,(s)$$

называется *взаимной корреляционной функцией* процессов $\xi_1(t)$ и $\xi_2(t)$. Если $b_{kk}(t, s)$ $(k = 1, 2)$ — корреляционная функция процесса $\xi_k(t)$, то матричная функция

$$\begin{bmatrix} b_{11}\,(t,\ s) & b_{12}\,(t,\ s) \\ b_{12}\,(t,\ s) & b_{22}\,(t,\ s) \end{bmatrix}$$

положительно определена, т. е. для всех $x_1, x_2, \ldots, x_k$; $y_1, y_2, \ldots, y_k$; $t_1, t_2, \ldots, t_k$

$$\sum_{i,\,j=1}^{k} (b_{11}\,(t_i,\ t_j)\,x_i x_j + b_{12}\,(t_i,\ t_j)\,(x_i y_j + x_j y_i) + \\ + b_{22}\,(t_i,\ t_j)\,y_i y_j) \geqslant 0. \qquad (1.1)$$

Всякая симметрическая положительно определенная операторная функция является корреляционной функцией некоторого процесса. Поэтому выполнение условия (1.1) необходимо и достаточно, чтобы $b_{12}(t, s)$ была взаимной корреляционной функцией процессов $\xi_1(t)$ и $\xi_2(t)$.

9.1.3. Стохастическая непрерывность. Пусть случайный процесс $\xi(t)$ задан на некотором интервале T. Он называется *стохастически непрерывным в некоторой точке* $t_0 \in T$, если для всякого $\varepsilon > 0$

$$\lim_{t \to t_0} \mathsf{P}\,\{\,|\,\xi\,(t) - \xi\,(t_0)\,| > \varepsilon\} = 0.$$

Если процесс стохастически непрерывен в каждой точке интервала T, то говорят, что он *стохастически непрерывен на интервале* T. (Это определение относится не только к числовым, но и к векторным процессам. В последнем случае $|\cdot|$ обозначает норму вектора.) Пусть процесс $\xi(t)$ стохастически непрерывен на T. Тогда справедливы следующие утверждения:

1. *Если $f(t, x)$ — непрерывная функция при $t \in T$, $x \in X$, где X — область значений $\xi(t)$, то $f(t, \xi(t))$ является также стохастически непрерывным процессом на T.*

2. *Пусть при некотором $\delta > 0$*

$$\sup_t \mathsf{M}\,|f(t, \xi(t))|^{1+\delta} < \infty,$$

где f — такая функция, как в утверждении 1; тогда функция $\mathsf{M}f(t, \xi(t))$ непрерывна по t.

3. *Пусть f — такая, как в утверждении 1, и числовая неотрицательная функция $\lambda(h) \uparrow +\infty$ при $h \to +\infty$. Если*

$$\sup_t \mathsf{M}f\,(t, \xi(t))\,\lambda\,(|f(t, \xi(t))|) < \infty,$$

то $\mathsf{M}f(t, \xi(t))$ — непрерывная функция.

4. *Если при некотором $\delta > 0$ $\sup \mathsf{M}|\xi(t)|^{k+\delta} < \infty$, то моментные функции $m_j(t_1, \ldots, t_j)$, $j \leqslant k$, процесса $\xi(t)$ непрерывны по совокупности переменных.*

5. *Если процесс $\xi(t)$ стохастически непрерывен на ограниченном замкнутом множестве T, то он равномерно стохастически непрерывен, т. е. для всех $\varepsilon > 0$*

$$\lim_{h \downarrow 0} \sup_{\substack{t_1, t_2 \in T \\ |t_1 - t_2| \leqslant h}} \mathsf{P}\{|\xi(t_1) - \xi(t_2)| > \varepsilon\} = 0.$$

6. *Процесс $\xi(t)$ называется ограниченным по вероятности на множестве T, если*

$$\lim_{c \to +\infty} \sup_{t \in T} \mathsf{P}\{|\xi(t)| > c\} = 0.$$

Если процесс стохастически непрерывен на ограниченном замкнутом множестве T, то он ограничен по вероятности.

9.1.4. Процессы с дискретным фазовым пространством. Во многих задачах область значений процесса является счетным множеством. Для таких процессов конкретный вид фазового пространства не существен. Пусть область возможных значений X состоит из элементов $\{x_1, x_2, \ldots\}$, T — область определения процесса. Конечномерные распределения процесса тогда удобно задавать с помощью вероятностей

$$\mathsf{P}_{t_1, \ldots, t_n}(k_1, \ldots, k_n) = \mathsf{P}\{\xi(t_1) = x_{k_1}, \ldots, \xi(t_n) = x_{k_n}\}.$$

Зная эти вероятности, можно определить и конечномерные распределения процесса по формуле

$$F_{t_1, \ldots, t_n}(A_1, \ldots, A_n) = \sum_{x_{k_1} \in A_1, \ldots, x_{k_n} \in A_n} \mathsf{P}_{t_1, \ldots, t_n}(k_1, \ldots, k_n).$$

9.1.5. Процессы с дискретным временем. Если множество T, на котором определен процесс, есть либо последовательность неотрицательных целых чисел, либо последовательность всех целых чисел, то тогда $\xi(t)$ называется *процессом с дискретным временем* или *случайной последовательностью*.

Пусть $T = \{0, 1, 2, \ldots\}$. Будем писать ξ_n вместо $\xi(n, \omega)$. Конечномерные распределения последовательности $\{\xi_n\}$ полностью определяются функциями распределения

$$F_n(A_0, \ldots, A_n) = \mathsf{P}\{\xi_0 \in A_0, \ldots, \xi_n \in A_n\}.$$

Вместо этих функций распределения иногда удобнее задавать условные функции распределения ξ_n при заданных $\xi_0, \ldots, \xi_{n-1}$

$$F_n(A \mid x_0, \ldots, x_{n-1});$$

это такие функции, что с вероятностью 1

$$\mathsf{P}\{\xi_n \in A \mid \xi_0, \ldots, \xi_{n-1}\} = F_n(A \mid \xi_0, \ldots, \xi_{n-1}).$$

Если $T = \{0, \pm 1, \pm 2, \ldots\}$, то используются конечномерные распределения

$$F_n(A_{-n}, \ldots, A_0, \ldots, A_n) = \\ = \mathsf{P}\{\xi_0 \in A_0, \xi_1 \in A_1, \xi_{-1} \in A_{-1}, \ldots, \xi_n \in A_n, \xi_{-n} \in A_{-n}\}.$$

Их также можно задавать с помощью условных распределений

$$\mathsf{P}\{\xi_n \in A \mid \xi_0, \xi_1, \xi_{-1}, \ldots, \xi_{n-1}, \xi_{-n+1}\};$$

$$\mathsf{P}\{\xi_{-n} \in A \mid \xi_0, \xi_1, \xi_{-1}, \ldots, \xi_{n-1}, \xi_{-n+1}, \xi_n\}.$$

Примеры случайных процессов.

1. На вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, где Ω есть $[0, 1]$; $\mathfrak{S}$ — σ-алгебра борелевских множеств этого отрезка; P — мера Лебега на $[0, 1]$, процесс $\xi(t, \omega)$ при $t \in [0, 1]$ определяется равенством

$$\xi(t, \omega) = \begin{cases} 1, & t > \omega, \\ 0, & t \leqslant \omega. \end{cases}$$

Конечномерные распределения процесса (его фазовое пространство состоит из двух точек: 0 и 1) определяются соотношениями:

при $t_1 < t_2 < \ldots < t_n$

$$\mathsf{P}\{\xi(t_1) = 0, \ldots, \xi(t_{i-1}) = 0, \quad \xi(t_i) = 1, \ldots, \xi(t_n) = 1\} = \\ = t_i - t_{i-1};$$

при $1 < i \leqslant n$

$$\mathsf{P}\{\xi(t_1) = 0, \ldots, \xi(t_n) = 0\} = 1 - t_n,$$

$$\mathsf{P}\{\xi(t_1) = 1, \ldots, \xi(t_n) = 1\} = t_1.$$

Во всех остальных случаях $\mathsf{P}\{\xi(t_1) = k_1, \ldots, \xi(t_n) = k_n\} = 0$ ($k_1, \ldots, k_n$ принимают значения 0 и 1).

Процесс $\xi(t)$ является стохастически непрерывным: при $\varepsilon < 1$, $t_1 < t_2$

$$\mathsf{P}\{|\xi(t_2) - \xi(t_1)| > \varepsilon\} = \mathsf{P}\{\xi(t_1) = 0, \quad \xi(t_2) = 1\} = t_2 - t_1.$$

Однако почти все выборочные функции процесса являются разрывными.

Этот пример показывает, что стохастическая непрерывность не влечет непрерывности выборочных функций.

2. *Процесс Пуассона* — это процесс $\xi(t)$, значениями которого служат неотрицательные целые числа, определенный при $t \geqslant 0$, если

его конечномерные распределения при $0 = t_0 < t_1 < \ldots < t_n$ заданы равенством

$$\mathsf{P}\{\xi(t_1) = k_1,\ \xi(t_2) = k_2,\ \ldots,\ \xi(t_n) = k_n\} =$$
$$= \begin{cases} \prod\limits_{j=1}^{n} \dfrac{[a(t_j - t_{j-1})]^{k_j - k_{j-1}}}{(k_j - k_{j-1})!} e^{-a(t_j - t_{j-1})}, & \text{если} \\ & 0 = k_0 \leqslant k_1 \leqslant \ldots \leqslant k_2 \leqslant \ldots \leqslant k_n; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Этот процесс описывает число редких событий, происходящих за время t (например, число космических частиц, зарегистрированных счетчиком, или число вызовов, поступивших на телефонную станцию, и т. п.). Число $a > 0$ называют *параметром процесса*: $a = \mathsf{M}\xi(t)/t$.

Процесс Пуассона можно построить следующим образом. Пусть $\eta_1, \eta_2, \ldots$ — последовательность независимых одинаково распределенных неотрицательных величин, для которых

$$\mathsf{P}\{\eta_k > t\} = e^{-at}.$$

Если $\varepsilon(z) = 1$ при $z \geqslant 0$, $\varepsilon(z) = 0$ при $z < 0$, то функция

$$\xi(t) = \sum_{k=1}^{\infty} \varepsilon\left(t - \sum_{l=1}^{k} \eta_l\right)$$

является процессом Пуассона (последняя формула задает процесс как функцию t и ω, поскольку от ω зависят величины η_l).

3. *Процесс чистого роста.* Пусть η_k такие же, как и в предыдущем примере, а $\lambda_k > 0$ — некоторая последовательность, для которой $\sum\limits_{k=1}^{\infty} \dfrac{1}{\lambda_k} = +\infty$. Процесс вида

$$\xi(t) = \sum_{k=1}^{\infty} \varepsilon\left(t - \sum_{i=1}^{k} \frac{\eta_l}{a\lambda_t}\right)$$

называется *процессом чистого роста*. Выборочные функции этого процесса — неубывающие целочисленные ступенчатые функции, все скачки которых равны 1, $\xi(0) = 0$.

4. *Одномерное броуновское движение* (*винеровский процесс*) — процесс $w(t)$, определенный при $t \geqslant 0$; его конечномерные распределения определяются совместными плотностями распределения величин $w(t_1), w(t_2), \ldots, w(t_n)$ при $t_1 < t_2 < \ldots < t_n$, которые имеют вид

$$f_{t_1, \ldots, t_n}(x_1, \ldots, x_n) =$$
$$= \prod_{j=1}^{n} [2\pi(t_j - t_{j-1})]^{-1/2} \exp\left\{-\frac{1}{2} \cdot \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}}\right\},$$

где $t_0 = 0$, $x_0 = 0$, а $x_1, x_2, \ldots, x_n$ — вещественные переменные.

Процесс $w(t)$ может служить вероятностной моделью явлений диффузии или броуновского движения ($w(t)$ — одна из координат диффундирующей частицы).

9.2. Измеримость и интегрируемость случайных процессов

9.2.1. Измеримые случайные процессы. Случайный процесс $\xi(t, \omega)$, заданный на борелевском множестве T с фазовым пространством X, называется *измеримым*, если $\xi(t, \omega)$ измеримы относительно σ-алгебры $\mathfrak{B}_T \times \mathfrak{S}$, где $\mathfrak{B}_T$ — σ-алгебра борелевских множеств на T; $\mathfrak{S}$ — σ-алгебра событий вероятностного пространства $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, на котором определен случайный процесс. Это означает, что для всякого борелевского множества $A \subset X$

$$\{(t, \omega): \xi(t, \omega) \in A\} \in \mathfrak{B}_T \times \mathfrak{S}$$

(произведение σ-алгебр $\mathfrak{B}$ и $\mathfrak{S}$ есть минимальная σ-алгебра, содержащая множества $B \times S$, где $B \in \mathfrak{B}$, $S \in \mathfrak{S}$).

Если процесс $\xi(t, \omega)$ измерим, то почти для всех ω выборочные функции $\xi(\cdot, \omega)$ являются борелевскими функциями t.

Процесс, который строится по конечномерным распределениям в теореме Колмогорова (см. § 9.1), не будет измеримым. Возникает вопрос: при каких условиях по данным конечномерным распределениям можно построить измеримый процесс?

Воспользуемся понятием *стохастической эквивалентности* двух случайных процессов (оно отличается от понятия *стохастической эквивалентности в широком смысле*, введенного в § 9.1). *Два процесса* $\xi_1(t)$ *и* $\xi_2(t)$, определенных на одном и том же вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ и заданных на одном и том же множестве T, называются стохастически эквивалентными, если

$$\mathsf{P}\{\xi_1(t) = \xi_2(t)\} = 1, \quad \forall t \in T.$$

Очевидно, что стохастически эквивалентные процессы имеют одинаковые конечномерные распределения.

Теорема. *Если процесс* $\xi(t)$ *стохастически непрерывен на борелевском множестве* T, *то существует измеримый процесс* $\xi'(t)$, *стохастически эквивалентный* $\xi(t)$.

Этот процесс $\xi'(t)$ можно определить как предел процессов $\xi_n(t) = \xi(t_{nk})$ при $t \in [t_{nk}, t_{nk+1})$, где $t_{nk} \in T$ и $\max\limits_{n \to \infty} |t_{nk+1} - t_{nk}| \to 0$. Для тех пар (t, ω), для которых указанный предел не существует, полагаем $\xi'(t) = 0$.

Следствие. *Если* $\xi(t)$ *имеет не более чем счетное множество точек разрыва, то существует измеримый процесс* $\xi'(t)$, *стохастически эквивалентный* $\xi(t)$.

Отметим некоторые важные свойства измеримых процессов.

1) Пусть $\varphi(t, x)$ — функция, измеримая относительно $\mathfrak{B}_T \times \mathfrak{B}_X$, где $\mathfrak{B}_X$ — σ-алгебра борелевских множеств в фазовом пространстве X рассматриваемого процесса. Если

$$\mathsf{M}\,|\varphi(t_1, \xi(t_1))| \ldots |\varphi(t_k, \xi(t_k))| < \infty \quad \forall t_1, \ldots, t_k \in T,$$

то функция

$$g(t_1, \ldots, t_k) = \mathsf{M}\varphi(t_1, \xi(t_1)) \ldots \varphi(t_k, \xi(t_k))$$

является борелевской, т. е. измеримой относительно $\mathfrak{B}_T^k$.

В частности, все моментные функции измеримого процесса являются борелевскими.

2) Если $\varphi(t, x)$ — ограниченная $\mathfrak{B}_T \times \mathfrak{B}_X$-измеримая функция, то

$$\int_T \varphi(t, \xi(t, \omega))\, dt$$

существует для почти всех ω; этот интеграл — $\mathfrak{S}$-измеримая величина, и

$$\mathsf{M} \int_T \varphi(t, \xi(t, \omega))\, dt = \int_T \mathsf{M}\varphi(t, \xi(t, \omega))\, dt$$

(это утверждение является следствием теоремы Фубини).

3) Если $\mathsf{M}|\xi(t, \omega)| < \infty$ и $\int_T \mathsf{M}|\xi(t, \omega)|\, dt < \infty$, то для почти всех ω существует $\int_T \xi(t, \omega)\, dt$.

9.2.2. Интегрирование случайных процессов. Пусть $\xi(t)$ — измеримый вещественный стохастически непрерывный случайный процесс на отрезке $[a, b]$. Приведем условия, при которых выборочные функции $\xi(t)$ с вероятностью 1 принадлежат $L_p[a, b]$, где $0 < p \leqslant 2$, т. е.

$$\mathsf{P}\left\{\int_a^b |\xi(t)|^p\, dt < \infty\right\} = 1. \tag{2.1}$$

Обозначим через $\chi(g)$ характеристический функционал

$$\chi(g) = \mathsf{M} \exp\left\{i \int_a^b \xi(t)\, dg(t)\right\}$$

процесса $\xi(t)$, определенный на ступенчатых функциях g, заданных на $[a, b]$.

Пусть $t_{nk} = a + \frac{k}{n}(b - a)$ $(k = 0, \ldots, n)$; $\eta_0, \eta_1, \ldots$ — независимые случайные величины, равномерно распределенные на $[0, 1]$; $\zeta_0, \zeta_1, \ldots$ — случайные величины, не зависящие друг от друга и от $\eta_0, \eta_1, \ldots$, для которых

$$\mathsf{M}e^{is\zeta_k} = e^{-|s|^p}$$

(т. е. ζ_k имеют симметричное устойчивое распределение с показателем p).

Введем случайный процесс на $[a, b]$

$$\nu_n^p(t) = \sum_{j=0}^{n-1} \frac{\zeta_j}{n^{1/p}}\, \varepsilon\left(t - t_{nj} - \frac{b-a}{n}\eta_j\right),$$

где $\varepsilon(t) = 0$ при $t < 0$, $\varepsilon(t) = 1$ при $t \geqslant 0$.

Поскольку $\nu_n^p(t)$ — ступенчатая функция на $[a, b]$, определена величина $\chi\left(\lambda\nu_n^p\right)$. Поскольку $\nu_n^p(\cdot)$ — случайная функция, то и $\chi\left(\lambda\nu_n^p\right)$ будет случайной величиной.

Необходимым и достаточным условием для выполнения (2.1) является условие: для всех положительных λ существует предел

$$\psi(\lambda) = \lim_{n\to\infty} \mathsf{M}\chi\left(\lambda\nu_n^p\right),$$

удовлетворяющий соотношению $\psi(0+) = 1$. При этом

$$\psi(\lambda) = \mathsf{M}\exp\left\{-\frac{\lambda^p}{b-a}\int_a^b |\xi(t)|^p\,dt\right\}.$$

Пусть $\xi(t)$ с вероятностью 1 принадлежит $L_p[a, b]$, $p \geqslant 1$. Тогда для всякой ограниченной измеримой функции $\varphi(t)$ на $[a, b]$ с вероятностью 1 определен интеграл

$$\int_a^b \xi(t)\,\varphi(t)\,dt.$$

Поэтому для такого процесса будет также определен характеристический функционал

$$\chi_1(\varphi) = \mathsf{M}\exp\left\{i\int_a^b \xi(t)\,\varphi(t)\,dt\right\}.$$

Характеристические функционалы $\chi_1(\varphi)$ и $\chi(g)$ связаны между собой следующими соотношениями:

1) $\chi_1(\varphi) = \lim\limits_{n\to\infty} \mathsf{M}\chi(\nu_n)$,

где ν_n — последовательность случайных функций на $[a, b]$ вида

$$\nu_n = \sum_{j=0}^{n-1} \frac{b-a}{n}\,\varphi\left(t_{nj} + \frac{b-a}{n}\eta_j\right)\varepsilon\left(t - t_{nj} - \frac{b-a}{n}\eta_j\right),$$

где η_j — последовательность независимых, равномерно распределенных на $[0, 1]$ величин;

2) если $\xi(t)$ — стохастически непрерывный процесс, то

$$\chi(g) = \lim_{n\to\infty} \chi_1(\varphi_n),$$

где φ_n — такая последовательность измеримых функций, что

$$\int_a^t \varphi_n(s)\,ds \to g(t) - g(a).$$

Рассмотрим в качестве примера процесс броуновского движения $w(t)$. Для него

$$\chi(g) = \exp\left\{-\frac{1}{2}\int_0^1\int_0^1 \min[t, s]\, dg(t)\, dg(s)\right\},$$

откуда

$$\chi_1(\varphi) = \exp\left\{-\frac{1}{2}\int_0^1\int_0^1 \min[t, s]\, \varphi(t)\,\varphi(s)\, dt\, ds\right\}.$$

9.3. Сепарабельность. Свойства выборочных функций

9.3.1. Определение сепарабельного процесса. Пусть $\xi(t)$ — случайный процесс с фазовым пространством X на вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, заданный на множестве T. Он называется *сепарабельным*, если существует такое счетное плотное в T подмножество $I \subset T$ и такое множество $\Lambda \in \mathfrak{S}$, $\mathsf{P}(\Lambda) = 0$, что для всех замкнутых множеств $F \subset X$ и для всякого интервала (α, β)

$$\{\omega\colon \xi(t, \omega) \in F,\quad t \in (\alpha, \beta) \cap I\} - \{\omega\colon \xi(t, \omega) \in F,$$
$$t \in (\alpha, \beta) \cap T\} \subset \Lambda.$$

Множество I называется множеством сепарабельности процесса.

Если, например, $\xi(t)$ — сепарабельный числовой случайный процесс на $[a, b]$ и $I = \{t_1, \ldots, t_k, \ldots\}$, то

$$\mathsf{P}\{\sup_{a \leqslant t \leqslant b} \xi(t) \leqslant x\} = \mathsf{P}\{\sup_{t_k} \xi(t_k) \leqslant x\} =$$
$$= \lim_{n \to \infty} \mathsf{P}\{\xi(t_1) \leqslant x, \ \ldots, \ \xi(t_n) \leqslant x\},$$

Оказывается, переходя к стохастически эквивалентным процессам, можно всегда процесс превратить в сепарабельный. Это утверждает следующая теорема Дж. Л. Дуба.

Теорема 1. *Пусть $\xi(t)$ — случайный процесс с фазовым пространством X, являющимся конечномерным евклидовым пространством, и $\bar{X}$ — некоторое компактное расширение X. Тогда существует сепарабельный процесс $\xi'(t)$ со значениями в $\bar{X}$, стохастически эквивалентный $\xi(t)$.*

Поскольку X — локально компактное пространство, такое компактное расширение $\bar{X}$ всегда существует. Если, например, X — прямая, то, чтобы получить компактное расширение, следует добавить к X точки $\pm\infty$.

Хотя в общем случае невозможно указать для данного процесса множество сепарабельности, однако для стохастически непрерывных процессов в качестве множества сепарабельности можно взять любое счетное всюду плотное множество $I \subset T$.

9.3.2. Непрерывные процессы. Пусть $\{F_T, \mathfrak{S}_T\}$ — измеримое пространство, где F_T — множество всех функций со значениями в X, определенных на T, а $\mathfrak{S}_T$ — σ-алгебра, порожденная цилиндрическими множествами. Если C_T — множество всех непрерывных функций на T $(C_T \subset F_T)$, то для несчетных T $C_T \notin \mathfrak{S}_T$. Поэтому, если процесс построен так, как указано в теореме Колмогорова (§ 9.1), то говорить о непрерывности выборочных функций процесса не имеет смысла. Для сепарабельного процесса и замкнутого множества T множество непрерывных выборочных функций измеримо, так как

$$\{\omega\colon \xi(\cdot, \omega) \in C_T\} = \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{\substack{t, s \in I \\ |t-s| \leqslant 1/k}} \left\{\omega\colon |\xi(t, \omega) - \xi(s, \omega)| \leqslant \frac{1}{m}\right\}.$$

Здесь I — множество сепарабельности процесса. Для того чтобы выборочные функции сепарабельного процесса были с вероятностью 1 непрерывны, необходимо и достаточно, чтобы для некоторого счетного плотного в T множества I выполнялось соотношение

$$\mathsf{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{\substack{t, s \in I \\ |t-s| \leqslant 1/k}} \left\{\omega\colon |\xi(t, \omega) - \xi(s, \omega)| \leqslant \frac{1}{m}\right\}\right) = 1.$$

Это соотношение требует для своей проверки знания всех конечномерных распределений и, как правило, не проверяемо. Следующая теорема Колмогорова дает удобные достаточные условия непрерывности процесса.

Теорема 2. *Пусть $\xi(t)$ — сепарабельный процесс, заданный на $[a, b]$. Если существуют такие $\alpha > 0$, $\beta > 0$ и K, что для всех t, $s \in [a, b]$*

$$\mathsf{M}\,|\xi(t) - \xi(s)|^{\alpha} \leqslant K\,|t - s|^{1+\beta},$$

то $\xi(t)$ с вероятностью 1 непрерывен.

Рассмотрим применение этой теоремы к вопросу о непрерывности винеровского процесса. Для него при $t > s$ величина $w(t) - w(s)$ имеет нормальное распределение со средним 0 и дисперсией $t - s$. Поэтому

$$\mathsf{P}\{w(t) - w(s) < x\} = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{x} \exp\left\{-\frac{u^2}{2(t-s)}\right\} du;$$

$$\mathsf{M}\,|w(t) - w(s)|^{\alpha} = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{\infty} |u|^{\alpha} \exp\left\{-\frac{u^2}{2(t-s)}\right\} du = |t - s|^{\alpha/2} \int_{-\infty}^{\infty} \frac{|u|^{\alpha}}{\sqrt{2\pi}} e^{-u^2/2}\, du.$$

Условия теоремы Колмогорова выполнены, если $\alpha > 2$, $\beta = \alpha/2 - 1$. Значит, сепарабельный винеровский процесс с вероятностью 1 непрерывен.

9.3.3. Процессы без разрывов второго рода. Пусть процесс $\xi(t, \omega)$ с фазовым пространством X, являющимся конечномерным евклидовым пространством, задан на отрезке $[a, b]$. Обозначим через D множество функций $x(t)$ со значениями из X, определенных на $[a, b]$, для которых при $t \in [a, b)$ определены пределы справа, а при $t \in (a, b]$ — пределы слева. Функции из D являются функциями без разрывов второго рода. Для того чтобы функция $x(t)$ не имела разрывов второго рода, необходимо и достаточно, чтобы

$$\lim_{h \to 0} \sup_{a \leqslant t \leqslant s \leqslant u \leqslant t+h \leqslant b} \min (|x(s) - x(t)|, \quad |x(u) - x(s)|) = 0.$$

Пусть $\xi(t, \omega)$ — сепарабельный процесс, I — множество сепарабельности. Тогда с вероятностью 1

$$\sup_{a \leqslant t \leqslant s \leqslant u \leqslant t+h \leqslant b} \min (|\xi(s, \omega) - \xi(t, \omega)|, |\xi(u, \omega) - \xi(s, \omega)|) =$$
$$= \sup_{\substack{a \leqslant t \leqslant s \leqslant u \leqslant t+h \leqslant b \\ t, s, u \in I}} \min (|\xi(s, \omega) - \xi(t, \omega)|, |\xi(u, \omega) - \xi(s, \omega)|).$$

Поэтому для сепарабельного процесса необходимым и достаточным условием того, что выборочные функции процесса принадлежат D, т. е. процесс с вероятностью 1 не имеет разрывов второго рода, является условие

$$\mathsf{P}\Bigg(\bigcap_{r=1}^{\infty} \bigcup_{l=1}^{\infty} \bigcap_{\substack{t, s, u \in I \\ a \leqslant t \leqslant s \leqslant u \leqslant t+1/l \leqslant b}} \left\{\omega: |\xi(s, \omega) - \xi(t, \omega)| \leqslant \frac{1}{r}\right\} \bigcup$$
$$\bigcup \left\{\omega: |\xi(u, \omega) - \xi(s, \omega)| \leqslant \frac{1}{r}\right\}\Bigg) = 1.$$

Для проверки принадлежности D можно использовать число ε-колебаний функции. Говорят, что функция $x(t)$ имеет на $[a, b]$ не менее k ε-колебаний, если существуют такие точки $t_0 < t_1 < \ldots < t_k$, что

$$|x(t_{i+1}) - x(t_i)| \geqslant \varepsilon, \quad i = 0, \ldots, k-1.$$

Для того чтобы $x(\cdot) \in D$, необходимо и достаточно, чтобы для всех $\varepsilon > 0$ функция $x(t)$ имела на $[a, b]$ конечное число ε-колебаний.

Если $\xi(t, \omega)$ — сепарабельный процесс, I — множество сепарабельности, то $\xi(\cdot, \omega) \in D$ с вероятностью 1, если $\xi(\cdot, \omega)$ для всякого $\varepsilon > 0$ имеет с вероятностью 1 конечное число ε-колебаний на I. Пусть $I = \bigcup\limits_n I_n$, где I_n — возрастающая последовательность конечных множеств. Если ν_ε — число ε-колебаний на I, а ν_ε^n — число ε-колебаний на I_n, то $\nu_\varepsilon = \lim\limits_{n \to \infty} \nu_\varepsilon^n$. Зная совместное распределение величин $\xi(t_1), \ldots, \xi(t_n)$, где $I_n = \{t_1, \ldots, t_n\}$, можем вычислить распределение ν_ε^n, затем с помощью предельного перехода — распределение ν_ε и проверить условие $\mathsf{P}\{\nu_\varepsilon < \infty\} = 1$.

Приведенные выше условия отсутствия у процесса разрывов второго рода требуют знания всех конечномерных распределений процесса и вычисления вероятностей весьма сложных событий.

Приведем некоторые достаточные легко проверяемые условия отсутствия разрывов второго рода.

Теорема 3. *Пусть $\xi(t)$ — сепарабельный процесс на $[a, b]$, для которого существуют такие $\alpha > 0$, $\beta > 0$ и K, что при $t < s < u$*

$$\mathsf{M}\left(|\xi(u)-\xi(s)|\,|\xi(s)-\xi(t)|\right)^{\alpha} \leqslant K(u-t)^{1+\beta}.$$

Тогда процесс $\xi(t)$ с вероятностью 1 не имеет разрывов второго рода.

В качестве примера рассмотрим пуассоновский процесс с параметром a. Для него величины $\xi(u)-\xi(s)$ и $\xi(s)-\xi(t)$ независимы. Взяв $\alpha = 1$, будем иметь

$$\mathsf{M}\,|\xi(u)-\xi(s)|\,|\xi(s)-\xi(t)| = \mathsf{M}\,|\xi(u)-\xi(s)|\,\mathsf{M}\,|\xi(s)-\xi(t)| = a^2(u-s)(s-t) \leqslant a^2(u-t)^2.$$

Таким образом, условия теоремы выполнены при $\alpha = 1$, $\beta = 1$, $K = a^2$. Значит, сепарабельный пуассоновский процесс не имеет разрывов второго рода.

Следующая теорема использует условные распределения процесса. Пусть $\xi(t)$ — некоторый процесс. Обозначим через $\mathfrak{F}_t^{\xi}$ минимальную σ-алгебру, относительно которой измеримы величины $\xi(u)$ при $u \leqslant t$.

Теорема 4. *Пусть $\xi(t)$ — сепарабельный процесс, заданный на $[a, b]$. Если существует такая (неслучайная) функция $\varphi_\varepsilon(h)$ ($\varepsilon > 0$, $h > 0$), что $\varphi_\varepsilon(h) \downarrow 0$ при $h \downarrow 0$ и с вероятностью 1*

$$\mathsf{P}\left\{|\xi(t+h)-\xi(t)| > \varepsilon \mid \mathfrak{F}_t^{\xi}\right\} \leqslant \varphi_\varepsilon(h),$$

то процесс $\xi(t)$ с вероятностью 1 не имеет разрывов второго рода.

Последняя теорема удобна для применения в случае процессов с *независимыми приращениями*. Так называется процесс $\xi(t)$, для которого величины $\xi(t_0)$, $\xi(t_1)-\xi(t_0)$, ..., $\xi(t_k)-\xi(t_{k-1})$ независимы при $t_0 < t_1 < \ldots < t_k$. Для процессов с независимыми приращениями величина $\xi(t+h)-\xi(t)$ не зависит от σ-алгебры $\mathfrak{F}_t^{\xi}$ ($h > 0$); поэтому

$$\mathsf{P}\left\{|\xi(t+h)-\xi(t)| > \varepsilon \mid \mathfrak{F}_t^{\xi}\right\} = \mathsf{P}\{|\xi(t+h)-\xi(t)| > \varepsilon\}.$$

Поскольку стохастически непрерывный на $[a, b]$ процесс будет и равномерно стохастически непрерывным на $[a, b]$, то

$$\lim_{h \downarrow 0} \sup_{a \leqslant t < t+h \leqslant b} \mathsf{P}\{|\xi(t+h)-\xi(t)| > \varepsilon\} = 0.$$

Следовательно, всякий стохастически непрерывный на $[a, b]$ сепарабельный процесс с независимыми приращениями с вероятностью 1 не имеет разрывов второго рода.

9.3.4. Условия непрерывности для процессов без разрывов второго рода. Пусть $x(t)$ — функция из D. Выберем последовательность разбиений отрезка $[a, b]$: $a = t_{n0} < t_{n1} < \ldots < t_{nn} = b$, для которой $\max_k (t_{nk+1} - t_{nk}) \to 0$. Если n_ε — число разрывов процесса $x(t)$, превосходящих $\varepsilon > 0$, то

$$n_\varepsilon \leqslant \varliminf_{n\to\infty} \sum_{k=0}^{n-1} \chi_{(\varepsilon,\infty)}(|x(t_{nk+1}) - x(t_{nk})|),$$

где χ_A — индикатор множества A. Поэтому, если число разрывов случайного процесса $\xi(t, \omega)$, превосходящих $\varepsilon > 0$, обозначим ν_ε, то

$$\mathsf{M}\nu_\varepsilon \leqslant \varliminf_{n\to\infty} \sum_{k=0}^{n-1} \mathsf{M}\chi_{(\varepsilon,\infty)}(|\xi(t_{nk+1}) - \xi(t_{nk})|) =$$

$$= \varliminf_{n\to\infty} \sum_{k=0}^{\infty} \mathsf{P}\{|\xi(t_{nk+1}) - \xi(t_{nk})| > \varepsilon\}.$$

Чтобы процесс был непрерывен, необходимо и достаточно, чтобы $\nu_\varepsilon = 0$ для всякого $\varepsilon > 0$. Таким образом, для непрерывности процесса, который с вероятностью 1 не имеет разрывов второго рода, достаточно, чтобы для всякого $\varepsilon > 0$

$$\varliminf_{n\to\infty} \sum_{k=0}^{n-1} \mathsf{P}\{|\xi(t_{nk+1}) - \xi(t_{nk})| > \varepsilon\} = 0. \tag{3.1}$$

Поскольку сепарабельный процесс с независимыми приращениями не имеет разрывов второго рода, если только он стохастически непрерывен, а условие (3.1) влечет стохастическую непрерывность, то (3.1) достаточно, чтобы сепарабельный процесс $\xi(t)$ с независимыми приращениями был непрерывен с вероятностью 1. Оказывается, что это условие и необходимо для непрерывности процесса с независимыми приращениями.

Условие (3.1) *необходимо и достаточно, чтобы сепарабельный процесс с независимыми приращениями* $\xi(t)$ *был с вероятностью* 1 *непрерывен.*

9.4. Потоки σ-алгебр и согласованные с ними процессы

9.4.1. Поток σ-алгебр. При рассмотрении случайных процессов удобно бывает пользоваться следующей схемой. Пусть $\{\Omega, \mathfrak{F}, \mathsf{P}\}$ — основное вероятностное пространство. В σ-алгебре $\mathfrak{F}$ для каждого $t \geqslant 0$ выделяется под-σ-алгебра $\mathfrak{F}_t$ событий, наблюденных до момента времени t включительно. Совокупность σ-алгебр $\{\mathfrak{F}_t\}_{t \geqslant 0}$ образует поток σ-алгебр (определение потока будет приведено ниже).

Процесс $\xi(t)$, определенный для $t \geqslant 0$, называется *согласованным с этим потоком*, если для всех $t \geqslant 0$ величина $\xi(t)$ является $\mathfrak{F}_t$-измеримой. Естественно рассматривать только согласованные процессы.

С каждым процессом $\xi(t)$, определенным на $[0, \infty)$, можно связать поток, с которым он согласован. Минимальный в некотором смысле поток получится, если считать, что ничего, кроме случайного процесса $\xi(t)$, не наблюдается, так что до момента t мы наблюдаем только траектории процесса на $[0, t]$. Тогда в качестве $\mathfrak{F}_t$ естественно взять σ-алгебру, порожденную событиями $\{\xi(s) \in B\}$, $s \leqslant t$, $B \in \mathfrak{B}$, где $(X, \mathfrak{B})$ — фазовое пространство процесса $\xi(t)$.

Совокупность σ-алгебр $\mathfrak{F}_t \subset \mathfrak{F}$, определенных для $t \geqslant 0$, называется *потоком σ-алгебр*, если выполнены следующие условия:

а) σ-алгебра $\mathfrak{F}$ полна *) относительно меры P и $\mathfrak{F}_0$ содержит все события вероятности 0;

б) для $0 \leqslant s < t$ $\mathfrak{F}_s \subset \mathfrak{F}_t$ ($\mathfrak{F}_t$ монотонно возрастает с t);

в) для всех $t \geqslant 0$ $\mathfrak{F}_t = \bigcap\limits_{s>t} \mathfrak{F}_s$ ($\mathfrak{F}_t$ непрерывно справа по t).

Заметим, что построенная выше по процессу $\xi(t)$ совокупность σ-алгебр удовлетворяет лишь условию б) и не будет потоком. Однако ее легко превратить в поток, пополняя σ-алгебры множествами меры 0 и рассматривая вместо $\mathfrak{F}_t$ σ-алгебры $\mathfrak{F}_{t+} = \bigcap\limits_{s>t} \mathfrak{F}_s$. Эти σ-алгебры будут уже потоком.

9.4.2. Моменты остановки и связанные с ними σ-алгебры. Неотрицательная случайная величина τ, которая может принимать значение $+\infty$ и определена на $\{\Omega, \mathfrak{F}, \mathsf{P}\}$, называется *моментом остановки* (м. о.) *относительно потока* $\{\mathfrak{F}_t\}_{t \geqslant 0}$, если для всех $t \geqslant 0$ событие $\{\tau \leqslant t\}$ принадлежит $\mathfrak{F}_t$. Моменты остановки называются еще *марковскими моментами* или *моментами, не зависящими от будущего*.

Если поток $\{\mathfrak{F}_t\}_{t \geqslant 0}$ порожден некоторым случайным процессом $\xi(t)$, то моменты остановки определяются поведением процесса $\xi(t)$ на отрезке $[0, \tau]$. Например, в случае непрерывного процесса $\xi(t)$ моментами остановки будут моменты достижения процессом некоторого замкнутого множества.

Пусть τ — м. о. относительно $\{\mathfrak{F}_t\}$. Свяжем с τ σ-алгебру $\mathfrak{F}_\tau$ тех множеств $A \in \mathfrak{F}$, для которых $\{\tau \leqslant t\} \cap A \in \mathfrak{F}_t$ для всех $t \geqslant 0$; σ-алгебра $\mathfrak{F}_\tau$ — это события, которые наблюдаются до момента τ (включительно). Определим теперь σ-алгебру $\mathfrak{F}_{\tau-}$ как σ-алгебру, порожденную событиями вида $A_t \cap \{\tau > t\}$, где $A_t \in \mathfrak{F}_t$. Это σ-алгебра событий, которые наблюдаются во все моменты, предшествующие τ.

Отметим некоторые свойства м. о. и связанных с ними σ-алгебр.

1) Если τ_1 и τ_2 — м. о., то $\tau_1 \vee \tau_2$ и $\tau_1 \wedge \tau_2$ также м. о.

2) Если τ_n $(n = 1, 2, \ldots)$ — м. о., то $\varlimsup\limits_{n\to\infty} \tau_n$ и $\varliminf\limits_{n\to\infty} \tau_n$ также м. о.

3) Если τ и σ — м. о. и $\tau < \sigma$ при $\sigma < \infty$, то $\mathfrak{F}_{\tau-} \subset \mathfrak{F}_\tau \subset \mathfrak{F}_{\sigma-} \subset \subset \mathfrak{F}_\sigma$.

4) Если τ и σ — м. о. и $\tau \leqslant \sigma$, то $\mathfrak{F}_{\tau-} \subset \mathfrak{F}_{\sigma-}$, $\mathfrak{F}_\tau \subset \mathfrak{F}_\sigma$.

5) Если τ_n — м. о. и $\tau_n \downarrow \tau$, то $\mathfrak{F}_\tau = \bigcap\limits_n \mathfrak{F}_{\tau_n}$; если же $\tau_n \uparrow \tau$, то $\mathfrak{F}_{\tau-} = \bigvee\limits_n \mathfrak{F}_{\tau_n -}$.

6) Если τ_n—м. о. и $\tau_n \downarrow \tau$, $\tau_n > \tau$ при $\tau < \infty$, то $\mathfrak{F}_\tau = \bigcap\limits_n \mathfrak{F}_{\tau_n -}$; если же $\tau_n \uparrow \tau$, $\tau_n < \tau$ при $\tau < \infty$, то $\mathfrak{F}_{\tau-} = \bigvee\limits_n \mathfrak{F}_{\tau_n}$.

*) $\mathfrak{F}$ полна относительно меры P, если из того, что $A \in \mathfrak{F}$, $\mathsf{P}(A) = 0$ и $B \subset A$, следует, что $B \in \mathfrak{F}$.

7) Если τ, σ — м. о., то для $A \in \mathfrak{F}_\sigma$

$$A \cap \{\sigma \leqslant \tau\} \in \mathfrak{F}_\tau, \quad A \cap \{\sigma < \tau\} \in \mathfrak{F}_{\tau-}.$$

Эти свойства показывают, что σ-алгебры, связанные с моментами остановки, ведут себя так, как и σ-алгебры $\mathfrak{F}_t$, где t неслучайны (при этом $\mathfrak{F}_{t-} = \bigvee_{s<t} \mathfrak{F}_s$).

Момент остановки τ называется *предсказуемым*, если существует последовательность м. о. τ_n такая, что $\tau_n < \tau$ и $\tau_n \to \tau$.

Всякий момент остановки, не являющийся предсказуемым, называется *непредсказуемым*.

Пример. Пусть $\{\mathfrak{F}_t\}$ — поток, порожденный случайным процессом $\xi(t, \omega) = I_{\{t > \omega\}}$, заданным на вероятностном пространстве $\{R_+, \mathfrak{B}_+, P\}$, где $R_+ = [0, \infty)$, $\mathfrak{B}_+$ — борелевская σ-алгебра R_+, P — некоторая вероятностная мера на $\mathfrak{B}_+$, не имеющая атомов. Тогда величина $\tau(\omega) = \omega$ является моментом остановки, но непредсказуемым.

Свойства предсказуемых моментов остановки (п. м. о).

8) Если σ — п. м. о., τ — м. о., тогда при $A \in \mathfrak{F}_{\sigma-}$

$$A \cap \{\sigma \leqslant \tau\} \in \mathfrak{F}_{\tau-}, \quad \{\sigma \leqslant \tau\} \in \mathfrak{F}_{\tau-}, \quad \{\sigma = \tau\} = \mathfrak{F}_{\tau-}.$$

9) Если τ_n — п. м. о. и $\tau_n \uparrow \tau$, то τ — п. м. о.; если же $\tau_n \downarrow \tau$ и $\tau_n = \tau$ для достаточно больших n (такое n, вообще говоря, случайно), то τ — п. м. о.

Пусть τ — некоторый момент остановки $A \in \mathfrak{F}$. Полагаем

$$\tau_A = \begin{cases} \tau, & \omega \in A, \\ +\infty, & \omega \notin A. \end{cases}$$

10) Если τ — п. м. о, то τ_A — п. м. о. тогда и только тогда, когда $A \in \mathfrak{F}_{\tau-}$.

Поток σ-алгебр называется *квазинепрерывным слева*, если $\mathfrak{F}_\tau = \mathfrak{F}_{\tau-}$ для п. м. о. τ.

9.4.3. Основные σ-алгебры и измеримость относительно этих δ-алгебр. Согласованный с потоком $\{\mathfrak{F}_t\}_{t \geqslant 0}$ случайный процесс $\xi(t) = \xi(t, \omega)$ с измеримым фазовым пространством $(X, \mathfrak{B})$ называется *прогрессивно измеримым* (относительно $\{\mathfrak{F}_t\}_{t \geqslant 0}$), если для каждого $t > 0$ функция $\xi(s, \omega)$, рассматриваемая на $[0, t] \times \Omega$, измерима относительно $\mathfrak{B}_{[0, t]} \otimes \mathfrak{F}_t$, где $\mathfrak{B}_{[0, t]}$ — σ-алгебра борелевских множеств на $[0, t]$.

Утверждение 1. *Если $\xi(t)$ — прогрессивно измеримый процесс и τ — м. о., то $\xi(\tau)$ определено и является $\mathfrak{F}_\tau$-измеримой величиной.*

Утверждение 2. *Если $\xi(t)$ — ограниченный числовой прогрессивно измеримый процесс, то $\int_0^t \xi(s)\, ds$ является $\mathfrak{F}_t$-измеримой величиной.*

Будем рассматривать подмножества σ-алгебры $\mathfrak{B}_+ \otimes \mathfrak{F}$ ($\mathfrak{B}_+$ — σ-алгебра борелевских множеств R_+). Такие множества называются *случайными*. Если $C \in \mathfrak{B}_+ \otimes \mathfrak{F}$, то $I_C(t, \omega)$ является измеримым случайным процессом. Множество C называется *согласованным*, если этот процесс согласован и прогрессивно измерим. Прогрессивная измеримость C означает, что

$$C \cap [0, t] \times \Omega \in \mathfrak{B}_{[0,\, t]} \otimes \mathfrak{F}_t$$

для всех $t \geqslant 0$. Прогрессивно измеримые множества образуют σ-алгебру Π. Измеримость функции $\xi(t, \omega)$ относительно Π эквивалентна прогрессивной измеримости $\xi(t, \omega)$.

Дебютом множества C называется величина $\operatorname{deb} C(\omega) = \inf \{t\colon (t, \omega) \in C\}$ (считаем $\inf \varnothing = +\infty$). Если $\mathfrak{F}$ полна относительно P, то $\operatorname{deb} C(\omega)$ есть случайная величина. Справедливо следующее утверждение.

Утверждение 3. *Для всех* $C \in \pi$ $\operatorname{deb} C$ *есть м. о.*

Условия прогрессивной измеримости случайного процесса дает следующая теорема.

Теорема 1. *Пусть фазовое пространство X является сепарабельным полным метрическим пространством. Тогда для прогрессивной измеримости процесса достаточно, чтобы он был непрерывен справа (слева) для всех $t \geqslant 0$.*

Среди случайных множеств выделим стохастические интервалы $[\![\sigma, \tau]\!]$, $]\!]\sigma, \tau]\!]$, $[\![\sigma, \tau[\![$, $]\!]\sigma, \tau[\![$, где σ и τ — м. о., каждый из интервалов содержит те пары (t, ω), для которых t при всех ω удовлетворяет соответствующим неравенствам, скажем $\sigma \leqslant t < \tau$ для первого интервала.

Пусть $\mathscr{W}$ — σ-алгебра, порожденная интервалами $[\![\sigma, \tau[\![$, где σ, τ — произвольные м. о.; она называется *σ-алгеброй вполне измеримых множеств*. Так как $I_{[\![\sigma,\, \tau[\![}$ — непрерывный справа процесс, то он прогрессивно измерим. Поэтому $\mathscr{W} \subset \Pi$; σ-алгебра $\mathscr{W}$ порождается множествами $\{(\theta, \omega)\colon \omega \in A,\ A \in \mathfrak{F}_0\}$ и $[s, t[\times A$, где $s < t$ — неслучайные числа, $A \in \mathfrak{F}_s$.

Если функция $\xi(t, \omega)$ является $\mathscr{W}$-измеримой, то процесс $\xi(t, \omega)$ называется *вполне измеримым* или *опциональным*.

Теорема 2. *В условиях теоремы 1 всякий непрерывный справа (слева) процесс $\mathscr{W}$-измерим.*

Требование вполне измеримости процесса более сильное, чем требование измеримости. Ниже будет приведен пример прогрессивно измеримого, но не вполне измеримого процесса. Два процесса $\xi(t)$ и $\xi'(t)$ называются *неразличимыми*, если

$$\mathsf{P}\left\{\bigcup_{t \geqslant 0} \{\xi(t) \neq \xi'(t)\}\right\} = 0.$$

Теорема 3. *Если для двух вполне измеримых процессов $\xi_1(t)$ и $\xi_2(t)$:*

1) $\mathsf{P}\{\xi_1(\tau) = \xi_2(\tau)\} = 1$

для любого м. о., то $\xi_1(t)$ и $\xi_2(t)$ неразличимы;

2) *для любой ограниченной измеримой функции $f(x)$ из X в $\mathbf{R}$ и любого м. о. τ*

$$\mathsf{M} f(\xi_1(\tau)) = \mathsf{M} f(\xi_2(\tau)),$$

то $\xi_1(t)$ и $\xi_2(t)$ неразличимы.

Пример. Пусть $w(t)$ — винеровский процесс. Множество, где $w(t) \neq 0$, есть объединение счетного числа интервалов. Обозначим через L множество левых концов этих интервалов и $\xi(t) = I_L(t)$.

Тогда $\xi(t)$ П-измерим относительно потока, порожденного винеровским процессом (зная процесс $w(s)$ для $s \leqslant t$, мы знаем множество $L \cap [0, t)$, а $\mathsf{P}\{t \in L\} = 0$ для всех $t > 0$), но он не $\mathscr{W}$-измерим, так как $\mathsf{P}\{\tau \in L\} = 0$ для всякого м. о. τ, а, значит, $\mathsf{P}\{\xi(\tau) = 0\} = 1$, и, если бы $\xi(t)$ был $\mathscr{W}$-измерим, он по теореме 3 был бы неотличим от нуля, что не так, поскольку L непусто.

Обозначим через $\mathscr{P}$ σ-алгебру, порожденную стохастическими интервалами $[\![\sigma, \tau[\![$, где σ и τ — произвольные предсказуемые м. о.; $\mathscr{P}$ называется *предсказуемой* σ-алгеброй $\mathscr{P} \subset \mathscr{W}$. σ-алгебра $\mathscr{P}$ — наименьшая σ-алгебра, относительно которой измеримы все непрерывные согласованные процессы. Еще $\mathscr{P}$ можно определить как σ-алгебру, порожденную множествами $\{(0; \omega), \omega \in A_0\}$, $A_0 \in \mathfrak{F}_0$ и $]s, t] \times A_s$, где $A_s \in \mathfrak{F}_s$, $s < t$.

$\mathscr{P}$-измеримые случайные процессы называются *предсказуемыми*.

Теорема 4. *Всякий непрерывный слева согласованный процесс в полном сепарабельном метрическом пространстве является предсказуемым.*

Для предсказуемых процессов справедлив такой аналог теоремы 3.

Теорема 3'. *Если $\xi_1(t)$ и $\xi_2(t)$ — предсказуемые процессы, то утверждения теоремы 3 остаются в силе, если условия этой теоремы выполнены лишь для предсказуемых м. о.*

Связь между предсказуемостью и измеримостью устанавливает следующая теорема.

Теорема 5. *Для всякого вполне измеримого процесса $\xi(t)$ существует такой предсказуемый процесс $\xi'(t)$ и такая последовательность моментов остановки $\tau_1, \tau_2, \ldots$, что*

$$\mathsf{P}\left\{ \bigcap_{\substack{t \neq \tau_k \\ k=1,\, 2,\, \ldots}} \{\xi(t) = \xi'(t)\} \right\} = 1.$$

9.4.4. Существование измеримых модификаций. Пусть $\xi(t)$ $(t \geqslant 0)$ — процесс в измеримом фазовом пространстве $(X, \mathfrak{B})$ со счетно порожденной σ-алгеброй $\mathfrak{B}$, согласованный с потоком σ-алгебр $\{\mathfrak{F}_t\}_{t \geqslant 0}$. Этот поток будем называть *непрерывным слева*, если $\mathfrak{F}_{t-} = \bigvee_{s<t} \mathfrak{F}_s = \mathfrak{F}_t$ для всех $t > 0$. Процесс $\xi(t)$ будем называть *счетно порожденным*, если пространство случайных величин $\{I_B(\xi(t)), B \in \mathfrak{B}, t \geqslant 0\}$ с топологией сходимости в среднеквадратическом является сепарабельным.

Теорема 6. *Если процесс $\xi(t)$ счетно порожден и для любых A и $B \in \mathfrak{B}$ функция*

$$\mathsf{M} I_A(\xi(t))\, I_B(\xi(s))$$

измерима по совокупности переменных s, t, то:

1) *существует вполне измеримая модификация процесса $\xi(t)$;*

2) *если, кроме того, поток $\{\mathfrak{F}_t\}_{t \geqslant 0}$ непрерывен слева, то существует предсказуемая модификация процесса $\xi(t)$.*

Как следствие этой теоремы получаем существование в условиях теоремы прогрессивно измеримой модификации.

9.4.5. Измеримые проекции. Рассмотрим числовой процесс $\xi(t, \omega)$ на вероятностном пространстве $\{\Omega, \mathfrak{F}, \mathsf{P}\}$, на котором задан поток

σ-алгебр $(\mathfrak{F}_t)_{t\geqslant 0}$. Будем предполагать, что $\xi(t, \omega)$ ограничен или неотрицателен, измерим, но не обязательно согласован с потоком $(\mathfrak{F}_t)_{t\geqslant 0}$.

Процесс ${}^o\xi(t)$ называется *вполне измеримой* (*опциональной*) *проекцией* процесса $\xi(t, \omega)$, если он вполне измерим и для всякого момента остановки τ

$$\mathsf{M}\xi(\tau, \omega) I_{\{\tau<\infty\}} = \mathsf{M}^o\xi(\tau) I_{\{\tau<\infty\}}.$$

Процесс ${}^p\xi(t)$ называется *предсказуемой проекцией* процесса $\xi(\tau, \omega)$, если он предсказуем и для всякого предсказуемого момента остановки τ

$$\mathsf{M}\xi(\tau, \omega) I_{\{\tau<\infty\}} = \mathsf{M}^p\xi(\tau) I_{\{\tau<\infty\}}.$$

Вполне измеримая и предсказуемая проекции процесса определяются однозначно. Предсказуемая проекция процесса совпадает с предсказуемой проекцией его опциональной проекции.

Теорема 7. *Всякий ограниченный или неотрицательный измеримый процесс имеет опциональную и предсказуемую проекции. Существует не более чем счетное множество моментов остановки $\{\tau_n\}$ таких, что ${}^o\xi(t) = {}^p\xi(t)$ при $t \neq \tau_n$ для всех n.*

9.5. Абсолютная непрерывность мер, соответствующих случайным процессам

9.5.1. Меры, соответствующие случайным процессам. Пусть $\xi(t)$ — случайный процесс, заданный на множестве T с фазовым пространством X; F_T — пространство всех функций $x(t)$, определенных на T со значениями из X, $\mathfrak{S}_T$ — минимальная σ-алгебра подмножеств F_T, содержащая все цилиндрические множества. Мера μ_ξ, определенная на $\mathfrak{S}_T$ соотношением

$$\mu_\xi(A) = \mathsf{P}\{\xi(\cdot) \in A\}, \quad A \in \mathfrak{S}_T,$$

называется мерой, *соответствующей процессу* $\xi(\cdot)$. Иногда вместо F_T рассматривают некоторое другое множество функций (например, $\mathcal{M}_T$ — пространство измеримых функций, D_T — пространство функций без разрывов второго рода, C_T — пространство непрерывных функций) такое, что выборочные функции процесса с вероятностью 1 принадлежат этому множеству. Всякая $\mathfrak{S}_T$-измеримая функция $\varphi(x(\cdot))$ определяет некоторый функционал от процесса — случайную величину $\varphi(\xi(\cdot))$. Пусть процесс с вероятностью 1 принадлежит D_T. Тогда такими функционалами будут, например,

$$\sup_{t\in T} |\xi(t)|, \qquad \int_T f(t, \xi(t))\,dt,$$

где $f(t, x)$ — измеримая ограниченная функция.

Зная меру, соответствующую процессу, можно определить распределение любого функционала от процесса. Для этого можно использовать формулу: если $\varphi(x(\cdot))$ — ограниченный $\mathfrak{S}_T$-измеримый функционал, то

$$\mathsf{M}\varphi(\xi(\cdot)) = \int_{F_T} \varphi(x)\,\mu_\xi(dx).$$

Поэтому для любого $\mathfrak{S}_T$-измеримого функционала $\varphi(x(\cdot))$ характеристическая функция величины $\varphi(\xi(\cdot))$ определяется равенством

$$\mathsf{M}e^{iz\varphi(\xi(\cdot))} = \int_{F_T} e^{iz\varphi(x)} \mu_\xi(dx).$$

9.5.2. Абсолютная непрерывность мер. Одним из вопросов, рассматриваемых в теории мер, соответствующих случайным процессам, является вопрос об абсолютной непрерывности этих мер. Приведем необходимые определения. Пусть X — некоторое множество, на котором выделена σ-алгебра подмножеств $\mathfrak{B}$. Пара $(X, \mathfrak{B})$ называется *измеримым пространством*. Предположим, что на $(X, \mathfrak{B})$ заданы две меры μ и ν. Говорят, что мера μ *абсолютно непрерывна* относительно меры ν (обозначается $\mu \ll \nu$), если $\mu(A) = 0$ для всех $A \in \mathfrak{B}$, для которых $\nu(A) = 0$. Если $\mu \ll \nu$ и $\nu \ll \mu$, то говорят, что μ и ν *эквивалентны* $(\mu \sim \nu)$. Меры μ и ν *сингулярны* (*ортогональны*) $(\mu \perp \nu)$, если существует такое множество $S \in \mathfrak{B}$, что $\mu(S) = 0$, $\nu(X \setminus S) = 0$. Каковы бы ни были μ и ν, всегда возможно представление $\mu = \mu_1 + \mu_2$, $\nu = \nu_1 + \nu_2$, где $\mu_1 \sim \nu_1$, $\mu_2 \perp \nu$, $\nu_2 \perp \mu$.

Теорема Радона — Никодима. *Если μ и ν — конечные меры и $\mu \ll \nu$, то существует $\mathfrak{B}$-измеримая функция $\rho(x)$ такая, что для всех $A \in \mathfrak{B}$*

$$\mu(A) = \int_A \rho(x)\, \nu(dx).$$

Функция $\rho(x)$ определяется однозначно с точностью до множеств нулевой меры ν.

Функция $\rho(x)$ называется *плотностью меры μ относительно меры ν* или *производной* и обозначается $\frac{d\mu}{d\nu}(x)$.

Для случайных процессов изучаются условия: 1) абсолютной непрерывности одной меры относительно другой; 2) взаимной сингулярности (ортогональности) мер; 3) в абсолютно непрерывном случае вычисляется плотность одной меры относительно другой. Укажем на те применения, которые могут быть получены при изучении вопросов абсолютной непрерывности.

Если известно, что $\mu_{\xi_1} \ll \mu_{\xi_2}$ для двух случайных процессов $\xi_1(t)$ и $\xi_2(t)$, то всякое событие, которое имеет вероятность 1 для процесса $\xi_2(t)$, будет иметь такую же вероятность и для $\xi_1(t)$. В частности, если выборочные функции процесса $\xi_2(t)$ с вероятностью 1 обладают некоторым свойством (непрерывны, не имеют разрывов второго рода, измеримы, дифференцируемы и т. п.), то существует процесс $\xi_1'(t)$, стохастически эквивалентный $\xi_1(t)$, выборочные функции которого с вероятностью 1 обладают тем же свойством. Если, кроме того, известна плотность $\frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(x(\cdot))$, то вычисление математических ожиданий функционалов от процесса $\xi_1(\cdot)$ можно свести к вычислению математических ожиданий функционалов от процесса $\xi_2(\cdot)$, используя формулу

$$\mathsf{M}\varphi(\xi_1(\cdot)) = \mathsf{M}\varphi(\xi_2(\cdot)) \frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(\xi_2(\cdot)).$$

Эта формула позволяет вычислять и распределения функционалов:

$$\mathsf{P}\{\varphi(\xi_1(\cdot)) < \lambda\} = \mathsf{M}\chi_{(-\infty,\lambda)}(\varphi(\xi_1(\cdot))) = \\ = \mathsf{M}\chi_{(-\infty,\lambda)}(\varphi(\xi_2(\cdot)))\frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(\xi_2(\cdot)),$$

где $\chi_{(-\infty,\lambda)}$ — индикатор интервала $(-\infty, \lambda)$.

Если установлено, что $\mu_{\xi_1} \perp \mu_{\xi_2}$, и указано то множество S, для которого $\mu_{\xi_1}(S) = 1$, $\mu_{\xi_2}(S) = 0$, то можно решать следующую статистическую задачу. Наблюдается процесс $\xi(t)$ $(t \in T)$, конечномерные распределения которого неизвестны. Известно лишь, что соответствующая ему мера — либо мера μ_{ξ_1}, либо μ_{ξ_2}. Нужно по наблюдению определить, какая именно из этих мер отвечает $\xi(t)$. (Например, при обнаружении сигнала на фоне случайного шума μ_{ξ_1} — распределение чистого шума, μ_{ξ_2} — распределение сигнала с шумом. Нужно по наблюдению определить наличие или отсутствие сигнала.) Решение задачи, очевидно, такое: если $\xi(\cdot) \in S$, считаем, что $\mu_\xi = \mu_{\xi_1}$; если $\xi(\cdot) \notin S$, то $\mu_\xi = \mu_{\xi_2}$.

9.5.3. Абсолютная непрерывность мер, соответствующих случайным процессам. Предположим, что процессы $\xi_i(t)$ определены и стохастически непрерывны на множестве T, а T_n — возрастающая последовательность конечных подмножеств такая, что $\bigcup T_n$ плотно в T. Обозначим через $\mathfrak{S}_{T_n}$ σ-алгебру, порожденную цилиндрическими множествами с основаниями в T_n. Если $T_n = \{t_{n1}, \ldots, t_{nk_n}\}$, то множества A из $\mathfrak{S}_{T_n}$ имеют вид

$$\{x(\cdot)\colon (x(t_{n1}), \ldots, x(t_{nk_n})) \in B_{k_n}\},$$

где B_{k_n} — произвольное борелевское множество из $\mathbf{R}^{k_n}$ — k_n-мерного евклидова пространства, $(x_1, \ldots, x_{k_n})$ — точка этого пространства.

Обозначим через $\mu_{\xi_i}^{T_n}$ сужение меры μ_{ξ_i} на σ-алгебру $\mathfrak{S}_{T_n}$. Мера $\mu_{\xi_i}^{T_n}$ однозначно определяется функцией

$$F^{(i)}_{t_{n1},\ldots,t_{nk_n}}(A_1, \ldots, A_{k_n}) = \mathsf{P}\{\xi_i(t_{n1}) \in A_1, \ldots, \xi_i(t_{nk_n}) \in A_{k_n}\}.$$

Имеют место следующие утверждения.

1. *Если* $\mu_{\xi_1} \ll \mu_{\xi_2}$, *то* $\mu_{\xi_1}^{T_n} \ll \mu_{\xi_2}^{T_n}$ *для всех* n,

$$F^{(1)}_{t_{n1},\ldots,t_{nk_n}}(A_1, \ldots, A_{k_n}) = \\ = \int_{A_1}\cdots\int_{A_{k_n}} g_n(x_1, \ldots, x_{k_n})\, F^{(2)}_{t_{n1},\ldots,t_{nk_n}}(dx_1, \ldots, dx_{k_n}), \quad (5.1)$$

при этом

$$\frac{d\mu_{\xi_1}^{T_n}}{d\mu_{\xi_2}^{T_n}}(x(\cdot)) = g_n\left(x(t_{n1}), \ldots, x\left(t_{nk_n}\right)\right).$$

2. *Пусть* $\mathfrak{F}_{T_n}^{\xi_2}$ — *σ-алгебра, порожденная величинами* $\xi_2(t_{ni})$ $(i = 1, \ldots, k_n)$. *Тогда*

$$\frac{d\mu_{\xi_1}^{T_n}}{d\mu_{\xi_2}^{T_n}}(\xi_2(\cdot)) = \mathsf{M}\left(\frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(\xi_2(\cdot)) \,\Big|\, \mathfrak{F}_{T_n}^{\xi_2}\right).$$

3. *Пусть для всех n существует функция g_n такая, что выполнено* (5.1) *для всех борелевских множеств* $A_1, \ldots, A_{k_n}$ *из* $\mathbb{R}$. *Тогда с вероятностью* 1 *существует предел*

$$\rho = \lim_{n\to\infty} g_n\left(\xi_2(t_{n1}), \ldots, \xi_2\left(t_{nk_n}\right)\right).$$

Если при этом $\mathsf{M}\rho = 1$, *то* $\mu_{\xi_1} \ll \mu_{\xi_2}$ *и*

$$\frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(\xi_2(\cdot)) = \rho.$$

4. *Пусть существует функция g_n такая, что выполнено* (5.1), $\rho_n = g_n\left(\xi_2(t_{n1}), \ldots, \xi_2\left(t_{nk_n}\right)\right)$ *и последовательность ρ_n равномерно интегрируема; тогда* $\mathsf{M}\rho = 1$.

В частности, это будет выполнено, если для некоторой функции $\psi(\lambda)$, для которой $\psi(\lambda) \uparrow +\infty$ при $\lambda \uparrow +\infty$, $\sup \mathsf{M}\rho_n\psi(\rho_n) < \infty$ $\left(\text{например, } \sup\limits_n \mathsf{M}\rho_n^\alpha < \infty \text{ при некотором } \alpha > 1\right)$.

5. *Предположим, что функции g_n, удовлетворяющие* (5.1), *положительные. Тогда с вероятностью* 1 *существует предел*

$$\rho_1 = \lim_{n\to\infty}\left[g_n\left(\xi_1(t_{n1}), \ldots, \xi_1\left(t_{nk_n}\right)\right)\right]^{-1}.$$

Если $\mathsf{P}\{\rho_1 = 0\} = 0$, *то* $\mu_{\xi_1} \sim \mu_{\xi_2}$ *и*

$$\frac{d\mu_{\xi_1}}{d\mu_{\xi_2}}(\xi_2(\cdot)) = \rho, \quad \rho_1 = \frac{d\mu_{\xi_2}}{d\mu_{\xi_1}}(\xi_1(\cdot)).$$

Пример. Пусть $\xi_2(t) = w(t)$, $\xi_1(t) = w(t) + a(t)$, где $w(t)$ — винеровский процесс, $a(t)$ — некоторая непрерывная функция, $a(0) = 0$. Найдем условия, при которых $\mu_{\xi_1} \ll \mu_{\xi_2}$, если $T = [0, 1]$.

Пусть $T_n = \left\{\frac{k}{2^n},\ k = 0, 1, \ldots, 2^n\right\}$. Используя то обстоятельство, что $\xi_i\left(\frac{k+1}{2^n}\right) - \xi_i\left(\frac{k}{2^n}\right)$ при $k = 0, \ldots, 2^n - 1$ независимы между собой и имеют нормальное распределение с дисперсией

$1/2^n$ и средними 0 при $i=2$ и $a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)$ при $i=1$, убеждаемся, что

$$\frac{d\mu_{\xi_1}^{T_n}}{d\mu_{\xi_2}^{T_n}}(w(\cdot))=\exp\left\{\sum_{k=0}^{2^n-1}2^n\left[a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)\right]\times\right.$$

$$\times\left[w\left(\frac{k+1}{2^n}\right)-w\left(\frac{k}{2^n}\right)\right]-\frac{1}{2}\sum_{k=0}^{2^n-1}\frac{1}{2^n}\times$$

$$\left.\times\left\{2^n\left[a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)\right]\right\}^2\right\}.$$

Эта величина имеет отличный от нуля предел, если

$$\varlimsup_{n\to\infty}\sum_{k=0}^{2^n-1}\frac{1}{2^n}\left\{2^n\left[a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)\right]\right\}^2<\infty.$$

Последнее условие влечет существование интегрируемой с квадратом производной $a'(t)$, при этом

$$\lim_{n\to\infty}\sum_{k=0}^{2^n-1}\frac{1}{2^n}\left\{2^n\left[a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)\right]\right\}^2=\int_0^1[a'(t)]^2\,dt,$$

$$\lim_{n\to\infty}\sum_{k=0}^{2^n-1}2^n\left[a\left(\frac{k+1}{2^n}\right)-a\left(\frac{k}{2^n}\right)\right]\left[w\left(\frac{k+1}{2^n}\right)-w\left(\frac{k}{2^n}\right)\right]=$$

$$=\int_0^1 a'(t)\,dw(t)$$

(определение этого интеграла дано в п. 19.2). Таким образом,

$$\rho=\exp\left\{\int_0^1 a'(t)\,dw(t)-\frac{1}{2}\int_0^1[a'(t)]^2\,dt\right\}. \qquad (5.2)$$

Поскольку $\int_0^1 a'(t)\,dw(t)$ — нормально распределенная величина со средним 0 и дисперсией $\int_0^1[a'(t)]^2\,dt$, то $M\rho=1$. Итак, правая часть (5.2) дает плотность $d\mu_{\xi_1}/d\mu_{\xi_2}$ в случае существования интегрируемой с квадратом $a'(t)$. В противном случае $\mu_{\xi_2}\perp\mu_{\xi_1}$.

Литература: [19, 27, 28, 36, 68].

Глава 10. $\mathscr{L}_2$-ТЕОРИЯ

10.1. Пространство гильбертовых случайных величин $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$

10.1.1. Определение. Сходимость. Совокупность комплекснозначных случайных величин ξ, заданных на вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ с конечным вторым моментом $\mathsf{M}|\xi|^2 < \infty$, образует линейное нормированное гильбертово пространство $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ со скалярным произведением

$$(\xi, \eta) = \mathsf{M}\xi\bar{\eta} \tag{1.1}$$

и нормой

$$\|\xi\| = [\mathsf{M}|\xi|^2]^{1/2}. \tag{1.2}$$

С помощью нормы определяется расстояние между случайными величинами из $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$

$$\rho(\xi, \eta) = \|\xi - \eta\|.$$

Случайные величины ξ, принадлежащие $\mathscr{L}_2 = \mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$, называют *гильбертовыми случайными величинами.*

Гильбертовы случайные величины ξ и η называют ортогональными, если $\mathsf{M}\xi\bar{\eta} = 0$.

Приведенное определение $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ сохраняется и в более общей ситуации для случайных элементов со значениями в измеримом полном гильбертовом пространстве $\mathscr{H}$. При этом $\xi\bar{\eta}$ означает скалярное произведение в $\mathscr{H}$, $|\xi|^2 = \xi\bar{\xi}$.

Сходимость в пространстве $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ определяется в *среднеквадратическом*:

$$\underset{n\to\infty}{\mathrm{l.\,i.\,m.}}\, \xi_n = \xi,$$

если $\lim\limits_{n\to\infty} \|\xi_n - \xi\| = 0$, или в эквивалентной форме

$$\lim_{n\to\infty} \mathsf{M}|\xi_n - \xi|^2 = 0.$$

Из сходимости в среднеквадратическом (в с. к.) следует сходимость по вероятности. Обратное неверно.

Однако если $|\xi_n| < \eta \in \mathscr{L}_2$, то из сходимости ξ_n по вероятности следует сходимость и в среднеквадратическом.

10.1.2. Ковариация, характеристическое свойство. *Гильбертова случайная функция* $\{\zeta(x),\ x \in \mathscr{X}\}$ задается совокупностью гильбертовых случайных величин $\zeta(x)$, зависящих от параметра x, принимающего значения в некотором параметрическом множестве $\mathscr{X}$.

Ковариацией $B(x, y)$ гильбертовой случайной функции $\{\zeta(x),\ x \in \mathscr{X}\}$ называется функция

$$B(x, y) = \mathsf{M}\zeta(x)\bar{\zeta}(y); \quad x, y \in \mathscr{X}.$$

Ковариация $B(x, y)$ является *положительно определенной функцией*:

$$\sum_{k, r=1}^{n} B(x_k, x_r) z_k \bar{z}_r \geqslant 0$$

для любых $n \geqslant 1$, $x_k \in \mathscr{X}$ и комплексных чисел z_k. При этом

$$\sum_{k,r=1}^{n} B(x_k, x_r) z_k \bar{z}_r = \mathsf{M} \left| \sum_{k=1}^{n} \zeta(x_k) z_k \right|^2.$$

Положительная определенность является характеристическим свойством ковариации.

Теорема 1. *Для того чтобы функция $B(x, y)$ была ковариационной, необходимо и достаточно, чтобы она была положительно определенной.*

Свойства гильбертовых случайных функций, выраженные в терминах свойств ковариационной функции, называют *ковариационными свойствами* или *свойствами второго порядка.*

10.1.3. Непрерывность случайной функции. Пусть $\mathscr{X}$ — метрическое пространство с метрикой ρ.

Гильбертова случайная функция $\{\zeta(x),\ x \in \mathscr{X}\}$ называется *непрерывной* (*в среднеквадратическом*) в точке x_0, если

$$\mathsf{M}\, |\zeta(x) - \zeta(x_0)|^2 \to 0, \quad \rho(x, x_0) \to 0.$$

Теорема 2. *Для непрерывности $\zeta(x)$ в точке x_0 необходима и достаточна непрерывность ковариации $B(x, y) = \mathsf{M}\zeta(x)\bar{\zeta}(y)$ в точке (x_0, x_0).*

Следствие. *Если ковариация $B(x, y)$ непрерывна в каждой диагональной точке $(x_0, x_0) \in \mathscr{X} \times \mathscr{X}$, то она непрерывна и во всех точках $(x, y) \in \mathscr{X} \times \mathscr{X}$.*

Заметим, что из непрерывности (в с. к.) случайной функции $\zeta(x)$ не следует непрерывность ее выборочных функций.

10.1.4. Дифференцируемость случайной функции. Пусть $\mathscr{X} = (a, b)$ — интервал вещественной оси.

Гильбертова случайная функция $\{\zeta(x),\ x \in \mathscr{X}\}$ *дифференцируема* (*в с. к.*) *в точке* x_0, если существует

$$\zeta'(x_0) = \underset{h \to 0}{\mathrm{l.\,i.\,m.}} \frac{1}{h} [\zeta(x_0 + h) - \zeta(x_0)]; \quad x_0,\ x_0 + h \in (a, b).$$

Случайная величина $\zeta'(x_0)$ называется *производной* (в с. к.) случайной функции $\zeta(x)$ в точке x_0.

Теорема 3. *Гильбертова случайная функция $\{\zeta(x),\ x \in (a, b)\}$ дифференцируема в каждой точке x_0 интервала (a, b) тогда и только тогда, когда существует обобщенная смешанная производная второго порядка ковариации*

$$\left. \frac{\partial^2 B(x, y)}{\partial x\, \partial y} \right|_{x=y} = \lim_{h, h_1 \to 0} \frac{1}{h h_1} [B(x_0 + h, x_0 + h_1) - B(x_0, x_0 + h_1) - B(x_0 + h, x_0) + B(x_0, x_0)].$$

При этом

$$\mathsf{M}\zeta'(x)\, \bar{\zeta}'(y) = \frac{\partial^2 B(x, y)}{\partial x\, \partial y},$$

$$\mathsf{M}\zeta'(x)\, \bar{\zeta}(y) = \frac{\partial B(x, y)}{\partial x}.$$

10.1.5. Интегрирование случайной функции. Пусть $\mathscr{X}$ — полное сепарабельное метрическое пространство с σ-конечной мерой $m(dx)$ и $m(\mathscr{X}) < \infty$.

Для измеримой случайной функции $\{\zeta(x),\ x \in \mathscr{X}\}$ *интеграл Лебега* определяется так:

$$\int_{\mathscr{X}} \zeta(x)\, m(dx) = \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \int_{\mathscr{X}} \zeta_n(x)\, m(dx), \tag{1.3}$$

где $\zeta_n(x)$ — монотонно неубывающая последовательность случайных функций, принимающих конечное число значений и таких, что $\zeta(x) = \lim_{n\to\infty} \zeta_n(x)$ с вероятностью 1.

Интеграл Лебега (1.3) можно также определить как предел (в с. к.) лебеговых интегральных сумм

$$\int_{\mathscr{X}} \zeta(x)\, m(dx) = \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \sum_{k=1}^{n} \zeta(x_k)\, m(\Delta x_k), \tag{1.4}$$

где

$$\mathscr{X} = \sum_{k=1}^{n} \Delta x_k, \quad x_k \in \Delta x_k.$$

Теорема 4. *Если конечен интеграл*

$$\int_{\mathscr{X}} B(x, x)\, m(dx) < \infty \tag{1.5}$$

и $m(\mathscr{X}) < \infty$, то для измеримой случайной функции $\{\zeta(x),\ x \in \mathscr{X}\}$ интеграл Лебега (1.3) *с вероятностью* 1 *конечен и может быть определен либо соотношением* (1.4), *либо для каждой реализации $\zeta(x)$. При этом*

$$\mathsf{M} \int_{\mathscr{X}} |\zeta(x)|^2\, m(dx) = \int_{\mathscr{X}} B(x, x)\, m(dx).$$

Следствие. *Пусть функции $f_i(x)$ $(i = 1, 2)$ принадлежат $\mathscr{L}_2(\mathscr{X}, \mathfrak{X}, m)$ и выполнено условие* (1.5). *Тогда с вероятностью* 1 *существуют интегралы*

$$\eta_i = \int_{\mathscr{X}} f_i(x)\, \zeta(x)\, m(dx), \qquad i = 1, 2,$$

$$\mathsf{M}\eta_1\bar{\eta}_2 = \int_{\mathscr{X}}\int_{\mathscr{X}} f_1(x)\, B(x, y)\, \bar{f}_2(y)\, m(dx)\, m(dy).$$

Замечание. Несобственный (в с. к.) интеграл определяется следующим образом:

$$\int_a^{\infty} \zeta(t)\, dt = \underset{N\to\infty}{\mathrm{l.\,i.\,m.}} \int_a^{N} \zeta(t)\, dt. \tag{1.6}$$

Для существования несобственного интеграла (1.6) необходимо и достаточно существование

$$\lim_{N, M \to \infty} \int_a^N \int_a^M B(t, s)\, dt\, ds.$$

10.1.6. Разложение в ортогональные ряды. Пусть $\{\zeta(x), x \in [a, b]\}$ — непрерывная гильбертова случайная функция с ковариацией $B(x, y)$. Согласно теории интегральных уравнений, ядро $B(x, y)$ может быть разложено в равномерно сходящийся ряд по своим собственным функциям $\varphi_n(x)$:

$$B(x, y) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \overline{\varphi_n(y)}, \tag{1.7}$$

где

$$\lambda_n \varphi_n(x) = \int_a^b B(x, y) \varphi_n(y)\, dy, \quad \int_a^b \varphi_n(x) \overline{\varphi_m(x)}\, dx = \delta_{nm}, \tag{1.8}$$

причем собственные числа λ_n положительны.

Положим

$$\xi_n = \int_a^b \zeta(x) \overline{\varphi_n(x)}\, dx. \tag{1.9}$$

Тогда (см. следствие теоремы 4)

$$\mathsf{M}\xi_n\bar{\xi}_m = \int_a^b \int_a^b B(x, y) \varphi_n(x) \overline{\varphi_m(y)}\, dx\, dy = \lambda_n \delta_{nm}, \tag{1.10}$$

$$\mathsf{M}\zeta(x)\, \xi_n = \int_a^b B(x, y) \varphi_n(y)\, dy = \lambda_n \varphi_n(x). \tag{1.11}$$

Так что последовательность случайных величин ξ_n $(n \geqslant 1)$ ортогональна.

Теорема 5. *Измеримая и непрерывная* (в с.к.) *гильбертова случайная функция $\zeta(x)$ на $[a, b]$ представима рядом*

$$\zeta(x) = \sum_{n=1}^{\infty} \xi_n \varphi_n(x), \tag{1.12}$$

который сходится в $\mathscr{L}_2$ при каждом $x \in [a, b]$.

В этом разложении $\{\xi_n, n \geqslant 1\}$ — ортогональная последовательность случайных величин с $\mathsf{M}|\xi_n|^2 = \lambda_n$, λ_n — собственные числа, $\varphi_n(x)$ — собственные функции ковариации $B(x, y)$ случайной функции $\zeta(x)$.

Замечание. Если случайная функция $\zeta(x)$ имеет гауссовское распределение при каждом x, тогда случайные величины ξ_n в разложении (1.12) являются независимыми гауссовскими величинами и ряд (1.12) сходится с вероятностью 1.

Пример. Процесс броуновского движения $\zeta(t)$ на $[0, 1]$ с $\zeta(0) = 0$, $\mathsf{M}\zeta(t) = 0$, $\mathsf{D}\zeta(t) = t$ и $B(t, s) = \mathsf{M}\zeta(t)\zeta(s) = \min(t, s)$ представим в виде ортогонального ряда

$$\zeta(t) = \sqrt{2} \sum_{n=0}^{\infty} \xi_n \frac{\sin(n + 1/2)\pi t}{(n + 1/2)\pi},$$

где ξ_n — последовательность независимых гауссовских случайных величин с параметрами $\mathsf{M}\xi_n = 0$, $\mathsf{D}\xi_n = 1$.

10.2. Стохастические меры и интегралы

10.2.1. Определение стохастического интеграла. Пусть $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ — вероятностное пространство, E — некоторое множество, $\mathfrak{M}$ — полукольцо подмножеств E.

Семейство гильбертовых случайных величин $\{\zeta(\Delta), \Delta \in \mathfrak{M}\}$, удовлетворяющее условиям: 1) $\zeta(\Delta_1 \cup \Delta_2) = \zeta(\Delta_1) + \zeta(\Delta_2) \pmod{\mathsf{P}}$, если $\Delta_1 \cap \Delta_2 = \varnothing$; 2) $\mathsf{M}\zeta(\Delta_1)\bar{\zeta}(\Delta_2) = m(\Delta_1 \cap \Delta_2)$, где $m(\Delta)$ — функция множеств на $\mathfrak{M}$; 3) $\zeta(\varnothing) = 0$, $m(\varnothing) = 0$, называется *элементарной ортогональной стохастической мерой*, а $m(\Delta)$ — ее *структурной функцией*.

Ортогональность стохастической меры $\zeta(\Delta)$ следует из свойства 2): $\mathsf{M}\zeta(\Delta_1)\overline{\zeta(\Delta_2)} = 0$, если $\Delta_1 \cap \Delta_2 = \varnothing$.

Структурная функция $m(\Delta)$ является элементарной мерой на полукольце $\mathfrak{M}$, так как она неотрицательна: $m(\Delta) = \mathsf{M}|\zeta(\Delta)|^2 \geqslant 0$, $m(\varnothing) = 0$, и аддитивна: $m(\Delta_1 \cup \Delta_2) = m(\Delta_1) + m(\Delta_2)$, если $\Delta_1 \cap \Delta_2 = \varnothing$.

Стохастический интеграл от простой функции $f(x) = \sum_{k=1}^{n} c_k \chi_{\Delta_k}(x)$, $\Delta_k \in \mathfrak{M}$, заданной на E, по элементарной стохастической мере $\zeta(\Delta)$ определяется соотношением

$$\int f(x)\,\zeta(dx) = \sum_{k=1}^{n} c_k \zeta(\Delta_k). \tag{2.1}$$

Предположим, что элементарная мера $m(\Delta)$ полуаддитивна; тогда она может быть продолжена до полной меры $\{E, \mathscr{E}, m\}$.

Введем гильбертово пространство $\mathscr{L}_2(E, \mathscr{E}, m)$ функций со скалярным произведением

$$(f, g) = \int f(x)\,\overline{g(x)}\,m(dx).$$

Стохастический интеграл от $f(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$ по элементарной стохастической мере $\zeta(\Delta)$ определяется соотношением

$$\int f(x)\,\zeta(dx) = \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \int f_n(x)\,\zeta(dx) \tag{2.2}$$

для произвольной последовательности простых функций $f_n(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$ таких, что

$$\int |f(x) - f_n(x)|^2 m(dx) \to 0, \quad n \to \infty. \tag{2.3}$$

Теорема 1. *Для произвольной последовательности функций $f_n(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$ таких, что выполняется условие (2.3), имеет место соотношение (2.2). Для любых $f(x)$ и $g(x)$ из $\mathscr{L}_2(E, \mathscr{E}, m)$ имеют место равенства*

$$\int [\alpha f(x) + \beta g(x)] \zeta(dx) = \alpha \int f(x) \zeta(dx) + \beta \int g(x) \zeta(dx), \tag{2.4}$$

где α, β — произвольные постоянные, и

$$\mathsf{M} \int f(x) \zeta(dx) \overline{\int g(x) \zeta(dx)} = \int f(x) \overline{g(x)} m(dx). \tag{2.5}$$

Замечание. Равенство (2.5) означает изометрическое соответствие между $\mathscr{L}_2(E, \mathscr{E}, m)$ и $\mathscr{L}_2(\zeta)$ — гильбертовым пространством случайных величин $\eta = \int f(x) \zeta(dx)$ с $f(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$.

Изометрическое соотношение между $\mathscr{L}_2(E, \mathscr{E}, m)$ и $\mathscr{L}_2(\zeta)$ можно положить в основу определения стохастического интеграла.

Пусть L_0 — класс всех множеств $A \in \mathscr{E}$, для которых $m(A) < \infty$. Случайная функция множеств

$$\tilde{\zeta}(A) = \int \chi_A(x) \zeta(dx) = \int_A \zeta(dx) \tag{2.6}$$

является стохастической ортогональной мерой на L_0, удовлетворяющей условиям:

а) $\tilde{\zeta}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \tilde{\zeta}(A_n)$,

если $\bigcup_{n=1}^{\infty} A_n \in L_0$ и $A_k \cap A_r = \varnothing$ при $k \neq r$;

б) $\mathsf{M}\tilde{\zeta}(A)\overline{\tilde{\zeta}(B)} = m(A \cap B), \quad A, B \in L_0$;

в) $\tilde{\zeta}(\Delta) = \zeta(\Delta), \Delta \in \mathfrak{M}$.

Теорема 2. *Если структурная функция $m(\Delta)$ элементарной стохастической меры $\zeta(\Delta)$ полуаддитивна, то $\{\zeta(\Delta), \Delta \in \mathfrak{M}\}$ может быть продолжена до стохастической меры $\{\tilde{\zeta}(A), A \in L_0\}$, причем*

$$\int f(x) \zeta(dx) = \int f(x) \tilde{\zeta}(dx). \tag{2.7}$$

10.2.2. Свойства стохастического интеграла. Пусть ζ — ортогональная стохастическая мера со структурной функцией m, являющейся полной мерой на $\{E, \mathscr{E}\}$. Положим для $g(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$

$$\lambda(A) = \int \chi_A(x) g(x) \zeta(dx), \quad A \in \mathscr{E}. \tag{2.8}$$

1) Тогда $\lambda(A)$ — ортогональная стохастическая мера на $\{E, \mathscr{E}\}$ со структурной функцией

$$l(A) = \int_A |g(x)|^2 m(dx). \tag{2.9}$$

2) Если $f(x) \in \mathscr{L}_2(E, \mathscr{E}, l)$, то $f(x) g(x) \in \mathscr{L}_2(E, \mathscr{E}, m)$ и

$$\int f(x) \lambda(dx) = \int f(x) g(x) \zeta(dx). \tag{2.10}$$

3) Если $m(A) < \infty$, то

$$\zeta(A) = \int \frac{\chi_A(x)}{g(x)} \lambda(dx). \tag{2.11}$$

4) Пусть T — конечный или бесконечный отрезок вещественной прямой; B — σ-алгебра борелевских подмножеств T; l — мера Лебега.

Теорема 3. *Для борелевской функции* $g(t, x) \in \mathscr{L}_2(T \times E, B \times \mathscr{E}, l \times m)$ *и* $g(t, x) \in \mathscr{L}_2(E, \mathscr{E}, m)$ *при каждом* $t \in T$ *стохастический интеграл*

$$\xi(t) = \int g(t, x) \zeta(dx) \tag{2.12}$$

можно определить как функцию от t таким образом, чтобы процесс $\xi(t)$ был измерим.

5) Если $g(t, s)$ и $h(t)$ — борелевские функции,

$$\int_a^b \int_{-\infty}^{\infty} |g(t, s)|^2 dt\, m(ds) < \infty, \int_a^b |h(t)|^2 dt < \infty, \tag{2.13}$$

то

$$\int_a^b h(t) \int_{-\infty}^{\infty} g(t, s) \zeta(ds) dt = \int_{-\infty}^{\infty} \zeta(ds) \int_a^b h(t) g(t, s) dt. \tag{2.14}$$

Замечание. Соотношение (2.14) имеет место и при $a = -\infty$, $b = +\infty$, если существует несобственный интеграл $\int_{-\infty}^{\infty} h(t) g(t, s) dt$ в смысле сходимости в $\mathscr{L}_2(E, \mathscr{E}, m)$.

6) Пусть $\{\xi(t),\ t \in [a, b]\}$ — *процесс с ортогональными приращениями*:

$$\mathsf{M}\,(\xi(t_2) - \xi(t_1))\,\overline{(\xi(t_4) - \xi(t_3))} = 0$$

для любых $t_1 < t_2 < t_3 < t_4$, принадлежащих $[a, b)$, непрерывный (в с. к.) слева:

$$\mathsf{M}\,|\xi(t) - \xi(s)|^2 \to 0, \quad s \uparrow t.$$

Пусть $\mathfrak{M}$ — класс всех полуинтервалов $\Delta = [t_1, t_2)$, $a \leqslant t_1 < t_2 \leqslant b$. Определим элементарную стохастическую меру

$$\zeta([t_1, t_2)) = \xi(t_2) - \xi(t_1) \tag{2.15}$$

со структурной функцией

$$m([t_1, t_2)) = \mathsf{M}\,|\,\xi(t_2) - \xi(t_1)\,|^2 = F(t_2) - F(t_1), \tag{2.16}$$

где

$$F(t) = \mathsf{M}\,|\,\xi(t) - \xi(a)\,|^2. \tag{2.17}$$

Функция $F(t)$ монотонно неубывающая непрерывная слева. Поэтому структурная функция (2.16) допускает продолжение до полной меры на $[a, b)$. Следовательно, определен стохастический интеграл Стилтьеса по процессу с независимыми приращениями равенством

$$\int_a^b f(t)\,d\xi(t) = \int_a^b f(t)\,\zeta(dt) \tag{2.18}$$

для произвольной борелевской функции $f(t) \in \mathscr{L}_2(F)$. Определение интеграла (2.18) сохраняется и при $b = +\infty$.

10.2.3. Стохастический интеграл по векторной мере. Стохастический интеграл обобщается на *векторные стохастические меры.* Пусть $\zeta(\Delta) = \{\zeta^1(\Delta), \zeta^2(\Delta), \ldots, \zeta^p(\Delta)\}$ — векторная стохастическая (ортогональная) мера на $\mathfrak{M}$ со *структурной матрицей* $m(\Delta) = \{m_j^k(\Delta)\} = \mathsf{M}\zeta(\Delta)\,\zeta^T(\Delta)$ $\left(m_j^k(\Delta) = \mathsf{M}\zeta^k(\Delta)\,\overline{\zeta^j(\Delta)},\ 1 \leqslant j, k \leqslant p\right)$ удовлетворяет условиям:

1) $\zeta(\Delta_1 \cup \Delta_2) = \zeta(\Delta_1) + \zeta(\Delta_2) \pmod{\mathsf{P}}$, если $\Delta_1 \cap \Delta_2 = \varnothing$;

2) $\mathsf{M}\zeta^k(\Delta_1)\,\overline{\zeta^j(\Delta_2)} = m_j^k(\Delta_1 \cap \Delta_2)$; $\Delta_1, \Delta_2 \in \mathfrak{M}$, $1 \leqslant k, j \leqslant p$;

3) $\mathsf{M}\,|\,\zeta(\Delta)\,|^2 = \mathsf{M}\zeta(\Delta)\,\overline{\zeta(\Delta)} = \mathsf{M}\sum\limits_{k=1}^{p} |\,\zeta^k(\Delta)\,|^2 < \infty$, $\zeta(\varnothing) = 0$.

Положим $m_0(\Delta) = \operatorname{Sp} m(\Delta) = \sum\limits_{k=1}^{p} m_k^k(\Delta)$.

Если след $m_0(\Delta)$ матрицы $m(\Delta)$ — полуаддитивная функция на $\mathfrak{M}$, то $m_j^k(\Delta)$ могут быть продолжены до счетно-аддитивных функций множеств на $\{E, \mathscr{E}\}$. Полная матричная мера $m(\Delta)$ на $\{E, \mathscr{E}\}$ обладает свойством положительной определенности:

$$\sum_{j,k=1}^{n} z_j^T m\left(\Delta_j \cap \Delta_k\right) z_k = \mathsf{M}\left|\sum_{k=1}^{n} \zeta^T\left(\Delta_k\right) z_k\right|^2 \geqslant 0 \tag{2.19}$$

для любых векторов $z_k = \left\{z_k^1, z_k^2, \ldots, z_k^p\right\}$ и любых $\Delta_k \in \mathscr{E}$ $(1 \leqslant k \leqslant p)$.

Здесь $\zeta^T(\Delta)$ — вектор-строка с компонентами $\bar{\zeta}^j(\Delta)$ $(j = 1, 2, \ldots, p)$.

Введем пространство $\mathscr{L}_0\{m_0\}$ простых функций $f(x) = \sum\limits_{k=1}^{n} c_k \chi_{\Delta_k}$ $(\Delta_k \in m_0)$ со скалярным произведением

$$(f, g) = \int f(x)\,\overline{g(x)}\,m_0(dx).$$

Далее определим пространство $\mathscr{L}_0^p(\zeta)$ случайных векторов $\eta=\sum_{k=1}^{n} c_k\zeta(\Delta_k)$ $(\Delta_k\in\mathfrak{M})$ со скалярным произведением $(\eta_1,\eta_2)=\mathrm{M}\eta_1\bar\eta_2$.

Замыкание (в смысле с. к. сходимости) пространства случайных величин $\mathscr{L}_0^p\{\zeta\}$ обозначим через $\mathscr{L}_2^p(\zeta)$, а пополнение $\mathscr{L}_0\{m_0\}$ — через $\mathscr{L}_2(m_0)$.

Равенство

$$\eta=\int f(x)\,\zeta(dx)=\sum_{k=1}^{n} c_k\zeta(\Delta_k) \tag{2.20}$$

для $f(x)=\sum_{k=1}^{n} c_k\chi_{\Delta_k}(x)$ устанавливает изометрическое отображение $\eta=\mathfrak{F}(f)$ пространства $\mathscr{L}_0(m_0)$ на $\mathscr{L}_0^p(\zeta)$, которое может быть продолжено до изометрического соответствия $\eta=\mathfrak{F}(f)$ $\mathscr{L}_2(m_0)$ на $\mathscr{L}_2^p(\zeta)$. При этом случайный вектор $\eta=\mathfrak{F}(f)$ называют *стохастическим интегралом* и пишут

$$\eta=\int f(x)\,\zeta(dx),\quad f(x)\in\mathscr{L}_2(m_0). \tag{2.21}$$

Приведенные выше свойства стохастического интеграла сохраняются и в данном случае.

10.2.4. Интегральное представление случайных функций. С помощью стохастических интегралов можно получать интегральные представления различных классов случайных функций.

Теорема. *Пусть задана p-мерная векторная случайная функция $\{\xi(x),\ x\in\mathscr{X}\}$ с ковариационной матрицей $B(x_1,x_2)=\mathrm{M}\xi(x_1)\times$ $\times\xi^T(x_2)=\{B_j^k(x_1,x_2)\}$, $B_j^k(x_1,x_2)=\mathrm{M}\xi^k(x_1)\overline{\xi^j(x_2)}$ $(1\leqslant j,k\leqslant p)$, допускающей интегральное представление*

$$B(x_1,x_2)=\int g(x_1,u)\overline{g(x_2,u)}\,m(du), \tag{2.22}$$

в котором $m(\Delta)$ — положительно определенная матричная мера на $\{\mathfrak{A},\mathfrak{B}\}$ с $m_0(u)=\mathrm{Sp}\,m(u)$, $g(x,u)$ — скалярная функция, удовлетворяющая условиям: 1) $g(x,u)\in\mathscr{L}_2\{\mathfrak{A},\mathfrak{B},m_0\}$ при каждом $x\in\mathscr{X}$; 2) семейство функций $\{g(x,u),\ x\in\mathscr{X}\}$ полно в $\mathscr{L}_2\{\mathfrak{A},\mathfrak{B},m_0\}$.

Тогда существует стохастическая ортогональная векторная мера $\{\zeta(B),\ B\in\mathfrak{B}\}$ со структурной матричной функцией $m(B)=$ $=\mathrm{M}\zeta(B)\zeta^T(B)$ такая, что с вероятностью 1 при каждом x случайная функция $\{\xi(x),\ x\in\mathscr{X}\}$ представима в виде

$$\xi(x)=\int g(x,u)\,\zeta(du). \tag{2.23}$$

При этом стохастическая мера $\{\zeta(B),\ B\in\mathfrak{B}\}$ подчинена случайной функции $\{\xi(x),\ x\in\mathscr{X}\}$ в том смысле, что $\zeta(B)\in\mathscr{L}_2^p(\xi)$ при каждом $B\in\mathfrak{B}$.

Стохастическая мера $\{\zeta(B),\ B\in\mathfrak{B}\}$ определяется с помощью изометрического соответствия между пространствами $\mathscr{L}_2(\xi)$ и $\mathscr{L}_2(m)$, при котором:

а) $\xi(x) \longleftrightarrow g(x, u)$, $\zeta(B) \longleftrightarrow \chi_B(u)$;

б) если $\eta_i \longleftrightarrow f_i(u)$ $(i = 1, 2)$, то

$$\eta_i = \int f_i(u)\,\zeta(du); \quad \mathsf{M}\eta_1\eta_2^T = \int f_1(u)\,\overline{f_2(u)}\,m(du).$$

10.3. Линейный прогноз и фильтрация гильбертовых случайных функций

Гильбертова случайная функция $\{\xi(x),\ x\in\mathscr{X}\}$ со значениями в измеримом пространстве $\{E, \mathscr{E}\}$ порождает гильбертово пространство случайных величин $\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}$, являющееся замкнутой (в смысле с. к. сходимости) линейной оболочкой семейства случайных величин $\{\xi(x),\ x\in\mathscr{X}\}$ и констант.

Гильбертово пространство $\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}$ является подпространством гильбертова пространства $\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ всех гильбертовых случайных величин, заданных на том же основном вероятностном пространстве $\{\Omega, \mathfrak{S}, \mathsf{P}\}$, что и семейство гильбертовых случайных величин $\{\xi(x),\ x\in\mathscr{X}\}$.

Наилучшее линейное приближение (*оценка*) $\tilde\zeta$ гильбертовой случайной величины $\zeta\in\mathscr{L}_2\{\Omega, \mathfrak{S}, \mathsf{P}\}$ в пространстве $\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}$ определяется условием:

$$\mathsf{M}\,|\tilde\zeta-\zeta|^2 \leqslant \mathsf{M}\,|\zeta'-\zeta|^2, \quad \zeta'\in\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}. \tag{3.1}$$

Условие (3.1) означает, что оценка $\tilde\zeta$ имеет *минимальную среднеквадратическую погрешность.*

Из теории гильбертовых пространств следует, что элемент $\tilde\zeta$ является *проекцией* ζ на подпространство $\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}$ и определяется единственным (mod P) образом системой линейных уравнений:

$$\mathsf{M}\tilde\zeta\bar\xi(x) = \mathsf{M}\zeta\bar\xi(x), \quad x\in\mathscr{X}. \tag{3.2}$$

Среднеквадратическая погрешность δ приближенного равенства $\tilde\zeta\approx\zeta$ равна длине перпендикуляра, опущенного из конца вектора ζ на подпространство $\mathscr{L}_2\{\xi(x),\ x\in\mathscr{X}\}$ и дается формулой

$$\delta^2 = \mathsf{M}\,|\tilde\zeta-\zeta|^2. \tag{3.3}$$

В частности, выполняется условие *несмещенности оценки*:

$$\mathsf{M}\tilde\zeta = \mathsf{M}\zeta.$$

Наилучшая линейная оценка $\tilde\zeta$, определяемая системой (3.2), является линейной функцией от $\xi(x)$ с конечной дисперсией.

Задача построения оценки $\tilde\zeta$ возникает при *линейном прогнозе* (*экстраполяции*) случайного процесса $\{\xi(t),\ t\in T\}$, когда требуется оценить значение $\xi(t^*)$ в некоторый момент времени t^* по значениям процесса $\xi(t)$ на множестве моментов времени T, предшествовавших t^*.

Другим примером построения оценки $\tilde{\zeta}$ является *задача линейной фильтрации* случайного процесса, которая состоит в следующем. Наблюдается процесс $\xi(t) = \zeta(t) + \eta(t)$, представляющий собой сумму *полезного сигнала* $\zeta(t)$ и *шума* $\eta(t)$. Требуется отделить сигнал от шума, т. е. для данного t^* нужно найти наилучшие линейные приближения $\tilde{\zeta} \in \mathscr{L}_2\{\xi(t),\ t \in T\}$ для сигнала $\zeta(t^*)$.

Разумеется, линейная оценка $\tilde{\zeta}$ не всегда является приемлемой с практической точки зрения. Однако в весьма важном случае, когда все конечномерные распределения случайных величин $\{\xi(x),\ x \in \mathscr{X}\}$ нормальны и $\mathsf{M}\xi(x) = 0$, $\mathsf{M}\zeta = 0$, наилучшая линейная оценка в $\mathscr{L}_2\{\xi(x),\ x \in \mathscr{X}\}$ является также наилучшей в с. к. смысле.

В этом случае

$$\tilde{\zeta} = \mathsf{M}\,[\zeta \mid \sigma(\xi(x),\ x \in \mathscr{X})], \tag{3.4}$$

где $\sigma(\xi(x),\ x \in \mathscr{X})$ — σ-алгебра, порожденная совокупностью случайных величин $\{\xi(x),\ x \in \mathscr{X}\}$.

Примеры. 1. Пусть задана конечная совокупность линейно независимых гильбертовых случайных величин $\{\xi_k,\ k = 1, 2, \ldots, n\}$. Решение системы линейных уравнений (3.2) определяется формулой

$$\tilde{\zeta} = \frac{1}{\Gamma}\begin{vmatrix} (\xi_1, \xi_1) & \ldots & (\xi_1, \xi_n) & \xi_1 \\ \cdot & \cdot & \cdot & \cdot \\ (\xi_n, \xi_1) & \ldots & (\xi_n, \xi_n) & \xi_n \\ (\zeta, \xi_1) & \ldots & (\zeta, \xi_n) & 0 \end{vmatrix}, \tag{3.5}$$

где $\Gamma = \Gamma(\xi_1, \xi_2, \ldots, \xi_n)$ — определитель Грама системы величин $\{\xi_k,\ k = 1, 2, \ldots, n\}$:

$$\Gamma(\xi_1, \xi_2, \ldots, \xi_n) = \begin{vmatrix} (\xi_1, \xi_1) & \ldots & (\xi_1, \xi_n) \\ \cdot & \cdot & \cdot \\ (\xi_n, \xi_1) & \ldots & (\xi_n, \xi_n) \end{vmatrix}. \tag{3.6}$$

Среднеквадратическая погрешность $\delta^2 = \mathsf{M}|\tilde{\zeta} - \zeta|^2$ определяется равенством

$$\delta^2 = \frac{\Gamma(\xi_1, \xi_2, \ldots, \xi_n, \zeta)}{\Gamma(\xi_1, \xi_2, \ldots, \xi_n)}. \tag{3.7}$$

2. Пусть задан гильбертов непрерывный случайный процесс $\{\xi(t),\ t \in [a, b]\}$ с корреляционной функцией

$$R(t, s) = \mathsf{M}\,[(\xi(t) - a(t))\,\overline{(\xi(s) - a(s))}], \quad a(t) = \mathsf{M}\xi(t).$$

Процесс $\{\xi(t),\ t \in [a, b]\}$ может быть разложен в ортогональный ряд

$$\xi(t) = \sum_{n=1}^{\infty} \xi_n \varphi_n(t), \tag{3.8}$$

в котором $\mathsf{M}\xi_n\bar{\xi}_m = \lambda_n\delta_{nm}$; λ_n и $\varphi_n(t)$ — собственные значения и собственные функции корреляционной функции $R(t, s)$ на $[a, b]$:

$$\int_a^b R(t, s)\,\varphi_n(s)\,ds = \lambda_n\varphi_n(t).$$

В этом случае наилучшая линейная оценка $\tilde{\zeta}$ определяется соотношениями

$$\tilde{\zeta} = \sum_{n=1}^{\infty} c_n \xi_n, \quad \xi_n = \int_a^b \overline{\varphi_n(t)}\, \xi(t)\, dt, \tag{3.9}$$

$$c_n = \frac{1}{\lambda_n} \mathsf{M}\zeta\bar{\xi}_n = \frac{1}{\lambda_0} \int_a^b R_{\zeta\xi}(t)\, \varphi_n(t)\, dt, \tag{3.10}$$

где

$$R_{\zeta\xi}(t) = \mathsf{M}\zeta\overline{\xi(t)}.$$

Среднеквадратическая погрешность оценки определяется формулой

$$\delta^2 = \mathsf{M}\,|\zeta|^2 - \sum_{n=0}^{\infty} |c_n|^2. \tag{3.11}$$

Практическое применение приведенных формул возможно при условии знания собственных чисел и собственных функций ядра $R(t, s)$.

Литература: [19, 20, 50, 60, 70, 98].

Глава 11. СТАЦИОНАРНЫЕ ПРОЦЕССЫ

Важное место в теории случайных процессов занимают процессы, те или иные характеристики которых остаются неизменными при сдвигах временно́го или пространственного параметра, или, в более общем случае, процессы, определенные характеристики которых инвариантны относительно некоторой группы или полугруппы преобразований.

Такие процессы обладают определенными свойствами неизменности, стационарности. Чаще всего в качестве характеристик, инвариантных относительно заданной группы или полугруппы преобразований, выступают либо моменты, либо конечномерные распределения.

В первом случае говорят о стационарности r-го порядка, если свойством инвариантности обладают моменты до r-го порядка включительно. Наиболее важный класс составляют процессы, обладающие свойством стационарности второго порядка, обычно называемые *стационарными в широком смысле.*

Если в качестве инвариантных характеристик выбираются конечномерные распределения, то говорят о стационарности в узком смысле и соответствующие процессы называют *стационарными в узком смысле.*

11.1. Стационарные в широком смысле случайные процессы

11.1.1. Основные определения. Пусть $(\Omega, \mathfrak{S}, \mathsf{P})$ — фиксированное вероятностное пространство, на котором рассматриваются случайные процессы $\{\xi(t),\ t \in T\}$, где T — одно из множеств вида $(-\infty, \infty)$, $[0, \infty)$ (непрерывное время) либо $\{0, \pm1, \pm2, \ldots\}$, $\{0, 1, 2, \ldots\}$ (дискретное время).

Процесс $\xi(t)$ может принимать значения либо в $\mathbf{R}=(-\infty, \infty)$ (действительный скалярный процесс), либо в комплексной плоскости $\mathbf{Z}$ (комплексный скалярный процесс), либо в евклидовом пространстве $\mathbf{R}^k$ (k-мерный действительный процесс), либо в k-мерном комплексном пространстве $\mathbf{Z}^k$ (k-мерный комплексный процесс).

Если $\{\xi(t),\ t\in T\}$ — скалярный процесс и существует $\mathsf{M}|\xi(t)|^2$ для $t\in T$, то функция $C(t, s)=\mathsf{M}\xi(t)\overline{\xi(s)}$ называется *ковариационной функцией*, а функция $B(t,s)=\mathsf{M}[\xi(t)-\mathsf{M}\xi(t)]\overline{[\xi(s)-\mathsf{M}\xi(s)]}$ — *корреляционной функцией* процесса $\{\xi(t),\ t\in T\}$.

Если $\xi(t)=(\xi_1(t), \xi_2(t), \ldots, \xi_k(t))$, $t\in T$, — векторный процесс, для которого существуют $\mathsf{M}\xi_l(t)\overline{\xi_m(t)}$ $(l, m=1, \ldots, k)$, то матрица $C(t, s)=\{c_{lm}(t, s);\ l, m=1, \ldots, k\}$, где $c_{lm}(t, s)= \mathsf{M}\xi_l(t)\overline{\xi_m(s)}$, называется *ковариационной матрицей*, а матрица $B(t, s)=\{B_{lm}(t, s);\ l, m=1, \ldots, k\}$, где $B_{lm}(t, s)=\mathsf{M}[\xi_l(t)-\mathsf{M}\xi_l(t)]\overline{[\xi_m(s)-\mathsf{M}\xi_m(s)]}$, — *корреляционной матрицей* процесса $\{\xi(t),\ t\in T\}$. Ковариационную матрицу $C(t, s)$ можно представить в виде

$$C(t, s)=\mathsf{M}\xi(t)\xi^T(s), \tag{1.1}$$

где знак ^{T} обозначает транспонирование вектор-столбца $\xi(s)$ и переход к комплексно-сопряженным элементам.

В скалярном случае удобно считать, что $\xi^T(s)=\overline{\xi(s)}$, поэтому формулу (1.1) можно использовать как в скалярном, так и в векторном случае. В скалярном и векторном случаях $C(t, s)$ и $B(t, s)$ часто называют *ковариационной* и *корреляционной функцией* соответственно.

Случайный процесс $\{\xi(t),\ t\in T\}$ называется *стационарным в широком смысле*, если его математическое ожидание $\mathsf{M}\xi(t)=a$ не зависит от t, а корреляционная функция $B(t, s)$ зависит лишь от разности $(t-s)$, т. е. если

$$B(t, s)=B(t-s). \tag{1.2}$$

Таким образом, если $\{\xi(t),\ t\in T\}$ — стационарный в широком смысле процесс, то процесс $\{\xi_u(t),\ t\in T\}$, где $\xi_u(t)=\xi(t+u)$ и $u\in T$ фиксировано, имеет одинаковое с $\xi(t)$ математическое ожидание $\mathsf{M}\xi_u(t)=a$ и корреляционную функцию $B(t)=B(t, 0)$.

Дисперсия стационарного в широком смысле процесса совпадает с $B(0)$: $\mathsf{M}[\xi(t)-a][\xi(t)-a]^T=B(0)$.

Корреляционная и ковариационная функции связаны соотношением

$$C(t, s)=B(t, s)+aa^T;$$

следовательно, для стационарных процессов $C(t, s)=C(t-s)$.

Если $\mathsf{M}\xi(t)=0$, то корреляционная и ковариационная функции совпадают. Более того, если $C(t)=C(t, 0)$, то формула

$$C(t)=B(t)+aa^T$$

показывает, что, не уменьшая общности, можно рассматривать процесс с нулевым математическим ожиданием, так как всегда можно перейти к процессу $\xi_0(t)=\xi(t)-\mathsf{M}\xi(t)$, для которого это свойство выполнено.

Все рассматриваемые в этой главе процессы с непрерывным временем предполагаются среднеквадратично (с. к.) непрерывными справа, т. е.

$$\lim_{s \downarrow t} \mathsf{M}\|\xi(s) - \xi(t)\|^2 = 0, \quad s, t \in T.$$

Пусть $\{\xi(t),\ t \in T\}$ и $\{\eta(t),\ t \in T\}$ — стационарные в широком смысле процессы с корреляционными функциями $B_\xi(t)$ в $B_\eta(t)$ соответственно.

Процессы $\{\xi(t),\ t \in T\}$ и $\{\eta(t),\ T \in T\}$ называются *стационарно связанными*, если их *взаимная корреляционная функция* $B_{\xi\eta}(ts) = \mathsf{M}\xi(t)\eta^T(s)$ зависит лишь от разности $(t-s)$. В связи с этим $B_\xi(t)$ и $B_\eta(t)$ называют иногда *автокорреляционными функциями*.

Корреляционная функция $B(t)$ стационарного процесса обладает следующими свойствами.

1) Эрмитовость:

$$B(t) = \begin{cases} \overline{B(-t)} & \text{(скалярный процесс)}, \\ B^T(-t) & \text{(векторный процесс)}. \end{cases} \tag{1.3}$$

2) Неотрицательная определенность:

$$\sum_{l,\, m=1}^{N} B(t_l - t_m)\lambda_l\overline{\lambda}_m \geqslant 0 \quad \text{(скалярный процесс)}, \tag{1.4}$$

каковы бы ни были $N \geqslant 1$, $t_l \in T$ и комплексные числа λ_l $(l = 1, \ldots, N)$;

$$\sum_{l,\, m=1}^{N} z_l^T B(t_l - t_m) z_m \geqslant 0 \quad \text{(векторный процесс)}, \tag{1.4'}$$

каковы бы ни были $N \geqslant 1$, $t_l \in T$ и векторы z_l $(l = 1, \ldots, N)$.

3) $|B(t)| \leqslant B(0)$ (скалярный процесс),

$$|B_{lm}(t)|^2 \leqslant B_{ll}(0)\,B_{mm}(0) \quad (l, m = 1, \ldots, k) \quad \text{(векторный процесс)}.$$

4) Если (в случае непрерывного времени) корреляционная функция $B(t)$ непрерывна в точке $t = 0$, то она непрерывна и в любой другой точке $t \in T$.

5) Если $\xi(t) = (\xi_1(t), \ldots, \xi_k(t))$ — векторный процесс, то функции

$$\rho_{lm}(t) = \frac{B_{lm}(t)}{\sqrt{B_{ll}(0)\,B_{mm}(0)}},$$

называемые *коэффициентами взаимной корреляции* компонент $\xi_l(t)$ и $\xi_m(t)$, удовлетворяют неравенству

$$-1 \leqslant \rho_{lm}(t) \leqslant 1, \quad l, m = 1, \ldots, k,$$

и определяют степень линейной зависимости процессов $\xi_l(t)$ и $\xi_m(t)$.

Различные стационарные процессы могут иметь одинаковые математическое ожидание и ковариационную функцию.

Если $\{\xi(t),\ t \in T\}$ — стационарный процесс и для любого $n \geqslant 1$ и любого набора $\{t_k \in T,\ k = 1, \ldots, n\}$ вектор $(\xi(t_1),\ \xi(t_2), \ldots$

..., $\xi(t_n)$) имеет многомерное нормальное распределение, то $\{\xi(t), t \in T\}$ называется *стационарным гауссовским* (или *стационарным нормальным*) процессом.

Стационарный гауссовский процесс определяется своим математическим ожиданием и ковариационной функцией.

Обратно, каждая функция $m(t) \equiv \text{const}$ и функция $B(t)$, обладающая свойствами (1.3) и (1.4), задают некоторый стационарный гауссовский процесс.

Пусть $\{\xi(t), t \in T\}$ — скалярный (действительный или комлексный) стационарный процесс, $\xi^{(N)} = \sum_{l=1}^{N} \lambda_l \xi(t_l)$, где $N \geqslant 1$ — любое целое число, $t_l \in T$, λ_l — произвольные числа. Линейная оболочка $H_0\{\xi\}$ всех таких случайных величин, построенных по процессу $\{\xi(t), t \in T\}$, является подпространством гильбертова пространства $\mathscr{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$ интегрируемых с квадратом $\mathfrak{S}$-измеримых функций на Ω по мере P. Введем в $H_0\{\xi\}$ скалярное произведение по формуле

$$(\xi, \eta) = \mathsf{M}\xi\bar{\eta}. \tag{1.5}$$

В случае векторных процессов аналогично определим линейную оболочку $H_0\{\xi\}$ случайных векторов, построенных по процессу $\{\xi(t), t \in T\}$, в которой вводится скалярное произведение

$$(\xi, \eta) = \operatorname{Sp} \mathsf{M}\xi\eta^T, \tag{1.6}$$

где $\operatorname{Sp} B$ означает след матрицы B; $H_0\{\xi\}$ является подпространством гильбертова пространства $\mathscr{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$ интегрируемых с квадратом нормы по мере P $\mathfrak{S}$-измеримых векторнозначных функций на Ω. Гильбертово пространство H_ξ, являющееся замыканием $H_0\{\xi\}$ по норме, порожденной скалярным произведением (1.5) (соответственно (1.6)), называется *пространством значений стационарного в широком смысле процесса* $\{\xi(t), t \in T\}$. С геометрической точки зрения процесс $\{\xi(t), t \in T\}$ является кривой в пространстве $\mathscr{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$ и H_ξ есть пересечение всех подпространств в $\mathscr{L}_2(\Omega, \mathfrak{S}, \mathsf{P})$, содержащих эту кривую.

11.1.2. Примеры. 1. Пусть ζ_m $(m = 1, \ldots, p)$ — набор некоррелированных случайных величин таких, что

$$\mathsf{M}\zeta_m = 0, \quad \mathsf{M}\zeta_m\bar{\zeta}_l = \delta_{ml}\sigma_m^2,$$

где δ_{ml} — символ Кронекера, $\sigma_m^2 < \infty$. Если $p = \infty$, то пусть $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$ и пусть λ_m $(m = 1, \ldots, p)$ — набор произвольных действительных чисел.

Процесс $\{\xi(t), t \in T\}$, где $\xi(t) = \sum_{m=1}^{p} e^{i\lambda_m t} \zeta_m$, является стационарным в широком смысле (см. § 10.1), так как $\mathsf{M}\xi = 0$ и $B(t, s) = \sum_{m=1}^{p} e^{i\lambda_m (t-s)} \sigma_m^2 = B(t-s)$.

2. Пусть ζ_m $(m = 1, \ldots, p)$ — набор некоррелированных случайных k-мерных векторов таких, что $\mathsf{M}\zeta_m = 0$, $\mathsf{M}\zeta_m\zeta_l^T = \delta_{ml}G_m$, где G_m — эрмитова неотрицательно определенная $k \times k$-матрица $\left(\sum_{m=1}^{p} G_m < \infty, \text{ если } p = \infty\right)$, и пусть λ_m $(m = 1, \ldots, p)$ — набор произвольных действительных чисел. Процесс $\{\xi(t),\ t \in T\}$, где $\xi(t) = \sum_{m=1}^{p} e^{i\lambda_m t}\zeta_m$, является стационарным в широком смысле, так как $\mathsf{M}\xi(t) = 0$, $B(t, s) = \sum_{m=1}^{p} e^{i\lambda_m (t-s)} G_m = B(t-s)$.

3. Пусть $\xi(t) = \cos(t\eta + \varphi)$, где φ — случайная величина, равномерно распределенная на $[0, 2\pi]$, а случайная величина η не зависит от φ и имеет функцию распределения $G(x)$. Процесс $\{\xi(t),\ t \in T\}$ является (действительным) стационарным в широком смысле, так как

$$\mathsf{M}\xi(t) = 0, \quad B(t, s) = \frac{1}{2}\int_{-\infty}^{\infty} \cos[(t-s)x]\, G(dx) = B(t-s).$$

4. Пусть $T = [0, \infty)$ и $\{w(t),\ t \in T\}$ — стандартный винеровский процесс. Процесс $\{\xi(t),\ t \in T\}$, где $\xi(t) = w(t+h) - w(t)$ и $h > 0$ — фиксированное число, является (действительным) стационарным в широком смысле, так как $\mathsf{M}\xi(t) = 0$ и

$$B(t, s) = \begin{Bmatrix} 0, & \text{если } |t-s| \geqslant h, \\ h - |t-s|, & \text{если } |t-s| < h \end{Bmatrix} = B(t-s).$$

5. Пусть $T = \{0, \pm 1, \pm 2, \ldots\}$ и $\zeta(t)$ $(t \in T)$ — стандартная последовательность *некоррелированных случайных величин*, т. е. $\mathsf{M}\zeta(t) = 0$, $\mathsf{M}\zeta(m)\overline{\zeta(l)} = \delta_{ml}$, и пусть последовательность комплексных чисел $c(t)$ $(t \in T)$ такова, что $\sum_{t \in T} |c(t)|^2 < \infty$.

Случайная последовательность $\{\xi(t),\ t \in T\}$, где $\xi(t) = \sum_{s \in T} c(t-s)\,\zeta(s)$, является стационарной в широком смысле; так как $\mathsf{M}\xi(t) = 0$, то

$$B(t, s) = \sum_{m \in T} c(t-s+m)\,\overline{c(m)} = B(t-s).$$

6. Пусть $T = (-\infty, \infty)$ и $\{\zeta(t),\ t \in T\}$ — *стандартный k-мерный процесс с ортогональными приращениями*, т. е. $\mathsf{M}\zeta(t) = 0$, $\mathsf{M}[\zeta(t_1) - \zeta(t_2)][\zeta(s_1) - \zeta(s_2)]^T = 0$ для $t_2 < t_1 \leqslant s_2 < s_1$ и $\mathsf{M}[\zeta(t) - \zeta(s)][\zeta(t) - \zeta(s)]^T = |t-s|\,I$, где I — единичная $k \times k$-матрица, и пусть матричнозначная функция $C(t)$ $(t \in T)$ такая, что $\int_{-\infty}^{\infty} \operatorname{Sp} C(t)\, C^T(t)\, dt < \infty$.

Процесс $\{\xi(t),\ t \in T\}$, где $\xi(t) = \int_{-\infty}^{\infty} C(t-s)\,d\zeta(s)$ и интеграл понимается как предел в среднеквадратичном смысле, является стационарным в широком смысле; так как $\mathsf{M}\xi(t) = 0$, то

$$B(t,\ s) = \int_{-\infty}^{\infty} C(t-s+u)\,C^T(u)\,du = B(t-s).$$

Процессы в примерах 5 и 6 называются *процессами скользящего суммирования.*

11.2. Спектральное представление корреляционных функций

11.2.1. Теорема Бохнера — Хинчина. *Пусть $\{\xi(t),\ t \in T\}$ — стационарный в широком смысле процесс с корреляционной функцией $B(t)$.*

а) *Если $\{\xi(t),\ t \in T\}$ — скалярный процесс с дискретным временем, то*

$$B(t) = \begin{cases} \displaystyle\int_{-\pi}^{\pi} e^{i\lambda t}\,dF(\lambda), \\ \displaystyle\int_{0}^{\pi} [\cos\lambda t\,dC(\lambda) + \sin\lambda t\,dQ(\lambda)] \ \textit{в случае действительного процесса,} \end{cases} \tag{2.1}$$

где $F(\lambda)$ — неотрицательная неубывающая функция, определяемая по $B(t)$ однозначно, если потребовать, чтобы $F(-\pi) = 0$ и $F(\lambda)$ была непрерывной справа, $C(\lambda)$ — действительная четная неубывающая функция ограниченной вариации, $Q(\lambda)$ — действительная нечетная функция ограниченной вариации.

б) *Если $\{\xi(t),\ t \in T\}$ — векторный процесс с дискретным временем, то для $B(t)$ имеют место представления (2.1), где $F(\lambda)$ — матрица, приращения которой $F(\lambda_1) - F(\lambda_2)$ $(\lambda_1 \geqslant \lambda_2)$ эрмитовы и неотрицательно определены, $C(\lambda)$ — вещественная симметричная матрица, приращения которой $C(\lambda_1) - C(\lambda_2)$ $(\lambda_1 \geqslant \lambda_2)$ неотрицательно определены, $Q(\lambda)$ — вещественная кососимметричная матрица. Матрица $F(\lambda)$ определяется однозначно по $B(t)$, если потребовать, чтобы $F(-\pi) = 0$ (нулевая матрица) и $F(\lambda)$ была непрерывной справа (в смысле поэлементной сходимости).*

в) *Если $\{\xi(t),\ t \in T\}$ — скалярный процесс с непрерывным временем, то*

$$B(t) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} e^{i\lambda t}\,dF(\lambda) \\ \displaystyle\int_{0}^{\infty} [\cos t\lambda\,dC(\lambda) + \sin t\lambda\,dQ(\lambda)] \ \textit{в случае действительного процесса,} \end{cases} \tag{2.2}$$

где функции $F(\lambda)$, $C(\lambda)$ *и* $Q(\lambda)$ *определяются так же, как и в случае* а), *за исключением условия* $F(-\infty) = 0$.

г) *Если* $\{\xi(t), t \in T\}$ — *векторный процесс с непрерывным временем, то для* $B(t)$ *имеют место представления* (2.2), *где матрицы* $F(\lambda)$, $C(\lambda)$ *и* $Q(\lambda)$ *определяются так же, как и в случае* б), *за исключением условия* $F(-\infty) = 0$ (*нулевая матрица*).

Матрица $F(\lambda)$ называется (матричной) *спектральной функцией* процесса $\{\xi(t), t \in T\}$, а мера, порожденная спектральной функцией $F(\lambda)$, называется *спектральной мерой*. Матрицы $C(\lambda)$ и $Q(\lambda)$ называются соответственно (матричной) *коспектральной* и (матричной) *квадратурной спектральной* функциями.

Имеют место равенства:

$$B(0) = \begin{cases} F(\pi) & \text{(дискретное время)}, \\ F(\infty) & \text{(непрерывное время)}, \end{cases} \tag{2.3}$$

$$dC(\lambda) = \begin{cases} dF(\lambda), & \lambda = 0, \\ 2\,\mathrm{Re}\,dF(\lambda), & \lambda \neq 0, \end{cases} \tag{2.4}$$

$$dQ(\lambda) = \begin{cases} 0, & \lambda = 0, \\ 2\,\mathrm{Im}\,dF(\lambda), & \lambda \neq 0. \end{cases} \tag{2.5}$$

11.2.2. Примеры. 1. Пусть $T = (-\infty, \infty)$, $\xi(t) = \eta e^{it\xi}$, где случайные величины η и ξ независимы, $\mathsf{M}\eta = 0$, $\mathsf{D}\eta = \sigma_\eta^2$, случайная величина ξ имеет функцию распределения $G_\xi(x)$. Стационарный процесс $\{\xi(t), t \in T\}$ имеет корреляционную функцию $B_\xi(t) = \int_{-\infty}^{\infty} e^{it\lambda}\sigma_\eta^2\,dG_\xi(\lambda)$; следовательно, спектральная функция $F_\xi(\lambda)$ равна $F_\xi(\lambda) = \sigma_\eta^2 G_\xi(\lambda)$.

Пример показывает, что существуют стационарные процессы с любой наперед заданной спектральной функцией.

2. Пусть $T = \{0, \pm 1, \pm 2, \ldots\}$ и $\{\xi(t), t \in T\}$ — последовательность некоррелированных случайных величин таких, что $\mathsf{M}\xi(t) = 0$, $\mathsf{M}\xi(t)\overline{\xi(s)} = \delta_{ts}\sigma^2$. Тогда

$$B(t) = \begin{cases} \sigma^2, & t = 0; \\ 0, & t \neq 0, \end{cases}$$

следовательно, $F(\lambda) = \dfrac{\pi + \lambda}{2\pi}\sigma^2$.

3. Пусть $T = [0, \infty)$ и $\{\xi(t), t \in T\}$ — стационарный в широком смысле и *марковский в широком смысле* процесс. Последнее означает, что $a(s, u) = a(s, t)a(t, u)$, $s < t < u$, где

$$a(s, t) = \begin{cases} \dfrac{\mathsf{M}\xi(t)\overline{\xi(s)}}{\mathsf{M}|\xi(s)|^2}, & \text{если } \mathsf{M}|\xi(s)|^2 > 0, \\ 0, & \text{если } \mathsf{M}|\xi(s)|^2 = 0. \end{cases}$$

Корреляционная функция $B(t)$ имеет вид

$$B(t) = e^{-\alpha|t|}\sigma^2, \quad \alpha > 0, \quad \text{либо} \quad B(t) = e^{i\beta t}\sigma^2$$

(β — действительное число).

В первом случае спектральная функция

$$F(\lambda) = \frac{\sigma^2}{\pi} \operatorname{arctg} \frac{\lambda}{\alpha} + \frac{\sigma^2}{2}.$$

Во втором случае

$$F(\lambda) = \begin{cases} \sigma^2, & \lambda \geqslant \beta, \\ 0, & \lambda < \beta. \end{cases}$$

11.2.3. Спектральная плотность и моменты. Если $S_m = \int\limits_{\Lambda} \lambda^m dF(\lambda) < \infty$, где Λ совпадает с $[-\pi, \pi]$ в случае дискретного времени и с $(-\infty, \infty)$ в случае непрерывного времени, то S_m называется *m-м спектральным моментом.*

Теорема. $\int\limits_{\Lambda} \lambda^{2m} dF(\lambda) < \infty$ *тогда и только тогда, когда корреляционная функция* $B(t)$ *имеет производную порядка* $2m$ *в нуле.*

В частности, если (в случае непрерывного времени) $F(\lambda)$ изменяется лишь на ограниченном интервале, то существуют спектральные моменты любого порядка и корреляционная функция имеет в нуле производные любого порядка.

Для спектральной функции $F(\lambda)$ имеет место *разложение Лебега*:

$$F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda). \tag{2.6}$$

Здесь $F_1(\lambda)$ абсолютно непрерывна относительно меры Лебега, т. е.

$$F_1(\lambda) = \begin{cases} \int\limits_{-\pi}^{\lambda} f(\lambda)\, d\lambda & \text{(дискретное время)}, \\ \int\limits_{-\infty}^{\lambda} f(\lambda)\, d\lambda & \text{(непрерывное время)}; \end{cases}$$

$F_2(\lambda)$ может меняться только скачками в конечном или счетном множестве точек λ; $F_3(\lambda)$ непрерывна и имеет почти всюду по мере Лебега равную нулю производную. Производная $F_1'(\lambda) = f(\lambda)$ абсолютно непрерывной компоненты спектральной функции $F(\lambda)$ называется (матричной) *спектральной плотностью.*

Пусть $f_{lm}(\lambda)$ есть (l, m)-й элемент матричной спектральной плотности $f(\lambda)$ векторного стационарного процесса $\xi(t) = (\xi_1(t), \ldots, \xi_k(t))$. Величина

$$\rho_{lm}(\lambda) = \frac{f_{lm}(\lambda)}{\sqrt{f_{ll}(\lambda) f_{mm}(\lambda)}}$$

является комплексным аналогом коэффициента взаимной корреляции компонент $\xi_l(t)$ и $\xi_m(t)$ на частоте λ. В случае действительного процесса $\xi(t)$, согласно (2.4), (2.5),

$$|\rho_{lm}(\lambda)|^2 = \frac{c_{lm}^2(\lambda) + q_{lm}^2(\lambda)}{c_{ll}(\lambda)\, c_{mm}(\lambda)},$$

где $\quad c_{lm}(\lambda) = C'_{lm}(\lambda), \quad q_{lm}(\lambda) = Q'_{lm}(\lambda) \quad (l, m = 1, \ldots, k).$
Величина $\rho_{lm}(\lambda)$, называемая *коэффициентом когерентности*, описывает максимальную взаимную корреляцию компонент $\xi_l(t)$ и $\xi_m(t)$, достижимую за счет сдвига по фазе одной компоненты относительно другой. При этом величина соответствующего фазового сдвига равна

$$\theta_{lm}(\lambda) = \operatorname{arctg} \frac{q_{lm}(\lambda)}{c_{lm}(\lambda)}.$$

Если $\{\xi(t), t \in T\}$ — векторный процесс с абсолютно непрерывной спектральной функцией $F(\lambda)$ и если его матричная спектральная плотность $f(\lambda) = F'(\lambda)$ имеет при почти всех λ ранг r $(1 \leqslant r \leqslant k)$, где k — размерность процесса, то $\{\xi(t), t \in T\}$ называется *процессом ранга* r. Если $r = k$, т. е. $\det f(\lambda) \neq 0$ для почти всех λ, то $\{\xi(t), t \in T\}$ называется *процессом максимального ранга.*

11.3. Спектральное представление стационарных процессов

11.3.1. Спектральное представление. Спектральным представлениям корреляционной функции $B(t)$ вида (2.1)—(2.5) соответствуют спектральные представления самого процесса $\{\xi(t), t \in T\}$.

Теорема 1. *Для всякого стационарного в широком смысле случайного процесса со спектральной функцией $F(\lambda)$ имеет место спектральное представление*

$$\xi(t) = \int_\Lambda \begin{cases} e^{i\lambda t}\, d\zeta(\lambda), \\ [\cos t\lambda\, d\eta(\lambda) + \sin t\lambda\, d\theta(\lambda)] \ \text{в случае действительного процесса}, \end{cases} \tag{3.1}$$

где $\Lambda = [-\pi, \pi]$, если время t дискретно; $\Lambda = (-\infty, \infty)$, если время t непрерывно; интегралы понимаются как среднеквадратичные пределы последовательностей сумм Римана — Стилтьеса; $\zeta(\lambda)$, $\eta(\lambda)$, $\theta(\lambda)$ — процессы с ортогональными приращениями такие, что

$$\mathsf{M}\zeta(\lambda) = \mathsf{M}\eta(\lambda) = \mathsf{M}\theta(\lambda) = 0,$$

$$\mathsf{M}\, d\zeta(\lambda)\, d\zeta^T(\lambda) = dF(\lambda), \quad \mathsf{M}\, d\eta(\lambda)\, d\eta^T(\lambda) = dC(\lambda),$$

$$\mathsf{M}\, d\theta(\lambda)\, d\theta^T(\lambda) = \begin{cases} dQ(\lambda), & \lambda = 0, \\ dC(\lambda), & \lambda \neq 0, \end{cases}$$

$$\mathsf{M}\, d\eta(\lambda)\, d\theta^T(\lambda) = -\mathsf{M}\, d\theta(\lambda)\, d\eta^T(\lambda) = dQ(\lambda).$$

Если потребовать, чтобы процесс $\zeta(\lambda)$ был непрерывен справа в среднеквадратичном, то он определяется единственным образом с точностью до подмножества множества Ω нулевой вероятности.

Процесс $\zeta(\lambda)$ в представлении (3.1) называется *спектральным процессом*, соответствующим стационарному процессу $\{\xi(t), t \in T\}$.

Пусть Γ — произвольное борелевское множество из Λ. Положим $\Phi(\Gamma)=\int\limits_{\Gamma} d\zeta(\lambda)$. Случайная (векторнозначная) функция множества $\Phi(\Gamma)$ обладает следующими свойствами:

1) $\Phi(\Gamma_1\cup\Gamma_2)=\Phi(\Gamma_1)+\Phi(\Gamma_2)$, если $\Gamma_1\cap\Gamma_2=\varnothing$;

2) $M\Phi(\Gamma)\Phi^T(\Gamma)=\int\limits_{\Gamma} dF(\lambda)$;

3) $M\Phi(\Gamma_1)\Phi^T(\Gamma_2)=0$, если $\Gamma_1\cap\Gamma_2=\varnothing$,

4) если $\Gamma=\bigcup\limits_{l=1}^{\infty}\Gamma_l$, $\Gamma_l\cap\Gamma_m=\varnothing$, то $\Phi(\Gamma)=\sum\limits_{l=1}^{\infty}\Phi(\Gamma_l)$ и ряд справа сходится в среднеквадратичном ($\Phi(\cdot)$ называют *спектральной мерой*).

Свойства спектральной меры позволяют дать эквивалентное спектральное представление стационарного процесса

$$\xi(t)=\int\limits_{\Lambda} e^{i\lambda t}\,d\zeta(\lambda)=\int\limits_{\Lambda} e^{i\lambda t}\,\Phi(d\lambda).$$

Если $\{\xi(t),\ t\in T\}$ — скалярный процесс, то элемент $e^{it\lambda}d\zeta(\lambda)$ представляет собой гармоническое колебание с угловой частотой λ, случайной амплитудой и случайной фазой, определяемыми случайной величиной

$$d\zeta(\lambda)=|d\zeta(\lambda)|\exp\{i\arg(d\zeta(\lambda))\}.$$

Спектральное представление показывает, каким образом процесс $\xi(t)$ складывается из элементарных гармонических колебаний.

Спектральные процессы $\zeta(\lambda)$ в представлении (3.1) подчинены процессу $\{\xi(t),\ t\in T\}$ в том смысле, что $\zeta\in H_\xi$, где H_ξ — пространство значений процесса $\{\xi(t),\ t\in T\}$, а ζ есть предел в среднеквадратичном выражений вида $\zeta^N=\sum\limits_{1}^{N}\alpha_k\zeta(\lambda_k)$.

Пусть $\{\xi(t),\ t\in T\}$ — скалярный процесс со спектральной функцией $F(\lambda)$. Обозначим через $\mathscr{L}_2\{F\}$ гильбертово пространство функций $\varphi(\lambda)$, интегрируемых с квадратом по мере, порожденной $F(\lambda)$, скалярное произведение в котором имеет вид

$$(\varphi,\ \psi)=\int\limits_{\Lambda}\varphi(\lambda)\overline{\psi(\lambda)}\,dF(\lambda),$$

а интегрирование ведется по $[-\pi,\ \pi]$ либо по $(-\infty,\ \infty)$. В $\mathscr{L}_2\{F\}$ не различаются функции $\varphi_1(\lambda)$ и $\varphi_2(\lambda)$, для которых

$$\int[\varphi_1(\lambda)-\varphi_2(\lambda)]\overline{[\varphi_1(\lambda)-\varphi_2(\lambda)]}\,dF(\lambda)=0.$$

Непосредственным следствием спектрального представления (3.1) является изометрический изоморфизм H_ξ и $\mathscr{L}_2\{F\}$, так как для всякого $\eta\in H_\xi$ найдется единственная (с точностью до определенной выше эквивалентности) функция $\varphi(\lambda) \in \mathscr{L}_2\{F\}$ такая, что $\eta = \int\limits_\Lambda \varphi(\lambda)\, d\zeta(\lambda)$, и, обратно, если $\varphi(\lambda) \in \mathscr{L}_2\{F\}$, то $\int\limits_\Lambda \varphi(\lambda)\, d\zeta(\lambda) = \eta \in H_\xi$. Соответствие $\eta \leftrightarrow \varphi(\lambda)$ является изометрическим: если $\eta_i \leftrightarrow \varphi_i(\lambda)$ $(i = 1, 2)$, то

$$(\eta_1, \eta_2) = \mathsf{M}\eta_1\bar{\eta}_2 = \int\limits_\Lambda \varphi_1(\lambda)\, \overline{\varphi_2(\lambda)}\, \mathsf{M}\, |d\zeta(\lambda)|^2 = \int\limits_\Lambda \varphi_1(\lambda)\, \overline{\varphi_2(\lambda)}\, dF(\lambda) = (\varphi_1, \varphi_2).$$

Для векторных процессов аналогичным образом определяется гильбертово пространство $\mathscr{L}_2\{F\}$ $m \times k$-матриц $\varphi(\lambda)$ (m произвольно, но фиксировано, k — размерность процесса), в котором скалярное произведение имеет вид

$$(\varphi, \psi) = \int\limits_\Lambda [\operatorname{Sp} \varphi(\lambda)\, dF(\lambda)\, \bar{\psi}^T(\lambda)],$$

и если $m = k$, то H_ξ и $\mathscr{L}_2\{F\}$ изометрично изоморфны.

11.3.2. Процессы с абсолютно непрерывной спектральной функцией. Разложению Лебега (2.6) спектральной функции $F(\lambda)$ соответствует разложение процесса $\{\xi(t),\ t \in T\}$ вида

$$\xi(t) = \xi_1(t) + \xi_2(t) + \xi_3(t) \tag{3.2}$$

на три взаимно ортогональных стационарных процесса.

Процесс $\xi_1(t)$ имеет абсолютно непрерывную спектральную функцию $F_1(\lambda)$. Такие процессы характеризуются следующим образом.

Теорема 2. *Стационарный в широком смысле случайный процесс $\{\xi(t),\ t \in T\}$ имеет абсолютно непрерывную спектральную функцию тогда и только тогда, когда он является процессом скользящего суммирования, т. е. найдутся функции (матрицы) $C(t)$ такие, что:*

а) *в случае дискретного времени*

$$\xi(t) = \sum_{-\infty}^{\infty} C(t - s)\, \zeta_0(s), \tag{3.3}$$

где $\sum\limits_{-\infty}^{\infty} |C(t)|^2 < \infty$ *(скалярный процесс) либо* $\sum\limits_{-\infty}^{\infty} \operatorname{Sp} C(t)\, C^T(t) < \infty$ *(векторный процесс) и $\zeta_0(t)$ — стандартная стационарная последовательность с некоррелированными значениями, при этом*

$$f(\lambda) = F'(\lambda) = \frac{1}{2\pi} C(\lambda)\, C^T(\lambda), \quad C(\lambda) = \sum_{-\infty}^{+\infty} e^{-i\lambda t}\, C(t);$$

б) *в случае непрерывного времени*

$$\xi(t) = \int\limits_{-\infty}^{\infty} C(t - s)\, d\zeta_0(s), \tag{3.4}$$

где $\int\limits_{-\infty}^{\infty} |C(t)|^2\,dt < \infty$ *(скалярный процесс) либо* $\int\limits_{-\infty}^{\infty} \operatorname{Sp} C(t)\, C^T(t) \times$ $\times\, dt < \infty$ *(векторный процесс) и* $\zeta_0(t)$ *— стандартный процесс с ортогональными приращениями, при этом* $f(\lambda) = F'(\lambda) = \frac{1}{2\pi} C(\lambda)\, C^T(\lambda)$, $C(\lambda) = \int\limits_{-\infty}^{+\infty} e^{-i\lambda t}\, C(t)\, dt$.

Для векторных процессов можно указать другую характеризацию, учитывающую тот факт, что спектральная плотность $f(\lambda)$ может иметь различные ранги при различных λ.

Теорема 3. *Векторный стационарный процесс* $\{\xi(t),\ t \in T\}$ *имеет абсолютно непрерывную спектральную функцию тогда и только тогда, когда он может быть представлен в виде суммы не более чем* k *(k — размерность процесса) взаимно ортогональных процессов скользящего суммирования*

$$\xi(t) = \sum_{l=1}^{k} \xi_l(t), \tag{3.5}$$

где в случае дискретного времени

$$\xi_l(t) = \sum_{s \in T} C_l(t-s)\, \zeta_l(s), \qquad \sum_{t \in T} \operatorname{Sp} C_l(t)\, C_l^T(t) < \infty,$$

$C_l(t)$ *—* $k \times l$*-матрицы,* $\zeta_l(s)$ *— стационарные некоррелированные и взаимно некоррелированные* l*-мерные случайные последовательности, а в случае непрерывного времени*

$$\xi_l(t) = \int\limits_{-\infty}^{\infty} C_l(t-s)\, d\zeta_l(s), \qquad \int\limits_{-\infty}^{\infty} \operatorname{Sp} C_l(t)\, C_l^T(t)\, dt < \infty,$$

$C_l(t)$ *—* $k \times l$*-матрицы,* $\zeta_l(s)$ *— стандартные взаимно ортогональные процессы с ортогональными приращениями.*

В частности, если процесс $\xi(t)$ имеет абсолютно непрерывную спектральную плотность и постоянный ранг r, то $\xi(t) = \xi_r(t)$, где $\xi_r(t)$ — один из описанных выше процессов.

11.4. Аналитические свойства стационарных процессов и их траекторий

11.4.1. Среднеквадратичная непрерывность и дифференцируемость стационарных процессов. Пусть $\{\xi(t),\ t \in T\}$ — скалярный процесс с непрерывным временем, $B(t)$ и $F(\lambda)$ — его корреляционная и спектральная функции соответственно. Среднеквадратичная непрерывность и дифференцируемость стационарных процессов, являющихся частным случаем гильбертовых случайных процессов, определяются так же, как для последних (см. п. 10.1.3).

Теорема 1. *Для того чтобы процесс $\{\xi(t),\ t \in T\}$ был среднеквадратично непрерывным, необходимо и достаточно, чтобы его корреляционная функция $B(t)$ была непрерывна в нуле. Для того чтобы процесс $\{\xi(t),\ t \in T\}$ имел среднеквадратичную производную порядка m, необходимо и достаточно, чтобы существовала $2m$-я производная его корреляционной функции $B(t)$ в нуле либо, что эквивалентно, чтобы существовал конечный $2m$-й спектральный момент*

$$S_{2m} = \int_{-\infty}^{\infty} \lambda^{2m}\, dF(\lambda).$$

В частности, если действительный скалярный процесс $\{\xi(t),\ t \in T\}$ имеет конечный второй спектральный момент $S_2 = \int_{-\infty}^{\infty} \lambda^2\, dF(\lambda)$, то процесс $\eta(t) = (\xi'(t),\ \xi(t)),\ t \in T$, где $\xi'(t)$ означает среднеквадратичную производную в t, стационарен в широком смысле, и его матричная корреляционная функция $B_\eta(t)$ имеет вид

$$B_\eta(t) = \begin{bmatrix} \int_{-\infty}^{\infty} e^{it\lambda} \lambda^2\, dF(\lambda) & \int_{-\infty}^{\infty} e^{it\lambda} i\lambda\, dF(\lambda) \\ \int_{-\infty}^{\infty} e^{it\lambda} i\lambda\, dF(\lambda) & \int_{-\infty}^{\infty} e^{it\lambda}\, dF(\lambda). \end{bmatrix}.$$

11.4.2. Аналитические свойства траекторий. Свойства траекторий стационарных процессов с непрерывным временем описываются следующими теоремами.

Теорема 2. *Если*

$$2B(0) - B(t) - B(-t) = O\left(\frac{|t|}{|\ln|t||^q}\right) \tag{4.1}$$

при $t \to 0$ для некоторого $q > 3$, то процесс $\{\xi(t),\ t \in T\}$ с такой корреляционной функцией стохастически эквивалентен процессу, траектории которого с вероятностью 1 непрерывны на любом конечном интервале.

Условие (4.1) выполнено, в частности, если $B(t)$ имеет вторую производную в нуле.

Замечание. Для гауссовских стационарных процессов утверждение теоремы 2 верно, если вместо (4.1) выполнено условие

$$B(t) = B(0) - O\left(\frac{1}{|\ln|t||^q}\right), \quad t \to 0.$$

Теорема 3. *Если*

$$6B(0) - 4B(t) - 4B(-t) + B(2t) + B(-2t) = O\left(\frac{|t|^3}{|\ln|t||^q}\right) \quad (4.2)$$

при $t \to 0$ для некоторого $q > 3$, то процесс $\{\xi(t),\ t \in T\}$ с такой корреляционной функцией стохастически эквивалентен процессу, траектории которого с вероятностью 1 непрерывно дифференцируемы.

Условие (4.2), в частности, выполнено, если $B(t)$ имеет четвертую производную в нуле.

Замечание. Для гауссовских стационарных процессов утверждение теоремы 3 верно, если вместо (4.2) выполнено условие

$$B(t) = B(0) + Ct^2 + O\left(\frac{|t|^2}{|\ln|t||^q}\right),$$

где C — константа.

Аналогичным образом существование производных высших порядков у траекторий стационарных процессов связано с поведением корреляционной функции в нуле.

Теорема 4. *Если спектральная функция $F(\lambda)$ стационарного процесса отлична от нуля лишь на конечном интервале, то существует стохастически эквивалентный данному процесс, траектории которого с вероятностью 1 аналитичны.*

11.5. Эргодическая теорема

Пусть $\{\xi(t),\ t \in T\}$ стационарный в широком смысле процесс, $B(t)$ и $F(\lambda)$ — его корреляционная и спектральная функции соответственно, $\xi(t) = \int\limits_{\Lambda} e^{i\lambda t}\, d\zeta(\lambda)$ — его спектральное представление.

Величины

$$\hat{\xi}_s = \begin{cases} \dfrac{1}{s}\sum\limits_{t=0}^{s-1} \xi(t), & T = \{0, 1, 2, \dots\},\ s \text{ — целые положительные числа;} \\ \dfrac{1}{2s+1}\sum\limits_{t=-s}^{s} \xi(t), & T = \{0, \pm 1, \pm 2, \dots\},\ s \text{ — целые положительные числа;} \\ \dfrac{1}{s}\int\limits_0^s \xi(t)\,dt, & T = [0, \infty),\ s \text{ — положительные числа;} \\ \dfrac{1}{2s}\int\limits_{-s}^{s} \xi(t)\,dt, & T = (-\infty, \infty),\ s \text{ — положительные числа,} \end{cases}$$

называются *временными средними.*

Пусть

$$\hat{B}_s = \begin{cases} \dfrac{1}{s}\sum\limits_{t=0}^{s-1} B(t), & T=\{0,\ 1,\ 2,\ \ldots\}, \\ \dfrac{1}{2s+1}\sum\limits_{t=-s}^{s} B(t), & T=\{0,\ \pm 1,\ \pm 2,\ \ldots\}, \\ \dfrac{1}{s}\int\limits_0^s B(t)\,dt, & T=[0,\ \infty], \\ \dfrac{1}{2s}\int\limits_{-s}^{s} B(t)\,dt, & T=(-\infty,\ \infty). \end{cases}$$

Существование среднеквадратичных пределов временных средних $\hat{\xi}_s$ при $s \to \infty$ составляет содержание *эргодических теорем*, или *закона больших чисел* для стационарных процессов.

Теорема 1.

$$\underset{s\to\infty}{\text{l.i.m.}}\ \hat{\xi}_s = \zeta(0) - \zeta(0-), \qquad \lim_{s\to\infty} \hat{B}_s = F(0) - F(0-).$$

Теорема 2. *Для того чтобы* $\underset{s\to\infty}{\text{l.i.m.}}\ \hat{\xi}_s = \mathsf{M}\xi(t) = 0$, *необходимо и достаточно, чтобы спектральная функция* $F(\lambda)$ *была непрерывна в нуле.*

Достаточным для непрерывности $F(\lambda)$ в нуле является условие $\lim\limits_{t\to\infty} B(t) = 0$. При достаточно быстром стремлении $B(t)$ к нулю при $t \to \infty$ временные средние $\hat{\xi}_s$ могут сходиться к $\mathsf{M}\xi(t) = 0$ не только в среднеквадратичном, а значит, и по вероятности, но и с вероятностью 1.

Теорема 3. *Если существуют константы* $K > 0$ *и* $\alpha > 0$ *такие, что*

$$\frac{1}{s^2}\sum_{t,\ u=0}^{s-1} B(t-u) = \frac{1}{s}\sum_{t=-s+1}^{s-1} B(t)\left[1 - \frac{|t|}{s}\right] \leqslant Ks^{-\alpha}$$

в случае дискретного времени и

$$\frac{1}{s^2}\int\limits_0^s\int\limits_0^s B(t-u)\,du\,dt = \frac{1}{s}\int\limits_{-s}^{s} B(t)\left[1 - \frac{|t|}{s}\right]dt \leqslant Ks^{-\alpha}$$

в случае непрерывного времени, то $\hat{\xi}_s$ *сходится к* $\mathsf{M}\xi(t) = 0$ *с вероятностью* 1.

Замечание. Теорема 3 сформулирована для скалярного процесса. В векторном случае вместо $B(t)$ надо взять $\operatorname{Sp} B(t)$.

Достаточным для выполнения условий теоремы 3 является условие $B(t) \leqslant \gamma|t|^{-\alpha}$, $\gamma > 0$, $\alpha > 0$ — некоторые константы.

11.6. Линейные преобразования (фильтры)

11.6.1. Определение линейного фильтра. Пусть $\{\xi(t),\ t \in T\}$ — стационарный в широком смысле процесс, $B_\xi(t)$ и $F_\xi(\lambda)$ — его (матричные) корреляционная и спектральная функции соответственно. Представим, что процесс $\xi(t)$ (как функция времени t) поступает на вход некоторого физического устройства и преобразуется им так, что на выходе устройства получается новый (преобразованный) процесс $\{\eta(t),\ t \in T\}$.

Преобразование A процесса $\xi(t)$ в процесс $\eta(t) = A\xi(t)$ называется *допустимым линейным фильтром* (или просто *фильтром*), если процесс $\eta(t)$ представим в виде

$$\eta(t) = \begin{cases} \sum\limits_{-\infty}^{\infty} h(t-s)\,\xi(s) & \text{(дискретное время)},\\ \int\limits_{-\infty}^{\infty} h(t-s)\,\xi(s)\,ds & \text{(непрерывное время)}, \end{cases} \tag{6.1}$$

где $h(t)$ — $r \times k$-матрицы (r — размерность процесса $\eta(t)$, k — размерность процесса $\xi(t)$) такие, что

$$\begin{aligned} &\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} h(t_1)\,B(t_2-t_1)\,h^T(t_2) < \infty && \text{(дискретное время)},\\ &\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h(t_1)\,B(t_2-t_1)\,h^T(t_2)\,dt_1\,dt_2 < \infty && \text{(непрерывное время)}, \end{aligned} \tag{6.2}$$

а сумма и интеграл в (6.1) понимаются как среднеквадратичные пределы соответственно сумм $\sum\limits_{-a}^{b} h(t-s)\,\xi(s)$ и интегралов $\int\limits_{-a}^{b} h(t-s)\,\xi(s)\,ds$ при $a,\ b \to \infty$. Функция $h(t)$ называется (матричной) *импульсной переходной функцией* фильтра A.

Замечание. Название это связано с тем, что если на вход фильтра подана импульсная функция (в случае непрерывного времени дельта-функция Дирака с особенностью в нуле), то на выходе фильтра будет $h(t)$.

Пусть

$$H(i\lambda) = \begin{cases} \sum\limits_{-\infty}^{\infty} e^{-it\lambda} h(t) & \text{(дискретное время)},\\ \int\limits_{-\infty}^{\infty} e^{-it\lambda} h(t)\,dt & \text{(непрерывное время)} \end{cases}$$

есть преобразование Фурье импульсной переходной функции $h(t)$. Условие (6.2) эквивалентно условию $H(i\lambda) \in \mathscr{L}_2\{F_\xi\}$. Функция $H(i\lambda)$ называется (матричной) *частотной характеристикой* фильтра A.

Если процесс $\xi(t)$ на входе фильтра A имеет спектральное представление $\xi(t) = \int\limits_{\Lambda} e^{i\lambda t}\, d\zeta(\lambda)$, то процесс $\eta(t)$ на выходе фильтра имеет спектральное представление $\eta(t) = \int\limits_{\Lambda} e^{i\lambda t} H(i\lambda)\, d\zeta(\lambda)$.

В частности, если $\xi(t)$ и $\eta(t)$ — скалярные процессы, то $H(i\lambda) = |H(i\lambda)|\, e^{i\psi(\lambda)}$, и $|H(i\lambda)|$ называется *коэффициентом усиления фильтра* (на частоте λ), а $\psi(\lambda)$ — *фазой фильтра.*

Процесс $\eta(t) = A\xi(t)$ является стационарным, причем

$$B_\eta(t) = \begin{cases} \sum\limits_{-\infty}^{\infty}\sum\limits_{-\infty}^{\infty} h(t+n)\, B_\xi(m-n)\, h^T(m) = \\ = \int\limits_{-\pi}^{\pi} e^{it\lambda} H(i\lambda)\, dF_\xi(\lambda)\, H^T(i\lambda) \quad \text{(дискретное время)}, \\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h(t+u)\, B_\xi(v-u)\, h^T(v)\, du\, dv = \\ = \int\limits_{-\infty}^{\infty} e^{it\lambda} H(i\lambda)\, dF_\xi(\lambda)\, H^T(i\lambda) \quad \text{(непрерывное время)}, \end{cases} \tag{6.3}$$

$$dF_\eta(\lambda) = H(i\lambda)\, dF_\xi(\lambda)\, H^T(i\lambda). \tag{6.4}$$

Если существуют спектральные плотности $f_\xi(\lambda)$ и $f_\eta(\lambda)$, то

$$f_\eta(\lambda) = H(i\lambda)\, f_\xi(\lambda)\, H^T(i\lambda). \tag{6.5}$$

Пусть $\{\xi(t),\ t \in T\}$ и $\{\eta(t),\ t \in T\}$ — два произвольных стационарных в широком смысле процесса со спектральными функциями $F_\xi(\lambda)$ и $F_\eta(\lambda)$. Ответ на вопрос, является ли процесс $\eta(t)$ линейным преобразованием процесса $\xi(t)$, дает следующая теорема.

Теорема 1. *Пусть процессы $\{\xi(t),\ t \in T\}$ и $\{\eta(t),\ t \in T\}$ совместно стационарны, т. е. (объединенный) векторный процесс $\{(\xi(t), \eta(t)),\ t \in T\}$ является стационарным в широком смысле. Для того чтобы процесс $\{\eta(t),\ t \in T\}$ мог быть получен из процесса $\{\xi(t),\ t \in T\}$ с помощью фильтра с частотной характеристикой $H(i\lambda)$, необходимо и достаточно, чтобы соответствующие спектральные и взаимные спектральные функции были связаны соотношениями*

$$dF_\eta(x) = H(i\lambda)\, dF_\xi(\lambda)\, H^T(i\lambda), \quad dF_{\xi\eta}(\lambda) = dF_\xi(\lambda)\, H^T(i\lambda). \tag{6.6}$$

11.6.2. Примеры фильтров. 1. *Полосовой фильтр* пропускает, не изменяя, лишь гармонические составляющие скалярного процесса $\xi(t)$ с частотами в заданном интервале (a, b). Его частотная характеристика:

$$H(i\lambda) = \chi_{(a,\, b)}(\lambda) = \begin{cases} 1, & \lambda \in (a, b), \\ 0, & \lambda \notin (a, b). \end{cases}$$

Импульсная переходная функция $h(t)$ (для конечных a и b):

$$h(t)=\frac{e^{ibt}-e^{iat}}{2\pi it}.$$

В соответствии с расположением интервала (a, b) полосовой фильтр называют *низкочастотным, среднечастотным* или *высокочастотным.*

Если $a=-\infty$ или (и) $b=\infty$, то импульсной переходной функции не существует.

2. *Дифференцирование.* Операцию $A=\sum_{l=0}^{m} B_l \frac{d^{m-l}}{dt^{m-l}}$ к процессу $\{\xi(t),\ t\in T\}$ с непрерывным временем можно применить тогда и только тогда, когда конечен $2m$-й спектральный момент

$$S_{2m}=\int_{-\infty}^{\infty} \lambda^{2m}\, dF_{\xi}(\lambda).$$

Частотная характеристика $H(i\lambda)=\sum_{l=0}^{m}(i\lambda)^{m-l}B_l$. В частности, если $A=d/dt$, то $H(i\lambda)=i\lambda$. Импульсной переходной функции не существует.

3. *Дифференциальные уравнения.* Рассмотрим фильтр, определяемый линейным дифференциальным уравнением с постоянными коэффициентами $L\eta(t)=M\xi(t)$, где L и M — дифференциальные операторы:

$$L=\sum_{j=0}^{l} C_j \frac{d^{n-j}}{dt^{n-j}},\quad M=\sum_{j=0}^{m} B_j \frac{d^{m-j}}{dt^{m-j}}.$$

Предполагается, что существует $2m$-й спектральный момент процесса $\xi(t)$.

Если $M(i\lambda)/L(i\lambda)\in\mathscr{L}_2\{F_\xi\}$, где $L(i\lambda)=\sum_{j=0}^{l}(i\lambda)^{l-j}C_j$ и $M(i\lambda)=\sum_{j=0}^{m}(i\lambda)^{m-j}B_j$, то существует фильтр, соответствующий рассматриваемому дифференциальному уравнению, с частотной характеристикой $H(i\lambda)=M(i\lambda)/L(i\lambda)$.

11.6.3. Физически осуществимые фильтры. В фильтрах, определяемых равенством (6.1), значения процесса $\{\eta(t),\ t\in T\}$ на выходе в момент t могут зависеть как от прошлых моментов времени $(s<t)$, так и от будущих $(s>t)$.

Реальные физические устройства лишены возможности предвосхищать будущее. Поэтому, если фильтр моделирует реальный объект, его импульсная переходная функция $h(t)$ должна удовлетворять *условию физической осуществимости*:

$$h(t)=0,\quad t<0. \tag{6.7}$$

Фильтры, удовлетворяющие условию (6.7), называются *физически осуществимыми.*

Теорема 2. *Для того чтобы скалярный процесс* $\{\xi(t),\ t \in T\}$ *со спектральной функцией* $F(\lambda)$ *был реакцией физически осуществимого фильтра, на вход которого поступает стандартная некоррелированная последовательность* $\{\zeta_0(t),\ t \in T\}$ *(в случае дискретного времени) либо стандартный процесс с ортогональными приращениями* $\{\zeta_0(t),\ t \in T\}$ *(в случае непрерывного времени), необходимо и достаточно, чтобы его спектральная функция* $F(\lambda)$ *была абсолютно непрерывной, а спектральная плотность* $f(\lambda)$ *удовлетворяла условию*

$$\begin{gathered}\int_{-\pi}^{\pi} \ln f(\lambda)\, d\lambda > -\infty \quad (\textit{дискретное время}),\\ \int_{-\infty}^{\infty} \frac{\ln f(\lambda)\, d\lambda}{1+\lambda^2} > -\infty \quad (\textit{непрерывное время}).\end{gathered} \tag{6.8}$$

При выполнении этих условий процесс $\{\xi(t),\ t \in T\}$ *на выходе физически осуществимого фильтра можно представить в виде*

$$\begin{gathered}\xi(t) = \sum_{-\infty}^{t} h(t-s)\,\zeta_0(s), \quad \sum_{t=0}^{\infty} |h(t)|^2 < \infty \quad (\textit{дискретное время});\\ \xi(t) = \int_{-\infty}^{t} h(t-s)\, d\zeta_0(s), \quad \int_0^{\infty} |h(t)|^2\, dt < \infty \quad (\textit{непрерывное время}).\end{gathered} \tag{6.9}$$

Замечание. Второе из равенств (6.9) записывают иногда в виде

$$\xi(t) = \int_{-\infty}^{t} h(t-s)\,\varepsilon(t)\, dt, \tag{6.10}$$

где $\varepsilon(t)$ — *процесс белого шума*, представляющий собой обобщенный (см. § 12.5) стационарный в широком смысле процесс; $\mathsf{M}\varepsilon(t) = 0$, $\mathsf{M}\varepsilon(t)\varepsilon(s) = \delta(t-s)$, где $\delta(t)$ — дельта-функция Дирака. Это позволяет интерпретировать $\xi(t)$ как реакцию физически осуществимого фильтра на белый шум.

11.6.4. Факторизация спектральной плотности. В случае дискретного времени спектральная плотность $f(\lambda)$, удовлетворяющая первому из условий (6.8), допускает факторизацию

$$f(\lambda) = \frac{1}{2\pi} |\varphi(e^{-i\lambda})|^2, \tag{6.11}$$

где $\varphi(e^{-i\lambda})$ является граничным значением аналитической функции

$$\varphi(z) = \sqrt{2\pi} \exp\left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln f(\lambda) \frac{e^{-i\lambda}+z}{e^{-i\lambda}-z}\, d\lambda \right\},$$

т. е.

$$\varphi\left(e^{-i\lambda}\right)=\lim_{\rho\uparrow 1}\varphi\left(\rho e^{-i\lambda}\right),$$

при этом

$$h(t)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i\lambda t}\varphi\left(e^{-i\lambda}\right)d\lambda,\quad \varphi\left(e^{-i\lambda}\right)=\sum_{-\infty}^{+\infty}h(t)\,e^{-i\lambda t}.$$

В случае непрерывного времени спектральная плотность $f(\lambda)$, удовлетворяющая второму из условий (6.8), допускает факторизацию

$$f(\lambda)=\frac{1}{2\pi}\left|\varphi(i\lambda)\right|^2, \tag{6.12}$$

где $\varphi(i\lambda)$ является граничным значением аналитической в правой полуплоскости функции

$$\varphi(z)=\sqrt{\pi}\exp\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\ln f(\lambda)}{1+\lambda^2}\,\frac{i+\lambda z}{\lambda+iz}\,d\lambda\right\},$$

при этом

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\lambda t}\varphi(i\lambda)\,d\lambda.$$

Векторный аналог теоремы 2 имеет наиболее простой вид в том случае, когда процесс $\{\xi(t),\ t\in T\}$ на выходе фильтра имеет максимальный ранг, т. е. спектральная плотность $f(\lambda)$ имеет почти всюду по мере Лебега отличный от нуля детерминант.

Теорема 3. *Для того чтобы векторный процесс $\{\xi(t),\ t\in T\}$ максимального ранга со спектральной функцией $F(\lambda)$ был реакцией физически осуществимого фильтра, на вход которого поступает стандартная последовательность некоррелированных случайных векторов $\{\zeta_0(t),\ t\in T\}$ (в случае дискретного времени) либо стандартный k-мерный процесс с ортогональными приращениями $\{\zeta_0(t),\ t\in T\}$ (в случае непрерывного времени), необходимо и достаточно, чтобы спектральная функция $F(\lambda)$ была абсолютно непрерывной, а спектральная плотность $f(\lambda)$ удовлетворяла условию*

$$\begin{aligned}&\int_{-\pi}^{\pi}\ln\det[f(\lambda)]\,d\lambda>-\infty\quad(\textit{дискретное время}),\\&\int_{-\infty}^{\infty}\frac{\ln\det[f(\lambda)]}{1+\lambda^2}\,d\lambda>-\infty\quad(\textit{непрерывное время}).\end{aligned}\tag{6.13}$$

При выполнении этих условий процесс $\{\eta(t),\ t\in T\}$ *на выходе физически осуществимого фильтра можно представить в виде*

$$\eta(t)=\sum_{-\infty}^{t} h(t-s)\,\zeta_0(s),\qquad \sum_{t=0}^{\infty}\operatorname{Sp} h(t)\,h^T(t)<\infty\ (\textit{дискретное время}),$$

$$\eta(t)=\int_{-\infty}^{t} h(t-s)\,\zeta_0(ds),\qquad \int_{0}^{\infty}\operatorname{Sp} h(t)\,h^T(t)\,dt<\infty\ (\textit{непрерывное время}). \tag{6.14}$$

В случае дискретного времени спектральная плотность $f(\lambda)$, удовлетворяющая первому из условий (6.13), допускает факторизацию

$$f_\eta(\lambda)=\frac{1}{2\pi}\,\varphi\left(e^{-i\lambda}\right)\varphi^T\left(e^{-i\lambda}\right), \tag{6.15}$$

где матрица $\varphi(e^{-i\lambda})$ размерности $k\times k$ является граничным значением аналитической внутри единичного круга матрицы $\varphi(z)$, однозначно определяемой условиями

$$\lim_{\rho\uparrow 1}\varphi\left(\rho e^{-i\lambda}\right)\varphi^T\left(\rho e^{-i\lambda}\right)=2\pi f_\eta(\lambda),$$

$$|\det\varphi(0)|^2=(2\pi)^k\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\det\left[f_\eta(\lambda)\right]d\lambda\right\}.$$

11.7. Процессы с дробно-рациональными спектральными плотностями

11.7.1. Теоремы о факторизации. Спектральная плотность $f(\lambda)$ называется *дробно-рациональной*, если $f(\lambda)$ либо ее элементы $f_{ml}(\lambda)$ (в том случае, когда $f(\lambda)$ — матрица) допускают представление в виде $P\left(e^{-i\lambda}\right)/Q\left(e^{-i\lambda}\right)$ (дискретное время) либо $P(i\lambda)/Q(i\lambda)$ (непрерывное время), где $P(z)$ и $Q(z)$ — некоторые многочлены.

Процессы с дробно-рациональными спектральными плотностями представимы в виде процессов скользящего суммирования, а векторные процессы, кроме того, имеют постоянный ранг.

Основной результат для этого класса стационарных процессов содержится в теоремах о факторизации.

Теорема 1. *Если $f(\lambda)$ — дробно-рациональная спектральная плотность некоторого стационарного процесса с дискретным временем, то она допускает факторизацию вида*

$$f(\lambda)=\begin{cases}\dfrac{1}{2\pi}\left|\dfrac{P\left(e^{-i\lambda}\right)}{Q\left(e^{-i\lambda}\right)}\right|^2 & (\textit{скалярный процесс}),\\[2ex] \dfrac{1}{2\pi}B\left(e^{-i\lambda}\right)B^T\left(e^{-i\lambda}\right) & (\textit{векторный процесс}),\end{cases} \tag{7.1}$$

где многочлены $P(z)=\sum_{l=0}^{p} p_l z^l$ *и* $Q(z)=\sum_{l=0}^{q} q_l z^l$ *не имеют нулей внутри единичного круга, причем, если* $f(\lambda)=f(-\lambda)$, *то коэффициенты* p_l, $l=1, \ldots, p$, *и* q_l, $l=1, \ldots, q$, *могут быть выбраны действительными;* $B(z)$ — $k\times r$*-матрица* (k — *размерность процесса, а* r — *его ранг*), *элементы которой дробно-рациональны относительно* z, *причем матрица* $B(z)$ *аналитична в единичном круге.*

Теорема 2. *Если* $f(\lambda)$ — *дробно-рациональная спектральная плотность некоторого стационарного процесса с непрерывным временем, то она допускает факторизацию вида*

$$f(\lambda)=\begin{cases} \dfrac{1}{2\pi}\left|\dfrac{P(i\lambda)}{Q(i\lambda)}\right|^2 & (\textit{скалярный процесс}),\\ \dfrac{1}{2\pi}B(i\lambda)\,B^T(i\lambda) & (\textit{векторный процесс}),\end{cases} \tag{7.2}$$

где полиномы $P(z)=\sum_{l=0}^{p} p_l z^l$ *и* $Q(z)=\sum_{l=0}^{q} q_l z^l$ *не имеют нулей в нижней полуплоскости, а если к тому же* $f(\lambda)=f(-\lambda)$, *то полиномы* $P(iz)$ *и* $Q(iz)$ *имеют действительные коэффициенты;* $B(z)$ — $k\times r$*-мерная матрица* (k — *размерность процесса,* r — *его ранг*), *элементы которой дробно-рациональны относительно* z *и матрица* $B(z)$ *аналитична в нижней полуплоскости.*

Теоремы о факторизации дают следующие спектральные представления процессов с дробно-рациональными спектральными плотностями:

$$\begin{aligned} \xi(t) &= \int_{-\pi}^{\pi} e^{i\lambda t}\frac{P(e^{-i\lambda})}{Q(e^{-i\lambda})}\,d\zeta_0(\lambda) && \text{(скалярный процесс с дискретным временем)},\\ \xi(t) &= \int_{-\infty}^{\infty} e^{i\lambda t}\frac{P(i\lambda)}{Q(i\lambda)}\,d\zeta_0(\lambda) && \text{(скалярный процесс с непрерывным временем)}, \end{aligned} \tag{7.3}$$

где $\zeta_0(\lambda)$ — стандартный процесс с ортогональными приращениями;

$$\begin{aligned} \xi(t) &= \int_{-\pi}^{\pi} e^{i\lambda t}B(e^{-i\lambda})\,d\zeta_0(\lambda) && \text{(векторный процесс с дискретным временем)},\\ \xi(t) &= \int_{-\infty}^{\infty} e^{i\lambda t}B(i\lambda)\,d\zeta_0(\lambda) && \text{(векторный процесс с непрерывным временем)}, \end{aligned} \tag{7.4}$$

где $\zeta_0(\lambda)$ — r-мерный стандартный процесс с ортогональными приращениями, r — ранг процесса $\{\xi(t),\ t\in T\}$.

11.7.2. Примеры. Следующие ниже примеры скалярных стационарных процессов с дискретным временем $T=\{0, \pm 1, \pm 2, \ldots\}$ полностью описывают класс процессов с дробно-рациональными спектральными плотностями.

1. *Процессы скользящего среднего.* Пусть $\{\zeta_0(t),\ t \in T\}$ — стандартная некоррелированная последовательность и a_l, $l = 0, \ldots, p$, — произвольный набор действительных величин. Процесс $\{\xi(t),\ t \in T\}$, где $\xi(t) = a_0\zeta_0(t) + a_1\zeta_0(t-1) + \ldots + a_p\zeta_0(t-p)$, называется процессом *скользящего среднего порядка* p.

Спектральная плотность:

$$f(\lambda) = \frac{1}{2\pi} | a_0 + a_1 e^{-i\lambda} + \ldots + a_p e^{-ip\lambda} |^2.$$

Спектральное представление:

$$\xi(t) = \int_{-\pi}^{\pi} e^{i\lambda t} \left[a_0 + a_1 e^{-i\lambda} + \ldots + a_p e^{-ip\lambda}\right] d\zeta_0(\lambda).$$

В частности, если $a_l = \sigma/(p+1)$ $(l = 0, \ldots, p)$, то

$$f(\lambda) = \frac{\sigma^2}{2\pi (p+1)^2} \cdot \frac{\sin^2((p+1)\lambda/2)}{\sin^2(\lambda/2)},$$

$$\xi(t) = \frac{\sigma}{(p+1)} \int_{-\pi}^{\pi} \frac{1 - e^{-i(p+1)\lambda}}{1 - e^{-i\lambda}} e^{i\lambda t} d\zeta_0(\lambda).$$

2. *Процессы авторегрессии.* Пусть $\{\zeta_0(t),\ t \in T\}$ — стандартная некоррелированная последовательность. Рассмотрим уравнение в конечных разностях для определения процесса $\{\xi(t),\ t \in T\}$:

$$\xi(t) + b_1 \xi(t-1) + \ldots + b_q \xi(t-q) = \sigma^2 \zeta_0(t). \tag{7.5}$$

Решение этого уравнения, если оно существует как стационарный в широком смысле процесс, называется процессом *авторегрессии порядка* q. Стационарное решение уравнения (7.5) существует, если нули полинома $Q(z) = 1 + b_1 z + \ldots + b_q z^q$ лежат вне единичного круга.

Спектральная плотность:

$$f(\lambda) = \frac{1}{2\pi | 1 + b_1 e^{-i\lambda} + \ldots + b_q e^{-iq\lambda} |^2}.$$

Спектральное представление:

$$\xi(t) = \int_{-\pi}^{\pi} e^{i\lambda t} \frac{d\zeta_0(\lambda)}{1 + b_1 e^{-i\lambda} + \ldots + b_q e^{-iq\lambda}},$$

где $\zeta_0(\lambda)$ — стандартный процесс с ортогональными приращениями.

З а м е ч а н и е. Процесс авторегрессии первого порядка является марковским в широком смысле.

3. *Смешанная модель авторегрессии и скользящего среднего.* Комбинирование моделей авторегрессии и скользящего среднего приводит к уравнению

$$\xi(t) + b_1 \xi(t-1) + b_2 \xi(t-2) + \ldots + b_q \xi(t-q) =$$
$$= a_0 \zeta_0(t) + a_1 \zeta_0(t-1) + \ldots + a_p \zeta_0(t-p). \tag{7.6}$$

Если нули полинома $Q(z) = 1 + b_1 z + \ldots + b_q z^q$ лежат вне единичного круга, то уравнение (7.6) имеет стационарное решение $\{\xi(t),\ t \in T\}$, называемое *смешанным процессом авторегрессии и скользящего среднего порядка* $(q,\ p)$.

Спектральная плотность:

$$f(\lambda) = \left| \frac{a_0 + a_1 e^{-i\lambda} + \ldots + a_p e^{-ip\lambda}}{1 + b_1 e^{-i\lambda} + \ldots + b_q e^{-iq\lambda}} \right|^2.$$

Спектральное представление:

$$\xi(t) = \int_{-\pi}^{\pi} e^{i\lambda t} \frac{a_0 + a_1 e^{-i\lambda} + \ldots + a_p e^{-ip\lambda}}{1 + b_1 e^{-i\lambda} + \ldots + b_q e^{-iq\lambda}} \, d\zeta_0(\lambda).$$

Примечателен тот факт, что адекватное описание наблюдаемых на практике явлений, моделируемых стационарными процессами, достигается (смешанными) моделями авторегрессии и скользящего среднего, порядок которых не превышает 2.

11.8. Прогнозирование, интерполирование и фильтрация стационарных процессов

11.8.1. Общие задачи прогноза, интерполирования и фильтрации. 1) *Прогнозирование (экстраполяция)*. Предположим, что стационарный в широком смысле процесс $\{\xi(t),\ t \in T\}$ наблюдается в моменты времени $t \in T_0$, где $T_0 = \{t \in T\colon t \leqslant t_0\}$ либо $T_0 = \{t \in T\colon t_0 - h \leqslant t \leqslant t_0\}$, $h > 0$. Требуется на основе этих наблюдений дать наилучший среднеквадратичный прогноз этого процесса в некоторый будущий момент $t^* = t_0 + \tau$ $(\tau > 0)$, т. е. требуется найти такой функционал $\eta(t^*) = g_{t^*}(\xi(t),\ t \in T_0)$ от значений процесса $\xi(t)$ в моменты $t \in T_0$, чтобы

$$\mathsf{M}\,\|\xi(t^*) - \eta(t^*)\|^2 \leqslant \mathsf{M}\,\|\xi(t^*) - \eta_1(t^*)\|^2, \tag{8.1}$$

где $\eta_1(t^*)$ — любой другой функционал от значений процесса $\xi(t)$ в моменты $t \in T_0$.

2) *Интерполирование*. Пусть процесс $\{\xi(t),\ t \in T\}$ наблюдается в моменты $t \in T_0 \subset T$, и пусть $t^* \in T$ — такой момент времени, что $t^* \notin T_0$ и найдутся $t_i \in T_0$ $(i = 1,\ 2)$, для которых $t_1 < t^* < t_2$. Требуется на основе этих наблюдений процесса $\{\xi(t),\ t \in T\}$ наилучшим в смысле среднеквадратичного критерия образом проинтерполировать значение процесса $\{\xi(t),\ t \in t\}$ в момент t^*, т. е. найти функционал $\eta(t^*) = g_{t^*}(\xi(t),\ t \in T_0)$ от значений процесса $\xi(t)$ в моменты $t \in T_0$, для которых имеет место соотношение (8.1).

3) *Фильтрация*. Пусть в моменты $t \in T_0 \subset T$ наблюдается процесс $\xi(t) = s(t) + \theta(t)$, представляющий собой сумму *полезного сигнала* $s(t)$ и *шума* $\theta(t)$, где $\{s(t),\ t \in T\}$ и $\{\theta(t),\ t \in T\}$ — стационарные и некоррелированные процессы.

Требуется отделить (отфильтровать) шум $\theta(t)$ от сигнала $s(t)$, т. е. найти функционал $\eta(t^*) = g_{t^*}(\xi(t),\ t \in T_0)$ от значений процесса $\xi(t)$ в моменты $t \in T_0$ такой, что

$$\mathsf{M}\,\|s(t^*) - \eta(t^*)\|^2 \leqslant \mathsf{M}\,\|s(t^*) - \eta_1(t^*)\|^2, \tag{8.2}$$

где $\eta_1(t^*)$ — любой другой функционал от значений наблюдаемого процесса $\xi(t)$ в моменты $t \in T_0$.

Левые части (8.1) и (8.2) называют соответственно *ошибкой прогноза и интерполирования* и *ошибкой фильтрации*.

Общее решение всех сформулированных задач дается следующей теоремой (верной, впрочем, для любых гильбертовых случайных процессов).

Теорема. *Пусть $\mathfrak{F}_{T_0}$ — σ-алгебра, порожденная значениями процесса $\{\xi(t),\ t \in T\}$ в моменты $t \in T_0$. Наилучший в среднеквадратичном смысле функционал, решающий задачу прогноза, интерполяции или фильтрации, имеет вид*

$$\eta(t^*) = \mathsf{M}\{\xi(t^*) \mid \mathfrak{F}_{T_0}\}. \tag{8.3}$$

К сожалению, практическая ценность этой теоремы невелика, так как эффективное вычисление правой части (8.3) является чрезвычайно трудной задачей.

11.8.2. Задачи линейного прогнозирования, интерполирования и фильтрации рассматриваются в более простой постановке: функционал $\eta(t^*)$ ищется в классе линейных функционалов от значений процесса $\{\xi(t),\ t \in T\}$ в моменты $t \in T_0$, т. е.

$$\eta(t^*) = \sum_{s \in T_0} C(s)\, \xi(s) \quad \text{(дискретное время)} \tag{8.4}$$

либо

$$\eta(t^*) = \int_{T_0} C(s)\, \xi(s)\, ds \quad \text{(непрерывное время)}. \tag{8.5}$$

В случае непрерывного времени даже для сравнительно простых классов процессов функция $C(s)$ в (8.5) оказывается обобщенной. Рассмотрение линейных функционалов вида

$$\eta(t^*) = \int_{-\infty}^{\infty} e^{i\lambda t^*} C(i\lambda)\, d\zeta(\lambda), \tag{8.6}$$

где $\zeta(\lambda)$ — спектральный процесс, соответствующий процессу $\{\xi(t),\ t \in T\}$, т. е. $\int_{-\infty}^{\infty} e^{i\lambda t}\, d\zeta(\lambda) = \xi(t)$, позволяет проводить анализ, не оперируя непосредственно обобщенными функциями.

Задачи линейного прогнозирования, линейного интерполирования и линейной фильтрации допускают простую геометрическую интерпретацию.

Пусть H_ξ — пространство значений процесса $\{\xi(t),\ t \in T\}$ и $H_\xi(T_0)$ — замкнутое подпространство в H_ξ, являющееся замыканием в H_ξ линейной оболочки случайных величин (векторов) $\xi(t_j)$ $(t_j \in T_0,\ j = 1, \ldots, N)$. Задачи линейного прогнозирования и интерполирования состоят в нахождении величин $\hat{\eta}(t^*)$, являющихся проекциями неизвестных значений $\xi(t^*)$ как элементов из H_ξ на подпространство $H_\xi(T_0)$; задача линейной фильтрации состоит в нахождении величин $\hat{\eta}(t^*)$, являющихся проекциями неизвестных значений $s(t^*)$ из подпространства $H_s \subset H_\xi$ на подпространство $H_\xi(T_0)$.

Пусть стационарный процесс $\{\xi(t),\ t \in T\}$ с корреляционной функцией $B_\xi(t)$ наблюдается в момент $t \in T_0 \subset T$, и пусть $\{\eta(t),\ t \in T\}$ — стационарный и стационарно связанный с $\{\xi(t),\ t \in T\}$ процесс, а $B_{\eta\xi}(t)$ — их взаимная корреляционная функция.

Если предположить, что оценка $\hat\eta(t^*)$ значения ненаблюдаемого процесса $\{\eta(t),\ t \in T\}$ в момент $t^* \in T$ имеет вид

$$\hat\eta(t^*) = \begin{cases} \sum\limits_{s \in T_0} C_{t^*}(s)\,\xi(s) & \text{(дискретное время)},\\ \int\limits_{T_0} C_{t^*}(s)\,\xi(s)\,ds & \text{(непрерывное время)}, \end{cases}$$

то функция $C_{t^*}(t)$ $(t \in T_0)$, называемая *импульсной переходной функцией оптимального линейного фильтра*, может быть найдена как решение линейного (интегрального) уравнения Фредгольма первого рода с эрмитовым ядром:

$$\begin{aligned} &\sum_{s \in T_0} C_{t^*}(s)\,B_\xi(s-t) = B_{\eta\xi}(t^*-t), \quad t \in T_0 \quad \text{(дискретное время)},\\ &\int_{T_0} C_{t^*}(s)\,B_\xi(s-t)\,ds = B_{\eta\xi}(t^*-t), \quad t \in T_0 \quad \text{(непрерывное время)}. \end{aligned} \tag{8.7}$$

Так, например, если $T_0 = \{t \in (-\infty,\ \infty)\colon t \leqslant t_0\}$, $t^* = t_0 + \tau$, то второе из уравнений (8.7) принимает вид

$$\int_{-\infty}^{t_0} C_{t^*}(s)\,B_\xi(s-u)\,ds = B_{\eta\xi}(t^*-u), \quad u \leqslant t_0;$$

после замены $t_0 - u = v$, $t_0 - s = z$ последнее уравнение переходит в

$$\int_0^\infty C_{t^*}(t_0 - z)\,B_\xi(v-z)\,dz = B_{\eta\xi}(\tau + v), \quad v \geqslant 0. \tag{8.8}$$

Если решение интегрального уравнения (8.8) существует и единственно, то $C_\tau(z) = C_{t^*}(t_0 - z)$ не зависит от t_0, откуда

$$\int_0^\infty C_\tau(s)\,B_\xi(t-s)\,ds = B_{\eta\xi}(\tau + t), \quad t \geqslant 0, \tag{8.9}$$

и прогнозирующий процесс имеет вид

$$\hat\eta_\tau(t) = \int_{-\infty}^{t} C_\tau(s)\,\xi(s)\,ds = \int_0^\infty C_\tau(s)\,\xi(t-s)\,ds$$

с ошибкой прогноза

$$\sigma_\tau^2 = B_\xi(0) - \int_0^\infty \int_0^\infty C_\tau(s)\, B_\xi(t-s)\, \overline{C_\tau(t)}\, ds\, dt =$$

$$= B_\eta(0) - \int_{-\infty}^{\infty} |C_\tau(i\lambda)|^2\, dF_\xi(\lambda),$$

где $F_\xi(\lambda)$ — спектральная функция процесса $\xi(t)$, $C_\tau(i\lambda) = \int_0^\infty C_\tau(t) \times$ $\times e^{-i\lambda t}\, dt$ — *частотная характеристика оптимального фильтра.*

11.8.3. Метод Винера. При некоторых дополнительных предположениях интегральное уравнение (8.9) может быть решено методом, предположенным Н. Винером. Именно, пусть процесс $\{\xi(t),\ t \in T\}$, $T = [0, \infty)$, абсолютно непрерывен и его спектральная плотность $f_\xi(\lambda)$ допускает факторизацию вида

$$f_\xi(\lambda) = \frac{1}{2\pi} |\varphi(i\lambda)|^2, \quad \varphi(z) = \frac{1}{2\pi} \int_0^\infty e^{-zt} h(t)\, dt, \quad \operatorname{Re} z \geqslant 0,$$

и пусть $f_{\eta\xi}(\lambda)$ — взаимная спектральная плотность процессов $\{\xi(t),$ $t \in T\}$ и $\{\eta(t),\ t \in T\}$, причем функция $k(i\lambda) = f_{\eta\xi}(\lambda)/\overline{\varphi(i\lambda)}$ интегрируема с квадратом, так что

$$B_{\eta\xi}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f_{\lambda\xi}(\lambda)\, d\lambda = \int_{-\infty}^{\infty} e^{i\lambda t} k(i\lambda)\, \overline{\varphi(i\lambda)}\, d\lambda =$$

$$= \int_0^\infty b(t+s)\, \overline{h(s)}\, ds,$$

где

$$b(t) = \int_{-\infty}^{\infty} e^{i\lambda t} k(i\lambda)\, d\lambda.$$

Уравнение (8.9) может быть записано в виде

$$\int_0^\infty \left[b(\tau + t + s) - \int_0^\infty C_\tau(u)\, h(t + s - u)\, du \right] \overline{h(s)}\, ds = 0, \quad t > 0.$$

Последнее равенство выполняется, если

$$b(\tau + t) = \int_0^\infty C_\tau(u)\, h(t-u)\, du, \quad t > 0, \tag{8.10}$$

либо

$$b(\tau+t)=\int_0^t C_\tau(u)\,h(t-u)\,du,\quad t>0.$$

Уравнение (8.10) решается с помощью преобразования Лапласа:

$$C_\tau(t)=\int_{-\infty}^{\infty} e^{i\lambda t}\,\frac{B_\tau(i\lambda)}{\varphi(i\lambda)}\,d\lambda, \tag{8.11}$$

где

$$B_\tau(z)=\int_0^\infty b(\tau+x)\,e^{-zx}\,dx=\int_{-\infty}^{\infty} e^{i\tau\lambda}\,\frac{f_{\eta\xi}(\lambda)}{\overline{\varphi(i\lambda)}\,(z-i\lambda)}\,d\lambda.$$

11.8.4. Метод Яглома. Как отмечено в п. 11.8.2, импульсная переходная функция $C_\tau(t)$ оптимального фильтра может не существовать (точнее, может существовать лишь как обобщенная функция). В таких случаях оказывается естественным обращение к частотной характеристике $C_\tau(i\lambda)$ соответствующего оптимального фильтра.

Так, если $T_0=\{t\in(-\infty,\ \infty):\ t\leqslant t_0\}$, $t^*=t_0+\tau$, $\xi(t)=\int_{-\infty}^{\infty} e^{i\lambda t}\,d\zeta_\xi(\lambda)$, $\eta(t)=\int_{-\infty}^{\infty} e^{i\lambda t}\,d\zeta_\eta(\lambda)$ — стационарно связанные процессы, имеющие спектральные плотности $f_\xi(\lambda)$ и $f_\eta(\lambda)$ соответственно, и $f_{\eta\xi}(\lambda)$ — их взаимная спектральная плотность, причем процесс $\xi(t)$ наблюдается на T_0, то ищется *частотная характеристика* $C_\tau(i\lambda)$ оптимального линейного фильтра, т. е. такая функция, что для $t^*=t_0+\tau$

$$\hat\eta(t^*)=\int_{-\infty}^{\infty} e^{i\lambda t^*}C_\tau(i\lambda)\,d\zeta_\xi(\lambda),\qquad \int_{-\infty}^{\infty}|C_\tau(i\lambda)|^2 f_\xi(\lambda)\,d\lambda<\infty.$$

Метод Яглома дает процедуру, позволяющую найти частотную характеристику как функцию, однозначно определяемую некоторыми условиями.

Теорема Яглома. *Пусть спектральная плотность $f_\xi(\lambda)$ ограничена и выполнены условия:*

а) $\int_{-\infty}^{\infty}|C_\tau(i\lambda)|^2 f_\xi(\lambda)\,d\lambda<\infty$;

б) $C_\tau(i\lambda)$ *является граничным значением аналитической в правой полуплоскости функции $C_\tau(z)$, возрастающей при $|z|\to\infty$ не быстрее некоторой степени $|z|$;*

в) *функция $\psi(i\lambda)=e^{i\lambda\tau}f_{\eta\xi}(\lambda)-C_\tau(i\lambda)f_\xi(\lambda)$ является граничным значением аналитической в левой полуплоскости функции $\psi(z)$, для которой $\int_{-\infty}^{\infty}|\psi(x+iy)|^2\,dy<\infty$ для $x<0$.*

Тогда однозначно определяется частотная характеристика $C_\tau(i\lambda)$ оптимального фильтра, оценивающего величину

$$\eta(t^*) = \eta(t_0+\tau).$$

При этом среднеквадратичная погрешность оптимальной оценки

$$\sigma_\tau^2 = \mathrm{M}\,|\,\eta(t_0+\tau) - \hat{\eta}(t_0+\tau)\,|^2 = B_\eta(0) - \int_{-\infty}^{\infty} |\,C_\tau(i\lambda)\,|^2 f_\xi(\lambda)\,d\lambda.$$

Пример. Рассмотрим задачу чистого прогноза процесса $\{\xi(t), t\in T\}$, спектральная плотность которого дробно-рациональна, т. е. $\xi(t) = \eta(t)$,

$$f_\xi(\lambda) = \frac{|\,P(i\lambda)\,|^2}{2\pi\,|\,Q(i\lambda)\,|^2},$$

где $P(z)$ и $Q(z)$ — многочлены степени p и q соответственно, нули которых лежат в левой полуплоскости.

Если $z_j^{(p)}$ и $z_j^{(q)}$ — нули многочленов $P(z)$ и $Q(z)$ соответственно, то имеют место представления

$$P(z) = a\prod_{j=1}^{m}\left(z - z_j^{(p)}\right)^{\alpha_j}, \quad \sum_{j=1}^{m}\alpha_j = p,$$

$$Q(z) = b\prod_{j=1}^{n}\left(z - z_j^{(q)}\right)^{\beta_j}, \quad \sum_{j=1}^{n}\beta_j = q.$$

Пусть

$$P_1(z) = (-1)^p a\prod_{j=1}^{m}\left(z + \bar{z}_j^{(p)}\right)^{\alpha_j},$$

$$Q_1(z) = (-1)^n b\prod_{j=1}^{n}\left(z + \bar{z}_j^{(q)}\right)^{\beta_j}.$$

Аналитическое продолжение функции $\psi(i\lambda) = \left[e^{i\lambda\tau} - C_\tau(i\lambda)\right] f_\xi(\lambda)$ имеет вид

$$\psi(z) = \left[e^{z\tau} - C_\tau(z)\right]\frac{P(z)\,P_1(z)}{Q(z)\,Q_1(z)}.$$

Условия теоремы Яглома требуют, чтобы функция $C_\tau(z)$ имела вид

$$C_\tau(z) = M_\tau(z)/P(z),$$

где $M_\tau(z)$ — многочлен степени $m_1 \leqslant p-1$ такой, что

$$\left.\frac{d^l M_\tau(z)}{dz^l}\right|_{z=z_k^{(q)}} = \left.\frac{d^l\left(e^{z\tau}p(z)\right)}{dz^l}\right|_{z=z_k^{(q)}},$$

$$l = 0, \ldots, \beta_k - 1, \quad k = 1, \ldots, n.$$

11.9. Разложение стационарного процесса

11.9.1. Разложение Вольда. Пусть $T_0 = \{t \in T: \quad t \leqslant t_0\}$, $H_\xi\{t_0\}$ — замыкание линейной оболочки случайных величин, порождаемых значениями скалярного процесса $\xi(t)$ для $t \in T_0$ (см. п. 11.1.1). Обозначим $H_\xi^s = \bigcap\limits_{t_0 \in T} H_\xi\{t_0\}$. Для подпространства H_ξ^s имеют место следующие возможности:

$$H_\xi^s = H_\xi \quad \text{или} \quad H_\xi^s \neq H_\xi.$$

В последнем случае крайней является ситуация, когда $H_\xi^s = 0$ (тривиальное пространство, состоящее из нулевого вектора).

Если $H_\xi^s = H_\xi$, то процесс $\{\xi(t),\ t \in T\}$ называется *сингулярным* (или *детерминированным*).

Если H_ξ^s есть собственное подпространство пространства H_ξ, то процесс $\{\xi(t),\ t \in T\}$ называется *недетерминированным*.

Если $H_\xi^s = 0$, то процесс $\{\xi(t),\ t \in T\}$ называется *регулярным* (или *вполне недетерминированным*). С точки зрения задач (линейного) прогноза сингулярность процесса $\{\xi(t),\ t \in T\}$ означает, что его линейный прогноз $\hat\eta_\tau(t) = \hat\eta(t^*)$ $(t^* = t + \tau,\ \tau > 0)$ на любое время τ вперед безошибочен, т. е.

$$\hat\eta_\tau(t) = \xi(t + \tau).$$

Если же процесс $\{\xi(t),\ t \in T\}$ регулярен, то наилучший линейный прогноз бесконечно удаленного будущего состоит лишь в указании среднего, т. е.

$$\lim_{\tau \to \infty} \hat\eta_\tau(t) = \mathrm{M}\xi(t) = 0.$$

Пусть $\{\xi(t),\ t \in T\}$ и $\{\eta(t),\ t \in T\}$ — стационарные процессы с пространствами значений H_ξ и H_η соответственно. Процесс $\{\eta(t),\ t \in T\}$ *вполне подчинен* процессу $\{\xi(t),\ t \in T\}$, если $H_\eta\{t\} \subset H_\xi\{t\}$ для всех $t \in T$.

Теорема 1 (разложение Вольда). *Всякий стационарный процесс $\{\xi(t),\ t \in T\}$ может быть представлен и притом единственным образом в виде*

$$\xi(t) = \xi^s(t) + \xi^r(t), \tag{9.1}$$

где $\xi^s(t)$ и $\xi^r(t)$ — некоррелированные между собой процессы, вполне подчиненные процессу $\{\xi(t),\ t \in T\}$, причем процесс $\xi^s(t)$ сингулярен, процесс $\xi^r(t)$ регулярен.

Величины $\xi^r(t)$ являются перпендикулярами в H_ξ, опущенными из $\xi(t)$ на подпространство H_ξ^s, а величины $\xi^s(t)$ — соответствующими проекциями.

11.9.2. Регулярная и сингулярная составляющие стационарного процесса. Пусть $F(\lambda)$ — спектральная функция процесса $\{\xi(t),\ t \in T\}$ и $F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda)$ — ее разложение Лебега, где $F_1(\lambda)$ абсолютно непрерывна, $F_2(\lambda)$ кусочно-постоянна, $F_3(\lambda)$ непрерывна и почти всюду по мере Лебега имеет равную нулю производную. Функция $F_1(\lambda)$ является спектральной функцией *регулярной* *составляющей* $\xi^r(t)$, а $F_2(\lambda)+F_3(\lambda)$ — спектральной функцией *сингулярной составляющей* $\xi^s(t)$. Сингулярная составляющая может быть прогнозирована в принципе безошибочно.

Пример. Пусть $\{\xi(t),\ t\in T\}$ — стационарный процесс с дискретным временем, спектральная функция которого кусочно-постоянна: $F_\xi(\lambda)=F_2(\lambda)$, т. е. процесс сингулярен.

Ищется линейный прогноз $\hat\eta_\tau(t_0)$ процесса $\{\xi(t),\ t\in T\}$ по наблюдениям в моменты $s\leqslant t_0$, т. е. требуется найти α_k $(k\geqslant 0)$ такие, чтобы ошибка прогноза

$$\mathsf{M}\,|\,\xi(t_0+\tau)-\hat\eta_\tau(t_0)\,|^2=\mathsf{M}\left|\,\xi(t_0+\tau)-\sum_{k=0}^{\infty}\alpha_k\xi(t_0-k)\right|^2$$

была минимальной.

Так как ошибка прогноза может быть представлена в виде

$$B_\xi(0)-\int_{-\pi}^{\pi}\sum_{k,\,l=0}\alpha_k\bar\alpha_l e^{i(l-k)\lambda}\,dF_\xi(\lambda)$$

и $F_\xi(\lambda)$ кусочно-постоянна, то можно указать α_k такие, что

$$B_\xi(0)-\int_{-\pi}^{\pi}\sum_{k,\,l=0}\alpha_k\bar\alpha_l e^{i(l-k)\lambda}\,dF_\xi(\lambda)=0,$$

т. е. прогноз может быть сделан в принципе безошибочно.

Сингулярными являются скалярные процессы, у которых $f(\lambda)=\dfrac{d}{d\lambda}F_1(\lambda)$ обращается в нуль на множестве положительной меры Лебега, либо

$$\int_{-\pi}^{\pi}\ln F_1'(\lambda)\,d\lambda=-\infty\quad\text{(дискретное время)},$$

$$\int_{-\infty}^{\infty}\frac{\ln F_1'(\lambda)}{1+\lambda^2}\,d\lambda=-\infty\quad\text{(непрерывное время)}.$$

Если же $\displaystyle\int_{-\pi}^{\pi}\ln F_1'(\lambda)\,d\lambda>-\infty\ \left(\int_{-\infty}^{\infty}\frac{\ln F_1'(\lambda)}{1+\lambda^2}\,d\lambda>-\infty\right)$, то регулярная и сингулярная составляющие таких процессов равны

$$\xi^r(t)=\int_{\bar\rho_0}e^{i\lambda t}\,d\zeta(\lambda),\qquad \xi^s(t)=\int_{\rho_0}e^{i\lambda t}\,d\zeta(\lambda),$$

где $\xi(t)=\int e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное разложение процесса $\{\xi(t),\ t\in T\}$; $\rho_0\subset\Lambda$ — множество нулевой лебеговой меры, на котором сосредоточены точки разрыва (роста) функции $F_2(\lambda)+F_3(\lambda)$; $\bar\rho_0$ — дополнение ρ_0 в Λ. Любой векторный стационарный процесс ранга 1 либо регулярен, либо сингулярен.

Так как сингулярная составляющая стационарного процесса может быть представлена в принципе безошибочно по бесконечно удаленному прошлому, наибольший интерес в задачах прогнозирования, интерполяции и фильтрации представляют регулярные процессы.

Теорема 2. *Для регулярности скалярного стационарного процесса необходимо и достаточно, чтобы он был реакцией физически осуществимого фильтра, на вход которого поступает стандартная некоррелированная последовательность (в случае дискретного времени) либо стандартный процесс с ортогональными приращениями (в случае непрерывного времени).*

Условие теоремы 2 необходимо и достаточно для регулярности векторного процесса максимального ранга. Процессы с дробно-рациональными спектральными плотностями регулярны.

11.10. Решение задач линейного прогнозирования, интерполирования и фильтрации

11.10.1. Линейное прогнозирование (экстраполяция). Пусть $\{\xi(t), t\in T\}$ — скалярный регулярный процесс с дискретным временем, $f(\lambda)=\frac{1}{2\pi}|\varphi(e^{-i\lambda})|^2$ — его спектральная плотность, $\xi(t)= \int\limits_{-\pi}^{\pi} e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное представление, $\xi(t)=\sum\limits_{s=-\infty}^{t} C(t-s)\,\zeta_0(s)$ — представление в виде процесса скользящего суммирования, $\sum\limits_{t=0}^{\infty}|C(t)|^2<\infty$.

Теорема 1. *Наилучший линейный прогноз $\hat{\xi}_\tau(t)$ значения $\xi(t+\tau)$ $(\tau>0)$ по наблюдениям $\xi(s)$ $(s\leqslant t)$ на время τ вперед дается формулой*

$$\hat{\xi}_\tau(t)=\int\limits_{-\pi}^{\pi} e^{i(t+\tau)\lambda}\,\frac{\varphi_\tau(e^{-i\lambda})}{\varphi(e^{-i\lambda})}\,d\zeta(\lambda), \tag{10.1}$$

где

$$\varphi(e^{-i\lambda})=\sum_{t=0}^{\infty} C(t)\,e^{-i\lambda t},\quad \varphi_\tau(e^{-i\lambda})=\sum_{t=\tau}^{\infty} C(t)\,e^{-i\lambda t}.$$

Ошибка прогноза $\sigma_\tau^2=\mathrm{M}\,|\xi(t+\tau)-\hat{\xi}_\tau(t)|^2$ равна

$$\sigma_\tau^2=\sum_{s=0}^{\tau-1} C(s)=2\pi\exp\left\{\frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\ln f(\lambda)\,d\lambda\right\}\sum_{s=0}^{\tau-1}|d_s|^2, \tag{10.2}$$

где d_n определяются из соотношения

$$\exp\left\{\frac{1}{\pi}\sum_{n=1}^{\infty} z^n\int\limits_{-\pi}^{\pi} e^{in\lambda}\ln f(\lambda)\,d\lambda\right\}=\sum_{n=0}^{\infty} d_n z^n.$$

В частности,

$$\sigma_1^2 = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda)\,d\lambda\right\}, \tag{10.3}$$

т. е. $\sigma_1^2/(2\pi)$ есть (непрерывное) среднее геометрическое спектральной плотности.

Пример 1. Пусть $\{\xi(t),\ t \in T\}$ $(T = 0, \pm 1, \pm 2, \ldots)$ — марковский в широком смысле процесс с корреляционной функцией $B(t) = \sigma^2 e^{-\alpha|t|}$ $(\alpha > 0)$. Спектральная плотность $f(\lambda)$ имеет вид

$$f(\lambda) = \frac{\sigma^2(1-\beta^2)}{2\pi|1-\beta e^{-i\lambda}|^2}, \quad \beta = e^{-\alpha}.$$

Наилучший линейный прогноз на время τ вперед дается формулой

$$\hat{\xi}_\tau(t) = \int_{-\pi}^{\pi} e^{i\lambda t} e^{-\alpha\tau}\,d\zeta(\lambda) = e^{-\alpha\tau}\xi(t),$$

где $\zeta(\lambda)$ — спектральный процесс, соответствующий процессу $\{\xi(t),\ t \in T\}$. Ошибка прогноза

$$\sigma_\tau^2 = \sigma^2\left(1 - e^{-2\alpha\tau}\right).$$

Пусть $\{\xi(t),\ t \in T\}$ — векторный регулярный процесс максимального ранга с дискретным временем, $f(\lambda) = \frac{1}{2\pi}\varphi\left(e^{-i\lambda}\right)\varphi^T\left(e^{-i\lambda}\right)$ — спектральная (матричная) плотность, $\xi(t) = \int_{-\pi}^{\pi} e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное представление и матрица $\varphi(z)$ разлагается в ряд

$$\varphi(z) = \sum_{n=0}^{\infty} b(n) z^n.$$

Теорема 2. *Наилучший линейный прогноз $\hat{\xi}_\tau(t)$ на время τ вперед дается формулой*

$$\hat{\xi}_\tau(t) = \int_{-\pi}^{\pi} e^{i\lambda(t+\tau)}\left[\varphi\left(e^{-i\lambda}\right) - \sum_{n=0}^{\tau-1} b(n)\,e^{-in\lambda}\right]\varphi^{-1}\left(e^{-i\lambda}\right)d\zeta(\lambda). \tag{10.4}$$

Матрица ошибок прогноза на один шаг вперед

$$G = \varphi(0)\,\varphi^T(0). \tag{10.5}$$

Пример 2. Пусть $\{\xi(t) = (\xi_1(t), \ldots, \xi_k(t)),\ t \in T\}$ — смешанный процесс авторегрессии и скользящего среднего, спектральная плотность которого имеет вид

$$f(\lambda) = \frac{1}{2\pi} B^{-1}\left(e^{-i\lambda}\right) A\left(e^{-i\lambda}\right) G A^T\left(e^{-i\lambda}\right) B^{T^{-1}}\left(e^{-i\lambda}\right),$$

где $A(z)$ и $B(z)$ — матричные многочлены, $A(0)=B(0)=I$ (единичная $k\times k$-матрица), $\det G\neq 0$. Здесь $\varphi(e^{-i\lambda})=B^{-1}(e^{-i\lambda})\times A(e^{-i\lambda})G^{1/2}$. Наилучший линейный прогноз $\hat{\xi}_\tau(t)$ дается формулой

$$\hat{\xi}_\tau(t)=\sum_{n=0}^{\infty}C_n\xi(t-n),$$

где C_n определяются как коэффициенты разложения $\sum_{n=0}^{\infty}C_nz^n=\left(\sum_{n=\tau}^{\infty}b_nz^n\right)\varphi^{-1}(z)$. Матрица ошибок прогноза на один шаг вперед совпадает с G.

В частности, для процесса авторегрессии ($A(z)\equiv I$) наилучший прогноз на один шаг вперед определяется по значениям $\xi(t)$, $\xi(t-1),\ldots,\xi(t-q)$, где q — степень многочлена $B(z)$.

Пусть $\xi(t)$ — скалярный регулярный процесс с непрерывным временем, $f(\lambda)=\frac{1}{2\pi}|\varphi(i\lambda)|^2$ — его спектральная плотность, $\xi(t)=\int_{-\infty}^{\infty}e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное представление, $\xi(t)=\int_{-\infty}^{t}C(t-s)\,d\zeta(s)$ — представление в виде процесса скользящего суммирования.

Теорема 3. *Наилучший линейный прогноз $\hat{\xi}_\tau(t)$ на время τ вперед дается формулой*

$$\hat{\xi}_\tau(t)=\int_{-\infty}^{\infty}e^{i\lambda(t+\tau)}\frac{\varphi_\tau(i\lambda)}{\varphi(i\lambda)}\,d\zeta(\lambda), \tag{10.6}$$

где $\varphi(i\lambda)=\int_0^{\infty}C(s)e^{-i\lambda s}\,ds$, $\varphi_\tau^{(i\lambda)}=\int_\tau^{\infty}C(s)e^{-i\lambda s}\,ds$. *Ошибка прогноза*

$$\sigma_\tau^2=\int_0^{\tau}|a(s)|^2\,ds. \tag{10.7}$$

Пример 3. Пусть $\{\xi(t),\ t\in T\}$ — процесс авторегрессии, спектральная плотность которого имеет вид

$$f(\lambda)=\frac{1}{2\pi|Q(i\lambda)|^2},$$

где $Q(z)$ — полином степени $q>1$, корни которого β_j ($j=1,\ldots,q$) простые и имеют положительные действительные части. Наилучший линейный прогноз на время τ вперед дается формулой

$$\hat{\xi}_\tau(t) = \int_{-\infty}^{\infty} e^{it} \sum_{j=1}^{q} e^{-\beta_j \tau} \prod_{l \neq j} \frac{\beta_l + i\lambda}{\beta_l - \beta_j} \, d\zeta(\lambda) =$$

$$= \sum_{j=1}^{q} e^{-\beta_j \tau} \prod_{l \neq j} \frac{1}{\beta_l - \beta_j} \left(\beta_l + \frac{d}{dt}\right) \xi(t)$$

11.10.2. Интерполяция пропущенного значения. Пусть $T = \{0, \pm 1, \pm 2, \ldots\}$ и $\{\xi(t),\ t \in T\}$ — стационарный процесс с абсолютно непрерывной спектральной функцией $F(\lambda)$. Предположим, что известны значения $\xi(s)$ $(s \neq t_0)$. Требуется найти наилучшую (линейную) интерполяцию $\hat{\xi}(t_0)$ пропущенного значения $\xi(t_0)$. Величина

$$\mathsf{M}\,[\xi(t_0) - \hat{\xi}(t_0)]\,[\xi^T(t_0) - \hat{\xi}^T(t_0)] = \begin{cases} \sigma^2 & \text{(скалярный процесс)}, \\ G & \text{(векторный процесс)} \end{cases}$$

называется *ошибкой (матрицей ошибок) интерполяции пропущенного значения.*

Теорема 4. *Если $\{\xi(t),\ t \in T\}$ — скалярный процесс и $\int_{-\pi}^{\pi} \frac{d\lambda}{f(\lambda)} < \infty$, где $f(\lambda)$ — его спектральная плотность, то наилучшая линейная интерполяция $\hat{\xi}(t_0)$ пропущенного значения $\xi(t_0)$ дается формулой*

$$\hat{\xi}(t_0) = \int_{-\pi}^{\pi} e^{i\lambda t_0} \left[1 - 2\pi \left(f(\lambda) \int_{-\pi}^{\pi} \frac{d\mu}{f(\mu)}\right)^{-1}\right] d\zeta(\lambda), \tag{10.8}$$

где $\zeta(\lambda)$ — спектральный процесс для $\{\xi(t),\ t \in T\}$. При этом ошибка интерполяции

$$\sigma^2 = 4\pi^2 \left(\int_{-\pi}^{\pi} \frac{d\lambda}{f(\lambda)}\right)^{-1}. \tag{10.9}$$

Пусть $\{\xi(t),\ t \in T\}$ — векторный процесс и $f(\lambda)$ — его матричная спектральная плотность. Обозначим через $f^{(-1)}(\lambda)$ обратную матрицу к $f(\lambda)$ (если $\det f(\lambda) \neq 0$) либо обобщенную обратную (если $\det f(\lambda) = 0$). Последнее означает, что $f^{(-1)}(\lambda) = [f(\lambda) + \Pi(\lambda)]^{-1} - \Pi(\lambda)$, где $\Pi(\lambda)$ однозначно определяется соотношениями $f(\lambda)\Pi(\lambda) = \Pi(\lambda)f(\lambda) = 0$ и $\Pi(\lambda) = \Pi^2(\lambda)$.

Теорема 5. *Если $\{\xi(t),\ t \in T\}$ — векторный процесс и матрица $f^{(-1)}(\lambda)$ интегрируема, то наилучшая линейная интерполяция $\hat{\xi}(t_0)$ пропущенного значения $\xi(t_0)$ дается формулой*

$$\hat{\xi}(t_0) = \int_{-\pi}^{\pi} e^{i\lambda t_0} \left[I - \left\{\frac{1}{2\pi} \int_{-\pi}^{\pi} f^{(-1)}(\lambda)\, d\lambda\right\}^{(-1)} f^{(-1)}(\lambda)\right] d\zeta(\lambda), \tag{10.10}$$

где $\zeta(\lambda)$ — спектральный процесс для $\{\xi(t),\ t \in T\}$. При этом матрица ошибок интерполяции

$$G = 4\pi^2 \left\{ \int_{-\pi}^{\pi} f^{(-1)}(\lambda)\, d\lambda \right\}^{(-1)}. \tag{10.11}$$

Пример 4. Пусть $\{\xi(t),\ t \in T\}$ — марковский в широком смысле процесс со спектральной плотностью

$$f(\lambda) = \frac{1-\beta^2}{2\pi\, |1-\beta e^{-i\lambda}|^2}, \quad \beta = e^{-\alpha}, \quad \alpha > 0.$$

Наилучшая линейная интерполяция $\hat{\xi}(t_0)$ пропущенного значения дается формулой

$$\hat{\xi}(t_0) = \int_{-\pi}^{\pi} e^{i\lambda t_0} \frac{\beta}{1+\beta^2} \left[e^{i\lambda} + e^{-i\lambda}\right] d\zeta(\lambda) =$$

$$= \frac{\beta}{1+\beta^2}\, \xi(t_0+1) + \frac{\beta}{1+\beta^2}\, \xi(t_0-1).$$

Ошибка интерполяции

$$\sigma^2 = (1-\beta^2)/(1+\beta^2).$$

11.10.3. Интерполирование значений стационарного процесса с непрерывным временем по наблюдениям в равноотстоящие дискретные моменты. Пусть $\{\xi(t),\ t \in T\}$ — скалярный стационарный процесс с непрерывным временем, спектральная функция $F(\lambda)$ которого абсолютно непрерывна. Предположим, что наблюдаются значения $\xi(n)$ ($n = 0, \pm 1, \pm 2, \ldots$).

Теорема 6. *Наилучшая линейная интерполяция $\hat{\xi}(t)$ процесса $\{\xi(t),\ t \in T\}$ по наблюдениям $\xi(n)$ ($n = 0, \pm 1, \pm 2, \ldots$) дается формулой*

$$\hat{\xi}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} \left(\sum_{l=-\infty}^{\infty} f(\lambda + 2\pi l)\, e^{it(\lambda + 2\pi l)} \right) \times$$

$$\times \left(\sum_{l=-\infty}^{\infty} f(\lambda + 2\pi l) \right)^{-1} d\zeta(\lambda), \tag{10.12}$$

где $f(\lambda)$ — спектральная плотность процесса $\{\xi(t),\ t \in T\}$ и $\zeta(\lambda)$ — его спектральный процесс. Ошибка интерполяции $\sigma^2 = \mathsf{M}\, |\xi(t) - \hat{\xi}(t)|^2$ равна

$$\sigma^2 = \int_{-\infty}^{\infty} \left| 1 - \left(\sum_{l=-\infty}^{\infty} f(\lambda + 2\pi l)\, e^{it2\pi l} \right) \times \right.$$

$$\left. \times \left(\sum_{l=-\infty}^{\infty} f(\lambda + 2\pi l) \right)^{-1} \right| f(\lambda)\, d\lambda. \tag{10.13}$$

В частности, если $\sigma^2 = 0$, то процесс $\{\xi(t),\ t \in T\}$ может быть безошибочно проинтерполирован по значениям $\xi(n)$ $(n = 0, \pm 1, \pm 2, \ldots)$. Для этого необходимо и достаточно, чтобы $f(\lambda)$ обращалась в нуль вне отрезка $[-\pi, \pi]$. В этом случае имеет место *формула Котельникова — Шеннона*

$$\hat{\xi}(t) = \sum_{-\infty}^{\infty} \frac{\sin \pi (j - t)}{\pi (j - t)} \xi(j). \tag{10.14}$$

11.10.4. Линейная фильтрация. Задача фильтрации — выделение сигнала $s(t)$ по наблюдениям стационарного процесса $\xi(t) = s(t) + \theta(t)$ — имеет простое решение в том случае, если наблюдению доступны значения процесса $\xi(t)$ на всем временно́м интервале.

Пусть $\{\xi(t),\ t \in T\}$ — скалярный стационарный процесс, имеющий абсолютно непрерывную спектральную функцию $F_\xi(\lambda)$, и пусть $f_\xi(\lambda)$, $f_s(\lambda)$ и $f_\theta(\lambda)$ — спектральные плотности процессов $\xi(t)$, $s(t)$ и $\theta(t)$ соответственно.

Теорема 7. *Частотная характеристика оптимального фильтра для выделения сигнала $s(t)$ имеет вид*

$$H(i\lambda) = \frac{f_s(\lambda)}{f_\xi(\lambda)}, \tag{10.15}$$

а среднеквадратичная ошибка фильтрации равна

$$\int_\Lambda \frac{f_s(\lambda)}{f_\xi(\lambda)} f_\theta(\lambda)\, d\lambda. \tag{10.16}$$

Пусть $\{\xi(t),\ t \in T\}$ — векторный стационарный процесс максимального ранга с абсолютно непрерывной (матричной) спектральной функцией $F_\xi(\lambda)$, и пусть $\hat{f}_\xi(\lambda)$, $\hat{f}_s(\lambda)$ и $\hat{f}_\theta(\lambda)$ — производные спектральных функций $F_\xi(\lambda)$, $F_s(\lambda)$ и $F_\theta(\lambda)$ по мере $\mu(d\lambda) = \mathrm{Sp}\, dF_\xi(\lambda)$, т. е.

$$\int_\Gamma dF_\xi(\lambda) = \int_\Gamma \hat{f}_\xi(\lambda)\, \mu(d\lambda), \quad \int_\Gamma dF_s(\lambda) = \int_\Gamma \hat{f}_s(\lambda)\, \mu(d\lambda),$$

$$\int_\Gamma dF_\theta(\lambda) = \int_\Gamma \hat{f}_\theta(\lambda)\, \mu(d\lambda).$$

Теорема 8. *Частотная характеристика оптимального фильтра для выделения сигнала $s(t)$ имеет вид*

$$H(i\lambda) = \hat{f}_s(\lambda)\, \hat{f}_\xi^{-1}(\lambda), \tag{10.17}$$

а матрица ошибок фильтрации равна

$$\int_\Lambda \hat{f}_s(\lambda)\, \hat{f}_\xi^{-1}(\lambda)\, \hat{f}_\theta(\lambda)\, \mu(d\lambda). \tag{10.18}$$

Пусть $\xi(t) = s(t) + \theta(t)$ — регулярный стационарный процесс (имеющий максимальный ранг в данном случае), $f_\xi(\lambda)$, $f_s(\lambda)$ —

спектральные плотности процессов $\xi(t)$ и $s(t)$ соответственно, $f_{s\xi}(\lambda)$ — взаимная спектральная плотность и $g(\lambda)=f_{s\xi}(\lambda)f_{\xi}^{-1}(\lambda)$. Через $\hat{s}_\tau(t)$ обозначена наилучшая в среднеквадратичном оценка сигнала $s(t+\tau)$ по наблюдениям процесса $\xi(u)$ $(u\leqslant t)$. В случае $\tau>0$ говорят о *фильтрации с прогнозом*, а в случае $\tau<0$ — о *фильтрации с запаздыванием*. Теоремы 7 и 8 описывают оптимальные фильтры в случае сколь угодно большого запаздывания.

Теорема 9. *Пусть частотная характеристика $H_\tau(i\lambda)$ оптимального фильтра для оценки сигнала $s(t+\tau)$ по наблюдениям $\xi(u)$ $(u\leqslant t)$ имеет вид*

$$H_\tau(i\lambda)=\begin{cases}\displaystyle\sum_{s=0}^{\infty} a(s+\tau)\,e^{-i\lambda s}\varphi^{-1}\left(e^{-i\lambda}\right) & (\textit{дискретное время}),\\ \displaystyle\left[\int_0^{\infty} e^{-i\lambda s}\,a(s+\tau)\,ds\right]\varphi^{-1}(i\lambda) & (\textit{непрерывное время}),\end{cases} \tag{10.19}$$

где $\varphi\left(e^{-i\lambda}\right)$, $\varphi(i\lambda)$ — компоненты факторизации спектральной плотности $f_\xi(\lambda)$,

$$a(s)=\begin{cases}\displaystyle\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\lambda s}g(\lambda)\,\varphi\left(e^{-i\lambda}\right)d\lambda & (\textit{дискретное время}),\\ \displaystyle\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda s}g(\lambda)\,\varphi(i\lambda)\,d\lambda & (\textit{непрерывное время}).\end{cases}$$

Тогда, если $\xi(t)=\int_\Lambda e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное представление процесса $\{\xi(t),\ t\in T\}$, то

$$\hat{s}_\tau(t)=\int_\Lambda e^{i\lambda t}H_\tau(i\lambda)\,d\zeta(\lambda). \tag{10.20}$$

11.11. Стационарные в узком смысле случайные процессы

11.11.1. Определение. Пусть $\{\xi(t),\ t\in T\}$ — случайный процесс со значениями в измеримом пространстве $\{\mathfrak{X},\mathfrak{B}\}$, T — одно из множеств вида $\{0,\pm1,\pm2,\ldots\}$, $\{0,1,2,\ldots\}$ (дискретное время) либо $(-\infty,\infty)$, $[0,\infty)$ (непрерывное время).

Процесс $\{\xi(t),\ t\in T\}$ называется *стационарным в узком смысле*, если для любых t, $t_k\in T$ $(k=1,\ldots,n,\ n\geqslant1)$ таких, что $t_k+t\in T$, совместное распределение в $\{\mathfrak{X}^n,\mathfrak{B}^n\}$ случайных величин $\{\xi(t_1+t),\ \xi(t_2+t),\ \ldots,\ \xi(t_n+t)\}$ не зависит от t.

Иными словами, процесс стационарен в узком смысле, если его конечномерные распределения не меняются при допустимых $(t_k+t\in T,\ k=1,\ldots,n)$ сдвигах времени.

Процесс $\{\xi(t),\ t\in T\}$ стационарен в узком смысле, если для произвольной $\mathfrak{B}^n$-измеримой функции $f(x_1, x_2, \ldots, x_n)$ $(x_k\in\mathfrak{X})$ математическое ожидание $\mathsf{M}f(\xi(t_1+t),\ \xi(t_2+t), \ldots, \xi(t_n+t))$ не зависит от t, каковы бы ни были t, t_k $(k=1, \ldots, n,\ n\geqslant 1)$ такие, что $t_k+t\in T$.

Если $\xi(t)$ — стационарный в узком смысле случайный процесс, $f(x)$ — $\mathfrak{B}$-измеримая векторнозначная функция, для которой $\mathsf{M}|f(\xi(t))|^2<\infty$, $t\in T$, то процесс $\eta(t)=f(\xi(t))$ стационарен в широком смысле.

11.11.2. Примеры. 1. Пусть $T=\{0, \pm1, \pm2, \ldots\}$ либо $T=\{0, 1, 2, \ldots\}$ и $\{\xi(t),\ t\in T\}$ — последовательность независимых и одинаково распределенных случайных величин. Процесс $\{\xi(t),\ t\in T\}$ является стационарной в узком смысле случайной последовательностью.

2. Пусть $\{\xi(t),\ t\in T\}$ — стационарная последовательность, определенная в примере 1, и пусть α_t $(t\in T)$ — последовательность действительных или комплексных чисел такая, что ряд $\sum\limits_{s\in T}\alpha_s\xi(t+s)$ $(s+t\in T)$ сходится по вероятности (и, следовательно, в силу независимости $\xi(s)$ с вероятностью 1). Процесс $\{\eta(t),\ t\in T\}$, где $\eta(t)=\sum\limits_{s\in T}\alpha_s\xi(t+s)$, является стационарной в узком смысле случайной последовательностью.

3. Пусть $T=\{0, 1, 2, \ldots\}$ либо $T=[0, \infty)$ и $\{\xi(t),\ t\in T\}$ — марковский процесс со значениями в $\{\mathfrak{X}, \mathfrak{B}\}$ ($\mathfrak{X}$ — компактное метрическое пространство), имеющий стационарное распределение $\rho(\cdot)$ (инвариантную меру).

Если распределение $\xi(0)$ совпадает с распределением $\rho(\cdot)$, то $\{\xi(t),\ t\in T\}$ — стационарный в узком смысле случайный процесс.

4. Пусть $\{\eta(t),\ t\in T\}$, $T=[0, \infty)$, — однородный процесс с независимыми приращениями. Если $f(u, x)$ — непрерывная функция, для которой

$$\int\limits_0^\infty \mathsf{M}\,|f(u, \eta(u))|\,du<\infty,$$

то процесс

$$\{\xi(t),\ t\in T\},\quad \text{где}\quad \xi(t)=\int\limits_0^\infty f(u, \eta(t+u)-\eta(t))\,du,$$

является стационарным в узком смысле.

5. Гауссовский стационарный в широком смысле процесс является стационарным и в узком смысле.

6. Пусть $\{\xi(t),\ t\in T\}$ — последовательность случайных векторов, стационарная в узком и широком смысле одновременно.

Процесс $\{\eta_h(t),\ t\in T\}$, где $\eta_h(t)=\xi(t)\xi^T(t+h)$, $t+h\in T$ и h фиксировано, является стационарной в узком смысле последовательностью случайных матриц.

11.11.3. Преобразования, сохраняющие меру. Пусть $\{\xi(t), t\in T\}$ — стационарный в узком смысле случайный процесс со значениями в $\{\mathfrak{X}, \mathfrak{B}\}$. Через $\mathfrak{X}^T$ обозначим пространство последовательностей $x=\{\ldots, x_{-1}, x_0, x_1, \ldots\}$ в случае дискретного времени либо пространство функций $x(t)$, $t \in (-\infty, \infty)$, в случае непрерывного времени T; через $\mathfrak{L}$ обозначим минимальную σ-алгебру в $\mathfrak{X}^T$, содержащую цилиндрические множества, и пусть P_ξ — мера на $\mathfrak{L}$, индуцируемая процессом $\{\xi(t),\ t \in T\}$, определенная на цилиндрических множествах из $\mathfrak{L}$ равенством: для любых $A_k \in \mathfrak{B}$ $(t_k \in T,\ k = 1, \ldots$ $\ldots, n,\ \ n \geqslant 1)$

$$\mathsf{P}_\xi\left\{x \in \mathfrak{X}^T\colon\ x(t_1) \in A_1, \ldots, x(t_n) \in A_n\right\} = \\ = \mathsf{P}\{\xi(t_1) \in A_1, \ldots, \xi(t_n) \in A_n\}.$$

Пространство $\{\mathfrak{X}^T, \mathfrak{L}, \mathsf{P}_\xi\}$ называется *представлением процесса* $\{\xi(t),\ t \in T\}$.

Для любого $t \geqslant 0$, $t \in T$, преобразование S_t пространства $\mathfrak{X}^T$ в себя, называемое *оператором сдвига времени*, определим следующим равенством:

$$\forall x \in \mathfrak{X}^T \quad x_t = S_t x, \quad x_t(s) = x(t + s).$$

Имеют место равенства

$$S_t S_s = S_{t+s}, \quad S_0 = I, \tag{11.1}$$

где I — тождественное преобразование.

Условие стационарности процесса $\{\xi(t),\ t \in T\}$ означает, что для произвольного цилиндрического множества $C \in \mathfrak{L}$

$$\mathsf{P}_\xi(C) = \mathsf{P}_\xi(S_h C).$$

Пусть $\{\mathcal{Y}, \mathfrak{F}, \mu\}$ — некоторое пространство с мерой и S — измеримое отображение $\{\mathcal{Y}, \mathfrak{F}\}$ в $\{\mathcal{Y}, \mathfrak{F}\}$. Преобразование S называется *сохраняющим меру*, если для любого $A \in \mathfrak{F}$

$$\mu(S^{-1}A) = \mu(A), \tag{11.2}$$

где $S^{-1}A$ есть полный прообраз множества A при отображении S.

Преобразование S называется *обратимым*, если существует такое измеримое преобразование S^{-1}, что $SS^{-1} = S^{-1}S = I$. Преобразование S^{-1} называется *обратным* к S.

Процесс $\{\xi(t),\ t \in T\}$ стационарен, если оператор сдвига времени S_t в $\mathfrak{X}^T$ сохраняет меру P_ξ.

Пусть $(\Omega, \mathfrak{S}, \mathsf{P})$ — основное вероятностное пространство, на котором определен процесс $\{\xi(t),\ t \in T\}$. Поставив каждому $\omega \in \Omega$ в соответствие траекторию $\xi(\cdot, \omega)$ процесса $\{\xi(t) = \xi(t, \omega),\ t \in T\}$, можно определить измеримое отображение T_ξ пространства $(\Omega, \mathfrak{S})$ в пространство $(\mathfrak{X}^T, \mathfrak{L})$.

Если T_ξ^{-1} — отображение из $(\mathfrak{X}^T, \mathfrak{L})$ в $(\Omega, \mathfrak{S})$, областью определения которого является область значений преобразования T_ξ, то преобразование S_t пространства $\mathfrak{X}^T$ индуцирует взаимно однозначное преобразование $\hat{S}_t$ пространства Ω по формуле

$$\hat{S}_t = T_\xi^{-1} S_t T_\xi. \tag{11.3}$$

Преобразование $\hat{S}_t$ сохраняет меру P. В свою очередь, преобразование $\hat{S}_t$ порождает преобразование случайных величин со значениями в $(\mathfrak{X}, \mathfrak{B})$: для любой случайной величины $\xi(\omega) \in \mathfrak{X}$

$$\left[\hat{S}_t \xi\right](\omega) = \xi\left(\hat{S}_t^{-1}\omega\right). \tag{11.4}$$

Всякий стационарный в узком смысле процесс может быть представлен в виде

$$\xi(t) = \widehat{S}_t \xi(0). \tag{11.5}$$

В частности, если время t процесса $\xi(t)$ дискретно, то

$$\xi(t) = \widehat{S}_1^t \xi(0). \tag{11.6}$$

11.11.4. Эргодические теоремы. Теория стационарных в узком смысле случайных процессов в той ее части, которая существенно использует оператор сдвига, может рассматриваться как частный случай эргодической теории сохраняющих меру преобразований некоторого пространства с мерой в себя.

Одним из наиболее важных свойств стационарных в узком смысле случайных процессов $\{\xi(t),\ t \in T\}$ является существование предела временны́х средних:

$$\begin{gathered}
\eta^+(t) = \frac{1}{t} \sum_{s=0}^{t-1} f[\xi(s)], \quad T = \{0, \pm 1, \pm 2, \ldots\} \text{ либо } T = \{0, 1, 2, \ldots\}, \\
\eta^\pm(t) = \frac{1}{2t+1} \sum_{-t}^{t} f[\xi(s)], \quad T = \{0, \pm 1, \pm 2, \ldots\}, \\
\eta^+(t) = \frac{1}{t} \int_0^t f[\xi(s)]\, ds, \quad T = (-\infty, \infty) \text{ либо } T = [0, \infty), \\
\eta^\pm(t) = \frac{1}{2t} \int_{-t}^{t} f[\xi(s)]\, ds, \quad T = (-\infty, \infty)
\end{gathered} \tag{11.7}$$

при $t \to \infty$, где f — $\mathfrak{B}$-измеримая (векторнозначная) функция и $\mathsf{M}\,|f[\xi(0)]| < \infty$.

Обозначим через $\mathfrak{S}^t \subset \mathfrak{S}$ и $\mathfrak{S}^{\pm t} \subset \mathfrak{S}$ σ-алгебры, порожденные случайными величинами $\xi(s)$ для $s \geqslant t$ и $|s| \geqslant t \geqslant 0$ соответственно, и пусть

$$\mathfrak{S}^\infty = \bigcap_{t \in T} \mathfrak{S}^t, \quad \mathfrak{S}^{\pm\infty} = \bigcap_{t \in T} \mathfrak{S}^{\pm t}.$$

Теорема Биркгофа. *Если $\{\xi(t),\ t \in T\}$ — стационарный в узком смысле случайный процесс, то с вероятностью 1 существуют пределы*

$$\lim_{t \to \infty} \eta^+(t) = \eta^+, \quad \lim_{t \to \infty} \eta^\pm(t) = \eta^\pm, \tag{11.8}$$

причем

$$\eta^+ = \mathsf{M}\,\{f[\xi(0)] \mid \mathfrak{S}^\infty\}, \quad \eta^\pm = \mathsf{M}\,\{f[\xi(0)] \mid \mathfrak{S}^{\pm\infty}\}, \tag{11.9}$$

$$\mathsf{M}\eta^+ = \mathsf{M}f[\xi(0)], \quad \mathsf{M}\eta^\pm = \mathsf{M}f[\xi(0)]. \tag{11.10}$$

Событие $A \in \mathfrak{S}$ называется *инвариантным относительно преобразования* $\hat{S}_t$ (*$\hat{S}_t$-инвариантным*), если

$$\mathsf{P}\left\{\hat{S}_t^{-1}(A)\,\Delta A\right\} = 0,$$

где Δ — символ симметрической разности множеств.

Класс всех $\hat{S}_t$-инвариантных множеств содержится в σ-алгебрах $\mathfrak{S}^\infty$ или $\mathfrak{S}^{\pm\infty}$ (последнее в случае $T = (-\infty, \infty)$ либо $T = \{0, \pm 1, \pm 2, \ldots\}$).

Если любое $\hat{S}_t$-инвариантное множество имеет вероятность 0 или 1, то преобразование $\hat{S}_t$ называется *метрически транзитивным* (эргодическим).

Очевидно, что если $\hat{S}_t$ — метрически транзитивное преобразование, то в условиях теоремы Биркгофа имеют место равенства

$$\eta^+ = \mathsf{M}f[\xi(0)], \quad \eta^\pm = \mathsf{M}f[\xi(0)].$$

Стационарный в узком смысле процесс $\{\xi(t),\ t \in T\}$ называется *эргодическим*, если σ-алгебры $\mathfrak{S}^\infty$, $\mathfrak{S}^{\pm\infty}$ тривиальны, т. е. содержат лишь события, вероятности которых равны 0 или 1.

Теорема 1. *Для того чтобы стационарный процесс $\{\xi(t),\ t \in T\}$ был эргодическим, необходимо и достаточно, чтобы выполнялось любое из двух условий:*

1) *преобразование $\hat{S}_t$ метрически транзитивно;*

2) *для любой $\mathfrak{B}$-измеримой функции $f(x)$ такой, что $\mathsf{M}|f(\xi(0))| < \infty$, функция*

$$\bar{f} = \begin{cases} \lim\limits_{t\to\infty} \dfrac{1}{t} \displaystyle\sum_{m=0}^{t-1} f\left(\hat{S}_1^m \xi(0)\right) & (\textit{дискретное время}), \\ \lim\limits_{t\to\infty} \dfrac{1}{t} \displaystyle\int_0^t f(\hat{S}_u \xi(0))\, du & (\textit{непрерывное время}) \end{cases} \tag{11.11}$$

с вероятностью 1 постоянна.

Непосредственным следствием теоремы Биркгофа является следующая теорема.

Теорема 2 (усиленный закон больших чисел для стационарных в узком смысле процессов). *Если $\{\xi(t),\ t \in T\}$ — эргодический стационарный в узком смысле случайный процесс, то с вероятностью*

$$\lim_{t\to\infty} \eta^+(t) = \mathsf{M}f[\xi(0)], \quad \lim_{t\to\infty} \eta^\pm(t) = \mathsf{M}f[\xi(0)].$$

11.11.5. Примеры эргодических процессов. 1. Последовательность независимых одинаково распределенных случайных величин $\{\xi(t),\ t \in T\}$ с $\mathsf{M}|\xi(0)| < \infty$ — эргодическая последовательность.

2. Пусть $\{\xi(t),\ t \in T\}$ — стационарный гауссовский процесс с корреляционной функцией $B(t)$. Если $\lim\limits_{t\to\infty} B(t) = 0$, то $\{\xi(t),\ t \in T\}$ — эргодический процесс.

3. Пусть $\mathfrak{X} = \{1, 2, \ldots, n\}$, $\{\xi(t),\ t \in T\}$, $T = \{0, 1, 2, \ldots\}$ — цепь Маркова со значениями в $\mathfrak{X}$ и матрицей переходных вероятностей

$$P = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

(циклическая цепь Маркова).

Если распределение $\xi(0)$ имеет вид $\mathsf{P}\{\xi(0) = k\} = 1/n$, то $\{\xi(t),\ t \in T\}$ — эргодическая стационарная последовательность.

4. Если процесс $\{\xi(t),\ t \in T\}$ эргодичен, то также эргодичен процесс $\{\eta(t),\ t \in T\}$, где $\eta(t) = f(\xi(t_1 + t), \xi(t_2 + t), \ldots, \xi(t_n + t))$, $t,\ t_k \in T$, $k = 1, \ldots, n$, $n \geqslant 1$ и $f(x_1, x_2, \ldots, x_n)$ — произвольная $\mathfrak{B}^n$-измеримая функция, $\mathsf{M}\,|f(\xi(t_1), \ldots, \xi(t_n))| < \infty$.

В частности, если $\{\xi(t),\ t \in T\}$ — последовательность независимых одинаково распределенных случайных величин, $T = \{0, 1, 2, \ldots\}$ и $f(x) = \chi_A(x)$ — индикатор множества $A \in \mathfrak{B}$, то $\frac{1}{n} \sum_{s=0}^{n-1} \chi_A(\xi(s)) = \frac{\nu_n(A)}{n}$ есть частота наступления события $\xi(s) \in A$. По теореме 2 с вероятностью 1

$$\lim_{n \to \infty} \frac{\nu_n(A)}{n} = \mathsf{M}\chi_A(\xi(0)) = \mathsf{P}\{\xi(0) \in A\}.$$

Это утверждение известно как усиленный закон больших чисел (см. п. 2.5.3).

11.11.6. Перемешивание. Пусть $\hat{\xi}_s$ — временные средние, введенные в п. 11.11.5. Говорят, что к стационарному (многомерному) процессу $\{\xi(t),\ t \in T\}$ применима центральная предельная теорема, если существуют пределы

$$\lim_{s \to \infty} \mathsf{M} s \hat{\xi}_s \hat{\xi}_s^T = C$$

для $T = [0, \infty)$ либо $T = \{0, 1, 2, \ldots\}$ или

$$\lim_{s \to \infty} \mathsf{M} 2s \hat{\xi}_s \hat{\xi}_s^T = C$$

для $T = (-\infty, \infty)$ либо $T = \{0, \pm 1, \pm 2, \ldots\}$ и если

$$\lim_{s \to \infty} F_{\hat{\xi}(s)}(x) = \Phi(x),$$

где

$$F_{\hat{\xi}(s)}(x) = \begin{cases} \mathsf{P}\{\sqrt{s}\,\hat{\xi}_s \leqslant x\}, & T = [0, \infty) \text{ либо } T = \{0, 1, 2, \ldots\}, \\ \mathsf{P}\{\sqrt{2s}\,\hat{\xi}_s \leqslant x\}, & T = (-\infty, \infty) \text{ либо } T = \{0, \pm 1, \pm 2, \ldots\} \end{cases}$$

(неравенство в случае многомерного процесса понимается поэлементно), $\Phi(x)$ — функция нормального (многомерного) распределения с нулевым средним и дисперсией (ковариационной матрицей) C.

Наличие временных средних в $F_{\xi(s)}(x)$ предполагает эргодичность процессов, к которым применима центральная предельная теорема. Однако даже при наличии всех необходимых моментов здесь требуются дополнительные условия, более жесткие, чем эргодичность.

Простейший и в достаточной мере показательный пример эргодического стационарного процесса, имеющего все моменты, к которому тем не менее не применима центральная предельная теорема, дает пример 3 из предыдущего пункта (циклическая цепь Маркова). Причина — в зависимости слагаемых в сумме

$$\sqrt{s}\,\hat{\xi}_s = \frac{1}{\sqrt{s}} \sum_{t=0}^{s-1} \xi(t).$$

Пусть $\mathfrak{S}^t_{-\infty}$ и $\mathfrak{S}^{\infty}_{t+\tau}$ $(\tau > 0)$ — σ-алгебры, порожденные процессом $\xi(t)$: $\mathfrak{S}^t_{-\infty} = \sigma\{\xi(s),\ s < t\}$, $\mathfrak{S}^{\infty}_{t+\tau} = \sigma\{\xi(s),\ s \geqslant t+\tau\}$; $\mathfrak{S}^t_{-\infty}$ интерпретируется как прошлое, а $\mathfrak{S}^{\infty}_{t+\tau}$ — как будущее процесса $\xi(t)$. Применимость центральной предельной теоремы к стационарным процессам в существенной мере связана с выполнением условий, обеспечивающих уменьшение зависимости прошлого $\mathfrak{S}^t_{-\infty}$ от будущего $\mathfrak{S}^{\infty}_{t+\tau}$ при возрастании τ. Одним из таких условий является следующее:

$$\alpha(\tau) = \sup_{\substack{A \in \mathfrak{S}^t_{-\infty} \\ B \in \mathfrak{S}^{\infty}_{t+\tau}}} |\mathsf{P}(AB) - \mathsf{P}(A)\,\mathsf{P}(B)| \to 0$$

при $\tau \to \infty$, называемое *условием сильного перемешивания*. Функция $\alpha(\tau)$ называется *коэффициентом* (*сильного*) *перемешивания*.

Процессы, удовлетворяющие условию сильного перемешивания, являются эргодическими, поскольку условие эргодичности может быть записано в виде чезаровских пределов:

$$\frac{1}{s} \sum_{\tau=1}^{s} \alpha(\tau, A, B) \to 0 \text{ при } s \to \infty \quad \text{(дискретное время)},$$

$$\frac{1}{s} \int_0^s \alpha(\tau, A, B)\, d\tau \to 0 \text{ при } s \to \infty \quad \text{(непрерывное время)}$$

для любых $A \in \mathfrak{S}^t_{-\infty}$, $B \in \mathfrak{S}^{\infty}_{t+\tau}$, где $\alpha(\tau, A, B) = \mathsf{P}(AB) - \mathsf{P}(A)\,\mathsf{P}(B)$.

Эффективные критерии проверки условия сильного перемешивания имеются для гауссовских процессов.

Пусть $\xi(t)$ — гауссовский стационарный процесс, $H^t_{-\infty}$ и $H^{\infty}_{t+\tau}$ — гильбертовы пространства, порожденные случайными величинами $\xi(s)$ для $s \leqslant t$ и $s \geqslant t+\tau$ соответственно (см. гл. 10).

Положим

$$\rho(\tau) = \sup_{\substack{\xi \in H^t_{-\infty} \\ \bar\eta \in H^\infty_{t+\tau}}} M\xi\bar\eta.$$

Теорема Колмогорова—Розанова. *Для гауссовских стационарных процессов*

$$\alpha(\tau) \leqslant \rho(\tau) \leqslant 2\pi\alpha(\tau).$$

В случае стационарных гауссовских последовательностей для выполнения условия сильного перемешивания достаточно, чтобы спектральная плотность $f(\lambda)$ была непрерывной и положительной, т. е. $f(\lambda) > C > 0$.

11.11.7. Центральная предельная теорема. Пусть $\xi(t)$ — случайный процесс, стационарный в узком смысле, и $\hat\xi_s$ — его временные средние.

Теорема 3. *Если процесс $\xi(t)$ удовлетворяет условию сильного перемешивания, $M|f[\xi(0)]|^2 < \infty$ и процесс $\zeta(t) = f[\xi(t)]$ имеет ограниченную непрерывную спектральную плотность $f_\zeta(\lambda)$, причем (в случае многомерного процесса $\zeta(t)$) $\det f_\zeta(0) \neq 0$, то для применимости к процессу $\xi(t)$ центральной предельной теоремы необходимо и достаточно следующее условие: для любого $\varepsilon > 0$ существуют числа N_ε и T_ε такие, что*

$$\int\limits_{\|x\| > N_\varepsilon} \|x\|^2 \, dF_{\hat\xi(s)}(x) \leqslant \varepsilon$$

при $s > T_\varepsilon$. При этом дисперсия (ковариационная матрица) предельного нормального распределения есть $C = 2\pi f_\zeta(0)$.

Более легко проверяемые достаточные условия применимости центральной предельной теоремы связаны либо с дополнительными предположениями о стремлении коэффициента перемешивания $\alpha(\tau)$ к нулю, либо с условиями, еще более жесткими, чем условие сильного перемешивания.

Теорема 4. *Пусть*

1) *процесс $\xi(t)$ удовлетворяет условию сильного перемешивания, причем*

$$\alpha(\tau) = O\left(\tau^{-(1+\varepsilon)}\right),$$

и $M\|\zeta(t)\|^{2+\sigma} < \infty$, $\zeta(t) = f[\xi(t)]$ при некоторых $\varepsilon > 0$, $\sigma > 4/\varepsilon$;

2) *спектральная плотность $f_\zeta(\lambda)$ процесса $\zeta(t)$ ограничена, непрерывна и (в случае многомерного процесса) $\det f_\zeta(0) \neq 0$.*

Тогда к процессу $\zeta(t)$ применима центральная предельная теорема. При этом дисперсия (ковариационная матрица) предельного нормального распределения есть $C = 2\pi f_\zeta(0)$.

Литература: [45, 46, 50, 70, 92, 93, 95, 102, 103].

Глава 12. СЛУЧАЙНЫЕ ПОЛЯ

12.1. Основные определения

12.1.1. Определение случайного поля. Конечномерные распределения. *Случайным полем* называется случайная функция нескольких вещественных переменных. Приведем более точное определение. Пусть $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ — некоторое вероятностное пространство, D — некоторое множество в $\mathbf{R}^m$. Функция $\xi(\omega, x_1, \ldots, x_m) = \xi(\omega, \boldsymbol{x})$, определенная при $\omega \in \Omega$, $(x_1, \ldots, x_m) = \boldsymbol{x} \in D$, называется *случайным полем*, определенным на множестве D, если при фиксированных $x_1, \ldots, x_m$ она $\mathfrak{S}$-измерима по ω.

Если $\mathbf{R}$ — область значений случайной функции $\xi(\omega, x_1, \ldots, x_m)$, то говорят о скалярном поле, если же $\mathbf{R}^n$ — о векторном. Наиболее естественный пример случайного поля — это поле, заданное на $D \times [0, T]$, где D — некоторая область в трехмерном пространстве; отрезок $[0, T]$ интерпретируется как отрезок времени. С помощью таких полей можно описывать случайную эволюцию непрерывных сред (например, распределение тепла при случайной теплопроводности и случайных источниках тепла, фильтрацию в случайной среде и т. п.). Функции $\xi(\omega, x_1, \ldots, x_m)$ при всевозможных фиксированных ω являются *выборочными функциями* случайного поля.

Конечномерные распределения случайного поля $\xi(\omega, \boldsymbol{x})$ $(\boldsymbol{x} \in D \subset \mathbf{R}^m)$ — это набор распределений

$$F_{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k}(A_1, \ldots, A_k) = \mathsf{P}\left\{\bigcap_{j=1}^{k} \{\xi(\omega, \boldsymbol{x}^j) \in A_j\}\right\},$$
$$\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k \in D, \quad k = 1, 2, \ldots$$

($A_1, \ldots, A_k$ — борелевские множества из области значений $\xi(\cdot)$).

При фиксированном k это k-мерные распределения случайного поля.

Конечномерные распределения случайного поля удовлетворяют условиям согласования.

1) Для любой перестановки $i_1, \ldots, i_k$ чисел $1, \ldots, k$

$$F_{\boldsymbol{x}^{i_1}, \ldots, \boldsymbol{x}^{i_k}}(A_{i_1}, \ldots, A_{i_k}) = F_{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k}(A_1, \ldots, A_k).$$

2) Для любого k

$$F_{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k}(A_1, \ldots, A_{k-1}, \mathbf{R}^n) = F_{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^{k-1}}(A_1, \ldots, A_{k-1})$$

($\mathbf{R}^n$ — область значений функции $\xi(\cdot)$).

Теорема Колмогорова (см. п. 9.1.1) утверждает, что для всякого согласованного семейства конечномерных распределений существует случайное поле $\xi(\omega, \boldsymbol{x})$ с данными конечномерными распределениями.

12.1.2. Моментные функции. Рассмотрим случайное поле с числовыми значениями $\xi(\omega, \boldsymbol{x})$. Функция

$$m_k(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k) = \mathsf{M}\xi(\omega, \boldsymbol{x}^1) \ldots, \xi(\omega, \boldsymbol{x}^k)$$

(если она определена при $\boldsymbol{x}^i \in D$ $(i = 1, \ldots, k)$) называется *моментной функцией k-го порядка случайного поля* $\xi(\omega, \boldsymbol{x})$.

Моментная функция первого порядка

$$m_1(x) = \mathsf{M}\xi(\omega, x) = a(x)$$

называется *средним значением случайного поля*. Функция

$$\tilde{m}_k(x^1, \ldots, x^k) = \mathsf{M}(\xi(\omega, x^1) - a(x^1)) \ldots (\xi(\omega, x^k) - a(x^k))$$

называется *центральной моментной функцией k-го порядка случайного поля* $\xi(\omega, x)$. Центральная моментная функция второго порядка называется *корреляционной функцией случайного поля*:

$$B(x, y) = \mathsf{M}\xi(\omega, x)\,\xi(\omega, y) - \mathsf{M}\xi(\omega, x)\,\mathsf{M}\xi(\omega, y).$$

Пусть $\xi(\omega, x)$ — случайное поле со значениями в $\mathbf{R}^n$, $\xi(\omega, x) = (\xi_1(\omega, x), \ldots, \xi_n(\omega, x))$. Векторная функция

$$\boldsymbol{a}(x) = (a_1(x), \ldots, a_n(x)) = (\mathsf{M}\xi_1(\omega, x), \ldots, \mathsf{M}\xi_n(\omega, x))$$

называется *средним значением векторного случайного поля*. Функция $B(\boldsymbol{x}, \boldsymbol{y})$, определенная на $D \times D$ и принимающая матричные значения, для которой элементы матрицы $B(x, y)$ определяются равенствами

$$b_{ij}(x, y) = \mathsf{M}\xi_i(x)\,\xi_j(y) - a_i(x)\,a_j(y),$$

называется *корреляционной матричной функцией векторного случайного поля*. Старшие моментные функции векторного поля задаются равенствами

$$m_{m_1, \ldots, m_k}(x_1, \ldots, x_k) = \mathsf{M}\xi^{m_1}(x_1) \ldots \xi^{m_k}(x_k),$$

где $m_j = \left(m_1^{(j)}, \ldots, m_n^{(j)}\right)$, $\xi^m = \xi_1^{m_1}\xi_2^{m_2} \ldots \xi_n^{m_n}$. Однако эти функции редко используются.

12.1.3. Представление векторных полей ортогональными рядами. Пусть D — ограниченное множество, $\xi(x)$ — числовое случайное поле, корреляционная функция которого $B(x, y)$ существует, и

$$\int_D B(x, x)\,dx < \infty$$

(интеграл по мере Лебега). Тогда существует полная ортогональная последовательность функций $\varphi_k(x)$ на D, являющихся собственными функциями интегрального оператора с ядром $B(x, y)$,

$$\lambda_k\varphi_k(x) = \int B(x, y)\,\varphi_k(y)\,dy,$$

при этом числа λ_k неотрицательны и $\sum\limits_k \lambda_k < \infty$.

Случайное поле $\xi(x)$ представимо в виде ряда

$$\xi(x) = a(x) + \sum_{k=1}^{\infty} \eta_k\varphi_k(x), \tag{1.1}$$

где η_k — последовательность некоррелированных случайных величин, $\mathrm{M}\eta_k = 0$, $\mathrm{D}\eta_k = \lambda_k$. Ряд (1.1) сходится в следующем смысле:

$$\lim_{N\to\infty} \int_D \mathrm{M} \left| \xi(x) - a(x) - \sum_{k=1}^{N} \eta_k \varphi_k(x) \right|^2 dx = 0.$$

Случайные величины η_k определяются из соотношений

$$\eta_k = \int \xi(x)\,\varphi_k(x)\,dx.$$

Аналогичный результат справедлив и для векторных случайных полей. Пусть поле $\xi(x)$ имеет корреляционную матричную функцию $B(x, y) = \| b_{ij}(x, y) \|$, для которой

$$\int_D \sum b_{ij}^2(x, x)\,dx < \infty.$$

Тогда существует такая последовательность $\lambda_k \downarrow 0$, для которой система интегральных уравнений

$$\lambda\varphi^i(x) = \sum \int b_{ij}(x, y)\,\varphi^j(y)\,dy, \quad i = 1, \ldots, n, \tag{1.2}$$

имеет решение при $\lambda = \lambda_k$. Если

$$\varphi_k(x) = \left(\varphi_k^1(x), \ldots, \varphi_k^n(x)\right)$$

является решением системы (1.2) при $\lambda = \lambda_k$, для которого

$$\int \sum_{i=1}^{n} |\varphi_k^i(x)|^2\,dx = 1,$$

то

$$\xi(x) = a(x) + \sum_{k=1}^{\infty} \eta_k \varphi_k(x), \tag{1.3}$$

где η_k — последовательность некоррелированных случайных величин, для которых $\mathrm{M}\eta_k = 0$, $\mathrm{D}\eta_k = \lambda_k$. Сходимость в (1.3) покоординатно такая же, как и в (1.1).

Примеры. 1. *Гауссовские случайные поля.* Поле $\xi(x)$ называется гауссовским, если для любых $x_1, \ldots, x_k \in D$ совместное распределение величин $\xi(x_1), \ldots, \xi(x_k)$ является гауссовским. Аналогично векторное поле $\xi(x)$ называется *гауссовским*, если совместное распределение величин $\xi_1(x_1), \ldots, \xi_n(x_1), \ldots, \xi_1(x_k), \ldots, \xi_n(x_k)$ является гауссовским. Конечномерные распределения для гауссовских полей $\xi(x)$ полностью определяются функциями $a(x)$ и $B(x, y)$. Функция $a(x)$ может быть произвольной, а корреляционная матричная функция должна быть положительно определена: каковы бы ни были $x_1, \ldots, x_k \in D$ и комплексные числа $\alpha_1^1, \ldots, \alpha_1^n, \ldots, \alpha_k^1, \ldots$ $\ldots, \alpha_k^n$, должно выполняться неравенство

$$\sum_{i, j, k, l} b_{ij}(x_k, x_l)\,\alpha_k^i \bar{\alpha}_l^j \geqslant 0$$

($\bar{\alpha}$ — комплексно-сопряженное число к α).

2. *Поля с независимыми приращениями.* Пусть $\xi(x)$ — случайное поле, заданное на $\mathbf{R}_+^m = \{x\colon x_i \geqslant 0,\ i = 1, \ldots, m\}$. Обозначим для прямоугольника $\Pi = \{x\colon a_i \leqslant x_i \leqslant b_i,\ i = 1, \ldots, m\}$

$$\Delta_\Pi F(x) = \Delta^{(1)}_{[a_1, b_1]} \Delta^{(2)}_{[a_2, b_2]} \ldots \Delta^{(m)}_{[a_m, b_m]} F(x),$$

где $\Delta^{(k)}_{[a, b]} = F(x_1, \ldots, x_{k-1}, b, x_{k+1}, \ldots, x_m) - F(x_1, \ldots, x_{k-1}, a, x_{k+1}, \ldots, x_m)$.

Если для непересекающихся прямоугольников $\Pi_1, \ldots, \Pi_N$, каково бы ни было N, величины $\Delta_{\Pi_1}\xi(x), \ldots, \Delta_{\Pi_N}\xi(x)$ независимы между собой, то $\xi(x)$ называется *полем с независимыми приращениями.* В том случае, когда $\xi(x) = 0$ на границе $\mathbf{R}_+^m$, стохастически непрерывное поле $\xi(x)$ имеет распределение, определяемое характеристической функцией

$$\mathsf{M}e^{i\lambda\xi(x)} = \exp\left\{ i\lambda a(x) + \int\left(e^{i\lambda z} - 1 - \frac{i\lambda z}{1+z^2}\right)\frac{1+z^2}{z^2}\, G(\Pi_x, dz)\right\}, \tag{1.4}$$

где $G(\cdot, dz)$ — мера на $\mathbf{R}_+^m \times (-\infty, \infty)$, $\Pi_a = \{x\colon 0 \leqslant x_i \leqslant a_i,\ i = 1, \ldots, m\}$ для $a \in \mathbf{R}_+^m$ (подынтегральная функция в (1.4) при $z = 0$ считается равной $-\lambda^2/2$). В частности, гауссовское поле с независимыми приращениями в этих предположениях имеет характеристическую функцию

$$\mathsf{M}e^{i\lambda\xi(x)} = \exp\left\{ i\lambda a(x) - \frac{\lambda^2}{2}\mu_2(\Pi_x)\right\}, \tag{1.5}$$

где $a(x)$ — функция, а μ_2 — мера на $\mathbf{R}_+^m$.

12.2. Свойства выборочных функций

12.2.1. Измеримые случайные поля. Интегрирование. Случайное поле $\xi(x)$ с заданными конечномерными распределениями, построенное в теореме Колмогорова (см. аналогичный результат для случайных процессов в п. 9.1.1), имеет в качестве выборочных функций все функции на D. Интерес представляет вопрос: при каких условиях существует случайное поле с заданными конечномерными распределениями, выборочные функции которых принадлежат заданному классу (измеримы, непрерывны, дифференцируемы и т. п.)?

Будем говорить, что два случайных поля $\xi(x)$ и $\eta(x)$, заданных на одном и том же множестве $D \subset \mathbf{R}^m$, называются *стохастически эквивалентными,* если

$$\forall x \in D\ \mathsf{P}\{\xi(x) = \eta(x)\} = 1.$$

Пусть D — измеримое (борелевское) множество. Случайное поле $\xi(\omega, x)$ называется *измеримым*, если функция $\xi(\omega, x)$ измерима относительно $\mathfrak{S} \times \mathfrak{B}^m$, где $\mathfrak{S}$ — σ-алгебра в вероятностном пространстве, $\mathfrak{B}^m$ — σ-алгебра борелевских множеств в R^m.

Случайное поле $\xi(x)$ *стохастически непрерывно в точке* $x_0 \in D$, если $\forall \varepsilon$

$$\lim_{x \to x_0} \mathsf{P}\{|\xi(x) - \xi(x_0)| > \varepsilon\} = 0.$$

Теорема 1. *Пусть D — измеримое множество и случайное поле $\xi(x)$ стохастически непрерывно на D. Тогда существует измеримое случайное поле $\xi'(x)$, стохастически эквивалентное $\xi(x)$.*

Пусть $\mu(dx)$ — некоторая мера, определенная на борелевских подмножествах D. Для измеримого поля $\xi'(x)$ можно рассмотреть вопрос о существовании интеграла

$$\int_D \xi'(x)\, \mu(dx).$$

Множество тех ω, для которых этот интеграл определен (т. е. множество тех ω, для которых $\int_D |\xi'(x)|\, \mu(dx) < \infty$), $\mathfrak{S}$-измеримо, и само значение интеграла также $\mathfrak{S}$-измеримо. Поле интегрируемо по мере μ, если

$$\mathsf{P}\left\{\int_D |\xi'(x)|\, \mu(dx) < \infty\right\} = 1.$$

Достаточным условием интегрируемости поля по мере μ является условие

$$\int_D \mathsf{M}\, |\xi'(x)|\, \mu(dx) < \infty. \tag{2.1}$$

Если (2.1) выполнено, то (в силу теоремы Фубини)

$$\mathsf{M} \int_D \xi'(x)\, \mu(dx) = \int_D \mathsf{M}\xi'(x)\, \mu(dx). \tag{2.2}$$

Если $\xi'(x)$ — гауссовское случайное поле и $\int_D \xi'(x)\, \mu(dx)$ определен, то этот интеграл также является гауссовской случайной величиной. При этом

$$\mathsf{M} \int_D \xi'(x)\, \mu(dx) = \int_D a(x)\, \mu(dx),$$

$$\mathsf{D}\left[\int_D \xi'(x)\, \mu(dx)\right] = \int_D \int_D B(x, y)\, \mu(dx)\, \mu(dy), \tag{2.3}$$

где $a(x)$ и $B(x, y)$ — соответственно среднее значение и корреляционная функция поля $\xi'(x)$. Заметим, что формула (2.3) справедлива для любых полей, для которых выполнено (2.1) и интеграл справа в (2.3) сходится (абсолютно).

12.2.2. Сепарабельное случайное поле. Поле $\xi(x)$ называется *сепарабельным относительно множества* $\Lambda \subset D$, если Λ счетно и плотно в D и существует такое множество $\Gamma \in \mathfrak{S}$, что $\mathsf{P}(\Gamma) = 1$, и для всякой сферы $S \subset \mathbf{R}^m$

$$\{\omega\colon \sup_{x \in \Lambda \cap S} \xi(x) = \sup_{x \in D \cap S} \xi(x)\} \subset \Omega \setminus \Gamma,$$

$$\{\omega\colon \inf_{x \in \Lambda \cap S} \xi(x) = \inf_{x \in D \cap S} \xi(x)\} \subset \Omega \setminus \Gamma.$$

Теорема 2. *Для всякого случайного поля $\xi(x)$ существует сепарабельное поле $\xi'(x)$, стохастически эквивалентное $\xi(x)$.*

Как и для случайных процессов, понятие сепарабельности удобно использовать при исследовании свойств выборочных функций случайного поля $\xi(x)$. Пусть D компактно. Для того чтобы случайное поле $\xi(x)$ было с вероятностью 1 непрерывным (т. е. чтобы почти все его выборочные функции были непрерывными), необходимо и достаточно выполнения условий:

1) $\xi(x)$ сепарабельно относительно некоторого множества $\Lambda \subset D$;
2) $\xi(x)$ равномерно непрерывно на Λ с вероятностью 1.

Теорема Ченцова. *Пусть сепарабельное случайное поле задано на прямоугольнике* $\Pi_a = \{x\colon 0 \leqslant x_i \leqslant a_i,\ i = 1, \ldots, m\}$. *Если оно непрерывно на сторонах* $\Gamma_i = \Pi_a \cap \{x\colon x_i = a_i\}$ *и существуют такие* $\alpha > 0$, $\beta > 0$ *и* $\gamma > 0$, *что для всякого прямоугольника* $\Pi \subset \Pi_a$

$$\mathsf{M}\,|\Delta_\Pi \xi(x)|^\alpha \leqslant \gamma\,[\mu_L(\Pi)]^{1+\beta}$$

(Δ_Π *определено в* 12.1, μ_L — *мера Лебега в* $\mathbf{R}^m$), *то поле* $\xi(x)$ *с вероятностью* 1 *непрерывно.*

Эта теорема представляет собой обобщение теоремы Колмогорова о непрерывности случайного процесса на случайные поля.

12.2.3. Непрерывность гауссовских полей. Для гауссовских случайных полей можно получить условия непрерывности случайного поля через его корреляционную функцию.

Пусть $\xi(x)$ — гауссовское случайное поле на компактном множестве D, $a(x)$ — его среднее значение, $B(x, y)$ — корреляционная функция. Пусть выполнены условия: а) $\xi(x)$ сепарабельно; б) $a(x)$ — непрерывная функция; в) существуют такие постоянные γ и $\delta > 0$, что при $\rho(x, x_0) < 1/2$ (ρ — расстояние в $\mathbf{R}^m$)

$$B(x_0, x_0) + B(x, x) - 2B(x, x_0) \leqslant \gamma \left[\ln \frac{1}{\rho(x, x_0)}\right]^{-(1+\varepsilon)}.$$

Тогда гауссовское поле $\xi(x)$ с вероятностью 1 непрерывно.

12.2.4. Дифференцируемость случайных полей. Поле $\xi(x)$ *дифференцируемо в точке* x_0 (в среднеквадратическом), если существуют такие случайные величины

$$\frac{\partial}{\partial x_k} \xi(x_0), \quad k = 1, 2, \ldots, m,$$

что

$$\lim_{\Delta x\to 0}\left(\frac{1}{\rho(x_0, x_0+\Delta x)}\right)^2 \mathsf{M}\left[\xi(x_0+\Delta x)-\xi(x_0)-\sum_{k=1}^{m}\frac{\partial}{\partial x_k}\xi(x_0)\Delta x_k\right]^2=0.$$

Величины $\frac{\partial}{\partial x_k}\xi(x_0)$ при этом определяются как среднеквадратические пределы

$$\lim\frac{\Delta_k\xi(x_0)}{\Delta x_k},$$

где $\Delta_k\xi(x)=\xi(x_1, \ldots, x_k+\Delta x_k, \ldots, x_m)-\xi(x_1, \ldots, x_k, \ldots, x_m)$. Для дифференцируемости $\xi(x)$ достаточно, чтобы были дифференцируемы функции $a(x)$ и $B(x, y)$ (по каждому переменному). Если окажется, что частные производные $\frac{\partial}{\partial x_k}\xi(x)$ (это также некоторые случайные поля) непрерывны с вероятностью 1, то тогда выборочные функции случайного поля $\xi(x)$ будут дифференцируемыми функциями x почти при всех ω. Аналогично определяется кратная дифференцируемость случайного поля. Достаточным условием k-кратной дифференцируемости случайного поля $\xi(x)$ является существование k-го дифференциала у функции $a(x)$ и у функции $B(x, y)$ по каждому переменному.

12.3. Однородные случайные поля

12.3.1. Определение однородного поля. Случайное поле $\xi(x)$, определенное на $D\subset\mathrm{R}^m$, называется *однородным*, если 1) D является полугруппой по сложению (если $x\in D$, $y\in D$, то $x+y\in D$); 2) $\mathsf{M}\xi(x)$ постоянно, $\mathsf{M}\xi(x)\xi(y)$ зависит только от $x-y$. Наиболее широко используются случайные поля, для которых D — группа всех целочисленных точек R^m и $D=\mathrm{R}^m$. Первые будем называть полями дискретного аргумента. Аналогично определение однородного векторного случайного поля.

12.3.2. Числовые поля дискретного аргумента. Пусть $\xi(x)$ —такое поле. Поскольку $\mathsf{M}\xi(x)=a$, где a постоянно, можно рассмотреть новое поле $\xi'(x)=\xi(x)-a$. Оно также будет однородным. Поэтому, не ограничивая общности, можно считать среднее значение поля равным 0. Корреляционная функция поля $B(x, y)$ имеет вид $B(x-y)$, где $B(z)$ — функция, заданная на D и удовлетворяющая условию положительной определенности

$$\sum_{k, j=1}^{n}B(z_k-z_j)\alpha_k\bar\alpha_j\geqslant 0, \tag{3.1}$$

каковы бы ни были n, $z_1, \ldots, z_n\in D$, комплексные числа $\alpha_1, \ldots, \alpha_n$ ($\bar\alpha_j$ — сопряженное к α_j комплексное число). Из условия (3.1) следует спектральное представление для $B(z)$:

$$B(z)=\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}\exp\left\{i\sum_{k=1}^{m}z_k\lambda_k\right\}F(d\lambda),\tag{3.2}$$

где $F(d\lambda)$ — некоторая конечная мера на прямоугольнике

$$[-\pi,\ \pi)^m=\{\lambda:\ -\pi\leqslant\lambda_j<\pi,\quad j=1,\ \ldots,\ m\};$$

$F(d\lambda)$ называется *спектральной мерой* случайного поля. Если эта мера абсолютно непрерывна относительно меры Лебега на $[-\pi,\ \pi)^m$ и

$$F(C)=\int_C f(\lambda)\,d\lambda$$

(f — плотность $F(d\lambda)$ относительно меры Лебега), то $f(\lambda)$ называется *спектральной плотностью* случайного поля. В этом случае (3.2) принимает вид

$$B(z)=\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}\exp\left\{i\sum_{k=1}^{m}z_k\lambda_k\right\}f(\lambda)\,d\lambda.\tag{3.3}$$

При определенных условиях спектральную плотность можем выразить через корреляционную функцию

$$f(\lambda)=(2\pi)^{-m}\sum_{z\in D}\exp\left\{-i\sum_{k=1}^{m}z_k\lambda_k\right\}B(z)\tag{3.4}$$

(ряд справа сходится в среднеквадратическом и равенство (3.4) имеет место, если $\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}f^2(\lambda)\,d\lambda<\infty$, или, что то же самое, $\sum_{z\in D}B^2(z)<\infty$).

Существует комплекснозначная случайная мера с ортогональными значениями $\psi(d\lambda)$ на $[-\pi,\ \pi)^m$, для которой

$$\mathsf{M}\psi(A)\,\overline{\psi(B)}=F(A\cap B),$$

такая, что для случайного поля $\xi(x)$ справедливо представление

$$\xi(x)=\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}\exp\left\{i\sum_{k=1}^{m}\lambda_k x_k\right\}\psi(d\lambda).\tag{3.5}$$

Закон больших чисел для случайного поля. *Пусть $\xi(x)$ — однородное поле в R^m, для которого $M\xi(x) = 0$. Тогда существует предел в среднеквадратическом*

$$\lim_{N\to\infty}\left(\frac{1}{2N+1}\right)^m \sum_{|x_i|\leqslant N,\ i=1,\ \dots,\ m} \xi(x), \tag{3.6}$$

равный $\psi(\{0\})$, где $\{0\}$ — множество в R^m, состоящее из одной точки $0\in R^m$. Для того чтобы предел (3.6) с вероятностью 1 равнялся $0 = M\xi(x)$ необходимо и достаточно, чтобы $M|\psi(\{0\})|^2 = F(\{0\}) = 0$. Это же будет выполнено, если

$$\lim_{\delta\downarrow 0}\ \lim_{N\to\infty} \sum_{|z_i|\leqslant N,\ i=1,\ \dots,\ m}\left(\prod_{k=1}^{m}\frac{\sin z_k\delta}{z_k}\right)B(z) = 0. \tag{3.7}$$

12.3.3. Векторные поля дискретного аргумента. Пусть $\xi(x)$ — однородное поле на целочисленной решетке R^m, принимающее значения из R^l, а $\xi^1(x), \dots, \xi^l(x)$ — компоненты поля. Будем предполагать, что $M\xi^i(x) = 0$. Положим

$$M\xi^i(x)\,\xi^j(y) = b_{ij}(x-y)$$

и обозначим матрицу $\|b_{ij}(x)\| = B(x)$. Для того чтобы $l\times l$-матрица $B(x)$ была корреляционной матричной функцией однородного векторного поля дискретного аргумента, необходимо и достаточно, чтобы для любых n, комплексных чисел $\alpha_1^1, \dots, \alpha_1^l, \dots, \alpha_n^1, \dots, \alpha_n^l$ и точек $z_1, \dots, z_n \in D$ выполнялось неравенство

$$\sum_{k,\ j=1}^{n}\ \sum_{p,\ q=1}^{l} b_{pq}(z_k - z_j)\,\alpha_k^p\overline{\alpha_j^q} \geqslant 0 \tag{3.8}$$

(это условие положительной определенности матричной функции).

Корреляционная матричная функция имеет следующее *спектральное представление*: существуют такие знакопеременные меры $F_{pq}(d\lambda)$ $(p, q = 1, \dots, l)$ на $[-\pi, \pi)^m$, что

$$b_{pq}(z) = \underbrace{\int_{-\pi}^{\pi}\dots\int_{-\pi}^{\pi}}_{m} \exp\left\{i\sum_{k=1}^{m} z_k\lambda_k\right\} F_{pq}(d\lambda), \tag{3.9}$$

при этом матрица

$$\|F_{pq}(A)\|_{p,\ q=1,\ \dots,\ l} \tag{3.10}$$

неотрицательно определена для всякого борелевского множества A из $[-\pi, \pi)^m$. Матрица (3.10) называется *спектральной матричной мерой* векторного поля. Если функция $F_{pq}(A)$ абсолютно непрерывна относительно меры Лебега на $[-\pi, \pi)^m$, т. е.

$$F_{pq}(A) = \int_A f_{pq}(\lambda)\,d\lambda, \tag{3.11}$$

то матрица $\|f_{pq}(\lambda)\|_{p,\,q=1,\,\ldots,\,l}$ называется *спектральной матричной плотностью* векторного поля. Она также неотрицательно определена.

В случае существования спектральной матричной плотности формула (3.9) принимает вид

$$b_{pq}(z)=\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}\exp\left\{i\sum_{k=1}^{m}z_k\lambda_k\right\}f_{pq}(\lambda)\,d\lambda. \tag{3.12}$$

Если спектральная матричная плотность существует, то она может быть определена с помощью следующей формулы:

$$\|f_{pq}(\lambda)\|=$$

$$=\lim_{\varepsilon\to 0}(2\pi)^{-m}\sum_{z\in D}\exp\left\{-i\sum_{k=1}^{m}z_k\lambda_k\right\}B(z)\,(1-\varepsilon)^{|z_1|+\ldots+|z_k|}. \tag{3.13}$$

Само случайное векторное поле также допускает *спектральное представление.* Существует набор комплекснозначных стохастических мер $\psi_p(d\lambda)$ на $[-\pi,\pi)^m$, удовлетворяющих условиям

$$\mathsf{M}\psi_p(A)=0,\quad \mathsf{M}\psi_p(A)\,\overline{\psi_q(B)}=F_{pq}(A\cap B)$$

для любых борелевских множеств A и B из $[-\pi,\pi)^m$, при этом

$$\xi^p(x)=\underbrace{\int_{-\pi}^{\pi}\ldots\int_{-\pi}^{\pi}}_{m}\exp\left\{i\sum_{k=1}^{m}z_k\lambda_k\right\}\psi_p(d\lambda). \tag{3.14}$$

12.3.4. Числовые поля непрерывного аргумента. Будем использовать те же обозначения, что и в случае числовых полей дискретного аргумента. Условие положительной определенности также имеет вид (3.1), однако $z_1,\ldots,z_n$ — теперь уже произвольные точки из $\mathbb{R}^m$. Корреляционная функция допускает спектральное представление

$$B(z)=\underbrace{\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}}_{m}\exp\left\{i\sum_{k=1}^{m}\lambda_k z_k\right\}F(d\lambda), \tag{3.15}$$

где $F(A)$ — конечная мера на $\mathbb{R}^m$. Эта мера называется *спектральной.* Если она абсолютно непрерывна, то ее плотность относительно меры Лебега $f(\lambda)$ называется *спектральной плотностью.* В качестве спектральной плотности может выступать любая неотрицательная функция, интегрируемая на $\mathbb{R}^m$. Соответствующая ей корреляционная функция выражается формулой

$$B(z)=\underbrace{\int_{-\infty}^{+\infty}\ldots\int_{-\infty}^{+\infty}}_{m}\exp\left\{i\sum_{k=1}^{m}\lambda_k z_k\right\}f(\lambda)\,d\lambda. \tag{3.16}$$

В том случае, когда спектральная плотность существует, она выражается через корреляционную функцию формулой

$$f(\lambda) = \lim_{\varepsilon \to 0} \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} \exp\left\{-i \sum_{k=1}^{m} \lambda_k z_k - \varepsilon \sum_{k=1}^{m} z_k^2\right\} B(z)\, dz. \tag{3.17}$$

Случайное поле также имеет спектральное представление

$$\xi(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} \exp\left\{i \sum_{k=1}^{m} \lambda_k x_k\right\} \psi(d\lambda), \tag{3.18}$$

где $\psi(d\lambda)$ — комплекснозначная стохастическая мера на R^m, для которой $M\psi(A) = 0$, $M\psi(A)\overline{\psi(B)} = F(A \cap B)$ для всех борелевских множеств A и $B \subset R^m$.

Закон больших чисел. *Существует предел в среднеквадратическом*

$$\lim_{T \to \infty} \left(\frac{1}{2T}\right)^m \underbrace{\int_{-T}^{T} \cdots \int_{-T}^{T}}_{m} \xi(x)\, dx = \psi(\{0\}). \tag{3.19}$$

Для того чтобы этот предел с вероятностью 1 *равнялся нулю, необходимо и достаточно, чтобы* $F(\{0\}) = 0$; *последнее эквивалентно условию*

$$\lim_{\delta \downarrow 0} \lim_{T \to \infty} \underbrace{\int_{-T}^{T} \cdots \int_{-T}^{T}}_{m} \left(\prod_{k=1}^{m} \frac{\sin \delta z_k}{z_k}\right) B(z)\, dz = 0. \tag{3.20}$$

12.3.5. Векторные поля непрерывного аргумента. Пусть однородное поле $\xi(x)$ определено на R^m и принимает значения из R^l; $\xi^1(x), \ldots, \xi^l(x)$ — его компоненты; $M\xi^j(x) = 0$, $M\xi^p(x)\,\xi^q(y) = b_{pq}(x - y)$.

Обозначим через $B(z) = \| b_{pq}(z) \|_{p,\, q=1,\, \ldots,\, l}$ корреляционную матричную функцию однородного поля. Условие (3.8), как и в случае дискретного аргумента, является необходимым и достаточным, чтобы непрерывная функция $B(z)$ с матричными значениями была корреляционной матричной функцией однородного векторного поля. На R^m существуют такие знакопеременные меры ограниченной вариации $F_{pq}(d\lambda)$, что справедливо спектральное представление

$$b_{pq}(z) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} \exp\left\{i \sum_{k=1}^{m} z_k \lambda_k\right\} F_{pq}(d\lambda), \tag{3.21}$$

при этом для всякого борелевского множества $A \subset \mathbf{R}^m$ матрица $\| F_{pq}(A) \|_{p,\ q=1,\ \ldots,\ l}$ неотрицательно определена. Матричная мера $\| F_{pq}(d\lambda) \|_{p,\ q=1,\ \ldots,\ l}$ называется *матричной спектральной мерой* векторного поля. Если эта мера абсолютно непрерывна относительно меры Лебега на $\mathbf{R}^m$, то матрица $\| f_{pq}(\lambda) \|_{p,\ q=1,\ \ldots,\ l}$, составленная из плотностей $f_{pq}(\lambda) = F_{pq}(d\lambda)/d\lambda$, называется *матричной спектральной плотностью* (эта функция при каждом λ является неотрицательно определенной матрицей). Матричная функция $\| f_{pq}(\lambda) \|_{p,\ q=1,\ \ldots,\ l}$ может выступать в качестве спектральной плотности, если при каждом λ она эрмитово симметрична, неотрицательна и

$$\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} | f_{pq}(\lambda) | \, d\lambda < \infty \text{ при } p,\ q = 1,\ \ldots,\ l.$$

Спектральное представление векторного поля имеет вид

$$\xi^p(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} \exp\left\{ i \sum_{k=1}^{m} \lambda_k z_k \right\} \psi_p(d\lambda), \tag{3.22}$$

где $\psi_1(d\lambda),\ \ldots,\ \psi_l(d\lambda)$ — стохастические комплекснозначные меры на $\mathbf{R}^m$, для которых

$$\mathsf{M}\psi_p(A)\overline{\psi_q(B)} = F_{pq}(A \cap B).$$

12.3.6. Дифференцирование однородных полей. Поскольку отдельные компоненты векторного поля являются числовыми полями, достаточно рассмотреть дифференцируемость лишь числового поля. Пусть $\xi(x)$ — числовое однородное поле и $F(d\lambda)$ — его спектральная мера. Для того чтобы существовала частная производная $\partial\xi(x)/\partial x_k$ в смысле сходимости в среднеквадратическом, необходимо и достаточно, чтобы существовала непрерывная производная $\partial^2 B(x) / \partial x_k^2$ или чтобы

$$\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} | \lambda_k |^2 F(d\lambda) < \infty.$$

При выполнении этих условий $\partial\xi(x)/\partial x_k$ также будет однородным полем с корреляционной функцией $\partial^2 B(x) / \partial x_k^2$ и спектральной мерой

$$F_1(A) = \int \cdots \int_A | \lambda_k |^2 F(d\lambda).$$

Для нахождения $\partial\xi(x)/\partial x_k$ можно дифференцировать спектральное представление (3.18):

$$\frac{\partial \xi(x)}{\partial x_k} = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} (i\lambda_k) \exp\left\{ i \sum_{k=1}^{m} \lambda_k x_k \right\} \psi(d\lambda). \tag{3.23}$$

Условия существования частных производных заключаются в следующем. Для того чтобы существовала производная

$$\frac{\partial^n \xi(x)}{\partial x_1^{n_1} \ldots \partial x_m^{n_m}}, \quad n = n_1 + \ldots + n_m,$$

в смысле сходимости в среднеквадратическом, непрерывная в этом же смысле, необходимо и достаточно выполнения одного из условий:

а) существует $\dfrac{\partial^{2n} B(x)}{\partial x_1^{2n_1} \ldots \partial x_m^{2n_m}}$, и эта производная непрерывна;

б) $\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} |\lambda_1|^{2n_1} \ldots |\lambda_m|^{2n_m} F(d\lambda) < \infty.$

Если эти условия выполнены, то $\dfrac{\partial^n \xi(x)}{\partial x_1^{n_1} \ldots \partial x_m^{n_m}}$ будет также однородным полем; корреляционная функция этого поля будет

$$(-1)^n \frac{\partial^{2n} B(x)}{\partial x_1^{2n_1} \ldots \partial x_m^{2n_m}},$$

а спектральная мера этого поля имеет вид

$$F_{n_1, \ldots, n_m}(A) = \int \cdots \int_A \lambda_1^{2n_1} \ldots \lambda_m^{2n_m} F(d\lambda). \tag{3.24}$$

Спектральное представление поля $\dfrac{\partial^n \xi(x)}{\partial x_1^{n_1} \ldots \partial x_m^{n_m}}$ можно получить дифференцированием формулы (3.18):

$$\frac{\partial^n \xi(x)}{\partial x_1^{n_1} \ldots \partial x_m^{n_m}} = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} i^n \lambda_1^{n_1} \ldots \lambda_m^{n_m} \exp\left\{ i \sum_{k=1}^{m} \lambda_k x_k \right\} \psi(d\lambda). \tag{3.25}$$

12.3.7. Дифференциальные операторы с постоянными коэффициентами. Пусть $L\left(\dfrac{\partial}{\partial x_1}, \ldots, \dfrac{\partial}{\partial x_m}\right)$ — дифференциальный оператор с постоянными коэффициентами; здесь $L(t_1, \ldots, t_m)$ — полином от $t_1, \ldots, t_m$ степени n.

Если поле $\xi(x)$ таково, что существуют частные производные до порядка n включительно и они непрерывны в среднеквадратическом, тогда можно к полю $\xi(x)$ применить дифференциальный оператор $L\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right)$. Обозначим

$$\eta(x) = L\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right)\xi(x). \tag{3.26}$$

Тогда $\eta(x)$ — также однородное случайное поле с корреляционной функцией

$$B_1(x) = L\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right)L\left(-\frac{\partial}{\partial x_1}, \ldots, -\frac{\partial}{\partial x_m}\right)B(x) \tag{3.27}$$

и спектральной мерой (A — борелевское подмножество $\mathbf{R}^m$)

$$F_1(A) = \int \ldots \int_A |L(i\lambda_1, \ldots, i\lambda_m)|^2 F(d\lambda). \tag{3.28}$$

Если $\xi(x)$ имеет спектральное представление (3.18), то $\eta(x)$ имеет вид

$$\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{m} L(i\lambda_1, \ldots, i\lambda_m) \exp\left\{i\sum_{k=1}^{m}\lambda_k x_k\right\}\psi(d\lambda). \tag{3.29}$$

Рассмотрим дифференциальное уравнение

$$L\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right)\xi(x) = \eta(x), \tag{3.30}$$

где $\eta(x)$ — некоторое однородное поле. Пусть $\eta(x)$ имеет спектральную меру $F_\eta(d\lambda)$ и его спектральное представление

$$\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{m} \exp\left\{i\sum_{k=1}^{m}\lambda_k x_k\right\}\psi_k(d\lambda).$$

Тогда (3.30) имеет единственное решение, являющееся однородным полем тогда и только тогда, когда

$$\underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{m} \frac{1}{|L(i\lambda_1, \ldots, i\lambda_m)|^2} F_\eta(d\lambda) < \infty.$$

В этом случае решение (3.30) имеет вид

$$\xi(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} \frac{1}{L(i\lambda_1, \ldots, i\lambda_m)} \exp\left\{ i \sum_{k=1}^{m} \lambda_k x_k \right\} \psi_\eta(d\lambda), \tag{3.31}$$

а спектральная мера $F_\xi(\cdot)$ поля $\xi(x)$ — вид

$$F_\xi(A) = \int \cdots \int_A \frac{1}{|L(i\lambda_1, \ldots, i\lambda_m)|^2} F_\eta(d\lambda). \tag{3.32}$$

12.3.8. Интегральные преобразования числовых однородных полей. Пусть $\xi(x)$ — числовое поле с корреляционной функцией $B(x)$ и спектральной мерой $F_\xi(d\lambda)$. Общее интегральное преобразование имеет вид

$$\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} K(x, y)\, \xi(y)\, dy,$$

где $K(x, y)$ — такая функция, что для всех x существует

$$\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} K(x, y_1) K(x, y_2) B(y_1 - y_2)\, dy_1 dy_2 < \infty.$$

Если $K(x, y) = K(x - y)$ и $K(z)$ — интегрируемая функция, то тогда функция

$$\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} K(x - y)\, \xi(y)\, dy \tag{3.33}$$

будет также однородным полем. Корреляционная функция $B_\eta(x)$ этого поля задается формулой

$$B_\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{m} K(y_1) K(y_2) B(x - y_1 + y_2)\, dy_1 dy_2, \tag{3.34}$$

а спектральная мера поля $\eta(x)$ имеет вид

$$F_\eta(A) = \int \cdots \int_A |\tilde{k}(\lambda)|^2 F_\xi(d\lambda), \tag{3.35}$$

где

$$\tilde{k}(\lambda) = \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{m} \exp\left\{-i \sum_{k=1}^{m} \lambda_k x_k\right\} K(x)\, dx. \qquad (3.36)$$

Если существует спектральная плотность $f_\xi(\lambda)$ для поля $\xi(x)$, то тогда и поле $\eta(x)$ имеет спектральную плотность

$$f_\eta(\lambda) = |\tilde{k}(\lambda)|^2 f_\xi(\lambda). \qquad (3.37)$$

12.3.9. Интегральное преобразование векторного однородного поля. Пусть $\xi(x)$ — однородное поле со значениями в $\mathbf{R}^l$, $B_\xi(x)$ — матричная корреляционная функция поля $\xi(x)$, $F_\xi(d\lambda)$ — матричная спектральная функция $\xi(x)$. Пусть далее $K(x)$ — матричная функция, отображающая $\mathbf{R}^l$ в $\mathbf{R}^p$ (т. е. матрица порядка $l \times p$). Если все элементы этой матрицы абсолютно интегрируемы на $\mathbf{R}^m$, то определен интеграл

$$\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{m} K(x-y)\, \xi(y)\, dy \qquad (3.38)$$

(результат применения K к вектору ξ из $\mathbf{R}^p$), при этом $\eta(x)$ является однородным векторным полем со значениями в $\mathbf{R}^p$. Корреляционная матричная функция $B_\eta(x)$ имеет вид

$$B_\eta(x) = \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{m} K(y_1)\, B_\xi(x - y_1 + y_2)\, K^*(y_1)\, dy_1 dy_2, \qquad (3.39)$$

а матричная спектральная мера — вид

$$F_\eta(A) = \int \dots \int_A \tilde{K}(\lambda)\, dF_\xi(d\lambda)\, \tilde{K}^*(\lambda)\, d\lambda, \qquad (3.40)$$

где K^* — матрица, сопряженная с K; $\tilde{K}$ — матрица с элементами

$$\tilde{k}_{ij}(\lambda) = \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{m} \exp\left\{-i \sum_{k=1}^{m} \lambda_k x_k\right\} k_{ij}(x)\, dx,$$

где k_{ij} — элементы матрицы K; $\tilde{K}^*$ — эрмитово сопряженная с $\tilde{K}$. Если существует спектральная матричная плотность $\|f_\xi\|(\lambda)$ поля $\xi(x)$, то существует и спектральная матричная плотность для $\eta(x)$:

$$\|f_\eta\|(\lambda) = \tilde{K}(\lambda) \|f_\xi\|(\lambda)\, \tilde{K}^*(\lambda) \qquad (3.41)$$

($\|f_\eta\|(\lambda)$ — $l \times l$-матрица).

12.4. Изотропные случайные поля

12.4.1. Однородные изотропные поля. Однородное случайное поле $\xi(x)$, определенное на R^m со значениями в R^l, называется *изотропным*, если его корреляционная матричная функция $B(x)$ зависит лишь от $|x|$, где $|x|$ — норма вектора в R^m. Пусть $B(|x|)$ — корреляционная функция числового однородного и изотропного поля. Тогда $B(r)$ допускает следующее представление:

$$B(r) = 2^{(m-2)/2}\Gamma\left(\frac{m}{2}\right)\int_0^\infty J_{(m-2)/2}(\lambda r)(\lambda r)^{(m-2)/2}\,d\Phi(\lambda), \qquad (4.1)$$

где $J_k(\lambda)$ — бесселева функция первого рода порядка k, а $\Phi(\lambda)$ — неубывающая на $[0, \infty)$ функция ограниченной вариации. Она выражается через спектральную функцию однородного поля по формуле

$$\Phi(\lambda) = \int_{|v|<\lambda} F(dv). \qquad (4.2)$$

Функция $\Phi(\lambda)$ также называется *спектральной функцией однородного изотропного поля.* Она выражается через $B(r)$ следующим образом:

$$\Phi(\lambda) = \frac{1}{2^{(m-2)/2}\,\Gamma(m/2)}\int_0^\infty J_{m/2}(\lambda r)(\lambda r)^{m/2}\frac{B(r)}{r}\,dr. \qquad (4.3)$$

Запишем теперь спектральное представление для самого поля. Пусть $(r, \theta_1, \ldots, \theta_{m-2}, \varphi)$ — сферические координаты точки в R^m ($r \in [0, \infty)$, $\theta_i \in [-\pi/2, \pi/2]$, $\varphi \in [0, 2\pi)$). Обозначим через

$$S_n^k(\theta_1, \ldots, \theta_{m-2}, \varphi),\ k = 1, 2, \ldots, h(n, m) = \frac{(2n+m-2)(n+m-3)!}{(m-2)!\,n!}$$

ортонормированную последовательность сферических гармоник степени n. Тогда, если $M\xi(x) = 0$, то

$$\xi(x) = c_m\sum_{n=0}^\infty\sum_{k=1}^{h(n,m)} S_n^k(\theta_1, \ldots, \theta_{m-2}, \varphi)\times$$
$$\times\int_0^\infty J_{n+(m-2)/2}(\lambda r)(\lambda r)^{(2-m)/2}\,\psi_n^k(d\lambda), \qquad (4.4)$$

где $c_m^2 = 2^{m-1}\Gamma(m/2)\,\pi^{m/2}$, а $\psi_n^k(d\lambda)$ $(n = 0, 1, \ldots; k = 1, \ldots, h(n, m))$ — комплекснозначные случайные меры с ортогональными значениями на $[0, \infty)$, причем для любых борелевских множеств Λ_1 и Λ_2 в $[0, \infty)$

$$\mathsf{M}\psi_n^k(\Lambda_1)\,\psi_p^j(\Lambda_2) = \delta_n^p \delta_k^j \int_{\Lambda_1 \cap \Lambda_2} d\Phi(\lambda),$$

$$\mathsf{M}\psi_n^k(\Lambda_1) = 0.$$

12.4.2. Однородные изотропные поля со значениями в $\mathbf{R}^l$. Пусть $B(|z|)$ — матричная корреляционная функция однородного изотропного поля, $b_{ij}(r)$ — элементы матрицы $B(r)$. Тогда

$$b_{ij}(r) = 2^{(m-2)/2}\Gamma\left(\frac{m}{2}\right)\int_0^\infty J_{(m-2)/2}(\lambda r)\,(\lambda r)^{(m-2)/2}\,d\Phi_{ij}(\lambda), \quad (4.5)$$

где функции $\Phi_{ij}(\lambda)$ определены на $[0, \infty)$, причем при $\lambda_1 < \lambda_2$ матрица $\|\Phi_{ij}(\lambda_2) - \Phi_{ij}(\lambda_1)\|$ неотрицательно определена. Матрица $\|\Phi_{ij}(\lambda)\| = \Phi(\lambda)$ называется *спектральной матрицей однородного изотропного поля.* Формула (4.5) может быть записана в матричном виде так:

$$B(r) = 2^{(m-2)/2}\Gamma\left(\frac{m}{2}\right)\int_0^\infty J_{(m-2)/2}(\lambda r)\,(\lambda r)^{(m-2)/2}\,d\Phi(\lambda). \quad (4.6)$$

Спектральное представление такого поля имеет вид

$$\xi(x) = c_m \sum_{n=0}^{\infty} \sum_{k=1}^{h(n,m)} S_n^k(\theta_1, \ldots, \theta_{m-2}, \varphi) \times$$
$$\times \int_0^\infty J_{n+(m-2)/2}(\lambda r)\,(\lambda r)^{(2-m)/2}\,\psi_n^k(d\lambda), \quad (4.7)$$

где c_m и S_n^k те же, что и в формуле (4.4), а $\psi_n^k(d\lambda)$ — попарно ортогональные случайные меры со значениями в $\mathbf{R}^l$, для которых

$$\mathsf{M}_i\psi_n^k(\Lambda_1)_j\,\psi_n^k(\Lambda_2) = \int_{\Lambda_1 \cap \Lambda_2} d\Phi_{ij}(\lambda), \quad \mathsf{M}\psi_n^k(\Lambda_1) = 0 \quad (4.8)$$

$\left({}_1\psi_n^k, \ldots, {}_l\psi_n^k\right.$ — координаты вектора $\left.\psi_n^k\right)$.

12.4.3. Изотропные поля на сфере. Случайное поле $\xi(x)$, заданное на сфере S пространства $\mathbf{R}^m$ и принимающее значения из $\mathbf{R}^l$, называется *изотропным*, если $\mathsf{M}\xi(x)$ постоянно, а корреляционная матричная функция $B(x, y)$ зависит лишь от углового расстояния между точками сферы x, y.

Пусть $\xi(x)$ — числовое изотропное поле на сфере S. Тогда его корреляционная функция имеет вид $B(\cos\theta)$, где θ — угловое расстояние между точками на сфере. Существуют такие неотрицательные коэффициенты b_n, что

$$B(\cos\theta) = \frac{1}{\omega_m}\sum_{n=0}^{\infty} b_n \frac{C_n^{(m-2)/2}(\cos\theta)}{C_n^{(m-2)/2}(1)} h(n, m), \tag{4.9}$$

где ω_m — лебегов $(m-1)$-мерный объем единичной сферы в R^m, $C_n^k(\cdot)$ — многочлены Гегенбауэра и ряд $\sum_{n=0}^{\infty} b_n h(n, m)$ сходится.

Случайное поле при $M\xi(x) = 0$ допускает следующее представление:

$$\xi(x) = \sum_{n=0}^{\infty}\sum_{k=1}^{h(n,m)} \xi_n^k S_n^k(x), \tag{4.10}$$

где ξ_n^k — случайные величины, для которых $M\xi_n^k = 0$, $M\xi_n^k\xi_p^j = b_n\delta_j^k\delta_p^n$.

Пусть теперь $\xi(x)$ — векторное изотропное поле на сфере S. Его матричная корреляционная функция имеет вид

$$B(\cos\theta) = \frac{1}{\omega_n}\sum_{n=0}^{\infty} h(n, m)\frac{C_n^{(m-2)/2}(\cos\theta)}{C_n^{(m-2)/2}(1)} B_n, \tag{4.11}$$

где B_n — последовательность неотрицательных симметрических матриц порядка $l\times l$, для которых ряд $\sum_{n=0}^{\infty} h(n, m) B_n$ сходится. Если $M\xi(x) = 0$, то само поле имеет вид

$$\xi(x) = \sum_{n=0}^{\infty}\sum_{k=1}^{h(n,m)} \xi_n^k S_n^k(x), \tag{4.12}$$

где ξ_n^k — последовательность попарно некоррелированных случайных векторов, для которых

$$M\xi_n^k = 0, \quad M_i\xi_{nj}^k\xi_n^k = b_n^{ij}$$

$\left({}_1\xi_n^k, \ldots, {}_l\xi_n^k\right.$ — компоненты вектора ξ_n^k, а b_n^{ij} — элементы матрицы $\left.B_n\right)$.

12.5. Обобщенные случайные поля

12.5.1. Определение обобщенного случайного поля. Обобщенное случайное поле в R^m определяется следующим образом. Пусть D — пространство вещественных функций $\varphi(x)$ $(x \in R^m)$, финитных и бесконечно дифференцируемых, с топологией, задаваемой окрестностями вида $\left\{\varphi\colon \sup_x \left|\varphi^{(k)}_{x_{i_1},\ldots,x_{i_k}}(x)\right| < \alpha\right\} \cap \{\varphi\colon \varphi(x) = 0,\ |x| > \beta\}$, где $\alpha, \beta > 0$; $k = 0, 1, 2, \ldots$; $\varphi^{(0)} = \varphi$; $\varphi^{(k)}_{x_{i_1}, \ldots, x_{i_k}}(x) = \dfrac{\partial^k}{\partial x_{i_1} \ldots \partial x_{i_k}} \varphi(x)$; $(x_1, \ldots, x_m)$ — координаты точки.

Пусть далее $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ — фиксированное вероятностное пространство. Семейство случайных величин $\xi(\varphi)$, определенных на этом вероятностном пространстве для всех $\varphi \in D$, называется *обобщенным случайным полем*, если выполнены следующие условия:

а) $\mathsf{P}\{\xi_1(d\varphi + \beta\psi) = \alpha\xi(\varphi) + \beta\xi(\psi)\} = 1$ для всех $\alpha, \beta \in \mathbf{R}$, $\varphi, \psi \in D$;

б) $\xi(\varphi_n) \to 0$ по вероятности, если $\varphi_n \to 0$ в топологии D.

Конечномерные распределения обобщенного случайного поля $\xi(\varphi)$ — это набор функций распределений $F_{\varphi_1, \ldots, \varphi_n}(t_1, \ldots, t_n)$, где $\varphi_1, \ldots, \varphi_n \in D$, $t_1, \ldots, t_n \in \mathbf{R}$ $(n = 1, 2, \ldots)$. Они согласованы: при одинаковых перестановках знаков $\varphi_1, \ldots, \varphi_n$ и $t_1, \ldots, t_n$ значения функции не меняются и

$$\lim_{t_n \uparrow +\infty} F_{\varphi_1, \ldots, \varphi_n}(t_1, \ldots, t_{n-1}, t_n) = F_{\varphi_1, \ldots, \varphi_{n-1}}(t_1, \ldots, t_{n-1}).$$

Дополнительные условия согласования возникают в силу а) и б). Их удобнее формулировать в терминах характеристического функционала обобщенного случайного поля, который очевидным образом выражается через одномерные распределения:

$$\chi(\varphi) = \mathsf{M} e^{i\xi(\varphi)} = \int e^{it}\, dF_{\varphi}(t). \tag{5.1}$$

а′) $\displaystyle\int \exp\left\{ i \sum_{k=1}^{n} \lambda_k t_k \right\} dF_{\varphi_1, \ldots, \varphi_n}(t_1, \ldots, t_n) = \chi\left(\sum_{k=1}^{n} \lambda_k \varphi_k \right) = \int e^{it}\, dF_{\sum_{k=1}^{n} \lambda_k \varphi_k}(t)$;

б′) $\chi(\varphi_n) \to 1$, если $\varphi_n \to 0$ в топологии D.

Теорема 1. *Если заданы конечномерные распределения* $F_{\varphi_1, \ldots, \varphi_n}$, *удовлетворяющие условиям* а′) *и* б′), *то существует на некотором вероятностном пространстве обобщенное случайное поле с этими конечномерными распределениями.*

(Это аналог теоремы Колмогорова о существовании случайной функции с заданными конечномерными распределениями; см. п. 9.1.1, теорема 1.)

Если для некоторого натурального r $\mathsf{M}|\xi(\varphi)|^r < \infty$, $\varphi \in D$, то определены моментные функции обобщенного случайного поля

$$m_k(\varphi_1, \ldots, \varphi_k) = \mathsf{M}\xi(\varphi_1) \ldots \xi(\varphi_k), \quad \varphi_1, \ldots, \varphi_k \in D, \quad k \leqslant r.$$

В частности,

$$m_1(\varphi) = m(\varphi) = \mathsf{M}\xi(\varphi)$$

называется *средним значением* обобщенного случайного поля, а

$$b(\varphi_1, \varphi_2) = \mathsf{M}\xi(\varphi_1)\xi(\varphi_2) - m(\varphi_1) m(\varphi_2)$$

его *корреляционной функцией*; $m(\varphi)$ — неслучайная обобщенная функция на D, $m_2(\varphi_1, \varphi_2)$ — положительно определенная билинейная функция на D^2.

Примеры. 1. Обобщенное случайное поле $\xi(\varphi)$ называется *гауссовским*, если $\xi(\varphi)$ для всех $\varphi \in D$ имеет гауссовское распределение. Пусть $m(\varphi) = \mathsf{M}\xi(\varphi)$ и $b(\varphi) = \mathsf{M}\xi^2(\varphi) - m^2(\varphi)$. Эти две функции полностью определяют гауссовское поле. Они называются *средним значением* и *дисперсией* этого поля, $m(\varphi)$ — произвольный непрерывный линейный функционал на D, а $b(\varphi)$ — непрерывный квадратический неотрицательный функционал на D. Последнее означает, что $b(\varphi_n) \to 0$, когда $\varphi_n \to 0$ в топологии D, и что билинейная форма $b(\varphi_1, \varphi_2)$, определяемая из равенства

$$b(\alpha_1\varphi_1 + \alpha_2\varphi_2) = \alpha_1^2 b(\varphi_1) + 2\alpha_1\alpha_2 b(\varphi_1, \varphi_2) + \alpha_2^2 b(\varphi_2), \tag{5.2}$$

является неотрицательно определенной.

2. *Обобщенное поле с независимыми значениями.* Так называются поля $\xi(\varphi)$, для которых величины $\xi(\varphi_1), \ldots, \xi(\varphi_n)$ независимы, каковы бы ни были $\varphi_1, \ldots, \varphi_n \in D$ такие, что $\varphi_i\varphi_j = 0$ при $i \neq j$. Примером поля с независимыми значениями может служить поле, задаваемое равенством

$$\xi(\varphi) = \int \varphi(x)\,\mu(dx), \tag{5.3}$$

где μ — случайная мера с независимыми значениями на непересекающихся множествах; интеграл в правой части (5.3) является стохастическим (см. § 10.2).

12.5.2. Однородные в широком смысле обобщенные случайные поля. В дальнейшем мы будем рассматривать лишь L_2-теорию полей, в которой заданы лишь первые два момента поля $\xi(\varphi)$: $m(\varphi) = \mathsf{M}\xi(\varphi)$ и $b(\varphi) = \mathsf{M}\xi^2(\varphi) - m^2(\varphi)$. Корреляционная функция поля вычисляется с помощью формулы (5.2). Введем в D оператор сдвига θ_a: $\theta_a\varphi(x) = \varphi(x + a)$, $\varphi \in D$, $a \in \mathbf{R}^m$.

Обобщенное случайное поле $\xi(\varphi)$ называется *однородным*, если для всех $y \in D$, $a \in \mathbf{R}^m$

$$m(\varphi) = m(\theta_a\varphi), \qquad b(\varphi) = b(\theta_a\varphi).$$

В случае $\mathbf{R}^m = \mathbf{R} = (-\infty, \infty)$ обобщенное однородное случайное поле с независимыми значениями (пример 2) называется *процессом обобщенного белого шума.*

Теорема 2. *Пусть $\xi(\varphi)$ — обобщенное однородное случайное поле. Тогда существуют такая постоянная $m_0 \in \mathbf{R}$ и мера $F(d\lambda)$ в $\mathbf{R}^m$, для которых при некотором $p > 0$*

$$\int_{\mathbf{R}^m} (1 + |\lambda|^p)^{-1} F(d\lambda) < \infty \tag{5.4}$$

такие, что

$$m(\varphi) = m_0 \int \varphi(x)\,dx, \tag{5.5}$$

$$b(\varphi) = \int |\tilde{\varphi}(\lambda)|^2 F(d\lambda), \tag{5.6}$$

где $\tilde{\varphi}(\lambda) = \int e^{i(\lambda, x)} \varphi(x)\, dx$ $((\cdot, \cdot)$ — *скалярное произведение в* $\mathbf{R}^m)$. *Постоянная* m_0 *и мера* $F(d\lambda)$ *определяются однозначно. Формулы* (5.5) *и* (5.6) *при условии, что* $F(d\lambda)$ *удовлетворяет* (5.4) *при некотором* $p > 0$, *задают среднее значение и дисперсию некоторого однородного обобщенного случайного поля.*

Замечания. 1. В том случае, когда $F(\mathbf{R}^m) < \infty$, найдется такое однородное случайное поле $\eta(x)$ (в смысле § 12.3), что

$$\xi(\varphi) = \int \eta(x)\, \varphi(x)\, dx.$$

2. Если $\xi(\varphi)$ — комплекснозначное поле, то для его корреляционной функции

$$b(\varphi_1, \varphi_2) = \mathsf{M}\xi(\varphi_1)\overline{\xi(\varphi_2)} - \mathsf{M}\xi(\varphi_1)\overline{\mathsf{M}\xi(\varphi_2)}$$

справедливо представление

$$b(\varphi_1, \varphi_2) = \int \tilde{\varphi}_1(\lambda)\overline{\tilde{\varphi}_2(\lambda)}\, F(d\lambda),$$

где $\tilde{\varphi}_k(\lambda) = \int \varphi_k(x)\, e^{i(\lambda, x)}\, dx$, а мера F такая же, как в теореме 1.

Рассмотрим теперь векторные обобщенные случайные поля. Поле со значениями в $\mathbf{R}^l$ задается набором l числовых обобщенных случайных полей: $\xi(\varphi) = (\xi_1(\varphi), \ldots, \xi_l(\varphi))$.

В L_2-теории считаются заданными вектор

$$a(\varphi) = (\mathsf{M}\xi_1(\varphi), \ldots, \mathsf{M}\xi_l(\varphi)) = (a_1(\varphi), \ldots, a_l(\varphi))$$

и матрица $B(\varphi) = (b_{ij}(\varphi))$ $(i, j = 1, \ldots, l)$,

$$b_{kj}(\varphi) = \mathsf{M}\xi_k(\varphi)\,\xi_j(\varphi) - \mathsf{M}\xi_k(\varphi)\,\mathsf{M}\xi_j(\varphi).$$

Матрица $B(\varphi)$ называется *корреляционной матрицей* поля ξ. При этом $a_1(\varphi), \ldots, a_l(\varphi)$ — произвольные линейные непрерывные функционалы на D, а $b_{ij}(\varphi)$ — непрерывные квадратические функционалы, удовлетворяющие условию симметричности: $b_{ij} \equiv b_{ji}$, положительной определенности: для любых $\alpha_1, \ldots, \alpha_l \in \mathbf{R}$

$$\sum b_{ij}(\varphi)\, \alpha_i \alpha_j \geqslant 0.$$

Обобщенное векторное поле $\xi(\varphi)$ со значениями в $\mathbf{R}^l$ называется *однородным*, если для всех $\varphi \in D$ и $a \in \mathbf{R}^m$ $a(\varphi) = a(\theta_a\varphi)$, $B(\varphi) = B(\theta_a\varphi)$.

Теорема 3. *Пусть* $\xi(\varphi)$ — *некоторое обобщенное векторное случайное поле со средним значением* $a(\varphi)$ *и корреляционной матрицей* $B(\varphi)$. *Тогда существуют такой неслучайный вектор* a *и мера* $F(d\lambda)$ *на борелевских множествах* $\mathbf{R}^m$, *значениями которой служат неотрицательные симметрические матрицы и для которой при некотором* $p > 0$

$$\int_{\mathbf{R}^m} \frac{\operatorname{Sp} F(d\lambda)}{1 + |\lambda|^p} < \infty \tag{5.7}$$

(Sp — *след матрицы*), *что*

$$a(\varphi) = a\int \varphi(x)\,dx, \tag{5.8}$$

$$b_{kj}(\varphi) = \int |\tilde{\varphi}(\lambda)|^2 F_{ij}(d\lambda), \tag{5.9}$$

где $F_{ij}(d\lambda)$ — *элемент матрицы* $F(d\lambda)$. *Вектор* a *и мера* $F(d\lambda)$ *определяются по полю однозначно. Формулы* (5.8) *и* (5.9) *при условии* (5.7) *всегда определяют среднее значение и корреляционную функцию некоторого однородного обобщенного случайного поля.*

Замечание. Если $\operatorname{Sp} F(\mathbf{R}^m) < \infty$, то существует такое обычное однородное векторное поле $\eta(x) = (\eta_1(x), \ldots, \eta_l(x))$, что

$$\xi_k(\varphi) = \int \eta_k(x)\,\varphi(x)\,dx, \qquad k = 1, \ldots, l.$$

Теоремы 2 и 3 дают спектральное представление для дисперсии или корреляционной матрицы числового (векторного) обобщенного однородного поля. Эти представления дают возможность получать спектральное представление для самого поля.

Теорема 2′. *В условиях теоремы 2 справедливо следующее представление для* $\xi(\varphi)$:

$$\xi(\varphi) = m\int \varphi(x)\,dx + \int \tilde{\varphi}(\lambda)\,\mu(d\lambda), \tag{5.10}$$

где $\mu(d\lambda)$ — *комплекснозначная случайная мера на борелевских множествах* $\mathbf{R}^m$, *для которой* $\mathsf{M}\mu(A) = 0$, $\mathsf{M}\mu(A)\,\overline{\mu(B)} = F(A \cap B)$. *Эта мера однозначно определяется по полю* $\xi(\varphi)$.

Теорема 3′. *В условиях теоремы 3 справедливо следующее представление для* $\xi_k(\varphi)$:

$$\xi_k(\varphi) = a_k\int \varphi(x)\,dx + \int \tilde{\varphi}(\lambda)\,\mu_k(d\lambda), \tag{5.11}$$

где $\mu_k(d\lambda)$ $(k = 1, \ldots, l)$ — *комплекснозначные случайные меры на борелевских множествах* $\mathbf{R}^m$, *для которых* $\mathsf{M}\mu_k(A) = 0$, $\mathsf{M}\mu_k(A) \times \times \overline{\mu_j(B)} = F_{kj}(A \cap B)$. *Эти меры однозначно определяются по полю* $\xi(\varphi)$.

Замечание. Обобщенные случайные поля можно дифференцировать по правилам дифференцирования обобщенных функций, при этом опять будут получаться обобщенные случайные поля $\left(\text{например, } \dfrac{\partial\xi}{\partial x_k}(\varphi) = -\xi\left(-\varphi'_{x_k}\right)\right)$.

Формулы (5.10) и (5.11), дающие представления однородных случайных полей, можно дифференцировать по x сколько угодно раз. При этом в правую часть вместо $\tilde{\varphi}(\lambda)$ подставляется преобразование Фурье соответствующей производной с соответствующим знаком. Так, например, формула (5.10) дифференцируется следующим образом:

$$\frac{\partial^n}{\partial x_1^{k_1} \ldots \partial x_m^{k_m}}\,\xi(\varphi) = \int \tilde{\varphi}(\lambda)\,(-i)^n \lambda_1^{k_1} \ldots \lambda_m^{k_m} \mu(d\lambda),$$

$$n = k_1 + k_2 + \ldots + k_m.$$

12.5.3. Обобщенные поля с однородными приращениями. Обозначим через D_0 множество тех $\varphi \in D$, для которых $\int \varphi(x)\,dx = 0$. Поле $\xi(\varphi)$ называется *полем с однородными приращениями*, если оно определено при $\varphi \in D_0$, $\mathsf{M}\xi(\theta_a\varphi) = \mathsf{M}\xi(\varphi)$, $\mathsf{M}\xi^2(\theta_a\varphi) = \mathsf{M}\xi^2(\varphi)$ для всех $a \in \mathrm{R}^m$. Векторное поле $\xi(\varphi) = (\xi_1(\varphi), \ldots, \xi_l(\varphi))$ называется *векторным полем с однородными приращениями*, если $\mathsf{M}\xi_k(\theta_a(\varphi)) = \mathsf{M}\xi_k(\varphi)$, $\mathsf{M}\xi_k(\theta_a\varphi)\,\xi_j(\theta_a\varphi) = \mathsf{M}\xi_k(\varphi)\,\xi_j(\varphi)$ для всех $a \in \mathrm{R}^m$, $\varphi \in D_0$, $k, j \leqslant m$

Теорема 4. *Пусть $\xi(\varphi) = (\xi_1(\varphi), \ldots, \xi_l(\varphi))$, $\varphi \in D_0$, — поле с однородными приращениями со средним значением $a(\varphi)$ и корреляционной матрицей $B(\varphi)$. Тогда существуют оператор из R^n в R^l — $l \times n$-матрица M, мера $F(d\lambda)$, определенная на борелевских множествах R^m, значениями которой служат неотрицательные $l \times l$-матрицы, $F(\{0\}) = 0$ и при некотором $p > 0$*

$$\int \frac{|\lambda|^2 \operatorname{Sp} F(d\lambda)}{(1+|\lambda|^2)^{p+1}} < \infty, \tag{5.12}$$

а также набор матриц A^{kj} порядка m $(k, j \leqslant l)$, удовлетворяющих условиям $A^{kj} = A^{jk}$, и $\sum \alpha_k \alpha_j A^{kj}$ для всех $\alpha_1, \ldots, \alpha_l \in \mathrm{R}$ — неотрицательная симметрическая матрица, такие, что

$$a(\varphi) = \int \varphi(x)\,(Mx)\,dx, \tag{5.13}$$

$$b_{ij}(\varphi) = \int |\tilde{\varphi}(\lambda)|^2 F_{ij}(d\lambda) + \iint (A^{ij}x, y)\,\varphi(x)\,\varphi(y)\,dx\,dy, \tag{5.14}$$

где $F_{ij}(d\lambda)$ — элемент матрицы $F(d\lambda)$. Матрицы M, A^{kj} $(k, j = 1, \ldots, l)$ и мера $F(d\lambda)$ определяются однозначно. При сформулированных предположениях формулы (5.13) и (5.14) определяют среднее значение и корреляционную матрицу обобщенного поля с однородными приращениями.

Спектральное представление корреляционной матрицы дает возможность получить и спектральное представление для самого обобщенного поля с однородными приращениями.

Теорема 4′. *В условиях теоремы 4 имеет место представление*

$$\xi_k(\varphi) = \int \tilde{\varphi}(\lambda)\,Z_k(d\lambda) + \int \varphi(x)\,(\eta_k, x)\,dx, \tag{5.15}$$

где $Z_1(d\lambda), Z_2(d\lambda), \ldots, Z_l(d\lambda)$ — случайные меры на R^m, для которых $Z_k(\{0\}) = 0$, $\mathsf{M}Z_k(A) = 0$, $\mathsf{M}Z_j(A)\overline{Z_k(B)} = F_{jk}(A \cap B)$; $\eta_1, \ldots, \eta_l$ — случайные векторы в R^m, для которых $\mathsf{M}Z_j(A)\,\eta_k = 0$ $(k, j = 1, \ldots, l)$; если $\eta_k = (\eta_k^1, \ldots, \eta_k^m)$, то $\mathsf{M}\eta_k^r = m_k^r$, где

$$\begin{bmatrix} a_1^1 & \ldots & a_1^m \\ \cdot & \cdot\ \cdot & \cdot \\ a_l^1 & \ldots & a_l^m \end{bmatrix} = M, \quad \mathsf{M}\eta_s^k \eta_r^j = a_s^k a_r^j = b_{sr}^{kj}, \quad \begin{bmatrix} b_{sr}^{11} & \ldots & b_{sr}^{1m} \\ b_{sr}^{m1} & \ldots & b_{sr}^{mm} \end{bmatrix} = A^{sr}.$$

Случайные меры $Z_1, \ldots, Z_l$ и векторы $\eta_1, \ldots, \eta_l$ однозначно определяются по векторному полю с однородными приращениями $\xi(\varphi)$.

12.5.4. Обобщенные однородные и изотропные случайные поля. Обозначим через G группу всех ортогональных преобразований $\mathbf{R}^m$. Для $g \in G$ через $g\varphi(x)$ будем обозначать $\varphi(g^{-1}x)$. Поле $\xi(\varphi)$ называется *изотропным*, если $\mathsf{M}\xi(\varphi) = \mathsf{M}\xi(g\varphi)$, $\mathsf{M}\xi^2(\varphi) = \mathsf{M}\xi^2(g\varphi)$. В этом пункте рассматриваются однородные обобщенные поля, являющиеся дополнительно изотропными.

Теорема 5. *Для того чтобы однородное случайное поле $\xi(\varphi)$ было изотропным, необходимо и достаточно, чтобы мера $F(d\lambda)$ в формуле (5.6) была инвариантна относительно ортогональных преобразований $\mathbf{R}^m$, т. е. чтобы для всякого борелевского множества $B \subset \mathbf{R}^m$ выполнялось равенство $F(g(B)) = F(B)$. Для того чтобы $b(\varphi_1, \varphi_2)$ была корреляционной функцией однородного изотропного поля в $\mathbf{R}^m$, необходимо и достаточно, чтобы она была представима в виде*

$$b(\varphi_1, \varphi_2) = \int_0^\infty \int_{\mathbf{R}^m} \int_{\mathbf{R}^m} \varphi_1(x) \overline{\varphi_2(y)}\, Y_m(|x-y|\lambda)\, dx\, dy\, d\Phi(\lambda), \qquad (5.16)$$

где $Y_n(t) = 2^{(n-2)/2}\Gamma(n/2)\, t^{-(n-2)/2} J_{(n-2)/2}(t)$, $J_{(n-2)/2}$ — функция Бесселя порядка $(n-2)/2$, а $\Phi(\lambda)$ — неубывающая функция на $[0, \infty)$, для которой

$$\int_0^\infty (1+\lambda^p)^{-1}\, d\Phi(\lambda) < \infty.$$

Непрерывная слева функция $\Phi(\lambda)$, для которой $\Phi(0+) = \Phi(0) = 0$, определяется однозначно.

Замечание. Теорема сформулирована так, чтобы она была справедливой и для комплекснозначного поля; определение изотропности здесь очевидно.

Рассмотрим теперь многомерный случай — l-мерное поле $\xi(\varphi) = (\xi_1(\varphi), \ldots, \xi_l(\varphi))$. Предположим, что каждому $g \in G$ соответствует ортогональное преобразование U_g пространства $\mathbf{R}^l$, так что $U_{g_1g_2} = U_{g_1}U_{g_2}$ (т. е. задано представление группы G ортогональными $l \times l$-матрицами). Будем говорить, что однородное l-мерное поле $\xi(\varphi)$ является *изотропным*, если

$$\mathsf{M}\xi(g\varphi) = \mathsf{M}U_g\xi(\varphi), \quad \mathsf{M}(\xi(g\varphi), z)^2 = \mathsf{M}(U_g\xi(\varphi), z)^2$$

для всех $g \in G$, $\varphi \in D$, $z \in \mathbf{R}^l$; $(\cdot, \cdot)$ — скалярное произведение в $\mathbf{R}^l$.

Теорема 6. *Пусть $U_g = I$ (единичная матрица) для всех $g \in G$. Тогда для того, чтобы однородное поле $\xi(\varphi)$ было изотропным, необходимо и достаточно, чтобы матричная мера $F(d\lambda)$ в формуле (5.9) была инвариантна относительно ортогональных преобразований $\mathbf{R}^m$, т. е. чтобы для всякого борелевского множества $B \subset \mathbf{R}^m$ $F(B) = F(g(B))$.*

Для того чтобы матрица $B(\varphi) = (b_{ij}(\varphi))$ *была корреляционной матрицей изотропного однородного случайного поля* $\xi(\varphi)$, *необходимо и достаточно, чтобы она допускала представление*

$$b_{ij}(\varphi) = \int_0^\infty \int_{R^m} \int_{R^m} \varphi(x)\,\bar{\varphi}(y)\,Y_n(|x-y|\lambda)\,dx\,dy\,\Phi_{ij}(d\lambda), \quad (5.17)$$

где $Y_n(t)$ *такое же, как и в формуле* (5.16), *а* $\Phi_{ij}(\lambda)$—*элементы матричной симметрической возрастающей функции* $\Phi(\lambda)$, *для которой*

$$\int_0^\infty (1+\lambda^p)^{-1}\,d\Phi_{ii}(\lambda) < \infty, \qquad i = 1, \ldots, m.$$

Если $\Phi(\lambda)$ *удовлетворяет перечисленным условиям, то* (5.17) *задает корреляционный оператор однородного изотропного векторного поля.*

Второй случай, который мы здесь рассмотрим, это $m = l$, $U_g = g$.

Теорема 7. *Пусть* $U_g = g$. *Тогда* $m(\varphi) = 0$, *а корреляционная матрица* $B(\varphi) = (b_{ij}(\varphi))$ *имеет вид*

$$b_{ij}(\varphi) = \int_0^\infty \int_{R^m} \int_{R^m} \varphi(x)\,\bar{\varphi}(y)\left\{Y_m^{(1)}(|x-y|\lambda)\,\chi_i\chi_j + \right.$$
$$\left. + Y_m^{(2)}(|x-y|\lambda)\,\delta_{ij}\right\} dx\,dy\,d\Phi_1(\lambda) +$$
$$+ \int_0^\infty \int_{R^m} \int_{R^m} \varphi(x)\,\bar{\varphi}(y)\left\{Y_m^{(3)}(|x-y|\lambda)\,\chi_i\chi_j + \right.$$
$$\left. + Y_m^{(4)}(|x-y|\lambda)\,\delta_{ij}\right\} dx\,dy\,d\Phi_2(\lambda), \quad (5.18)$$

где

$$\chi = \frac{x-y}{|x-y|} = (\chi_1, \ldots, \chi_m),$$
$$Y_n^{(1)}(t) = -2^{(n-2)/2}\Gamma(n/2)\,t^{-(n-2)/2}J_{(n+2)/2}(t),$$
$$Y_n^{(2)}(t) = 2^{(n-2)/2}\Gamma(n/2)\,t^{-n/2}J_{n/2}(t),$$
$$Y_n^{(3)}(t) = 2^{(n-2)/2}\Gamma(n/2)\,t^{-(n-2)/2}J_{(n+2)/2}(t),$$
$$Y_n^{(4)}(t) = 2^{(n-2)/2}\Gamma(n/2)\left[t^{-(n-2)/2}J_{(n-2)/2}(t) - t^{-n/2}J_{n/2}(t)\right],$$

а $\Phi_1(\lambda)$, $\Phi_2(\lambda)$ *не убывают на* $[0, \infty)$, $\Phi_1(0) = \Phi_2(0)$, $\Phi_1(+0) = \Phi_2(+0)$ *и при некотором* $p > 0$

$$\int_0^\infty (1+\lambda^p)^{-1}\,d\Phi_i(\lambda) < \infty, \qquad i = 1, 2.$$

Наоборот, если $\Phi_1(\lambda)$ и $\Phi_2(\lambda)$ удовлетворяют этим условиям, то формула (5.18) *определяет корреляционный оператор векторного однородного и изотропного обобщенного поля.*

Литература: [16, 104, 105].

Глава 13. МАРТИНГАЛЫ

13.1. Определения. Общие свойства

Мартингалы и полумартингалы образуют важный класс процессов, обобщающий процессы с независимыми приращениями. Он включает в себя широкий класс процессов — марковские процессы, решения стохастических уравнений, управляемые случайные процессы и др. Разработана специальная «мартингальная» техника изучения случайных процессов. Основы ее излагаются здесь.

13.1.1. Определения. Пусть $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ — полное вероятностное пространство, на котором задан поток σ-алгебр $(\mathfrak{F}_t)_{t \in T}$, где T — некоторое подмножество вещественной прямой $\mathbf{R}$. Это означает, что $\mathfrak{F}_t$ монотонно зависит от t: при $t_1 < t_2$, $t_1, t_2 \in T$ будет $\mathfrak{F}_{t_1} \subset \mathfrak{F}_{t_2} \subset \mathfrak{S}$. (В гл. 9 дано несколько другое определение потока, мы будем им пользоваться при рассмотрении мартингалов с непрерывным временем.)

Семейство вещественных величин $\xi(t)$ $(t \in T)$ называется *мартингалом*, если $\forall t \in T$ $\mathsf{M}|\xi(t)| < \infty$ и

$$\mathsf{M}(\xi(t_2) \mid \mathfrak{F}_{t_1}) = \xi(t_1), \quad t_1 < t_2, \quad t_1, \quad t_2 \in T. \tag{1.1}$$

Семейство случайных величин $\xi(t)$ $(t \in T)$ называется *супермартингалом*, если $\forall t$ $\mathsf{M}\xi^-(t) = \frac{1}{2}\mathsf{M}(|\xi(t)| - \xi(t)) < \infty$ и

$$\mathsf{M}(\xi(t_2) \mid \mathfrak{F}_{t_1}) \leqslant \xi(t_1), \quad t_1 < t_2, \quad t_1, \quad t_2 \in T. \tag{1.2}$$

Семейство случайных величин $\xi(t)$ $(t \in T)$ называется *субмартингалом*, если $-\xi(t)$ является супермартингалом.

Мартингал является одновременно и супермартингалом, и субмартингалом; верно и обратное: если $\xi(t)$ $(t \in T)$ является супермартингалом и субмартингалом одновременно, то $\xi(t)$ — мартингал. Термины «мартингал», «субмартингал», «супермартингал» относятся также и к случайным процессам. Более точно, называя процесс мартингалом (супермартингалом), следует указывать поток, относительно которого свойства (1.1) или (1.2) выполнены. Поэтому используется также термин «$(\mathfrak{F}_t)_{t \in T}$-мартингал» («супермартингал»). Иногда термин «мартингал» относят к процессам, для которых (1.1) выполнено для потока, порождаемого самим процессом $\xi(t)$, т. е. когда $\mathfrak{F}_t = \sigma(\xi(s), s \leqslant t, s \in T)$ — наименьшая σ-алгебра, относительно которой измеримы все указанные в скобках величины.

В дальнейшем относительно σ-алгебр $\mathfrak{F}_t$ будет предполагаться, что они содержат все множества P-меры нуль. Тогда всякая модификация мартингала (супермартингала) также будет мартингалом (супермартингалом) относительно того же потока.

Супермартингалы и субмартингалы относятся к классу полумартингалов (или семимартингалов); общее содержательное определение полумартингала будет дано ниже в случае непрерывного времени.

Вместо свойств (1.1), (1.2) и аналогичного свойства субмартингала можно использовать интегральные неравенства:

$$\int_A \xi(t_2)\,\mathsf{P}(d\omega) \leqslant \int_A \xi(t_1)\,\mathsf{P}(d\omega) \quad \text{(супермартингал)},$$

$$\int_A \xi(t_2)\,\mathsf{P}(d\omega) = \int_A \xi(t_1)\,\mathsf{P}(d\omega) \quad \text{(мартингал)}, \tag{1.3}$$

$$\int_A \xi(t_2)\,\mathsf{P}(d\omega) \geqslant \int_A \xi(t_1)\,\mathsf{P}(d\omega) \quad \text{(субмартингал)},$$

$A \in \mathfrak{F}_{t_1}$, $t_1 < t_2$, $t_1, t_2 \in T$.

13.1.2. Примеры. 1. Пусть $\{\Omega, \mathfrak{S}, \mathsf{P}\}$ — произвольное вероятностное пространство, $(\mathfrak{F}_t)_{t \in T}$ — поток σ-алгебр на этом пространстве, η — произвольная вещественная величина, для которой $\mathsf{M}\,|\,\eta\,| < \infty$. Тогда процесс

$$\xi(t) = \mathsf{M}(\eta\,|\,\mathfrak{F}_t)$$

является мартингалом.

2 Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых случайных величин, для которых $\mathsf{M}\,|\,\xi_k\,| < \infty$ $(k = 1, 2, \ldots)$, $\zeta_k \doteq \xi_1 + \ldots + \xi_k$, $\mathfrak{F}_k = \sigma(\xi_1, \ldots, \xi_k)$; $\{\zeta_t, t \in T\}$ $(T = \{1, 2, \ldots\})$ является мартингалом тогда и только тогда, когда $\mathsf{M}\xi_k = 0$ $(k = 1, 2, \ldots)$, и супермартингалом, если $\mathsf{M}\xi_k \leqslant 0$ $(k = 1, 2, \ldots)$.

3. Если в условиях предыдущего примера $\mathsf{M}\xi_k = 0$ и $\mathsf{D}\xi_k < \infty$, то $\left\{\zeta_t^2, t \in T\right\}$ является субмартингалом.

4. Если $\xi(t)$ $(t \geqslant 0)$ — процесс с независимыми приращениями, то он будет мартингалом при условии, что $\mathsf{M}\xi(t)$ $(t \geqslant 0)$ существует и постоянно.

5. Пусть $\xi(t)$ — однородный процесс с независимыми приращениями, для которого при некотором λ $\mathsf{M}e^{\lambda\xi(1)} = a(\lambda) < \infty$. Тогда $\exp\{\lambda\xi(t) - t\ln a(\lambda)\}$ является мартингалом.

6. $\eta_1, \eta_2, \ldots$ — однородная цепь Маркова в произвольном фазовом пространстве X, $h(x)$ — эксцессивная функция для этой цепи, т. е. функция, удовлетворяющая неравенству

$$h(x) \leqslant \int_X h(y)\,P(x, dy),$$

где $P(x, dy)$ — вероятности перехода для цепи. Тогда процесс $\xi_n = h(\eta_n)$ является супермартингалом относительно потока $\{\mathfrak{F}_k, k \geqslant 1\}$, где $\mathfrak{F}_k = \sigma(\eta_1, \ldots, \eta_k)$.

7. Пусть $\eta(t)$ $(t \in T)$ — марковский процесс с переходной вероятностью $P(s, x, t, A)$. Если T содержит максимальный элемент t_m, то процесс

$$\xi(t) = P(t, \eta(t), t_m, A)$$

является мартингалом относительно потока $\mathfrak{F}_t = \sigma(\eta(s),\ s \leqslant t,\ s \in T)$.

8. Пусть $\{\zeta_k,\ k \geqslant 1\}$ — последовательность вещественных случайных величин такая, что для всех n существует совместная плотность распределения $p_n(x_1, \ldots, x_n)$ величин $\zeta_1, \ldots, \zeta_n$. Если $q_n(x_1, \ldots, x_n)$ — произвольная последовательность функций, для которых при $n \geqslant 1$

$$\int |q_n(x_1, \ldots, x_n)|\, dx_1 \ldots dx_n < \infty, \quad \int q_n(x_1, \ldots, x_n)\, dx_n = q_{n-1}(x_1, \ldots, x_{n-1}),$$

то последовательность

$$\xi_k = \frac{q_k(\zeta_1, \ldots, \zeta_k)}{p_k(\zeta_1, \ldots, \zeta_k)}$$

является мартингалом относительно потока $\mathfrak{F}_k = \sigma(\zeta_1, \ldots, \zeta_k)$.

В частности, q_k могут быть k-мерными распределениями некоторой последовательности случайных величин, и тогда ξ_k есть отношение правдоподобия.

9. Пусть $w(t)$ $(t \geqslant 0)$ — винеровский процесс в $\mathbf{R}^m$, $\mathfrak{F}_t = \sigma(w(s),\ s \leqslant t)$ и $\varphi(x)$ — вещественная непрерывная функция на $\mathbf{R}^m$. Процесс $\varphi(w(t))$ будет мартингалом относительно потока $(\mathfrak{F}_t)_{t \geqslant 0}$, если $\varphi(x)$ — гармоническая функция; он будет супермартингалом, если $\varphi(x)$ — супергармоническая функция, и субмартингалом, если $\varphi(x)$ — субгармоническая функция. Этот пример указывает на связь теории мартингалов с классической теорией потенциала.

13.1.3. Общие свойства. 1) Функция $\mathsf{M}\xi(t)$ $(t \in T)$: а) постоянна, если $\xi(t)$ — мартингал, б) не возрастает, если $\xi(t)$ — супермартингал, и в) не убывает, если $\xi(t)$ — субмартингал.

2) Если $\xi(t)$ — супермартингал (субмартингал), то для того, чтобы $\xi(t)$ был мартингалом, необходимо и достаточно, чтобы $\mathsf{M}\xi(t)$ было постоянным на области определения.

3) Если $\xi_1(t)$ и $\xi_2(t)$ — два мартингала (супермартингала) относительно одного и того же потока σ-алгебр $(\mathfrak{F}_t)_{t \in T}$, то $a\xi_1(t) + b\xi_2(t)$ — также мартингал $(a, b \in \mathbf{R})$ (супермартингал $(a, b \in \mathbf{R}_+)$).

4) Если $\xi_1(t)$ и $\xi_2(t)$ — два супермартингала относительно одного и того же потока, то $\xi_1(t) \wedge \xi_2(t)$ — также супермартингал относительно того же потока; в частности, для всех $c \in \mathbf{R}$ $\xi_1(t) \wedge c$ является супермартингалом, если $\xi_1(t)$ — супермартингал.

5) Если $\xi_1(t)$ и $\xi_2(t)$ — два субмартингала относительно одного и того же потока, то $\xi_1(t) \vee \xi_2(t)$ — субмартингал относительно того же потока; в частности, для всех $c \in \mathbf{R}$ $\xi_1(t) \vee c$ есть субмартингал, если только $\xi_1(t)$ — субмартингал.

6) Пусть $\xi(t)$ — мартингал и f — вещественная, определенная на $\mathbf{R}$ выпуклая вниз функция. Тогда процесс $f(\xi(t))$ является субмартингалом, если $\mathsf{M}(|f(\xi(t))| + f(\xi(t))) < \infty$ $(t \in T)$. В частности, субмартингалами будут $|\xi(t)|$, $|\xi(t)|^p$ $(p > 1)$, $(\xi(t) - c)^+$ (здесь $(x)^+ = (|x| + x)/2$) при условии существования математических ожиданий этих величин.

7) Пусть $\xi(t)$ — супермартингал и $f(x)$: $\mathbf{R} \to \mathbf{R}$ выпукла вниз и не возрастает. Тогда $f(\xi(t))$ — супермартингал, если только $\mathsf{M}(|f(\xi(t))| - f(\xi(t))) < \infty$ для всех $t \in T$

Приведем несколько свойств, относящихся к возможным изменениям потока σ-алгебр, при которых мартингалы остаются мартингалами.

8) Пусть $\xi(t)$ $(t \in T)$ является мартингалом (супермартингалом) относительно потока $(\mathfrak{F}_t)_{t \in T}$. Если $(\tilde{\mathfrak{F}}_t)_{t \in T}$ — такой поток, что $\tilde{\mathfrak{F}}_t \subset \mathfrak{F}_t$ $\forall t \in T$ и $\xi(t)$ $\tilde{\mathfrak{F}}_t$-измерим, то $\xi(t)$ будет мартингалом (супермартингалом) относительно потока $(\tilde{\mathfrak{F}}_t)_{t \in T}$.

9) Пусть $\xi(t)$ $(t \in T)$ является мартингалом (супермартингалом) относительно потока $(\mathfrak{F}_t)_{t \in T}$ и $(\tilde{\mathfrak{F}}_t)_{t \in T}$ — такой поток, что $\tilde{\mathfrak{F}}_t \supset \mathfrak{F}_t$ и $\tilde{\mathfrak{F}}_t \subset \mathfrak{F}_t \vee \mathfrak{S}_0$, где $\mathfrak{S}_0$ — σ-алгебра, порожденная множествами P-меры 0. Тогда $\xi(t)$ будет мартингалом (супермартингалом) относительно потока $(\tilde{\mathfrak{F}}_t)_{t \in T}$.

10) Пусть T — интервал прямой, $\xi(t)$ $(t \in T)$ — стохастически непрерывный мартингал (супермартингал) относительно потока $(\mathfrak{F}_t)_{t \in T}$. Тогда $\xi(t)$ будет мартингалом (супермартингалом) относительно потока $(\mathfrak{F}_{t+})_{t \in T}$, где $\mathfrak{F}_{t+} = \bigcap\limits_{s \in T,\, s > t} \mathfrak{F}_s$.

Свойства 9) и 10) показывают, что при рассмотрении мартингалов непрерывного аргумента естественно требовать от потока, чтобы он был непрерывен справа и пополнен множествами P-меры 0 (именно эти требования входили в определение потока в § 9.4).

13.2. Мартингалы с дискретным параметром

В этом параграфе рассматриваются процессы $\xi(t)$ $(t \in T)$, когда T — одно из трех множеств: $\mathbf{Z}$ — все целые числа, $\mathbf{Z}_+$ — неотрицательные целые числа, $\mathbf{Z}_-$ — неположительные целые числа. Относительно потока σ-алгебр $(\mathfrak{F}_t)_{t \in T}$ здесь не требуется никаких дополнительных предположений (полнота, непрерывность справа).

13.2.1. Моменты остановки. Моментом остановки относительно потока $(\mathfrak{F}_t)_{t \in T}$ называется такая случайная величина τ со значениями в T (т. е. $\mathbf{Z}$, $\mathbf{Z}_+$ или $\mathbf{Z}_-$), что событие $\{\tau = t\} \in \mathfrak{F}_t$ для всех $t \in T$. Через $\mathfrak{F}_\tau$ обозначается σ-алгебра множеств A, для которых $A \cap \{\tau = t\} \in \mathfrak{F}_t$ $\forall t \in T$.

Последовательность ζ_k, определенная при $k \in \mathbf{Z}_+$, называется *предсказуемой*, если для всех k она $\mathfrak{F}_{k-1}$-измерима, при этом ζ_0 должно быть постоянно.

Справедливы следующие утверждения.

1. *Пусть ζ_k — предсказуемая последовательность, ξ_k — мартингал; положим*

$$\eta_n = \zeta_0\xi_0 + (\xi_1 - \xi_0)\zeta_1 + \ldots + (\xi_n - \xi_{n-1})\zeta_n.$$

Тогда $(\eta_n,\ n \geqslant 0)$ является мартингалом, если только $\mathsf{M}|\eta_n| < \infty$ для всех $n \geqslant 0$.

2. *Пусть ξ_k $(k \geqslant 0)$ — мартингал (супермартингал) и τ — момент остановки, для которого $\mathsf{P}\{\tau \leqslant N\} = 1$ при некотором N. Тогда $\mathsf{M}\xi_\tau = \mathsf{M}\xi_0$ $(\mathsf{M}\xi_\tau \leqslant \mathsf{M}\xi_0)$.*

3. *Пусть ξ_k $(k \geqslant 0)$ — мартингал (супермартингал), τ и σ — моменты остановки, для которых $\mathsf{P}\{\sigma \leqslant \tau \leqslant N\} = 1$ при некотором N. Тогда $\mathsf{M}(\xi_\tau \mid \mathfrak{F}_\sigma) = \xi_\sigma$ $(\mathsf{M}(\xi_\tau \mid \mathfrak{F}_\sigma) \leqslant \xi_\sigma)$.*

4. *Пусть* ξ_k $(k \geqslant 0)$ — *супермартингал и* τ — *момент остановки, для которого* $\mathsf{P}\{0 \leqslant \tau \leqslant N\} = 1$. *Тогда*

$$\mathsf{M}\xi_0 \geqslant \mathsf{M}\xi_\tau \geqslant \mathsf{M}\xi_N,$$

$$\mathsf{M}\,|\xi_\tau| \leqslant \mathsf{M}\,|\xi_0| + 2\mathsf{M}\,(\xi_N)^- \leqslant 3 \sup_{n \leqslant N} \mathsf{M}\,|\xi_n|.$$

5. *Пусть* σ — *некоторый конечный момент остановки, такой, что для мартингала* ξ_k $(k \geqslant 0)$ $\mathsf{M}\,|\xi_\sigma| < \infty$. *Тогда совокупность случайных величин* $\{\xi_\tau,\ \tau \leqslant \sigma\}$, *где* τ — *моменты остановки, равномерно интегрируема* *). *При этом* $\xi_\tau = \mathsf{M}\,(\xi_\sigma \mid \mathfrak{F}_\tau)$ *для всех моментов остановки* $\tau \leqslant \sigma$.

6. *Пусть* ξ_k $(k \geqslant 0)$ — *положительный супермартингал. Тогда* $\mathsf{M}\,|\xi_\sigma| < \infty$ *для всякого конечного момента остановки* σ. *Если* σ *и* τ — *два конечных момента остановки, для которых* $\sigma \leqslant \tau$, *то* $\mathsf{M}\,(\xi_\tau \mid \mathfrak{F}_\sigma) \leqslant \xi_\sigma$.

Далее рассмотрим случай, когда момент остановки может принимать значение $+\infty$. Для этого понадобится понятие замыкания мартингала (супермартингала) справа.

Мартингал (супермартингал) ξ_k $(k \geqslant 0)$ называется *замкнутым справа*, если существует такая величина ξ_∞, что $\xi_k = \mathsf{M}\,(\xi_\infty \mid \mathfrak{F}_k)$ $(\mathsf{M}\,(\xi_\infty)^- < +\infty$ и $\mathsf{M}\,(\xi_\infty \mid \mathfrak{F}_k) \leqslant \xi_k)$.

7. *Если* ξ_k — *замкнутый справа мартингал, положим* $\xi_\tau = \xi_\infty$ *при* $\tau = +\infty$, *где* ξ_∞ — *величина, входящая в определение замыкания мартингала. Тогда совокупность случайных величин* $\{\xi_\tau,\ \tau \leqslant \infty\}$, *где* τ — *всевозможные моменты остановки, равномерно интегрируема и для всех моментов остановки* τ

$$\xi_\tau = \mathsf{M}\,(\xi_\infty \mid \mathfrak{F}_\tau)$$

$(\mathfrak{F}_\infty = \bigvee_k \mathfrak{F}_k \vee \mathfrak{S}(\xi_\infty)$, *где* $\mathfrak{S}(\xi_\infty)$ — σ-*алгебра, порожденная величиной* ξ_∞; $\mathfrak{F}_\tau$ *содержит все события* A *из* $\mathfrak{F}_\infty$, *для которых* $A \cap \{\tau \leqslant t\} \in \mathfrak{F}_t$ *при* $t < \infty$). *Если* $0 \leqslant \sigma \leqslant \tau \leqslant +\infty$ — *два момента остановки, то*

$$\mathsf{P}\{\xi_\tau \mid \mathfrak{F}_\sigma\} = \xi_\sigma.$$

8. *Если* ξ_k — *замкнутый справа супермартингал, то, используя те же соглашения, что и в предыдущем утверждении, имеем* $\mathsf{M}\,|\xi_\tau| < \infty$ *для всех моментов остановки* τ, *при этом*

$$\xi_\tau \geqslant \mathsf{M}\,(\xi_\infty \mid \mathfrak{F}_\tau),$$

и если $0 \leqslant \sigma \leqslant \tau \leqslant \infty$ — *два момента остановки, то*

$$\xi_\sigma \geqslant \mathsf{M}\,(\xi_\tau \mid \mathfrak{F}_\sigma).$$

9. *Пусть* ξ_k $(k \geqslant 0)$ — *замкнутый справа мартингал (супермартингал), а* τ_n $(n \geqslant 0)$ — *произвольная неубывающая последовательность моментов остановки. Тогда последовательность* $\hat{\xi}_n = \xi_{\tau_n}$

---

*) Множество случайных величин $\{\xi_\alpha\}$ называется *равномерно интегрируемым*, если $\lim\limits_{\delta \to 0} \sup\limits_{\mathsf{P}(A) < \delta} \sup\limits_{\alpha} \int\limits_A |\xi_\alpha|\, d\mathsf{P} = 0$.

($n \geqslant 0$) является мартингалом (супермартингалом) относительно потока $\{\hat{\mathfrak{F}}_n\}_{n \geqslant 0}$, *где* $\hat{\mathfrak{F}}_n = \hat{\mathfrak{F}}_{\tau_n}$.

10. *Пусть* ξ_k ($k \geqslant 0$) — *замкнутый справа мартингал (супермартингал),* τ — *момент остановки,* $\tau \leqslant \infty$. *Тогда* $\hat{\xi}_n = \xi_{n \wedge \tau}$ ($n \geqslant 0$) *является мартингалом (супермартингалом).*

13.2.2. Основные неравенства. Мартингальные неравенства делятся на две группы: неравенства, относящиеся к оценке распределения максимума или супремума процесса, и неравенства для распределения числа пересечений процессом полосы.

Рассмотрим процесс ξ_t ($t \in \mathbf{Z}_+$).

1. Пусть ξ_t ($t \in \mathbf{Z}_+$) — супермартингал, $\lambda > 0$. Тогда

а) $$\lambda \mathsf{P}\left\{\sup_{t \leqslant k} \xi_t \geqslant \lambda\right\} \leqslant \mathsf{M}\xi_0 - \int\limits_{\{\sup_{t \leqslant k} \xi_t < \lambda\}} \xi_k \, d\mathsf{P} \leqslant \mathsf{M}\xi_0 + \frac{1}{2}\mathsf{M}[|\xi_k| - \xi_k];$$

б) $$\lambda \mathsf{P}\left\{\inf_{t \leqslant k} \xi_t \leqslant -\lambda\right\} \leqslant -\int\limits_{\{\inf_{t \leqslant k} \xi_t \leqslant -\lambda\}} \xi_k \, d\mathsf{P} \leqslant \frac{1}{2}\mathsf{M}[|\xi_k| - \xi_k];$$

в) $$\lambda \mathsf{P}\left\{\sup_{t \leqslant k} |\xi_t| \geqslant \lambda\right\} \leqslant 3 \sup_{t \leqslant k} \mathsf{M}|\xi_t|.$$

Эта группа неравенств иногда еще называется максимальной леммой.

2. Пусть ξ_t ($t \in \mathbf{Z}_+$) — мартингал, $\lambda > 0$. Тогда

а) $$\lambda \mathsf{P}\left\{\sup_{t \leqslant k} \xi_t \geqslant \lambda\right\} \leqslant \frac{1}{2}\mathsf{M}[|\xi_k| + \xi_k];$$

б) $$\lambda \mathsf{P}\left\{\inf_{t \leqslant k} \xi_t \leqslant -\lambda\right\} \leqslant \frac{1}{2}\mathsf{M}[|\xi_k| - \xi_k];$$

в) $$\lambda \mathsf{P}\left\{\sup_{t \leqslant k} |\xi_t| \geqslant \lambda\right\} \leqslant \mathsf{M}|\xi_k|;$$

г) для $\alpha > 1$ $$\lambda^\alpha \mathsf{P}\left\{\sup_{t \leqslant k} |\xi_t| \geqslant \lambda\right\} \leqslant \mathsf{M}|\xi_k|^\alpha, \quad \lambda > 0.$$

3. Пусть ξ_k ($t \in \mathbf{Z}_+$) — супермартингал, $\xi_t \geqslant 0$. Тогда

а) $$\lambda \mathsf{P}\left\{\sup_{t \leqslant k} \xi_t \geqslant \lambda\right\} = \lambda \mathsf{P}\left\{\sup_{t \leqslant k} |\xi_t| \geqslant \lambda\right\} \leqslant \mathsf{M}\xi_0;$$

б) для $\alpha > 1$ $$\lambda^\alpha \mathsf{P}\left\{\sup_{t \leqslant k} \xi_t \geqslant \lambda\right\} = \lambda^\alpha \mathsf{P}\left\{\sup_{t \leqslant k} |\xi_t| > \lambda\right\} \leqslant \sup_{t \leqslant k} \mathsf{M}|\xi_t|^\alpha.$$

4. Пусть ξ_t ($t \in \mathbf{Z}_+$) — мартингал, $\xi_t \geqslant 0$. Тогда

а) $$\mathsf{M} \sup_{t \geqslant 0} |\xi_t| \leqslant 2\left(1 + \sup_n \mathsf{M}|\xi_n| \ln(|\xi_n| \vee 1)\right);$$

б) для $\alpha > 1$ $$\mathsf{M}\left(\sup_{t \geqslant 0} |\xi_t|\right)^\alpha \leqslant \left(\frac{\alpha}{\alpha - 1}\right)^\alpha \sup_{n \geqslant 0} \mathsf{M}|\xi_n|^\alpha.$$

5. Пусть ξ_t $(t \in \mathbf{Z}_+)$ — субмартингал, $\xi_t \geqslant 0$. Тогда

а) $\mathsf{M} \sup\limits_{t>0} \xi_t \leqslant 2\left(1 + \sup\limits_n \mathsf{M}\xi_n \ln(|\xi_n| \vee 1)\right)$;

б) для $\alpha > 1$ $\mathsf{M} \sup\limits_{t \geqslant 0} \xi_t^\alpha \leqslant \left(\dfrac{\alpha}{\alpha-1}\right)^\alpha \sup\limits_{n \geqslant 0} \mathsf{M}\xi_n^\alpha$.

6. Пусть ξ_n $(n \geqslant 0)$ — мартингал, v_n $(n \geqslant 0)$ — предсказуемая последовательность, $|v_n| \leqslant 1$,

$$\eta_n = v_0\xi_0 + v_1(\xi_1 - \xi_0) + \ldots + v_n(\xi_n - \xi_{n-1}).$$

Тогда

а) $\lambda \mathsf{P}\left\{\sup\limits_n |\eta_n| \geqslant \lambda\right\} \leqslant 18 \sup\limits_n \mathsf{M}|\xi_n|$;

б) если $\xi_n \geqslant 0$, то

$$\lambda \mathsf{P}\left\{\sup_n |\eta_n| \geqslant \lambda\right\} \leqslant 9 \sup \mathsf{M}|\xi_n|;$$

в) для всякого $\alpha > 1$ существует такая постоянная c_α, что

$$\mathsf{M} \sup_n |\eta_n|^\alpha \leqslant c_\alpha \sup_n \mathsf{M}|\xi_n|^\alpha$$

(это *неравенство Буркхольдера*).

Рассмотрим теперь супермартингалы (мартингалы), определенные на $\mathbf{Z}_-$ $(\mathbf{Z})$. Для них справедливы аналогичные неравенства; приведем лишь основные из них.

7. Пусть ξ_t — супермартингал, $t \in \mathbf{Z}_-$ $(t \in \mathbf{Z})$. Тогда

а) $\lambda \mathsf{P}\left\{\sup\limits_{-k \leqslant t \leqslant 0} |\xi_t| \geqslant \lambda\right\} \leqslant 3 \sup\limits_{-k \leqslant t \leqslant 0} \mathsf{M}|\xi_t|$;

б) $\lambda \mathsf{P}\left\{\sup\limits_{|t| \leqslant k} |\xi_t| \geqslant \lambda\right\} \leqslant 3 \sup\limits_{|t| \leqslant k} \mathsf{M}|\xi_t|$;

в) если $\xi_t \geqslant 0$, $t \in \mathbf{Z}_-$, то

$$\lambda \mathsf{P}\left\{\sup_{-k \leqslant t \leqslant 0} \xi_t \geqslant \lambda\right\} \leqslant \mathsf{M}\xi_{-k};$$

для $\alpha > 1$ $\lambda^\alpha \mathsf{P}\left\{\sup\limits_{-k \leqslant t \leqslant 0} \xi_t \geqslant \lambda\right\} \leqslant \sup\limits_{-k \leqslant t \leqslant 0} \mathsf{M}|\xi_t|^\alpha$;

г) если $\xi_t \geqslant 0$ $(t \in \mathbf{Z})$, то

$$\lambda \mathsf{P}\left\{\sup_{|t| \leqslant k} \xi_t \geqslant \lambda\right\} \leqslant \mathsf{M}\xi_{-k},$$

для $\alpha > 1$ $\lambda^\alpha \mathsf{P}\left\{\sup\limits_{|t| \leqslant k} \xi_t \geqslant \lambda\right\} \leqslant \sup\limits_{|t| \leqslant k} |\xi_t|^\alpha$.

8. Пусть ξ_t — субмартингал, $\xi_t \geqslant 0$. Тогда

а) $\mathsf{M} \sup \xi_t \leqslant 2(1 + \sup \mathsf{M}\xi_t \ln(|\xi_t| \vee 1))$;

б) при $\alpha > 1$ $\mathsf{M} \sup \xi_t^\alpha \leqslant \left(\dfrac{\alpha}{\alpha-1}\right)^\alpha \sup \mathsf{M}\xi_t^\alpha$.

Здесь супремум берется либо по $\mathbf{Z}_-$, либо по $\mathbf{Z}$ в зависимости от области определения субмартингала.

9. Если ξ_t — мартингал, то неравенства 8 остаются справедливыми, если в них заменить ξ_t на $|\xi_t|$.

Рассмотрим теперь неравенства, относящиеся к числу пересечений процессом полосы. Пусть a_t ($t \in \mathbf{Z}$) — некоторая последовательность вещественных чисел, $V \subset \mathbf{Z}$ — отрезок целых чисел, $\alpha < \beta \in \mathbf{R}$. Говорят, что последовательность a_t пересекает полосу $[\alpha, \beta]$ на отрезке V сверху вниз m раз, если можно указать такие точки $t_1 < t_2 < \ldots < t_{2m}$ из отрезка V, что $a_{t_1} \geqslant \beta$, $a_{t_2} \leqslant \alpha$, ..., $a_{t_{2m-1}} \geqslant \beta$, $a_{t_{2m}} \leqslant \alpha$, и нельзя указать $2m+2$ точки, чтобы для них выполнялись аналогичные неравенства. В этом случае m называется *числом пересечений* сверху вниз полосы $[\alpha, \beta]$ на отрезке V. Аналогично определяется число пересечений снизу вверх. Пусть ξ_t — случайный процесс, заданный на $\mathbf{Z}_+$, $\mathbf{Z}_-$ или $\mathbf{Z}$. Обозначим через $\nu_V[\alpha, \beta]$ число пересечений им снизу вверх полосы $[\alpha, \beta]$ на отрезке V.

10. *Неравенства Дуба*:

а) $$(\beta-\alpha)\mathsf{M}\nu_V[\alpha, \beta] \leqslant \frac{1}{2}\mathsf{M}(|\xi_k-\alpha|-(\xi_k-\alpha)) \leqslant \mathsf{M}|\xi_k-\alpha|,$$

где $k = \sup[t, t \in V]$, если V ограничено сверху;

б) $$(\beta-\alpha)\mathsf{M}\nu_{\mathbf{Z}_+}[\alpha, \beta] \leqslant \sup_{k\in\mathbf{Z}_+} \frac{1}{2}\mathsf{M}(|\xi_k-\alpha|-(\xi_k-\alpha)) \leqslant \sup_{k\in\mathbf{Z}_+} \mathsf{M}|\xi_k-\alpha|;$$

аналогичные неравенства справедливы для $\mathbf{Z}_-$ и $\mathbf{Z}$.

11. Если V ограничено сверху и $k = \sup[t, t \in V]$, то

$$(\beta-\alpha)\mathsf{P}\{\nu_V[\alpha, \beta] > n\} \leqslant \int\limits_{\{\nu_V[\alpha, \beta]=n\}} (|\xi_k-\alpha|-(\xi_k-\alpha))\,d\mathsf{P}$$

(это *неравенство Дубинса*).

Для субмартингалов можно получить аналогичные неравенства для пересечений сверху вниз, воспользовавшись тем, что $-\xi_t$ является супермартингалом.

13.2.3. Теоремы сходимости. Одним из важнейших свойств супермартингалов является существование предела при весьма общих предположениях об их ограниченности.

Теорема 1. *Пусть* ξ_n ($n \geqslant 0$) *является супермартингалом, для которого* $\sup\limits_n \mathsf{M}|\xi_n| < \infty$. *Тогда с вероятностью* 1 *существует* $\lim\limits_{n\to\infty} \xi_n$.

Аналогичное утверждение справедливо для супермартингала ξ_n ($n \leqslant 0$).

Теорема 2. *Пусть* ξ_n ($n \geqslant 0$) *является супермартингалом. Тогда*

а) *если последовательность* $\xi_n^- = (|\xi_n| - \xi_n)/2$ ($n \geqslant 0$) *равномерно интегрируема, то предел* $\xi = \lim\limits_{n\to\infty} \xi_n$ *существует и замыкает супермартингал справа*;

б) *наоборот, если супермартингал замкнут справа (т. е. существует такая случайная величина* ξ_∞, *что* $\xi_n \geqslant \mathsf{M}(\xi_\infty | \mathfrak{F}_n)$), *то* ξ_n^- *равномерно интегрируемы и, значит, существует с вероятностью* 1 *предел* $\xi = \lim \xi_n$.

Теорема 3. *Пусть* ξ_n $(n \leqslant 0)$ — *супермартингал относительно потока* $(\mathfrak{F}_n)_{n\leqslant 0}$, *для которого* $\lim\limits_{n\to-\infty} \mathsf{M}\xi_n^- < +\infty$. *Тогда*

а) *последовательность* ξ_n *равномерно интегрируема;*

б) *с вероятностью* 1 *существует* $\lim\limits_{n\to-\infty} \xi_n = \xi$;

в) $\lim\limits_{n\to-\infty} \mathsf{M}\,|\xi_n - \xi\,| = 0$;

г) ξ *замыкает* ξ_n *слева, т. е.* $\xi \geqslant \mathsf{M}(\xi_n\,|\,\mathfrak{F}_{-\infty})$, *где* $\mathfrak{F}_{-\infty} = \bigwedge\limits_n \mathfrak{F}_n$.

Теорема 4. *Пусть* ξ_n $(n \geqslant 0)$ — *мартингал. Тогда*

а) *если* ξ_n *равномерно интегрируем, то с вероятностью* 1 *существует* $\lim\limits_{n\to\infty} \xi_n = \xi$, *при этом* $\lim\limits_{n\to\infty} \mathsf{M}\,|\xi_n - \xi\,| = 0$ *и* $\xi_n = \mathsf{M}(\xi\,|\,\mathfrak{F}_n)$, *т. е.* ξ *замыкает мартингал* ξ_n *справа;*

б) *наоборот, если мартингал* ξ_n *замкнут справа (т. е.* $\xi_n = \mathsf{M}(\eta\,|\,\mathfrak{F}_n)$, *где* η — *некоторая величина с* $\mathsf{M}\,|\eta\,| < \infty$), *то* ξ_n *равномерно интегрируем, существует* $\lim\limits_{n\to\infty} \xi_n = \mathsf{M}(\eta\,|\,\bigvee\limits_n \mathfrak{F}_n) = \xi$ *и* $\lim\limits_{n\to\infty} \mathsf{M}\,|\xi_n - \xi\,| = 0$.

Теорема 5. *Пусть* ξ_n $(n \leqslant 0)$ — *мартингал относительно потока* $(\mathfrak{F}_n)_{n\leqslant 0}$. *Тогда* ξ_n *равномерно интегрируем, существует* $\lim\limits_{n\to-\infty} \xi_n = \xi_{-\infty}$, *при этом* $\lim\limits_{n\to-\infty} \mathsf{M}\,|\xi_n - \xi_{-\infty}\,| = 0$ *и* $\xi_{-\infty}$ *замыкает мартингал слева, т. е.* $\xi_{-\infty} = \mathsf{M}(\xi_n\,|\,\bigwedge\limits_k \mathfrak{F}_k)$. *Если, кроме того,* $\mathsf{M}(\xi_0)^\alpha < \infty$ $(\alpha > 1)$, *то* $\lim\limits_{n\to-\infty} \mathsf{M}\,|\,\xi_{-\infty} - \xi_n\,|^\alpha = 0$.

13.2.4. Теоремы о разложении. Супермартингал ζ_n $(n \geqslant 0)$ называется *потенциалом*, если $\zeta_n \geqslant 0$ и $\lim\limits_{n\to\infty} \zeta_n = 0$.

Теорема 6. *Для всякого супермартингала* ζ_n $(n \geqslant 0)$, *для которого* $\lim\limits_{n\to\infty} \mathsf{M}\xi_n > -\infty$, *справедливо представление*

$$\xi_n = \zeta_n + \eta_n,$$

где ζ_n *является потенциалом, а* η_n — *мартингалом. Такое разложение супермартингала единственно. При этом* η_n *является наибольшим субмартингалом, ограниченным сверху супермартингалом* ξ_n.

Теорема 7. *Если* ξ_n $(n \geqslant 0)$ — *супермартингал, то следующие три утверждения эквивалентны:*

а) $\lim\limits_{n\to\infty} \mathsf{M}\xi_n > -\infty$;

б) ξ_n *ограничен снизу субмартингалом;*

в) ξ_n *ограничен снизу мартингалом.*

Теорема 8 (разложение Крикеберга для мартингала). *Если* ξ_n $(n \geqslant 0)$ — *мартингал, для которого* $\sup\limits_n \mathsf{M}\,|\,\xi_n\,| < \infty$, *тогда*

$$\xi_n = \xi_n^{(1)} - \xi_n^{(2)},$$

где $\xi_n^{(1)}$ *и* $\xi_n^{(2)}$ — *неотрицательные мартингалы, причем* $\sup\limits_n \mathsf{M}\,|\,\xi_n\,| = \sup\limits_n \mathsf{M}\xi_n^{(1)} + \sup\limits_n \mathsf{M}\xi_n^{(2)}$, $\xi_n^{(1)}$ *является наименьшим неотрицательным* *мартингалом, ограничивающим* ξ_n *сверху, а* $\xi_n^{(2)}$ — *наименьшим неотрицательным, ограничивающим сверху* $-\xi_n$.

Теорема 9 (разложение Дуба). *Для всякого супермартингала* ξ_n $(n \geqslant 0)$ *можно указать возрастающую предсказуемую последовательность* α_n $(\alpha_0 = 0)$ *и мартингал* η_n *такие, что* $\xi_n = \eta_n - \alpha_n$. *Если* $\lim\limits_{n \to \infty} \mathsf{M}\xi_n > -\infty$, *то* α_n *ограничено и* $\alpha_\infty = \lim\limits_{n \to \infty} \alpha_n$ *таково, что* $\mathsf{M}\alpha_\infty < \infty$. *Если* ξ_n — *равномерно интегрируемый супермартингал, то* α_n *ограничено,* $\mathsf{M}\alpha_\infty < \infty$, *а мартингал* η_n *равномерно интегрируем.*

Справедливо и обратное: если $\mathsf{M}\alpha_\infty < \infty$ *и* η_n — *равномерно интегрируемый мартингал, то таким будет и супермартингал* ξ_n.

Теорема 10. *Всякий потенциал* ζ_n *представим в виде*

$$\zeta_n = \mathsf{M}\,[\alpha_\infty - \alpha_n \mid \mathfrak{F}_n],$$

где α_n — *предсказуемая возрастающая последовательность,* $\alpha_0 = 0$, $\alpha_\infty = \lim\limits_{n \to \infty} \alpha_n$.

13.3. Мартингалы с непрерывным временем

В этом параграфе рассматриваются мартингалы (супермартингалы и субмартингалы), определенные на R_+ (иногда на $[0, \infty]$) относительно потока σ-алгебр $(\mathfrak{F}_t)_{t \geqslant 0}$ (иногда от этой совокупности σ-алгебр будет требоваться лишь монотонность). Основная цель здесь — распространение результатов для последовательностей на процессы с непрерывным временем.

13.3.1. Теоремы о непрерывных справа супермартингалах.

Теорема 1. *Пусть* $\xi(t)$ — *супермартингал, заданный на* R_+ *относительно монотонной совокупности* σ-*алгебр* $(\mathfrak{F}_t)_{t \geqslant 0}$, D — *счетное и плотное в* R_+ *множество. Тогда*

а) *для почти всех* ω *по мере* P *у функции* $\xi(t)$ *для всех* $t \in \mathrm{R}_+$ *существует предел справа*

$$\xi(t+) = \lim_{s \in D,\ s \downarrow t} \xi(s)$$

и для всех $t > 0$ — *предел слева*

$$\xi(t-) = \lim_{s \in D,\ s \uparrow t} \xi(s);$$

б) *для всех* $t \in \mathrm{R}_+$ *существует* $\mathsf{M}\,|\xi(t+)|$ *и с вероятностью* 1

$$\xi(t) \geqslant \mathsf{M}\,(\xi(t+) \mid \mathfrak{F}_t)$$

(*неравенство превращается в равенство во всех точках, где* $\mathsf{M}\xi(t)$ *непрерывно справа*). *Процесс* $\xi_+(t) = \xi(t+)$ *является супермартингалом относительно совокупности* σ-*алгебр* $(\mathfrak{F}_{t+})_{t \geqslant 0}$, *где* $\mathfrak{F}_{t+} = \bigcap\limits_{s > t} \mathfrak{F}_s$; *он является мартингалом, если* $\xi(t)$ — *мартингал;*

в) *для всех* $t > 0$ $\mathsf{M}\,|\xi(t-)| < \infty$ *и*

$$\xi(t-) \geqslant \mathsf{M}\,(\xi(t) \mid \mathfrak{F}_{t-}), \quad \mathfrak{F}_{t-} = \bigvee_{s<t} \mathfrak{F}_s$$

(неравенство превращается в равенство, если $\mathsf{M}\xi(t)$ *непрерывно слева). Процесс* $\xi_-(t) = \xi(t-)$ *является супермартингалом относительно совокупности* σ*-алгебр* $(\mathfrak{F}_{t-})_{t\geqslant 0}$ $(\mathfrak{F}_{0-} = \mathfrak{F}_0)$, $\xi_-(0) = \xi(0)$; *он будет мартингалом, если таков* $\xi(t)$.

Очевидно, что $\xi_+(t)$ — непрерывный справа процесс; он будет стохатически эквивалентен исходному, если $\mathsf{M}\xi(t)$ непрерывно справа. Поэтому имеет смысл рассматривать непрерывные справа супермартингалы.

Теорема 2. *Пусть* $\xi(t)$ $(t \in \mathsf{R}_+)$ — *супермартингал по отношению к* $(\mathfrak{F}_t)_{t\geqslant 0}$, *для которого* $\xi(t)$ *с вероятностью* 1 *непрерывен справа. Тогда*

а) $\xi(t)$ *не имеет с вероятностью* 1 *разрывов второго рода;*

б) $\xi(t)$ *является супермартингалом относительно потока* $\hat{\mathfrak{F}}_t = \mathfrak{F}_{t+} \vee \mathfrak{S}_0$, *где* $\mathfrak{S}_0$ — σ*-алгебра, порожденная множествами* $\mathfrak{F}$ P*-меры* 0 $(\{\Omega, \mathfrak{F}, \mathsf{P}\}$ — *основное вероятностное пространство).*

Теорема 3. *Пусть* $\xi(t)$ $(t \in \mathsf{R}_+)$ — *супермартингал, непрерывный справа. Тогда*

а) $\lambda\mathsf{P}\{\sup\limits_{t\in\Delta}|\xi(t)| > \lambda\} \leqslant 3 \sup\limits_{t\in\Delta}\mathsf{M}|\xi(t)|, \quad \Delta \subset \mathsf{R}_+;$

б) *если* $\xi(t)$ — *мартингал, то*

$$\lambda\mathsf{P}\{\sup_{t\in\Delta}|\xi(t)| > \lambda\} \leqslant \sup_{t\in\Delta}\mathsf{M}|\xi(t)|,$$

при $\alpha > 1$

$$\mathsf{M}\sup_{t\in\Delta}|\xi(t)|^\alpha \leqslant \left(\frac{\alpha}{\alpha-1}\right)^\alpha \sup_{t\in\Delta}\mathsf{M}|\xi(t)|^\alpha;$$

в) *если* $\nu_\Delta^+([a, b])$ *обозначает число пересечений процессом* $\xi(t)$ *полосы* $[a, b]$ *на множестве* Δ, *то*

$$\mathsf{M}\nu_\Delta^+([a, b]) \leqslant \frac{1}{b-a}\left(|a| + \sup_{t\in\Delta}|\xi(t)|\right).$$

Далее будем предполагать, что $(\mathfrak{F}_t)_{t\geqslant 0}$ является потоком.

Теорема 4. *Пусть* $\xi(t)$ — *непрерывный справа супермартингал. Тогда*

а) *если* $\sup\limits_t \mathsf{M}|\xi(t)| < \infty$, *то с вероятностью* 1 *существует* $\lim\limits_{t\to\infty}\xi(t) = \xi_\infty$, *причем* $\mathsf{M}|\xi_\infty| < \infty$;

б) *утверждение* а) *имеет место, если супермартингал замкнут справа величиной* η, *при этом* $\xi_\infty \geqslant \mathsf{M}(\eta\,|\,\mathfrak{F}_{\infty-})$, *где* $\mathfrak{F}_{\infty-} = \bigvee\limits_t \mathfrak{F}_t$; *в случае мартингала* $\xi_\infty = \mathsf{M}(\eta\,|\,\mathfrak{F}_{\infty-})$;

в) *если супермартингал* $\xi(t)$ *(мартингал) равномерно интегрируем, то* ξ_∞ *замыкает* $\xi(t)$ *справа.*

Теорема 5. *Пусть* $\xi(t)$ — *непрерывный справа супермартингал, определенный для* $t > 0$. *Если* $\lim\limits_{t\to 0}\mathsf{M}\xi(t) < \infty$, *то с вероятностью* 1 *существует* $\lim\limits_{t\to 0}\xi(t) = \xi_0$, *причем* $\lim\limits_{t\to 0}\mathsf{M}|\xi(t) - \xi_0| = 0$.

Рассмотрим теперь значения процесса в моменты остановки.

Теорема 6. *Пусть* $\xi(t)$ $(t \in \mathrm{R}_+)$ — *непрерывный справа супермартингал, замкнутый справа величиной* ξ_∞, *измеримой относительно* $\mathfrak{F}_\infty \supset \mathfrak{F}_t$ $(t \in \mathrm{R}_+)$. *Тогда*

а) *если* τ *и* σ — *два момента остановки, для которых* $\tau \leqslant \sigma$, *то существуют* $\mathsf{M}\,|\xi(\sigma)|$ *и* $\mathsf{M}\,|\xi(\tau)|$ *и*

$$\mathsf{M}(\xi(\sigma)\,|\,\mathfrak{F}_\tau) \leqslant \xi(\tau);$$

б) *если* τ *и* σ — *произвольные моменты остановки, то*

$$\mathsf{M}(\xi(\sigma)\,|\,\mathfrak{F}_\tau) \leqslant \xi(\sigma \wedge \tau);$$

в) *если* τ — *момент остановки, то*

$$\xi^\tau(t) = \xi(t \wedge \tau)$$

является супермартингалом (мартингалом, если $\xi(t)$ — *мартингал).*

Будем обозначать для момента остановки τ через $\mathfrak{F}_{\tau-}$ σ-алгебру, порожденную событиями вида $A_t \cap \{\tau < t\}$, где $A_t \in \mathfrak{F}_t$.

Теорема 7. *Пусть* $\xi(t)$ $(t \geqslant 0)$ — *непрерывный справа супермартингал, замкнутый справа величиной* ξ_∞. *Будем считать, что* $\xi(0-) = \xi(0)$. *Тогда*

а) *если* σ *и* τ — *два предсказуемых момента остановки, для которых* $\sigma \leqslant \tau$, *то* $\xi(\sigma-)$ *и* $\xi(\tau-)$ *имеют математические ожидания и*

$$\xi(\sigma-) \geqslant \mathsf{M}(\xi(\tau-)\,|\,\mathfrak{F}_{\sigma-}) \geqslant \mathsf{M}(\xi(\tau)\,|\,\mathfrak{F}_{\sigma-})$$

(для мартингала неравенство превращается в равенство);

б) *пусть* τ — *предсказуемый момент остановки,* $\xi^{\tau-}(t) = \xi(t)$ $t < \tau$, $\xi^{\tau-}(t) = \xi(\tau-)$, $t \geqslant \tau$; *тогда, если* $\xi(t)$ — *непрерывный справа супермартингал (мартингал), то* $\xi^{\tau-}(t)$ — *также супермартингал (мартингал);*

в) *если* $\xi(t)$ — *неотрицательный супермартингал, непрерывный справа,* $\tau = \inf[t, \xi(t) = 0]$, *то* $\xi(t) = 0$ *при* $t \geqslant \tau$.

Теорема 8. *Пусть* $\xi_n(t)$ — *возрастающая последовательность непрерывных справа супермартингалов,* $\xi(t) = \sup\limits_n \xi_n(t)$. *Тогда* $\xi(t)$ *в предположении, что* $\xi(0)$ *конечно, является процессом без разрывов второго рода, непрерывным справа.*

13.3.2. Супермартингалы класса (*D*). Локальные мартингалы и супермартингалы. Пусть $\eta(t)$ $(t \geqslant 0)$ — измеримый случайный процесс. Он называется *ограниченным* в L^1 (по отношению к потоку $(\mathfrak{F}_t)_{t \geqslant 0}$), если конечна величина

$$\|\eta\|_1 = \sup_\tau \mathsf{M}\,|\eta(\tau)|\,I_{\{\tau<\infty\}},$$

где τ — пробегает все моменты остановки относительно потока $(\mathfrak{F}_t)_{t \geqslant 0}$. Если семейство случайных величин $\eta(\tau)\,I_{\{\tau<\infty\}}$ равномерно интегрируемо, то говорят, что $\eta(t)$ принадлежит классу (D).

Теорема 9. а) *Равномерно интегрируемый непрерывный справа мартингал принадлежит классу* (D);

б) *замкнутый справа неотрицательный субмартингал принадлежит классу* (D);

в) *для непрерывного справа супермартингала* $\xi(t)$ *существует последовательность моментов остановки* τ_n *такая, что* $\tau_n \uparrow \infty$ *и* $\xi(t \wedge \tau_n)$ *для всех* n *принадлежит классу* (D);

г) *пусть $\xi(t)$ — неотрицательный непрерывный справа супермартингал, $\zeta_n = \inf[t: \xi(t) \geqslant n]$; для того чтобы $\xi(t)$ принадлежал классу (D), необходимо и достаточно, чтобы*

$$\lim_{n\to\infty} \mathsf{M}\xi(\zeta_n) I_{\{\zeta_n<\infty\}} = 0.$$

Приведем пример неотрицательного равномерно интегрируемого супермартингала, не принадлежащего классу (D).

Пусть $w(t)$ — трехмерное броуновское движение, выходящее из точки $x \neq 0$. Тогда процесс $\xi(t) = |w(t)|^{-1}$ является даже потенциалом; кроме того, он непрерывен с вероятностью 1. Пусть $\zeta_n = \inf\{t: \xi(t) = n\}$. При $\zeta_n < \infty$ $\xi(\zeta_n) = n$, а $\mathsf{P}\{\zeta_n < \infty\} = |x|/n$ при $|x| < n$, так что $\mathsf{M}\xi(\zeta_n) I_{\{\zeta_n<\infty\}} = |x|$ для достаточно больших n. Из утверждения г) следует, что $\xi(t)$ не принадлежит классу (D).

Пусть $\xi(t)$ $(t \geqslant 0)$ — непрерывный справа процесс, согласованный с потоком σ-алгебр $(\mathfrak{F}_t)_{t\geqslant 0}$. Говорят, что $\xi(t)$ является локальным мартингалом (относительно потока $(\mathfrak{F}_t)_{t\geqslant 0}$), если существует такая возрастающая последовательность моментов остановки τ_n, что $\xi(t\wedge\tau_n)$ является равномерно интегрируемым мартингалом и $\tau_n \to \infty$. Последовательность моментов остановки называется *приводящей локальный мартингал.* Момент остановки τ *приводит* локальный мартингал $\xi(t)$, если $\xi(t\wedge\tau)$ — мартингал.

Всякий мартингал является локальным мартингалом.

Теорема 10. *Пусть $\xi(t)$ — локальный мартингал. Тогда*

а) *если τ, σ — моменты остановки, τ приводит $\xi(t)$ и $\sigma \leqslant \tau$, то σ также приводит $\xi(t)$;*

б) *сумма двух локальных мартингалов является локальным мартингалом;*

в) *если τ и σ — два момента остановки, приводящих $\xi(t)$, то таким будет и момент остановки $\tau \vee \sigma$;*

г) *для всякого момента остановки τ процесс $\xi^\tau(t) = \xi(t\wedge\tau)$ является также локальным мартингалом;*

д) *если $\xi(t)$ $(t \geqslant 0)$ — непрерывный справа, согласованный с потоком $(\mathfrak{F}_t)_{t\geqslant 0}$ процесс, для которого существует такая последовательность моментов остановки τ_n, что $\sup\limits_n \tau_n = +\infty$ и $\xi^{\tau_n}(t) = \xi(t\wedge\tau_n)$ является локальным мартингалом для всех n, то $\xi(t)$ — также локальный мартингал.*

13.3.3. Теоремы о разложениях для мартингалов и квазимартингалов.

Теорема 11. *Пусть $\xi(t)$ — непрерывный справа мартингал. Он ограничен в L^1 (т. е. $\|\xi\|_1 < \infty$) тогда, когда он представим в виде*

$$\xi(t) = \xi^+(t) - \xi^-(t),$$

где $\xi^+(t)$ и $\xi^-(t)$ — два неотрицательных непрерывных справа мартингала, причем эти мартингалы могут быть выбраны таким образом, чтобы

$$\|\xi\|_1 = \|\xi^+\|_1 + \|\xi^-\|_1 = \mathsf{M}\xi^+(0) + \mathsf{M}\xi^-(0). \quad (3.1)$$

При этом $\xi^+(t)$ является наименьшим мартингалом, ограничивающим сверху $\xi(t)$, а $\xi^-(t)$ — наименьшим мартингалом, ограничивающим сверху $-\xi(t)$.

Теорема 12. *Пусть $\xi(t)$ — локальный мартингал, непрерывный справа. Он допускает единственное представление в виде $\xi(t) = \xi^+(t) - \xi^-(t)$, где $\xi^+(t)$ и $\xi^-(t)$ — неотрицательные непрерывные справа локальные мартингалы, такие, что выполнено* (3.1).

Если $\xi(t) = \xi^1(t) - \xi^2(t)$, где ξ^1 и ξ^2 — неотрицательные супермартингалы, то $\xi^1(t) - \xi^+(t)$ и $\xi^2(t) - \xi^-(t)$ являются неотрицательными супермартингалами.

Пусть $\xi(t)$ $(t \geqslant 0)$ — некоторый непрерывный справа случайный процесс, согласованный с потоком $(\mathfrak{F}_t)_{t \geqslant 0}$, для которого $\mathsf{M}\,|\xi(t)| < \infty$ для всех $t \geqslant 0$. Для замкнутого отрезка $[a, b]$ введем величину

$$\operatorname{Var}_{[a,\, b]}(\xi) = \sup \mathsf{M} \sum_{j=0}^{n-1} |\mathsf{M}(\xi(t_{j+1}) - \xi(t_j) \mid \mathfrak{F}_{t_j})|,$$

где $a = t_0 < t_1 < \ldots < t_n = b$; супремум берется по всем разбиениям отрезка $[a, b]$. Если величина $\operatorname{Var}_{[a,\, b]}(\xi) < \infty$, то $\xi(t)$ называется квазимартингалом (относительно потока $(\mathfrak{F}_t)_{t \geqslant 0}$) на отрезке $[a, b]$. Если $\operatorname{Var}_{[0,\, \infty[}(\xi) = \sup\limits_t \operatorname{Var}_{[0,\, t]}(\xi) < \infty$, то $\xi(t)$ называется *квазимартингалом*. Если определен $\xi(\infty)$, то определена $\operatorname{Var}_{[0,\, \infty]}(\xi)$, и в случае конечности этой величины $\xi(t)$ — квазимартингал на $[0, \infty]$. Если $\xi(t)$ определен лишь при $t < \infty$, то, полагая $\xi^0(t) = \xi(t)$ при $t < \infty$, $\xi^0(\infty) = 0$ и $\operatorname{Var}^*_{[0,\, \infty]}(\xi) = \operatorname{Var}_{[0,\, \infty]}(\xi^0)$, будем говорить, что $\xi(t)$ — квазимартингал вплоть до бесконечности, если $\operatorname{Var}^*_{[0,\, \infty]}(\xi) < \infty$.

Теорема 13. *Мартингал $\xi(t)$ будет квазимартингалом вплоть до бесконечности тогда и только тогда, когда он ограничен в L^1, при этом*

$$\operatorname{Var}^*_{[0,\, \infty]}(\xi) = \|\xi\|_1 = \sup_t \mathsf{M}\,|\xi(t)|.$$

Теорема 14. *Пусть $\xi(t)$ определен при $t \geqslant 0$. Он будет квазимартингалом вплоть до бесконечности тогда и только тогда, когда он представим как разность двух неотрицательных непрерывных справа супермартингалов: $\xi(t) = \xi^+(t) - \xi^-(t)$. В этом случае существует такое единственное разложение указанного выше вида, что*

$$\operatorname{Var}^*_{[0,\, \infty]}(\xi) = \operatorname{Var}^*_{[0,\, \infty]}(\xi^+) + \operatorname{Var}^*_{[0,\, \infty]}(\xi^-) = \mathsf{M}(\xi^+(0) + \xi^-(0)).$$

Для всякого другого разложения $\xi(t) = \xi^1(t) - \xi^2(t)$ на два неотрицательных непрерывных справа супермартингала процессы $\xi^1(t) - \xi^+(t)$ и $\xi^2(t) - \xi^-(t)$ будут неотрицательными супермартингалами.

Теорема 15. *Пусть $\xi(t)$ — квазимартингал на $[0, \infty[$. Тогда он допускает представление в виде*

$$\xi(t) = \xi_0(t) + \xi^+(t) - \xi^-(t),$$

где $\xi_0(t)$ — *мартингал,* $\xi^{\pm}(t)$ — *неотрицательные супермартингалы и все процессы* ξ_0, ξ^+, ξ^- *непрерывны справа.*

13.3.4. Возрастающие процессы. Под *возрастающим процессом* $\alpha(t)$ $(t \geqslant 0)$ понимают согласованный неотрицательный непрерывный справа процесс, выборочные функции которого являются неубывающими. Он называется *интегрируемым,* если $M\alpha(\infty -) = \lim\limits_{t \to \infty} M\alpha(t) < \infty$. Разность двух возрастающих интегрируемых процессов называется *процессом интегрируемой вариации.* Возрастающий процесс $\alpha(t)$ называется *непрерывным,* если $\alpha(0) = 0$ и $\alpha(t)$ — непрерывная функция t с вероятностью 1.

Теорема 16. *Пусть* $\alpha(t)$ — *возрастающий процесс. Тогда существуют непрерывный возрастающий процесс* $\alpha^c(t)$, *последовательность моментов остановки* τ_r *и (неслучайных) постоянных* λ_n *такие, что*

$$\alpha(t) = \alpha^c(t) + \sum_n \lambda_n I_{\{\tau_n \leqslant t\}}.$$

Если $\alpha(t)$ предсказуем, моменты τ_n можно выбрать предсказуемыми. Процесс $\alpha^c(t)$ определяется однозначно; он называется *непрерывной частью* $\alpha(t)$, а процесс

$$\alpha^d(t) = \alpha(t) - \alpha^c(t)$$

называется чисто *разрывной (скачкообразной) частью* $\alpha(t)$.

Теорема 17. *Пусть* $\xi(t)$ — *измеримый неотрицательный процесс,* ${}^o\xi(t)$ *и* ${}^p\xi(t)$ — *соответственно его опциональная и предсказуемая проекции (см. п. 9.4.5). Тогда*

$$M \int\limits_{[0,\,\infty[} \xi(s)\,d\alpha(s) = M \int\limits_{[0,\,\infty[} {}^o\xi(s)\,d\alpha(s),$$

и в том случае, когда $\alpha(t)$ *предсказуем,*

$$M \int\limits_{[0,\,\infty[} \xi(s)\,d\alpha(s) = M \int\limits_{[0,\,\infty[} {}^p\xi(s)\,d\alpha(s).$$

Теорема 18. *Пусть* $\alpha(t)$ — *возрастающий не обязательно согласованный интегрируемый процесс* $(M\alpha(\infty -) < \infty)$. *Если для каждого ограниченного процесса* $\xi(t)$

$$M \int\limits_{[0,\,\infty[} \xi(s)\,d\alpha(s) = M \int\limits_{[0,\,\infty[} {}^o\xi(s)\,d\alpha(s),$$

то $\alpha(t)$ *вполне измерим и, значит, является возрастающим процессом в обычном смысле; если при тех же условиях*

$$M \int\limits_{[0,\,\infty[} \xi(s)\,d\alpha(s) = M \int\limits_{[0,\,\infty[} {}^p\xi(s)\,d\alpha(s),$$

то $\alpha(t)$ *предсказуем.*

Пусть $\alpha(t)$ $(t \geqslant 0)$ — возрастающий интегрируемый не обязательно согласованный процесс. Опциональный возрастающий процесс

$\beta(t)$ называется *опциональной дуальной проекцией* процесса $\alpha(t)$, если

$$\mathsf{M}\int\limits_{[0,\,\infty[}\xi(s)\,d\beta(s)=\mathsf{M}\int\limits_{[0,\,\infty[}{}^{o}\xi(s)\,d\alpha(s)$$

для всякого ограниченного измеримого процесса $\xi(t)$.

Предсказуемый возрастающий процесс $\gamma(t)$ называется *предсказуемой дуальной проекцией процесса* $\alpha(t)$, если

$$\mathsf{M}\int\limits_{[0,\,\infty[}\xi(s)\,d\gamma(s)=\mathsf{M}\int\limits_{[0,\,\infty[}{}^{p}\xi(s)\,d\alpha(s)$$

для тех же $\xi(t)$.

Теорема 19. *Пусть $\alpha(t)$ и $\beta(t)$ — два возрастающих интегрируемых не обязательно согласованных процесса.*

а) *Для того чтобы они имели одинаковые опциональные дуальные проекции, необходимо и достаточно, чтобы для любого момента остановки τ* $\mathsf{M}[\alpha(\infty-)-\alpha(\tau-)\mid\mathfrak{F}_\tau]=\mathsf{M}[\beta(\infty-)-\beta(\tau-)\mid\mathfrak{F}_\tau]$ $(\alpha(0-)=\beta(0-)=0)$.

б) *Для того чтобы они имели одинаковые предсказуемые дуальные проекции, необходимо и достаточно, чтобы*

$$\mathsf{M}(\alpha(0)\mid\mathfrak{F}_0)=\mathsf{M}(\beta(0)\mid\mathfrak{F}_0),\quad \mathsf{M}(\alpha(\infty-)-\alpha(t)\mid\mathfrak{F}_t)=$$
$$=\mathsf{M}(\beta(\infty-)-\beta(t)\mid\mathfrak{F}_t)$$

с вероятностью 1.

Пусть $\alpha(t)$ $(t\geqslant 0)$ — возрастающий (согласованный) процесс. Его предсказуемая дуальная проекция называется *компенсатором*.

Возрастающий не обязательно согласованный процесс $\alpha(t)$ $(t\geqslant 0)$ называется *локально интегрируемым*, если существует возрастающая последовательность моментов остановки τ_n, $\tau_n\to+\infty$, для которой $\mathsf{M}(\alpha(\tau_n)-\alpha(0))<\infty$ и $\mathsf{M}(\alpha(0)\mid\mathfrak{F}_0)<\infty$ с вероятностью 1. Процесс, представимый в виде разности двух возрастающих локально интегрируемых процессов, называется *процессом локально ограниченной вариации*.

Очевидны определения опциональной дуальной проекции, предсказуемой дуальной проекции и компенсатора для локально интегрируемого возрастающего процесса, а также для процесса локально ограниченной вариации.

Теорема 20. *Всякий интегрируемый (локально интегрируемый) возрастающий процесс имеет компенсатор, который также является интегрируемым (локально интегрируемым).*

Теорема 21. *Если $\xi(t)$, $\xi(0)=0$, есть локальный мартингал конечной вариации, тогда он представим в виде $\xi(t)=\eta(t)-\zeta(t)$, где $\eta(t)$ есть процесс локально ограниченной вариации, а $\zeta(t)$ — его компенсатор.*

13.3.5. Потенциалы, порождаемые возрастающими процессами. Пусть $\beta(t)$ $(t\in[0,\infty])$ — возрастающий интегрируемый не обязательно согласованный процесс. Обозначим через $\nu(t)$ непрерывную справа модификацию мартингала $\nu(t)=\mathsf{M}(\beta(\infty)\mid\mathfrak{F}_t)$; пусть далее $\alpha(t)$ — предсказуемая дуальная проекция $\beta(t)$.

Супермартингал $\pi(t)=\nu(t)-\alpha(t)$ является потенциалом; он называется *потенциалом, порождаемым возрастающим процессом* $\beta(t)$.

Если $\beta(t)$ обладает теми же свойствами и $\alpha^*(t)$ — его опциональная дуальная проекция, то супермартингал

$$\pi^*(t) = \mathsf{M}(\beta(\infty) \mid \mathfrak{F}_t) - \alpha^*(t-)$$

также является потенциалом; он называется *левым потенциалом*, порожденным возрастающим процессом $\beta(t)$.

Теорема 22. *Если $\alpha(t)$ $(t \geqslant 0)$ — предсказуемый возрастающий интегрируемый процесс и $\pi(t)$ — порожденный им потенциал, то*

$$\mathsf{M}[\alpha(\infty)]^2 = \mathsf{M} \int\limits_{[0,\ \infty[} (\pi(s) + \pi(s-))\, d\alpha(s).$$

Теорема 23. *Если $\alpha(t)$ $(t \geqslant 0)$ — опциональный возрастающий интегрируемый процесс и $\pi^*(t)$ — порожденный им левый потенциал, то*

$$\mathsf{M}[\alpha(\infty)]^2 = \mathsf{M} \int\limits_{[0,\ \infty[} (\pi^*(s) + \pi^*(s+))\, d\alpha(s).$$

Теорема 24. *Пусть $\alpha(t)$ — возрастающий (опциональный) процесс, который порождает потенциал (левый потенциал), ограниченный сверху постоянной c. Тогда*

а) *для всех $n \geqslant 1$*

$$\mathsf{M}[\alpha(\infty)]^n \leqslant n!\, c^n;$$

б) *для всех $0 \leqslant \lambda \leqslant 1/c$*

$$\mathsf{M} \exp\{\lambda\alpha(\infty)\} \leqslant 1/(1 - \lambda c).$$

Теорема 25. *Пусть $\alpha(t)$ — возрастающий интегрируемый предсказуемый процесс, а $\eta(t)$ — неотрицательный измеримый не обязательно согласованный процесс. Предположим, что для всякого предсказуемого ограниченного момента остановки τ выполнено*

$$\mathsf{M}\eta(\tau) \leqslant \mathsf{M}\alpha(\tau-), \quad \alpha(0-) = 0;$$

тогда для всех $\lambda > 0$ и предсказуемого ограниченного момента остановки τ

а) $\mathsf{M}(\eta(\tau) \wedge \lambda) \leqslant \mathsf{M}(\alpha(\tau-) \wedge \lambda) + \lambda\mathsf{P}\{\alpha(\tau-) \geqslant \lambda\}$;

б) $\mathsf{M}(\eta(\tau) \wedge \lambda) \leqslant 2\mathsf{M}(\alpha(\tau-) \wedge \lambda)$;

в) *если $\eta(t)$ также предсказуем, то*

$$\lambda\mathsf{P}\{\sup_{t \leqslant \tau} \eta(t) \geqslant \lambda\} \leqslant \mathsf{M}(\alpha(\tau-) \wedge \lambda) + \lambda\mathsf{P}\{\alpha(\tau-) > \lambda\}.$$

13.3.6. Разложение Дуба для супермартингалов. Разложение Дуба для последовательностей приведено в п. 13.2.4. Обобщение этого разложения на супермартингалы непрерывного аргумента принадлежит П. А. Мейеру.

Теорема 26. *Пусть $\zeta(t)$ — неотрицательный супермартингал класса (D). Тогда существует интегрируемый предсказуемый возрастающий процесс $\alpha(t)$, определенный при $t \in [0, \infty]$, для которого $\alpha(0) = 0$ (но не обязательно $\alpha(\infty) = \alpha(\infty-)$) и*

$$\zeta(t) = \mathsf{M}(\alpha(\infty) - \alpha(t) \mid \mathfrak{F}_t) \qquad (3.1)$$

с вероятностью 1; такой процесс $\alpha(t)$ определяется единственным образом с точностью до неотличимости (т. е. два таких процесса с вероятностью 1 совпадают для всех t). Наоборот, если $\alpha(t)$ удовлетворяет перечисленным выше условиям, то $\zeta(t)$ — супермартингал класса (D).

Процесс $\alpha(t)$ называется *возрастающим процессом, ассоциированным с супермартингалом* $\zeta(t)$.

Теорема 27. *Пусть $\alpha(t)$ — возрастающий процесс, ассоциированный с супермартингалом $\zeta(t)$ класса (D), $\Delta\alpha(t)=\alpha(t)-\alpha(t-)$. Тогда $\Delta\alpha(t)=\zeta(t-)-{}^p\zeta(t)$, где ${}^p\zeta(t)$ — предсказуемая проекция процесса $\zeta(t)$. В частности, $\alpha(t)$ — непрерывный процесс, если ${}^p\zeta(t)=\zeta(t-)$. Это условие эквивалентно следующему: для всякой возрастающей последовательности моментов остановки τ_n, для которой $\lim\tau_n<\infty$, будет $\lim\limits_{n\to\infty}\mathsf{M}\zeta(\tau_n)=\mathsf{M}\zeta(\lim\tau_n)$.*

Теорема 28. *Пусть $\zeta(t)$ — неотрицательный супермартингал класса (D). Тогда*

$$\zeta(t)=\mu(t)-\alpha(t),$$

где $\mu(t)$ — замкнутый справа (а следовательно, равномерно интегрируемый, или, что то же самое, класса (D)) мартингал, а $\alpha(t)$ — предсказуемый интегрируемый возрастающий процесс (ассоциированный с $\zeta(t)$ возрастающий процесс). Для того чтобы супермартингал допускал указанное разложение, необходимо и достаточно, чтобы супермартингал $\zeta(t)$ был класса (D).

Теорема 29. *Пусть $\zeta(t)$ — супермартингал. Тогда $\zeta(t)$ имеет единственное разложение вида*

$$\zeta(t)=\mu(t)-\alpha(t),$$

где $\mu(t)$ является локальным мартингалом, а $\alpha(t)$ — предсказуемым возрастающим процессом, для которого $\alpha(0)=0$ и $\mathsf{M}\alpha(t)<\infty$ для всех $t\geqslant 0$; $\alpha(t)$ — интегрируемый процесс тогда и только тогда, когда $\lim\limits_{t\to\infty}\mathsf{M}\zeta(t)>-\infty$.

Это разложение называется *разложением Дуба* для супермартингала; процесс $\alpha(t)$ называется *возрастающим процессом, ассоциированным с супермартингалом* $\zeta(t)$.

Теорема 30. *Пусть $\eta(t)$ — неотрицательный супермартингал. Тогда $\eta(t)$ допускает представление*

$$\eta(t)=\zeta(t)+\nu(t),$$

где $\zeta(t)$ — потенциал класса (D), а $\nu(t)$ — локальный мартингал.

Семейство случайных процессов $\{\xi^\theta(t),\ \theta\in\Theta\}$, согласованных с потоком $(\mathfrak{F}_t)_{t\geqslant 0}$, равномерно принадлежит классу (D), если семейство случайных величин $\{\xi^\theta(\tau),\ \theta\in\Theta,\ \tau$ — конечный момент остановки$\}$ равномерно интегрируемо.

Теорема 31. *Пусть $\xi_n(t)$ $(n\geqslant 0,\ t\in[0,\infty[)$ — последовательность неотрицательных супермартингалов и $\alpha_n(t)$ $(n\geqslant 0)$ — ассоциированные с ними возрастающие процессы. Предположим, что*

а) *$\xi_n(t)$ равномерно принадлежит классу (D);*

б) *для каждого момента остановки τ $\mathsf{M}\,|\,\xi_0(\tau)-\xi_n(\tau)\,|\to 0$.*

Тогда для всякого момента остановки τ

$$\mathsf{M}\,|\,\alpha_0(\tau)-\alpha_n(\tau)\,|\to 0.$$

Если дополнительно $\sup\limits_{t\geqslant 0} |\xi_n(t)-\xi_0(t)|\to 0$ *по вероятности, то*

$$\lim_{n\to\infty}\sup_t \mathsf{M}\,|\alpha_0(t)-\alpha_n(t)| = 0.$$

13.4. Семимартингалы и стохастические интегралы

13.4.1. Семимартингалы. Основные свойства. Согласованный с потоком $(\mathfrak{F}_t)_{t\geqslant 0}$ случайный процесс $\xi(t)$ $(t\geqslant 0)$ называется *семимартингалом*, если он допускает представление $\xi(t)=\xi(0)+\alpha(t)+\nu(t)$, где $\alpha(t)$ — процесс конечной вариации, $\alpha(0)=0$, а $\nu(t)$ — локальный мартингал, $\nu(0)=0$. Если можно указать такое представление, что $\alpha(t)$ предсказуем, то $\xi(t)$ называется *специальным семимартингалом.*

Примером семимартингала, не являющегося специальным, может служить такой: $\xi(t)=0$, $t<1$, $\xi(t)=\xi(1)$, $t\geqslant 1$, $\mathfrak{F}_t=\mathfrak{F}_0$, $t<1$, где $\mathfrak{F}_0$ — тривиальная σ-алгебра, $\mathfrak{F}_t=\mathfrak{F}_1$, $t\geqslant 1$, $\mathfrak{F}_1$ порождена $\xi(1)$, $\mathsf{M}\xi(1)$ не существует. Всякий предсказуемый процесс $\alpha(t)$ в точке $t=1$ должен иметь неслучайное значение; тогда $\xi(t)-\alpha(t)$ не может быть локальным мартингалом. Условия, при которых семимартингал $\xi(s)$ является специальным, даются следующими утверждениями.

1. *Пусть* $\xi(t)$ $(t\geqslant 0)$ — *семимартингал. Следующие утверждения эквивалентны:*

а) $\xi(t)$ — *специальный семимартингал;*

б) *существует представление* $\xi(t)=\xi(0)+\alpha(t)+\nu(t)$, *где* $\alpha(t)$ — *процесс локально интегрируемой вариации;*

в) *для всякого представления* $\xi(t)=\xi(0)+\alpha(t)+\nu(t)$, *где* $\nu(t)$ — *локальный мартингал,* $\alpha(t)$ — *процесс локально интегрируемой вариации;*

г) *возрастающий процесс* $\xi^*(t)=\sup\limits_{s\leqslant t}|\xi(s)|$ *локально интегрируем;*

д) *возрастающий процесс* $\delta(t)=\sup\limits_{s\leqslant t}|\xi(s)-\xi(s-)|$ *локально интегрируем.*

2. *Если для процесса* $\xi(t)$ *существует такая возрастающая последовательность моментов остановки* τ_n, *что* $\tau_n\to\infty$ *с вероятностью* 1 *и процесс* $\xi(t\wedge\tau_n)$ *есть (специальный) семимартингал, то* $\xi(t)$ — *также (специальный) семимартингал.*

3. *Пусть* $(\tau_t)_{t\geqslant 0}$ — *семейство конечных моментов остановки относительно потока* $(\mathfrak{F}_t)_{t\geqslant 0}$, *монотонно зависящих от* t *и непрерывных справа по* t. *Тогда*

а) *семейство* σ-*алгебр* $\left(\hat{\mathfrak{F}}_t=\mathfrak{F}_{\tau_t}\right)_{t\geqslant 0}$ *будет также потоком;*

б) *если* $\xi(t)$ — *семимартингал относительно* $(\mathfrak{F}_t)_{t\geqslant 0}$, *то* $\hat{\xi}(t)=\xi(\tau_t)$ *будет семимартингалом относительно* $(\hat{\mathfrak{F}}_t)_{t\geqslant 0}$.

4. *Произведение двух семимартингалов является семимартингалом.*

5. *Если* $\xi(t)$ — *семимартингал,* f — *выпуклая функция, то* $f(\xi(t))$ — *также семимартингал.*

Пусть $\xi(t)$ и $\eta(t)$ — два локальных мартингала. Если $\xi(t)\eta(t)$ есть специальный семимартингал, то единственный предсказуемый процесс $\langle\xi, \eta\rangle_t$, локально ограниченный, равный нулю при $t=0$, для которого $\xi(t)\eta(t)-\langle\xi, \eta\rangle_t$ является локальным мартингалом, называется *взаимной характеристикой* ξ и η. Для нее используется обозначение $\langle\xi, \eta\rangle$.

6. *Пусть $\xi(t)$ и $\eta(t)$ — два локальных мартингала. Тогда существует единственный процесс локально ограниченной вариации $[\xi, \eta]_t$ такой, что $\xi(t)\eta(t)-[\xi, \eta]_t$ — локальный мартингал.*

Квадратные скобки можно определить для любых семимартингалов, считая, что они линейны по любому аргументу и симметричны, и используя представление семимартингалов $\xi(t)=\alpha(t)+\nu_1(t)$, $\eta(t)=\beta(t)+\nu_2(t)$, где α и β конечной вариации, ν_1, ν_2 — локальные мартингалы. При этом

$$[\xi, \eta]=[\alpha, \beta]+[\alpha, \nu_2]+[\nu_1, \beta]+[\nu_1, \nu_2];$$

здесь $[\nu_1, \nu_2]$ уже определено, а $[\gamma, \delta]$, где γ не имеет разрывов второго рода и непрерывен справа и δ имеет конечную вариацию, определяется равенством

$$[\gamma, \delta]_t=\int_0^t[\gamma(s)-\gamma(s-)]\,d\delta(s).$$

Эта формула позволяет определить $[\alpha, \beta]$, $[\alpha, \nu_2]=[\nu_2, \alpha]$ и $[\nu_1, \beta]$.

Пример. Пусть $\xi(t)$ — однородный процесс с независимыми приращениями, имеющий представление

$$\xi(t)=\sigma w(t)+\int f(u)\,q(du\times dt);$$

здесь $\sigma\in\mathbf{R}$, $w(t)$ — винеровский процесс, $f(u)$ — функция из $\mathbf{R}_+$ в $\mathbf{R}$, для которой при всех $c<\infty$ $\int I_{\{|f|\leqslant c\}}\,du<\infty$, а $q(du\times dt)$ — центрированная пуассоновская мера на $\mathbf{R}_+^2$: $q(du\times dt)=p(du\times dt)-du\,dt$, где $p(du\times dt)$ имеет распределение Пуассона с параметром $du\,dt$. Тогда

$$[\xi, \xi]=\sigma^2+\int f^2(u)\,p(du\times dt),$$

величина $\langle\xi, \xi\rangle$ определена, если $\int f^2(u)\,du<\infty$, и $\langle\xi, \xi\rangle_t=(\sigma^2+\int f^2(u)\,du)\,t$.

7. *Пусть $\xi(t)$ — семимартингал. Для того чтобы он был специальным, необходимо и достаточно, чтобы возрастающий процесс $[\xi, \xi]_t$ был локально интегрируемым.*

13.4.2. Локально квадратически интегрируемые мартингалы. Пусть $\nu(t)$ — согласованный с потоком $(\mathfrak{F}_t)_{t\geqslant 0}$ непрерывный справа процесс. Он называется *локально квадратически интегрируемым мартингалом*, если существует возрастающая к ∞ последовательность моментов остановки τ_n такая, что $\nu(t\wedge\tau_n)$ — мартингал и $\mathsf{M}\nu^2(t\wedge\tau_n)<\infty$. Будем обозначать множество локально квадратически интегрируемых мартингалов через $\mathcal{M}_2^l$.

Справедливы следующие утверждения о локально квадратически интегрируемых мартингалах.

1. *Если* $\nu \in \mathcal{M}_2^l$, *то определен процесс* $\langle\nu\rangle_t = \langle\nu, \nu\rangle_t$, *называемый (квадратической) характеристикой* $\xi(t)$. *При* ν_1 *и* $\nu_2 \in \mathcal{M}_2^l$ *определен процесс* $\langle\nu_1, \nu_2\rangle$. *Если* $\langle\nu_1, \nu_2\rangle_t = 0$ $\forall t \geqslant 0$, *то* ν_1 *и* ν_2 *ортогональны:* $\nu_1 \perp \nu_2$.

2. *Пусть* $\nu_1(t), \nu_2(t) \in \mathcal{M}_2^l$, $\xi_1(t), \xi_2(t)$ — *измеримые процессы. Тогда*

$$\int_0^\infty |\xi_1(t)\xi_2(t)|\,|d\langle\nu_1, \nu_2\rangle_t| \leqslant \left(\int_0^\infty \xi_1^2(t)\,d\langle\nu_1\rangle_t\right)^{1/2} \times \left(\int_0^\infty \xi_2^2(t)\,d\langle\nu_2\rangle_t\right)^{1/2}$$

(*это неравенство Кунита — Ватанабе*).

3. *Пусть* $\nu(t) \in \mathcal{M}_2^l$. *Для того чтобы* $\langle\nu\rangle_t$ *был непрерывным процессом, необходимо и достаточно, чтобы для любого предсказуемого момента остановки* τ *с вероятностью* 1 $\nu(\tau) = \nu(\tau-)$. *Это условие эквивалентно следующему: для всякого предсказуемого момента остановки* τ, *для которого* $\mathsf{M}\nu^2(\tau) < \infty$, *будет* $\mathsf{M}\nu^2(\tau-) = \mathsf{M}\nu^2(\tau)$.

4. *Пусть* $\nu(t) \in \mathcal{M}_2^l$ *и* $\nu(t)$ *непрерывен с вероятностью* 1. *Тогда*

$$\langle\nu\rangle_t = \lim_{\lambda\to 0}\sum_{k=1}^n [\nu(t_k) - \nu(t_{k-1})]^2$$

по вероятности, где $0 = t_0 < t_1 < \ldots < t_n = t$, $\lambda = \max(t_k - t_{k-1})$.

5. *Пусть* $\nu(t) \in \mathcal{M}_2^l$ *и* $\langle\nu\rangle_t$ — *непрерывная функция. Тогда*

$$\langle\nu\rangle_t = \lim_{\lambda\to 0}\sum_{k=1}^n \mathsf{M}\left([\nu(t_k) - \nu(t_{k-1})]^2 \mid \mathfrak{F}_{t_{k-1}}\right).$$

6. *Пусть* $\nu(t) \in \mathcal{M}_2^l$, $\nu(0) = 0$ *и* $\langle\nu\rangle_t$ — *непрерывная функция. Тогда для* $\varepsilon > 0$ *и* $c > 0$

$$\mathsf{P}\left\{\sup_{0\leqslant s\leqslant t}|\nu(s)| \geqslant \varepsilon\right\} \leqslant \frac{c}{\varepsilon^2} + \mathsf{P}\{\langle\nu\rangle_t \geqslant c\}.$$

7. *Пусть* $\nu(t) \in \mathcal{M}_2^l$, $\nu(0) = 0$, *и* $\langle\nu\rangle_t$ — *непрерывная функция. Тогда для всякого момента остановки* τ

а) $\mathsf{M}\sup\limits_{0\leqslant s\leqslant\tau}|\nu(s)| \leqslant 3\mathsf{M}\langle\nu\rangle_\tau^{1/2}$;

б) *для всех* $\alpha \in (0, 2)$ *существует постоянная* c_α *такая, что*

$$\mathsf{M}\sup_{0\leqslant s\leqslant\tau}|\nu(s)|^\alpha \leqslant c_\alpha \mathsf{M}\langle\nu\rangle_\tau^{\alpha/2};$$

в) *для* $\alpha > 1$

$$\mathsf{M}\sup_{0\leqslant s\leqslant\tau}|\nu(s)|^\alpha \leqslant \mathsf{M}\left[3\langle\nu\rangle_\tau^{1/2}+\sup_{s\leqslant\tau}|\nu(s)-\nu(s-)|\right]^\alpha;$$

г) *если* $\sup\limits_s|\nu(s)-\nu(s-)|\leqslant a$, *то при* $\alpha\geqslant 1$

$$\mathsf{M}\sup_{0\leqslant s\leqslant\tau}|\nu(s)|^\alpha \leqslant \mathsf{M}\left[3\left\langle\nu_\tau\right\rangle^{1/2}+a\right]^\alpha.$$

Теорема Леви. *Пусть* $\xi(t)$ *является непрерывным локальным мартингалом с характеристикой* $\langle\xi\rangle_t = t$, *тогда* $\xi(t)$ — *винеровский процесс, т. е. однородный процесс с независимыми приращениями, для которого* $\xi(t+h)-\xi(t)$ *имеет нормальное распределение со средним* 0 *и дисперсией* h.

13.4.3. Стохастические интегралы по мартингалам. Пусть $\eta(t)\in\mathcal{M}_2^l$, $\varphi(t)$ — $\mathscr{P}$-измеримая (предсказуемая) функция, для которой $\int_0^t\varphi^2(s)\,d\langle\eta\rangle_s<\infty$ $\forall t$. Тогда определен стохастический интеграл

$$J_t(\varphi,\eta)=\int_0^t\varphi(s)\,d\eta(s),$$

удовлетворяющий следующим условиям:

1) если $\Phi_2\langle\eta\rangle$ обозначает указанное множество функций $\varphi(t)$, то с вероятностью 1

$$J_t(\alpha\varphi_1+\beta\varphi_2,\eta)=\alpha J_t(\varphi_1,\eta)+\beta J_t(\varphi_2,\eta),\quad \varphi_1,\varphi_2\in\Phi_2\langle\eta\rangle;$$

2) $J_t(\varphi,\eta)\in\mathcal{M}_2^l$ и при $\varphi_1,\varphi_2\in\Phi_2\langle\eta\rangle$

$$\langle J(\varphi_1,\eta),J(\varphi_2,\eta)\rangle_t=\int_0^t\varphi_1(s)\varphi_2(s)\,d\langle\eta\rangle_s;$$

3) если $\varphi(t)=I_{[t_1,t_2]}(t)\varphi_1$, где φ_1 — $\mathfrak{F}_{t_1}$-измерима, то

$$\int_0^t I_{[t_1,t_2]}(s)\varphi_1\,d\eta(s)=\varphi_1[\eta(t\wedge t_2)-\eta(t\wedge t_1)].$$

Эти свойства определяют стохастический интеграл однозначно с точностью до неотличимости (два процесса ξ_1 и ξ_2 *неотличимы*, если $\mathsf{P}\left\{\sup\limits_t|\xi_1(t)-\xi_2(t)|=0\right\}=1$).

Свойства стохастических интегралов.

1) Пусть $\eta_1,\eta_2\in\mathcal{M}_2^l$ и $\varphi_k\in\Phi_2\left\langle\eta_k\right\rangle$ $(k=1,2)$. Тогда

$$\langle J(\varphi_1,\eta_1),J(\varphi_2,\eta_2)\rangle_t=\int_0^t\varphi_1(s)\varphi_2(s)\,d\langle\eta_1,\eta_2\rangle_s.$$

2) Для произвольных η_1 и $\eta_2 \in \mathcal{M}_2^l$ справедливо представление

$$\eta_2(t) = \int_0^t \alpha(s)\, d\eta_1(s) + \tilde{\eta}_2(t),$$

где $\tilde{\eta}_2 \perp \eta_1$, а $\alpha(s) \in \Phi_2(\langle\eta_1\rangle)$, $\alpha(s) = \frac{d\langle\eta_1, \eta_2\rangle_s}{d\langle\eta_1, \eta_1\rangle_s}$.

3) Пусть $\eta \in \mathcal{M}_2^l$, $\langle\eta\rangle_t$ непрерывно и $\varphi \in \Phi_2\langle\eta\rangle$. Тогда

а) $\mathsf{M} J_t^2(\varphi, \eta) = \mathsf{M} \int_0^t \varphi^2(s)\, d\langle\eta\rangle_s$;

б) $\mathsf{M} \sup\limits_{s \leq t} |J_t(\varphi, \eta)|^2 \leq 4\mathsf{M} \int_0^t \varphi^2(s)\, d\langle\eta\rangle_s$;

в) для всех $\alpha \in (0, 2)$ существует c_α такая, что

$$\mathsf{M} \sup_{s<t} |J_s(\varphi, \eta)|^\alpha \leq c_\alpha \mathsf{M} \left(\int_0^t \varphi^2(s)\, d\langle\eta\rangle_s \right)^{\alpha/2}.$$

13.4.4. Случайные меры и стохастические интегралы. Пусть $(\Theta, \mathfrak{C})$ — некоторое измеримое пространство и $\mathfrak{B}_+$ — σ-алгебра борелевских множеств в $\mathbf{R}_+$. Будем рассматривать случайные меры, определенные на $\mathfrak{C} \otimes \mathfrak{B}_+$. Нас будут интересовать два вида мер: мартингальные и целочисленные. Будем говорить, что на $\mathfrak{C} \otimes \mathfrak{B}_+$ задана мартингальная мера μ, если для всякого множества $C \in \mathfrak{C} \otimes \mathfrak{B}_+$ задана случайная величина $\mu(C)$ так, что выполнены следующие условия:

1) $\mu(C \cap \Theta \times [0, t]) = \mu_t(C)$ является $\mathfrak{F}_t$-локально квадратично интегрируемым мартингалом;

2) если $C_1 \cap C_2 = \varnothing$, C_1, $C_2 \in \mathfrak{C} \otimes \mathfrak{B}$, то $\mu(C_1 \cup C_2) = \mu(C_1) + \mu(C_2)$ и локальные мартингалы $\mu_t(C_1)$ и $\mu_t(C_2)$ ортогональны;

3) $\mu(\varnothing) = 0$;

4) для всякой последовательности попарно непересекающихся множеств $C_k \in \mathfrak{C} \otimes \mathfrak{B}_+$

$$\mu\left(\bigcup_k C_k\right) = \sum_k \mu(C_k)$$

с вероятностью 1; ряд справа сходится по вероятности.

Для мартингальной меры μ существует ее (квадратическая) характеристика $\langle\mu\rangle(C)$, определенная на $\mathfrak{C} \otimes \mathfrak{B}_+$ и удовлетворяющая следующим условиям:

1′) $\langle\mu\rangle(C \cap \Theta \times [0, t]) = \langle\mu\rangle_t(C)$ — возрастающий предсказуемый процесс;

2′) если C_k — последовательность попарно непересекающихся $C_k \in \mathfrak{C} \otimes \mathfrak{B}_+$, то $\langle\mu\rangle\left(\bigcup_k C_k\right) = \sum_k \langle\mu\rangle(C_k)$; ряд справа сходится с вероятностью 1.

Пусть $\Phi_2(\langle\mu\rangle)$ — совокупность функций $\varphi(\theta, t, \omega)$, измеримых относительно $\mathfrak{L}\otimes\mathscr{P}$ ($\mathscr{P}$-предсказуемая σ-алгебра), для которых $\int \varphi^2(\theta, t, \omega)\langle\mu\rangle(d\theta\times dt) < \infty$.

Интеграл $\int \Psi(\theta, t, \omega)\langle\mu\rangle(d\theta\times dt)$, фигурирующий в определении $\Phi_2(\langle\mu\rangle)$, где $\Psi(\theta, t, \omega)$ — $\mathfrak{L}\otimes\mathscr{P}$-измеримая неотрицательная функция, определяется так, чтобы выполнялись свойства: а) аддитивность по Ψ и неотрицательность; б) при $\Psi(\theta, t, \omega) = \Psi_1 I_B I_{[t_1, t_2]}$, где Ψ_1 $\mathfrak{F}_{t_1}$-измеримо и неотрицательно,

$$\int \Psi(\theta, t, \omega)\langle\mu\rangle(d\theta\times dt) = \Psi_1\langle\mu\rangle(B\times[t_1, t_2]).$$

Для всех $\varphi\in\Phi_2(\langle\mu\rangle)$ можно определить стохастический интеграл

$$J_t(\varphi, \mu) = \int_0^t \int_\Theta \varphi(\theta, s, \omega)\,\mu(d\theta\times ds)$$

так, что выполняются условия:

1) он линеен по φ;

2) $J_t(\varphi, \mu)\in\mathscr{M}_2^l$ и при $\varphi_1, \varphi_2\in\Phi_2(\langle\mu\rangle)$

$$\langle J_t(\varphi_1, \mu), J_t(\varphi_2, \mu)\rangle = \int_0^t \int_\Theta \varphi_1(\theta, s, \omega)\,\varphi_2(\theta, s, \omega)\langle\mu\rangle(d\theta\times ds);$$

3) если $\varphi(\theta, s, \omega) = I_B I_{[t_1, t_2]}\varphi_1$, где $B\in\mathfrak{L}$, и φ_1 $\mathfrak{F}_{t_1-}$-измеримо, то $J_t(\varphi, \mu) = \varphi_1\mu(B\times[t_1, t_2]\cap[0, t])$.

Свойства 1)—3) определяют однозначно интеграл по μ (с точностью до неотличимости).

Случайная мера ν на $\mathfrak{L}\otimes\mathfrak{B}_+$ называется *целочисленной*, если $\nu(C)$ принимает целые неотрицательные значения и $\nu(\Theta\times\{t\}) \leqslant 1$ для всех $t\in R_+$. Для всякой целочисленной случайной меры ν существуют последовательности моментов остановки τ_k и Θ-значных $\mathfrak{F}_{\tau_k}$-измеримых величин θ_k такие, что

$$\nu(C) = \sum_k I_{\{(\theta_k;\,\tau_k)\in C\}}.$$

Интеграл по мере ν определен для всякой $\mathfrak{L}\otimes\mathfrak{B}_+\otimes\mathfrak{F}$-измеримой неотрицательной функции $f(\theta, t, \omega)$ с помощью формулы

$$\int f(\theta, t, \omega)\,\nu(d\theta\times dt) = \sum_k f(\theta_k, \tau_k, \omega)$$

и естественным образом распространяется на знакопеременные функции, если $\sum_k |f(\theta_k, \tau_k, \omega)| < \infty$.

Так как $\nu(C\cap\Theta\times[0, t])$ — непрерывный справа возрастающий согласованный процесс, все скачки которого не превосходят 1, то существует компенсатор этого процесса, который мы обозначим $\bar{\nu}(C\cap\Theta\times[0, t])$.

Случайная мера $\bar{\nu}(C)=\bar{\nu}(C\cap\Theta\times\mathbf{R})$ называется *компенсатором* целочисленной меры ν. Тогда $\hat{\nu}(C)=\nu(C)-\bar{\nu}(C)$ является мартингальной мерой, характеристика которой также совпадает с $\bar{\nu}(C)$, т. е. $\langle\bar{\nu}\rangle=\bar{\nu}$, $\hat{\nu}$ называется *компенсированной целочисленной мерой*. Можно убедиться, что для $\mathfrak{L}\otimes\mathcal{P}$-измеримых функций $f(\theta, t, \omega)$

$$\int f(\theta, t, \omega)\,\nu(d\theta\times dt)=\int f(\theta, t, \omega)\,\hat{\nu}(d\theta\times dt)+\\ +\int f(\theta, t, \omega)\,\langle\hat{\nu}\rangle(d\theta\times dt)$$

при условии, что все интегралы в этом равенстве определены.

При рассмотрении целочисленных мер очень важную роль играет следующая теорема.

Теорема 1. *Пусть ν — целочисленная случайная мера, компенсатор которой $\bar{\nu}$ неслучаен. Тогда ν — пуассоновская мера с независимыми значениями, т. е. $\nu(C_1), \ldots, \nu(C_k)$ независимы, если $C_1, \ldots, C_k$ попарно не пересекаются, и $\nu(C)$ имеет распределение Пуассона с параметром $\bar{\nu}(C)$.*

13.4.5. Точечные процессы. *Точечным процессом* называется последовательность моментов остановки τ_n ($n\geqslant 1$) (относительно фиксированного потока $(\mathfrak{F}_t)_{t\geqslant 0}$), удовлетворяющая двум требованиям: а) если $\tau_n<\infty$, то $\tau_{n+1}>\tau_n$; б) для всякого t найдется такое n, что $\tau_n>t$.

Обычно точечный процесс задается последовательностью величин τ_n, удовлетворяющих условиям а) и б) без требования, чтобы они были моментами остановки, и без фиксации определенного потока. Если такая последовательность задана, то наименьший поток, относительно которого τ_n будут моментами остановки, есть поток, для которого $\mathfrak{F}_t$ — σ-алгебра, порожденная событиями $\{\tau_n\in C\}$ $(n\geqslant 1)$, C — борелевское множество отрезка $[0, t]$.

Со всяким точечным процессом можно связать целочисленную меру на $\mathbf{R}_+$

$$\nu(C)=\sum_k I_{\{\tau_k\in C\}},$$

которая в силу условия б) конечна для всякого ограниченного множества C. Процесс $\nu([0, t])$ является возрастающим. Его компенсатор относительно минимального потока $(\mathfrak{F}_t)_{t\geqslant 0}$, порожденного этим процессом, определяется равенством

$$\bar{\nu}([0, t])=\sum_{k=1}^{\infty}\ln\frac{1}{1-F_k(t\wedge\tau_k\mid\tau_1, \ldots, \tau_{k-1})}I_{\{\tau_{k-1}\leqslant t\}},$$

где $F_k(s\mid\tau_1, \ldots, \tau_{k-1})$ — условная функция распределения величины τ_k.

Маркированным точечным процессом называется последовательность пар $(\tau_1; \theta_1)$, $(\tau_2; \theta_2)$, ..., где τ_n — последовательность моментов остановки, удовлетворяющая условиям а) и б), а θ_k — $\mathfrak{F}_{\tau_k}$-измеримые величины в некотором измеримом пространстве $(\Theta, \mathfrak{L})$.

Если имеется произвольная последовательность $(\tau_n; \theta_n)_{n \geqslant 1}$, то поток σ-алгебр $(\mathfrak{F}_t)_{t \geqslant 0}$, относительно которого эта последовательность будет маркированным точечным процессом, можно указать, полагая, что σ-алгебра $\mathfrak{F}_t$ порождена событиями вида $\{\tau_k \in B, \theta_k \in C\}$ $(k = 1, 2, \ldots)$, где $B \subset [0, t]$ — борелевское множество, $C \in \mathfrak{L}$.

С маркированным точечным процессом можно связать целочисленную случайную меру на $\mathfrak{L} \otimes \mathfrak{B}_+$ с помощью равенства

$$\nu(C) = \sum_k I_{\{(\theta_k;\, \tau_k) \in C\}}.$$

Компенсатор $\tilde{\nu}$ этой меры определяется равенством: для $C_1 \in \mathfrak{L}$

$$\tilde{\nu}(C_1 \times [0, t]) = \sum_k I_{\{\tau_{k-1} \leqslant t\}} \int_0^{t \wedge \tau_k} \frac{\mathsf{P}\{\theta_k \in C_1, \tau_k \in ds \mid \tau_1, \ldots, \tau_{k-1}, \theta_1, \ldots, \theta_{k-1}\}}{\mathsf{P}\{\tau_k > s \mid \tau_1, \ldots, \tau_{k-1}, \theta_1, \ldots, \theta_{k-1}\}};$$

под знаком интеграла стоят условные распределения при заданных $\tau_1, \ldots, \tau_{k-1}, \theta_1, \ldots, \theta_{k-1}$.

13.4.6. Представление мартингалов. Формула Ито. Пусть $\xi(t)$ — локально квадратично интегрируемый мартингал. Свяжем с ним целочисленную случайную меру на $\mathfrak{L} \otimes \mathfrak{B}_+$, где $\mathfrak{B}$ — σ-алгебра в $\mathbf{R}$, определяемую равенством

$$\nu(C) = \sum I_{\{\xi(s) \neq \xi(s-),\, (\xi(s) - \xi(s-);\, s) \in C\}}.$$

Это мера для множеств $C = \{x\colon |x| > \varepsilon\} \times [0, t]$, а значит, и для всякого измеримого подмножества этого множества. Если $\tilde{\nu}(C)$ — компенсатор этой меры, то

$$\xi(t) = \xi_0(t) + \int_{|x| \leqslant \varepsilon} x\,[\nu(dx \times [0, t]) - \tilde{\nu}(dx \times [0, t])] + \int_{|x| > \varepsilon} x\nu(dx \times [0, t]) + \alpha(t),$$

где $\xi_0(t)$ — непрерывный мартингал, $\alpha(t)$ — предсказуемый процесс ограниченной вариации.

Теорема 2 (обобщенные формулы Ито). *Пусть $\xi(t)$ — процесс, представимый в виде*

$$\xi(t) = \alpha(t) + \eta(t) + \int_0^t \int_{\Theta_1} \varphi(\theta, s, \omega)\,\hat{\nu}(d\theta \times ds) + \int_0^t \int_{\Theta_2} \varphi(\theta, s, \omega)\,\nu(d\theta \times ds),$$

где $\alpha(t)$ — *непрерывный процесс ограниченной вариации;* $\eta(t)$ — *непрерывный мартингал с характеристикой* $\langle\eta\rangle_t$; ν — *целочисленная мера на* $\Theta\times R_+$; $\bar\nu$ — *ее компенсатор, для которого* $\bar\nu(C\times[0,t])$ — *непрерывная функция* t, *если только она конечна;* $\Theta_1\cup\Theta_2=\Theta$, $\Theta_k\in\mathfrak{L}$, $\Theta_1\cap\Theta_2=\varnothing$, *а* $\varphi(\theta,s,\omega)$ — $\mathfrak{L}\otimes\mathscr{P}$-*измерима и*

$$\int_0^t\int_{\Theta_1}\varphi^2(\theta,s,\omega)\,\bar\nu(d\theta\times ds)<\infty\quad\forall t>0,\ \nu(\Theta_2\times[0,t])<\infty.$$

Если $g(x)$ — *дважды непрерывно дифференцируемая функция из* $\mathbf{R}$ *в* $\mathbf{R}$, *то*

$$g(\xi(t))=g(\xi(0))+\int_0^t g'(\xi(s))\,d\alpha(s)+\frac{1}{2}\int_0^t g''(\xi(s))\,d\langle\eta\rangle_s+$$
$$+\int_0^t\int_{\Theta_1}[g(\xi(s-)+\varphi(\theta,s,\omega))-g(\xi(s-))-$$
$$-g'(\xi(s-))\,\varphi(\theta,s,\omega)]\,\bar\nu(d\theta\times ds)+$$
$$+\int_0^t g'(\xi(s))\,d\eta(s)+\int_0^t\int_{\Theta_1}[g(\xi(s-)+\varphi(\theta,s,\omega))-g(\xi(s-))]\times$$
$$\times\hat\nu(d\theta\times ds)+\int_0^t\int_{\Theta_2}[g(\xi(s-)+\varphi(\theta,s,\omega))-g(\xi(s-))]\,\nu(d\theta\times ds).$$

Теорема 3. *Пусть* $\xi_k(t)$ $(k=1,\ldots,m)$ — *случайные процессы, представимые в виде*

$$\xi_k(t)=\alpha_k(t)+\eta_k(t)+$$
$$+\int_0^t\int_{\Theta_1}\varphi_k(\theta,s,\omega)\,\hat\nu(d\theta\times ds)+\int_0^t\int_{\Theta_2}\varphi_k(\theta,s,\omega)\,\nu(d\theta\times ds),$$

где $\alpha_k(t)$ — *непрерывный процесс ограниченной вариации;* $\eta_k(t)$ — *непрерывные мартингалы с непрерывными взаимными характеристиками* $\langle\eta_k,\eta_l\rangle$; ν, Θ_1, Θ_2 *такие же, как и в теореме* 1; *функции* $\varphi_k(\theta,s,\omega)$ $\mathfrak{L}\otimes\mathscr{P}$-*измеримы и для всех* t

$$\int_0^t\int_{\Theta_1}\sum_{k=1}^m\varphi_k^2(\theta,s,\omega)\,\bar\nu(d\theta\times ds)<\infty.$$

Если $g(x_1,\ldots,x_m)$ — *дважды непрерывно дифференцируемая функция из* $\mathbf{R}^m$ *в* $\mathbf{R}$, *то*

$$g(\xi_1(t), \ldots, \xi_m(t)) = g(\xi_1(0), \ldots, \xi_m(0)) +$$
$$+ \int_0^t \sum_k g'_{x_k}\big(\xi_1(s), \ldots, \xi_m(s)\big)\, d\alpha_k(s) +$$
$$+ \frac{1}{2} \sum_{k,\, l=1}^{m} \int_0^t g''_{x_k x_l}\big(\xi_1(s), \ldots, \xi_m(s)\big)\, d\big\langle \xi_k, \xi_l \big\rangle_s +$$
$$+ \int_0^t \int_{\Theta_1} \Big[g\big(\xi_1(s) + \varphi_1(\theta, s, \omega), \ldots, \xi_m(s) + \varphi_m(\theta, s, \omega)\big) -$$
$$- \varphi\big(\xi_1(s), \ldots, \xi_m(s)\big) - \sum_{r=1}^{m} g'_{x_r}\big(\xi_1(s), \ldots, \xi_m(s)\big)\, \varphi_r(\theta, s, \omega) \Big] \times$$
$$\times \tilde{\nu}(d\theta \times ds) + \sum_{k=1}^{m} \int_0^t g'_{x_k}(\xi_1(s-), \ldots, \xi_m(s-))\, d\eta_k(s) +$$
$$+ \int_0^t \int_{\Theta_1} \big[g\big(\xi_1(s-) + \varphi_1(\theta, s, \omega), \ldots, \xi_m(s-) + \varphi_m(\theta, s, \omega)\big) -$$
$$- g\big(\xi_1(s-), \ldots, \xi_m(s-)\big)\big]\, \hat{\nu}(d\theta \times d\bar{s}) + \int_0^t \int_{\Theta_2} \big[g\big(\xi_1(s-) +$$
$$+ \varphi_1(\theta, s, \omega)\big), \ldots, \xi_m(s-) + \varphi_m(\theta, s, \omega)\big) -$$
$$- g\big(\xi_1(s-), \ldots, \xi_m(s-)\big)\big]\, \nu(d\theta \times ds).$$

Отметим специально случай непрерывных процессов. Если $\xi_k(t) = \alpha_k(t) + \eta_k(t)$, где α_k, η_k и $g(x_1, \ldots, x_m)$ такие же, как и в теореме 2, тогда

$$g\big(\xi_1(t), \ldots, \xi_m(t)\big) =$$
$$= g\big(\xi_1(0), \ldots, \xi_m(0)\big) + \sum_{k=1}^{m} \int_0^t g'_{x_k}\big(\xi_1(s), \ldots, \xi_m(s)\big)\, d\alpha_k(s) +$$
$$+ \frac{1}{2} \sum_{k,\, l=1}^{m} \int_0^t g''_{x_k x_l}\big(\xi_1(s), \ldots, \xi_m(s)\big)\, d\big\langle \eta_k, \eta_l \big\rangle_s +$$
$$+ \sum_{k=1}^{m} \int_0^t g'_{x_k}\big(\xi_1(s-), \ldots, \xi_m(s-)\big)\, d\eta_k(s).$$

Литература: [18, 26, 27, 28, 38, 59, 62, 111, 116].

Глава 14. МАРКОВСКИЕ ПРОЦЕССЫ

14.1. Марковские случайные функции

14.1.1. Марковское свойство. В основе понятия марковского процесса лежит представление об эволюционирующих во времени стохастических системах, обладающих свойством «отсутствия последействия» («отсутствия памяти»). Такого рода процессы с дискретным временем, называемые цепями Маркова, рассматривались в гл. 8. В настоящей главе рассмотрены процессы Маркова с непрерывным временем. Мы будем предполагать, что время изменяется в некотором отрезке (интервале, полуинтервале) $\mathcal{T}$, содержащемся во множестве всех действительных неотрицательных чисел. В частности, возможно $\mathcal{T} = [0, \infty)$.

Пусть заданы:

а) некоторое вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$;

б) поток σ-алгебр $\{\mathfrak{F}_t, t \in \mathcal{T}\}$, т. е. такое семейство σ-алгебр $\mathfrak{F}_t$, $t \in \mathcal{T}$, что $\mathfrak{F}_t \subset \mathfrak{F}$ и $\mathfrak{F}_s \subset \mathfrak{F}_t$ при $s \leqslant t$, $s, t \in \mathcal{T}$;

в) функция двух переменных $\xi(t) = \xi(t, \omega)$, $t \in \mathcal{T}$, $\omega \in \Omega$, со значениями в некотором измеримом пространстве $(X, \mathfrak{B})$ такая, что $\xi(t)$ при каждом $t \in \mathcal{T}$ является измеримым отображением пространства $(\Omega, \mathfrak{F}_t)$ в $(X, \mathfrak{B})$.

Тем самым задан случайный процесс $\xi(t)$, подчиненный потоку σ-алгебр $\{\mathfrak{F}_t, t \in \mathcal{T}\}$. В частности, σ-алгебра $\mathfrak{F}_t$ может совпадать с минимальной σ-алгеброй событий, порожденной всеми событиями вида $\{\omega: \xi(s) \in \Gamma\}$, $\Gamma \in \mathfrak{B}$, $s \in \mathcal{T}$, $s \leqslant t$. В общем же случае — это более широкая σ-алгебра.

Случайный процесс $\xi(t)$, подчиненный потоку σ-алгебр $\mathfrak{F}_t$, обладает *марковским свойством* относительно этого потока, если при всех $s \leqslant t$ $(s, t \in \mathcal{T})$, $\Gamma \in \mathfrak{B}$ с вероятностью 1 выполняется соотношение

$$\mathsf{P}\{\xi(t) \in \Gamma \mid \mathfrak{F}_s\} = \mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}.$$

Такой процесс будем называть *марковской случайной функцией*. Термин «марковский процесс» зарезервируем для другого, более удобного (с точки зрения изучения марковского свойства) понятия, связанного с целым семейством марковских случайных функций (см. также гл. 8).

Если интерпретировать $\xi(t)$ как состояние (положение) некоторой системы (частицы) в момент времени t, то марковское свойство означает, что такая система обладает свойством отсутствия последействия: при прогнозировании (в среднем) поведения системы в «будущие» моменты времени по наблюдениям за системой во все «прошедшие» моменты времени вплоть до «настоящего» существенным является знание положения рассматриваемой системы лишь в «настоящий» момент времени.

Пусть $\{\xi(t), t \in \mathcal{T}\}$ — случайный процесс со значениями в некотором измеримом пространстве $(X, \mathfrak{B})$, подчиненный потоку σ-алгебр $\{\mathfrak{F}_t, t \in \mathcal{T}\}$. Обозначим через $\mathfrak{N}_t$ минимальную σ-алгебру событий, порожденную событиями вида $\{\omega: \xi(s) \in \Gamma\}$, $s \leqslant t$ $(s, t \in \mathcal{T})$, $\Gamma \in \mathfrak{B}$, а через $\mathfrak{N}^t$ минимальную σ-алгебру событий, содержащую все события вида $\{\omega: \xi(s) \in \Gamma\}$, $s \geqslant t$ $(s, t \in \mathcal{T})$, $\Gamma \in \mathfrak{B}$. Заметим, что $\mathfrak{N}_t \subset \mathfrak{F}_t$, и если $\xi(t)$ обладает марковским свойством относительно потока $\{\mathfrak{F}_t,\ t \in \mathcal{T}\}$, то $\xi(t)$ обладает тем же свойством относительно потока $\{\mathfrak{N}_t,\ t \in \mathcal{T}\}$.

Марковское свойство процесса $\{\xi(t),\ t \in \mathcal{T}\}$ относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in \mathcal{T}\}$ эквивалентно любому из следующих свойств:

1) для любой ограниченной $\mathfrak{B}$-измеримой функции $f(x)$ $(x \in X)$ и любых $s \leqslant t$ $(s,\ t \in \mathcal{T})$

$$\mathsf{M}\{f(\xi(t)) \mid \mathfrak{F}_s\} = \mathsf{M}\{f(\xi(t)) \mid \xi(s)\} \quad \text{п. н. } \mathsf{P};$$

2) для любой ограниченной $\mathfrak{N}^t$-измеримой случайной величины η и любых $s \leqslant t$ $(s,\ t \in \mathcal{T})$

$$\mathsf{M}\{\eta \mid \mathfrak{F}_s\} = \mathsf{M}\{\eta \mid \xi(s)\} \quad \text{п. н. } \mathsf{P};$$

3) для любых событий $A \in \mathfrak{N}^t$ и $B \in \mathfrak{F}_t$

$$\mathsf{P}\{A \cap B \mid \xi(t)\} = \mathsf{P}\{A \mid \xi(t)\}\, \mathsf{P}\{B \mid \xi(t)\} \quad \text{п. н. } \mathsf{P}.$$

Последнее свойство означает, что для процесса, обладающего марковским свойством, события из «будущего» и «прошлого» при фиксированном «настоящем» условно независимы.

14.1.2. Вероятность перехода. Пусть $\{\xi(t),\ t \in \mathcal{T}\}$ — марковская случайная функция со значениями в $(X,\ \mathfrak{B})$ относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in \mathcal{T}\}$. Следствием марковского свойства и формулы полной вероятности является следующее соотношение:

$$\mathsf{P}\{\xi(t_3) \in \Gamma \mid \xi(t_1)\} = \mathsf{M}\{\mathsf{P}\{\xi(t_3) \in \Gamma \mid \xi(t_2)\} \mid \xi(t_1)\},$$

справедливое почти наверное относительно меры P при всех $\Gamma \in \mathfrak{B}$ и $t_1 \leqslant t_2 \leqslant t_3$ $(t_1,\ t_2,\ t_3 \in \mathcal{T})$. Это соотношение называется *уравнением Колмогорова — Чепмена.*

Особый интерес представляет случай, когда для условной вероятности $\mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}$ существует регулярная условная вероятность, т. е. такая функция $P(s,\ x,\ t,\ \Gamma)$, $x \in X$, $\Gamma \in \mathfrak{B}$, $s < t$ $(s,\ t \in \mathcal{T})$, что выполнены условия:

а) при фиксированных s, x, t функция $P(s,\ x,\ t,\ \Gamma)$ является вероятностной мерой на $(X,\ \mathfrak{B})$;

б) при фиксированных s, t, Γ функция $P(s,\ x,\ t,\ \Gamma)$ $\mathfrak{B}$-измерима;

в) с вероятностью 1 при всех s, t, Γ

$$P(s,\ \xi(s),\ t,\ \Gamma) = \mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}.$$

Если для данной марковской случайной функции существует функция $P(s,\ x,\ t,\ \Gamma)$, удовлетворяющая условиям а) — в), то она называется *вероятностью перехода.*

В терминах вероятности перехода уравнение Колмогорова — Чепмена запишется в виде

$$P(t_1,\ \xi(t_1),\ t_3,\ \Gamma) = \int_X P(t_2,\ y,\ t_3,\ \Gamma)\, P(t_1,\ \xi(t_1),\ t_2,\ dy) \quad \text{п. н. } \mathsf{P},$$

где $t_1 < t_2 < t_3$ $(t_1,\ t_2,\ t_3 \in \mathcal{T})$, $\Gamma \in \mathfrak{B}$.

Мы всегда будем предполагать выполненным несколько более сильное соотношение для вероятности перехода, называемое также уравнением Колмогорова — Чепмена. Именно будем предполагать, что при всех $x \in X$, $\Gamma \in \mathfrak{B}$ и $t_1 < t_2 < t_3$ (t_1, t_2, $t_3 \in \mathcal{T}$)

$$P(t_1, x, t_3, \Gamma) = \int_X P(t_2, y, t_3, \Gamma)\, P(t_1, x, t_2, dy).$$

При весьма широких предположениях о фазовом пространстве марковской случайной функции (об измеримом пространстве $(X, \mathfrak{B})$) вероятность перехода существует и удовлетворяет последнему соотношению. Эти предположения сводятся к тому, чтобы σ-алгебра $\mathfrak{B}$ не была слишком богатой. Именно, если марковская случайная функция принимает значения в фазовом пространстве $(X, \mathfrak{B})$, изоморфном борелевскому подмножеству некоторого полного сепарабельного метрического пространства с σ-алгеброй его борелевских подмножеств, то для такой марковской случайной функции вероятность перехода существует и удовлетворяет уравнению Колмогорова — Чепмена при всех $x \in X$.

14.1.3. Стохастическая эквивалентность. Конечномерные распределения марковской случайной функции не определяются одной лишь вероятностью перехода. Если в $\mathcal{T}$ существует минимальный элемент t_0, то, зная начальное распределение

$$\mu(\Gamma) = \mathsf{P}\{\xi(t_0) \in \Gamma\}, \quad \Gamma \in \mathfrak{B},$$

и вероятность перехода $P(s, x, t, \Gamma)$ марковской случайной функции, можем определить все ее конечномерные распределения. Действительно, для $t_0 < t_1 < \ldots < t_n$, $\Gamma_0, \Gamma_1, \ldots, \Gamma_n \in \mathfrak{B}$ имеем

$$\mathsf{P}\{\xi(t_0) \in \Gamma_0, \xi(t_1) \in \Gamma_1, \ldots, \xi(t_n) \in \Gamma_n\} =$$
$$= \int_{\Gamma_0} \mu(dx_0) \int_{\Gamma_1} P(t_0, x_0, t_1, dx_1) \ldots \int_{\Gamma_n} P(t_{n-1}, x_{n-1}, t_n, dx_n).$$

Таким образом, если для двух марковских случайных функций, заданных, возможно, на разных вероятностных пространствах, совпадают начальные распределения и вероятности перехода, то у них совпадают и все конечномерные распределения. Это означает, что такие две марковские случайные функции стохастически эквивалентны.

Если в Г существует минимальный элемент t_0 и если заданы вероятностная мера μ, определенная на $(X, \mathfrak{B})$, и функция $P(s, x, t, \Gamma)$, $t_0 \leqslant s < t$, s, $t \in \mathcal{T}$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяющая условиям а), б) п. 14.1.2 и уравнению Колмогорова — Чепмена (см. последнее соотношение в п. 14.1.2), то всегда можно построить на некотором вероятностном пространстве марковскую случайную функцию, для которой мера μ была бы начальным распределением, а функция $P(s, x, t, \Gamma)$ — вероятностью перехода.

Предположим теперь, что в $\mathcal{T}$ не существует минимального элемента. Если $\{\xi(t), t \in \mathcal{T}\}$ — некоторая марковская случайная функция, то положим

$$\mu_t(\Gamma) = \mathsf{P}\{\xi(t) \in \Gamma\}, \quad t \in \mathcal{T}, \quad \Gamma \in \mathfrak{B}.$$

Вероятностная мера μ_t на $(X, \mathfrak{B})$ называется *законом входа рассматриваемой марковской случайной функции.* Закон входа связан с вероятностью перехода следующим соотношением:

$$\mu_t(\Gamma) = \int_X \mu_s(dx)\, P(s, x, t, \Gamma), \quad s < t \ (s, t \in \mathcal{T}), \quad \Gamma \in \mathfrak{B}. \quad (1.1)$$

Зная закон входа и вероятность перехода марковской случайной функции, можем определить все ее конечномерные распределения

$$\mathsf{P}\{\xi(t_1) \in \Gamma_1, \ldots, \xi(t_n) \in \Gamma_n\} = \\ = \int_{\Gamma_1} \mu_{t_1}(dx_1) \int_{\Gamma_2} P(t_1, x_1, t_2, dx_2) \ldots \int_{\Gamma_n} P(t_{n-1}, x_{n-1}, t_n, dx_n),$$

где $t_1 < t_2 < \ldots < t_n$ $(t_1, t_2, \ldots, t_n \in \mathcal{T})$, $\Gamma_1, \ldots, \Gamma_n \in \mathfrak{B}$.

Таким образом, марковская случайная функция в рассматриваемом случае определяется своими законом входа и вероятностью перехода однозначно с точностью до стохастической эквивалентности.

Обратно, если заданы семейство вероятностных мер $\mu_t(\Gamma)$, $t \in \mathcal{T}$, $\Gamma \in \mathfrak{B}$, и функция $P(s, x, t, \Gamma)$, $s < t$ $(s, t \in \mathcal{T})$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяющая условиям а), б) п. 14.1.2 и уравнению Колмогорова — Чепмена, такие, что выполнено соотношение (1.1), то на некотором вероятностном пространстве существует марковская случайная функция, для которой закон входа совпадает с μ_t, а вероятность перехода — с $P(s, x, t, \Gamma)$.

14.1.4. Обрывающиеся марковские случайные функции. Иногда приходится иметь дело с системами, для описания которых приведенное определение марковской случайной функции недостаточно.

Пусть, например, $\xi(t)$ означает число особей, имеющихся к моменту времени t в некоторой биологической популяции. Тогда $\xi(t)$ — случайный процесс, фазовым пространством которого служат все натуральные числа. Может оказаться, что интенсивность размножений в данной популяции столь велика, что по истечении некоторого конечного (вообще говоря, случайного) времени число особей в рассматриваемой популяции становится бесконечно большим. Таким образом, мы здесь сталкиваемся с ситуацией, в которой процесс $\xi(t)$ определен лишь на некотором случайном промежутке времени, по истечении которого происходит исчезновение процесса из фазового пространства (обрыв, взрыв, гибель). Приведенное ниже определение марковской случайной функции учитывает такую возможность.

Будем считать, что $\mathcal{T} = [t_0, \infty]$. Пусть даны:

а) вероятностное пространство $(\Omega, \mathfrak{A}, \mathsf{P})$;

б) случайная величина $\zeta = \zeta(\omega)$, $\omega \in \Omega$, принимающая значения в расширенном интервале $[t_0, \infty]$;

в) при каждом $t \in \mathcal{T}$ σ-алгебра $\mathfrak{F}_t$ подмножеств множества $\Omega_t = \{\omega\colon \zeta(\omega) > t\}$, причем $\mathfrak{F}_s[\Omega_t] \subset \mathfrak{F}_t \subset \mathfrak{A}$ при $s \leqslant t$ $(s, t \in \mathcal{T})$, где $\mathfrak{F}_s[\Omega_t]$ — след σ-алгебры $\mathfrak{F}_s$ на множестве Ω_t, т. е. совокупность всех множеств вида $A \cap \Omega_t$, $A \in \mathfrak{F}_s$;

г) функция двух переменных $\xi(t) = \xi(t, \omega)$, $\omega \in \Omega$, $t \in [t_0, \zeta(\omega))$, со значениями в некотором измеримом пространстве $(X, \mathfrak{B})$ такая, что при всех $t \in \mathcal{T}$ и $\Gamma \in \mathfrak{B}$ выполнено включение $\{\omega\colon \xi(t, \omega) \in \Gamma\} \in \mathfrak{F}_t$.

Система объектов а) — г) определяет *обрывающуюся марковскую случайную функцию*, если при любых $s \leqslant t$ $(s,\ t \in \mathscr{T})$, $\Gamma \in \mathfrak{B}$ выполнено соотношение

$$\mathsf{P}\{\xi(t) \in \Gamma \mid \mathfrak{F}_s\} = \mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}$$

почти наверное относительно меры P на множестве Ω_s. (Другими словами, это соотношение выполняется для почти всех $\omega \in \Omega_s$ относительно меры P.)

Момент времени $\zeta(\omega)$ называется *моментом обрыва* марковской случайной функции, а величина $\zeta(\omega) - t_0$ называется ее *временем жизни*.

Правая часть последнего равенства есть условная вероятность относительно σ-алгебры подмножеств множества Ω_s, порожденной множествами вида $\{\omega: \xi(s) \in \Gamma\}$, $\Gamma \in \mathfrak{B}$.

Предположим, что для $\mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}$ существует регулярная условная вероятность, т. е. функция $P(s, x, t, \Gamma)$, $x \in X$, $s < t$ $(s,\ t \in \mathscr{T})$, $\Gamma \in \mathfrak{B}$, удовлетворяющая условиям:

1) при фиксированных $s,\ t,\ \Gamma$ функция $P(s, x, t, \Gamma)$ является $\mathfrak{B}$-измеримой функцией от x;

2) при фиксированных $s,\ x,\ t$ функция $P(s, x, t, \Gamma)$ является мерой на $(X, \mathfrak{B})$ (не обязательно вероятностной, так как $P(s, x, t, X) \leqslant 1$);

3) при всех $s \leqslant t$ $(s,\ t \in \mathscr{T})$, $\Gamma \in \mathfrak{B}$

$$\mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\} = P(s, \xi(s), t, \Gamma)$$

почти наверное относительно P на множестве Ω_s (п. н. Ω_s, P).

В этом случае функция $P(s, x, t, \Gamma)$ называется *вероятностью перехода марковской случайной функции*. Она интерпретируется как условная вероятность события $\{\xi(t) \in \Gamma\}$ при условии $\xi(s) = x$. В частности, величина $1 - P(s, x, t, X)$ есть условная вероятность того, что к моменту времени t марковская случайная функция оборвалась (исчезла из фазового пространства) при условии $\xi(s) = x$.

Заметим, что если $\zeta(\omega) \equiv +\infty$, то определение обрывающейся марковской случайной функции превращается в определение необрывающейся марковской случайной функции, заданной на $\mathscr{T} = [t_0, \infty)$.

Обрывающуюся марковскую случайную функцию всегда можно превратить в необрывающуюся. С этой целью расширим пространство X, добавив к нему некоторую «несобственную» точку $a \notin X$. Положим $X^{(a)} = X \cup \{a\}$. Обозначим через $\mathfrak{B}^{(a)}$ σ-алгебру подмножеств множества $X^{(a)}$, состоящую из множеств $\Gamma \in \mathfrak{B}$ и множеств вида $\Gamma \cup \{a\}$, $\Gamma \in \mathfrak{B}$.

Далее, для $\omega \in \{\zeta(\omega) < +\infty\}$ положим $\xi^{(a)}(t, \omega) = a$ при $t \geqslant \zeta$. При $t \in [t_0, \zeta(\omega))$ положим $\xi^{(a)}(t, \omega) = \xi(t, \omega)$. Наконец, определим σ-алгебру $\mathfrak{F}_t^{(a)}$, $t_0 \leqslant t < \infty$, как наименьшую σ-алгебру подмножеств Ω, содержащую все σ-алгебры $\mathfrak{F}_s$ при $s \leqslant t$. Тогда σ-алгебры $\mathfrak{F}_t^{(a)}$, $t \in \mathscr{T}$, образуют поток, и процесс $\{\xi^{(a)}(t),\ t \in \mathscr{T}\}$ подчинен этому потоку. Нетрудно проверить, что $\{\xi^{(a)}(t),\ t \in \mathscr{T}\}$ есть необрывающаяся марковская случайная функция относительно потока $\{\mathfrak{F}_t^{(a)},\ t \in \mathscr{T}\}$. Вероятность перехода $\tilde{P}(s, x, t, \Gamma)$ этой случайной функции равна

$$\tilde{P}(s, x, t, \Gamma) = \begin{cases} P(s, x, t, \Gamma \cap X) + \chi_\Gamma(a)\,[1 - P(s, x, t, X)], & x \neq a, \\ \chi_\Gamma(a), & x = a, \end{cases}$$

где $P(s, x, t, \Gamma)$ — вероятность перехода исходной (обрывающейся) случайной функции. Для процесса $\{\xi^{(a)}(t),\ t \in \mathcal{T}\}$ состояние a является *поглощающим*. Это означает, что, попав в это состояние, процесс никогда из него не выйдет. Разумеется, это не единственный способ превратить обрывающуюся марковскую случайную функцию в необрывающуюся.

14.2. Марковские процессы. Определение и основные свойства

14.2.1. Определение. Как уже отмечалось в гл. 8, при изучении марковского свойства случайных процессов удобно не фиксировать начальное распределение процесса (да и сам начальный момент времени), а рассматривать целое семейство марковских случайных функций, «начинающихся» в произвольный момент времени в произвольной точке фазового пространства. С точки зрения теории вероятностей это означает, что на вероятностном пространстве имеется уже не одна фиксированная вероятностная мера, а семейство мер P_{sx}, зависящих от временнóй и фазовой переменных и связанных между собой марковским свойством. При этом мера P_{sx} интерпретируется как условная вероятность некоторого события, которое может произойти после момента времени s при условии, что $\xi(s) = x$. Будем предполагать, что $\mathcal{T} = [0, \infty)$.

Пусть даны:

а) измеримое пространство $(\Omega, \mathfrak{A})$, называемое пространством элементарных событий;

б) семейство σ-алгебр $\mathfrak{F}_t^s$ $(0 \leqslant s \leqslant t \leqslant \infty)$ такое, что $\mathfrak{F}_t^s \subset \mathfrak{F}_{t_1}^{s_1} \subset \mathfrak{A}$ при $0 \leqslant s_1 \leqslant s \leqslant t \leqslant t_1 \leqslant \infty$; условимся писать $\mathfrak{F}_t$ вместо $\mathfrak{F}_t^0$ и $\mathfrak{F}^s$ вместо $\mathfrak{F}_\infty^s$;

в) функция двух переменных $\xi(t) = \xi(t, \omega)$, $t \in \mathcal{T}$, $\omega \in \Omega$, со значениями в некотором измеримом пространстве $(X, \mathfrak{B})$ такая, что при любых $0 \leqslant s \leqslant t$ отображение $\xi(t, \cdot)$ пространства $\left(\Omega, \mathfrak{F}_t^s\right)$ в пространство $(X, \mathfrak{B})$ измеримо; предполагается, что σ-алгебра $\mathfrak{B}$ содержит все одноточечные множества;

г) семейство вероятностных мер $\{\mathsf{P}_{sx},\ s \in \mathcal{T},\ x \in X\}$ на σ-алгебре $\mathfrak{F}^s$.

Систему объектов а) — г) будем называть (*необрывающимся*) *марковским процессом*, если выполнены условия:

1) при любых $0 \leqslant s \leqslant t$, $\Gamma \in \mathfrak{B}$ функция

$$P(s, x, t, \Gamma) = \mathsf{P}_{sx}\{\xi(t) \in \Gamma\}$$

есть $\mathfrak{B}$-измеримая функция от x, причем $P(s, x, s, \Gamma) = \chi_\Gamma(x)$, где $\chi_\Gamma(x)$ — индикатор множества Γ;

2) при любых $0 \leqslant s \leqslant t_1 \leqslant t_2$, $x \in X$, $\Gamma \in \mathfrak{B}$ с P_{sx}-вероятностью 1 выполнено соотношение

$$\mathsf{P}_{sx}\left\{\xi(t_2) \in \Gamma \mid \mathfrak{F}_{t_1}^s\right\} = P(t_1, \xi(t_1), t_2, \Gamma).$$

Марковский процесс обозначается $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$. Пространство $(X, \mathfrak{B})$ называется *фазовым пространством процесса*, функция

$P(s, x, t, \Gamma)$ — *вероятностью перехода*. Заметим, что $P(s, x, t, X) = 1$ и, как следует из условия 2),

$$P(s, x, t_2, \Gamma) = \int_X P(s, x, t_1, dy)\, P(t_1, y, t_2, \Gamma)$$

при всех $0 \leqslant s \leqslant t_1 \leqslant t_2$, $x \in X$, $\Gamma \in \mathfrak{B}$, т. е. вероятность перехода удовлетворяет уравнению Колмогорова — Чепмена.

Пусть $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$. Обозначим через $\mathfrak{N}_t^s$ минимальную σ-алгебру событий, включающую в себя все события вида $\{\xi(u) \in \Gamma\}$ при $u \in [s, t]$, $\Gamma \in \mathfrak{B}$, а через $\mathfrak{N}^s$ минимальную σ-алгебру событий, содержащую все σ-алгебры $\mathfrak{N}_t^s$ при $t \geqslant s$; как обычно, вместо $\mathfrak{N}_t^0$ будем писать $\mathfrak{N}_t$. Очевидно, что $\mathfrak{N}_t^s \subseteq \mathfrak{F}_t^s$, $\mathfrak{N}^s \subseteq \mathfrak{F}^s$, и процесс $(\xi(t), \mathfrak{N}_t^s, \mathsf{P}_{sx})$ также является марковским.

Следующие свойства марковского процесса являются простыми непосредственными следствиями его определения:

1) для любого $A \in \mathfrak{N}^s$ функция $\mathsf{P}_{sx}(A)$ $\mathfrak{B}$-измерима как функция x;

2) для любой ограниченной (неотрицательной) $\mathfrak{N}^s$-измеримой случайной величины η функция $\mathsf{M}_{sx}\eta$ $\mathfrak{B}$-измерима как функция x;

3) для любого $A \in \mathfrak{N}^t$ с P_{sx}-вероятностью 1

$$\mathsf{P}_{sx}\left\{A \mid \mathfrak{F}_t^s\right\} = \mathsf{P}_{t\xi(t)}(A), \quad s \leqslant t;$$

4) для любой ограниченной $\mathfrak{N}^t$-измеримой случайной величины η с P_{sx}-вероятностью 1

$$\mathsf{M}_{sx}\left\{\eta \mid \mathfrak{F}_t^s\right\} = \mathsf{M}_{t\xi(t)}\eta, \quad s \leqslant t;$$

5) если $A \in \mathfrak{F}_t^s$, $B \in \mathfrak{N}^t$, то

$$\mathsf{P}_{sx}(A \cap B) = \int_A \mathsf{P}_{t\xi(t)}(B)\, \mathsf{P}_{sx}(d\omega), \quad s \leqslant t.$$

14.2.2. Расширение основных σ-алгебр. Пусть $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$. Оказывается возможным (и часто бывает полезным) некоторое расширение основных σ-алгебр, входящих в определение процесса, с сохранением марковского свойства. Такое расширение целесообразно в связи с тем, что σ-алгебры $\mathfrak{N}_t^s$, $\mathfrak{N}^s$ не содержат многих важных с точки зрения теории вероятностей событий. Указанное расширение состоит в пополнении σ-алгебр по системам мер.

Пусть μ — некоторая конечная мера на $(X, \mathfrak{B})$. Обозначим через $\mathfrak{B}_\mu$ пополнение σ-алгебры $\mathfrak{B}$ по мере μ. Это означает, что $A \in \mathfrak{B}_\mu$, если существуют такие множества A_1, $A_2 \in \mathfrak{B}$, что $A_1 \subseteq A \subseteq A_2$ и $\mu(A_1) = \mu(A_2)$. Через $\mathfrak{B}^*$ обозначим σ-алгебру, являющуюся пересечением σ-алгебр $\mathfrak{B}_\mu$ по всем конечным мерам μ, заданным на $\mathfrak{B}$. Множества из $\mathfrak{B}^*$ называются *универсально измеримыми множества-*

ми, порожденными σ-алгеброй $\mathfrak{B}$. Далее, σ-алгебры $\mathfrak{F}^s$ пополним по семейству мер $\{\mathsf{P}_{ux},\ u \leqslant s,\ x \in X\}$ и обозначим это пополнение через $\overline{\mathfrak{F}^s}$. Это означает, что $A \in \overline{\mathfrak{F}^s}$, если для каждых $u \leqslant s$, $x \in X$ найдутся множества $A_1, A_2 \in \mathfrak{F}^s$ такие, что $\mathsf{P}_{ux}(A_1) = \mathsf{P}_{ux}(A_2)$ и $A_1 \subseteq A \subseteq A_2$.

Пусть $\overline{\mathfrak{F}}_t^s$ обозначает σ-алгебру, содержащую все такие события $A \in \overline{\mathfrak{F}^s}$, что при каждых $u \leqslant s$, $x \in X$ найдется событие $A_1 \in \mathfrak{F}_t^s$, для которого $\mathsf{P}_{ux}(A \Delta A_1) = 0$ где $A \Delta A_1$ означает симметрическую разность множеств A и A_1. Аналогично пусть $\widetilde{\mathfrak{N}}^s$ обозначает пополнение σ-алгебры $\mathfrak{N}^s$ по семейству мер $\mathsf{P}_{u\mu}$, где $u \leqslant s$, μ — произвольная вероятностная мера на $(X, \mathfrak{B})$, а $\mathsf{P}_{u\mu}$ определяется формулой

$$\mathsf{P}_{u\mu}(A) = \int_X \mu(dx)\, \mathsf{P}_{ux}(A), \quad A \in \mathfrak{N}^s.$$

Наконец, через $\widetilde{\mathfrak{N}}_t^s$ обозначим σ-алгебру событий $A \in \widetilde{\mathfrak{N}}^s$ таких, что при каждых $u \leqslant s$ и вероятностной мере μ на $(X, \mathfrak{B})$ найдется событие $A_1 \in \mathfrak{N}_t^s$, для которого $\mathsf{P}_{u\mu}(A \Delta A_1) = 0$.

Можно доказать, что для всякой ограниченной $\widetilde{\mathfrak{N}}^s$-измеримой случайной величины η функция $\mathsf{M}_{sx}\eta$ $\mathfrak{B}^*$-измерима как функция x и что отображение $\xi(t, \cdot)$ пространства $\left(\Omega, \widetilde{\mathfrak{N}}_t^s\right)$ (а стало быть, и пространства $\left(\Omega, \overline{\mathfrak{F}}_t^s\right)$, поскольку $\widetilde{\mathfrak{N}}_t^s \subset \overline{\mathfrak{F}}_t^s$) в пространство $(X, \mathfrak{B}^*)$ является измеримым при всех $s \leqslant t$. Кроме того, для всех $s \leqslant t$ и $A \in \widetilde{\mathfrak{N}}^t$ с P_{sx}-вероятностью 1 выполнено соотношение

$$\mathsf{P}_{sx}\left\{A \mid \overline{\mathfrak{F}}_t^s\right\} = \mathsf{P}_{t\xi(t)}(A).$$

Таким образом, процесс $\left(\xi(t), \overline{\mathfrak{F}}_t^s, \mathsf{P}_{sx}\right)$ является марковским в фазовом пространстве $(X, \mathfrak{B}^*)$. Поэтому всегда, когда это удобно, можно считать, что $\mathfrak{B} = \mathfrak{B}^*$, $\mathfrak{F}_t^s = \overline{\mathfrak{F}}_t^s$, $\widetilde{\mathfrak{N}}_t^s = \mathfrak{N}_t^s$.

14.2.3. Условия регулярности. Пусть $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с пространством элементарных событий $(\Omega, \mathfrak{A})$. Зафиксируем некоторое $s_0 \geqslant 0$, и пусть μ — некоторая вероятностная мера на $(X, \mathfrak{B})$. Рассмотрим случайную функцию $\{\xi(t),\ t \geqslant s_0\}$, подчиненную потоку σ-алгебр $\left\{\mathfrak{N}_t^{s_0},\ t \geqslant s_0\right\}$ и определенную на вероятностном пространстве $\left(\Omega, \mathfrak{N}^{s_0}, \mathsf{P}_{s_0\mu}\right)$, где

$$\mathsf{P}_{s_0\mu}(A) = \int_X \mathsf{P}_{s_0x}(A)\, \mu(dx), \quad A \in \mathfrak{N}^{s_0}.$$

Она является марковской случайной функцией, определенной при $t \geqslant s_0$, с начальным распределением μ и вероятностью перехода $F(t_1, x, t_2, \Gamma)$, совпадающей с вероятностью перехода марковского процесса $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$.

Таким образом, имея процесс Маркова, мы можем построить бесчисленное множество марковских случайных функций, выбирая начало отсчета времени и задавая начальное распределение.

Два марковских процесса, заданных в одном и том же фазовом пространстве (возможно, на разных вероятностных пространствах), называются *эквивалентными*, если построенные по ним марковские случайные функции с одним и тем же началом отсчета и начальным распределением стохастически эквивалентны (т. е. имеют одинаковые конечномерные распределения). Отсюда следует, что марковские процессы эквивалентны тогда и только тогда, когда у них совпадают вероятности перехода. Таким образом, своей вероятностью перехода марковский процесс определяется однозначно с точностью до эквивалентности.

Рассмотрим теперь вопрос о том, всегда ли существует марковский процесс с заданной вероятностью перехода. Ответ на этот вопрос дает следующая теорема.

Теорема 1. *Пусть X — полное метрическое сепарабельное пространство, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X, $\mathcal{T} = [0, \infty)$. Предположим, что задана функция $P(s, x, t, \Gamma)$, $0 \leqslant s \leqslant t < \infty$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяющая условиям:*

а) *при фиксированных s, t, Γ $P(s, x, t, \Gamma)$ $\mathfrak{B}$-измерима;*

б) *при фиксированных s, t, x $P(s, x, t, \Gamma)$ является вероятностной мерой на $(X, \mathfrak{B})$;*

в) *при всех $x \in X$, $0 \leqslant s \leqslant t_1 \leqslant t_2 < \infty$*

$$P(s, x, t_2, \Gamma) = \int_X P(s, x, t_1, dy)\, P(t_1, y, t_2, \Gamma).$$

Тогда существует (необрывающийся) марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода $P(s, x, t, \Gamma)$.

Разумеется, как в определении марковского процесса, так и в этой теореме множество $\mathcal{T}$ может быть конечным отрезком или даже некоторым подмножеством множества $[0, \infty)$.

Естественно поставить вопрос: при каких условиях на вероятность перехода среди всех эквивалентных марковских процессов существует такой процесс, выборочные функции или траектории которого (т. е. функции $\xi(t, \omega)$ при фиксированном ω как функции t) обладают тем или иным свойством регулярности, скажем непрерывны, имеют пределы слева и т. д.? Прежде чем сформулировать ответ на этот вопрос, приведем некоторые определения.

Марковский процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ в метрическом фазовом пространстве $(X, \mathfrak{B})$ ($\mathfrak{B}$ — σ-алгебра борелевских множеств) называется *непрерывным* (*непрерывным справа*), если для любых $s \geqslant 0$ и $x \in X$ с P_{sx}-вероятностью 1 его траектории непрерывны (непрерывны справа) при всех $t \geqslant s$.

Марковский процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ не имеет разрывов второго рода, если для всех $s \geqslant 0$ и $x \in X$ с P_{sx}-вероятностью 1 его траектории не имеют разрывов второго рода при всех $t \geqslant s$.

Обозначим через $U_\varepsilon(x)$ шар в X радиуса ε с центром в точке x и положим $\overline{U}_\varepsilon(x) = X \setminus U_\varepsilon(x)$.

Теорема 2. *Пусть $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$, где X — полное локально компактное сепарабельное метрическое пространство, а $\mathfrak{B}$ — σ-алгебра борелевских множеств в X, с вероятностью перехода $P(s, x, t, \Gamma)$.*

1) *Если при любом* $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \sup_{\substack{0 \leqslant s \leqslant t \leqslant s+\delta \\ x \in X}} P(s, x, t, \overline{U}_\varepsilon(x)) = 0,$$

то процесс $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ *эквивалентен марковскому процессу, не имеющему разрывов второго рода и непрерывному справа.*

2) *Если при любом* $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \delta^{-1} \sup_{\substack{0 \leqslant s \leqslant t \leqslant s+\delta \\ x \in X}} P(s, x, t, \overline{U}_\varepsilon(x)) = 0,$$

то процесс $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ *эквивалентен непрерывному марковскому процессу.*

14.2.4. Строгая марковость. Пусть $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$. Случайную величину $\tau = \tau(\omega)$ со значениями в $[s, \infty)$ будем называть *марковским моментом* относительно потока σ-алгебр $\{\mathfrak{F}_t^s, t \geqslant s\}$ $(\{\mathfrak{N}_t^s, t \geqslant s\}, \{\widetilde{\mathfrak{N}}_t^s, t \geqslant s\})$, если при всех $t \geqslant s$ выполняется условие $\{\omega: \tau(\omega) \leqslant t\} \in \mathfrak{F}_t^s$ $(\{\omega: \tau(\omega) \leqslant t\} \in \mathfrak{N}_t^s, \{\omega: \tau(\omega) \leqslant t\} \in \widetilde{\mathfrak{N}}_t^s)$. Марковские моменты называют иногда величинами, не зависящими от будущего, так как наглядно условие $\{\tau \leqslant t\} \in \mathfrak{F}_t^s$ означает, что наступление или ненаступление события $\{\tau \leqslant t\}$ зависит лишь от явлений, наблюдаемых в течение времени от момента s до момента t.

Каждому марковскому моменту времени τ относительно потока $\{\mathfrak{F}_t^s, t \geqslant s\}$ $(\{\mathfrak{N}_t^s, t \geqslant s\}, \{\widetilde{\mathfrak{N}}_t^s, t \geqslant s\})$ можно поставить в соответствие σ-алгебру $\mathfrak{F}_\tau^s$ $(\mathfrak{N}_\tau^s, \widetilde{\mathfrak{N}}_\tau^s)$, определяемую как совокупность всех тех $A \in \mathfrak{F}^s$ $(A \in \mathfrak{N}^s, A \in \widetilde{\mathfrak{N}}^s)$, для которых $A \cap \{\tau \leqslant t\} \in \mathfrak{F}_t^s$ при любом $t \in [s, \infty)$ $(A \cap \{\tau \leqslant t\} \in \mathfrak{N}_t^s, A \cap \{\tau \leqslant t\} \in \widetilde{\mathfrak{N}}_t^s)$.

Очевидно, что величина τ $\mathfrak{F}_\tau^s$-измерима, и если τ_1 и τ_2 — два марковских момента относительно потока $\{\mathfrak{F}_t^s, t \geqslant s\}$, для которых $\{\tau_1 \leqslant \tau_2\}$, то $\mathfrak{F}_{\tau_1}^s \subset \mathfrak{F}_{\tau_2}^s$.

Часто бывает необходимо рассматривать значение случайной функции $\xi(t)$ в случайный момент времени τ. Для того чтобы в результате получилась случайная величина, нужно, чтобы функции $\xi(t)$ (как функции t) были измеримы. Более того, если мы желаем, чтобы $\xi(\tau)$ было $\mathfrak{F}_\tau^s$-измеримым, нужно потребовать прогрессивной измеримости процесса $\xi(t)$.

Марковский процесс $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ в фазовом пространстве $(X, \mathfrak{B})$ называется *прогрессивно измеримым*, если для любых s, t $(0 \leqslant s < t < \infty)$ отображение $\xi(u, \omega)$ пространства $([s, t] \times \Omega, \mathcal{T}_t^s \times \mathfrak{F}_t^s)$ в пространство $(X, \mathfrak{B})$ измеримо. Здесь $\mathcal{T}_t^s$ — σ-алгебра борелевских подмножеств отрезка $[s, t]$.

Заметим, что если процесс $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ непрерывен справа, то он прогрессивно измерим. Если марковский процесс прогрессивно измерим и τ — конечный марковский момент относительно потока $\{\mathfrak{F}_t^s,\ t \geqslant s\}$, то тогда случайный элемент $\xi(\tau) = \xi(\tau(\omega), \omega)$ со значениями в $(X, \mathfrak{B})$ $\mathfrak{F}_\tau^s$-измерим. Более того, если $\{\tau_t,\ t \geqslant s\}$ — семейство конечных марковских моментов относительно потока $\{\mathfrak{F}_t^s,\ t \geqslant s\}$ и τ_t при фиксированном ω представляет собой монотонно неубывающую непрерывную справа функцию от t, то случайный процесс $\eta(t) = \xi(\tau_t(\omega), \omega)$ будет прогрессивно измеримым относительно потока σ-алгебр $\{\mathfrak{F}_{\tau_t}^s,\ t \geqslant s\}$.

Прогрессивно измеримый марковский процесс $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ в фазовом пространстве $(X, \mathfrak{B})$ называется *строго марковским*, если выполнены следующие условия:

1) при фиксированном $\Gamma \in \mathfrak{B}$ вероятность перехода $P(s, x, t, \Gamma)$ процесса является $\mathscr{T}^0 \times \mathfrak{B} \times \mathscr{T}^0$-измеримой функцией от (s, x, t) на множестве $0 \leqslant s \leqslant t < \infty$, $x \in X$; здесь $\mathscr{T}^0$ — σ-алгебра борелевских подмножеств полуоси $[0, \infty)$;

2) для любых $s \geqslant 0$, $t \geqslant 0$, $x \in X$, $\Gamma \in \mathfrak{B}$ и произвольного марковского момента τ относительно потока σ-алгебр $\{\mathfrak{F}_u^s,\ u \geqslant s\}$ выполнено соотношение

$$\mathsf{P}_{sx}\{\xi(t+\tau) \in \Gamma \mid \mathfrak{F}_\tau^s\} = P(\tau, \xi(\tau), t+\tau, \Gamma)$$

почти наверное на множестве $\Omega_\tau = \{\omega\colon \tau(\omega) < \infty\}$ относительно меры P_{sx}.

Заметим, что в частном случае, когда $\tau(\omega) \equiv t_0$, т. е. когда $\tau(\omega)$ не случайно, условие 2) последнего определения совпадает с условием 2) определения марковского процесса (см. п. 14.2.1). Таким образом, понятие строго марковского процесса выделяет из совокупности всех марковских процессов те из них, которые обладают марковским свойством также и в некоторые случайные моменты времени, а именно в марковские моменты.

Если $(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx})$ — строго марковский процесс, а функция $f(x_1, x_2, \ldots, x_n)$, $x_1, x_2, \ldots, x_n \in X$, вещественная $\mathfrak{B}^n$-измеримая и ограниченная, то для любого марковского момента τ и для любых положительных $t_1, t_2, \ldots, t_n$ выполнено соотношение

$$\mathsf{M}_{sx}\{f(\xi(\tau+t_1), \xi(\tau+t_2), \ldots, \xi(\tau+t_n)) \mid \mathfrak{F}_\tau^s\} = \\ = \mathsf{M}_{\tau\xi(\tau)}\{f(\xi(\tau+t_1), \xi(\tau+t_2), \ldots, \xi(\tau+t_n))\}$$

почти наверное на множестве $\Omega_\tau = \{\tau < \infty\}$ относительно меры P_{sx}. При этом правую часть следует понимать так: положим

$$h(s, x, t_1, t_2, \ldots, t_n) = \mathsf{M}_{sx} f(\xi(t_1), \xi(t_2), \ldots, \xi(t_n)),$$

где $x \in X$, $s \leqslant \min(t_1, t_2, \ldots, t_n)$. Тогда

$$\mathsf{M}_{\tau\xi(\tau)} f(\xi(\tau+t_1), \xi(\tau+t_2), \ldots, \xi(\tau+t_n)) = \\ = h(\tau, \xi(\tau), \tau+t_1, \tau+t_2, \ldots, \tau+t_n).$$

Сформулируем теперь критерий строгой марковости процесса. Для этого определим сначала оператор R_λ, $\lambda > 0$, действующий на вещественную ограниченную измеримую функцию $f(t, x)$, $t \in [0, \infty)$, $x \in X$, по формуле

$$(R_\lambda f)(s, x) = \mathsf{M}_{sx} \int_0^\infty e^{-\lambda t} f(s+t, \xi(s+t))\, dt.$$

Можно показать, что если вероятность перехода $P(s, x, t, \Gamma)$ при фиксированном Γ является измеримой функцией переменных (s, x, t), то функция $(R_\lambda f)(s, x)$ измерима по паре переменных s, x и ограничена, каковы бы ни были $\lambda > 0$ и ограниченная измеримая функция $f(s, x)$.

Теорема 3. *Пусть X — метрическое пространство, $\mathfrak{B}$ — σ-алгебра универсально измеримых множеств пространства X. Предположим, что задан марковский процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ в фазовом пространстве $(X, \mathfrak{B})$, удовлетворяющий условиям:*

а) вероятность перехода $P(s, x, t, \Gamma)$ при фиксированных t и Γ измерима по паре переменных (s, x);

б) при любых $s \geqslant 0$, $x \in X$ почти наверное относительно P_{sx} траектории процесса (т. е. функции $\xi(t)$ как функции от t, $t \geqslant s$) непрерывны справа;

в) для любых $s \geqslant 0$, $x \in X$ и для любой ограниченной непрерывной функции $f(x)$, $x \in X$, почти наверное относительно меры P_{sx} траектории процесса $(R_\lambda f)(t, \xi(t))$, $t \geqslant s$, непрерывны справа.

Тогда процесс $\left(\xi(t), \mathfrak{F}_{t+}^s, \mathsf{P}_{sx}\right)$ является строго марковским. $\left(\text{Здесь } \mathfrak{F}_{t+}^s = \bigcap_{\varepsilon > 0} \mathfrak{F}_{t+\varepsilon}^s\right)$.

14.2.5. Стандартные процессы. Марковский процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ в фазовом пространстве $(X, \mathfrak{B})$, где X — локально компактное метрическое пространство, а $\mathfrak{B}$ — σ-алгебра борелевских подмножеств пространства X, называется *стандартным*, если выполнены следующие условия:

1) $\mathfrak{F}_t^s = \mathfrak{F}_{t+}^s = \overline{\mathfrak{F}}_t^s$ при s, t $(0 \leqslant s \leqslant t < \infty)$;

2) процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ непрерывен справа и имеет пределы слева;

3) процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ является строго марковским;

4) процесс $\left(\xi(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ квазинепрерывен слева; это означает, что для всяких $s \geqslant 0$, $x \in X$ и всякой неубывающей последовательности марковских моментов τ_n $(n = 1, 2, \ldots)$ относительно потока $\left\{\mathfrak{F}_t^s, t \geqslant s\right\}$ имеет место соотношение $\lim\limits_{n \to \infty} \xi(\tau_n) = \xi\left(\lim\limits_{n \to \infty} \tau_n\right)$ почти наверное на множестве $\{\omega: \lim\limits_{n \to \infty} \tau_n < \infty\}$ относительно меры P_{sx}.

Приведем теперь условия на вероятность перехода, при которых можно гарантировать существование стандартного процесса с заданной вероятностью перехода. Предварительно дадим определение феллеровской вероятности перехода.

Пусть X — сепарабельное локально компактное метрическое пространство, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X, $P(s, x, t, \Gamma)$ — вероятность перехода в $(X, \mathfrak{B})$, т. е. функция, удовлетворяющая условиям теоремы 1. Обозначим через $C_0(X)$ совокупность всех вещественных непрерывных функций, заданных на X и стремящихся к нулю, когда x выходит из всех компактов, содержащихся в X. Это означает, что для каждого $\varepsilon > 0$ существует такой компакт $K_\varepsilon \subset X$, что $|f(x)| < \varepsilon$ для всех $x \in X \setminus K_\varepsilon$. Заметим, что если X — компакт, то $C_0(X)$ совпадает с множеством всех непрерывных вещественнозначных функций, определенных при $x \in X$.

Вероятность перехода $P(s, x, t, \Gamma)$ называется *феллеровской*, если выполнены следующие условия:

1) при любых s, t $(0 \leqslant s < t < \infty)$ и $f \in C_0(X)$ функция

$$(T_{st}f)(x) = \int_X f(y) P(s, x, t, dy)$$

непрерывна в совокупности переменных (s, t, x), $0 \leqslant s \leqslant t$, $x \in X$;

2) для любой функции $f \in C_0(X)$

$$\lim_{t \downarrow s} \sup_{x \in X} |(T_{st}f)(x) - f(x)| = 0.$$

Теорема 4. *Если $P(s, x, t, \Gamma)$ — феллеровская вероятность перехода в сепарабельном локально компактном метрическом пространстве $(X, \mathfrak{B})$, то существует стандартный марковский процесс с вероятностью перехода $P(s, x, t, \Gamma)$.*

14.2.6. Обрывающиеся процессы. Пусть заданы:

а) измеримое пространство $(\Omega, \mathfrak{A})$, называемое пространством элементарных событий;

б) случайная величина $\zeta(\omega)$, принимающая значения в расширенном отрезке $[0, \infty]$;

в) для каждых s, t, $0 \leqslant s \leqslant t$, σ-алгебры $\mathfrak{F}_t^s \subset \mathfrak{A}$ в пространстве $\Omega_t = \{\omega: \zeta(\omega) > t\}$ такие, что если $s \leqslant t \leqslant u$ и $A \in \mathfrak{F}_t^s$, то $A \cap \Omega_u \in \mathfrak{F}_u^s$;

г) функция двух переменных $\xi(t) = \xi(t, \omega)$, $t \in [0, \zeta(\omega))$, $\omega \in \Omega$, со значениями в некотором измеримом пространстве $(X, \mathfrak{B})$ такая, что при любых $0 \leqslant s \leqslant t$ отображение $\xi(t, \cdot)$ пространства $(\Omega_t, \mathfrak{F}_t^s)$ в пространство $(X, \mathfrak{B})$ измеримо; предполагается, что σ-алгебра $\mathfrak{B}$ содержит все одноточечные множества;

д) для каждых $s \geqslant 0$, $x \in X$ вероятностные меры P_{sx} на σ-алгебре $\mathfrak{F}^s = \mathfrak{F}_\infty^s$.

Систему объектов а) — д) будем называть (*обрывающимся*) *марковским процессом*, если выполнены следующие условия:

1) при любых $0 \leqslant s \leqslant t$, $\Gamma \in \mathfrak{B}$ функция

$$P(s, x, t, \Gamma) = \mathsf{P}_{sx}\{\xi(t) \in \Gamma\}$$

есть $\mathfrak{B}$-измеримая функция от x, причем $P(s, x, s, X \setminus \{x\}) = 0$;

2) при любых $0 \leqslant s \leqslant t_1 \leqslant t_2$, $x \in X$, $\Gamma \in \mathfrak{B}$ выполнено соотношение

$$\mathsf{P}_{sx}\left\{\xi(t_2) \in \Gamma \mid \mathfrak{F}_{t_1}^s\right\} = P(t_1, \xi(t_1), t_2, \Gamma)$$

почти наверное на множестве Ω_{t_1} относительно меры P_{sx} (п. н. Ω_{t_1}, P_{sx}).

Момент времени ζ называется *моментом обрыва*. Обрывающийся марковский процесс будем обозначать $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$. Очевидно, что если $\zeta(\omega) \equiv +\infty$, то последнее определение превращается в определение необрывающегося марковского процесса.

Функция $P(s, x, t, \Gamma)$, определенная в условии 1) последнего определения, называется *вероятностью перехода марковского процесса*. Из условия 2) следует, что она удовлетворяет уравнению Колмогорова — Чепмена. Следует только иметь в виду, что $P(s, x, t, X) \leqslant 1$ и величина $1 - P(s, x, t, X)$ представляет собой вероятность того, что, выйдя из точки x в момент времени s, процесс оборвется к моменту времени t.

Вероятность перехода $P(s, x, t, \Gamma)$ называется *нормальной*, если при всех $s \geqslant 0$ и $x \in X$ $P(s, x, s+, X) = \lim\limits_{t \downarrow s} P(s, x, t, X) = 1$. Тот факт, что указанный предел существует, следует из неравенства $(s \leqslant t_1 < t_2)$

$$P(s, x, t_2, X) = \int\limits_X P(s, x, t_1, dy)\, P(t_1, y, t_2, X) \leqslant P(s, x, t_1, X),$$

означающего, что $P(s, x, t, X)$ монотонно не возрастает как функция от t при $t \geqslant s$.

Марковский процесс с нормальной вероятностью перехода называется *нормальным*.

Многие результаты, полученные для необрывающихся марковских процессов, переносятся с известными оговорками и на обрывающиеся процессы. Например, аналогом теоремы 1 (п. 14.2.3) является следующая теорема.

Теорема 1'. *Пусть X — полное метрическое пространство, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X. Предположим, что функция $P(s, x, t, \Gamma)$, $0 \leqslant s \leqslant t < \infty$, $x \in X$, $\Gamma \in \mathfrak{B}$, удовлетворяет условиям 1), 3) теоремы 1 и условию: 2') при фиксированных s, x, t $P(s, x, t, \Gamma)$ является мерой (не обязательно вероятностной) на $(X, \mathfrak{B})$, причем $P(s, x, t, \Gamma) \leqslant 1$ и $P(s, x, s+, X) = 1$.*

Тогда существует нормальный марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода $P(s, x, t, \Gamma)$.

Процессы с одной и той же вероятностью перехода называются *эквивалентными*. Если по эквивалентным процессам построить (обрывающиеся) марковские случайные функции так же, как в п. 14.2.3, то они будут стохастически эквивалентными (т. е. будут иметь одинаковые конечномерные распределения).

Определение строго марковского обрывающегося процесса совпадает с соответствующим определением необрывающегося процесса с единственным изменением, состоящим в том, что соотношение в условии 2) должно выполняться почти наверное на множестве $\Omega_\tau = \{\omega: \tau(\omega) < \zeta(\omega)\}$ относительно меры P_{sx}.

Определение стандартного обрывающегося марковского процесса вполне аналогично определению, данному в п. 14.2.5. Следует иметь в виду, что поскольку функции $\xi(t)$ заданы лишь на полуоткрытом интервале $[0, \zeta)$, то непрерывность справа процесса, а также существование пределов слева относятся к точкам $t \in [0, \zeta)$.

Аналогично в определении квазинепрерывности слева соотношение $\lim\limits_{n \to \infty} \xi(\tau_n) = \xi(\lim\limits_{n \to \infty} \tau_n)$ должно выполняться на множестве $\{\omega: \lim\limits_{n \to \infty} \tau_n < \zeta(\omega)\}$ почти наверное относительно P_{sx}.

Следующая теорема дает условия существования обрывающегося стандартного процесса.

Теорема 5. *Пусть X — локально компактное сепарабельное полное метрическое пространство, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X, а $P(s, x, t, \Gamma)$ — нормальная вероятность перехода (т. е. функция, удовлетворяющая условиям теоремы 1). Предположим, что выполнены условия:*

1) *какова бы ни была непрерывная ограниченная функция $f(x)$ $(x \in X)$ с вещественными значениями, функция*

$$(T_{st}f)(x) = \int\limits_X P(s, x, t, dy) f(y), \quad s \leqslant t, \quad x \in X,$$

обладает следующим свойством непрерывности:

$$\lim_{\substack{x \to z \\ s \downarrow s_0}} (T_{st}f)(x) = (T_{s_0t}f)(z);$$

2) $$\lim_{\delta \downarrow 0} \sup_{\substack{0 \leqslant s \leqslant t \leqslant s+\delta \\ x \in X}} P(s, x, t, \overline{U}_\varepsilon(x)) = 0,$$

где $U_\varepsilon(x)$ — шар в X с центром в точке x радиуса ε, а $\overline{U}_\varepsilon(x) = X \setminus U_\varepsilon(x)$.

Тогда существует нормальный стандартный марковский процесс с вероятностью перехода $P(s, x, t, \Gamma)$.

14.3. Мультипликативные функционалы от марковских процессов

14.3.1. Определение и свойства. Пусть $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$.

Семейство вещественнозначных случайных величин $\alpha_t^s = \alpha_t^s(\omega)$ $(0 \leqslant s \leqslant t,\ \omega \in \Omega_t)$ называется *мультипликативным функционалом* от марковского процесса, если выполнены следующие условия:

а) случайная величина α_t^s $\widetilde{\mathfrak{N}}_t^s$-измерима;

б) при любых $x \in X$, $0 \leqslant s \leqslant t \leqslant u$ почти наверное на множестве Ω_u относительно меры P_{sx} выполнено

$$\alpha_t^s \alpha_u^t = \alpha_u^s;$$

в) $0 \leqslant \alpha_t^s \leqslant 1$ при всех $0 \leqslant s \leqslant t$, $\omega \in \Omega_t$.

Мультипликативный функционал называется *непрерывным справа*, если при всех $t \geqslant s \geqslant 0$, $x \in X$, P_{sx}-почти наверное на множестве Ω_t $\lim\limits_{t_n \downarrow t} \alpha_{t_n}^s = \alpha_t^s$.

Приведем пример мультипликативного функционала. Предположим, что марковский процесс прогрессивно измерим относительно σ-алгебр $\widetilde{\mathfrak{N}}_t^s$ (это означает, что в определении п. 14.2.4 вместо σ-алгебры $\mathfrak{F}_t^s$ нужно поставить σ-алгебру $\widetilde{\mathfrak{N}}_t^s$, а вместо Ω — Ω_t).

Пусть $v(s, x)$ — неотрицательная функция, измеримая по паре переменных $(s, x) \in [0, \infty) \times X$. Положим

$$\alpha_t^s = \exp\left\{-\int_s^t v(u, \xi(u))\,du\right\}, \quad 0 \leqslant s \leqslant t, \quad \omega \in \Omega_t.$$

Если интеграл в этой формуле конечен при всех $0 \leqslant s \leqslant t < \zeta(\omega)$, то α_t^s является мультипликативным функционалом от процесса $\xi(t)$. Он называется *мультипликативным функционалом интегрального типа.* Заметим, что такой функционал непрерывен при $t \geqslant s$, $\omega \in \Omega_t$.

Пусть α_t^s и β_t^s — два мультипликативных функционала от марковского процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$. Они называются *стохастически эквивалентными,* если при всех $s \geqslant 0$, $x \in X$, $t > s$ $\mathsf{P}_{sx}\left\{\alpha_t^s \neq \beta_t^s\right\} = 0$.

Из определения следует, что если α_t^s — мультипликативный функционал, то $\alpha_s^s = \left(\alpha_s^s\right)^2$, так что α_s^s может принимать лишь два значения — 0 и 1. Более того, $\mathsf{P}_{sx}\left\{\alpha_s^s = 1\right\} = 0$ или 1, что следует из следующего закона нуля или единицы.

Теорема (закон нуля или единицы). *Если* $A \in \mathfrak{N}_s^s$, *то* $\mathsf{P}_{sx}(A) = 0$ *или* 1.

Для данного s обозначим через X_α^s совокупность всех тех $x \in X$, для которых $\mathsf{P}_{sx}\left\{\alpha_s^s = 1\right\} = 1$. Очевидно, что $X_\alpha^s \in \mathfrak{B}$. Положим

$$\widetilde{P}(s, x, t, \Gamma) = \mathsf{M}_{sx}\,\chi_\Gamma(\xi(t))\,\alpha_t^s,$$

где $0 \leqslant s \leqslant t < \infty$, $x \in X$, $\Gamma \in \mathfrak{B}$, $\chi_\Gamma(x)$ — индикатор множества Γ. При фиксированных s, x, t функция $\widetilde{P}(s, x, t, \Gamma)$ является мерой на $\mathfrak{B}$, а при фиксированных s, t, Γ она $\mathfrak{B}$-измерима. Кроме того $\widetilde{P}(s, x, t, \Gamma)$ удовлетворяет уравнению Колмогорова — Чепмена

$$\widetilde{P}(s, x, t_2, \Gamma) = \int_X \widetilde{P}(s, x, t_1, dy)\,\widetilde{P}(t_1, y, t_2, \Gamma),$$

$$0 \leqslant s \leqslant t_1 \leqslant t_2, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

При этом

$$\widetilde{P}(s, x, t, \Gamma) \leqslant P(s, x, t, \Gamma), \tag{3.1}$$

где $P(s, x, t, \Gamma)$ — вероятность перехода процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$. Если этот процесс нормален, то

$$\widetilde{P}(s, x, s, \Gamma) = \chi_{\Gamma \cap X_\alpha^s}(x).$$

Таким образом, каждый мультипликативный функционал от марковского процесса порождает некоторую вероятность перехода $\widetilde{P}(s, x, t, \Gamma)$, удовлетворяющую неравенству (3.1). При этом два мультипликативных функционала стохастически эквивалентны тогда и только тогда, когда они порождают одну и ту же вероятность перехода.

Следующая теорема показывает, что при некоторых условиях всякая вероятность перехода, удовлетворяющая неравенству (3.1), порождается некоторым функционалом.

Теорема 1. *Предположим, что заданы нормальный марковский процесс* $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$ *в фазовом пространстве* $(X, \mathfrak{B})$ *с вероятностью перехода* $P(s, x, t, \Gamma)$ *и некоторая вероятность перехода* $\tilde{P}(s, x, t, \Gamma)$, *удовлетворяющая неравенству* (3.1). *Пусть выполнены условия:*

1) *σ-алгебра* $\mathfrak{B}$ *порождается некоторым счетным семейством подмножеств* X;

2) *на полуоси* $[0, \infty)$ *существует такое счетное всюду плотное подмножество* J, *что σ-алгебры* $\mathfrak{N}_t^s$ *порождаются событиями вида* $\{\xi(u) \in \Gamma\}$ *при* $u \in J \cap [s, t]$, $\Gamma \in \mathfrak{B}$.

Тогда существует мультипликативный функционал α_t^s *от процесса* $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{st})$ *такой, что*

$$\tilde{P}(s, x, t, \Gamma) = \mathsf{M}_{sx}\chi_\Gamma(\xi(t))\,\alpha_t^s.$$

При этом, если поток σ-алгебр $\{\tilde{\mathfrak{N}}_t^s, t \geqslant s\}$ *непрерывен справа* (*это означает, что* $\tilde{\mathfrak{N}}_{t+}^s = \tilde{\mathfrak{N}}_t^s$) *и функция* $\tilde{P}(s, x, t, X)$ *также непрерывна справа по* t *в точке* $t = s$ *для всех* s *и* x, *то мультипликативный функционал* α_t^s *можно определить так, чтобы он был непрерывным справа.*

Замечание. Условия 1) и 2) теоремы выполнены, если X — сепарабельное метрическое пространство, $\mathfrak{B}$ — σ-алгебра борелевских множеств пространства X, а процесс $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{st})$ непрерывен справа.

В заключение этого пункта приведем интегральное уравнение, которому удовлетворяет вероятность перехода $\tilde{P}(s, x, t, \Gamma)$, если она порождается функционалом интегрального типа.

Пусть $v(s, x)$ — неотрицательная измеримая по паре переменных функция, заданная при $s \geqslant 0$, $x \in X$. Процесс предполагается прогрессивно измеримым относительно σ-алгебр $\tilde{\mathfrak{N}}_t^s$ (см. выше).

Положим для $0 \leqslant s \leqslant t$, $\omega \in \Omega_t = \{\omega: \zeta(\omega) > t\}$

$$\alpha_t^s = \exp\left\{-\int_s^t v(u, \xi(u))\,du\right\}.$$

Будем предполагать, что интеграл в показателе экспоненты конечен при всех $0 \leqslant s \leqslant t < \zeta(\omega)$. Если

$$\tilde{P}(s, x, t, \Gamma) = \mathsf{M}_{sx}\chi_\Gamma(\xi(t))\,\alpha_t^s,$$

то функция $\tilde{P}(s, x, t, \Gamma)$ удовлетворяет следующему интегральному уравнению:

$$\tilde{P}(s, x, t, \Gamma) = P(s, x, t, \Gamma) - \int_s^t \int_X P(s, x, u, dy)\, v(u, y) \times \\ \times \tilde{P}(u, y, t, \Gamma)\,du.$$

Если функция $v(s, x)$ ограничена, то единственным решением этого уравнения является функция

$$\widetilde{P}(s, x, t, \Gamma)=\sum_{n=0}^{\infty} \widetilde{P}_n(s, x, t, \Gamma),$$

где $\widetilde{P}_0(s, x, t, \Gamma) = P(s, x, t, \Gamma)$, а при $n = 0, 1, 2, \ldots$

$$\widetilde{P}_{n+1}(s, x, t, \Gamma)=-\int_s^t \int_X P(s, x, u, dy)\, v(u, y)\, \widetilde{P}_n(u, y, t, \Gamma)\, du.$$

Это следует из легко получаемых оценок

$$\widetilde{P}_n(s, x, t, \Gamma) \leqslant \frac{K^n}{n!}(t-s)^n, \quad n=0, 1, 2, \ldots,$$

где

$$K=\sup_{s \geqslant 0,\, x \in X} |v(s, x)|.$$

14.3.2. Подпроцессы. Пусть заданы два марковских процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ и $\left(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx}\right)$ в фазовом пространстве $(X, \mathfrak{B})$ с одним и тем же пространством элементарных событий Ω. Предположим, что выполнены условия:

а) $\tilde{\zeta}(\omega) \leqslant \zeta(\omega)$ при любом $\omega \in \Omega$;

б) $\tilde{\xi}(t) = \xi(t)$ при $0 \leqslant t < \tilde{\zeta}(\omega)$;

в) $\tilde{\mathfrak{F}}_t^s = \mathfrak{F}_t^s\left[\widetilde{\Omega}_t\right]$, где $\widetilde{\Omega}_t = \{\omega: \tilde{\zeta}(\omega) > t\}$, а $\mathfrak{F}_t^s\left[\widetilde{\Omega}_t\right]$ — след σ-алгебры $\mathfrak{F}_t^s$ на множестве $\widetilde{\Omega}_t$, т. е. совокупность событий вида $A \cap \widetilde{\Omega}_t$, $A \in \mathfrak{F}_t^s$.

Тогда говорят, что процесс $\left(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx}\right)$ получен путем сокращения времени жизни процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$.

Марковский процесс $\left(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx}\right)$ называют *подпроцессом* процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$, если первый из них может быть получен путем сокращения жизни некоторого процесса, эквивалентного второму.

Если $P(s, x, t, \Gamma)$ — вероятность перехода марковского процесса $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$, а $\widetilde{P}(s, x, t, \Gamma)$ — вероятность перехода некоторого его подпроцесса, то очевидно, что

$$\widetilde{P}(s, x, t, \Gamma) \leqslant P(s, x, t, \Gamma).$$

Теорема 2. *Если марковский процесс* $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ *удовлетворяет условиям теоремы* 1, *а процесс* $\left(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx}\right)$ *является его подпроцессом, то вероятность перехода* $\widetilde{P}(s, x, t, \Gamma)$ *подпроцесса порождается некоторым мультипликативным функционалом* α_t^s *от процесса* $\left(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ *в том смысле, что*

$$\widetilde{P}(s, x, t, \Gamma)=\mathsf{M}_{sx}\, \chi_\Gamma(\xi(t))\, \alpha_t^s.$$

В известном смысле справедливо и обратное утверждение.

Теорема 3. *Пусть* α_t^s — *непрерывный справа мультипликативный функционал от нормального марковского процесса* $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$. *Тогда существует такой подпроцесс* $(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx})$ *этого процесса, что его вероятность перехода* $\tilde{P}(s, x, t, \Gamma)$ *порождается функционалом* α_t^s, *т. е.*

$$\tilde{P}(s, x, t, \Gamma) = \tilde{\mathsf{M}}_{sx}\chi_\Gamma(\tilde{\xi}(t)) = \mathsf{M}_{sx}\chi_\Gamma(\xi(t))\,\alpha_t^s$$

при всех $0 \leqslant s \leqslant t$, $x \in X$, $\Gamma \in \mathfrak{B}$.

Если $\mathsf{P}_{sx}\{\alpha_s^s = 1\} = 1$ *при всех* $s \geqslant 0$ *и* $x \in X$, *то указанный подпроцесс является нормальным.*

Рассмотрим теперь подпроцесс $(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx})$ процесса $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$, который порождается мультипликативным функционалом интегрального типа

$$\alpha_t^s = \exp\left\{-\int_s^t v(u, \xi(u))\,du\right\}$$

с ограниченной измеримой неотрицательной функцией $v(u, x)$. Тогда

$$\tilde{P}(s, x, t, X) = \mathsf{M}_{sx}\alpha_t^s.$$

Положим $t = s + h$, и пусть $h \downarrow 0$. С точностью до бесконечно малых более высокого порядка будем иметь

$$P(s, x, s+h, X) - \tilde{P}(s, x, s+h, X) \approx \mathsf{M}_{sx}\int_s^{s+h} v(u, \xi(u))\,du.$$

Предположим, что функция $v(u, \xi(u))$ непрерывна справа, а процесс нормален. Тогда при $h \downarrow 0$

$$P(s, x, s+h, X) - \tilde{P}(s, x, s+h, X) \approx v(s, x)\,h$$

с точностью до бесконечно малых высшего порядка. Если подпроцесс $(\tilde{\xi}(t), \tilde{\zeta}, \tilde{\mathfrak{F}}_t^s, \tilde{\mathsf{P}}_{sx})$ получается путем сокращения времени жизни процесса $(\xi(t), \zeta, \mathfrak{F}_t^s, \mathsf{P}_{sx})$, то левая часть последнего соотношения может быть записана в виде

$$\mathsf{P}_{sx}\{\tilde{\zeta} \leqslant s + h < \zeta\},$$

что представляет собой вероятность того, что подпроцесс $\tilde{\xi}(t)$, выйдя из состояния x в момент времени s, оборвется до момента времени $s + h$, в то время как процесс $\xi(t)$ до этого момента не оборвется.

Литература: [18, 19, 27, 28, 29, 36, 53, 90, 107, 114].

Глава 15. ОДНОРОДНЫЕ МАРКОВСКИЕ ПРОЦЕССЫ

15.1. Определение и основные свойства

15.1.1. Определение. Марковский процесс называется *однородным*, если вероятность перехода $P(s, x, t, \Gamma)$ обладает тем свойством, что функция $P(s, x, s+h, \Gamma)$ не зависит от s. Если положить $P(h, x, \Gamma) = P(s, x, s+h, \Gamma)$, то для конечномерных распределений процесса будет иметь место равенство

$$\mathsf{P}_{sx}\{\xi(t_1)\in\Gamma_1, \ldots, \xi(t_n)\in\Gamma_n\} =$$
$$=\int_{\Gamma_1} P(t_1-s, x, dy_1)\int_{\Gamma_2} P(t_2-t_1, y_1, dy_2)\ldots$$
$$\ldots\int_{\Gamma_n} P(t_n-t_{n-1}, y_{n-1}, dy_n) =$$
$$=\mathsf{P}_{0x}\{\xi(t_1-s)\in\Gamma_1, \ldots, \xi(t_n-s)\in\Gamma_n\},$$
$$0\leqslant s<t_1<t_2<\ldots<t_n.$$

Таким образом, вместо семейства мер P_{sx}, зависящих от временно́й и пространственной переменных, в однородном случае достаточно рассматривать семейство мер $\mathsf{P}_x = \mathsf{P}_{0x}$, зависящих лишь от пространственной переменной. Другими словами, каждый раз, когда процесс выходит из состояния x в момент времени s, мы производим сдвиг времени так, чтобы точка s стала начальной (нулевой). Естественно, что мы должны иметь возможность сдвигать соответственно и все траектории процесса, а это значит, что пространство элементарных событий должно быть достаточно богатым.

Пусть даны:

а) пространство элементарных событий $(\Omega, \mathfrak{A})$;

б) случайная величина $\zeta(\omega)$ со значениями в расширенном отрезке $[0, \infty]$;

в) для каждого $t\geqslant 0$ σ-алгебра $\mathfrak{F}_t$ в пространстве $\Omega_t=\{\omega: \zeta(\omega)>t\}$, причем если $s\leqslant t$, то $\mathfrak{F}_s[\Omega_t]\subseteq\mathfrak{F}_t\subseteq\mathfrak{A}$, где $\mathfrak{F}_s[\Omega_t]$ — след σ-алгебры $\mathfrak{F}_s$ на множестве Ω_t, т. е. совокупность множеств вида $A\cap\Omega_t$, $A\in\mathfrak{F}_s$;

г) функция двух переменных $\xi(t)=\xi(t, \omega)$, $t\in[0, \zeta(\omega)]$, $\omega\in\Omega$, принимающая значения в некотором измеримом пространстве $(X, \mathfrak{B})$, такая, что при каждом $t\geqslant 0$ отображение $\xi(t, \cdot)$ пространства $(\Omega_t, \mathfrak{F}_t)$ в пространство $(X, \mathfrak{B})$ измеримо; предполагается, что σ-алгебра $\mathfrak{B}$ содержит все одноточечные множества;

д) для каждого $x\in X$ вероятностная мера P_x на некоторой σ-алгебре $\mathfrak{F}$ в пространстве Ω, содержащей все $\mathfrak{F}_t$, $t\geqslant 0$.

Система объектов а) — д) образует *однородный марковский процесс*, если выполнены условия:

1) при любых $t\geqslant 0$, $\Gamma\in\mathfrak{B}$ функция

$$P(t, x, \Gamma)=\mathsf{P}_x\{\xi(t)\in\Gamma\}$$

$\mathfrak{B}$-измерима как функция от x, причем $P(0, x, X\setminus\{x\})=0$;

2) для любых $s, t \geqslant 0$, $\Gamma \in \mathfrak{B}$

$$\mathsf{P}_x \{\xi(t+s) \in \Gamma \mid \mathfrak{F}_s\} = P(t, \xi(s), \Gamma)$$

почти наверное относительно меры P_x на множестве Ω_s;

3) для каждого $\omega \in \Omega_t$ найдется такое $\omega' \in \Omega$, что $\zeta(\omega') = \zeta(\omega) - t$ и $\xi(s, \omega') = \xi(s+t, \omega)$ при $0 \leqslant s < \zeta(\omega')$.

Условие 2) соединяет в себе свойство марковости процесса и свойство его однородности во времени. Условие 3) означает, что вместе с каждой траекторией процесса произвольный кусок ее после некоторого момента времени также является возможной траекторией. Расширяя, если это нужно, пространство Ω, всегда можно добиться того, чтобы условие 3) выполнялось.

Функция $P(t, x, \Gamma)$, определенная в условии 1), называется *вероятностью перехода.* Из условия 2) следует, что она удовлетворяет уравнению Колмогорова — Чепмена

$$P(s+t, x, \Gamma) = \int_X P(s, x, dy)\, P(t, y, \Gamma), \quad s, t \geqslant 0, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

При этом $P(s, x, X) \leqslant 1$. Если $P(+0, x, X) = \lim\limits_{t \downarrow 0} P(t, x, X) = 1$, то вероятность перехода называется *нормальной*, а соответствующий процесс — *нормальным*.

Однородный марковский процесс будем обозначать $(\xi(t), \zeta, \mathfrak{F}_t, \mathsf{P}_x)$. Если $\zeta \equiv +\infty$, то процесс называется *необрывающимся* и обозначается $(\xi(t), \mathfrak{F}_t, \mathsf{P}_x)$. Для необрывающегося процесса $P(t, x, X) \equiv 1$.

Обозначим через $\mathfrak{N}^0$ минимальную σ-алгебру подмножеств Ω, содержащую все множества вида $\{\xi(t) \in \Gamma\}$ при $t \geqslant 0$, $\Gamma \in \mathfrak{B}$, а через $\mathfrak{N}_t$ минимальную σ-алгебру подмножеств Ω_t, содержащую все множества вида $\{\xi(s) \in \Gamma\} \cap \Omega_t$ при $s \in [0, t]$, $\Gamma \in \mathfrak{B}$.

Пусть $\mathfrak{N}$ обозначает след σ-алгебры $\mathfrak{N}^0$ на множестве $\Omega_0 = \{\zeta > 0\}$. Очевидно, что $\mathfrak{N} \subseteq \mathfrak{N}^0 \subseteq \mathfrak{F}$, $\mathfrak{N}_t \subseteq \mathfrak{F}_t$, и вместе с процессом $(\xi(t), \zeta, \mathfrak{F}_t, \mathsf{P}_x)$ однородным марковским процессом будет также процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$.

Как уже отмечалось в § 14.2, σ-алгебры $\mathfrak{N}_t$ могут не содержать многих важных множеств. Например, множество $\{\omega: \xi(s) \in \Gamma$ для всех $s \in [u, t]\}$ может не входить в $\mathfrak{N}_t$, поскольку является пересечением несчетного числа цилиндрических множеств. Однако часто множества такого вида содержатся в пересечении пополнений σ-алгебры $\mathfrak{N}_t$ по системе мер P_x. Обозначим эту σ-алгебру через $\overline{\mathfrak{N}}_t$. В дальнейшем будем считать, что $\mathfrak{F}_t = \mathfrak{N}_t$ или $\mathfrak{F}_t = \overline{\mathfrak{N}}_t$.

15.1.2. Эквивалентные марковские процессы. Имея марковский процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ (слово «однородный» часто будем опускать, поскольку речь будет идти только о таких процессах) и вероятностную меру μ на σ-алгебре $\mathfrak{B}$, можем построить марковскую случайную функцию (см. § 14.1) $\xi(t)$, $t \in [0, \infty)$, с временем жизни ζ, потоком σ-алгебр $\mathfrak{N}_t$ и мерой $\mathsf{P}_\mu(A)$, $A \in \mathfrak{N}^0$ определяемой формулой

$$\mathsf{P}_\mu(A) = \int_X \mu(dx)\, \mathsf{P}_x(A).$$

(Заметим, что из условия 1) определения, данного в п. 15.1.1, следует, что при любом $A \in \mathfrak{N}^0$ функция $\mathsf{P}_x(A)$ $\mathfrak{B}$-измерима.) Конечномерные распределения этой случайной функции имеют вид

$$\mathsf{P}_\mu\{\xi(t_1)\in\Gamma_1,\ \xi(t_2)\in\Gamma_2,\ \ldots,\ \xi(t_n)\in\Gamma_n\} = \int\limits_X \mu(dx)\int\limits_{\Gamma_1} P(t_1, x, dy_1)\int\limits_{\Gamma_2} P(t_2-t_1, y_1, dy_2)\ldots \ldots\int\limits_{\Gamma_n} P(t_n-t_{n-1}, y_{n-1}, dy_n),$$

где $0<t_1<\ldots<t_n$, $\Gamma_1, \ldots, \Gamma_n\in\mathfrak{B}$, а $P(t, x, \Gamma)$ — вероятность перехода процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$. Мера μ называется *начальным распределением* построенной случайной функции. Если два марковских процесса имеют одну и ту же вероятность перехода, то соответствующие марковские случайные функции с одним и тем же начальным распределением стохастически эквивалентны (т. е. имеют одинаковые конечномерные распределения).

Марковские процессы с одной и той же вероятностью перехода называются *эквивалентными*.

В соответствии с теоремой 1′ п. 14.2.6 по всякой нормальной однородной вероятности перехода можно построить однородный марковский процесс. При этом под однородной вероятностью перехода в фазовом пространстве $(X, \mathfrak{B})$ мы понимаем функцию $P(t, x, \Gamma)$ $(t\geqslant 0,\ x\in X,\ \Gamma\in\mathfrak{B})$, являющуюся мерой по Γ при фиксированных t, x (с условием $P(t, x, X)\leqslant 1$, $P(0, x, X\setminus\{x\})=0$), представляющую собой $\mathfrak{B}$-измеримую функцию от x при фиксированных t, Γ и удовлетворяющую уравнению Колмогорова — Чепмена.

15.1.3. Операторы сдвига. Определим операцию сдвига θ_t $(t\geqslant 0)$ множеств из σ-алгебры $\mathfrak{N}$. Для множеств вида $\{\omega: \xi(s)\in\Gamma\}$ $(s\geqslant 0,\ \Gamma\in\mathfrak{B})$ положим

$$\theta_t\{\xi(s)\in\Gamma\}=\{\xi(t+s)\in\Gamma\}.$$

Кроме того, потребуем, чтобы операторы θ_t сохраняли все теоретико-множественные операции. Тем самым действие оператора θ_t на любое множество $A\in\mathfrak{N}$ определено однозначно. Например, $\theta_t\Omega_s=\Omega_{s+t}$, а цилиндрические множества

$$\{\xi(t_1)\in\Gamma_1, \ \ldots,\ \xi(t_n)\in\Gamma_n\}$$

под действием оператора θ_t переходят в множества

$$\{\xi(t_1+t)\in\Gamma_1, \ \ldots,\ \xi(t_n+t)\in\Gamma_n\}.$$

Можно определить операторы θ_t и на $\mathfrak{N}$-измеримых функциях от ω. Именно полагаем $(\theta_t\eta(\omega))=a$, если $\omega\in\theta_t\{\eta=a\}$. Очевидно, что $\theta_t\theta_s\eta=\theta_{t+s}\eta$, так что операторы θ_t образуют полугруппу. В частности, $\theta_t\zeta=\zeta-t$ для $\omega\in\Omega_t$.

В терминах операторов θ_t условие 2) определения, данного в п. 15.1.1, может быть записано в виде

$$\mathsf{P}_x\{\theta_s\{\xi(t)\in\Gamma\}\,|\,\mathfrak{F}_s\}=\mathsf{P}_{\xi(s)}\{\xi(t)\in\Gamma\}$$

почти наверное относительно P_x на множестве Ω_s (п. н. Ω_s, P_x).

Следующие свойства марковского процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ являются простыми следствиями определения, данного в п. 15.1.1.

1) Если $A \in \mathfrak{N}$, то

$$\mathsf{P}_x\{\theta_t A \mid \mathfrak{N}_t\} = \mathsf{P}_{\xi(t)}(A) \quad (\text{п. н. } \Omega_t, \mathsf{P}_x);$$

2) если $A \in \mathfrak{N}_t$, $B \in \mathfrak{N}$, то

$$\mathsf{P}_x(A \cap \theta_t B) = \int_A \mathsf{P}_{\xi(t)}(B)\, \mathsf{P}_x(d\omega);$$

3) если η — ограниченная $\mathfrak{N}$-измеримая случайная величина, то

$$\mathsf{M}_x\{\theta_t\eta \mid \mathfrak{N}_t\} = \mathsf{M}_{\xi(t)}\eta \quad (\text{п. н. } \Omega_t, \mathsf{P}_x);$$

4) если величина $\varkappa$ ограничена и $\mathfrak{N}_t$-измерима, а η ограничена и $\mathfrak{N}$-измерима, то

$$\mathsf{M}_x(\varkappa\theta_t\eta) = \mathsf{M}_x(\varkappa \mathsf{M}_{\xi(t)}\eta).$$

15.2. Полугруппы операторов, связанные с однородными марковскими процессами

15.2.1. Полугруппа операторов, соответствующая вероятности перехода. Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — однородный марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода $P(t, x, \Gamma)$. Обозначим через $B(X)$ банахово пространство всех вещественных ограниченных $\mathfrak{B}$-измеримых функций на X с нормой $\|f\| = \sup_{x \in X} |f(x)|$. Рассмотрим на $B(X)$ семейство операторов T_t $(t \geqslant 0)$, определяемых формулой

$$T_t f(x) = \int_X f(y)\, P(t, x, dy) = \mathsf{M}_x f(\xi(t)), \quad f \in B(X).$$

Из свойств вероятности перехода следуют свойства семейства операторов T_t $(t \geqslant 0)$:

1) при каждом $t \geqslant 0$ T_t является линейным ограниченным оператором, отображающим $B(X)$ в $B(X)$, причем $\|T_t\| \leqslant 1$;

2) при всех $s, t \geqslant 0$ $T_{t+s} = T_t T_s$;

3) если $f(x) \geqslant 0$ при всех $x \in X$, то $T_t f(x) \geqslant 0$ при всех $x \in X$ и $t \geqslant 0$;

4) если $f(x_0) = 0$, то $T_0 f(x_0) = 0$;

5) если при всех $x \in X$ $\lim_{n\to\infty} f_n(x) = f(x)$, где $f_n \in B(X)$, причем $\sup_n \|f_n\| < \infty$, то $\lim_{n\to\infty} T_t f_n(x) = T_t f(x)$.

Семейство операторов $\{T_t, t \geqslant 0\}$, удовлетворяющих условиям 1), 2), называется *сжимающей полугруппой операторов*. Свойство 3) означает, что оператор T_t оставляет инвариантным конус неотрицательных функций в $B(X)$.

Таким образом, всякий однородный марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ порождает сжимающую полугруппу операторов $\{T_t, t \geqslant 0\}$, удовлетворяющую условиям 3) — 5). При этом эквивалентные марковские процессы порождают одну и ту же полугруппу.

Можно показать, что всякая сжимающая полугруппа операторов, действующих в $B(X)$ и удовлетворяющих условиям 3) — 5), порождает однородную вероятность перехода, причём $P(t, x, \Gamma) = T_t\chi_\Gamma(x)$.

Таким образом, для изучения марковских процессов можно использовать теорию полугрупп.

15.2.2. Инфинитезимальный оператор. Пусть $(X, \mathfrak{B})$ — некоторое измеримое пространство и $\{T_t, t \geqslant 0\}$ — сжимающая полугруппа операторов, действующих в $B(X)$. Определим *инфинитезимальный оператор* A полугруппы T_t формулой $Af = g$, если

$$\lim_{t \downarrow 0} \left\| g - \frac{T_t f - f}{t} \right\| = 0.$$

Его область определения D_A состоит из всех тех функций $f \in B(X)$, для которых предел

$$\lim_{t \downarrow 0} \frac{T_t f - f}{t}$$

существует равномерно относительно $x \in X$. Очевидно, что $D_A \subset \subset B_0(X) = \{f: f \in B(X), \lim\limits_{t \downarrow 0} \| T_t f - f \| = 0\}$.

Отметим некоторые свойства инфинитезимального оператора:

1) замыкание множества D_A (в смысле сходимости по норме) совпадает с $B_0(X)$;

2) если $f \in D_A$, то $Af \in B_0(X)$ и

$$T_t f - f = \int_0^t T_s Af \, ds;$$

3) если $f \in D_A$, то функция $T_t f$ сильно дифференцируема по t $(t \geqslant 0)$ и

$$\frac{dT_t f}{dt} = AT_t f = T_t Af;$$

4) оператор A замкнут.

Для положительных чисел λ определим операторы R_λ на $B_0(X)$ формулой

$$R_\lambda g(x) = \int_0^\infty e^{-\lambda t} T_t g(x) \, dt, \quad g \in B_0(X).$$

(Заметим, что для таких g функция $T_t g$ является сильно непрерывной ограниченной, и потому интеграл в этой формуле всегда существует при $\lambda > 0$ и определяет некоторую функцию из $B(X)$.) Семейство операторов R_λ называется *резольвентой* полугруппы T_t. Свойства резольвенты:

1) при $\lambda, \mu > 0$ выполняется резольвентное уравнение

$$R_\lambda R_\mu = \frac{1}{\lambda - \mu} [R_\mu - R_\lambda];$$

2) $\|R_\lambda\| \leqslant 1/\lambda$

3) функция $f = R_\lambda g$ является единственным решением уравнения $\lambda f - Af = g$, $\lambda > 0$, $g \in B_0(X)$.

Таким образом, $R_\lambda = (\lambda I - A)^{-1}$, где I — единичный оператор, и R_λ взаимно однозначно отображает $B_0(X)$ на D_A.

Положим $R_\lambda(x, \Gamma) = R_\lambda \chi_\Gamma(x)$, $\lambda > 0$, $x \in X$, $\Gamma \in \mathfrak{B}$. Функция $R_\lambda(x, \Gamma)$ называется *резольвентным ядром*. Очевидно, что

$$R_\lambda g(x) = \int_X R_\lambda(x, dy)\, g(y), \quad g \in B_0(X).$$

Если процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ прогрессивно измерим, т. е. если при всех $t \geqslant 0$ измеримо отображение $\xi(s, \omega)$ измеримого пространства $\left([0, t] \times \Omega_t, \mathscr{T}_t^0 \times \mathfrak{N}_t\right)$ (здесь $\mathscr{T}_t^0$ — σ-алгебра борелевских подмножеств отрезка $[0, t]$) в измеримое пространство $(X, \mathfrak{B})$, то резольвента полугруппы операторов, отвечающей этому процессу, может быть выражена формулой

$$R_\lambda g(x) = \mathsf{M}_x \int_0^\zeta e^{-\lambda t} g(\xi(t))\, dt, \quad g \in B(X), \quad x \in X, \quad \lambda > 0.$$

15.2.3. Стохастически непрерывные процессы в топологических пространствах. Пусть $P(t, x, \Gamma)$ — некоторая однородная вероятность перехода в пространстве $(X, \mathfrak{B})$. Согласно п. 15.2.1, она порождает полугруппу операторов T_t, действующих в пространстве $B(X)$. Инфинитезимальный оператор A этой полугруппы будем называть *инфинитезимальным оператором вероятности перехода* $P(t, x, \Gamma)$. Согласно п. 15.2.2,

$$Af(x) = \lim_{t \downarrow 0} \frac{\int_X f(y) P(t, x, dy) - f(x)}{t},$$

причем $f \in D_A$, если этот предел существует равномерно относительно $x \in X$.

Возникает вопрос: при каких условиях вероятность перехода однозначно определяется своим инфинитезимальным оператором?

Прежде чем сформулировать теорему, отвечающую на поставленный вопрос, введем понятие стохастически непрерывной вероятности перехода.

Пусть X — топологическое пространство, а $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X. Однородная вероятность перехода $P(t, x, \Gamma)$ называется *стохастически непрерывной*, если для всякого $x \in X$ и всякой окрестности U точки x выполнено соотношение

$$\lim_{t \downarrow 0} P(t, x, U) = 1.$$

Обозначим через $C(X)$ пространство вещественных непрерывных ограниченных функций, заданных на X. Если $P(t, x, \Gamma)$ — стохастически непрерывная вероятность перехода в пространстве $(X, \mathfrak{B})$, а

T_t — соответствующая ей полугрупппа операторов в $B(X)$, то для всякой $f \in C(X)$

$$\lim_{t \downarrow 0} T_t f(x) = f(x),$$

каково бы ни было $x \in X$. Более того, если некоторая полугруппа T_t, порождаемая вероятностью перехода $P(t, x, \Gamma)$, обладает этим свойством, то $P(t, x, \Gamma)$ стохастически непрерывна.

Теорема 1. *Всякая стохастически непрерывная вероятность перехода в топологическом фазовом пространстве однозначно определяется своим инфинитезимальным оператором.*

Заметим, что стохастически непрерывная вероятность перехода нормальна. Поэтому если A — инфинитезимальный оператор стохастически непрерывной вероятности перехода, то он однозначно (с точностью до эквивалентности) определяет некоторый марковский процесс. Таким образом, задача об описании всех стохастически непрерывных процессов в $(X, \mathfrak{B})$ сводится к описанию всех таких операторов в $B(X)$, которые являются инфинитезимальными операторами стохастически непрерывных вероятностей перехода.

Может оказаться, что полугруппа T_t, порожденная некоторой вероятностью перехода $P(t, x, \Gamma)$, оставляет инвариантным некоторое подпространство $\bar{B}$ пространства $B(X)$. Рассматривая полугруппу T_t в пространстве $\bar{B}$, мы можем определить ее инфинитезимальный оператор в $\bar{B}$, который называется *$\bar{B}$-инфинитезимальным оператором вероятности перехода* $P(t, x, \Gamma)$; во многих случаях он однозначно определяет эту вероятность.

Пусть X — топологическое пространство, $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств. Однородная вероятность перехода $P(t, x, \Gamma)$ в фазовом пространстве $(X, \mathfrak{B})$ называется *феллеровской*, если при всех $t > 0$ и $f \in C(X)$ имеет место включение

$$T_t f(x) = \int_X P(t, x, dy) f(y) \in C(X).$$

Другими словами, вероятность перехода феллеровская, если соответствующая ей полугруппа оставляет инвариантным пространство $C(X)$.

Марковский процесс с феллеровской вероятностью перехода называется *феллеровским*. Полугруппу, порожденную феллеровской вероятностью перехода, можно рассмотреть в пространстве $C(X)$. Ее инфинитезимальный оператор в этом пространстве называется *C-инфинитезимальным оператором* соответствующей вероятности перехода.

Теорема 2. *Если топологическое пространство X удовлетворяет первой аксиоме счетности, то C-инфинитезимальный оператор стохастически непрерывной феллеровской вероятности перехода определяет эту вероятность перехода однозначно.*

15.2.4. Процессы на компактах и полукомпактах. Рассмотрим теперь вопрос о том, какие же операторы могут быть инфинитезимальными для некоторой вероятности перехода.

Пусть сначала X — компакт, $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств, а $P(t, x, \Gamma)$ — стохастически непрерывная феллеровская вероятность перехода в $(X, \mathfrak{B})$. Можно показать, что в таком случае $C(X) \subseteq B_0(X)$ (определение пространства $B_0(X)$ см. в п. 15.2.2). Это означает, что для любой $f \in C(X)$ $\|T_t f - f\| \to 0$ при $t \to 0$.

Всякая стохастически непрерывная феллеровская вероятность перехода в компакте $(X, \mathfrak{B})$ является равномерно стохастически непрерывной в следующем смысле. Пусть ρ — некоторая метрика в пространстве X, порождающая его топологию, $U_\varepsilon(x)$ — шар в X радиуса ε с центром в точке $x \in X$. Вероятность перехода $P(t, x, \Gamma)$ в $(X, \mathfrak{B})$ называется *равномерно стохастически непрерывной*, если

$$\lim_{t \downarrow 0} \sup_{x \in X} (1 - P(t, x, U_\varepsilon(x))) = 0$$

при любом $\varepsilon > 0$.

Таким образом, в соответствии с теоремой 5 п. 14.2.6 среди эквивалентных между собой марковских процессов на компакте со стохастически непрерывной феллеровской вероятностью перехода существует процесс, не имеющий разрывов второго рода и непрерывный справа. Следующая теорема дает описание инфинитезимальных операторов таких процессов.

Теорема 3. *Пусть X — компакт, A — линейный оператор в пространстве $C(X)$. Для того чтобы оператор A являлся C-инфинитезимальным оператором некоторой стохастически непрерывной феллеровской вероятности перехода в $(X, \mathfrak{B})$, необходимо и достаточно выполнения следующих условий:*

1) *область определения D_A оператора A всюду плотна в $C(X)$ (в смысле равномерной метрики);*

2) *уравнение*

$$\lambda f - Af = g$$

имеет решение $f \in D_A$ при любом $g \in C(X)$ и $\lambda > 0$;

3) *если $f \in D_A$, $f(x_0) \geqslant 0$ и $f(x_0) \geqslant f(x)$ для всех $x \in X$, то $Af(x_0) \leqslant 0$.*

Назовем вероятность перехода *консервативной*, если при всех $t \geqslant 0$ и $x \in X$ $P(t, x, X) = 1$. Вероятность перехода является консервативной тогда и только тогда, когда ее инфинитезимальный оператор A обладает следующими свойствами: $1 \in D_A$ и $A1 = 0$.

Если A — инфинитезимальный оператор некоторой консервативной вероятности перехода, то выполнено следующее свойство:

3') если $f \in D_A$ и $f(x_0) \geqslant f(x)$ для всех $x \in X$, то $Af(x_0) \leqslant 0$.

Поэтому линейный оператор A в пространстве $C(X)$, где X — компакт, является C-инфинитезимальным оператором некоторой консервативной стохастически непрерывной феллеровской вероятности перехода тогда и только тогда, когда выполнены условия 1), 2), 3') и условие: $1 \in D_A$ и $A1 = 0$.

Пусть теперь X — полукомпакт (т. е. локально компактное хаусдорфово пространство со счетной базой), $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств. Обозначим через $C_0(X)$ пространство всех вещественных непрерывных функций на X, стремящихся к нулю, когда x выходит из всех компактов (это означает, что при любом $\varepsilon > 0$ множество $\{x\colon x \in X, |f(x)| \geqslant \varepsilon\}$ является компактным в X). Пусть $P(t, x, \Gamma)$ — вероятность перехода в пространстве $(X, \mathfrak{B})$. Будем говорить, что она удовлетворяет условию C_0, если $T_t f \in C_0(X)$ при всех $t > 0$ и $f \in C_0(X)$. Инфинитезимальный оператор полугруппы T_t в пространстве $C_0(X)$ назовем *C_0-инфинитезимальным оператором вероятности перехода $P(t, x, \Gamma)$*.

Теорема 4. *Стохастически непрерывная вероятность перехода в полукомпакте X, удовлетворяющая условию C_0, однозначно определяется своим C_0-инфинитезимальным оператором.*

Вероятность перехода, удовлетворяющая условиям теоремы 4, обладает следующими свойствами:

а) она равномерно стохастически непрерывна на компактах;

б) для нее $C_0(X) \subseteq B_0(X)$.

Таким образом, марковский процесс с вероятностью перехода, удовлетворяющей условиям теоремы 4, и в этом случае можно выбрать непрерывным справа и не имеющим разрывов второго рода.

Следующая теорема описывает инфинитезимальные операторы таких процессов.

Теорема 5. *Пусть X — полукомпакт, A — линейный оператор в пространстве $C_0(X)$. Для того чтобы оператор A являлся C_0-инфинитезимальным оператором некоторой стохастически непрерывной вероятности перехода в X, удовлетворяющей условию C_0, необходимо и достаточно, чтобы он удовлетворял условиям* 1) — 3) *теоремы* 3 (*в условиях* 1), 2) *нужно заменить $C(X)$ на $C_0(X)$*).

15.3. Характеристические операторы строго марковских процессов

15.3.1. Строго марковские процессы. Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — однородный марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с пространством элементарных событий Ω. Функция $\tau = \tau(\omega)$ $(\omega \in \Omega)$ с числовыми значениями называется *марковским моментом*, если выполнены условия:

а) $0 \leqslant \tau(\omega) \leqslant \zeta(\omega)$ для всех $\omega \in \Omega$;

б) при всех $t \geqslant 0$ $\{\tau(\omega) \leqslant t < \zeta(\omega)\} \in \mathfrak{N}_t$.

Обозначим $\Omega_\tau = \{\omega: \tau(\omega) < \zeta(\omega)\}$. Для $\omega \notin \Omega_\tau$ $\tau(\omega) = \zeta(\omega)$.

Очевидно, положив $\tau(\omega) = t_0$ для $\omega \in \Omega_{t_0}$ и $\tau(\omega) = \zeta(\omega)$ для $\omega \notin \Omega_{t_0}$, мы получим марковский момент (здесь t_0 — неслучайное число).

Положим далее

$$\tau_\Gamma(\omega) = \inf\{s: 0 \leqslant s < \zeta(\omega), \xi(s, \omega) \notin \Gamma\}, \quad \Gamma \in \mathfrak{B},$$

если множество в фигурных скобках непусто; в противном случае полагаем $\tau_\Gamma(\omega) = \zeta(\omega)$. Величина τ_Γ называется *моментом первого выхода* из множества Γ. Если X — топологическое пространство, процесс непрерывен справа, а множество Γ открыто или замкнуто, то величина τ_Γ является марковским моментом.

Пусть теперь $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — прогрессивно измеримый марковский процесс (это означает, что отображение $\xi(s, \omega)$ пространства $\left([0, t] \times \Omega_t, \mathcal{T}_t^0 \times \mathfrak{N}_t\right)$ в пространство $(X, \mathfrak{B})$ измеримо при всех t). Если τ — марковский момент, то отображение $\xi(\tau(\omega), \omega)$ определяет измеримое отображение пространства $(\Omega_\tau, \mathfrak{N}_\tau)$ в пространство $(X, \mathfrak{B})$. (Напомним, что $\mathfrak{N}_\tau$ — совокупность всех таких $A \in \mathfrak{N}$, что $A \cap \{\tau \leqslant t < \zeta\} \in \mathfrak{N}_t$.)

Прогрессивно измеримый марковский процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ называется *строго марковским*, если для любых $t \geqslant 0$, $x \in X$, $\Gamma \in \mathfrak{B}$ и любого марковского момента τ выполняется соотношение

$$\mathsf{P}_x\{\xi(t+\tau) \in \Gamma \mid \mathfrak{N}_\tau\} = P(t, \xi(\tau), \Gamma) \quad (\text{п. н. } \Omega_\tau, \mathsf{P}_x).$$

Следующая теорема дает достаточное условие строгой марковости процесса.

Теорема 1. *Непрерывный справа феллеровский марковский процесс в топологическом фазовом пространстве является строго марковским.*

Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — строго марковский процесс. Определим операторы сдвига θ_τ для $A \in \mathfrak{N}$ и марковского момента τ формулой

$$\theta_\tau A = \bigcup_{t \geqslant 0} (\theta_t A) \cap \{\tau = t\}.$$

Если η — $\mathfrak{N}$-измеримая случайная величина, то положим

$$\theta_\tau \eta(\omega) = \theta_t \eta(\omega)$$

для $\omega \in \{\tau(\omega) = t < \zeta(\omega)\}$. Функция $\theta_\tau\eta$ определена лишь на множестве Ω_τ.

Следующие свойства строго марковского процесса являются простыми следствиями определения. Если $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — строго марковский процесс и τ — марковский момент, то:

1) $\mathsf{P}_x\{\theta_\tau A \mid \mathfrak{N}_\tau\} = \mathsf{P}_{\xi(\tau)}(A)$ (п. н. Ω_τ, P_x) при всех $A \in \mathfrak{N}$;

2) для всех $A \in \mathfrak{N}_\tau$ и $B \in \mathfrak{N}$

$$\mathsf{P}_x\{A \cap \theta_\tau B\} = \int_A \mathsf{P}_{\xi(\tau)}(B)\, \mathsf{P}_x(d\omega);$$

3) для любой $\mathfrak{N}$-измеримой ограниченной случайной величины η

$$\mathsf{M}_x\{\theta_\tau\eta \mid \mathfrak{N}_\tau\} = \mathsf{M}_{\xi(\tau)}\eta \quad (\text{п. н. } \Omega_\tau, \mathsf{P}_x);$$

4) для любой $\mathfrak{N}_\tau$-измеримой ограниченной случайной величины η и любой $\mathfrak{N}$-измеримой ограниченной величины $\varkappa$

$$\mathsf{M}_x\{\eta\theta_\tau\varkappa\} = \mathsf{M}_x\{\eta \mathsf{M}_{\xi(\tau)}\varkappa\}.$$

В соответствии с определением п. 14.2.5 нормальный однородный марковский процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ в фазовом пространстве $(X, \mathfrak{B})$, где X — полукомпакт, а $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств, называется *стандартным*, если выполнены условия:

а) $\mathfrak{N}_t = \overline{\mathfrak{N}_{t+}}$, $t \geqslant 0$;

б) процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ непрерывен справа и имеет пределы слева;

в) процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ строго марковский;

г) процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ квазинепрерывен слева.

Теорема 2. *Если $P(t, x, \Gamma)$ — стохастически непрерывная феллеровская вероятность перехода на компакте или стохастически непрерывная вероятность перехода на полукомпакте, удовлетворяющая условию C_0, то существует стандартный марковский процесс с вероятностью перехода $P(t, x, \Gamma)$.*

Рассмотрим пример марковского процесса, не являющегося строго марковским.

Пусть $X = (-\infty, \infty)$, $\mathfrak{B}$ — σ-алгебра борелевских множеств в X. Положим для $t > 0$, $x \in X$, $\Gamma \in \mathfrak{B}$

$$P(t, x, \Gamma) = \begin{cases} \dfrac{1}{\sqrt{2\pi t}} \displaystyle\int_\Gamma \exp\left\{-\dfrac{(y-x)^2}{2t}\right\} dy, & \text{если } x \neq 0; \\ \chi_\Gamma(x), & \text{если } x = 0. \end{cases}$$

Нетрудно видеть что по этой вероятности перехода можно построить однородный непрерывный необрывающийся марковский процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$. Для него (конечная!) случайная величина $\tau = \tau_{X \setminus \{0\}}$ — момент первого выхода из множества $X \setminus \{0\}$ — является марковским моментом. Пусть множество A состоит из тех ω, для которых существует такое t, что для всех $s \geqslant t$ $\xi(s, \omega) = 0$. Тогда $\mathsf{P}_x(A) = 0$ при $x \neq 0$, а $\mathsf{P}_0(A) = 1$. Далее, очевидно, что $\theta_t A = A$. Поэтому равенство

$$\mathsf{P}_x(\theta_\tau A) = \mathsf{M}_x \mathsf{P}_{\xi(\tau)}(A)$$

не может выполняться, так как при $x \neq 0$ левая часть равна нулю, а правая — единице. Значит, процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ не может быть строго марковским.

15.3.2. Характеристический оператор. Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный справа феллеровский (а значит, и строго марковский) процесс в полукомпакте X. Если A — инфинитезимальный оператор этого процесса и $f \in D_A$, то имеет место формула

$$\mathsf{M}_x f(\xi(t)) - f(x) = \mathsf{M}_x \int_0^t Af(\xi(s))\, ds$$

при всех $t \geqslant 0$, $x \in X$ (см. п. 15.2.2, свойство 2) инфинитезимального оператора).

Оказывается, что в некоторых случаях эта формула остается справедливой при замене t на марковский момент τ. Именно, пусть τ — марковский момент для процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$, и пусть $\mathsf{M}_x \tau < \infty$. Тогда, если $f \in D_A$, то

$$\mathsf{M}_x f(\xi(\tau)) - f(x) = \mathsf{M}_x \int_0^\tau Af(\xi(s))\, ds. \tag{3.1}$$

Точку $x_0 \in X$ назовем *поглощающей*, если $\mathsf{P}_{x_0}\{\xi(t) = x_0\} = 1$ при всех $t \geqslant 0$. Для поглощающей точки x_0 $T_t f(x_0) = f(x_0)$ при всех t и всех $f \in B(X)$. Если x — непоглощающая точка, то всегда найдется такая ее окрестность U, что $\mathsf{M}_x \tau_U < \infty$. Здесь τ_U — момент первого выхода из открытого множества U.

Будем говорить, что последовательность окрестностей U_n $(n = 1, 2, \ldots)$ точки x сходится к x $(U_n \downarrow x)$, если для любой окрестности U точки x найдется такое n_0, что при $n > n_0$ $U_n \subset U$.

Пусть точка $x \in X$ — непоглощающая точка процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ и $U_n \downarrow x$. Тогда из формулы (3.1) следует соотношение

$$Af(x) = \lim_{n \to \infty} \frac{\mathsf{M}_x f(\xi(\tau_n)) - f(x)}{\mathsf{M}_x \tau_n}, \tag{3.2}$$

где $\tau_n = \tau_{U_n}$, если только $f \in D_A$ и функция Af непрерывна в точке x. В случае, когда x — поглощающая точка, $\tau_U = \infty$ для любой окрестности U точки x. Поэтому можно считать, что правая часть формулы (3.2) в поглощающей точке обращается в нуль. То же самое происходит и с левой частью.

Таким образом, если A — C-инфинитезимальный оператор феллеровского непрерывного справа процесса и $f \in D_A$, то значение функции Af при всех $x \in X$ может быть вычислено по формуле (3.2).

Обозначим через $D_{\mathfrak{A}}^x$ совокупность всех тех функций $f \in B(X)$, для которых при данном $x \in X$ существует предел

$$\mathfrak{A}f(x) = \lim_{n \to \infty} \frac{\mathsf{M}_x f(\xi(\tau_n)) - f(x)}{\mathsf{M}_x \tau_n},$$

где U_n — произвольная последовательность окрестностей точки x, сходящаяся к x, а $\tau_n = \tau_{U_n}$. Для $f \in \bigcap_{x \in X} D_{\mathfrak{A}}^x$ соответствующий предел существует при всех $x \in X$ и определяет некоторую функцию $\mathfrak{A}f(x)$. Оператор $\mathfrak{A}$ называется *характеристическим оператором процесса*. Его область определения $D_{\mathfrak{A}} = \bigcap_{x \in X} D_{\mathfrak{A}}^x$ состоит из всех тех функций $f \in B(X)$, для которых предел в правой части формулы (3.2) существует при всех $x \in X$. Из предыдущего следует, что характеристический оператор феллеровского непрерывного справа процесса в полукомпакте X является расширением его C-инфинитезимального оператора. Это означает, что $D_A \subseteq D_{\mathfrak{A}}$ и для $f \in D_A$ имеет место равенство $Af(x) = \mathfrak{A}f(x)$.

Если U — некоторое открытое подмножество X, $\tau = \tau_U$ — момент первого выхода из U, то положим

$$\pi_U(x, \Gamma) = \mathsf{P}_x\{\xi(\tau) \in \Gamma\}, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

Вероятность $\pi_U(x, \Gamma)$ называется *вероятностью выхода* из множества U. В терминах вероятностей выхода характеристический оператор может быть записан в виде

$$\mathfrak{A}f(x) = \lim_{n \to \infty} \frac{1}{\mathsf{M}_x \tau_{U_n}} \left(\int_X f(y)\, \pi_{U_n}(x, dy) - f(x) \right),$$

$$f \in D_{\mathfrak{A}}, \quad U_n \downarrow x.$$

Таким образом, характеристический оператор определяется вероятностями выхода и средним временем до выхода из сколь угодно малых окрестностей начальной точки. Если процесс не обрывается и непрерывен, то $\xi(\tau_U)$ принадлежит границе множества U. Поэтому для непрерывных процессов характеристический оператор локален.

Заметим, что для эквивалентных процессов характеристические операторы совпадают. Однако если для двух процессов совпадают их характеристические операторы, то отсюда еще не следует, что процессы эквивалентны.

В некоторых случаях можно более точно описать связь между инфинитезимальным и характеристическим операторами.

Теорема 3. 1) *Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — необрывающийся стохастически непрерывный феллеровский процесс в компакте X, и пусть A — его C-инфинитезимальный оператор. Тогда*

$$D_A = D_{\mathfrak{A}} \cap C(X) \cap \{f\colon \mathfrak{A}f \in C(X)\}.$$

2) *Пусть* $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — *необрывающийся стохастически непрерывный процесс в полукомпакте* X, *удовлетворяющий условию* C_0, *а пусть* A — *его* C_0*-инфинитезимальный оператор. Тогда*

$$D_A = D_{\mathfrak{A}} \cap C_0(X) \cap \{f: \mathfrak{A}f \in C_0(X)\}.$$

15.4. Процессы со счетным множеством состояний

15.4.1. Классификация точек. Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$, удовлетворяющий условию (S): если $\xi(s, \omega) = x$ для всех рациональных значений s из интервала (α, β), то $\xi(s, \omega) = x$ и для всех $s \in (\alpha, \beta)$. (Это условие выполнено, если процесс в топологическом пространстве непрерывен справа.)

Обозначим через τ_x момент первого выхода из точки x. Тогда τ_x — марковский момент и

$$\mathsf{P}_x\{\tau_x > t\} = e^{-a(x)t},$$

где $a(x)$ — некоторая неотрицательная функция от x, возможно, в некоторых точках равная $+\infty$.

Таким образом, для каждой точки $x \in X$ имеется три возможности:

1) $a(x) = 0$; в этом случае $\tau_x = +\infty$ и точка x называется *поглощающей* (см. п. 15.3.2);

2) $a(x) = +\infty$; в этом случае $\tau_x = 0$ и точка x называется *пропускающей*;

3) $0 < a(x) < +\infty$; в этом случае $0 < \tau_x < +\infty$ и точка x называется *задерживающей*.

Если считать $1/0 = +\infty$ и $1/+\infty = 0$, то во всех трех случаях очевидно, что $a(x) = (\mathsf{M}_x\tau_x)^{-1}$.

Марковский процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ называется *скачкообразным*, если для каждых ω и $t \in [0, \zeta(\omega))$ существует такое положительное δ (зависящее от t и ω), что при всех $h \in [0, \delta)$ $\xi(t) = \xi(t+h)$. Момент t_0 называется *моментом скачка траектории* $\xi(t, \omega)$, если существует такая последовательность $t_n \uparrow t_0$, что $\xi(t_n, \omega) \neq \xi(t_0, \omega)$ $(n = 1, 2, \ldots)$. Инфинитезимальный оператор скачкообразного процесса является сужением оператора, ставящего в соответствие функции $f \in B(X)$ функцию

$$-a(x)f(x) + a(x)\mathsf{M}_x f(\xi(\tau_x)) = -a(x)f(x) + a(x)\int_X f(y)\pi(x, dy),$$

где $a(x) = (\mathsf{M}_x\tau_x)^{-1}$, $\pi(x, \Gamma) = \mathsf{P}_x\{\xi(\tau_x) \in \Gamma\}$ (очевидно, что у скачкообразного процесса нет пропускающих точек, и потому $a(x) \neq \infty$ при $x \in X$).

Скачкообразный процесс называется *ступенчатым*, если при любом ω множество моментов скачков не имеет предельных точек внутри полуинтервала $[0, \zeta(\omega))$. Если $P(t, x, \Gamma)$ — вероятность перехода в фазовом пространстве $(X, \mathfrak{B})$ и если равномерно относительно $x \in X$

$$\lim_{t \downarrow 0} P(t, x, \{x\}) = 1,$$

то существует ступенчатый процесс с вероятностью перехода $P(t, x, \Gamma)$.

15.4.2. Дифференцируемость вероятностей перехода и уравнения Колмогорова. Рассмотрим подробнее процессы со счетным числом состояний. Пусть X — некоторое счетное (или конечное) множество и $\mathfrak{B}$ — σ-алгебра всех его подмножеств. В этом случае достаточно рассматривать вероятность перехода в одноточечные множества $P(t, x, y) = P(t, x, \{y\})$, поскольку $P(t, x, \Gamma) = \sum_{y \in \Gamma} P(t, x, y)$. Функция $P(t, x, y)$ $(t \geqslant 0,\ x,\ y \in X)$ удовлетворяет условиям:

а) $P(t, x, y) \geqslant 0$;

б) $\sum_{y \in X} P(t, x, y) \leqslant 1$;

в) $P(s+t, x, y) = \sum_{z \in X} P(s, x, z)\, P(t, z, y), \quad s, t \geqslant 0, \quad x, y \in X.$

Вероятность перехода $P(t, x, y)$ называется *стохастически непрерывной*, если выполнено условие:

г) при всех x, y $\lim_{t \downarrow 0} P(t, x, y) = \delta(x, y)$,

где $\delta(x, y) = 0$ для $x \neq y$ и $\delta(x, y) = 1$ для $x = y$.

Числа $P(t, x, y)$ $(x, y \in X)$, удовлетворяющие условиям а) — г), при всех $t \geqslant 0$ образуют матрицу $P(t)$, которая называется *полустохастической*. Если вместо условия б) выполнено условие

$$\sum_{y \in X} P(t, x, y) = 1$$

при всех $t \geqslant 0$ и $x \in X$, то матрица $P(t)$ называется *стохастической*. Очевидно, что

$$P(t)\, P(s) = P(s)\, P(t) = P(s+t).$$

Теорема 1. *Пусть заданы функции $P(t, x, y)$ $(t \geqslant 0,\ x,\ y \in X)$, удовлетворяющие условиям* а) — г). *Тогда:*

1) *при всех $x \neq y$ существуют конечные пределы*

$$a(x, y) = \lim_{t \downarrow 0} \frac{P(t, x, y)}{t};$$

2) *при каждом $x \in X$ существует предел*

$$a(x) = \lim_{t \downarrow 0} \frac{1 - P(t, x, x)}{t},$$

возможно, равный $+\infty$;

3) *при всех $x \in X$*

$$\sum_{\substack{y \in X \\ y \neq x}} a(x, y) \leqslant a(x).$$

Если процесс с вероятностью перехода $P(t, x, y)$ удовлетворяет условию (S) п. 15.4.1 и τ_x — момент первого выхода из состояния x, то

$$\mathsf{P}_x\{\tau_x > t\} = e^{-\bar{a}(x)t},$$

где $0 \leqslant \bar{a}(x) \leqslant +\infty$. Можно показать, что в рассматриваемом случае $a(x) = \bar{a}(x)$, где $a(x)$ введена в утверждении 2) теоремы 1. Таким образом, по функции $a(x)$ мы можем классифицировать все точки пространства X на пропускающие $(a(x) = +\infty)$, поглощающие $(a(x) = 0)$ и задерживающие $(0 < a(x) < +\infty)$.

Непропускающая точка x называется *регулярной*, если

$$a(x) = \sum_{y \neq x} a(x, y).$$

Очевидно, что поглощающие точки регулярны. Процесс, у которого все точки регулярны, называется *локально регулярным*.

Если при некотором $x \in X$ $a(x) < +\infty$, то при всех $y \in X$, $t \geqslant 0$ выполнены неравенства

$$\frac{\partial P(t, x, y)}{\partial t} \geqslant \sum_{z \in X} a(x, z) P(t, z, y), \tag{4.1}$$

$$\frac{\partial P(t, y, x)}{\partial t} \geqslant \sum_{z \in X} P(t, y, z) a(z, x), \tag{4.2}$$

где положено $a(x, x) = -a(x)$.

Теорема 2. *Если для функции $P(t, x, y)$ $(t \geqslant 0,\ x, y \in X)$, удовлетворяющей условиям* а) — г), *все точки регулярны, то выполнена система уравнений*

$$\frac{\partial P(t, x, y)}{\partial t} = \sum_{z \in X} a(x, z) P(t, z, y), \quad t \geqslant 0,\ y, x \in X. \tag{4.3}$$

Система уравнений (4.3) носит название *первой системы уравнений Колмогорова*. Начальным условием для нее служит соотношение

$$\lim_{t \downarrow 0} P(t, x, y) = \delta(x, y), \quad x, y \in X.$$

Обозначим через A матрицу с элементами $a(x, y)$ $(x, y \in X)$. Уравнения (4.3) в матричной форме имеют вид

$$\frac{\partial P(t)}{\partial t} = AP(t),$$

где $\partial P(t)/\partial t$ — матрица с элементами $\partial P(t, x, y)/\partial t$ $(x, y \in X)$.

Систему уравнений (4.3) можно записать еще и так:

$$P(t, x, y) = \delta(x, y) e^{-a(x)t} + \int_0^t e^{-a(x)s} \sum_{z \neq x} a(x, z) P(t-s, z, y)\, ds.$$

Для $y \neq x$ положим $\pi(x, y) = \dfrac{a(x, y)}{a(x)}$, если $0 < a(x) < +\infty$ и $\pi(x, y) = 0$ при $a(x) = 0$. Тогда предыдущая система интегральных уравнений может быть записана в виде

$$P(t, x, y) = \int_0^t a(x) e^{-a(x)s} \sum_{z \neq x} \pi(x, z) P(t-s, z, y)\, ds,$$

$$x \neq y, \; x, \; y \in X,$$

$$P(t, x, x) = e^{-a(x)t} + \int_0^t a(x) e^{-a(x)s} \sum_{z \neq x} \pi(x, z) P(t-s, z, x)\, ds, \quad x \in X.$$

Этим равенствам легко дать вероятностное истолкование. Например, второе из них может быть истолковано так: выйдя из состояния x, система в момент времени t может оказаться в состоянии x, либо не выходя из x за все это время (вероятность этого события равна $e^{-(ax)_0}$), либо в момент времени s, впервые выйдя из состояния x (вероятность этого события равна $a(x)e^{-a(x)s}ds$), перейдет в состояние $z \neq x$ (вероятность этого события равна $\pi(x, z)$), а затем за оставшееся время $t-s$ перейдет из состояния z в состояние x (вероятность этого события равна $P(t-s, z, x)$). При этом мы должны просуммировать (проинтегрировать) произведение этих вероятностей по всем возможным моментам первого выхода из x (т. е. по s от 0 до t) и по всем состояниям $z \neq x$. Таким образом, введенные величины $\pi(x, y)$ трактуются как вероятности того, что в момент первого выхода из состояния x система окажется в состоянии y.

Далее, если в (4.2) знак неравенства заменить на равенство, мы получим *вторую систему уравнений Колмогорова*:

$$\frac{\partial P(t, x, y)}{\partial t} = \sum_{z \in X} P(t, x, z)\, a(z, y), \quad x, \; y \in X. \tag{4.4}$$

Однако даже в случае, когда все точки $x \in X$ регулярны, вообще говоря, нельзя утверждать, что вероятности $P(t, x, y)$ (т. е. функции, удовлетворяющие условиям а) — г)) удовлетворяют системе уравнений (4.4). Достаточное условие для того, чтобы вероятности $P(t, x, y)$ удовлетворяли системе уравнений (4.4), содержится в следующем утверждении.

Теорема 3. *Предположим, что числа $P(t, x, y)$ $(t \geqslant 0, x, y \in X)$ образуют стохастическую матрицу, для которой $\sup a(x) < \infty$. Тогда все точки $x \in X$ регулярны и вероятности $P(t, x, y)$ удовлетворяют второй системе уравннеий Колмогорова.*

Можно показать, что ограниченность функции $a(x)$ эквивалентна условию

$$\lim_{t \downarrow 0} \sup_{x \in X} (1 - P(t, x, x)) = 0.$$

В этом случае, как следует из п. 15.4.1, процесс с вероятностью перехода $P(t, x, y)$ эквивалентен ступенчатому процессу.

Если X конечно и $P(t)$ — стохастическая матрица, то функция $a(x)$ ограничена, все точки $x \in X$ регулярны и выполнена вторая система уравнений Колмогорова.

15.4.3. Минимальное решение. До сих пор вероятность перехода $P(t, x, y)$ $(t \geqslant 0,\ x,\ y \in X)$ считалась заданной, и по ней строилась матрица $A = \|a(x, y)\|$ $(x,\ y \in X)$. Рассмотрим теперь обратную задачу. Пусть задана матрица $A = \|a(x, y)\|$ $(x,\ y \in X)$, удовлетворяющая условию:

A) при $x \neq y$ $a(x, y) \geqslant 0$ и $-\infty < \sum\limits_{y \in X} a(x, y) \leqslant 0$ при всех $x \in X$.

Существует ли вероятность перехода $P(t, x, y)$, для которой $\dfrac{\partial P(0, x, y)}{\partial t} = a(x, y)$ при всех $x,\ y \in X$?

Для того чтобы ответить на этот вопрос, естественно рассмотреть две системы дифференциальных уравнений:

$$\frac{\partial Q}{\partial t}(t, x, y) = \sum_{z \in X} a(x, z)\, Q(t, z, y), \quad x,\ y \in X, \tag{4.5}$$

$$\frac{\partial Q}{\partial t}(t, x, y) = \sum_{z \in X} Q(t, x, z)\, a(z, y), \quad x,\ y \in X, \tag{4.6}$$

с начальным условием $\lim\limits_{t \downarrow 0} Q(t, x, y) = \delta(x, y)$.

Положим

$$P^{(0)}(t, x, y) = \delta(x, y)\, e^{-a(x)t},$$

$$P^{(n+1)}(t, x, y) = \sum_{z \neq x} \int_0^t e^{-a(x)(t-s)} a(x, z)\, P^{(n)}(s, z, y)\, ds,$$

$$n = 0, 1, 2, \ldots$$

Можно показать, что функции $P^{(n)}(t, x, y)$ могут быть определены и такой системой равенств:

$$P^{(n+1)}(t, x, y) = \sum_{z \neq y} \int_0^t e^{-a(y)(t-s)} a(z, y)\, P^{(n)}(s, x, z)\, ds,$$

$$n = 0, 1, 2, \ldots$$

Пусть

$$\bar{P}(t, x, y) = \sum_{n=0}^{\infty} P^{(n)}(t, x, y).$$

Теорема 4. *Для любой матрицы A, удовлетворяющей условию А), построенная выше функция $\bar{P}(t, x, y)$ $(t \geqslant 0,\ x,\ y \in X)$ является решением систем уравнений* (4.5) *и* (4.6), *удовлетворяющим условиям* а) — г).

Можно показать, что если $P(t, x, y)$ $(t \geqslant 0,\ x,\ y \in X)$ — произвольная функция, удовлетворяющая условиям а) — г) и соотношению

$$\frac{\partial P}{\partial t}(0, x, y) = a(x, y), \quad x,\ y \in X,$$

с заданной матрицей A, то $P(t, x, y) \geqslant \bar{P}(t, x, y)$, где $\bar{P}(t, x, y)$ — функция, построенная выше по той же матрице A. Поэтому $\bar{P}(t, x, y)$ называется *минимальным решением*, соответствующим матрице A.

Если для данной матрицы A, удовлетворяющей условию А), минимальное решение $\bar{P}(t, x, y)$ обладает тем свойством, что $\sum\limits_{y \in X} \bar{P}(t, x, y) = 1$ (т. е. матрица $\bar{P}(t)$ стохастична), то всякая функция $P(t, x, y)$, удовлетворяющая условиям а) — г) и условию $\frac{\partial P}{\partial t}(0, x, y) = a(x, y)$ $(x, y \in X)$, совпадает с $\bar{P}(t, x, y)$. В частности, в этом случае $\bar{P}(t, x, y)$ является единственным решением систем уравнений (4.5) и (4.6), причем матрица A обязана удовлетворять условию $\sum\limits_{y \in X} a(x, y) = 0$ вместо соответствующего неравенства в условии А).

15.4.4. **Регулярные процессы.** Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — необрывающийся локально регулярный процесс, удовлетворяющий условию (S), и пусть $P(t, x, y)$ — его вероятность перехода. Положим $\tau_1 = \inf\{t\colon \xi(t) \neq \xi(0)\}$, причем если $\xi(t) = \xi(0)$ при всех t, то полагаем $\tau_1 = +\infty$. Момент τ_1 называется *моментом первого скачка*. Определим теперь последовательность моментов τ_n формулой

$$\tau_{n+1} = \tau_n + \theta_{\tau_n}\tau_1, \qquad n = 1, 2, \ldots,$$

причем если $\tau_k(\omega) = +\infty$, то считаем $\tau_j(\omega) = +\infty$ при всех $j \geqslant k$. Очевидно, что τ_n — момент n-го скачка. Положим $\tau_0 = 0$ и $\zeta_n = \xi(\tau_n)$ $(n = 0, 1, 2, \ldots)$. При этом, если $\tau_k(\omega) = +\infty$, полагаем $\zeta_k(\omega) = \zeta_{k-1}(\omega)$. Тогда последовательность $\{\zeta_n, n = 0, 1, 2, \ldots\}$ образует однородную цепь Маркова. Ее вероятность перехода за один шаг определяется соотношениями:

$$\pi(x, y) = \frac{a(x, y)}{a(x)}, \quad \text{если} \quad x \neq y \quad \text{и} \quad a(x) > 0;$$

$$\pi(x, x) = 0, \quad \text{если} \quad a(x) > 0;$$

$$\pi(x, y) = \delta(x, y), \quad \text{если} \quad a(x) = 0.$$

Назовем процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ *регулярным*, если при всех $x \in X$

$$\mathsf{P}_x\{\lim_{n \to \infty} \tau_n = +\infty\} = 1.$$

Регулярный процесс имеет лишь конечное число скачков на каждом конечном промежутке времени. Можно показать, что процесс регулярен тогда и только тогда, когда для всех $x \in X$

$$\mathsf{P}_x\left\{\sum_{k=0}^{\infty} \frac{1}{a(\zeta_k)} = +\infty\right\} = 1$$

(при этом считаем, что сумма бесконечна, если $a(\xi_k) = 0$ хотя бы при одном k). Отсюда вытекают два достаточных условия регулярности процесса:

1) для регулярности процесса достаточно, чтобы функция $a(x)$ была ограниченной;

2) для регулярности процесса достаточно, чтобы цепь Маркова $(\zeta_n,\ n = 0,\ 1,\ \ldots)$ была возвратной.

Далее, пусть $\dfrac{\partial P}{\partial t}(0,\ x,\ y) = a(x,\ y)$ $(x,\ y \in X)$. Так как процесс $(\xi(t),\ \mathfrak{N}_t,\ \mathsf{P}_x)$ локально регулярен, то $\sum\limits_{y \in X} a(x,\ y) = 0$. Поэтому функция $P(t,\ x,\ y)$ удовлетворяет уравнению Колмогорова. По матрице $A = \|a(x,\ y)\|$ определим функции $P^{(n)}(t,\ x,\ y)$ $(n = 0,\ 1,\ 2,\ \ldots)$ и $\bar{P}(t,\ x,\ y)$ так, как это было сделано выше. Тогда

$$P^{(0)}(t,\ x,\ y) = \mathsf{P}_x\left\{\tau_1 > t,\ \xi(t) = y\right\},$$

$$P^{(n)}(t,\ x,\ y) = \mathsf{P}_x\left\{\tau_n \leqslant t < \tau_{n+1},\ \xi(t) = y\right\} \quad n = 1,\ 2,\ \ldots,$$

$$\bar{P}(t,\ x,\ y) = \mathsf{P}_x\left\{\lim_{n\to\infty} \tau_n > t,\ \xi(t) = y\right\}.$$

Поэтому если процесс регулярен, то $P(t,\ x,\ y) = \bar{P}(t,\ x,\ y)$ и $P(t,\ x,\ y)$ является единственным решением первой системы уравнений Колмогорова. Справедливо и обратное утверждение: если первая система уравнений Колмогорова имеет единственное решение, то процесс регулярен. Кроме того, вероятность перехода регулярного процесса удовлетворяет второй системе уравнений Колмогорова.

Пусть $B(X)$ — пространство всех ограниченных вещественных функций на X. Если в X ввести дискретную топологию, то $B(X) = C(X)$, где $C(X)$ — пространство непрерывных ограниченных функций на X с вещественными значениями. Для локально регулярного (сепарабельного) процесса определен характеристический оператор

$$\mathfrak{A}f(x) = \frac{\mathsf{M}_x f(\xi(\tau_1)) - f(x)}{\mathsf{M}_x \tau_1} = \sum_{y \in X} a(x,\ y) f(y),$$

где $f \in B(X)$. Таким образом, действие оператора $\mathfrak{A}$ на функцию f сводится к умножению матрицы A на вектор-столбец $f(y)$ $(y \in X)$. При этом функция $\mathfrak{A}f(x)$ может оказаться неограниченной. Обозначим $D_{\mathfrak{A}} = \{f\colon f \in B(X),\ \mathfrak{A}f \in B(X)\}$. Мы видели, что по матрице A однозначно восстанавливается процесс до момента первого накопления скачков, т. е. до момента $\eta = \lim\limits_{n\to\infty} \tau_n$. Поэтому характеристический оператор однозначно определяет процесс (с точностью до эквивалентности) по крайней мере в двух случаях: 1) когда процесс регулярен; 2) когда он обрывается в момент η. В первом случае вероятность перехода является единственным решением первой (и второй) системы уравнений Колмогорова. Во втором случае вероятность перехода является минимальным решением указанных систем уравнений. (Разумеется, если $\eta = +\infty$, то первый и второй случаи совпадают.) Заметим, что во втором случае, если $\eta(\omega) < +\infty$, то $\xi(t,\ \omega)$ выходит из всех компактов при $t \uparrow \eta(\omega)$. Компактами в X являются множества, состоящие из конечного числа точек.

Для того чтобы по заданному характеристическому оператору $\mathfrak{A}$ (т. е. по заданной матрице A, удовлетворяющей условию А) с равенством вместо соответствующего неравенства) построить необрывающийся марковский процесс, нужно задать распределения положений процесса в моменты накопления скачков. Это можно сделать неоднозначно. Вопросы, связанные с построением марковских процессов, для которых заданный оператор $\mathfrak{A}$ является характеристическим, входят в так называемую теорию границ для марковских процессов.

15.4.5. Примеры. 1. *Процесс чистого роста.* Пусть X состоит из целых неотрицательных чисел, и пусть задана числовая последовательность $a(x)$ $(x = 0, 1, 2, \ldots)$ такая, что $0 < a(x) < \infty$. Положим $a(x, x+1) = a(x)$ $(x = 0, 1, 2, \ldots)$, $a(x, x) = -a(x)$ и $a(x, y) = 0$, если $y \neq x+1$, $y \neq x$. Тем самым задана матрица A, удовлетворяющая условию А), причем $\sum_{y \in X} a(x, y) = a(x, x) + a(x, x+1) = 0$.

Первая система уравнений Колмогорова имеет вид

$$\frac{\partial P(t, x, y)}{\partial t} = -a(x) P(t, x, y) + a(x) P(t, x+1, y), \; x = 0, 1, 2, \ldots$$

Запишем вторую систему:

$$\frac{\partial P(t, x, y)}{\partial t} = -a(y) P(t, x, y) + a(y-1) P(t, x, y-1), \; y = 1, 2, \ldots$$

Переходя к преобразованию Лапласа, легко получить

$$\varphi_p(x, y) = \left(\prod_{k=x}^{y-1} a(k)\right) \prod_{k=x}^{y} \frac{1}{p + a(k)},$$

где $\varphi_p(x, y) = \int_0^\infty e^{-pt} P(t, x, y)\, dt$ $(x, y \in X,\; x \leqslant y,\; p > 0)$.

Если среди чисел $a(k)$ нет равных между собой, то $P(t, x, y) = 0$ при $y < x$, $P(t, x, x) = 1 - e^{-a(x)t}$, а при $y > x$

$$P(t, x, y) = \left(\prod_{k=x}^{y-1} a(k)\right) \sum_{k=x}^{y} \frac{e^{-a(k)t}}{b(k)},$$

где $b(k) = \prod_{\substack{r=x \\ r \neq k}}^{y} (a(r) - a(k))$. Это минимальное решение.

Момент первого накопления скачков $\zeta = \lim_{n \to \infty} \tau_n$ представляет собой сумму независимых показательно распределенных случайных величин τ_i' $(i = 1, 2, \ldots)$, причем $\mathsf{P}_x\{\tau_i' > t\} = e^{-a(x+i-1)t}$. Поэтому при всех $x \in X$ $\mathsf{P}_x\{\zeta = +\infty\} = 1$ тогда и только тогда, когда расходится ряд $\sum_{x \in X} [a(x)]^{-1}$. Процесс называется *процессом линейного роста*, если $a(x) = x\lambda$ $(x = 1, 2, \ldots)$, λ — положительное число. Очевидно, что процесс линейного роста регулярен, и его вероятности перехода совпадают с минимальным решением уравнения Колмогорова.

2. *Процессы размножения и гибели.* Пусть X такое же, как и в предыдущем примере. Положим

$$a(x, y) = \begin{cases} -(\lambda_x + \mu_x), & \text{если } y = x, \quad x = 0, 1, 2, \ldots; \\ \lambda_x, & \text{если } y = x + 1, \quad x = 0, 1, 2, \ldots; \\ \mu_x, & \text{если } y = x - 1, \quad x = 1, 2, \ldots; \\ 0, & \text{если } |y - x| > 1. \end{cases}$$

Уравнения Колмогорова имеют вид

$$\frac{\partial P(t, x, y)}{\partial t} = -(\lambda_x + \mu_x) P(t, x, y) + \mu_x P(t, x-1, y) + \lambda_x P(t, x+1, y), \quad x = 0, 1, 2, \ldots \quad (\mu_0 = 0),$$

$$\frac{\partial P(t, x, y)}{\partial t} = -(\lambda_y + \mu_y) P(t, x, y) + \lambda_{y-1} P(t, x, y-1) + \mu_{y+1} P(t, x, y+1), \quad y = 0, 1, 2, \ldots \quad (\mu_0 = \lambda_{-1} = 0)$$

— соответственно первая и вторая системы.

В прикладных вопросах важную роль играют так называемые *стационарные вероятности*, т. е. числа $p(x)$ $(x \in X)$, удовлетворяющие условиям: $p(x) \geqslant 0$, $\sum_{x \in X} p(x) = 1$ и $p(y) = \sum_{x \in X} p(x) P(t, x, y)$ при всех $y \in X$ и $t \geqslant 0$. Можно показать, что если $\mu_x > 0$ при $x = 1, 2, \ldots$, то необходимым и достаточным условием существования стационарного распределения является сходимость ряда

$$\sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_k}.$$

Если этот ряд сходится, то

$$p(x) = \frac{\lambda_0 \lambda_1 \ldots \lambda_{x-1}}{\mu_1 \mu_2 \ldots \mu_x} p(0), \quad x = 1, 2, \ldots,$$

$$p(0) = \left(1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_k}\right)^{-1}.$$

Если стационарные вероятности существуют, то $p(y) = \lim_{t \to +\infty} P(t, x, y)$ $(y \in X)$, каково бы ни было $x \in X$.

15.5. Функционалы от марковских процессов

15.5.1. Мультипликативные функционалы. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$. Семейство вещественных функций $\alpha_t = \alpha_t(\omega)$ $(t \geqslant 0, \omega \in \Omega)$ называется *однородным мультипликативным функционалом*, если выполнены условия:

M1) α_t представляет собой $\overline{\mathfrak{N}}_t$-измеримую случайную величину при всех $t \geqslant 0$;

М2) при всех $t \geqslant 0$ и $h > 0$ почти наверное относительно меры P_x выполнено соотношение

$$\alpha_{t+h} = \alpha_h \theta_h \alpha_t,$$

каково бы ни было $x \in X$.

В § 14.3 рассматривались мультипликативные функционалы, удовлетворяющие дополнительному условию:

М3) при всех $t \geqslant 0$ и $x \in X$ почти наверное относительно P_x

$$0 \leqslant \alpha_t \leqslant 1.$$

Если процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ прогрессивно измерим, то примером мультипликативного функционала может служить семейство величин

$$\alpha_t = \exp\left\{\int_0^t v(\xi(s))\,ds\right\},$$

где $v(x)$ $(x \in X)$ — ограниченная измеримая функция с вещественными значениями. Если к тому же $v(x) \leqslant 0$ при всех $x \in X$, то α_t удовлетворяет также и условию М3).

15.5.2. Аддитивные функционалы. Семейство случайных величин $\varphi_t = \varphi_t(\omega)$ $(t \geqslant 0,\ \omega \in \Omega)$ называется однородным аддитивным функционалом от процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$, если:

А1) при всех $t \geqslant 0$ случайная величина φ_t $\overline{\mathfrak{N}}_t$-измерима;

А2) при всех $t \geqslant 0$ и $h > 0$ почти наверное относительно P_x выполнено соотношение

$$\varphi_{t+h} = \varphi_h + \theta_h \varphi_t,$$

каково бы ни было $x \in X$.

Примером аддитивного функционала является интеграл

$$\varphi_t = \int_0^t v(\xi(s))\,ds, \quad t \geqslant 0,$$

если $v(x)$ — измеримая ограниченная функция на X, а процесс $\xi(t)$ прогрессивно измерим.

Аддитивный функционал φ_t (слово «однородный» будем опускать) называется *неотрицательным*, если при всех $t \geqslant 0$ и $x \in X$ $\mathsf{P}_x\{\varphi_t \geqslant 0\} = 1$.

Два функционала φ_t и $\bar{\varphi}_t$ называются *эквивалентными*, если при всех $x \in X$, $t \geqslant 0$ $\mathsf{P}_x\{\varphi_t = \bar{\varphi}_t\} = 1$. Функционал φ_t $(t \geqslant 0)$ называется *непрерывным*, если при всех $x \in X$ почти наверное относительно P_x функции $\varphi_t(\omega)$ непрерывны как функции t при фиксированном ω. Для непрерывных однородных аддитивных функционалов при всех $x \in X$

$$\mathsf{P}_x\{\theta_h \varphi_t = \varphi_{t+h} - \varphi_t \text{ при всех } t \geqslant 0 \text{ и } h > 0\} = 1.$$

Теорема 1. *Пусть φ_t — неотрицательный непрерывный однородный аддитивный функционал от непрерывного справа процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$. Положим $g(t, x) = \mathsf{M}_x e^{-\varphi_t}$. Тогда при всех $x \in X$ и $t > 0$*

$$\varphi_t = \lim_{h \downarrow 0} \int_0^t \frac{1 - g(h, \xi(s))}{h} \, ds,$$

где предел понимается в смысле сходимости по вероятности P_x.

Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный справа марковский процесс в топологическом фазовом пространстве $(X, \mathfrak{B})$, где $\mathfrak{B}$ — σ-алгебра борелевских множеств в X, и пусть φ_t — неотрицательный однородный непрерывный аддитивный функционал от этого процесса. Если $f(x)$ $(x \in X)$ — неотрицательная борелевская функция, то положив

$$I_t(f, \varphi) = \int_0^t f(\xi(s)) \, d\varphi_s,$$

получим новый функционал такого же типа, что и φ_t. Заметим, что интеграл здесь понимается в смысле Лебега — Стилтьеса. Если функция $f(x)$ ограничена и непрерывна, то

$$I_t(f, \varphi) = \lim_{\max\limits_k \Delta s_k \to 0} \sum_{k=0}^{n-1} f(\xi(s_{k+1})) [\varphi(s_{k+1}) - \varphi(s_k)],$$

где $0 = s_0 < s_1 < \ldots < s_n = t$; $\Delta s_k = s_{k+1} - s_k$.

15.5.3. W-функционалы. Непрерывный неотрицательный однородный аддитивный функционал φ_t называется *W-функционалом*, если при всех $t \geq 0$

$$\sup_{x \in X} \mathsf{M}_x \varphi_t < \infty.$$

Если φ_t — W-функционал, то при всех натуральных n и $t \geq 0$

$$\sup_{x \in X} \mathsf{M}_x (\varphi_t)^n \leq n! \left(\sup_{x \in X} \mathsf{M}_x \varphi_t \right)^n.$$

Функция

$$f_t(x) = \mathsf{M}_x \varphi_t, \quad t \geq 0, \quad x \in X,$$

называется *характеристикой* W-функционала. Можно показать, что своей характеристикой W-функционал определяется однозначно, причем

$$\varphi_t = \lim_{\max \Delta s_k \downarrow 0} \sum_{k=0}^{n-1} f_{\Delta s_k}(\xi(s_k)),$$

где $0 = s_0 < s_1 < \ldots < s_n = t$, $\Delta s_k = s_{k+1} - s_k$, а предел понимается в смысле сходимости в среднеквадратическом по мере P_x при любом $x \in X$.

Характеристика W-функционала удовлетворяет условиям:

W1) функция $f_t(x)$ при фиксированном $t \geqslant 0$ является $\mathfrak{B}$-измеримой функцией от x, а при фиксированном $x \in X$ она непрерывна по t и монотонно не убывает; кроме того,

$$\lim_{t \downarrow 0} f_t(x) = 0, \quad \sup_{x \in X} f_t(x) < \infty;$$

W2) для всех $t > 0$, $s > 0$ и $x \in X$

$$f_{t+s}(x) = f_t(x) + \mathsf{M}_x f_s(\xi(t)).$$

Всякая функция $f_t(x)$, удовлетворяющая условиям W1) и W2), называется *W-функцией*.

Следующая теорема содержит необходимое и достаточное условие для того, чтобы W-функции соответствовал W-функционал.

Теорема 2. *Пусть X — полное метрическое сепарабельное пространство, $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный справа процесс, а функция $f_t(x)$ удовлетворяет условиям* W1) *и* W2). *Для того чтобы существовал W-функционал φ_t, для которого*

$$f_t(x) = \mathsf{M}_x \varphi_t,$$

необходимо и достаточно, чтобы для всех $x \in X$ и $t \geqslant 0$

$$\lim_{\delta \downarrow 0,\, h \downarrow 0} \mathsf{M}_x \frac{1}{h} \int_0^t f_h(\xi(s)) f_\delta(\xi(s))\, ds = 0.$$

Простое достаточное условие дает следующая теорема.

Теорема 3. *Пусть процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ удовлетворяет условиям предыдущей теоремы, а W-функция $f_t(x)$ удовлетворяет условию*

$$\lim_{t \downarrow 0} \sup_{x \in X} f_t(x) = 0.$$

Тогда существует такой W-функционал φ_t, что $f_t(x) = \mathsf{M}_x \varphi_t$, причем

$$\varphi_t = \lim_{h \downarrow 0} \int_0^t \frac{f_h(\xi(s))}{h}\, ds,$$

где предел понимается в смысле среднеквадратической сходимости по мере P_x при всех $x \in X$.

Пусть теперь $f_t(x)$ — произвольная W-функция. Всегда существует предел $f(x) = \lim_{t \to \infty} f_t(x)$, возможно равный бесконечности. Если $f(x) < \infty$ при $x \in X$, то $f(x)$ удовлетворяет условиям:

Э1) $T_t f(x) \leqslant f(x)$ при всех $x \in X$, $t \geqslant 0$; $f(x) \geqslant 0$;

Э2) $\lim_{t \downarrow 0} T_t f(x) = f(x)$ при всех $x \in X$.

Всякую функцию, удовлетворяющую условиям Э1) и Э2), называют *эксцессивной*. Если конечная эксцессивная функция $f(x)$ является пределом при $t \uparrow +\infty$ характеристики $f_t(x)$ некоторого W-функционала φ_t, то существует предел

$$\varphi_\infty = \lim_{t \uparrow +\infty} \varphi_t.$$

При этом
$$f(x) = \mathsf{M}_x\varphi_\infty.$$

По функции $f(x)$ легко восстановить характеристику W-функционала φ_t. Именно,
$$f_t(x) = f(x) - T_t f(x).$$

Таким образом, если для W-функционала $\varphi_t(x)$ существует φ_∞ и $f(x) = \mathsf{M}_x\varphi_\infty < \infty$, то функционал φ_t однозначно определяется функцией $f(x)$. Из теоремы 3 следует, что эксцессивная функция $f(x)$ определяет W-функционал, если
$$\lim_{t\downarrow 0}\sup_{x\in X}[f(x) - T_t f(x)] = 0.$$

15.5.4. Примеры. Рассмотрим некоторые примеры W-функционалов от винеровского процесса. Пусть $X = \mathbf{R}^m$, где $\mathbf{R}^m$ — евклидово пространство, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств $\mathbf{R}^m$. Непрерывный однородный марковский процесс с вероятностью перехода
$$P(t, x, \Gamma) = (2\pi t)^{-m/2}\int_\Gamma \exp\left\{-\frac{|y-x|^2}{2t}\right\}dy, \; t>0, \; x\in\mathbf{R}^m, \; \Gamma\in\mathfrak{B},$$
называется *винеровским*.

1. Пусть сначала $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — одномерный винеровский процесс. Положим
$$f_t(x) = \int_0^t (2\pi s)^{-1/2}\exp\left\{-\frac{(x_0-x)^2}{2s}\right\}ds, \quad t\geqslant 0, \quad x\in\mathbf{R},$$
где x_0 — фиксированная точка из $\mathbf{R}$. Легко проверить, что $f_t(x)$ — W-функция. Так как $f_t(x)\leqslant\sqrt{2t/\pi}$, то, согласно теореме 3, существует W-функционал φ_t, для которого
$$f_t(x) = \mathsf{M}_x\varphi_t, \quad t\geqslant 0, \quad x\in\mathbf{R}.$$

Далее, поскольку $f_t(x)t^{-1}\to\delta(x-x_0)$ при $t\downarrow 0$, где $\delta(x-x_0)$ — δ-функция Дирака (т. е. функция, для которой
$$\int_\mathbf{R}\delta(x-x_0)\,h(x)\,dx = h(x_0)$$
при любой $h\in C(\mathbf{R})$), функционал φ_t можно символически записать в виде
$$\varphi_t = \int_0^t \delta(\xi(s)-x_0)\,ds.$$

Этот функционал называется *локальным временем в точке* x_0. Нетрудно найти распределение функционала φ_t. Оно равно

$$\mathsf{P}_x\{\varphi_t < a\} = \frac{2}{\sqrt{\pi}} \int_0^{(a+|x-x_0|)/\sqrt{2t}} e^{-u^2} du, \quad 0 < a < \infty, \ x \in \mathbf{R}, \ t \geqslant 0.$$

При $a \leqslant 0$ $\mathsf{P}_x\{\varphi_t < a\} = 0$.

Локальное время получает приращение лишь в те моменты времени, в которые процесс $\xi(t)$ попадает в точку x_0. Можно показать, что

$$\varphi_t = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \chi_\varepsilon(|\xi(s) - x_0|)\, ds,$$

где $\chi_\varepsilon(x)$ — индикатор отрезка $[0, \varepsilon]$.

2. Пусть теперь $m \geqslant 1$ и $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — винеровский процесс в $\mathbf{R}^m$. Положим $S = \{x\colon x \in \mathbf{R}^m \ (x, \nu) = 0\}$, где ν — фиксированный вектор из $\mathbf{R}^m$ с $|\nu| \neq 0$, а (x, ν) — скалярное произведение в $\mathbf{R}^m$. Обозначим через $r(x)$ расстояние от x до гиперплоскости S. Легко проверить, что функция

$$f_t(x) = \int_0^t (2\pi s)^{-1/2} \exp\left\{-\frac{r^2(x)}{2s}\right\} ds, \quad t \geqslant 0, \quad x \in \mathbf{R}^m,$$

удовлетворяет условиям теоремы 3 и, следовательно, существует W-функционал φ_t с характеристикой $f_t(x)$. Он называется *локальным временем на гиперплоскости* S. Поскольку $\lim\limits_{t \downarrow 0} t^{-1} f_t(x) = \delta_S(x)$, где $\delta_S(x)$ определяется соотношением

$$\int_{\mathbf{R}^m} \delta_S(x)\, h(x)\, dx = \int_S h(x)\, d\sigma \tag{5.1}$$

(здесь h — произвольная непрерывная финитная функция, а интеграл справа — поверхностный), то функционал φ_t естественно записать в виде

$$\varphi_t = \int_0^t \delta_S(\xi(\tau))\, d\tau.$$

Его распределение определяется формулой

$$\mathsf{P}_x\{\varphi_t < a\} = \frac{2}{\sqrt{\pi}} \int_0^{(a+r(x))/\sqrt{2t}} e^{-u^2} du, \quad a > 0, \quad x \in \mathbf{R}^m, \quad t \geqslant 0.$$

3. Аналогично можно определить W-функционал

$$\varphi_t = \int_0^t \delta_S(\xi(\tau))\,d\tau$$

от m-мерного винеровского процесса для сферы S с центром в начале координат радиуса R. Это W-функционал с характеристикой

$$\mathsf{M}_x\varphi_t = \int_0^1 d\tau \int_S (2\pi\tau)^{-m/2} \exp\left\{-\frac{|y-x|^2}{2\tau}\right\} d\sigma_y.$$

Здесь второй интеграл — поверхностный. Функционал $\delta_S(x)$ определяется соотношением (5.1), где интеграл справа берется по поверхности сферы S. Функционал φ_t называется *локальным временем на сфере* S. Если $m \geqslant 3$, то существует $\varphi_\infty = \lim\limits_{t\to\infty} \varphi_t$. Функция распределения случайной величины φ_∞ имеет вид

$$\mathsf{P}_x\{\varphi_\infty < a\} = \begin{cases} 1 - \exp\left\{-\dfrac{m-2}{2R}a\right\}, & \text{если } |x| \leqslant R; \\ 1 - \left(\dfrac{R}{|x|}\right)^{m-2} + \left(\dfrac{R}{|x|}\right)^{m-2} \times \\ \qquad \times \left(1 - \exp\left\{-\dfrac{m-2}{2R}a\right\}\right), & \text{если } |x| > R, \end{cases}$$

где $x \in \mathbf{R}^m$, $0 \leqslant a < \infty$. При $a < 0$ $\mathsf{P}_x\{\varphi_\infty < a\} = 0$. Отметим также формулу

$$\mathsf{M}_x\varphi_\infty = \begin{cases} \dfrac{2R}{m-2}, & \text{если } |x| \leqslant R; \\ \dfrac{2R^{m-1}}{(m-2)|x|^{m-2}}, & \text{если } |x| > R. \end{cases}$$

15.6. Преобразования марковских процессов

15.6.1. Преобразование, связанное с мультипликативными функционалами. Рассмотрим некоторые преобразования марковских процессов. Один тип преобразований уже рассматривался в § 14.3. Он связан с мультипликативными функционалами от процесса. С помощью такого функционала процесс мог быть преобразован в некоторый подпроцесс. Рассмотрим более подробно такое преобразование в однородном случае.

Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$, и пусть α_t — мультипликативный функционал от этого процесса, т. е. семейство вещественных случайных величин, удовлетворяющих условиям M1), M2) (п. 15.5.1). Если, кроме того,

$$\mathsf{M}_x\alpha_t \leqslant 1 \tag{6.1}$$

при всех $t \geqslant 0$, $x \in X$, то формула

$$\widetilde{P}(t, x, \Gamma) = \mathsf{M}_x \chi_\Gamma(\xi(t)) \alpha_t, \quad t \geqslant 0, \quad x \in X, \quad \Gamma \in \mathfrak{B}, \tag{6.2}$$

определяет вероятность перехода в $(X, \mathfrak{B})$.

Если функционал α_t удовлетворяет условию М3) (п. 15.5.1), то он удовлетворяет и неравенству (6.1), однако существуют мультипликативные функционалы α_t, для которых условие М3) не выполняется, тем не менее для них $\mathsf{M}_x \alpha_t \leqslant 1$ при всех $t \geqslant 0$, $x \in X$.

Примеры. 1. Для положительной измеримой функции $g(x)$ $(x \in X)$ положим

$$\alpha_t = \frac{g(\xi(t))}{g(\xi(0))}.$$

Семейство величин α_t, определенное таким образом, удовлетворяет условиям М1), М2). Оно будет удовлетворять неравенству (6.1) тогда и только тогда, когда при всех $t \geqslant 0$, $x \in X$ выполнено неравенство $\mathsf{M}_x g(\xi(t)) \leqslant g(x)$, или, что то же самое, $T_t g(x) \leqslant g(x)$. В частности, такому неравенству удовлетворяют эксцессивные функции (см. п. 15.5.3). В рассматриваемом примере формула (6.2) приводит к вероятности перехода

$$\widetilde{P}(t, x, \Gamma) = \frac{1}{g(x)} \int_\Gamma P(t, x, dy)\, g(y),$$

где $P(t, x, \Gamma)$ — вероятность перехода исходного процесса. Отсюда получаем формулы для полугруппы операторов $\widetilde{T}_t$, соответствующей преобразованному процессу, инфинитезимального оператора $\widetilde{A}$ полугруппы $\widetilde{T}_t$, ее резольвенты $\widetilde{R}_\lambda$:

$$\widetilde{T}_t = g^{-1} T_t g, \quad \widetilde{A} = g^{-1} A g, \quad \widetilde{R}_\lambda = g^{-1} R_\lambda g,$$

где g обозначает оператор умножения на функцию $g(x)$ (здесь T_t, A, R_λ — соответствующие характеристики исходного процесса).

2. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — m-мерный винеровский процесс (см. п. 15.5.4), и пусть $g(x)$ $(x \in \mathbb{R}^m)$ — измеримая функция со значениями в $\mathbb{R}^m$, для которой при всех $x \in \mathbb{R}^m$, $t \geqslant 0$

$$\mathsf{P}_x \left\{ \int_0^t |g(\xi(s))|^2\, ds < \infty \right\} = 1.$$

Положим для $t \geqslant 1$

$$\alpha_t = \exp\left\{ \int_0^t (g(\xi(s)),\ d\xi(s)) - \frac{1}{2} \int_0^t |g(\xi(s))|^2\, ds \right\},$$

где первый из интегралов под знаком экспоненты — стохастический интеграл Ито (см. § 19.2). Легко видеть, что α_t удовлетворяет условиям М1), М2) и не удовлетворяет условию М3). Тем не менее нетрудно установить, что $\mathsf{M}_x \alpha_t \leqslant 1$ при всех $x \in \mathbb{R}^m$, $t \geqslant 0$, и, стало быть, формула (6.2) определяет вероятность перехода в $\mathbf{R}^m$. Во многих случаях α_t обладает тем свойством, что $\mathsf{M}_x\alpha_t = 1$ при всех $t \geqslant 0$, $x \in \mathbf{R}^m$. Для этого достаточно, чтобы при всех $t \geqslant 0$, $x \in \mathbf{R}^m$

$$\mathsf{M}_x \exp\left\{\frac{1}{2}\int_0^t |g(\xi(s))|^2\,ds\right\} < \infty.$$

Этому условию удовлетворяют, например, функции, представимые в виде суммы ограниченной функции и функции, для которой

$$\int_{\mathbf{R}^m} |g(x)|^p\,dx < \infty$$

при некотором $p > m$ в случае $m \geqslant 2$; при $m = 1$ считаем $p \geqslant 2$. Если функция g представима таким образом, то можно показать, что вероятность перехода $\tilde{P}(t, x, \Gamma)$ преобразованного процесса абсолютно непрерывна относительно меры Лебега в $\mathbf{R}^m$ и ее плотность $G(t, x, y)$ $(t > 0,\ x, y \in \mathbf{R}^m)$ удовлетворяет следующему интегро-дифференциальному уравнению:

$$G(t, x, y) = (2\pi t)^{-m/2}\exp\left\{-\frac{|y-x|^2}{2t}\right\} + \\ + \int_0^t ds\int_{\mathbf{R}^m}(2\pi s)^{-m/2}\exp\left\{-\frac{|z-x|^2}{2s}\right\}(g(z), \nabla_z G(t-s, z, y))\,dz,$$

где $\nabla_z G$ — градиент функции G как функции z. В частности, если функция $g(z)$ ограничена и удовлетворяет условию

$$|g(x) - g(y)| \leqslant K|x-y|^\alpha, \quad x, y \in \mathbf{R}^m,$$

при некоторых положительных постоянных K и $\alpha \leqslant 1$, то функция $G(t, x, y)$ является фундаментальным решением уравнения $(t > 0,$ $x \in \mathbf{R}^m)$

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2}\Delta u(t, x) + (g(x), \nabla_x u(t, x))$$

(см. также п. 15.7.2, 19.3.5). Здесь Δ — оператор Лапласа по переменной $x \in \mathbf{R}^m$.

В заключение этого пункта рассмотрим преобразование (6.2) в том случае, когда функционал α_t представим в виде $e^{-\varphi_t}$, где φ_t — конечный неотрицательный аддитивный однородный непрерывный функционал от стандартного процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ в полукомпакте $(X, \mathfrak{B})$. Введем для $\lambda > 0$ операторы S_λ, действующие на функцию $f \in B(X)$ по формуле

$$S_\lambda f(x) = \mathsf{M}_x\int_0^\infty e^{-\lambda t} f(\xi(t))\,d\varphi_t,$$

где интеграл справа понимается в смысле п. 15.5.2. Очевидно, что

$$|S_\lambda f(x)| \leqslant a_\lambda(x) \|f\|,$$

где

$$a_\lambda(x) = \mathsf{M}_x \int_0^\infty e^{-\lambda t}\, d\varphi_t = \lambda \int_0^\infty e^{-\lambda t} f_t(x)\, dt,$$

$f_t(x) = \mathsf{M}_x \varphi_t$ — характеристика функционала φ_t. Таким образом, если $a_\lambda(x) < \infty$, то $S_\lambda f(x)$ имеет смысл для любой функции $f \in B(X)$. Обозначим через R_λ резольвенту исходного процесса, а через $\widetilde{R}_\lambda$ резольвенту процесса, порождаемого вероятностью перехода

$$\widetilde{P}(t, x, \Gamma) = \mathsf{M}_x \chi_\Gamma(\xi(t))\, e^{-\varphi_t}, \quad t \geqslant 0, \quad x \in X, \quad \Gamma \in \mathfrak{B}.$$

Тогда

$$R_\lambda f(x) = \widetilde{R}_\lambda f(x) + S_\lambda \widetilde{R}_\lambda f(x), \quad f \in B(X)$$

при всех $\lambda > 0$, $x \in X$, для которых $a_\lambda(x) < \infty$.

Если к тому же функционал φ_t таков, что

$$\|f_t\| = \sup_{x \in X} \mathsf{M}_x \varphi_t \to 0,$$

когда $t \downarrow 0$, то при достаточно больших λ справедливо соотношение

$$\lambda I - \widetilde{A} = (\lambda I - A)(I + S_\lambda), \tag{6.3}$$

где I — единичный оператор, а A, $\widetilde{A}$ — инфинитезимальные операторы исходного и преобразованного процессов соответственно.

Рассмотрим частный случай, когда функционал φ_t является функционалом интегрального типа, т. е.

$$\varphi_t = \int_0^t v(\xi(s))\, ds,$$

где $v(x)$ — неотрицательная измеримая функция на $(X, \mathfrak{B})$. Предположим, что функция $v(x)$ обладает тем свойством, что $v \cdot f \in B_0(X)$ для всякой $f \in B_0(X)$ (определение пространства $B_0(X)$ см. в п. 15.2.2). Тогда из соотношения (6.3) нетрудно вывести, что $D_{\widetilde{A}} = D_A$ и $\widetilde{A} = A - v$, где v — оператор умножения на функцию $v(x)$.

Замечание. Если процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ прогрессивно измерим, а функция $v(x)$ $(x \in X)$ неотрицательна, то, положив

$$\alpha_t = \exp\left\{-\int_0^t v(\xi(s))\, ds\right\}, \quad t \geqslant 0,$$

найдем, что функция $\widetilde{P}(t, x, \Gamma)$, определенная формулой (6.2), будет удовлетворять интегральному уравнению п. 14.3.1, которое в однородном случае примет вид

$$\widetilde{P}(t, x, \Gamma) = P(t, x, \Gamma) - \int_0^t \int_X P(s, x, dy)\, v(y)\, \widetilde{P}(t - s, y, \Gamma)\, ds.$$

Здесь $P(t, x, \Gamma)$ — вероятность перехода исходного процесса.

15.6.2. Случайная замена времени. С аддитивными функционалами от процесса связано другое преобразование марковского процесса, которое называется *случайной заменой времени.* Опишем вкратце это преобразование.

Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — строго марковский непрерывный справа процесс в топологическом пространстве $(X, \mathfrak{B})$, где $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X. Предположим, что задан аддитивный однородный непрерывный функционал φ_t от этого процесса, удовлетворяющий условию

$$\mathsf{P}_x\{\varphi_t > 0\} = 1$$

при всех $t > 0$ и $x \in X$. Тогда можно выделить некоторое множество $\widetilde{\Omega} \subset \Omega$ такое, что при всех $x \in X$ $\mathsf{P}_x(\widetilde{\Omega}) = 1$ и для $\omega \in \widetilde{\Omega}$ функция $\varphi_t(\omega)$ будет непрерывной монотонно возрастающей функцией. Будем считать, что $\Omega = \widetilde{\Omega}$.

Рассмотрим теперь семейство марковских моментов τ_t $(t \geqslant 0)$, которое определяется соотношением

$$\tau_t(\omega) = \inf\{s\colon\ \varphi_s(\omega) \geqslant t\}$$

при $t \in [0, \tilde{\zeta}(\omega))$, где $\tilde{\zeta}(\omega) = \varphi_\infty(\omega) = \lim\limits_{t \uparrow \infty} \varphi_t(\omega)$. Положим $\eta(t) = \eta(t, \omega) = \xi(\tau_t(\omega), \omega)$ для $t \in [0, \tilde{\zeta}(\omega))$. Тогда можно показать, что набор объектов $\left(\eta(t), \tilde{\zeta}, \mathfrak{N}_{\tau_t}, \mathsf{P}_x\right)$ образует марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ и с пространством элементарных событий Ω. Говорят, что процесс $\left(\eta(t), \tilde{\zeta}, \mathfrak{N}_{\tau_t}, \mathsf{P}_x\right)$ получен из процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ случайной заменой времени, отвечающей функционалу φ_t. Можно показать, что если исходный процесс стандартен, то и преобразованный также будет стандартным.

Предположим теперь, что исходный процесс феллеровский и стохастически непрерывный. Покажем, как можно найти резольвенту преобразованного процесса, если замена времени произведена с помощью аддитивного функционала

$$\varphi_t = \int_0^t v(\xi(s))\, ds, \tag{6.4}$$

где $v(x)$ $(x \in X)$ — положительная борелевская функция. Очевидно, что для $f \in C(X)$ имеем $(\lambda > 0)$

$$\widetilde{R}_\lambda^{(v)} f(x) = \mathsf{M}_x \int_0^{\tilde{\zeta}(\omega)} e^{-\lambda s} f(\eta(s))\, ds = \mathsf{M}_x \int_0^\infty e^{-\lambda \varphi_t} f(\xi(t))\, d\varphi_t =$$

$$= \mathsf{M}_x \int_0^\infty \exp\left\{-\lambda \int_0^t v(\xi(s))\, ds\right\} f(\xi(t))\, v(\xi(t))\, dt.$$

Для $f, g \in B(X)$ положим

$$Q_\lambda(t, x, f, g) = \mathsf{M}_x f(\xi(t)) \exp\left\{-\lambda \int_0^t g(\xi(s))\, ds\right\}.$$

где $t \geqslant 0$, $x \in X$, λ — положительное число. Функция $Q_\lambda(t, x, f, g)$ является единственным решением уравнения

$$Q_\lambda(t, x, f, g) = \int_X f(y) P(t, x, dy) - \\ - \lambda \int_0^t ds \int_X Q_\lambda(t-s, y, f, g) g(y) P(s, x, dy).$$

Поэтому

$$\widetilde{R}_\lambda^{(v)} f(x) = \int_0^\infty Q_\lambda(t, x, fv, v)\, dt.$$

Отметим связь характеристических операторов исходного и преобразованного процессов. Предположим, что процесс $\left(\eta(t), \widetilde{\zeta}, \mathfrak{N}_{\tau_t}, \mathsf{P}_x\right)$ получен из процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ случайной заменой времени, отвечающей функционалу φ_t. Пусть функционал φ_t определен формулой (6.4) с положительной непрерывной функцией $v(x)$ $(x \in X)$. Обозначим через $\mathfrak{A}$ и $\widetilde{\mathfrak{A}}$ характеристические операторы соответственно исходного и преобразованного процессов. Тогда в каждой точке $x \in X$

$$D_{\mathfrak{A}}^x = D_{\widetilde{\mathfrak{A}}}^x \quad \text{и} \quad \widetilde{\mathfrak{A}} f(x) = [v(x)]^{-1} \mathfrak{A} f(x).$$

15.6.3. Преобразование фазового пространства. Рассмотрим преобразования марковских процессов, связанные с преобразованием фазового пространства. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ с вероятностью перехода $P(t, x, \Gamma)$. Пусть γ — измеримое отображение пространства $(X, \mathfrak{B})$ в измеримое пространство $(\widetilde{X}, \widetilde{\mathfrak{B}})$ такое, что $\gamma X = \widetilde{X}$, и образ измеримого множества при отображении γ измерим. Предположим далее, что для всякого $\widetilde{\Gamma} \in \widetilde{\mathfrak{B}}$ при любых $x, y \in X$, для которых $\gamma x = \gamma y$, справедливо равенство

$$P\left(t, x, \gamma^{-1}\widetilde{\Gamma}\right) = P\left(t, y, \gamma^{-1}\widetilde{\Gamma}\right), \quad t \geqslant 0.$$

Положим $\widetilde{\xi}(t) = \gamma\xi(t)$ и обозначим через $\widetilde{\mathfrak{N}}_t$ минимальную σ-алгебру событий, порожденную событиями вида $\{\widetilde{\xi}(s) \in \widetilde{\Gamma}\}$ при $s \leqslant t$, $\widetilde{\Gamma} \in \widetilde{\mathfrak{B}}$. Для $A \in \widetilde{\mathfrak{N}} = \widetilde{\mathfrak{N}}_\infty$ определим меры $\widetilde{\mathsf{P}}_{\gamma x}(A) = \mathsf{P}_x(A)$. Тогда набор объектов $(\widetilde{\xi}(t), \widetilde{\mathfrak{N}}_t, \widetilde{\mathsf{P}}_x)$ образует марковский процесс с вероятностью перехода $\widetilde{P}(t, \widetilde{x}, \widetilde{\Gamma}) = P(t, x, \gamma^{-1}\widetilde{\Gamma})$, где x — произвольная точка из множества $\gamma^{-1}\widetilde{x}$, $t \geqslant 0$, $\widetilde{\Gamma} \in \widetilde{\mathfrak{B}}$. Будем говорить, что этот процесс получен преобразованием фазового пространства γ из процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$.

Преобразование γ индуцирует отображение γ^* пространства $B(\widetilde{X})$ в пространство $B(X)$, определяемое формулой

$$\gamma^* f(x) = f(\gamma x), \quad f \in B(\widetilde{X}), \quad x \in X.$$

Если T_t и $\widetilde{T}_t$ — полугруппы операторов, соответствующие исходному и преобразованному процессам, а A и $\widetilde{A}$ — их инфинитезимальные операторы соответственно, то имеют место равенства

$$\gamma^*\widetilde{T}_t = T_t\gamma^*, \quad \gamma^*\widetilde{A} = A\gamma^*,$$

причем $f \in D_{\widetilde{A}}$ тогда и только тогда, когда $\gamma^* f \in D_A$.

Если X — топологическое пространство, а преобразование γ непрерывно и открыто (т. е. образ открытого множества есть открытое множество), то феллеровские процессы переходят при преобразовании γ в феллеровские.

Пример. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — одномерный винеровский процесс, а преобразование γ действует по формуле $\gamma x = |x|$ $(x \in \mathbf{R})$. Тогда $\widetilde{X} = [0, \infty)$. Нетрудно проверить, что все требования, наложенные на преобразование γ и вероятность перехода, выполняются, и мы в результате преобразования получаем процесс $(\widetilde{\xi}(t), \widetilde{\mathfrak{N}}_t, \mathsf{P}_x)$ в $[0, \infty)$, который называется *винеровским процессом с отражением в нуле*. Его вероятность перехода определяется формулой

$$P(t, x, \Gamma) = (2\pi t)^{-1/2} \int_\Gamma \left[\exp\left\{-\frac{(y-x)^2}{2t}\right\} + \exp\left\{-\frac{(y+x)^2}{2t}\right\}\right] dy,$$

где $t > 0$, $x \in [0, \infty)$, Γ — борелевское подмножество на полуоси $[0, \infty)$.

15.6.4. Инвариантные процессы. Предположим, что преобразование γ взаимно однозначно отображает X в X. Тогда вероятность перехода $\widetilde{P}(t, x, \Gamma)$ преобразованного процесса связана с вероятностью перехода $P(t, x, \Gamma)$ исходного соотношением $\widetilde{P}(t, x, \Gamma) = P(t, \gamma^{-1}x, \gamma^{-1}\Gamma)$. Марковский процесс с вероятностью перехода $P(t, x, \Gamma)$ называется *инвариантным относительно преобразования* γ, если выполнены условия:

а) для каждого $\omega \in \Omega$ существует такое $\omega' \in \Omega$, что $\gamma\xi(s, \omega) = \xi(s, \omega')$ при всех $s \geq 0$;

б) при всех $t > 0$, $x \in X$, $\Gamma \in \mathfrak{B}$

$$P(t, x, \Gamma) = P(t, \gamma^{-1}x, \gamma^{-1}\Gamma).$$

Определим оператор θ_γ, отображающий σ-алгебру $\mathfrak{N}_\infty$ в $\mathfrak{N}_\infty$, положив $\theta_\gamma\{\xi(t) \in \Gamma\} = \{\gamma\xi(t) \in \Gamma\} = \{\xi(t) \in \gamma^{-1}\Gamma\}$ и потребовав, чтобы θ_γ сохранял все теоретико-множественные операции. Можно показать, что если $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — марковский процесс, инвариантный относительно преобразования γ, то при всех $A \in \mathfrak{N}$ и $x \in X$

$$\mathsf{P}_{\gamma^{-1}x}(\theta_\gamma A) = \mathsf{P}_x(A).$$

Легко видеть, что m-мерный винеровский процесс инвариантен относительно всех движений евклидова пространства $\mathbf{R}^m$.

Пусть $U_R(x_0)$ — шар в $\mathbf{R}^m$ радиуса R с центром в точке x_0. Обозначим через τ_R момент первого выхода винеровского процесса из шара $U_R(x_0)$. Тогда из инвариантности винеровского процесса относительно всех движений легко вывести, что $\xi(\tau_R)$ по мере P_{x_0} имеет равномерное распределение на сфере, ограничивающей шар $U_R(x_0)$.

Поэтому

$$\mathsf{M}_{x_0} f(\xi(\tau_R)) = \frac{1}{\sigma_R^{(m)}} \int\limits_{S_R(x_0)} f(y)\, d\sigma,$$

где $S_R(x_0) = \{y\colon y \in \mathbf{R}^m, |y - x_0| = R\}$, $\sigma_R^{(m)} = \frac{2\pi^{m/2}}{\Gamma(m/2)} R^{m-1}$ — площадь сферы $S_R(x_0)$, а интеграл — поверхностный интеграл по сфере $S_R(x_0)$. Другими словами, $\mathsf{M}_{x_0} f(\xi(\tau_R))$ есть среднее от функции $f(y)$ по сфере $S_R(x_0)$. Далее, нетрудно найти

$$\mathsf{M}_x \tau_R = \frac{1}{m}(R^2 - |x - x_0|^2), \quad x \in U_R(x_0).$$

Таким образом, если $\mathfrak{A}$ — характеристический оператор m-мерного винеровского процесса и $f \in D_{\mathfrak{A}}$, то

$$\mathfrak{A} f(x_0) = \lim_{R \downarrow 0} \frac{m}{R^2} \frac{1}{\sigma_R^{(m)}} \int\limits_{S_R(x_0)} [f(y) - f(x_0)]\, d\sigma.$$

Оператор, стоящий в правой части этого равенства, называется *оператором Бляшке — Привалова.*

15.7. Однородные диффузионные процессы в евклидовых пространствах

15.7.1. Определение. Пусть $\mathbf{R}^m$ — евклидово пространство, $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств. Для $x \in \mathbf{R}^m$ обозначим через D_2^x совокупность всех вещественных функций, каждая из которых определена и дважды непрерывно дифференцируема в некоторой окрестности точки x. Непрерывный строго марковский процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ в фазовом пространстве $(\mathbf{R}^m, \mathfrak{B})$ называется *диффузионным*, если при всех $x \in \mathbf{R}^m$ имеет место включение: $D_2^x \subset D_{\mathfrak{A}}^x$, где $\mathfrak{A}$ — характеристический оператор процесса. Другими словами, характеристический оператор диффузионного процесса определен на всякой дважды непрерывно дифференцируемой функции в окрестности точки $x \in \mathbf{R}^m$.

Предположим, что в $\mathbf{R}^m$ выбран некоторый базис, и пусть $x^1, x^2, \ldots, x^m$ — координаты точки $x \in \mathbf{R}^m$ в этом базисе. Положим для $x, x_0 \in \mathbf{R}^m$ $\Delta^{ij}(x) = (x^i - x_0^i)(x^j - x_0^j)$, $\Delta^i(x) = x^i - x_0^i$ $(i, j = 1, 2, \ldots, m)$, где $x_0^1, x_0^2, \ldots, x_0^m$ — координаты точки x_0 в выбранном базисе. Функции $\Delta_0(x) \equiv 1$, $\Delta^i(x)$, $\Delta^{ij}(x)$ $(i, j = 1, 2, \ldots, m)$ дважды непрерывно дифференцируемы в окрестности точки x_0. Поэтому, если $\mathfrak{A}$ — характеристический оператор диффузионного процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$, то определены функции

$$b^{ij}(x_0) = \mathfrak{A}\Delta^{ij}(x_0), \quad i, j = 1, 2, \ldots, m,$$
$$a^i(x_0) = \mathfrak{A}\Delta^i(x_0), \quad i = 1, 2, \ldots, m,$$
$$c(x_0) = -\mathfrak{A}\Delta_0(x_0) = -\mathfrak{A}1(x_0).$$

Отсюда легко вывести следующее: если $f \in D_2^{x_0}$, то

$$\mathfrak{A}f(x_0) = \frac{1}{2} \sum_{i,j=1}^{m} b^{ij}(x_0) \frac{\partial^2 f(x_0)}{\partial x^i \partial x^j} + \sum_{i=1}^{m} a^i(x_0) \frac{\partial f(x_0)}{\partial x^i} - c(x_0) f(x_0).$$

При этом $c(x_0) \geqslant 0$, а матрица $b(x_0)$ с коэффициентами $b^{ij}(x_0)$ ($i, j = 1, 2, \ldots, m$) неотрицательно определена в том смысле, что $(b(x_0)\theta, \theta) \geqslant 0$ для всех $\theta \in \mathbf{R}^m$.

Далее, нетрудно заметить, что при переходе к другому базису оператор $\mathfrak{A}$ будет иметь такой же вид. Поменяются лишь коэффициенты $b^{ij}(x_0)$ и $a^i(x_0)$.

Таким образом, если процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ в $(\mathbf{R}^m, \mathfrak{B})$ — диффузионный, то существуют матричная функция $b(x)$, векторная функция $a(x)$ и неотрицательная числовая функция $c(x)$ такие, что сужение оператора $\mathfrak{A}$ на дважды непрерывно дифференцируемые функции является эллиптическим дифференциальным оператором второго порядка указанного выше вида. Функция $c(x)$ называется *коэффициентом обрыва*, вектор $a(x)$ — *вектором переноса*, а матрица $b(x)$ — *матрицей диффузии*.

Следующая теорема содержит достаточные условия, при выполнении которых процесс будет диффузионным.

Теорема 1. *Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный процесс в $(\mathbf{R}^m, \mathfrak{B})$. Предположим, что для каждой точки $x \in \mathbf{R}^m$ выполнены условия:*

1) *существует окрестность U_0 точки x такая, что $\mathsf{M}_x \tau_{U_0} < \infty$, где τ_{U_0} — момент первого выхода из U_0;*

2) *в некоторой системе координат существуют пределы*

$$\lim_{t \downarrow 0} \frac{1}{t} [1 - P(t, x, \mathbf{R}^m)] = c(x),$$

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{\mathbf{R}^m} (y^i - x^i) P(t, x, dy) = a^i(x), \quad i = 1, 2, \ldots, m,$$

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{\mathbf{R}^m} (y^i - x^i)(y^j - x^j) P(t, x, dy) = b^{ij}(x), \; i, j = 1, 2, \ldots, m,$$

причем отношения, стоящие под знаком пределов, равномерно ограничены при $x \in \mathbf{R}^m$, $t \geqslant 0$, а функции $c(x)$, $a(x)$ и $b(x)$ непрерывны в точке x.

Тогда процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ диффузионный и для него $c(x)$ — коэффициент обрыва, $a(x) = (a^1(x), \ldots, a^m(x))$ — вектор переноса, $b(x) = \|b^{ij}(x)\|$ — матрица диффузии.

15.7.2. Построение диффузионных процессов. Предположим, что заданы функции $c(x)$, $a(x) = (a^1(x), \ldots, a^m(x))$ и матрица $b(x) = \|b^{ij}(x)\|$ ($i, j = 1, 2, \ldots, m$, $x \in \mathbf{R}^m$). Пусть выполнены условия:

1) функции $a^i(x)$, $b^{ij}(x)$ и $c(x)$ ограничены и удовлетворяют условию Гёльдера на $\mathbf{R}^m$ (функция $f(x)$ удовлетворяет условию Гёльдера на $\mathbf{R}^m$, если существуют положительные постоянные K и $\alpha \leqslant 1$ такие, что

$$|f(x) - f(y)| \leqslant K |x - y|^\alpha, \quad x, y \in \mathbf{R}^m);$$

2) существует постоянная $\rho > 0$ такая, что при всех $x \in \mathbf{R}^m$ и $\theta \in \mathbf{R}^m$

$$(b(x)\theta, \theta) = \sum_{i,j=1}^{m} b^{ij}(x)\theta^i\theta^j \geqslant \rho\,|\theta|^2;$$

3) при всех $x \in \mathbf{R}^m$ $c(x) \geqslant 0$.

Теорема 2. *Предположим, что на $\mathbf{R}^m$ заданы функции $a(x)$, $b(x)$ и $c(x)$, удовлетворяющие условиям 1) — 3). Тогда существует диффузионный процесс $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ в фазовом пространстве $(\mathbf{R}^m, \mathfrak{B})$, для которого*

$$\mathfrak{A}f(x) = \frac{1}{2}\sum_{i,j=1}^{m} b^{ij}\frac{\partial^2 f(x)}{\partial x^i\,\partial x^j} + \sum_{i=1}^{m} a^i(x)\frac{\partial f(x)}{\partial x^i} - c(x) f(x),$$

где $f \in D_2^x$, а $\mathfrak{A}$ — характеристический оператор процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$. Соответствующая процессу полугруппа оставляет инвариантным пространство $C_0(\mathbf{R}^m)$ ($C_0(\mathbf{R}^m)$ — пространство непрерывных функций на $\mathbf{R}^m$, стремящихся к нулю при $|x| \to \infty$). Для любой ограниченной дважды непрерывно дифференцируемой функции $f(x)$ ($x \in \mathbf{R}^m$) функция

$$u(t, x) = \mathsf{M}_x f(\xi(t)) = T_t f(x)$$

дважды непрерывно дифференцируема по x, дифференцируема по t и удовлетворяет уравнению

$$\frac{\partial u}{\partial t} = \mathfrak{A}u, \quad t > 0, \quad x \in \mathbf{R}^m$$

с начальным условием $\lim_{t \downarrow 0} u(t, x) = f(x)$. Для вероятности перехода $P(t, x, \Gamma)$ процесса существует плотность $G(t, x, y)$ ($t > 0$, $x, y \in \mathbf{R}^m$) относительно лебеговой меры в $\mathbf{R}^m$, причем $G(t, x, y)$ является фундаментальным решением уравнения

$$\frac{\partial u}{\partial t} = \mathfrak{A}u.$$

Доказательство этой теоремы основано на свойствах фундаментальных решений параболических уравнений. В гл. 19 будут рассмотрены другие методы построения диффузионных процессов.

Диффузионный процесс называется *каноническим*, если его вектор переноса, матрица диффузии и коэффициент обрыва удовлетворяют условиям 1) — 3).

15.7.3. Вероятностные представления решений некоторых дифференциальных уравнений. Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — стандартный непрерывный процесс в полукомпакте $(X, \mathfrak{B})$. Обозначим через $\mathfrak{U}(D)$ ($D \subseteq X$) совокупность всех открытых подмножеств D, имеющих компактное замыкание. Через τ_Γ обозначим момент первого выхода процесса из множества Γ (см. п. 15.3.1).

Пусть D — открытое в X множество. Вещественную функцию $f(x)$, заданную на D, назовем *гармонической* для процесса $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ на множестве D, если она

а) $\mathfrak{B}$-измерима;

б) ограничена на каждом множестве $U \in \mathfrak{U}(D)$;

в) удовлетворяет условию

$$\mathsf{M}_x f(\xi(\tau_U)) = f(x)$$

при всех $x \in D$, $U \in \mathfrak{U}(D)$.

Следующая теорема дает описание гармонических функций для непрерывного стандартного процесса.

Теорема 3. *Пусть* $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — *непрерывный стандартный процесс,* τ — *момент первого выхода из открытого множества* D. *Если* $\varphi(x)$ $(x \in X)$ — *произвольная ограниченная* $\mathfrak{B}$-*измеримая функция, то функция*

$$f(x) = \mathsf{M}_x \varphi(\xi(\tau)), \quad x \in D,$$

является гармонической для процесса на множестве D *и для любого* $x \in D$

$$\lim_{t \uparrow \tau} f(\xi(t)) = \varphi(\xi(\tau)) \quad (п.\ н.\ \Omega_\tau, \mathsf{P}_x). \tag{7.1}$$

Если $f(x)$ $(x \in D)$ *является гармонической для процесса на множестве* D *и удовлетворяет при любом* $x \in D$ *условию* (7.1) *и неравенству*

$$|f(x)| < c\mathsf{P}_x\{\tau < \zeta\}$$

(c — *некоторая постоянная*), *то*

$$f(x) = \mathsf{M}_x \lim_{t \uparrow \tau} f(\xi(t)).$$

Пусть теперь D — открытое множество. Обозначим через τ' момент первого после $+0$ выхода из множества D, т. е.

$$\tau'(\omega) = \inf\{s:\ 0 < s < \zeta(\omega),\ \xi(s, \omega) \notin D\}.$$

Точку x_0 границы ∂D множества D назовем *регулярной*, если $\mathsf{P}_{x_0}\{\tau' > 0\} = 0$.

Следующие две теоремы характеризуют решение задачи Дирихле для дифференциальных уравнений, связанных с каноническим диффузионным процессом.

Эллиптический дифференциальный оператор, фигурирующий в теореме 2, обозначим буквой $\mathfrak{L}$. Он называется *производящим дифференциальным оператором* диффузионного процесса.

Теорема 4. *Пусть* $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — *канонический диффузионный процесс в* $(\mathsf{R}^m, \mathfrak{B})$ *с характеристическим оператором* $\mathfrak{A}$ *и производящим дифференциальным оператором* $\mathfrak{L}$. *Совокупность всех функций, гармонических для этого процесса на открытом множестве* D, *совпадает с множеством всех непрерывных решений уравнения*

$$\mathfrak{A}f(x) = 0, \quad x \in D,$$

а также с множеством всех дважды непрерывно дифференцируемых функций, удовлетворяющих дифференциальному уравнению

$$\mathfrak{L}f(x) = 0, \quad x \in D. \tag{7.2}$$

Если все точки границы ∂D *множества* D *регулярны,* $D \cup \partial D$ — *компакт и функция* φ *непрерывна на* ∂D, *то функция*

$$f(x) = \mathsf{M}_x \varphi(\xi(\tau)), \quad x \in D$$

(τ — момент первого выхода из D) является единственным решением дифференциального уравнения (7.2), удовлетворяющим граничному условию

$$\lim_{x \to x_0} f(x) = \varphi(x_0), \quad x_0 \in \partial D.$$

Теорема 5. *Пусть $(\xi(t), \zeta, \mathfrak{N}_t, \mathsf{P}_x)$ — канонический диффузионный процесс с производящим дифференциальным оператором $\mathfrak{L}$, D — ограниченное открытое множество с регулярной границей, τ — момент первого выхода из D. Предположим, что заданы непрерывная функция φ на границе ∂D множества D и функции v и g на D, удовлетворяющие условию Гёльдера, причем $v \geqslant 0$.*

Тогда функция

$$f(x) = \mathsf{M}_x \int_0^{\tau} g(\xi(t)) \exp\left\{-\int_0^{t} v(\xi(s))\, ds\right\} dt +$$
$$+ \mathsf{M}_x \varphi(\xi(\tau)) \exp\left\{-\int_0^{\tau} v(\xi(s))\, ds\right\}, \quad x \in D,$$

дважды непрерывно дифференцируема на D и является единственным решением дифференциального уравнения

$$\mathfrak{L} f(x) - v(x) f(x) = -g(x), \quad x \in D,$$

удовлетворяющим граничному условию

$$\lim_{x \to x_0} f(x) = \varphi(x_0), \quad x_0 \in \partial D.$$

15.7.4. Пример, показывающий, что существуют непрерывные марковские процессы, не являющиеся диффузионными. Определим семейство операторов T_t $(t > 0)$, действующих в пространстве $B(\mathbf{R})$, по формуле

$$T_t f(x) = \int_{\mathbf{R}} \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} f(y)\, dy +$$
$$+ \frac{c}{\sqrt{2\pi t}} \int_0^{\infty} \exp\left\{-\frac{(y+|x|)^2}{2t}\right\} [f(y) - f(-y)]\, dy, \quad x \in \mathbf{R}, \quad f \in B(\mathbf{R}),$$

где c — некоторое вещественное число, $|c| \leqslant 1$. Можно показать, что семейство операторов $\{T_t, t > 0\}$ образует полугруппу операторов, которой соответствует непрерывный марковский процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ в пространстве $(\mathbf{R}, \mathfrak{B})$. При $c = 0$ это будет винеровский процесс. Если $c = 1$, то до момента первого попадания на полуось $(0, \infty)$ этот процесс ведет себя как винеровский процесс, а после этого момента — как винеровский процесс с отражением в нуле. Аналогичное описание справедливо и при $c = -1$. При $0 < |c| < 1$ получаем некоторые «промежуточные» процессы.

Нетрудно показать, что если функция $f(x)$ дважды непрерывно дифференцируема в окрестности точки $x_0 \neq 0$, то $f \in D_{\mathfrak{A}}^{x_0}$ и $\mathfrak{A}f(x_0) = (1/2) f''(x_0)$, где $\mathfrak{A}$ — характеристический оператор процесса. Если же $f(x)$ дважды непрерывно дифференцируема в окрестности точки $x = 0$, то при $c \neq 0$ $f \in D_{\mathfrak{A}}$ лишь в том случае, если $f'(0) = 0$, и при этом $\mathfrak{A}f(0) = (1/2)f''(0)$. Таким образом, рассматриваемый процесс при $c \neq 0$ не является диффузионным. Диффузионный характер движения нарушается в точке $x = 0$.

Рассмотренный процесс является обобщенным диффузионным процессом в следующем смысле. Положим

$$a(t, x) = \frac{1}{t} \mathsf{M}_x (\xi(t) - x) = \frac{1}{t} \int_{\mathbf{R}} (y - x) P(t, x, dy),$$

$$b(t, x) = \frac{1}{t} \mathsf{M}_x (\xi(t) - x)^2 = \frac{1}{t} \int_{\mathbf{R}} (y - x)^2 P(t, x, dy), \quad t > 0, \quad x \in \mathbf{R}.$$

Тогда для любой непрерывной финитной функции $\varphi(x)$ $(x \in \mathbf{R})$

$$\lim_{t \downarrow 0} \int_{\mathbf{R}} \varphi(x)\, a(t, x)\, dx = c\varphi(0).$$

Для функции $b(t, x)$ имеем соотношение

$$\lim_{t \downarrow 0} b(t, x) = 1$$

при всех $x \in \mathbf{R}$, причем $|b(t, x)| \leqslant K$ при всех $x \in \mathbf{R}$ и $t > 0$, где K — некоторая постоянная. Первое из этих соотношений означает, что «коэффициент переноса» рассматриваемого процесса равен $c\delta(x)$, где $\delta(x)$ — δ-функция Дирака. Второе соотношение показывает, что коэффициент диффузии равен единице.

15.8. Непрерывные процессы на прямой

15.8.1. Регулярные точки. Тот факт, что прямая $\mathbf{R}$ является упорядоченным множеством, позволяет описать все непрерывные строго марковские процессы со значениями в $\mathbf{R}$.

Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный строго марковский процесс в некотором интервале $\Delta \subset \mathbf{R}$. Точка $y \in \Delta$ называется *достижимой из точки* $x \in \Delta$, если $\mathsf{P}_x\{\tau_y < \infty\} > 0$, где τ_y — момент первого достижения точки y (это марковский момент). Обозначим через Δ_x совокупность всех тех $y \in \Delta$, которые достижимы из x. Тогда Δ_x — интервал (замкнутый, открытый или полуоткрытый, конечный или бесконечный). Точку $x \in \Delta$ назовем *регулярной*, если выполнены условия:

1) x является внутренней точкой интервала Δ_x;

2) существуют такие $x_1, x_2 \in \Delta_x$, что $x_1 < x < x_2$ и точка x достижима из точек x_1 и x_2.

О поведении процесса в регулярной точке можно судить по следующей теореме.

Теорема 1. *Если x — регулярная точка, то для всякого $\delta > 0$*

$$\mathsf{P}_x\left\{\sup_{0 \leqslant s \leqslant \delta} \xi(s) > x\right\} = 1, \quad \mathsf{P}_x\left\{\inf_{0 \leqslant s \leqslant \delta} \xi(s) < x\right\} = 1.$$

15.8.2. Процессы на замкнутом интервале. Опишем все непрерывные строго марковские процессы на интервале (α, β), предполагая, что все точки этого интервала регулярны, точки α и β достижимы изнутри, а произвольная точка $x \in (\alpha, \beta)$ достижима из точек α и β. Обозначим через ζ момент первого выхода из множества (α, β). Тогда либо $\xi(\zeta) = \alpha$, либо $\xi(\zeta) = \beta$. Положим $m(x) = \mathsf{P}_x\{\xi(\zeta) = \beta\}$.

Теорема 2. *Функция $m(x)$ непрерывна, строго монотонно возрастает и $m(\alpha) = 0$, $m(\beta) = 1$. Процесс $\{m(\xi(t \wedge \zeta)) - m(\xi(0)), \mathfrak{N}_t, \mathsf{P}_x\}$ является непрерывным мартингалом с интегрируемым квадратом, характеристика которого представляет собой W-функционал от процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$.*

Из этой теоремы следует, что если сделать преобразование фазового пространства при помощи функции $m(x)$, то мы получим процесс на отрезке $[0, 1]$, для которого $m(x) = x$.

Будем считать, что такая замена уже проделана и, следовательно, мы рассматриваем процесс на отрезке $[0, 1]$, для которого

$$\mathsf{P}_x\{\xi(\zeta) = 1\} = x, \quad \mathsf{P}_x\{\xi(\zeta) = 0\} = 1 - x.$$

Рассмотрим функцию $n(x) = \mathsf{M}_x\zeta$. Можно показать, что при любом $k = 1, 2, \ldots$ функция $\mathsf{M}_x\zeta^k$ ограничена на $[0, 1]$. Нетрудно проверить, что $n(x)$ — строго выпуклая вверх функция, и если τ — момент первого выхода из интервала $(x_0 - \varepsilon, x_0 + \delta)$ $(0 < x_0 - \varepsilon < x_0 + \delta < 1, \varepsilon, \delta > 0)$, то

$$\mathsf{M}_x\tau = n(x_0) - n(x_0 - \varepsilon)\frac{\delta}{\varepsilon + \delta} - n(x_0 + \delta)\frac{\varepsilon}{\varepsilon + \delta}.$$

Так как

$$\mathsf{M}_{x_0}f(\xi(\tau)) = f(x_0 - \varepsilon)\frac{\delta}{\delta + \varepsilon} + f(x_0 + \delta)\frac{\varepsilon}{\varepsilon + \delta},$$

то для характеристического оператора в точке $x_0 \in (0, 1)$ имеем соотношение

$$\mathfrak{A}f(x_0) = \lim_{\varepsilon \downarrow 0,\, \delta \downarrow 0} \frac{\varepsilon^{-1}[f(x_0 - \varepsilon) - f(x_0)] + \delta^{-1}[f(x_0 + \delta) - f(x_0)]}{\delta^{-1}[n(x_0) - n(x_0 + \delta)] - \varepsilon^{-1}[n(x_0 - \varepsilon) - n(x_0)]}.$$

Заметим, что существует производная $n'(x)$, представляющая собой убывающую функцию от x. Далее, можно показать, что если функция $f(x)$ абсолютно непрерывна и для нее существует такая непрерывная функция $f(t)$, что

$$f'(x) = f'(0) + \int_0^x g(t)\,dn'(t), \tag{8.1}$$

то при всех $x \in (0, 1)$

$$\mathfrak{A}f(x) = -g(x).$$

Функция $g(t)$, удовлетворяющая соотношению (8.1), является производной df'/dn'. Поэтому при $x \in (0, 1)$

$$\mathfrak{A}f(x) = -\frac{df'(x)}{dn'(x)},$$

причем характеристический оператор определен на всех таких функциях $f(x)$, которые абсолютно непрерывны и имеют непрерывную производную df'/dn'. Наконец, заметим, что если процесс обрывается в момент первого выхода из интервала $(0, 1)$, то его инфинитезимальный оператор A определен на всех абсолютно непрерывных функциях f, для которых df'/dn' непрерывно на $[0, 1]$, а $f(0) = f(1) = 0$. При этом $Af = \mathfrak{A}f$.

15.8.3. Процессы на интервале регулярности. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — непрерывный строго марковский процесс в интервале $\Delta \subset \mathbf{R}$. Пусть $x \in \Delta$ — регулярная точка для этого процесса. Положим

$$\alpha = \inf \{y\colon\ y \in \Delta_x,\ \ \mathsf{P}_y \{\tau_x < \infty\} > 0\},$$
$$\beta = \sup \{y\colon\ y \in \Delta_x,\ \ \mathsf{P}_y \{\tau_x < \infty\} > 0\}.$$

Так как x — регулярная точка, то интервал (α, β) непуст и содержит x. Он называется *интервалом регулярности* процесса, содержащим точку x.

Рассмотрим поведение процесса в интервале регулярности (α, β), предполагая, что точки α и β нерегулярны. Для границы α интервала (α, β) возможны четыре случая:

1) точка α достижима изнутри интервала и всякая точка $x \in (\alpha, \beta)$ достижима из точки α; в этом случае точка α называется *регулярной границей*;

2) точка α достижима изнутри, но из нее точки интервала недостижимы; в этом случае α называется *захватывающей границей*;

3) точка α недостижима изнутри, но из нее достижимы точки интервала; такая точка называется *выпускающей границей*;

4) точка α недостижима изнутри и из нее недостижимы точки интервала; такая точка называется *естественной границей*.

Такие же возможности имеются и у граничной точки β.

Пусть α — регулярная граница. Так как точка α не является регулярной (в смысле данного выше определения), то для нее при всех $y < \alpha$, $y \in \Delta$ либо $\mathsf{P}_y \{\tau_\alpha < \infty\} = 0$, либо $\mathsf{P}_\alpha \{\tau_y < \infty\} = 0$. В первом случае α называется *недостижимой слева*, во втором случае — *непроходимой налево*. Непроходимая налево регулярная граничная точка α называется *отражающей границей*. Следующая теорема указывает вид инфинитезимального оператора процесса с двумя отражающими границами.

Теорема 3. *Если $(\tilde{\xi}(t), \tilde{\mathfrak{N}}_t, \tilde{\mathsf{P}}_x)$ — процесс на $[0, 1]$, для которого $m(x) = x$ и точки 0 и 1 являются отражающими границами интервала регулярности $(0, 1)$, то во всех точках $x \in (0, 1)$*

$$Af(x) = -\frac{df'(x)}{dn'(x)},$$

а в граничных точках

$$Af(0) = \lim_{\varepsilon \downarrow 0} \frac{f(\varepsilon) - f(0)}{\mathsf{M}_0 \tau_\varepsilon}, \quad Af(1) = \lim_{\varepsilon \downarrow 0} \frac{f(1-\varepsilon) - f(1)}{\mathsf{M}_1 \tau_{1-\varepsilon}}.$$

При этом D_A совпадает со множеством функций, для которых Af непрерывно.

Далее, если границы интервала $(0, 1)$ являются захватывающими, то естественно считать, что процесс обрывается в момент первого выхода на границу. Поэтому инфинитезимальный оператор такого процесса будет иметь вид

$$Af(x) = -\frac{df'(x)}{dn'(x)},$$

причем $f \in D_A$, если $f(0) = f(1) = 0$, $f(x)$ абсолютно непрерывна и df'/dn' непрерывна на $[0, 1]$.

Допустим, что границы недостижимы. Тогда процесс всегда остается внутри интервала регулярности. Можно показать, что существует строго возрастающая непрерывная гармоническая функция $M(x)$ $(x \in (\alpha, \beta))$ такая, что любая другая гармоническая функция $g(x)$ определяется равенством $g(x) = c_1 M(x) + c_2$, где c_1 и c_2 — постоянные (функция $f(x)$ называется *гармонической* для процесса $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$, если процесс $(f(\xi(t)), \mathfrak{N}_t, \mathsf{P}_x)$ является мартингалом, т. е. $T_t f(x) = f(x)$). Преобразование фазового пространства $y = M(x)$ превращает исходный процесс в новый процесс, определенный на (конечном или бесконечном) интервале, для которого $M(x) = x$. Будем считать, что такое преобразование уже выполнено.

Далее, существует такая выпуклая вверх функция $N(x)$, что при $\alpha < x - \varepsilon < x < x + \delta < \beta$

$$\mathsf{M}_x \tau = N(x) - \frac{\delta}{\varepsilon + \delta} N(x - \varepsilon) - \frac{\varepsilon}{\varepsilon + \delta} N(x + \delta),$$

где τ — момент первого выхода из интервала $(x - \varepsilon, x + \delta)$. Теперь характеристический оператор процесса на интервале (α, β) с недостижимыми границами α и β (для которого $M(x) = x$) может быть записан в виде

$$\mathfrak{A} f(x) = -\frac{df'(x)}{dN'(x)}$$

для всех точек $x \in (\alpha, \beta)$. Поскольку интервал (α, β) является локально компактным пространством, то инфинитезимальный оператор процесса может быть определен в результате применения теоремы 3 § 15.3.

Недостижимая граница α называется *притягивающей*, если для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что при всех $x \in (\alpha, \alpha + \delta)$ $\mathsf{P}_x \left\{ \lim_{t \to \infty} \xi(t) = \alpha \right\} > 1 - \varepsilon$. Недостижимая граница α называется *отталкивающей*, если для всех $\alpha < x_1 < x$ $\mathsf{P}_{x_1} \{\tau_x < \infty\} = 1$.

Теорема 4. *Граница α недостижима, если $N(+\alpha) = -\infty$. При этом, если $M(+\alpha) > -\infty$, то α — притягивающая граница; если же $M(+\alpha) = -\infty$, то α — отталкивающая граница.*

15.8.4. Нерегулярные точки. Рассмотрим поведение процесса в нерегулярных точках. Если x — нерегулярная точка, то должно выполняться по крайней мере одно из условий:

1) при всех $y < x$ $\mathsf{P}_x \{\tau_y < \infty\} = 0$ (точка x непроходима налево);

2) при всех $y > x$ $\mathsf{P}_x \{\tau_y < \infty\} = 0$ (точка x непроходима направо);

3) при всех $y < x$ $\mathsf{P}_y \{\tau_x < \infty\} = 0$ (точка x недостижима слева);

4) при всех $y > x$ $\mathsf{P}_y \{\tau_x < \infty\} = 0$ (точка x недостижима справа).

Если точка x непроходима налево и направо, то она является *поглощающей* и для нее $\mathfrak{A}f(x) = Af(x) = 0$.

Пусть 1) выполнено, а 2) не выполнено (т. е. точка x непроходима налево, но проходима направо). Тогда, если τ^ε — момент первого выхода из интервала $(x-\varepsilon, x+\varepsilon)$, то $\tau^\varepsilon = \tau_{x+\varepsilon}$ п. н. P_x. Существует монотонная функция $g(y)$ такая, что $\mathsf{M}_x\tau^\varepsilon = g(x) - g(x+\varepsilon)$. Поэтому в этом случае

$$\mathfrak{A}f(x) = \lim_{\varepsilon\downarrow 0} \frac{f(x+\varepsilon)-f(x)}{g(x)-g(x+\varepsilon)}.$$

Аналогично, если 2) выполнено, а 1) не выполнено, то

$$\mathfrak{A}f(x) = \lim_{\varepsilon\downarrow 0} \frac{f(x-\varepsilon)-f(x)}{g(x)-g(x-\varepsilon)},$$

т. е. в обоих этих случаях вычисление характеристического оператора сводится к вычислению односторонней производной от функции $f(x)$ по некоторой монотонной функции.

Предположим, что условия 1) и 2) не выполнены. Если выполнено 3), а 4) не выполнено, то x является левой границей интервала регулярности, и поведение процесса в этом случае рассматривалось выше. Если, наоборот, выполнено 4), но не выполнено 3), то x — правая граница интервала регулярности.

Пусть теперь выполнены условия 3) и 4), но не выполнены 1) и 2). Тогда, если точка x незадерживающая, то процесс, выйдя из x, все время будет либо левее точки x, либо правее ее. Положим

$$p = \mathsf{P}_x\{\xi(t) > x \text{ при всех } t > 0\},$$
$$q = \mathsf{P}_x\{\xi(t) < x \text{ при всех } t > 0\}.$$

Для некоторого $\rho > 0$ определим функции

$$g_1(y) = \frac{1}{p}\mathsf{M}_y\tau^\rho\chi_{(y,\infty)}\left(\xi\left(\tau^\rho\right)\right), \quad g_2(y) = \frac{1}{q}\mathsf{M}_y\tau^\rho\chi_{(-\infty,y)}\left(\xi\left(\tau^\rho\right)\right).$$

Функция $g_1(y)$ убывает на промежутке $[x, x+\rho]$, а $g_2(y)$ возрастает на промежутке $[x-\rho, x]$. Характеристический оператор в точке x имеет вид

$$\mathfrak{A}f(x) = \lim_{\varepsilon\downarrow 0,\ \delta\downarrow 0} \frac{p[f(x+\delta)-f(x)]+q[f(x-\varepsilon)-f(x)]}{p[g_1(x)-g_1(x+\delta)]+q[g_2(x)-g_2(x-\varepsilon)]}.$$

15.9. Предельное поведение вероятностей перехода эргодических марковских процессов

15.9.1. Эргодичность. Однородный марковский процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ в фазовом пространстве $(X, \mathfrak{B})$ называется *эргодическим*, если он измерим (это означает, что отображение $\xi(t, \omega)$ пространства $\left([0,\infty)\times\Omega, \mathcal{F}^0_\infty\times\mathfrak{N}\right)$ в пространство $(X, \mathfrak{B})$ измеримо) и для всех $x \in X$ предел

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(\xi(s))\,ds \tag{9.1}$$

существует почти наверное относительно P_x, неслучаен и не зависит от x, какова бы ни была функция $f \in B(X)$.

Если процесс эргодичен, то для него существует единственная стационарная мера π, т. е. такая вероятностная мера на $(X, \mathfrak{B})$, что при всех $t \geqslant 0$, $\Gamma \in \mathfrak{B}$

$$\pi(\Gamma) = \int\limits_X P(t, x, \Gamma)\, \pi(dx).$$

При этом предел в (9.1) может быть записан в виде

$$\lim_{t \to \infty} \frac{1}{t} \int\limits_0^t f(\xi(s))\, ds = \int\limits_X f(x)\, \pi(dx).$$

Другими словами, среднее по времени вдоль траекторий процесса от ограниченной $\mathfrak{B}$-измеримой функции совпадает с ее пространственным средним по стационарной мере.

Теорема 1. *Следующие утверждения эквивалентны.*

1) *Процесс эргодичен.*

2) *При всех $x \in X$, $\Gamma \in \mathfrak{B}$ существует не зависящий от x предел*

$$\lim_{t \to \infty} \frac{1}{t} \int\limits_0^t P(s, x, \Gamma)\, ds,$$

представляющий собой как функция от Γ единственную стационарную меру π на $(X, \mathfrak{B})$.

3) *Процесс имеет стационарную меру и не имеет ограниченных гармонических функций, отличных от постоянных ($\mathfrak{B}$-измеримая функция $h(x)$ $(x \in X)$ называется гармонической, если $T_t f(x) = f(x)$ при всех $t \geqslant 0$, $x \in X$).*

15.9.2. Предельное поведение вероятностей перехода. Пусть $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — эргодический марковский процесс в фазовом пространстве $(X, \mathfrak{B})$ со стационарной мерой π. В этом пункте σ-алгебра $\mathfrak{B}$ считается счетно порожденной. Назовем процесс *периодическим*, если существуют вещественное число $\lambda \neq 0$ и комплексная $\mathfrak{B}$-измеримая функция φ такие, что $|\varphi(x)| = 1$ почти всюду относительно меры π и $\varphi(\xi(t)) = \varphi(\xi(0))\, e^{-i\lambda t}$ P_π-почти наверное для почти всех $t \geqslant 0$. Если процесс периодичен, то среди всех чисел λ, удовлетворяющих указанным условиям, существует наименьшее положительное число (обозначим его δ), и соответствующая функция φ допускает представление $\varphi(x) = e^{i\delta r(x)}$, где $r(x)$ — $\mathfrak{B}$-измеримая функция, однозначно определяемая из неравенств $0 \leqslant r(x) \leqslant 2\pi/\delta$ (здесь π — отношение длины окружности к ее диаметру). Число $\Delta = 2\pi/\delta$ называется *периодом*, а функция $r(x)$ — *функцией сдвига*.

В периодическом случае, разумеется, вероятность перехода $P(t, x, \Gamma)$ не имеет предела при $t \to \infty$, однако существуют пределы по некоторым подпоследовательностям, Теорема 2. *Если процесс $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ периодичен с периодом Δ, то при всех $s \geqslant 0$, $x \in X$, $\Gamma \in \mathfrak{B}$*

$$\lim_{n\to\infty} P(s+n\Delta,\ x,\ \Gamma) = \int_X P(r(y)-r(x)+s+\Delta,\ y,\ \Gamma)\,\pi(dy).$$

При этом сходимость равномерна относительно $\Gamma \in \mathfrak{B}$.

Это эквивалентно тому, что меры под знаком предела в левой части последнего соотношения сходятся к предельной мере по вариации.

Поведение вероятностей перехода в непериодическом случае характеризует следующая теорема (напоминаем, что процесс считается эргодическим).

Теорема 2. *Если процесс непериодичен, то*

$$\lim_{t\to\infty} \int_X P(t,\ x,\ dy)\,f(y) = \int_X f(y)\,\pi(dy) \tag{9.2}$$

для всякой функции $f \in B(X)$ такой, что процесс $f(\xi(t))$ $(t \geqslant 0)$ стохастически непрерывен по отношению к π-почти всем вероятностям P_x.

Среди непериодических процессов выделим так называемые *непериодические регулярные процессы.* Это те, для которых при некотором $t \geqslant 0$

$$\int_X \pi(dx) \int_X q_t(x,\ y)\,\pi(dy) > 0, \tag{9.3}$$

где $q_t(x, y)$ — $\mathfrak{B}\times\mathfrak{B}$-измеримый вариант плотности абсолютно непрерывной компоненты меры $P(t, x, dy)$ относительно меры $\pi(dy)$.

Теорема 3. *Для того чтобы при всех $x \in X$, $\Gamma \in \mathfrak{B}$ выполнялось соотношение*

$$\lim_{t\to\infty} P(t,\ x,\ \Gamma) = \pi(\Gamma), \tag{9.4}$$

необходимо и достаточно, чтобы процесс был непериодическим регулярным.

В непериодическом регулярном случае сходимость в (9.4) равномерна относительно $\Gamma \in \mathfrak{B}$, и, стало быть, вероятность перехода непериодического регулярного процесса сходится при $t \to \infty$ к стационарной мере по вариации.

Наконец, непериодический процесс назовем *непериодическим сингулярным*, если при всех $t \geqslant 0$ меры $\pi(dx)P(t, x, dy)$, заданные на $\mathfrak{B}\times\mathfrak{B}$, сингулярны с определенной там же мерой $\pi(dx)\pi(dy)$. Как это следует из теорем 2 и 3, для непериодического сингулярного процесса существуют множества $\Gamma \in \mathfrak{B}$ такие, что для них соотношение (9.4) не выполняется, однако эти процессы удовлетворяют соотношению (9.2).

Пример. Пусть $\{\tau_n,\ n \geqslant 1\}$ — последовательность неотрицательных независимых одинаково распределенных величин с функцией распределения $F(t)$. Предположим, что среднее $a = \int_0^\infty tF(dt)$ положительно и конечно. Положим

$$\xi(t) = \sum_{k=1}^{n} \tau_k - t,$$

если $\sum_{k=1}^{n-1} \tau_k \leqslant t < \sum_{k=1}^{n} \tau_k$ $(n = 1, 2, \ldots)$ (считаем $\sum_{k=1}^{0} \tau_k = 0$).

Пусть $\mathfrak{N}_t$ обозначает минимальную σ-алгебру событий, относительно которой измеримы величины $\xi(s)$ при $s \in [0, t]$, а через P_x, $x \geqslant 0$, обозначим регулярную условную вероятность на σ-алгебре $\mathfrak{N}_\infty$ при условии $\tau_1 = x$. Тогда $(\xi(t), \mathfrak{N}_t, \mathsf{P}_x)$ — эргодический марковский процесс в фазовом пространстве $(X, \mathfrak{B})$, где $X = [0, \infty)$, $\mathfrak{B}$ — σ-алгебра борелевских подмножеств X (точнее, возможные значения построенного процесса составляют интервал $[0, b]$, где b — точная верхняя грань тех $t > 0$, для которых $F(t + \varepsilon) - F(t - \varepsilon) > 0$ при любом $\varepsilon > 0$). Стационарная мера π этого процесса абсолютно непрерывна с плотностью $a^{-1}[1 - F(t)]$ относительно лебеговой меры.

Периодичность построенного процесса означает, что распределение F решетчато (см. п. 1.4.2); при этом период Δ совпадает с максимальным шагом распределения F; функция сдвига $r(x)$ определяется соотношением $r(x) = x - n\Delta$ при $x \in [n\Delta, n\Delta + \Delta)$ $(n = 0, 1, 2, \ldots)$.

Рассматриваемый процесс будет непериодическим регулярным в том и только том случае, когда n-кратная свертка функции F самой с собой при некотором $n \geqslant 1$ несингулярна с мерой Лебега на $[0, \infty)$.

Если же при всех $n \geqslant 1$ такие свертки сингулярны с лебеговой мерой и при этом распределение F нерешетчато, то процесс будет непериодическим сингулярным.

Литература: [18, 19, 27, 28, 29, 36, 53, 90, 107, 114].

Глава 16. ПРОЦЕССЫ С НЕЗАВИСИМЫМИ ПРИРАЩЕНИЯМИ

16.1. Определение и общие свойства

16.1.1. Определение. Примеры. Будем рассматривать процессы $\xi(t)$ со значениями в $\mathbf{R}^m$, определенные на некотором конечном или бесконечном интервале T. Процесс называется *процессом с независимыми приращениями*, если для всех $t_0 < t_1 < \ldots < t_n$ из T случайные величины

$$\xi(t_0),\ \xi(t_1) - \xi(t_0),\ \ldots,\ \xi(t_n) - \xi(t_{n-1})$$

независимы. Конечномерные распределения процесса $\xi(t)$ полностью определяются распределениями величин $\xi(t)$ $(t \in T)$ и $\xi(t_2) - \xi(t_1)$ $(t_1 < t_2,\ t_1,\ t_2 \in T)$.

Если $F(t, A) = \mathsf{P}\{\xi(t) \in A\}$, $G(t_1, t_2, A) = \mathsf{P}\{\xi(t_2) - \xi(t_1) \in A\}$ $(t_1 < t_2)$, то при $t_1 < t_2 < \ldots < t_n$

$$F_{t_1, \ldots, t_n}(A_1, \ldots, A_n) = \mathsf{P}\{\xi(t_1) \in A_1, \ldots, \xi(t_n) \in A_n\} =$$

$$= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \chi_{A_1}(x_1)\chi_{A_2}(x_1 + x_2)\ldots\chi_{A_n}(x_1 + \ldots + x_n) F_{t_1}(dx_1) \times$$
$$\times G(t_1, t_2, dx_2)\ldots G(t_{n-1}, t_n, dx_n).$$

Обозначим через $\varphi_t(z) = \mathsf{M}\exp\{i(z, \xi(t))\}$, $\psi_{t_1, t_2}(z) = \mathsf{M}\exp\{i(z, \xi(t_2) - \xi(t_1))\}$ $(z \in \mathbf{R}^m)$ характеристические функции значений процесса и его приращений. Тогда совместная характеристическая функция величин $\xi(t_1), \ldots, \xi(t_n)$ определяется формулой

$$\mathsf{M}\exp\left\{i\sum_{k=1}^{n}(\xi(t_k), z_k)\right\} =$$
$$= \varphi_{t_1}(z_1 + \ldots + z_n)\psi_{t_1, t_2}(z_2 + \ldots + z_n)\ldots\psi_{t_{n-1}, t_n}(z_n).$$

Функции $\varphi_t(z)$ и $\psi_{t_1, t_2}(z)$ связаны следующими уравнениями:

1) при $t_1 < t_2$

$$\varphi_{t_1}(z)\psi_{t_1, t_2}(z) = \varphi_{t_2}(z);$$

2) при $t_1 < t_2 < t_3$

$$\psi_{t_1, t_2}(z)\psi_{t_2, t_3}(z) = \psi_{t_1, t_3}(z).$$

Простейшим примером процесса с независимыми приращениями является произвольная неслучайная функция.

Укажем другой важный пример. Пусть $\{\xi_k^+\}$ и $\{\xi_k^-\}$ — такие последовательности независимых в совокупности случайных векторов из $\mathbf{R}^m$, что ряды $\sum \xi_{n_k}^{\pm}$ сходятся для любой последовательности различных натуральных чисел n_k (см. § 3.2), а t_i — произвольная последовательность вещественных чисел. Положим

$$\xi(t) = \sum_{t_i < t} \xi_i^+ + \sum_{t_i \leqslant t} \xi_i^-. \tag{1.1}$$

Этот процесс обладает следующими свойствами:

1) он имеет независимые приращения;

2) при $t \notin \{t_1, t_2, \ldots\}$ он стохастически непрерывен;

3) существуют для всех t_i пределы по вероятности $\xi(t_i - 0)$ и $\xi(t_i + 0)$ и

$$\mathsf{P}\{\xi(t_i) - \xi(t_i - 0) = \xi_i^-\} = 1, \quad \mathsf{P}\{\xi(t_i + 0) - \xi(t_i) = \xi_i^+\} = 1; \tag{1.2}$$

4) если $\xi(t)$ — сепарабельный процесс (поскольку соотношение (1.1) определяет процесс с точностью до стохастической эквивалентности, это всегда можно предположить), то с вероятностью 1 выбо-

рочные функции $\xi(t)$ непрерывны во всех точках, за исключением точек t_i; в точках t_i существуют пределы $\xi(t_i-0)$ и $\xi(t_i+0)$, и для этих пределов также выполнено (1.2).

Процессы вида (1.1) называются *дискретными процессами с независимыми приращениями.*

16.1.2. Разложение Леви. Для всякого процесса с независимыми приращениями $\xi(t)$ существует такая неслучайная функция $a(t)$, определенная на T и принимающая значения из $\mathbf{R}^m$, что процесс $\xi_1(t)=\xi(t)-a(t)$ обладает следующими свойствами: 1) с вероятностью 1 не имеет разрывов второго рода внутри T; 2) в каждой точке имеет пределы по вероятности справа и слева; 3) имеет не более чем счетное число точек стохастического разрыва, т. е. точек, в которых процесс не является стохастически непрерывным.

Функция $a(t)$, обладающая указанными свойствами, называется *центрирующей функцией* для процесса с независимыми приращениями. Она определяется неоднозначно: любые две центрирующие функции отличаются на функцию, не имеющую разрывов второго рода внутри T.

Если $\xi(t)$ — симметричный процесс, то центрирующую функцию можно взять тождественно равной нулю.

Пусть $a(t)$ — центрирующая функция для процесса $\xi(t)$ и $\xi_1(t)=\xi(t)-a(t)$. Обозначим через $\{t_1, t_2, \ldots\}$ множество точек стохастического разрыва $\xi_1(t)$, $\xi_k^+=\xi_1(t_k+0)-\xi_1(t_k)$, $\xi_k^-=\xi_1(t_k)-\xi_1(t_k-0)$. Тогда величины $\{\xi_k^+;\ \xi_k^-;\ k=1, 2, \ldots\}$ независимы в совокупности и $a(t)$ можно выбрать таким образом, что процесс

$$\xi_2(t)=\sum_{t_k<t}\xi_k^+ + \sum_{t_k\leqslant t}\xi_k^-$$

является дискретным процессом с независимыми приращениями. При этом процесс $\xi_3(t)=\xi_1(t)-\xi_2(t)$ не зависит от процесса $\xi_2(t)$; кроме того, $\xi_3(t)$ стохастически непрерывен. Таким образом, для всякого процесса с независимыми приращениями $\xi(t)$ можно указать неслучайную функцию $a(t)$, дискретный процесс с независимыми приращениями $\xi_д(t)$ и стохастически непрерывный процесс с независимыми приращениями $\xi_н(t)$ такие, что

$$\xi(t)=a(t)+\xi_д(t)+\xi_н(t), \tag{1.3}$$

при этом процессы $\xi_д(t)$ и $\xi_н(t)$ независимы. Представление (1.3) носит название *разложения Леви* для процесса с независимыми приращениями. Если центрирующая функция $a(t)$ выбрана, то остальные компоненты разложения определяются однозначно.

16.1.3. Некоторые неравенства. Пусть $\xi(t)$ — процесс с независимыми приращениями, для которого $\mathsf{M}(\xi(t), z)^2<\infty$ для всех $t\in T$, и $a(t)=\mathsf{M}\xi(t)$. Функция $a(t)$ является центрирующей. Обозначим через $B(t)$ симметричный оператор в $\mathbf{R}^m$, для которого

$$(B(t)z, z)=\mathsf{M}(\xi(t)-a(t), z)^2.$$

Тогда $B(t)$ — неотрицательный оператор и как функция t не убывает. Поэтому $B(t)$ ограничен на всяком замкнутом справа промежутке, лежащем в T, Теорема 1 (обобщение неравенства Колмогорова на процессы с независимыми приращениями). *Если $\xi(t)$ — сепарабельный процесс с независимыми приращениями и $[a, b] \in T$, то*

$$\mathsf{P}\left\{\sup_{t \in [a, b]} |\xi(t) - a(t)| > c\right\} \leqslant \frac{\operatorname{Sp} B(b)}{c^2}; \tag{1.4}$$

здесь $\operatorname{Sp} B$ *— след оператора* B: $\operatorname{Sp} B = \sum_{k=1}^{m} (Be_k, e_k)$, *где* $\{e_k, k = 1, \ldots, m\}$ *— базис в* $\mathbf{R}^m$.

Неравенство (1.4) можно распространить на T:

$$\mathsf{P}\left\{\sup_{t \in T} |\xi(t) - \xi(a)| > c\right\} \leqslant \frac{1}{c^2} \sup_{t \in T} \operatorname{Sp} B(t).$$

На сепарабельные процессы с независимыми приращениями переносятся и другие неравенства, известные для сумм независимых случайных величин.

Теорема 2. *Если $\xi(t)$ — симметричный сепарабельный процесс, то*

$$\mathsf{P}\left\{\sup_{t \in [a, b]} |\xi(t)| > c\right\} \leqslant 2\mathsf{P}\{|\xi(b)| > c\}.$$

Теорема 3. *Если при некотором $\alpha < 1$ для $t \in [a, b]$*

$$\mathsf{P}\{|\xi(b) - \xi(t)| > c\} \leqslant \alpha,$$

то для всех $x > 0$

$$\mathsf{P}\left\{\sup_{t \in [a, b]} |\xi(t)| > c + x\right\} \leqslant \frac{1}{1-\alpha} \mathsf{P}\{|\xi(b)| > x\}.$$

Теорема 4. *Если $\xi(t)$ не имеет разрывов второго рода и $\mathsf{P}\{\sup|\xi(t+0) - \xi(t-0)| \leqslant c\} = 1$, то для всех l и a выполнено неравенство*

$$\mathsf{M}e^{z\xi(t)} \leqslant \frac{e^{zl}}{\mathsf{P}\{\xi(t) \leqslant l\}} \left(1 - \frac{(a+c)z}{1 - 4\mathsf{P}\{|\xi(t)| > a\}}\right)^{-1},$$

если только $\mathsf{P}\{|\xi(t)| > a\} < 1/4$ *и* $|z| < \dfrac{1 - 4\mathsf{P}\{|\xi(t)| > a\}}{a + c}$.

16.2. Стохастически непрерывные процессы с независимыми приращениями

16.2.1. Свойства выборочных функций. 1) Сепарабельный стохастически непрерывный процесс с вероятностью 1 не имеет разрывов второго рода.

2) Для того чтобы сепарабельный процесс с независимыми приращениями $\xi(t)$, определенный на $[a, b]$, был с вероятностью 1 непрерывным, необходимо и достаточно, чтобы для всякого $\varepsilon > 0$ выполнялось условие

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \mathsf{P}\{|\xi(t_{nk}) - \xi(t_{nk-1})| > \varepsilon\} = 0,$$

где $a = t_{n0} < \ldots < t_{nn} = b$ и $\max_k (t_{nk} - t_{nk-1}) \to 0$ при $n \to \infty$.

3) Если $\xi(t)$ — сепарабельный процесс с независимыми приращениями на $[a, b]$, для которого $\mathsf{P}\{\xi(b)=\xi(a)\}>0$, то $\xi(t)$ — с вероятностью 1 ступенчатая функция, т. е. отрезок $[a, b]$ может быть разложен на конечное (случайное) число случайных интервалов, на каждом из которых $\xi(t)$ постоянно. Наоборот, если $\xi(t)$ — с вероятностью 1 ступенчатая функция на $[a, b]$, то $\mathsf{P}\{\xi(b)=\xi(a)\}>0$.

4) Для того чтобы числовой сепарабельный процесс с независимыми приращениями $\xi(t)$ был с вероятностью 1 не убывающим на $[a, b]$, необходимо и достаточно, чтобы $\mathsf{P}\{\xi(b)>\xi(a)\}=1$.

16.2.2. Формула Леви — Хинчина. Пусть $\xi(t)$ — стохастически непрерывный процесс с независимыми приращениями, определенный на $[a, b]$, со значениями в $\mathbf{R}^m$. Тогда для него существуют:

1) непрерывная функция $a(t)$ $(t\in[a, b])$ со значениями в $\mathbf{R}^m$;

2) непрерывная функция $B(t)$ $(t\in[a, b])$, значениями которой служат симметричные неотрицательные операторы в $\mathbf{R}^m$;

3) функция $\Pi(t, A)$ $(t\in[a, b])$, определенная для всех борелевских множеств A из $\mathbf{R}^m$ (лежащих на положительном расстоянии от точки 0) и обладающая свойствами:

а) $\Pi(t, A)$ — непрерывная неубывающая функция t;

б) при фиксированных $t\in[a, b]$ она счетно аддитивна по A;

в) интеграл $\displaystyle\int\frac{|x|^2}{1+|x|^2}\Pi(t, dx)$ по $\mathbf{R}^m$ с выкинутой точкой 0 конечен.

Характеристическая функция приращения процесса является безгранично делимой (см. п. 5.2.2) и выражается формулой: при $a\leqslant t<s\leqslant b$

$$\mathsf{M}\exp\{i(z, \xi(s)-\xi(t))\}=$$
$$=\exp\Big\{i(z, a(s)-a(t))-\frac{1}{2}([B(s)-B(t)]z, z)+$$
$$+\int\left(e^{i(z, x)}-1-\frac{i(z, x)}{1+|x|^2}\right)(\Pi(s, dx)-\Pi(t, dx))\Big\},\quad (2.1)$$

которая и носит название *формулы Леви — Хинчина*. Функции $a(t)$, $B(t)$ и $\Pi(t, A)$, входящие в формулу Леви — Хинчина, определяются однозначно.

16.2.3. Строение стохастически непрерывного процесса с независимыми приращениями. Пусть процесс $\xi(t)$ сепарабелен. Тогда он не имеет разрывов второго рода, и, следовательно, число скачков процесса, превосходящих по модулю ε, конечно на каждом замкнутом конечном интервале t. Обозначим через $\nu(t, A)$, где A — некоторое борелевское множество в $\mathbf{R}^m$, лежащее на положительном расстоянии от точки 0, число скачков процесса $\xi(s)$ (*скачком* в точке s называется величина $\xi(s+0)-\xi(s-0)$), происшедших до момента t и попавших во множество A; $\nu(t, A)$ как функция t является процессом Пуассона, т. е. $\nu(t, A)$ — стохастически непрерывный процесс с независимыми приращениями, имеющий при каждом t распределение Пуассона. При t фиксированном $\nu(t, A)$ является пуассоновской мерой с независимыми значениями. Это значит, что выполнены следующие условия:

1) $\nu\left(t, \bigcup_k A_k\right) = \sum_k \nu(t, A_k)$, если A_k попарно непересекающиеся множества и $\bigcup_k A_k$ лежит на положительном расстоянии от точки 0;

2) если $A_1, A_2, \ldots, A_k$ — попарно не пересекающиеся борелевские множества, то процессы $\nu(t, A_1), \ldots, \nu(t, A_k)$ независимы в совокупности.

Функция $\Pi(t, A) = \mathsf{M}\nu(t, A)$ является той функцией, которая входит в формулу Леви — Хинчина. Определим интегралы

$$\int_{|x|>1} x\nu(t, dx), \qquad \int_{0<|x|\leqslant 1} x[\nu(t, dx) - \Pi(t, dx)] = \lim_{\varepsilon\to 0} \int_{\varepsilon<|x|\leqslant 1} x \times$$

$\times [\nu(t, dx) - \Pi(t, dx)]$ (предел в смысле сходимости в среднеквадратическом). Тогда процесс

$$\xi_0(t) = \xi(t) - \int_{|x|>1} x\nu(t, dx) - \int_{0<|x|\leqslant 1} x[\nu(t, dx) - \Pi(t, dx)]$$

с вероятностью 1 непрерывен и не зависит от $\nu(t, A)$.

Пусть $\xi_0(t)$ — с вероятностью 1 непрерывный процесс с независимыми приращениями. Тогда $\xi_0(t)$ имеет гауссовские приращения, т. е. $\xi_0(t_2) - \xi_0(t_1)$ имеет нормальное распределение. Характеристическая функция приращения процесса $\xi_0(t)$ имеет вид

$$\mathsf{M}\exp\{i(z, \xi_0(t_2) - \xi_0(t_1))\} =$$
$$= \exp\left\{i(z, a(t_2) - a(t_1)) - \frac{1}{2}((B(t_2) - B(t_1))z, z)\right\}, \quad (2.2)$$

где $a(t) = \mathsf{M}(\xi_0(t_1) - \xi_0(t_0))$, $(B(t)z, z) = \mathsf{M}(\xi_0(t) - \xi_0(t_0) - a(t), z)^2$, а t_0 — минимальная точка в T.

Таким образом, для стохастически непрерывного процесса с независимыми приращениями $\xi(t)$ справедливо следующее представление:

$$\xi(t) = \xi_0(t) + \int_{|x|\leqslant 1} x[\nu(t, dx) - \Pi(t, dx)] + \int_{|x|>1} x\nu(t, dx), \quad (2.3)$$

где $\nu(t, A)$ — пуассоновская мера с независимыми значениями по A и пуассоновский процесс по t, $\Pi(t, A) = \mathsf{M}\nu(t, A)$, а $\xi_0(t)$ — непрерывный процесс с независимыми гауссовскими приращениями, характеристическая функция которых определяется формулой (2.2).

Укажем на некоторые связи между свойствами функций, входящих в правую часть формулы Леви — Хинчина, и свойствами выборочных функций процесса.

Теорема. 1) *Процесс $\xi(t)$ непрерывен тогда и только тогда, когда $\Pi(t, A) \equiv 0$ для всех $t \in T$ и борелевских множеств $A \subset \mathbf{R}^m$.*

2) *Пусть значение функции $\Pi(t, \mathbf{R}^m) = \lim_{\varepsilon\to 0} \Pi(t, \mathbf{R}^m \setminus S_\varepsilon)$, где S_ε — сфера в $\mathbf{R}^m$ радиуса ε с центром в точке 0. Этот предел может быть и бесконечным. Процесс $\xi(t)$ будет ступенчатым тогда и только тогда, когда:* а) *$a(t)$ постоянно;* б) *$B(t) = 0$ для всех $t \in T$;* в) *$\Pi(t, \mathbf{R}^m) < \infty$ для всех $t \in T$.*

3) *Процесс* $\xi(t)$ *с вероятностью* 1 *имеет ограниченную вариацию на отрезке* $[t_0, t_1]$ *тогда и только тогда, когда*: а) $a(t)$ *имеет ограниченную вариацию на* $[t_0, t_1]$; б) $B(t_1)-B(t_0)=0$; в) *интеграл* $\int\limits_{|x|\leqslant 1} |x| (\Pi(t_1, dx) - \Pi(t_0, dx)) < \infty$.

Если $\zeta(t)$ — *вариация* $\xi(t)$ *на отрезке* $[t_0, t]$, *то* $\zeta(t)$ *является также стохастически непрерывным процессом с независимыми приращениями, характеристическая функция* $\zeta(t)$ *имеет вид*

$$\mathsf{M}e^{i\lambda\zeta(t)} = \exp\left\{i\lambda\gamma(t) + \int \left(e^{i\lambda|x|} - 1\right)(\Pi(t, dx) - \Pi(t_0, dx))\right\},$$

где $\gamma(t) = \underset{[t_0, t]}{\operatorname{var}}\, a(s) + \int\limits_{|x|\leqslant 1} |x| (\Pi(t, dx) - \Pi(t_0, dx))$ (*первое слагаемое справа — вариация* $a(\cdot)$ *на отрезке* $[t_0, t]$).

4) *Пусть* K — *некоторый конус в* $\mathbb{R}^m$ *с центром в точке* 0. *Для того чтобы* $\xi(t) \in K$ $(t \in T)$ *с вероятностью* 1, *необходимо и достаточно, чтобы выполнялись условия*: а) $a(t) \in K$ *при* $t \in T$; б) $B(t) = 0$ *при* $t \in T$; в) *если* $A \cap K$ *пусто, то* $\Pi(t, A) = 0$ *для всех* $t \in T$.

5) *Пусть* $\xi(t)$ — *процесс со значениями в* $\mathbb{R}$. *Для того чтобы* $\xi(t)$ *был с вероятностью* 1 *неубывающей функцией, необходимо и достаточно, чтобы*: а) $\int \frac{|x|}{1+x^2} \Pi(t, dx) < \infty$ *для всех* t *и функция* $a(t) - \int \frac{x}{1+x^2} \Pi(t, dx)$ *являлась неубывающей*; б) $B(t) = 0$ *для всех* t; в) $\Pi(t, (-\infty, 0)) = 0$ *для всех* t.

16.3. Однородные процессы. Асимптотические свойства

16.3.1. Характеристическая функция однородного процесса. Процесс с независимыми приращениями $\xi(t)$ называется *однородным*, если он определен на $[0, \infty)$, $\xi(0) = 0$ и распределение $\xi(t+h) - \xi(t)$ совпадает с распределением $\xi(h)$ для всех $t > 0$ и $h > 0$. Всякий однородный процесс с независимыми приращениями $\xi(t)$, $t \in [0, \infty)$, представим в виде суммы $\xi(t) = a(t) + \xi_1(t)$, где $\xi_1(t)$ — стохастически непрерывный процесс с независимыми приращениями; $a(t)$ — неслучайная функция, удовлетворяющая условию: при всех $h > 0$, $t > 0$ $a(t+h) = a(t) + a(h)$.

Если $\xi(t)$ — однородный стохастически непрерывный процесс с независимыми приращениями в $\mathbb{R}^m$, то его характеристическая функция имеет вид

$$\mathsf{M}e^{i(z, \xi(t))} = \exp\left\{t\left[i(z, a) - \frac{1}{2}(Bz, z) + \int\limits_{|x|\leqslant 1} \left(e^{i(z, x)} - 1 - (z, x)\right)\Pi(dx) + \int\limits_{|x|>1} \left(e^{i(z, x)} - 1\right)\Pi(dx)\right]\right\}, \tag{3.1}$$

где $a \in \mathbf{R}^m$; B — симметричный неотрицательный оператор в $\mathbf{R}^m$; Π — мера в $\mathbf{R}^m$, для которой

$$\int \frac{(x, x)}{1+(x, x)} \Pi(dx) < \infty \text{ и } \Pi(\{0\}) = 0.$$

В силу однородности процесса правая часть (3.1) является также характеристической функцией приращения $\xi(t+h) - \xi(h)$ при любом $h > 0$. Как видно из формулы (3.1), величина

$$K(z) = \frac{1}{t} \ln \mathsf{M} e^{i(z, \xi(t))}$$

не зависит от t. Она называется *кумулянтой* однородного процесса с независимыми приращениями. Кумулянта процесса определяет все его конечномерные распределения.

16.3.2. Локальные свойства одномерных процессов. В этом пункте предполагается, что $\xi(t)$ — однородный процесс в $\mathbf{R}$. Характеристическая функция процесса имеет вид

$$\mathsf{M} e^{i\lambda\xi(t)} = \exp\left\{t\left[ia\lambda - \frac{b\lambda^2}{2} + \int\limits_{|x| \leqslant 1} (e^{i\lambda x} - 1 - i\lambda x)\Pi(dx) + \int\limits_{|x|>1} (e^{i\lambda x} - 1)\Pi(dx)\right]\right\}. \quad (3.2)$$

Будем изучать поведение $\xi(t)$ при $t \downarrow 0$.

1) Если хотя бы одна из величин $\overline{\lim\limits_{t \downarrow 0}} \frac{1}{t} \xi(t)$, $\underline{\lim\limits_{t \downarrow 0}} \frac{1}{t} \xi(t)$ конечна с положительной вероятностью, то $\xi(t)$ имеет с вероятностью 1 ограниченную вариацию на каждом конечном отрезке, и, следовательно (см. п. 16.2.3, 3)), $b = 0$ и $\int\limits_{|x| \leqslant 1} |x| \Pi(dx) < \infty$. В этом случае

$$\mathsf{P}\left\{\lim_{t \downarrow 0} \frac{1}{t} \xi(t) = a - \int_{-1}^{1} x\Pi(dx)\right\} = 1.$$

2) Если выполнено одно из условий:

а) $b > 0$; б) $\int\limits_{-1}^{1} |x| \Pi(dx) = +\infty$, то

$$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \frac{1}{t} \xi(t) = +\infty\right\} = \mathsf{P}\left\{\underline{\lim_{t \downarrow 0}} \frac{1}{t} \xi(t) = -\infty\right\} = 1.$$

3) $$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \left(\sqrt{2t \ln \ln \frac{1}{t}}\right)^{-1} \xi(t) = \sqrt{b}\right\} = \mathsf{P}\left\{\underline{\lim_{t \downarrow 0}} \left(\sqrt{2t \ln \ln \frac{1}{t}}\right)^{-1} \xi(t) = -\sqrt{b}\right\} = 1$$

(*локальный закон повторного логарифма*).

4) Пусть $\xi(t)$ — неубывающий процесс с кумулянтой

$$K(z) = \int_0^\infty (e^{izx} - 1)\, \Pi(dx)$$

и $g(x)$ — неубывающая функция, определенная при $x \geqslant 0$ и удовлетворяющая условиям: 1) $g(0) = 0$; 2) $g(x+y) \leqslant g(x) + g(y)$ (функция g полуаддитивна). Тогда

а) если $\int_0^\infty g(x)\, \Pi(dx) < \infty$, то

$$\mathsf{P}\left\{\lim_{t \downarrow 0} \frac{g(\xi(t))}{t} = 0\right\} = 1;$$

б) если $\int_0^1 g(x)\, \Pi(dx) = +\infty$, то

$$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \frac{g(\xi(t))}{t} = +\infty\right\} = 1.$$

5) Пусть $\varphi(t)$ возрастает на $[0, 1]$ и

$$\lim_{u \downarrow 1} \sup_t \left| \frac{\varphi(ut)}{\varphi(t)} - 1 \right| = 0.$$

Предположим, что $\xi(t)$ — такой однородный процесс с независимыми приращениями, что для всякого $\varepsilon > 0$

$$\sup_{0 \leqslant t \leqslant 1} \mathsf{P}\{\xi(t) < -\varepsilon\varphi(t)\} < 1.$$

Тогда

а) если $\int_0^1 \frac{1}{t} \mathsf{P}\{\xi(t) > \varphi(t)\}\, dt < \infty$, то

$$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \frac{\xi(t)}{\varphi(t)} \leqslant 1\right\} = 1;$$

б) если $\int_0^1 \frac{1}{t} \mathsf{P}\{\xi(t) > \varphi(t)\}\, dt = +\infty$, то

$$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \frac{\xi(t)}{\varphi(t)} \geqslant 1\right\} = 1.$$

6) Пусть $\xi(t)$ — устойчивый однородный процесс с независимыми приращениями, кумулянта которого имеет вид

$$K(z) = -c\,|z|^\alpha \left(1 - \frac{iz}{|z|}\, \omega(z, \alpha)\right),$$

где $\omega(z, \alpha) = \operatorname{tg}\frac{\pi}{2}\alpha$ при $\alpha \in (1, 2)$, $\omega(z, \alpha) = \frac{2}{\pi}\ln|z|$ при $\alpha = 1$. Этот процесс имеет только отрицательные скачки. Положим

$$\varphi(t) = \alpha\left((\alpha-1)^{\alpha-1}\left|\cos\frac{\pi}{2}\alpha\right|^{-1}\right)^{1/\alpha} t^{1/\alpha}\left[\ln\ln\frac{1}{t}\right]^{(\alpha-1)/\alpha}$$

при $\alpha \in (1, 2)$,

$$\varphi(t) = \frac{2ct}{\pi}\ln\frac{1}{t} \text{ при } \alpha = 1.$$

Тогда

$$\mathsf{P}\left\{\lim_{t\downarrow 0}\frac{\xi(t)}{\varphi(t)} = 1\right\} = 1.$$

7) Пусть $\xi(t)$ — устойчивый монотонный процесс с кумулянтой

$$K(z) = -c|z|^{\alpha}\left(1 - \frac{iz}{|z|}\operatorname{tg}\frac{\pi}{2}\alpha\right), \quad 0 < \alpha < 1.$$

Тогда

$$\mathsf{P}\left\{\varliminf_{t\downarrow 0}\xi(t)\left(t^{1/\alpha}\left[\ln\ln\frac{1}{t}\right]^{(1-\alpha)/\alpha}\right)^{-1} > 0\right\} = 1.$$

16.3.3. Поведение одномерных процессов при $t \to \infty$. Здесь используются обозначения предыдущего пункта.

I. Усиленный закон больших чисел. 1) *Если существует* $\mathsf{M}\xi(t) = \gamma t$, *то*

$$\mathsf{P}\left\{\lim_{t\to\infty}\frac{\xi(t)}{t} = \gamma\right\} = 1, \quad \gamma = a + \int_{-\infty}^{-1} x\Pi(dx) + \int_{1}^{\infty} x\Pi(dx).$$

2) *Пусть* $\int_{-\infty}^{-1} x\Pi(dx) > -\infty$, $\int_{1}^{\infty} x\Pi(dx) = +\infty$. *Тогда*

$$\mathsf{P}\left\{\lim_{t\to\infty}\frac{1}{t}\xi(t) = +\infty\right\} = 1.$$

3) *Пусть* $\int_{-\infty}^{-1} x\Pi(dx) = -\infty$, $\int_{1}^{\infty} x\Pi(dx) < +\infty$. *Тогда*

$$\mathsf{P}\left\{\lim_{t\to\infty}\frac{1}{t}\xi(t) = -\infty\right\} = 1.$$

II. Условия ограниченности процесса на $[0, \infty)$. 1) *Если существует* $\mathsf{M}\xi(t) < 0$, *то*

$$\mathsf{P}\left\{\sup_t \xi(t) < \infty\right\} = 1, \quad \mathsf{P}\left\{\inf_t \xi(t) = -\infty\right\} = 1.$$

2) *Если* $\mathsf{M}\xi(t) > 0$, *то*

$$\mathsf{P}\left\{\sup_t \xi(t) = +\infty\right\} = 1, \quad \mathsf{P}\left\{\inf_t \xi(t) > -\infty\right\} = 1.$$

3) *Если* $\mathsf{M}\xi(t) = 0$ *и* $\xi(t) \neq 0$ *тождественно, то*

$$\mathsf{P}\left\{\sup_t \xi(t) = +\infty\right\} = \mathsf{P}\left\{\inf_t \xi(t) = -\infty\right\} = 1.$$

4) *Для того чтобы* $\mathsf{P}\left\{\sup_t \xi(t) < +\infty\right\} = 1$, *необходимо и достаточно, чтобы*

$$\int_0^\infty \frac{1}{t} \mathsf{P}\{\xi(t) > 0\}\, dt < \infty.$$

5) *Для того чтобы* $\mathsf{P}\left\{\inf_t \xi(t) > -\infty\right\} = 1$, *необходимо и достаточно, чтобы*

$$\int_0^\infty \frac{1}{t} \mathsf{P}\{\xi(t) < 0\}\, dt < \infty.$$

III. *Пусть* $\xi(t)$ — *неубывающий процесс с кумулянтой*

$$K(z) = \int_0^\infty \left(e^{izx} - 1\right) \Pi(dx),$$

а $g(x)$ — *неубывающая функция, для которой* $g(x+y) \leqslant g(x) + g(y)$ *при* $x > 0$, $y > 0$.

1) *Если* $\int_1^\infty g(x)\, \Pi(dx) < \infty$, *то*

$$\mathsf{P}\left\{\lim_{t\to\infty} \frac{1}{t} g(\xi(t)) = 0\right\} = 1.$$

2) *Если* $\int_1^\infty g(x)\, \Pi(dx) = +\infty$, *то*

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}} \frac{1}{t} g(\xi(t)) = +\infty\right\} = 1.$$

IV. Закон повторного логарифма. *Пусть* $\mathsf{M}\xi(t) = 0$ *и* $\mathsf{D}\xi(t) < \infty$. *Тогда* $\mathsf{D}\xi(t) = tc$, *где* $c = b + \int x^2 \Pi(dx)$, *и*

$$\mathsf{P}\left\{\overline{\lim_{t\to\infty}} \frac{\xi(t)}{\sqrt{2ct \ln\ln t}} = 1\right\} = \mathsf{P}\left\{\underline{\lim_{t\to\infty}} \frac{\xi(t)}{\sqrt{2ct \ln\ln t}} = -1\right\} = 1.$$

16.4. Функционалы от процессов с независимыми приращениями

16.4.1. Интегро-дифференциальное уравнение процесса. Пусть $\xi(t)$ — процесс с независимыми приращениями на $[t_0, t_1]$ со значениями в $\mathbf{R}^m$, характеристическая функция которого имеет вид

$$\mathsf{M}\exp\{i(z, \xi(t))\} = \exp\left\{\int_{t_0}^{t}\left[i(\hat{a}(s), z) - \frac{1}{2}(\widehat{B}(s)z, z) + \right.\right.$$
$$+ \int_{|x|\leqslant 1}(e^{i(z,x)} - 1 - i(z, x))\,\widehat{\Pi}(s, dx) +$$
$$\left.\left. + \int_{|x|>1}(e^{i(z,x)} - 1)\,\widehat{\Pi}(s, dx)\right]ds\right\}, \tag{4.1}$$

где $\hat{a}(s) \in \mathbf{R}^m$, $\widehat{B}(s)$ — симметричный неотрицательный оператор в $\mathbf{R}^m$, $\widehat{\Pi}(s, A)$ — мера в $\mathbf{R}^m$, для которой

$$\int_{|x|\leqslant 1}|x|^2\,\widehat{\Pi}(s, dx) < \infty \text{ при } s \in [t_0, t_1].$$

Характеристическую функцию процесса можно записать в виде (4.1), если функции $a(t)$, $B(t)$, $\Pi(t, A)$, входящие в формулу (2.1), абсолютно непрерывны по t. Предположим, что $\hat{a}(t)$, $\widehat{B}(t)$, $\widehat{\Pi}(t, A)$ и $\int_{|x|\leqslant 1}|x|^2\,\widehat{\Pi}(t, dx)$ непрерывны по t. Тогда функция

$$\mathsf{M}f(x + \xi(t)) = u(t, x)$$

удовлетворяет следующему интегро-дифференциальному уравнению при $t \in [t_0, t_1]$:

$$\frac{\partial u(t, x)}{\partial t} + L_t u(t, x) = 0 \tag{4.2}$$

и граничному условию $\lim\limits_{t \uparrow t_1} u(t, x) = f(x)$, какова бы ни была дважды непрерывно дифференцируемая функция f с ограниченными производными; здесь

$$L_t u(t, x) = \sum_{i=1}^{m}\hat{a}^i(t)\frac{\partial u(t, x)}{\partial x^i} + \frac{1}{2}\sum_{i,j=1}^{m}\hat{b}^{ij}(t)\frac{\partial^2 u(t, x)}{\partial x^i\,\partial x^j} +$$
$$+ \int_{|y|\leqslant 1}\left[f(x + y) - f(x) - \sum_{j=1}^{m}\frac{\partial f(x)}{\partial x^j}y^j\right]\widehat{\Pi}(t, dx) +$$
$$+ \int_{|y|>1}[f(x + y) - f(x)]\,\widehat{\Pi}(t, dx); \tag{4.3}$$

$\hat{a}^i$, x^i — координаты векторов $\hat{a}$ и x; $\hat{b}^{ij}$ — элементы матрицы оператора $\hat{B}$ в некотором ортонормированном базисе в $\mathbb{R}^m$.

Оператор L_t можно использовать и при вычислении распределений функционалов интегрального вида. Пусть

$$\varphi(t, x) = \int_t^{t_1} g(s, \xi(s) - \xi(t) + x)\, ds, \tag{4.4}$$

где $g(s, x)$ — непрерывная ограниченная дважды непрерывно дифференцируемая по x с ограниченными производными функция, и

$$v_\lambda(t, x) = \mathsf{M}e^{\lambda\varphi(t, x)}.$$

Тогда $v_\lambda(t, x)$ удовлетворяет интегро-дифференциальному уравнению

$$\frac{\partial}{\partial t} v_\lambda(t, x) + L_t v_\lambda(t, x) + \lambda g(t, x)\, v_\lambda(t, x) = 0 \tag{4.5}$$

и граничному условию $\lim\limits_{t \uparrow t_1} v_\lambda(t, x) = 1$.

Если функция $v_\lambda(t, x)$ найдена, то

$$\mathsf{M}\exp\left\{\lambda \int_{t_0}^{t_1} g(s, \xi(s))\, ds\right\} = v_\lambda(t_0, 0).$$

16.4.2. Одномерные однородные процессы с отрицательными скачками. Пусть $\xi(t)$ — одномерный однородный процесс с кумулянтой

$$K(z) = i\gamma z - \frac{bz^2}{2} + \int_{-\infty}^{-1} \left(e^{izx} - 1\right)\Pi(dx) + \int_{-1}^{0} \left(e^{izx} - 1 - izx\right)\Pi(dx), \tag{4.6}$$

т. е. $\xi(t)$ может иметь лишь отрицательные скачки. Если $\gamma + \int_{-\infty}^{-1} x\Pi(dx) \geqslant 0$, то $\mathsf{P}\left\{\sup\limits_t \xi(t) = +\infty\right\} = 1$. Значит, процесс не ограничен сверху и величина

$$\tau_a = \inf\,[t\colon\ \xi(t) > a]$$

с вероятностью 1 конечна. В силу отсутствия положительных скачков $\xi(\tau_a) = a$. Величина τ_a называется *моментом первого достижения уровня* a (*моментом первого прохождения через уровень* a), τ_a — марковский момент для процесса $\xi(t)$ (см. п. 14.2.4). Укажем некоторые свойства τ_a.

1) Момент τ_a как функция a является однородным процессом с независимыми приращениями. Этот процесс с вероятностью 1 не убывает.

2) Обозначим

$$K_+(z) = \gamma z + \frac{bz^2}{2} + \int_{-\infty}^{-1} (e^{zx} - 1)\,\Pi(dx) + \int_{-1}^{0} (e^{zx} - 1 - zx)\,\Pi(dx),$$
$$\psi(\lambda) = \frac{1}{a} \ln \mathsf{M} e^{-\lambda\tau_a}, \quad \lambda > 0. \tag{4.7}$$

Тогда $\psi(\lambda)$ является единственным корнем уравнения

$$K_+(-\psi(\lambda)) = \lambda.$$

Заметим, что при $\operatorname{Re} z > 0$ существует $\mathsf{M} e^{z\xi(t)}$. Эта функция аналитична, и

$$\mathsf{M} e^{z\xi(t)} = e^{tK_+(z)}.$$

Функция $\psi(\lambda)$ аналитична при $\operatorname{Re}\lambda \geqslant 0$, и

$$\mathsf{M} e^{i\lambda\tau_a} = e^{a\psi(-i\lambda)}.$$

3) Укажем на связь между распределениями процесса $\xi(t)$ и величины τ_a:

$$\frac{d}{ds} \int_0^x \mathsf{P}\{\tau_y < s\}\, dy = \frac{1}{s} \int_0^x y\, d_y \mathsf{P}\{\xi(s) < y\}.$$

Предположим, что $\xi(s)$ имеет плотность распределения $f_\xi(s, x) = \frac{d}{dx} \mathsf{P}\{\xi(s) < x\}$. Тогда и величина τ_a имеет плотность $f_\tau(a, s) = \frac{d}{ds} \mathsf{P}\{\tau_a < s\}$ и $f_\tau(a, s) = \frac{a}{s} f_\xi(s, a)$.

4) Зная распределение τ_a, можем найти распределение максимума процесса:

$$\mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T} \xi(t) < a\right\} = \mathsf{P}\{\tau_a > T\}.$$

16.4.3. Распределение максимума и минимума однородного процесса. Пусть $\xi(t)$ — одномерный однородный процесс, $F(t, x)$ — функция распределения $\xi(t)$. Обозначим

$$Q_+(t, x) = \mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} \xi(s) < x\right\},$$

$$q_+(\lambda, x) = \lambda \int_0^\infty e^{-\lambda t} Q_+(t, x)\, dt, \quad \tilde{q}_+(\lambda, z) = \int_{-\infty}^{\infty} e^{izx}\, d_x q_+(\lambda, x).$$

Тогда

$$\tilde{q}_{+}(\lambda, z) = \exp\left\{\int_0^\infty \frac{e^{-\lambda t}}{t} \int_0^\infty (e^{izx} - 1)\, d_x F(t, x)\, dt\right\}.$$

Обозначим далее

$$Q_{-}(t, x) = \mathsf{P}\left\{\inf_{0 \leqslant s \leqslant t} \xi(s) < x\right\},$$

$$q_{-}(\lambda, x) = \lambda \int_0^\infty e^{-\lambda t} Q_{-}(t, x)\, dt, \quad \tilde{q}_{-}(\lambda, z) = \int_{-\infty}^{\infty} e^{izx}\, d_x q_{-}(\lambda, x).$$

Тогда

$$\tilde{q}_{-}(\lambda, z) = \exp\left\{\int_0^\infty \frac{e^{-\lambda t}}{t} \int_{-\infty}^{0} (e^{izx} - 1)\, d_x F(t, x)\, dx\right\}.$$

16.4.4. Распределение момента и величины перескока. Введем величины: при $a > 0$

$$\tau_a = \inf\,[t\colon \xi(t) > a], \quad \gamma_a = \xi(\tau_a + 0) - a,$$

величина τ_a называется *моментом первого перескока через уровень* a, γ_a — *величиной перескока*. Если $\sup_t \xi(t) \leqslant a$, считаем $\tau_a = +\infty$; γ_a в этом случае не определено. Положим $M(x) = \int_x^\infty \Pi(dy)$ $(x > 0)$, где Π — мера, входящая в характеристическую функцию процесса $\xi(t)$ (см. формулу (3.1)). Тогда совместное преобразование Лапласа величин τ_a и γ_a определяется соотношением

$$\mathsf{M} e^{-\lambda \tau_a - \mu \gamma_a} = 1 - q_{+}(\lambda, a) - \\ - \frac{\mu}{\lambda} \int_0^a \left\{\int_{-\infty}^{0} \left[\int_0^\infty e^{-\mu y} M(a + y - u - v)\, dy\right] dq_{-}(\lambda, u)\right\} dq_{+}(\lambda, v)$$

($q_{\pm}(\lambda, x)$ определены в п. 16.4.3).

Совместное распределение величин τ_a и γ_a задается формулой: при $y > 0$

$$\mathsf{P}\{\tau_a < t, \gamma_a > y\} = \int_0^a M(a + y - u)\, d_u Q_{+}(t, u) + \\ + \int_0^a \int_{-\infty}^{0} M(a + y - z - u)\, d_z \int_0^t d_u Q_{+}(t - s, u)\, d_s Q_{-}(s, z).$$

Если $a < 0$, $\tau_a = \inf[t\colon \xi(t) < a]$, $\gamma_a = \xi(\tau_a + 0) - a$, то совместное преобразование Лапласа величин τ_a и γ_a определяется формулой

$$\mathsf{M}e^{-\lambda\tau_a + \mu\gamma_a} = q_-(\lambda, a) -$$
$$- \frac{\mu}{\lambda}\int_a^0 \left[\int_0^\infty \left[\int_{-\infty}^0 e^{\mu y} N(a + y - u - v)\, dy\right] dq_+(\lambda, u)\right] dq_-(\lambda, v),$$

где $N(z) = \int_{-\infty}^{z} \Pi(dz)$ для $z < 0$.

Совместное распределение этих величин можно записать так: при $y < 0$

$$\mathsf{P}\{\tau_a < t,\ \gamma_a < y\} = \int_a^0 N(a + y - u)\, d_u Q_-(t + u) +$$
$$+ \int_a^0 \int_0^\infty N(a + y - z - u)\, d_z \int_0^t d_u Q_-(t - s, u)\, d_s Q_-(s, z).$$

16.4.5. Совместное распределение supremum'а, infimum'а и значения процесса. Положим для $a < 0 < b$, $(\alpha, \beta) \subset (a, b)$

$$Q(t;\ a, b;\ \alpha, \beta) = \mathsf{P}\left\{\inf_{s \leqslant t} \xi(s) \geqslant a,\ \sup_{s \leqslant t} \xi(s) \leqslant b,\ \xi(t) \in (\alpha, \beta)\right\},$$

$$\Gamma(x, dt, dy) = \mathsf{P}\{\tau_x \in dt,\ \gamma_x \in dy\};$$

при $x > 0$ эта мера по dy сосредоточена на $[0, \infty)$, при $x < 0$ — на $(-\infty, 0)$.

Рассмотрим также преобразования Лапласа по t этих функций:

$$\Gamma^{(\lambda)}(x, A) = \int_0^\infty e^{-\lambda t}\Gamma(x, dt, A),$$

$$q(\lambda;\ a, b;\ \alpha, \beta) = \int_0^\infty Q(t;\ a, b;\ \alpha, \beta)\, e^{-\lambda t}\, dt.$$

Пусть далее

$$r_\lambda(A) = \int_0^\infty e^{-\lambda t}\mathsf{P}\{\xi(t) \in A\}\, dt.$$

Заметим, что функцию $\Gamma^{(\lambda)}(x, A)$ можно определить, используя результаты предыдущего пункта. При $x > 0$, $y > 0$

$$\Gamma^{(\lambda)}(x, (y, \infty)) = \frac{1}{\lambda}\int_0^x \int_{-\infty}^0 M(x + y - u - v)\, dq_-(\lambda, u)\, dq_+(\lambda, v),$$

а при $x < 0$, $y < 0$

$$\Gamma^{(\lambda)}(x, (-\infty, y)) = \frac{1}{\lambda} \int_x^0 \int_0^\infty N(x + y - u - v)\, dq_+(\lambda, u)\, dq_-(\lambda, v).$$

Для $x \in [b, \infty)$ положим

$$G^{(\lambda)}(x, A) = \int \Gamma^{(\lambda)}(a - b - x, dy)\, \Gamma^{(\lambda)}(b - a - y, A_{-b}),$$

где $A_{-b} = \{x\colon x + b \in A\}$. Для $x \in (-\infty, a]$ положим

$$G^{(\lambda)}(x, A) = \int \Gamma^{(\lambda)}(b - a - x, dy)\, \Gamma^{(\lambda)}(a - b - y, A_{+a}),$$

$A_{+a} = A_{-|a|}$. Наконец, при $a < x < b$ полагаем $G^{(\lambda)}(x, A) = 0$. Пусть далее

$$H^{(\lambda)}(\mu, x, A) = \chi_A(x) + \sum_{k=1}^{\infty} \mu^k \int \ldots \int G^{(\lambda)}(x, dx_1) \ldots G^{(\lambda)}(x_{k-1}, A).$$

Тогда

$$q(\lambda; a, b; \alpha, \beta) = r_\lambda((\alpha, \beta)) - \iint \left(\Gamma^{(\lambda)}(a, dy) + \Gamma^{(\lambda)}(b, dy)\right) \times$$
$$\times H^{(\lambda)}(1, y, dx)\, r_\lambda((\alpha - x, \beta - x)) +$$
$$+ \iiint \left[\Gamma^{(\lambda)}(b, dy)\, H^{(\lambda)}(1, y, dz)\, \Gamma^{(\lambda)}(a - b - z, dx) +\right.$$
$$\left.+ \Gamma^{(\lambda)}(a, dy)\, H^{(\lambda)}(1, y, dz)\, \Gamma^{(\lambda)}(b - a - z, dx)\right] r_\lambda((\alpha - x, \beta - x)).$$

Рассмотрим также совместное распределение значения процесса и его supremum'а. Если $0 < x \leqslant a$, то

$$\mathsf{P}\left\{\sup_{s \leqslant t} \xi(s) < a,\ \xi(t) < x\right\} =$$
$$= \mathsf{P}\{\xi(t) < x\} + \int_0^t \int_0^\infty \Gamma(a, ds, dy)\, \mathsf{P}\{\xi(t - s) < x - a - y\};$$

если же $0 < a < x$, то

$$\mathsf{P}\left\{\sup_{s \leqslant t} \xi(s) < a,\ \xi(t) < x\right\} = Q_+(t, a).$$

16.4.6. Распределение supremum'а процесса на бесконечном промежутке. Предположим, что $\int_0^\infty \frac{1}{t} \mathsf{P}\{\xi(t) > 0\}\, dt < \infty$, и, следовательно (см. § 3.3), $\mathsf{P}\left\{\sup_t \xi(t) < \infty\right\} = 1$.

Тогда

$$\int_0^\infty e^{izx}\, d_x \mathsf{P}\left\{\sup_t \xi(t) < x\right\} = \exp\left\{\int_0^\infty \frac{1}{t}\int_0^\infty \left(e^{izx}-1\right) d_x F(t, x)\, dt\right\},$$

где $F(t, x) = \mathsf{P}\{\xi(t) < x\}$.

Рассмотрим случай, когда кумулянта процесса имеет вид (4.6), т. е. процесс имеет лишь отрицательные скачки. Тогда

$$\mathsf{P}\left\{\sup_t \xi(t) < x\right\} = 1 - e^{-kx},$$

где k — положительный корень уравнения $K_+(k) = 0$, $K_+(z)$ задается равенством (4.7).

16.5. Процесс Пуассона

16.5.1. Определение однородного процесса Пуассона. Однородный процесс с независимыми приращениями $\xi(t)$ называется однородным процессом Пуассона, если $\xi(t)$ имеет распределение Пуассона. Тогда существует такое $a > 0$, что для всех $k \geqslant 0$

$$\mathsf{P}\{\xi(h) = k\} = \mathsf{P}\{\xi(t+h) - \xi(t) = k\} = \frac{(ah)^k}{k!}e^{-ah}. \qquad (5.1)$$

Характеристическая функция процесса Пуассона имеет вид

$$\varphi(t, z) = \mathsf{M}e^{iz\xi(t)} = \exp\left\{at\left[e^{iz}-1\right]\right\}.$$

Укажем одну общую ситуацию, когда явления описываются с помощью процесса Пуассона.

Пусть в некотором эксперименте наблюдаются появления некоторых событий. Если 1) число событий, которые произошли на промежутке времени $[t, t+h]$, не зависит от того, сколько и когда произошло событий в промежутке $[0, t]$; 2) вероятность того, что на промежутке $[t, t+h]$ произойдет одно событие, равна $ah + o(h)$; 3) вероятность того, что на промежутке $[t, t+h]$ произойдет более одного события, равна $o(h)$, то величина $\xi(t)$, равная числу событий, которые произошли на промежутке $[0, t]$, как функция t будет процессом Пуассона.

Всякий однородный ступенчатый процесс с независимыми приращениями, все скачки которого равны 1, является процессом Пуассона.

16.5.2. Некоторые свойства процесса Пуассона. Рассмотрим некоторые свойства процесса $\xi_\gamma(t) = \gamma t + \xi(t)$, где $\xi(t)$ — процесс Пуассона с распределениями, задаваемыми формулой (5.1).

1) Если $\gamma + a > 0$, то

$$\mathsf{P}\left\{\sup_t \xi_\gamma(t) = +\infty\right\} = \mathsf{P}\left\{\inf_t \xi_\gamma(t) > -\infty\right\} = 1;$$

если $\gamma + a < 0$, то

$$\mathsf{P}\left\{\sup_t \xi_\gamma(t) < +\infty\right\} = \mathsf{P}\left\{\inf_t \xi_\gamma(t) = -\infty\right\} = 1;$$

если $\gamma + a = 0$, то

$$\mathsf{P}\left\{\sup_t \xi_\gamma(t) = +\infty\right\} = \mathsf{P}\left\{\inf_t \xi_\gamma(t) = -\infty\right\} = 1.$$

2) Пусть $\gamma < 0$, $\gamma + a > 0$. Тогда для $x < 0$

$$\mathsf{P}\left\{\inf_t \xi_\gamma(t) < x\right\} = e^{kx}, \tag{5.2}$$

где k — положительный корень уравнения

$$a(e^{-k} - 1) - k\gamma = 0. \tag{5.3}$$

3) Пусть $\gamma < 0$, $\gamma + a < 0$. Тогда для $x > 0$

$$\mathsf{P}\left\{\sup_t \xi_\gamma(t) > x\right\} = 1 - \left(1 + \frac{a}{\gamma}\right) \sum_{k=0}^{[x]} \frac{(x-k)^k}{k!} \left(\frac{a}{\gamma}\right)^k e^{-a(x-k)/\gamma}, \tag{5.4}$$

где $[x]$ — целая часть x.

4) Предположим, что $c < 0 < d$. Обозначим через $p(c, d)$ вероятность того, что процесс $\xi_\gamma(t)$ достигнет уровня c раньше, чем попадет в интервал (d, ∞). Тогда при $\gamma < 0$

$$p(c, d) = \sum_{k=0}^{[d]} \left(\frac{a}{\gamma}\right)^k \frac{1}{k!} e^{ka/\gamma} (d-k)^k \times$$
$$\times \left(e^{ac/\gamma} \sum_{k=0}^{[d-c]} \left(\frac{a}{\gamma}\right)^k \frac{1}{k!} e^{ka/\gamma} (d-c-k)^k\right)^{-1}; \tag{5.5}$$

здесь $[x]$ — целая часть x.

16.5.3. Неоднородный процесс Пуассона. Это стохастически непрерывный процесс с независимыми приращениями $\xi(t)$, для которого приращения $\xi(t+h) - \xi(t)$ имеют распределения Пуассона. Для такого процесса существует неубывающая функция $a(t)$ такая, что

$$\mathsf{P}\{\xi(t+h) - \xi(t) = k\} = \frac{[a(t+h) - a(t)]^k}{k!} e^{-[a(t+h) - a(t)]}. \tag{5.6}$$

Процесс Пуассона описывает число появлений некоторых случайных событий, если выполнены условия:

1) в каждом конечном промежутке происходит с вероятностью 1 конечное число событий;

2) число появлений событий на непересекающихся промежутках независимо;

3) вероятность появления хотя бы одного события на некотором интервале стремится к нулю, если длина интервала стремится к нулю;

4) вероятность появления одновременно двух и более событий равна нулю.

Если выполнены эти условия и $\xi(t)$ обозначает число событий, появившихся на промежутке $[t_0, t]$, то $\xi(t)$ — процесс Пуассона.

Если функция $a(t)$, входящая в формулу (5.6), строго монотонна, то процесс простым преобразованием можно превратить в однородный. Пусть $\xi(t)$ определен на $[t_0, \infty]$ и $a(t_0) = 0$. Обозначим через $\lambda(t)$ функцию, обратную к $a(t)$: $a(\lambda(t)) = t$. Функция $\lambda(t)$ определена на полуинтервале $[0, a(+\infty))$. Пусть $\xi_1(t) = \xi(\lambda(t))$. Тогда при $0 \leqslant t < t+h < a(+\infty)$

$$\mathsf{P}\{\xi_1(t+h) - \xi_1(t) = k\} = \mathsf{P}\{\xi(\lambda(t+h)) - \xi(\lambda(t)) = k\} =$$
$$= \frac{[a(\lambda(t+h)) - a(\lambda(t))]^k}{k!} \exp\{-[a(\lambda(t+h)) - a(\lambda(t))]\} =$$
$$= \frac{h^k}{k!} e^{-h}.$$

Таким образом, $\xi_1(t)$ — однородный пуассоновский процесс с параметром 1. Указанное преобразование позволяет сводить решение многих задач для общего процесса Пуассона к решению задач для однородного процесса.

16.6. Винеровский процесс

16.6.1. Определение и некоторые свойства. *Винеровским процессом* в $\mathbf{R}^m$ называется однородный процесс с независимыми приращениями, для которого $\xi(t)$ имеет гауссовское распределение с плотностью

$$p_t(x) = (2\pi t)^{-m/2} \exp\left\{-\frac{(x, x)}{2t}\right\}. \tag{6.1}$$

Этот процесс также называется *m-мерным винеровским процессом*. Характеристическая функция процесса имеет вид

$$\varphi_t(z) = \mathsf{M} \exp\{i(z, \xi(t))\} = e^{-t(z, z)/2}. \tag{6.2}$$

Отметим некоторые свойства многомерного винеровского процесса.

1) Сепарабельный винеровский процесс с вероятностью 1 непрерывен.

2) *Локальный закон повторного логарифма*:

$$\mathsf{P}\left\{\overline{\lim_{t \downarrow 0}} \frac{\xi(t)}{\sqrt{2t \ln\ln 1/t}} = 1\right\} = \mathsf{P}\left\{\underline{\lim_{t \downarrow 0}} \frac{\xi(t)}{\sqrt{2t \ln\ln 1/t}} = -1\right\} = 1.$$

3) *Закон повторного логарифма*:

$$\mathsf{P}\left\{\overline{\lim_{t \to \infty}} \frac{\xi(t)}{\sqrt{2t \ln\ln t}} = 1\right\} = \mathsf{P}\left\{\underline{\lim_{t \to \infty}} \frac{\xi(t)}{\sqrt{2t \ln\ln t}} = -1\right\} = 1.$$

4) Если размерность пространства $m \geqslant 3$, то

$$\mathsf{P}\left\{\lim_{t \to +\infty} |\xi(t)| = +\infty\right\} = 1,$$

при этом для всех $\lambda > 1$

$$\mathsf{P}\left\{\underline{\lim_{T \to \infty}} \frac{(\ln T)^{\lambda/m - 1/2}}{\sqrt{T}} \inf_{t > T} |\xi(t)| = +\infty\right\} = 1.$$

5) Если $z \in \mathbf{R}^m$ и $|z| = 1$, то процесс $(z, \xi(t)) = \xi_z(t)$ является одномерным винеровским процессом. Пусть далее $\{e_1, \ldots e^m\}$ — ортонормированный базис в $\mathbf{R}^m$. Тогда процессы $\xi_{e_1}(t), \ldots, \xi_{e_m}(t)$ — независимые между собой одномерные винеровские процессы.

6) Пусть $a \in \mathbf{R}^m$, C — линейный оператор в $\mathbf{R}^m$, $\xi(t)$ — m-мерный винеровский процесс. Тогда

$$\xi_1(t) = ta + C\xi(t) \tag{6.3}$$

является однородным гауссовским процессом с независимыми приращениями. Характеристическая функция его имеет вид

$$\mathsf{M}e^{i(z, \xi_1(t))} = \exp\left\{it(z, a) - \frac{t}{2}(Bz, z)\right\}, \tag{6.4}$$

где $B = CC^*$ (C^* — оператор, сопряженный к C). Всякий гауссовский однородный процесс с независимыми приращениями (его характеристическая функция обязательно имеет вид (6.4)) представим в виде (6.3); в качестве C можно взять оператор $B^{1/2}$ (положительный квадратный корень из неотрицательного оператора). Поэтому для всякого гауссовского однородного процесса с независимыми приращениями $\xi_1(t)$ существуют такие векторы a, e_1, e_2, ..., e_m и независимые однородные винеровские процессы $w_1(t)$, ..., $w_m(t)$ и числа $\beta_1, \ldots, \beta_m$, что

$$\xi_1(t) = ta + \sum_{k=1}^{m} \beta_k w_k(t) e_k. \tag{6.5}$$

В качестве векторов e_k можно взять собственные векторы оператора B, $\beta_k = \sqrt{(Be_k, e_k)}$.

16.6.2. Метод дифференциальных уравнений. Пусть $\xi(t)$ — m-мерный винеровский процесс. Обозначим через Δ оператор Лапласа в $\mathbf{R}^m$:

$$\Delta u = \sum_{k=1}^{m} \frac{\partial^2 u}{(\partial x^k)^2},$$

где $x^1, \ldots, x^m$ — координаты точки x в фиксированном ортонормированном базисе в $\mathbf{R}^m$.

Теорема 1. *Функция* $\mathsf{M}f(x + \xi(t)) = u(t, x)$, *где* f — *непрерывная ограниченная функция, удовлетворяет параболическому уравнению*

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2}\Delta u(t, x)$$

с начальным условием

$$\lim_{t \downarrow 0} u(t, x) = f(x).$$

Теорема 2. *Пусть* $g(x)$ — *ограниченная непрерывная функция. Тогда*

$$\mathsf{M}f(x + \xi(t)) \exp\left\{\int_0^t g(x + \xi(s))\, ds\right\} = v(t, x)$$

удовлетворяет уравнению

$$\frac{\partial v(t, x)}{\partial t} = \frac{1}{2}\Delta v(t, x) + g(x) v(t, x)$$

с начальным условием

$$\lim_{t \downarrow 0} v(t, x) = f(x).$$

Замечание. Теоремы 1 и 2 остаются справедливыми, если вместо ограниченности f и g требовать, чтобы выполнялось условие

$$\lim_{|x| \to \infty} \frac{\ln(1 + |f(x)|) + g(x)}{|x|^2} = 0.$$

Теорема 3. *Пусть $G \subset R^m$ — связная область с гладкой границей Γ. Обозначим через τ_x момент первого достижения процессом $x + \xi(t)$ $(x \in G)$ границы Γ:*

$$\tau_x = \inf [t\colon x + \xi(t) \notin G].$$

Величина τ_x может принимать значение $+\infty$. Пусть далее $\varphi(x)$ — произвольная непрерывная функция, определенная на Γ. Тогда:

а) *функция*

$$u(x) = \mathsf{M}\varphi(x + \xi(\tau_x))\, \chi_{\{\tau_x < \infty\}}$$

$\left(\chi_{\{\tau_x<\infty\}} = 1,\right.$ *если* $\tau_x \lessgtr \infty$, *и* $\chi_{\{\tau_x<\infty\}} = 0$, *если* $\left.\tau_x = +\infty\right)$ *удовлетворяет уравнению и граничному условию*

$$\Delta u(x) = 0, \quad u(x) = \varphi(x) \text{ при } x \in \Gamma,$$

т. е. $u(x)$ — гармоническая функция в Γ с данным граничным значением φ;

б) *функция*

$$v(x) = \mathsf{M} \int_0^{\tau_x} g(x + \xi(s))\, ds,$$

где $g(x)$ — непрерывна и ограничена в G, удовлетворяет уравнению и граничному условию:

$$\Delta v(x) = -2g(x), \quad v(x) = 0, \quad x \in \Gamma;$$

в) *функция*

$$w(x) = \mathsf{M}\varphi(x + \xi(\tau_x)) \exp\left\{\int_0^{\tau_x} g(x + \xi(s))\, ds\right\}$$

удовлетворяет уравнению и граничному условию

$$\frac{1}{2}\Delta w(x) + g(x) w(x) = 0, \quad w(x) = \varphi(x), \quad x \in \Gamma;$$

г) *функция*

$$u(t, x) = \mathsf{M} f(x + \xi(t)) \exp\left\{\int_0^t g(x + \xi(s))\, ds\right\} \chi_{\{\tau_x > t\}}$$

$\left(\chi_{\{\tau_x > t\}} = 1\right.$, *если* $\tau_x > t$, *и* $\chi_{\{\tau_x > t\}} = 0$, *если* $\left.\tau_x \leqslant t\right)$, *удовлетворяет уравнению и граничным условиям*

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2} \Delta u(t, x) + g(x) u(t, x), \quad x \in G; \quad u(0, x) = f(x)$$

$$x \in G, \quad u(t, x) = 0, \quad x \in \Gamma, \quad t > 0.$$

Функции f и g непрерывные в замыкании G; в случае неограниченного G они должны удовлетворять условию, сформулированному в замечании.

16.6.3. Одномерный винеровский процесс. Рассмотрим распределения некоторых функционалов от одномерного винеровского процесса $w(t)$.

1. *Распределение максимума.* Для $x > 0$

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} w(s) < x\right\} = \frac{2}{\sqrt{2\pi t}} \int_0^x e^{-u^2/(2t)}\, du,$$

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} w(s) > x\right\} = 2\mathsf{P}\{w(t) > x\}.$$

2. *Распределение времени первого прохождения.* Пусть $x > 0$, $\tau_x = \inf[t\colon w(t) > x]$. Тогда величина τ_x имеет плотность распределения: при $s > 0$

$$\frac{d}{ds} \mathsf{P}\{\tau_x < s\} = \frac{x}{\sqrt{2\pi s^3}} e^{-x^2/(2s)}.$$

3. *Совместное распределение максимума и значения процесса.* При $x < a$, $a > 0$

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} w(s) < a, \quad w(t) < x\right\} =$$

$$= \mathsf{P}\{w(t) < x\} - \mathsf{P}\{w(t) > 2a - x\} =$$

$$= \mathsf{P}\{w(t) < x\} - \mathsf{P}\{w(t) < x - 2a\} = \frac{1}{\sqrt{2\pi t}} \int_{x-2a}^{x} e^{-u^2/(2t)}\, du.$$

4. *Распределение максимума модуля:*

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} |w(s)| < a\right\} =$$

$$= \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} (-1)^k \int_{-a}^{a} \exp\left\{-\frac{1}{2t}(u - 2ka)^2\right\} du.$$

5. *Совместное распределение максимума модуля и значения процесса.* При $[c, d] \subset [-a, a]$

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant t} |w(s)| < a, \quad w(t) \in [c, d]\right\} =$$

$$= \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} (-1)^k \int_{c}^{d} \exp\left\{-\frac{1}{2t}(u - 2ka)^2\right\} du.$$

6. *Совместное распределение максимума, минимума и значения процесса.* Пусть $a < 0 < b$, $(\alpha, \beta) \subset (a, b)$. Тогда

$$\mathsf{P}\left\{\min_{0 \leqslant s \leqslant t} w(s) > a, \quad \max_{0 \leqslant s \leqslant t} w(s) < b, \quad w(t) \in (\alpha, \beta)\right\} =$$

$$= \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} \int_{\alpha}^{\beta} \left[\exp\left\{-\frac{1}{2t}(x + 2k(b-a))^2\right\} - \right.$$

$$\left. - \exp\left\{-\frac{1}{2t}(x - 2a + 2k(b-a))^2\right\}\right] dx.$$

7. Обозначим через τ момент первого выхода процесса из отрезка $[a, b]$; $a < 0 < b$:

$$\tau = \min[t\colon w(t) \notin [a, b]].$$

Тогда

$$\mathsf{P}\{w(\tau\} = b\} = \frac{-a}{b-a}; \quad \mathsf{P}\{w(t) = a\} = \frac{b}{b-a};$$

$$\mathsf{P}\{\tau < t, \quad w(\tau) = a\} =$$

$$= \frac{2}{\sqrt{2\pi t}} \int_{-b}^{b} \sum_{k=0}^{\infty} \exp\left\{-\frac{1}{2t}(x + (2k+1)(b-a))^2\right\} dx.$$

8. *Закон арксинуса.* Пусть $\varepsilon(x) = 1$ при $x > 0$, $\varepsilon(x) = 0$ при $x \leqslant 0$. Тогда при $0 < x < t$

$$\mathsf{P}\left\{\int_0^t \varepsilon(w(s))\, ds < x\right\} = \frac{2}{\pi} \arcsin\sqrt{\frac{x}{t}}.$$

16.6.4. Одномерный винеровский процесс со сносом. Пусть $\xi(t) = at + w(t)$, a — некоторое число, $w(t)$ — одномерный винеровский процесс. Рассмотрим некоторые функционалы от процесса.

1. Пусть $x > 0$, τ_x — момент первого попадания в точку x. Тогда при $a \geqslant 0$

$$\mathsf{M}e^{-\lambda\tau_x} = \exp\left\{-x\left(\sqrt{a^2 + 2\lambda} - a\right)\right\}, \quad \lambda > 0.$$

Если $a < 0$, то

$$\mathsf{P}\{\tau_x = +\infty\} = \mathsf{P}\left\{\sup_t \xi(t) < x\right\} = 1 - e^{2ax}.$$

2. Пусть $c < 0 < d$, $(\alpha, \beta) \subset (c, d)$, $Q(t;\, c, d;\, \alpha, \beta) =$

$$= \mathsf{P}\left\{\min_{s \leqslant t} \xi(s) > c, \ \max_{s \leqslant t} \xi(s) < d, \ \xi(t) \in (\alpha, \beta)\right\},$$

$$q(\lambda;\, c, d;\, \alpha, \beta) = \int_0^{\infty} e^{-\lambda t} Q(t;\, c, d;\, \alpha, \beta)\, dt.$$

Тогда

$$q(\lambda;\, c,\, d;\, \alpha,\, \beta) = \frac{1}{\sqrt{a^2+2\lambda}}\left[\int_\alpha^\beta \exp\left\{ay - \sqrt{a^2+2\lambda}\,|\,y\,|\right\} dy - \right.$$

$$- \frac{\operatorname{sh}\sqrt{a^2+2\lambda}\,c}{\operatorname{sh}\sqrt{a^2+2\lambda}\,(d-c)} \int_\alpha^\beta \exp\left\{ay - \sqrt{a^2+2\lambda}\,(d-y)\right\} dy -$$

$$\left. - \frac{\operatorname{sh}\sqrt{a^2+2\lambda}\,d}{\operatorname{sh}\sqrt{a^2+2\lambda}\,(d-c)} \int_\alpha^\beta \exp\left\{ay - \sqrt{a^2+2\lambda}\,(y-c)\right\} dy\right].$$

3. Пусть $c < 0 < d$, τ — момент первого выхода из $[c, d]$:

$$\tau = \inf\,[t\colon\ \xi(t) \notin [c,\, d]].$$

Тогда

$$\mathsf{P}\{\xi(\tau) = c\} = \frac{1 - e^{-2ad}}{e^{-2ac} - e^{-2ad}}.$$

Литература: [19, 20, 23, 55, 68, 81, 82, 87].

Глава 17. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ

17.1. Ветвящиеся поцессы с одним типом частиц (дискретное время)

17.1.1. Определение. Ветвящиеся процессы являются моделью многих реальных явлений размножения, гибели и превращения частиц в биологии, физике, технике, демографии и др.

Однородная цепь Маркова $\xi(t)$ $(t = 0, 1, 2, \ldots)$ с неотрицательными целочисленными значениями называется *ветвящимся процессом с одним типом частиц* или *процессом Гальтона — Ватсона*, если ее переходные вероятности $p_{ij}(t) = \mathsf{P}\{\xi(t) = j\,|\,\xi(0) = i\}$ за время t удовлетворяют условиям

$$p_{ij}(t) = \begin{cases} \delta_{0j}, & i = 0, \\ \sum\limits_{j_1+\ldots+j_i=j} p_{1j_1}(t)\, p_{1j_2}(t) \ldots p_{1j_i}(t). \end{cases} \tag{1.1}$$

Принята следующая теорминология. Модель, описываемую ветвящимся процессом, часто называют *популяцией*. Значение ветвящегося процесса $\xi(t)$ в момент t называют *числом частиц* или *индивидуумов* в популяции в поколении с номером t. Говорят также, что $\xi(t)$ есть *общее число потомков* $\xi(0)$ *частиц нулевого поколения* в поколении с номером t.

Первое равенство в (1.1) означает отсутствие самовозрождения популяции после того, как все частицы исчезли, либо отсутствие иммиграции (притока частиц извне).

Второе равенство в (1.1), означающее, что $p_{ij}(t)$ при $i \geqslant 1$ является i-кратной сверткой распределения $p_{1j}(t)$ $(j = 0, 1, 2, \ldots)$ с собой, эквивалентно предположению, что каждая из i первоначальных частиц эволюционирует (гибнет, превращается в новые частицы того же типа) независимо от других. Второе равенство в (1.1) называют *условием ветвления*.

Ветвящийся процесс может быть описан в терминах суммирования независимых одинаково распределенных неотрицательных целочисленных случайных величин.

Пусть ζ_{tk} $(t, k = 1, 2, \ldots)$ — независимые одинаково распределенные случайные величины, интерпретируемые как число потомков, даваемых любой из k частиц в момент превращения в t-м поколении, т. е. $\mathrm{P}\{\zeta_{tk} = j\} = p_{1j}$ $(j = 0, 1, 2, \ldots)$. Число частиц $\xi(t+1)$ в $(t+1)$-м поколении выражается через число частиц $\xi(t)$ в t-м поколении как

$$\xi(t+1) = \begin{cases} \sum\limits_{k=1}^{\xi(t)} \zeta_{tk}, & \zeta_{0k} = \delta_{0k}, \text{ если } \xi(0) = 1; \\ \sum\limits_{k=1}^{\xi(t)} \zeta_{tk}, & \zeta_{0k} = 1,\ k = 1, \ldots, i, \text{ если } \xi(0) = i. \end{cases} \tag{1.2}$$

17.1.2. Примеры. 1. *Выживание фамилий.* Фамилию наследуют только сыновья. Предположим, что каждый индивидуум с вероятностью p_j имеет j потомков мужского пола. Каждый индивидуум порождает первое поколение потомков, те в свою очередь второе и т. д. Общее число потомков в t-м поколении равно $\xi(t)$.

Особый интерес представляет исследование числа потомков в t-м поколении, т. е. распределение $\xi(t)$, а также определение вероятности вырождения фамилии $q = \mathrm{P}\{\xi(t) = 0$ для некоторого $t \mid \xi(0) > 0\}$.

2. *Электронный умножитель* представляет собой устройство для усиления слабого потока электронов. На пути потока электронов, испускаемых источником (число $\xi(0)$ таких электронов есть нулевое поколение), устанавливается последовательно ряд пластин. Каждый электрон, ударяясь о первую пластину, порождает случайное число новых электронов (первое поколение), которые в свою очередь ударяются о следующую пластину. Процесс $\xi(t)$ — число электронов, испущенных t-й пластиной, — представляет собой ветвящийся процесс.

3. *Нейтронная цепная реакция.* При взаимодействии с нейтроном ядро расщепляется, испуская случайное число новых нейтронов. Каждый из этих вторичных нейтронов может бомбардировать другие ядра, производя случайное число новых нейтронов, и т. д. Если первоначальное число нейтронов равнялось 1 (нулевое поколение), то первое поколение нейтронов, порожденных исходным нейтроном, есть случайная величина $\xi(1)$. Размер t-го поколения $\xi(t)$ складывается из случайного числа нейтронов, порожденных $\xi(t-1)$ нейтронами $(t-1)$-го поколения.

17.1.3. Уравнения для производящих функций. Целочисленность значений ветвящихся процессов и в особенности определяющие эти процессы равенства (1.1) и (1.2) делают аппарат производящих функций (см. гл. 3) основным при их исследовании.

Замечание. Для ветвящихся процессов $\xi(t)$ обычно предполагается (и это, как правило, делается ниже), что $\xi(0) = 1$; это не ограничивает общности, так как в силу определения ветвящегося процесса при $\xi(0) > 1$ мы имеем дело с $\xi(0)$ независимо развивающимися процессами, начинающимися от каждой из $\xi(0)$ первоначальных частиц.

Пусть $p_j = \mathsf{P}\{\xi(1) = j \mid \xi(0) = 1\}, p_{1j}(t) = \mathsf{P}\{\xi(t) = j \mid \xi(0) = 1\}$; $\Phi(s) = \mathsf{M}\left[s^{\xi(1)} \mid \xi(0) = 1\right]$ и $\Phi_t(s) = \mathsf{M}\left[s^{\xi(t)} \mid \xi(0) = 1\right]$ — производящие функции этих распределений, т. е.

$$\Phi_t(s) = \sum_{j=0}^{\infty} p_{1j}(t)\, s^j, \quad \Phi(s) = \Phi_1(s). \tag{1.3}$$

Функцию $\Phi_t(s)$ называют *производящей функцией* ветвящегося процесса $\xi(t)$ $(t = 0, 1, \ldots)$.

Теорема 1. *Производящая функция $\Phi_t(s)$ при любых t, $\tau \geqslant 0$ удовлетворяет основному функциональному уравнению*

$$\Phi_{t+\tau}(s) = \Phi_t(\Phi_\tau(s)) \tag{1.4}$$

и начальному условию

$$\Phi_0(s) = s. \tag{1.5}$$

Таким образом, $\Phi_t(s)$ есть t-кратная итерация производящей функции $\Phi(s)$: $\Phi_1(s) = \Phi(s)$, $\Phi_2(s) = \Phi(\Phi(s))$, $\Phi_3(s) = \Phi(\Phi_2(s)) = \Phi_2(\Phi(s)) = \Phi(\Phi(\Phi(s))), \ldots$, и вообще

$$\Phi_t(s) = \Phi(\Phi(\ldots\Phi(\Phi(s))\ldots)). \tag{1.6}$$

Если $\Phi^{(t)}(s_1, s_2, \ldots, s_t) = \mathsf{M}\left[s_1^{\xi(1)} \ldots s_t^{\xi(t)} \mid \xi(0) = 1\right]$ — совместная производящая функция случайных величин $\xi(1), \xi(2), \ldots, \xi(t)$, то

$$\Phi^{(t)}(s_1, \ldots, s_t) = \Phi(s_1\Phi(s_2\Phi(\ldots s_{t-1}\Phi(s_t)\ldots))).$$

Пусть $F_t(s)$ есть производящая функция суммы $\xi(0) + \ldots + \xi(t)$ — общего числа частиц в популяции за время $[0, t]$ $(t = 0, 1, \ldots)$, и пусть $F(s)$ — производящая функция суммы $\xi(0) + \xi(1) + \ldots$ — общего числа частиц в популяции. Тогда

$$F_{t+1}(s) = s\Phi(F_t(s)), \quad F(s) = s\Phi(F(s)). \tag{1.7}$$

17.1.4. Примеры. 1. *Процесс гибели.* Пусть $\xi(0) = 1$, $p_0 = \mathsf{P}\{\xi(1) = 0 \mid \xi(0) = 1\} = 1 - p$, $p_1 = \mathsf{P}\{\xi(1) = 1 \mid \xi(0) = 1\} = p$, $0 < p < 1$, $p_k = 0$, $k > 1$. Тогда $\Phi(s) = 1 - p + sp$, $\Phi_t(s) = 1 - p^t + sp^t$, откуда $p_{10}(t) = 1 - p^t$, $p_{11}(t) = p^t$, $p_{1k}(t) = 0$ $(k > 1)$.

2. *Дробно-линейные производящие функции.* Пусть $p_0 = \dfrac{1-(b+c)}{1-c}$, $p_k = bc^{k-1}$, $b, c > 0$, $b + c < 1$. Тогда $\Phi(s) = \dfrac{1-(b+c)}{1-c} + \dfrac{bs}{1-cs}$. Функция $\Phi(s)$ является дробно-линейной вида $(\alpha + \beta s)/(\gamma + \delta s)$. Пусть $m = \Phi'(1) = b/(1-c)^2$ и

$$q = \begin{cases} 1, & \text{если } m \leqslant 1; \\ p_0/c, & \text{если } m > 1. \end{cases}$$

Тогда

$$\Phi_t(s)= \begin{cases} \left(1-m^t\dfrac{1-q}{m^t-q}+m^t\left[\dfrac{1-q}{m^t-q}\right]^2 s\left(1-\left[\dfrac{m^t-1}{m^t-q}\right]s\right)\right)^{-1}, & \text{если } m\neq 1; \\ \dfrac{tc-[(t+1)c-1]s}{1+(t-1)c-tcs}, & \text{если } m=1. \end{cases}$$

Отсюда, например,

$$p_{10}(t)=\begin{cases} 1-m^t\left[\dfrac{1-q}{m^t-q}\right], & \text{если } m\neq 1; \\ \dfrac{tc}{1+(t-1)c}, & \text{если } m=1. \end{cases}$$

17.1.5. Моменты и классификация. Пусть $\xi(0)=1$, $m(t)=\mathsf{M}\xi(t)$, $m=m(1)$, $\sigma(t)=\mathsf{D}\xi(t)$, $\sigma^2=\sigma(1)$.

Непосредственным следствием основного функционального уравнения (1.4) для производящей функции ветвящегося процесса являются следующие выражения для $m(t)$ и $\sigma(t)$:

$$m(t)=m^t,\quad t=0,\ 1,\ \dots, \tag{1.8}$$

$$\sigma(t)=\begin{cases} \sigma^2 m^{t-1}\dfrac{m^t-1}{m-1}, & \text{если } m\neq 1; \\ \sigma^2 t, & \text{если } m=1. \end{cases} \tag{1.9}$$

Ветвящийся процесс с одним типом частиц называется *докритическим*, если $m<1$; *критическим*, если $m=1$, $\Phi''(1)>0$; *надкритическим*, если $m>1$.

Условие $\Phi''(1)>0$ в этом определении означает несингулярность, т. е. допредельную невырожденность соответствующего процесса.

Таким образом, для докритических процессов среднее $m(t)$ экспоненциально убывает, для критических процессов $m(t)$ постоянно и для надкритических процессов $m(t)$ экспоненциально растет.

17.1.6. Асимптотические свойства и предельные теоремы. Если в ветвящемся процессе $\xi(t)$ для некоторого $t_0>0$ $\xi(t_0)=0$, то говорят, что процесс $\xi(t)$ *выродился* к моменту t_0. Величину $q=\mathsf{P}\{\xi(t)=0$ для некоторого $t>0\,|\,\xi(0)=1\}$ называют *вероятностью вырождения*.

Если $q=1$, то процесс $\xi(t)$ называется *вырождающимся*.

Теорема 2. *Для того чтобы ветвящийся процесс был вырождающимся, необходимо и достаточно, чтобы он был докритическим или критическим.*

Теорема 3. *Вероятность вырождения q является наименьшим неотрицательным корнем уравнения*

$$\Phi(s)=s. \tag{1.10}$$

Вероятность вырождения q может быть найдена как один из следующих пределов:

$$q=\begin{cases}\lim\limits_{t\to\infty} p_{10}(t),\\ \lim\limits_{t\to\infty}\Phi_t(s), \quad |s|<1,\end{cases} \tag{1.11}$$

причем в последнем случае сходимость равномерна по всем $|s|\leqslant r$ $(r<1)$.

Асимптотическое поведение вероятностей $p_{10}(t)$ при $t\to\infty$ описывается следующим образом.

Теорема 4. 1) *Для докритического процесса $\xi(t)$ при $t\to\infty$*

$$m^{-t}(1-p_{10}(t))=c+o(1), \tag{1.12}$$

где

$$c=\prod_{n=0}^{\infty} h(\Phi_n(0)), \quad h(s)=\frac{1-\Phi(s)}{m(1-s)}, \tag{1.13}$$

причем $c>0$ тогда и только тогда, когда $\mathsf{M}\xi(1)\ln\xi(1)<\infty$.

2) *Для критического процесса, у которого $\Phi''(1)<\infty$,*

$$p_{10}(t)=1-\frac{2}{t\Phi''(1)}(1+o(1)). \tag{1.14}$$

3) *Для надкритического процесса*

$$p_{10}(t)=q-d\,[\Phi'(q)]^t+o\left([\Phi'(q)]^{2t}\right), \tag{1.15}$$

где $0<\Phi'(q)<1$, d — некоторая положительная константа.

Ветвящийся процесс сходится либо к нулю, либо к бесконечности, причем эта сходимость чрезвычайно неустойчива в том смысле, что если $m=\mathsf{M}\xi(1)<\infty$, то $\lim\limits_{t\to\infty} p_{1j}(t)=0$ $(j=1, 2, \ldots)$, и для любого $n\geqslant 1$

$$\lim_{t\to\infty}\mathsf{P}\{\xi(t)\geqslant n\,|\,\xi(0)=1\}=1-q.$$

Для докритических процессов существуют пределы

$$\lim_{t\to\infty}\frac{p_{1j}(t)}{1-p_{10}(t)}=\lim_{t\to\infty}\mathsf{P}\{\xi(t)=j\,|\,\xi(t)>0,\ \ \xi(0)=1\}=Q_j, \qquad j\geqslant 1, \tag{1.16}$$

и вероятности Q_1, Q_2, ... образуют распределение вероятностей, т. е.

$$\sum_{j=1}^{\infty} Q_j=1.$$

Теорема 5. *Производящая функция $Q(s)=\sum\limits_{j=1}^{\infty} Q_j s^j$ удовлетворяет функциональному уравнению*

$$1-Q(\Phi(s))=m(1-Q(s)). \tag{1.17}$$

Математическое ожидание распределения $Q_1, Q_2, \ldots$ *равно* $1/c$, *где* c *определяется равенством* (1.13).

Для критических процессов с конечным вторым моментом имеет место предельная теорема.

Теорема 6. *Если* $\xi(t)$ — *критический ветвящийся процесс с конечным вторым моментом, то*

$$\lim_{t\to\infty} \mathsf{P}\left\{\frac{\xi(t)}{\mathsf{M}[\xi(t)\mid\xi(t)>0,\ \xi(0)=1]} > x \mid \xi(t)>0,\ \xi(0)=1\right\} = e^{-x}. \tag{1.18}$$

Если $m < \infty$, то процесс $\eta(t) = m^{-t}\xi(t)$ является мартингалом, т. е.

$$\mathsf{M}[\eta(t+\tau)\mid\eta(t)] = \eta(t), \quad \tau \geqslant 0.$$

Пусть $\xi(t)$ — надкритический процесс. Из теоремы о сходимости мартингалов следует, что с вероятностью 1 процесс $\eta(t)$ сходится к некоторой случайной величине η.

Теорема 7. *Характеристическая функция* $\varphi(s) = \mathsf{M}e^{is\eta}$ *предельной случайной величины* η *удовлетворяет функциональному уравнению*

$$\varphi(ms) = \Phi(\varphi(s)), \tag{1.19}$$

имеющему единственное решение в классе характеристических функций с первым моментом, равным 1.

Если $0 < \sigma^2 < \infty$, то функция распределения $K(x) = \mathsf{P}\{\eta < x\}$ имеет скачок в нуле: $\mathsf{P}\{\eta = 0\} = q$. Условная функция распределения $\mathsf{P}\{\eta < x \mid \eta > 0\} = (K(x) - q)/(1-q)$ абсолютно непрерывна, и условная дисперсия $\mathsf{D}[\eta\mid\eta>0]$ положительна.

17.2. Ветвящиеся процессы с одним типом частиц (непрерывное время)

17.2.1. Определения. Однородная цепь Маркова $\xi(t)$ ($t \in [0, \infty)$) с неотрицательными целочисленными значениями называется *ветвящимся процессом с одним типом частиц*, если ее переходные вероятности $p_{ij}(t) = \mathsf{P}\{\xi(t) = j \mid \xi(0) = i\}$ удовлетворяют условиям:

1)

$$p_{ij}(t) = \begin{cases} \delta_{0j}, & i = 0, \\ \sum\limits_{j_1+\ldots+j_i=j} p_{1j_1}(t)\, p_{1j_2}(t) \ldots p_{1j_i}(t), & i \neq 0; \end{cases} \tag{2.1}$$

2)

$$\lim_{t\to 0} p_{ij}(t) = \delta_{ij}. \tag{2.2}$$

Пусть $\xi(0) = 1$ и переходные вероятности $p_{1j}(t)$ при значениях t, близких к нулю, удовлетворяют условию

$$\begin{aligned} p_{11}(t) &= 1 + q_1 t + o(t), & q_1 &< 0, \\ p_{1j}(t) &= q_j t + o(t), & j &\neq 1. \end{aligned} \tag{2.3}$$

Причем очевидно, что $q_j \geqslant 0$ для $j \neq 1$ (q_j в этом случае называют *плотностями перехода*).

Введем производящие функции

$$\Phi_t(s) = \sum_{j=0}^{\infty} p_{1j}(t)\, s^j = \mathsf{M}\,[s^{\xi(t)} | \xi(0) = 1], \quad f(s) = \sum_{j=0}^{\infty} q_j s^j.$$

Функцию $f(s)$ называют *инфинитезимальной* или *дифференциальной производящей функцией* ветвящегося процесса. Эволюция рассматриваемого ветвящегося процесса с непрерывным временем описывается следующим образом: каждая частица живет случайное время, распределенное по экспоненциальному закону с параметром $\lambda = \sum_{j \neq 1} q_j$. По завершении времени жизни она порождает случайное число ζ частиц того же типа, имеющее распределение

$$\mathsf{P}\{\zeta = j\} = q_j \Big(\sum_{j \neq 1} q_j \Big)^{-1}, \quad j = 0, 2, 3, \ldots$$

Простейшим примером ветвящегося процесса с непрерывным временем является процесс гибели и размножения, для которого

$$q_0 = \alpha, \quad q_2 = \beta, \quad q_1 = -(\alpha + \beta), \quad q_j = 0, \quad j = 3, 4, \ldots$$

Ветвящийся процесс $\xi(t)$ называется *регулярным*, если

$$\lim_{s \uparrow 1} \Phi_t(s) = 1. \tag{2.4}$$

Теорема 1. *Для того чтобы процесс $\xi(t)$ был регулярен, необходимо и достаточно, чтобы интеграл*

$$\int_{1-\varepsilon}^{1} \frac{du}{f(u)}$$

расходился при любом $\varepsilon > 0$.

Замечание. Если $\Phi_t(s)$ есть производящая функция регулярного ветвящегося процесса с непрерывным временем, то, полагая $\Phi(s) = \Phi_1(s)$ и ведя счет по моментам времени $t = 0, 1, 2, \ldots$, получаем производящую функцию ветвящегося процесса с дискретным временем.

При выполнении условий (2.3) и (2.4) производящая функция $\Phi_t(s)$ ветвящегося процесса равномерно по $|s| \leqslant 1$ удовлетворяет асимптотическому соотношению

$$\Phi_t(s) = s + tf(s) + o(t), \quad t \to 0. \tag{2.5}$$

Следствием (2.5) является следующая теорема, дающая аналог основного функционального уравнения для процессов с дискретным временем.

Теорема 2. *Производящая функция $\Phi_t(s)$ ветвящегося процесса с непрерывным временем удовлетворяет при $|s| \leqslant 1$:*

а) *обыкновенному (нелинейному) дифференциальному уравнению*

$$\frac{d\Phi_t(s)}{dt} = f(\Phi_t(s)) \tag{2.6}$$

с начальным условием

$$\Phi_0(s) = s; \tag{2.7}$$

б) *линейному уравнению в частных производных*

$$\frac{\partial \Phi_t(s)}{\partial t} = f(s)\frac{\partial \Phi_t(s)}{\partial s} \tag{2.8}$$

с тем же начальным условием (2.7);

в) *нелинейному интегральному уравнению*

$$\Phi_t(s) = \int_0^t h(\Phi_{t-u}(s))\,dG(u) + s(1 - G(t)), \tag{2.9}$$

где

$$G(t) = \begin{cases} 1 - e^{q_1 t}, & t \geqslant 0, \\ 0, & t < 0, \end{cases}$$

$$h(s) = \frac{f(s) - q_1 s}{-q_1} = -\sum_{j \neq 1} \frac{q_j}{q_1} s^j.$$

В уравнении (2.9) $G(t)$ интерпретируется как функция распределения времени жизни частицы, т. е. времени, прошедшему от ее рождения до первого превращения в 0, 2, 3, ... частиц, а $h(s)$ есть производящая функция условных вероятностей $\{-q_j|q_1\}$ ($j = 0$, 2, 3, ...) того, что частица превращается в j частиц, при условии что превращение имело место.

Решение уравнений (2.6), (2.8), (2.9) существует при любом $|s| < 1$ и является аналитической в круге $|s| < 1$ функцией, коэффициенты разложения которой в ряд по степеням s неотрицательны.

Для регулярных процессов решение указанных уравнений единственно.

17.2.2. Примеры. 1. Пусть $f(s) = q_0 + sq_1 + s^2 q_2$, т. е.

$$f(s) = a(s-1) + \frac{b}{2}(s-1)^2,$$

где $a = f'(1)$, $b = f''(1)$.

Уравнение (2.6) имеет вид

$$\frac{d\Phi_t(s)}{dt} = a[\Phi_t(s) - 1] + \frac{b}{2}[\Phi_t(s) - 1]^2.$$

Решение этого уравнения:

$$\Phi_t(s) = \begin{cases} 1 - \dfrac{e^{at}(1-s)}{\dfrac{b}{2a}\left(e^{at} - 1\right)(1-s) + 1}, & \text{если} \quad a \neq 0; \\[2ex] 1 - \dfrac{1-s}{\dfrac{bt}{2}(1-s) + 1}, & \text{если} \quad a = 0, \end{cases}$$

откуда, разлагая $\Phi_t(s)$ в ряд по степеням s, можно найти

$$p_{10}(t)=\begin{cases}1-\dfrac{e^{at}}{\dfrac{b}{2a}(e^{at}-1)+1}, & \text{если} \quad a\neq 0;\\[2ex] 1-\dfrac{1}{\dfrac{bt}{2}+1}, & \text{если} \quad a=0,\end{cases}$$

а для $j\neq 0$

$$p_{1j}(t)=\begin{cases}\left(\dfrac{b}{2a}(e^{at}-1)+1\right)^{-2}\left[\dfrac{\dfrac{b}{2a}(e^{at}-1)}{\dfrac{b}{2a}(e^{at}-1)+1}\right]^{j-1}, & \text{если } a\neq 0;\\[2ex] \left(\dfrac{2}{bt+2}\right)^2\left(\dfrac{bt}{bt+2}\right)^{j-1}, & \text{если } a=0.\end{cases}$$

2. Пусть $f(s)=a(s-1)+\lambda(1-s)^{1+\alpha}$ $(0<\alpha<1, \lambda>\max\{a,0\})$. Здесь второй момент процесса бесконечен и

$$\Phi_t(s)=\begin{cases}1-\left[\dfrac{\lambda}{a}(1-e^{-\alpha at})+e^{-\alpha at}(1-s)^{-\alpha}\right]^{-1/\alpha}, & \text{если} \quad a\neq 0;\\ 1-[\alpha\lambda t+(1-s)^{-\alpha}]^{-1/\alpha}, & \text{если} \quad a=0.\end{cases}$$

3. Пусть $f(s)=\lambda(s-s^{k+1})$, $\lambda>0$, k — целое положительное число. Здесь

$$\Phi_t(s)=s\left[e^{\lambda kt}-(e^{\lambda kt}-1)s^k\right]^{-1/k}.$$

4. Пусть $f(s)=\lambda[1-s][1+\ln(1-s)]$. Здесь

$$\Phi_t(s)=1-\exp\left\{e^{-\lambda t}-1+e^{-\lambda t}\ln(1-s)\right\}.$$

5. Пусть $f(s)=\lambda\left[1-s-(1-s)^{\alpha}\right]$, где $\lambda>0$, $0<\alpha<1$. Здесь

$$\Phi_t(s)=1-\left[1-e^{-(1-\alpha)\lambda t}+e^{-(1-\alpha)\lambda t}(1-s)^{1-\alpha}\right]^{1/(1-\alpha)}.$$

Это — пример нерегулярного процесса, так как

$$\lim_{s\uparrow 1}\Phi_t(s)=1-\left(1-e^{-(1-\alpha)\lambda t}\right)^{1/(1-\alpha)}<1.$$

17.2.3. Моменты и классификация. Конечность моментов $\mathsf{M}[\xi(t)]^k$ для ветвящегося процесса $\xi(t)$, $\xi(0)=1$, с непрерывным временем следует из конечности k-й производной $f(s)$ в единице.

Положим $a=f'(1)$ и $b=f''(1)$, и пусть $m(t)=\mathsf{M}\xi(t)$, $\sigma^2(t)=\mathsf{D}\xi(t)$.

Дифференцируя (2.9) по s и полагая $s=1$, можно получить следующие дифференциальные уравнения для $m(t)$ и $\sigma^2(t)$:

$$\frac{d}{dt}m(t)=am(t) \tag{2.10}$$

с начальным условием $m(0) = 1$;

$$\frac{d}{dt}\sigma^2(t) = \begin{cases} a\sigma^2(t) + (b-a)\,m^2(t), & \text{если } a \neq 0; \\ b, & \text{если } a = 0, \end{cases} \tag{2.11}$$

с начальным условием $\sigma^2(0) = 0$.

Отсюда $m(t) = e^{at}$ и

$$\sigma^2(t) = \begin{cases} \left(\frac{b}{a} - 1\right) e^{at}\left(e^{at} - 1\right), & \text{если } a \neq 0; \\ bt, & \text{если } a = 0. \end{cases} \tag{2.12}$$

Ветвящийся процесс с непрерывным временем называется *докритическим*, если $a < 0$; *критическим*, если $a = 0$, $b > 0$; *надкритическим*, если $a > 0$.

Из (2.10), в частности, следует, что среднее $m(t)$ экспоненциально убывает для докритических процессов, постоянно — для критических процессов и экспоненциально возрастает для надкритических процессов.

17.2.4. Асимптотические свойства и предельные теоремы. Пусть $q = \mathsf{P}\{\xi(t) = 0$ для некоторого $t > 0 \mid \xi(0) = 1\}$ — вероятность вырождения процесса $\xi(t)$.

Условия, при которых $q = 1$, для ветвящихся процессов с непрерывным временем те же, что и для процессов с дискретным временем (см. п. 17.1.6).

Теорема 3. *Вероятность вырождения q является наименьшим неотрицательным корнем уравнения*

$$f(s) = 0. \tag{2.13}$$

Вероятность вырождения q может быть найдена как один из следующих пределов:

$$q = \begin{cases} \lim\limits_{t \to \infty} p_{10}(t), \\ \lim\limits_{t \to \infty} \Phi_t(s), \quad |s| < 1, \end{cases} \tag{2.14}$$

причем в последнем случае сходимость равномерна по всем s, $|s| \leqslant r$, $r < 1$.

Асимптотическое поведение вероятностей $p_{10}(t)$ при $t \to \infty$ описывается следующим образом

Теорема 4. 1) *Для докритических процессов*

$$p_{10}(t) = 1 - ce^{at}(1 + o(1)), \tag{2.15}$$

если сходится интеграл $\displaystyle\int_0^1 \frac{au + f(1-u)}{uf(1-u)}\,du = -\ln c$.

2) *Для критических процессов*

$$p_{10}(t) = 1 - \frac{2}{bt}(1 + o(1)). \tag{2.16}$$

Для докритических процессов с непрерывным временем существуют пределы

$$\lim_{t\to\infty}\frac{p_{1j}(t)}{1-p_{10}(t)}=\lim_{t\to\infty}\mathsf{P}\{\xi(t)=j\mid\xi(t)>0,\ \xi(0)=1\}=Q_j\leqslant 1,\quad j\geqslant 1.$$

Вероятности $Q_1, Q_2, \ldots$ образуют распределение вероятностей, $\sum_{j=1}^{\infty}Q_j=1$, а производящая функция $Q(s)=\sum_{j=1}^{\infty}Q_js^j$ имеет вид

$$Q(s)=1-\exp\left\{a\int_0^s\frac{du}{f(u)}\right\}. \tag{2.17}$$

Если сходится интеграл $\int_0^1\frac{au+f(1-u)}{uf(1-u)}\,du=-\ln c$, то распределение с производящей функцией $Q(s)$ имеет математическое ожидание $1/c$. Для условных распределений

$$\mathsf{P}\left\{\frac{\xi(t)}{\mathsf{M}[\xi(t)\mid\xi(t)>0,\ \xi(0)=1]}>x\mid\xi(t)>0,\ \xi(0)=1\right\}$$

имеет место предельная теорема, аналогичная предельной теореме из п. 17.1.6.

Пусть $m=\mathsf{M}\xi(1)<\infty$ и $\eta(t)=e^{at}\xi(t)$, где $a=f'(1)$. Процесс $\eta(t)$ является мартингалом с непрерывным временем.

Пусть $\xi(t)$ — надкритический процесс. Из теоремы о сходимости мартингалов следует, что процесс $\eta(t)$ с вероятностью 1 при $t\to\infty$ сходится к некоторой случайной величине η.

Теорема 5. *Характеристическая функция* $\varphi(s)=\mathsf{M}e^{is\eta}$ *предельной случайной величины* η *удовлетворяет* (*нелинейному*) *дифференциальному уравнению*

$$\frac{d}{ds}\varphi(s)=\frac{f(\varphi(s))}{ds},\quad \varphi(0)=1,$$

либо эквивалентному интегральному уравнению

$$1-\varphi(s)=-is\exp\left\{\int_1^{\varphi(s)}\frac{f(u)-a(u-1)}{f(u)(u-1)}\,du\right\}.$$

При $b=f''(1)>0$ функция распределения $K(x)=\mathsf{P}\{\eta\leqslant x\}$ имеет скачок в нуле: $q=\mathsf{P}\{\eta=0\}$.

Условная функция распределения

$$\mathsf{P}\{\eta\leqslant x\mid\eta>0\}=\frac{K(x)-q}{1-q}$$

абсолютно непрерывна и имеет плотность, непрерывную при $x>0$.

17.3. Ветвящиеся процессы с конечным числом типов частиц (дискретное время)

17.3.1. Определение. Ветвящийся процесс с m ($m \geqslant 1$) типами частиц описывает популяцию частиц или индивидуумов, в которой частицы каждого типа могут порождать потомков каждого из m типов независимо от других частиц.

Фазовым пространством ветвящегося процесса, моделирующего популяцию из m типов частиц, является множество векторов $j = (j_1, j_2, \ldots, j_m)^T$, интерпретируемых как вектор-столбцы, где j_k — неотрицательные целые числа, соответствующие числу частиц k-го типа (^{T} означает транспонирование).

Пусть e_i — вектор-столбец, у которого i-я компонента равна 1, а все остальные равны 0.

Однородная цепь Маркова $\xi(t) = (\xi_1(t), \xi_2(t), \ldots, \xi_m(t))^T$ ($t = 0, 1, 2, \ldots$), значениями которой являются m-мерные векторы с неотрицательными целочисленными компонентами, называется *ветвящимся процессом с m типами частиц*, если ее переходные вероятности $p_{ij}(t) = \mathsf{P}\{\xi(t) = j \mid \xi(0) = i\}$, где $\xi(t) = j$ означает $\xi_k(t) = j_k$ ($k = 1, \ldots, m$), удовлетворяют условиям

$$p_{ij}(t) = \begin{cases} \delta_{0j}, & i = 0 \quad (0 = (0, \ldots, 0) \text{ — нулевой вектор}), \\ \left[p_{e_1 j}(t)\right]^{i_1^*} * \left[p_{e_2 j}(t)\right]^{i_2^*} * \ldots * \left[p_{e_m j}(t)\right]^{i_m^*}, & i \neq 0, \end{cases} \tag{3.1}$$

где $\left[p_{e_k j}(t)\right]^{i_k^*}$ означает i_k-кратную свертку распределения $p_{e_k j}(t)$ с собой.

Ветвящийся процесс с m типами частиц допускает простое описание в терминах сумм случайных величин. Пусть $\zeta_{tr} = \left\{\zeta_{tr}^{kl}, k, l = 1, \ldots, m\right\}$ ($t, r = 1, 2, \ldots$) — независимые одинаково распределенные случайные матрицы с независимыми неотрицательными целочисленными элементами. Для любого r случайные величины ζ_{tr}^{kl} интерпретируются как число частиц l-го типа, произведенных частицей k-го типа в момент превращения в t-м поколении. Предположим, что распределение k-й строки матрицы ζ_{tr} имеет вид

$$\mathsf{P}\left\{\zeta_{tr}^{k1} = j_1, \ldots, \zeta_{tr}^{km} = j_m\right\} = p_{e_k j}(1),$$

где $j = (j_1, j_2, \ldots, j_m)^T$. Имеют место следующие соотношения: если $\xi_k(t+1)$ — k-я компонента вектора $\xi(t+1) = (\xi_1(t+1), \xi_2(t+1), \ldots, \xi_m(t+1))$, где $\xi(t)$ — ветвящийся процесс с m типами частиц, то

$$\xi_k(t+1) = \sum_{r=1}^{\xi_1(t)} \zeta_{tr}^{1k} + \sum_{r=1}^{\xi_2(t)} \zeta_{tr}^{2k} + \ldots + \sum_{r=1}^{\xi_m(t)} \zeta_{tr}^{mk}. \tag{3.2}$$

В частности, если $m = 2$ и $\xi(0) = (1, 0)^T$, т. е. популяция состоит из частиц двух типов и в начальный момент имеется единственная частица 1-го типа, то

$$\xi(1)=\left(\zeta_1^{11},\ \zeta_1^{12}\right)^T$$

(первое поколение частиц состоит из ζ_1^{11} частиц первого типа и ζ_1^{12} частиц второго типа, порожденных частицей 1-го типа),

$$\xi(2)=\left(\sum_{r=1}^{\zeta_1^{11}}\zeta_{1r}^{11}+\sum_{r=1}^{\zeta_1^{12}}\zeta_{1r}^{21},\ \sum_{r=1}^{\zeta_1^{11}}\zeta_{1r}^{12}+\sum_{r=1}^{\zeta_1^{12}}\zeta_{1r}^{22}\right)^T$$

(второе поколение состоит из $\sum_{r=1}^{\zeta_1^{11}}\zeta_{1r}^{11}+\sum_{r=1}^{\zeta_1^{12}}\zeta_{1r}^{21}$ частиц первого типа и $\sum_{r=1}^{\zeta_1^{11}}\zeta_{1r}^{12}+\sum_{r=1}^{\zeta_1^{12}}\zeta_{1r}^{22}$ частиц второго типа) и т. д.

17.3.2. Уравнения для производящих функций. Пусть $s=(s_1, s_2, \ldots, s_m)$. Производящая функция переходных вероятностей $p_{ij}(t)$ ветвящегося процесса $\xi(t)$ с m типами частиц определяется равенством

$$\Phi_t(i,\ s)=\sum_{j\geqslant 0}p_{ij}(t)\,s_1^{j_1}s_2^{j_2}\ldots s_m^{j_m}=$$
$$=\mathsf{M}\left[s_1^{\xi_1(t)}s_2^{\xi_2(t)}\ldots s_m^{\xi_m(t)}\,\Big|\,\xi(0)=i\right],\tag{3.3}$$

где $\sum\limits_{j\geqslant 0}$ означает $\sum\limits_{j_1=0}^{\infty}\sum\limits_{j_2=0}^{\infty}\ldots\sum\limits_{j_m=0}^{\infty}$.

При фиксированном t функция $\Phi_t(i,\ s)$ есть скалярная функция векторных аргументов $i=(i_1, i_2, \ldots, i_m)^T$ и $s=(s_1, s_2, \ldots, s_m)$. Функции $\Phi_t(i,\ s)$ и $\Phi_t(e_k,\ s)$ $(k=1, \ldots, m)$ связаны соотношением

$$\Phi_t(i,\ s)=\prod_{k=1}^{m}\left[\Phi_t(e_k,\ s)\right]^{i_k}.\tag{3.4}$$

Пусть $\Phi_0(s)=\sum\limits_{i\geqslant 0}\mathsf{P}\{\xi(0)=i\}\,s_1^{i_1}s_2^{i_2}\ldots s_m^{i_m}$ — производящая функция начального распределения $p_0(i)=\mathsf{P}\{\xi(0)=i\}$; $\Phi_t(s)=$ $=\sum\limits_{j\geqslant 0}\mathsf{P}\{\xi(t)=j\}\,s_1^{j_1}s_2^{j_2}\ldots s_m^{j_m}$ — производящая функция значений процесса $\xi(t)$ в момент t, и пусть

$$\boldsymbol{\Phi}_t(s)=\left(\Phi_t(e_1,\ s),\ \Phi_t(e_2,\ s),\ \ldots,\ \Phi_t(e_m,\ s)\right)^T.\tag{3.5}$$

Функцию $\boldsymbol{\Phi}_t(s)$ называют *векторной производящей функцией* процесса $\xi(t)$. Имеет место равенство

$$\Phi_t(s)=\Phi_0(\boldsymbol{\Phi}_t(s)).\tag{3.6}$$

Теорема 1. *Производящие функции $\Phi_t(s)$ и $\boldsymbol{\Phi}_t(s)$ удовлетворяют следующим основным функциональным уравнениям*:

$$\Phi_{t+\tau}(s) = \Phi_t(\Phi_\tau(s)), \tag{3.7}$$

$$\boldsymbol{\Phi}_{t+\tau}(s) = \boldsymbol{\Phi}_t(\boldsymbol{\Phi}_\tau(s)). \tag{3.8}$$

Пусть $F(e_i, s)$ (возможно, $F(e_i, 1) < 1$) — производящая функция общего числа различных типов частиц во всех поколениях, если процесс $\xi(t)$ начался с одной частицы i-го типа, $\boldsymbol{F}(s) = (F(e_1, s), \ldots, F(e_m, s))^T$. Тогда $F(e_i, s) = s_i\Phi(e_i, \boldsymbol{F}(s))$.

17.3.3. Моменты и классификация. Обозначим через $M(t) = \{m_{ij}(t),\ i, j = 1, \ldots, m\}$ матрицу первых моментов $m_{ij}(t) = \mathsf{M}[\xi_i(t) \mid \xi(0) = e_j]$ ветвящегося процесса и через $B_k(t) = \left\{b_{ij}^{(k)}(t),\ i, j = 1, \ldots, m\right\}$ матрицу вторых моментов $b_{ij}^{(k)}(t) = \mathsf{M}\left[\xi_i(t)\,\xi_j(t) \mid \xi(0) = e_k\right]$, и пусть $\boldsymbol{M}e_k$ и $\boldsymbol{D}_k = \boldsymbol{B}_k - \boldsymbol{M}e_k e_k^T\boldsymbol{M}^T$, где $\boldsymbol{M} = \boldsymbol{M}(1)$, $\boldsymbol{B}_k = \boldsymbol{B}_k(1)$ — соответственно вектор среднего числа частиц и ковариационная матрица числа частиц, порождаемых частицей k-го типа, $t = 1$.

Из определения $\boldsymbol{\Phi}_t(s)$ следует, что

$$m_{ij}(t) = \lim_{s \uparrow 1} \frac{\partial\Phi_t(e_j, s)}{\partial s_i},$$
$$b_{ij}^{(k)}(t) = \lim_{s \uparrow 1} \frac{\partial^2\Phi_t(e_k, s)}{\partial s_i \partial s_j} + \delta_{ij} m_{jk}(t), \tag{3.9}$$

где $s \uparrow 1$ означает, что все компоненты вектора s стремятся, возрастая, к единице.

Матрицы моментов $M(t)$ и $\boldsymbol{B}_k(t)$ удовлетворяют разностным уравнениям

$$\boldsymbol{M}(t+1) = \boldsymbol{M}\boldsymbol{M}(t), \quad \boldsymbol{M}(0) = \boldsymbol{I},$$
$$\boldsymbol{B}_k(t+1) = \boldsymbol{M}\boldsymbol{B}_k(t)\boldsymbol{M}^T + \sum_{l=1}^{m}\left(e_l, \boldsymbol{M}e_k^T\right)\boldsymbol{D}_l, \quad \boldsymbol{B}_k(0) = e_k e_k^T, \tag{3.10}$$

где $(\cdot, \cdot)$ означает скалярное произведение, откуда

$$\boldsymbol{M}(t) = \boldsymbol{M}^t, \tag{3.11}$$
$$\boldsymbol{B}_k(t) = \boldsymbol{M}^t e_k e_k^T \left(\boldsymbol{M}^T\right)^t + \sum_{i=0}^{t-1}\sum_{l=1}^{m}\left(\bar{e}_l, \boldsymbol{M}^i e_k\right)\boldsymbol{M}^{t-i-1}\boldsymbol{D}_l\left(\boldsymbol{M}^T\right)^{t-i-1}.$$

Предположим, что все моменты m_{ij} матрицы $\boldsymbol{M}$ конечны и не все равны нулю. В силу того что $m_{ij} \geqslant 0$, по известной теореме Перрона — Фробениуса для неотрицательных матриц среди собственных чисел μ_i $(i = 1, \ldots, m)$ матрицы $\boldsymbol{M}$ имеется неотрицательное собственное число $\mu = \mu_{i_0}$ такое, что $\mu > \operatorname{Re}\mu_j$ $(j \neq i_0)$, называемое *перроновым корнем* матрицы $\boldsymbol{M}$.

Ветвящийся процесс $\xi(t)$ называется *неразложимым*, если геометрическая кратность перронова корня μ матрицы $\boldsymbol{M}$ равна 1, и *разложимым* — в противном случае.

Ветвящийся процесс $\xi(t)$ называется *положительно регулярным*, если найдется момент времени t_0 такой, что $m_{ij}(t_0) > 0$ для всех $i, j = 1, \ldots, m$.

Положительно регулярный процесс неразложим.

Пусть $\xi(t)$ — неразложимый ветвящийся процесс и μ — перронов корень матрицы M; u и v — соответственно правый и левый собственные векторы матрицы M, отвечающие перронову корню μ (которые по той же теореме Перрона — Фробениуса имеют неотрицательные компоненты) и нормированные условием

$$(v, u) = \sum_{i=1}^{m} v_i u_i = 1.$$

Неразложимый ветвящийся процесс $\xi(t)$ называется *докритическим*, если перронов корень $\mu < 1$; *критическим*, если перронов корень $\mu = 1$ и $b = \sum_{i, j, k=1}^{m} u_k b_{ij}^{(k)} v_i v_j > 0$; *надкритическим*, если перронов корень $\mu > 1$.

Условие $b = \sum_{i, j, k=1}^{m} u_k b_{ij}^{(k)} v_i v_j > 0$ обеспечивает несингулярность процесса $\xi(t)$, т. е. не все компоненты $\Phi(e_k, s)$ векторной производящей функции $\Phi(s)$ линейны по $s_1, s_2, \ldots, s_m$ и имеют нулевые свободные члены, и, следовательно, число частиц меняется со временем.

Говорят, что ветвящийся процесс является *периодическим* с периодом d, если наибольший общий делитель для всех t, для которых $m_{ii}(t) > 0$, равен d. Если $d = 1$, процесс называется *непериодическим*.

Положительно регулярный процесс является непериодическим.

17.3.4. Асимптотические свойства. Пусть $\xi(t)$ — неразложимый непериодический процесс; μ — перронов корень матрицы первых моментов M; u и v — соответственно правый и левый собственные векторы матрицы M, отвечающие перронову корню μ. Для матрицы $M(t)$ первых моментов имеет место асимптотическое представление при $t \to \infty$:

$$M(t) = \mu^t u v^T + o(\hat{\mu}^t),$$

где $uv^T = \{u_i v_j,\ i, j = 1, \ldots, m\}$, $|\hat{\mu}| < \mu$, $o(\hat{\mu}^t)$ имеет поэлементный смысл.

Пусть $q_i = \mathsf{P}\{\xi(t) = 0$ при некотором $t > 0 | \xi(0) = e_i\}$ есть вероятность вырождения процесса $\xi(t)$, у которого нулевое поколение состоит из единственной частицы i-го типа, и пусть $q = (q_1, q_2, \ldots, q_m)^T$ — вектор вероятностей вырождения.

Теорема 2. *Если положительно регулярный процесс $\xi(t)$ является докритическим или критическим, то*

$$q = 1 = (1, 1. 1, \ldots, 1)^T.$$

Пусть $s = (s_1, \ldots, s_m)$ и $|s| = \max_{1 \leqslant k \leqslant m} |s_k|$. Говорят, что вектор s неотрицателен, если все $s_k \geqslant 0$.

Теорема 3. *Пусть $\xi(t)$ — положительно регулярный процесс. Вектор вероятностей вырождения q является наименьшим по норме $|\cdot|$ неотрицательным решением уравнения*

$$\Phi(s) = s. \tag{3.12}$$

Пусть q^1 — произвольный неотрицательный вектор такой, что $|q^1| \leqslant 1$ и $q^1 \neq 1$. Вероятность вырождения может быть найдена как один из следующих пределов:

$$q_i = \begin{cases} \lim\limits_{t \to \infty} \mathsf{P}\{\xi(t) = 0 \mid \xi(0) = e_i\}, \\ \lim\limits_{t \to \infty} \Phi_t(e_i, q^1). \end{cases}$$

Отсюда следует, что в классе неотрицательных векторов s, $|s| \leqslant 1$, уравнение (3.12) имеет только два решения: q и 1.

Асимптотическое поведение вероятностей $p_{e_i 0}(t)$ при $t \to \infty$ описывается следующим образом.

Пусть $\xi(t)$ — положительно регулярный процесс; M — матрица математических ожиданий $m_{ij} = \mathsf{M}\,[\xi_i(1) \mid \xi(0) = e_j]$; μ — перронов корень матрицы M; $u = (u_1, \ldots, u_m)^T$ и $v = (v_1, \ldots, v_m)$ — соответственно правый и левый собственные векторы матрицы M, отвечающие перронову корню μ.

Теорема 4. 1) *Если $\xi(t)$ — докритический процесс, то*

$$\begin{aligned} &p_{e_i 0}(t) = 1 - c v_i \mu^t (1 + o(1)), \\ &\mathsf{P}\{\xi(t) \neq 0 \mid \xi(0) = i\} = (v, i)\,[c\mu^t (1 + o(1))], \end{aligned} \tag{3.13}$$

где $c = \lim\limits_{t \to \infty} \dfrac{1 - \Phi_t(e_i, 0)}{v_i \mu^t}$, причем для того, чтобы $0 < c < \infty$, необходимо и достаточно, чтобы

$$\mathsf{M}\,[\xi_j(1) \ln \xi_j(1) \mid \xi(0) = e_i] < \infty$$

для всех $i, j = 1, \ldots, m$.

2) *Если $\xi(t)$ — критический процесс, то*

$$\begin{aligned} &p_{e_i 0}(t) = 1 - \frac{2v_i}{tb}(1 + o(1)), \\ &\mathsf{P}\{\xi(t) \neq 0 \mid \xi(0) = i\} = \frac{2(v, i)}{tb}(1 + o(1)), \end{aligned} \tag{3.14}$$

где $b = \sum\limits_{i, j, k=1}^{m} u_k b_{ij}^{(k)} v_i v_j$.

17.3.5. Предельные теоремы. Пусть $\xi(t)$ — положительно регулярный процесс.

Теорема 5. *Если $\xi(t)$ — докритический процесс, то при $t \to \infty$ условные распределения*

$$\mathsf{P}\{\xi(t) = j \mid \xi(t) \neq 0, \quad \xi(0) = i\}, \quad i \neq 0,$$

сходятся к предельному распределению Q_j, $j \neq 0$, $\sum_{j\neq 0} Q_j = 1$, *производящая функция* $Q(s) = \sum_{j\neq 0} Q_j s_1^{j_1} \ldots s_m^{j_m}$ *которого удовлетворяет уравнению*

$$1 - Q(\Phi(s)) = \mu(1 - Q(s)),$$

и предельное распределение не зависит от вектора начальных состояний $i \neq 0$.

Распределение с производящей функцией $Q(s)$ имеет конечные математические ожидания

$$\lim_{s \uparrow 1} \frac{\partial Q(s)}{\partial s_i} = \frac{u_i}{c}, \quad \text{если } c > 0,$$

где c определено в соотношении (3.13).

Теорема 6. *Если* $\xi(t)$ — *положительно регулярный критический процесс,* $\xi(0) = e_i$ *и* $\zeta^{(e_i)}(t) = \left(\zeta_1^{(e_i)}(t), \ldots, \zeta_m^{(e_i)}(t)\right)$, *где*

$$\zeta_k^{(e_i)}(t) = \frac{2\xi_k(t)}{u_k bt},$$

то условное распределение процесса $\zeta^{(e_i)}(t)$ *при условии, что* $\zeta^{(e_i)}(t) \neq 0$, *сходится при* $t \to \infty$ *к распределению случайного вектора* $\zeta 1 = \zeta(1, 1, \ldots, 1)^T$, *не зависящего от* e_i, *где* ζ — *скалярная случайная величина, имеющая показательное распределение*

$$\mathsf{P}\{\zeta > x\} = e^{-x}.$$

Теорема 7. *Если* $\xi(t)$ — *положительно регулярный надкритический процесс, у которого конечны вторые моменты* $b_{ij}^{(k)}$ $(i, j, k = 1, \ldots, m)$ *и* μ — *перронов корень матрицы* M, *то случайный вектор* $\eta(t) = \mu^{-t}\xi(t)$ *в среднеквадратичном сходится при* $t \to \infty$ *к некоторому предельному случайному вектору* η *и с вероятностью* 1 *направление вектора* η *при* $\eta \neq 0$ *совпадает с направлением правого собственного вектора* u *матрицы* M, *отвечающего перронову корню* μ, *т. е.* $\eta = \zeta u$, *где* ζ — *скалярная случайная величина.*

Если $b = \sum_{i,j,k} u_k b_{ij}^{(k)} v_i v_j > 0$, то $q_k = \mathsf{P}\{\eta = 0 \mid \xi(0) = e_k\}$.

(Условная) характеристическая функция $\varphi(e_k, s) = \mathsf{M}\left[e^{i(s,\eta)} \mid \xi(0) = e_k\right]$ случайного вектора η удовлетворяет функциональным уравнениям

$$\varphi(e_k, \mu s) = \Phi(e_k, \varphi(s)),$$

где $\varphi(s) = (\varphi(e_1, s), \varphi(e_2, s), \ldots, \varphi(e_m, s))$.

17.4. Ветвящиеся процессы с конечным числом типов частиц (непрерывное время)

17.4.1. Определение. Однородная цепь Маркова $\xi(t) = (\xi_1(t), \xi_2(t), \ldots, \xi_m(t))^T$ $(t \in [0, \infty))$ со значениями во множестве m-мерных векторов с неотрицательными целочисленными компонентами называется *ветвящимся процессом с m типами частиц*, если ее переходные вероятности $p_{ij}(t) = \mathrm{P}\{\xi(t) = j \mid \xi(0) = i\}$ удовлетворяют условиям (3.1) и условию

$$\lim_{t \downarrow 0} p_{ij}(t) = \delta_{ij}.$$

Ветвящийся процесс $\xi(t)$, начальным состоянием которого является состояние e_i, т. е. нулевое поколение частиц состоит из единственной частицы i-го типа, эволюционирует следующим образом: по прошествии случайного времени τ_i частица i-го типа превращается в случайное число ζ_1^{ij} частиц j-го типа, $j = 1, \ldots, m$, каждая из которых независимо от других живет случайное время τ_j и превращается в случайное число ζ_2^{jk} частиц k-го типа, $k = 1, \ldots, m$, и т. д.

17.4.2. Уравнение для производящих функций. Пусть $\xi(0) = e_i$ и переходные вероятности ветвящегося процесса $\xi(t)$ удовлетворяют условиям

$$\begin{aligned} p_{e_ie_i}(t) &= 1 + q_{e_ie_i}t + o(t), \\ p_{e_ij}(t) &= q_{e_ij}t + o(t), \quad e_i \neq j, \quad t \to 0, \end{aligned} \tag{4.1}$$

$$\sum_{j \geqslant 0} q_{e_ij} = 0, \quad i = 1, \ldots, m. \tag{4.2}$$

Ясно, что $q_{e_ij} \geqslant 0$, $e_i \neq j$ (в этом случае q_{e_ij} называют *плотностью вероятности перехода* из e_i в j). Пусть

$$f(e_i, s) = \sum_{j \geqslant 0} q_{e_ij} s_1^{j_1} s_2^{j_2} \ldots s_m^{j_m},$$

$$f(s) = (f(e_1, s), \ldots, f(e_m, s))^T,$$

$$\Phi_t(e_i, s) = \sum_{j \geqslant 0} p_{e_ij}(t) s_1^{j_1} s_2^{j_2} \ldots s_m^{j_m} = \mathrm{M}\left[s_1^{\xi_1(t)} \ldots s_m^{\xi_m(t)} \,\middle|\, \xi(0) = e_i\right],$$

$$\Phi_t(s) = (\Phi_t(e_1, s), \ldots, \Phi_t(e_m, s))^T.$$

Функция $f(s)$ называется (векторной) *инфинитезимальной* или *дифференциальной производящей функцией*.

Производящая функция $\Phi_t(s)$ непрерывна по $t \in [0, \infty)$ равномерно по s, $|s| \leqslant 1$ и $\lim_{t \downarrow 0} \Phi_t(s) = s$. При выполнении условий (4.1), (4.2) равномерно по s, $|s| \leqslant 1$, имеет место асимптотическое представление

$$\Phi_t(s) = s + tf(s) + o(t), \quad t \to 0. \tag{4.3}$$

Следствием (4.3) является следующая теорема, дающая аналог основного функционального уравнения для процессов с непрерывным временем.

Теорема 1. *Производящая функция* $\Phi_t(s)$ *при* $|s| \leqslant 1$ *удовлетворяет следующим системам уравнений:*

а) *системе (нелинейных) обыкновенных дифференциальных уравнений*

$$\frac{d\Phi_t(s)}{dt} = f(\Phi_t(s)) \tag{4.4}$$

с начальными условиями

$$\Phi_0(s) = s; \tag{4.5}$$

б) *системе (линейных) уравнений в частных производных*

$$\frac{\partial \Phi_t(s)}{\partial t} = \sum_{i=1}^{m} f(e_i, s) \frac{\partial \Phi_t(s)}{\partial s_i} \tag{4.6}$$

с начальными условиями (4.5);

в) *системе (нелинейных) интегральных уравнений*

$$\Phi_t(e_i, s) = \int_0^t h_i(\Phi_{t-u}(s))\, dG_i(u) + s_i(1 - G_i(t)), \tag{4.7}$$

где

$$h_i(s) = \frac{f(e_i, s) - q_{e_i e_i} s_i}{-q_{e_i e_i}}, \quad i = 1, \ldots, m,$$

$$G_i(t) = \begin{cases} 1 - e^{q_{e_i e_i} t}, & t \geqslant 0, \\ 0, & t < 0. \end{cases}$$

Решения этих уравнений существуют, являются аналитическими функциями по s, $|s| < 1$, но $\Phi_t(e_i, s)$, вообще говоря, не обязаны быть производящими функциями собственных вероятностных распределений, т. е. в общем случае можно лишь утверждать, что

$$\lim_{s \uparrow 1} \Phi_t(e_i, s) \leqslant 1.$$

Если все производные $\dfrac{\partial f(e_i, s)}{\partial s_j}$ $(i, j = 1, \ldots, m)$ в точке $s = 1$ конечны, то решение указанных систем уравнений единственно при $|s| \leqslant 1$ и $\lim\limits_{s \uparrow 1} \Phi_t(e_i, s) = 1$ $(i = 1, \ldots, m)$.

Функция $G_i(t)$ в (4.8) интерпретируется как функция распределения времени жизни частицы i-го типа, а $h_i(s)$ интерпретируется как производящая функция плотностей превращения частицы i-го типа.

17.4.3. Моменты и классификация. Пусть $f(s) = (f(e_1, s), f(e_2, s), \ldots, f(e_m, s))^T$ — дифференциальная производящая функция ветвящегося процесса $\xi(t)$. В предположении, что существуют соответствующие пределы, примем обозначения

$$a_{ij} = \lim_{s \uparrow 1} \frac{\partial f(e_i, s)}{\partial s_j}, \quad c_{ij}^{(k)} = \lim_{s \uparrow 1} \frac{\partial^2 f(e_k, s)}{\partial s_i \partial s_j},$$
$$A = \{a_{ij},\ i, j = 1, \dots, m\}, \quad C_k = \{c_{ij}^{(k)},\ i, j = 1, \dots, m\}.$$

Пусть также $M(t) = \{m_{ij}(t),\ i, j = 1, \dots, m\}$, где $m_{ij}(t) = \mathbf{M}[\xi_i(t) \mid \xi(0) = e_j]$, $C_k(t) = \{c_{ij}^{(k)}(t), i, j = 1, \dots, m\}$, где $c_{ij}^{(k)}(t) = \mathbf{M}[\xi_j(t)\xi_i(t) \mid \xi(0) = e_k] - \delta_{ij} m_{jk}(t)$. Матрицы моментов $M(t)$ и $C_k(t)$ удовлетворяют дифференциальным уравнениям

$$\begin{gathered} \frac{d}{dt} M(t) = M(t) A, \quad M(0) = I, \\ \frac{d}{dt} C_k(t) = \sum (e_k, Ae_j) C_j(t) + M(t) C_k M^T(t), \quad C_k(0) = 0. \end{gathered} \tag{4.8}$$

Отсюда

$$M(t) = e^{At},$$
$$C_k(t) = \int_0^t M(u) \left[\sum_{j=1}^m (e_k, M(t-u) e_j) C_j \right] M^T(u)\, du. \tag{4.9}$$

Все недиагональные элементы матрицы A неотрицательны (в частности, положительны). Матрицы, обладающие этим свойством, называются *квазинеотрицательными* (соответственно *квазиположительными*). Спектральные свойства таких матриц сходны со спектральными свойствами неотрицательных (соответственно положительных) матриц. Например, среди всех собственных чисел α_i $(i = 1, \dots, m)$ матрицы A имеется действительное собственное число $\alpha = \alpha_{i_0}$ такое, что $\alpha > \operatorname{Re} \alpha_j$ $(j \neq i_0)$, называемое *перроновым корнем* матрицы A. Собственные векторы, отвечающие перронову корню α, имеют неотрицательные компоненты.

Ветвящийся процесс $\xi(t)$ называется *неразложимым*, если геометрическая кратность перронова корня α матрицы A равна 1, и *разложимым* — в противном случае.

Ветвящийся процесс $\xi(t)$ называется *регулярным*, если $a_{ii} < 0$ для всех $i = 1, \dots, m$.

Пусть $\xi(t)$ — неразложимый процесс, $u = (u_1, \dots, u_m)^T$ и $v = (v_1, \dots, v_m)$ — соответственно правый и левый собственные векторы матрицы A, отвечающие перронову корню α, нормированные условием $(u, v) = \sum_{i=1}^m u_i v_i = 1$.

Неразложимый ветвящийся процесс $\xi(t)$ называется *докритическим*, если $\alpha < 0$; *критическим*, если $\alpha = 0$, $b = \sum_{i,j,k=1}^m u_k c_{ij}^{(k)} v_i v_j > 0$; *надкритическим*, если $\alpha > 0$.

Условие $b = \sum_{i,\,j,\,k=1}^{m} u_k c_{ij}^{(k)} v_i v_j > 0$ обеспечивает несингулярность процесса $\xi(t)$. Последнее означает, что число частиц не остается неизменным во времени.

17.4.4. Асимптотические свойства. Пусть $\xi(t)$ — неразложимый ветвящийся процесс; α — перронов корень матрицы A; u и v — соответственно правый и левый собственные векторы матрицы A, отвечающие перронову корню α и нормированные условием $(u, v) = 1$. Для матрицы $M(t)$ первых моментов имеет место асимптотическое представление при больших t:

$$M(t) = e^{\alpha t} u v^T + o\left(e^{\hat{\alpha} t}\right),$$

где $uv^T = \left\{u_i v_j,\ i,\ j = 1, \ldots, m\right\}$, $|\hat{\alpha}| < \alpha$, $o\left(e^{\hat{\alpha} t}\right)$ понимается поэлементно.

Пусть q_i — вероятность вырождения процесса $\xi(t)$, у которого $\xi(0) = e_i$, и $q = (q_1, q_2, \ldots, q_m)^T$ — вектор вероятностей вырождения.

Теорема 2. *Если регулярный ветвящийся процесс $\xi(t)$ является докритическим или критическим, то*

$$q = 1 = (1, 1, \ldots, 1)^T.$$

Теорема 3. *Пусть $\xi(t)$ — неразложимый процесс. Вектор вероятностей вырождения q является ближайшим к 0 неотрицательным решением уравнения*

$$f(s) = 0, \quad s \geqslant 0, \quad |s| = \max_{1 \leqslant i \leqslant m} |s_i| \leqslant 1.$$

Асимптотическое поведение $p_{ij}(t)$ при $t \to \infty$ описывается следующим образом.

Теорема 4. 1) *Если неразложимый процесс $\xi(t)$ является докритическим, то*

$$e^{-\alpha t}\left(1 - p_{e_i 0}(t)\right) = c v_i + o(1),$$

$$e^{-\alpha t} \mathsf{P}\{\xi(t) \neq 0 \mid \xi(0) = i\} = (v, i)\, c + o(1),$$

где $c \geqslant 0$ — некоторая константа, отличная от нуля тогда и только тогда, когда $\mathsf{M}\left[\xi_j(t) \ln \xi_j(t) \mid \xi(0) = e_i\right] < \infty$ для всех $i, j = 1, \ldots, m$.

2) *Если неразложимый процесс $\xi(t)$ является критическим и конечны $c_{ij}^{(k)}$, то*

$$p_{e_i 0}(t) = 1 - \frac{2 v_i}{bt}(1 + o(1)),$$

$$\mathsf{P}\{\xi(t) \neq 0 \mid \xi(0) = i\} = \frac{2(v, i)}{bt}(1 + o(1)),$$

где $b = \sum_{i,\,j,\,k=1}^{m} u_k c_{ij}^{(k)} v_i v_j$.

17.5. Общие марковские ветвящиеся процессы

17.5.1. Определения. Общая модель марковского ветвящегося процесса учитывает наряду с числом частиц моделируемой популяции также такие характеристики, как положение частиц в пространстве, их размер, массу, энергию, возраст и т. п.

Важно отметить, что многие немарковские ветвящиеся процессы, описывающие число частиц в популяции, при привлечении дополнительной информации о частицах типа указанной выше могут быть изучены в рамках общих марковских ветвящихся процессов.

Предположим, что популяция характеризуется числом частиц и некоторым обобщенным случайным параметром η, интерпретируемым как положение частиц в определенном измеримом пространстве $(\mathfrak{X}, \mathfrak{A})$, называемом *фазовым пространством частиц*.

Пусть нулевое поколение популяции состоит из единственной частицы и ее положение в $\mathfrak{X}$ (например, масса или энергия) равно η_0. По истечении случайного времени τ частица превращается (например, дробится либо порождает новые частицы, отдавая им свою энергию) в случайное число ζ_1 частиц первого поколения, положения которых в $\mathfrak{X}$ равны соответственно $\eta_1', \eta_2', \ldots, \eta_{\zeta_1}'$, где $\eta_k' \in \mathfrak{X}$ $(k = 1, \ldots, \zeta_1)$ — одинаково распределенные независимые между собой и от ζ_1 случайные величины. Каждая частица первого поколения независимо от других ведет себя так же, как частица нулевого поколения, и т. д. Фазовое пространство ветвящегося процесса, моделирующего описанную схему, должно, очевидно, учитывать как число частиц в популяции в произвольный момент, так и их положение в фазовом пространстве $(\mathfrak{X}, \mathfrak{A})$.

Пусть в некоторый момент в популяции находится n частиц и их положения в $\mathfrak{X}$ равны соответственно $x_1, x_2, \ldots, x_n$. Состояние ветвящегося процесса можно описать набором $(x_1, x_2, \ldots, x_n)$, в котором порядок расположения несуществен, что соответствует неразличимости частиц в популяции.

Если $\mathfrak{X}^n$ есть n-кратное декартово (прямое) произведение пространства $\mathfrak{X}$ самого на себя, то обозначим через $\widetilde{\mathfrak{X}}_n$ пространство, получаемое из $\mathfrak{X}^n$ отождествлением всех точек $x^n = (x_1, x_2, \ldots, x_n)$, которые могут быть получены перестановками координат, и через $\widetilde{\mathfrak{A}}_n$ обозначим образ σ-алгебры $\mathfrak{A}^n$ при таком отображении.

Пусть $\widetilde{\mathfrak{X}}_0$ означает пространство, состоящее из единственной точки, обозначаемой тем же символом $\widetilde{\mathfrak{X}}_0$. Положим

$$\widetilde{\mathfrak{X}} = \bigcup_{n=0}^{\infty} \widetilde{\mathfrak{X}}_n,$$

и пусть $\widetilde{\mathfrak{A}}$ — наименьшая σ-алгебра, содержащая $\widetilde{\mathfrak{X}}_0$ и все σ-алгебры $\widetilde{\mathfrak{A}}_n$.

Общим марковским ветвящимся процессом с фазовым пространством частиц $(\mathfrak{X}, \mathfrak{A})$ называется однородный марковский процесс $\xi(t)$ $(t \in T,\ T = [0, \infty)$ либо $T = 0, 1, 2, \ldots)$ в фазовом пространстве $(\widetilde{\mathfrak{X}}, \widetilde{\mathfrak{A}})$, переходные вероятности

$$P_t\left(x^n, \widetilde{A}\right) = \mathsf{P}\left\{\xi(t) \in \widetilde{A} \mid \xi(0) = x^n\right\},$$
$$x^n \in \widetilde{\mathfrak{X}}_n,\ \widetilde{A} \in \widetilde{\mathfrak{A}}$$

которого удовлетворяют следующему уравнению Колмогорова:

$$P_{t+s}\left(x, \tilde{A}\right)=\sum_{n=0}^{\infty}\int_{\tilde{\mathfrak{X}}_n} P_t^{(n)}\left(x^n, dy^n\right) P_s\left(y^n, \tilde{A}\right), \quad x \in \tilde{\mathfrak{X}}, \tag{5.1}$$

где $P_t^{(n)}\left(x^n, \cdot\right)$ есть сужение меры $P_t\left(x^n, \cdot\right)$ на σ-алгебру $\tilde{\mathfrak{A}}_n$, и

$$P_0\left(x^n, \tilde{A}\right)=\chi_{\tilde{A}}\left(x^n\right)=\begin{cases}1, & x^n \in \tilde{A},\\ 0, & x^n \notin \tilde{A}.\end{cases} \tag{5.2}$$

$P_t^{(n)}\left(x^n, \tilde{\mathfrak{X}}_n\right)$ есть вероятность того, что в момент t в популяции имеется ровно n частиц при условии, что в начальный момент имелось n частиц и их положения в $\tilde{\mathfrak{X}}$ были $x_1, x_2, \ldots, x_n$.

17.5.2. Примеры. 1. Пусть $\mathfrak{X}$ — конечное множество и $x \in \mathfrak{X}$ интерпретируется как тип частицы. Соответствующий ветвящийся процесс является обычным ветвящимся процессом с конечным числом типов частиц.

2. Пусть $\mathfrak{X} = [0, \infty)$ и $x \in \mathfrak{X}$ означает возраст частицы, который меняется так, что $\Delta x = \Delta t$. Время жизни частицы задается некоторой функцией распределения $G(x)$. Каждая частица независимо от других порождает случайное число частиц нулевого возраста. Соответствующие этой модели *ветвящиеся процессы, зависящие от возраста*, описывают некоторые фазы эволюции колоний бактерий или других организмов.

3. *Одномерная модель ядерного реактора.* Пусть на отрезке $[a, b]$ (активная зона реактора) в обоих направлениях могут двигаться нейтроны, которые по достижении концов отрезка исчезают (уходят из активной зоны).

Положим $\mathfrak{X} = [a, b]$ и $x \in \mathfrak{X}$ есть положение нейтрона в момент рождения. Нейтрон, рожденный в точке x, с вероятностью 1/2 движется вправо или влево. В любом интервале длины dx из $\mathfrak{X}$ нейтрон с вероятностью αdx обращается в некоторое число новых нейтронов, каждый из которых независимо от других с вероятностью 1/2 движется вправо или влево. Эта модель описывается общим ветвящимся процессом и является наряду с ее двух- и трехмерными аналогами исходной в математической теории ядерных реакторов.

17.5.3. Уравнения для производящих функционалов. С ветвящимся процессом $\xi(t)$ $(t \in T)$ связаны следующие случайные меры $\xi_{x^n}(t, \cdot)$ и $\eta_x(\cdot)$ на $(\mathfrak{X}, \mathfrak{A})$ с целочисленными неотрицательными значениями: $\xi_{x^n}(t, A)$ — число частиц процесса $\xi(t)$, оказавшихся в момент t в множестве $A \in \mathfrak{A}$ при условии, что в начальный момент было n частиц и их положения в $\mathfrak{X}$ определялись точкой $x^n \in \tilde{\mathfrak{X}}_n$; $\eta_x(A)$ — число частиц-потомков в множестве $A \in \mathfrak{A}$ в момент превращения, если частица-предок в момент превращения находилась в точке $x \in \mathfrak{X}$.

Пусть $s(x)$ — $\mathfrak{A}$-измеримая функция такая, что $\sup\limits_{x \in \mathfrak{X}} |s(x)| \leqslant 1$,

$$\Phi_t\left(x^n, s(\cdot)\right) = \mathsf{M} \exp\left\{\int_{\mathfrak{X}} \ln s(y)\, \xi_{x^n}(t, dy)\right\} \tag{5.3}$$

— *производящий функционал случайной меры* $\xi_{x^n}(t, \cdot)$;

$$h(x, s(\cdot)) = \mathsf{M} \exp \left\{ \int_{\mathfrak{X}} \ln s(y)\, \eta_x(dy) \right\} \tag{5.4}$$

— *производящий функционал случайной меры* $\eta_x(\cdot)$.

Если $x^n = (x_1, x_2, \ldots, x_n)$, то

$$\Phi_t(x^n, s(\cdot)) = \prod_{k=1}^{n} \Phi_t(x_k, s(\cdot)). \tag{5.5}$$

Пусть $q_t(x, A)$ есть вероятность того, что частица, начавшая блуждание из точки $x \in \mathfrak{X}$, за время $[0, t]$ не испытает превращения и в момент t будет находиться в множестве $A \in \mathfrak{A}$; $K_x(t, A)$ — условная вероятность того, что время жизни частицы, находившейся в начальный момент в точке x, не превышает t, а точка, в которой находится эта частица в момент превращения, содержится в $A \in \mathfrak{A}$.

Теорема 1. *Производящий функционал* $\Phi_t(x, s(\cdot))$ $(x = x^1)$ *удовлетворяет следующим функциональным уравнениям*:

$$\Phi_{t+\tau}(x, s(\cdot)) = \Phi_t(x, \Phi_\tau(\cdot, s(\cdot))), \quad \Phi_{t+1}(x, s(\cdot)) = \Phi_t(x, h(\cdot, s(\cdot))),$$

$$\Phi_t(x, s(\cdot)) = \int_{\mathfrak{X}} s(y)\, q_x(t, dy) + \int_0^t K_x(du, dy)\, h(y, \Phi_{t-u}(\cdot, s(\cdot))),$$

Пусть $M(t, x^n, A) = \mathsf{M}\xi_{x^n}(t, A)$ — среднее число частиц, оказавшихся в момент t в множестве $A \in \mathfrak{A}$ при условии, что $\xi(0) = x^n$. $M(t, x^n, A)$ удовлетворяет уравнению

$$M(t + \tau, x^n, A) = \int_{\mathfrak{X}} M(t, y, A)\, M(\tau, x^n, dy). \tag{5.6}$$

Пример. Пусть $\xi(t)$ $(t \in [0, \infty))$ — ветвящийся процесс, у которого частицы не меняют своего положения между превращениями (скачкообразный ветвящийся процесс), причем

$$q_t(x, \mathfrak{X}) = e^{-qt}, \quad q > 0.$$

Параметр q называется *интенсивностью скачков частиц*.

Производящий функционал $\Phi_t(x, s(\cdot))$ такого процесса удовлетворяет дифференциальному уравнению

$$\frac{\partial}{\partial t} \Phi_t(x, s(\cdot)) + q\Phi_t(x, s(\cdot)) = qh(x, \Phi_t(\cdot, s(\cdot))),$$

а математическое ожидание $M(t, x, A)$ имеет вид

$$M(t, x, A) = \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} e^{-qt} L^{(n)}(x, A),$$

где

$$L^{(0)}(x, A) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

$$L^{(1)}(x, A) = \mathsf{M}\eta_x(A),$$

$$L^{(n)}(x, A) = \int_{\mathfrak{X}} L^{(1)}(x, dy)\, L^{(n-1)}(y, A).$$

17.5.4. Вероятность вырождения. Для общего ветвящегося процесса $\xi(t)$, $\xi(0) = x$, вероятность вырождения $q(x)$ определяется как любой из пределов:

$$q(x) = \begin{cases} \lim\limits_{t \to \infty} P_t^{(0)}(x, \mathfrak{X}), \\ \lim\limits_{t \to \infty} \Phi_t(x, 0). \end{cases}$$

$q(x)$ удовлетворяет функциональному уравнению

$$q(x) = h(x, q(\cdot)). \tag{5.7}$$

Если функция $s_0(\cdot)$ такова, что $0 \leqslant s_0(x) \leqslant 1$ $(x \in \mathfrak{X})$ удовлетворяет условию $h(x, s_0(\cdot)) \leqslant s_0(x)$ для всех $x \in \widetilde{\mathfrak{X}}$, то $q(x) \leqslant s_0(x)$.

Пусть $t = 0, 1, 2, \ldots$, $M(x) = \mathsf{M}\xi_x(1, \mathfrak{X})$, $B(x) = \mathsf{M}\xi_x^2(1, \mathfrak{X}) - M(x) = \mathsf{M}\xi_x(1, \mathfrak{X})[\xi_x(1, \mathfrak{X}) - 1]$.

Теорема 2. 1) *Если* $\sup\limits_{x \in \mathfrak{X}} M(x) < 1$, *то* $q(x) \equiv 1$.

2) *Если* $\inf\limits_{x \in \mathfrak{X}} M(x) > 1$ *и* $\sup\limits_{x \in \mathfrak{X}} B(x) < \infty$, *то* $\sup\limits_{x \in \mathfrak{X}} q(x) < 1$.

Литература: [20, 77, 91, 106, 115].

Глава 18. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

18.1. Слабая сходимость мер в метрических пространствах

18.1.1. Сходимость на множествах непрерывности предельной меры. Пусть $\{\mathscr{X}, \mathfrak{B}, \rho\}$ — метрическое пространство с борелевской σ-алгеброй $\mathfrak{B}$ и метрикой $\rho(x, y)$; $C(\mathscr{X})$ — пространство всех вещественных непрерывных и ограниченных функций, определенных на $\mathscr{X}$, с нормой $\|f\| = \sup\limits_{x \in \mathscr{X}} |f(x)|$.

Последовательность мер *) μ_n, определенных на $\mathfrak{B}$, называется *слабо сходящейся* к мере μ (обозначение: $\mu_n \Rightarrow \mu$), если выполняется соотношение

$$\lim_{n \to \infty} \int f(x)\, \mu_n(dx) = \int f(x)\, \mu(dx) \tag{1.1}$$

для всех $f \in C(\mathscr{X})$.

*) *Меры* μ_n, вообще говоря, не нормированы.

Так как значения интегралов $\int f(x)\,\mu(dx)$ для всех $f \in C(\mathscr{X})$ однозначно определяют меру μ, то из слабой сходимости $\mu_n \Rightarrow \mu$ и $\mu_n \Rightarrow \nu$ следует, что $\mu = \nu$.

Из определения слабой сходимости мер следует, что из сходимости по вероятности случайных элементов ζ_n со значениями в $\mathscr{X}$ к случайному элементу ζ следует слабая сходимость распределений P_n случайных элементов ζ_n к распределению P предельного случайного элемента ζ. Обратное утверждение не имеет места, за исключением случая, когда предельное распределение P сосредоточено в одной точке.

Введем обозначения: $\operatorname{Int} A$ — множество внутренних точек A; $[A]$ — замыкание множества A; A' — множество граничных точек множества A.

Лемма. *Если $\mu_n \Rightarrow \mu$, то для любого $A \in \mathfrak{B}$ имеют место неравенства*

$$\mu(\operatorname{Int} A) \leqslant \varliminf_{n\to\infty} \mu_n(A) \leqslant \varlimsup_{n\to\infty} \mu_n(A) \leqslant \mu([A]). \tag{1.2}$$

Множество A называется *множеством непрерывности меры* μ, если $\mu(A') = 0$.

Совокупность всех множеств непрерывности меры μ обозначим через $\mathfrak{A}_\mu$.

Теорема 1. *Для того чтобы последовательность мер μ_n слабо сходилась к мере μ, необходимо и достаточно, чтобы*

$$\lim_{n\to\infty} \mu_n(A) = \mu(A) \quad \text{для всех} \quad A \in \mathfrak{A}_\mu. \tag{1.3}$$

18.1.2. Условие слабой компактности семейства мер. Множество M мер, определенных на $\mathfrak{B}$, называется *слабо компактным*, если из всякой последовательности мер μ_n из M можно выбрать слабо сходящуюся последовательность.

Теорема Прохорова. *Пусть $\mathscr{X}$ — полное сепарабельное метрическое пространство. Для того чтобы множество M мер, определенных на $\mathfrak{B}$, было слабо компактным, необходимо и достаточно, чтобы выполнялись следующие два условия:*

а) $$\sup_{\mu\in M} \mu(\mathscr{X}) < \infty; \tag{1.4}$$

б) *для любого $\varepsilon > 0$ существует такой компакт K_ε, что*

$$\sup_{\mu\in M} \mu(\mathscr{X} \setminus K_\varepsilon) < \varepsilon. \tag{1.5}$$

Семейство мер M, для которого выполнены условия а), б), называется *плотным*.

Замечание. Полнота пространства $\mathscr{X}$ используется только при доказательстве необходимости условий а) и б) теоремы Прохорова.

При доказательстве слабой сходимости последовательности мер устанавливаются слабая компактность последовательности мер и единственность предельной меры.

Теорема 2. *Если последовательность мер μ_n, определенных на $\mathfrak{B}$ — σ-алгебре борелевских множеств полного сепарабельного метри-*

ческого пространства $\mathscr{X}$, такова, что для всех $f \in C(\mathscr{X})$ существует предел

$$L(f) = \lim_{n \to \infty} \int f(x)\,\mu_n(dx), \tag{1.6}$$

то существует такая мера μ, что

$$L(f) = \int f(x)\,\mu(dx),$$

т. е. последовательность мер μ_n слабо сходится к μ.

18.1.3. Условия слабой сходимости последовательности мер. Последовательность функций $f_n \in C(\mathscr{X})$ слабо сходится к f, если функции f_n ограничены в совокупности и для всех $x \in \mathscr{X}$ существует предел $\lim\limits_{n \to \infty} f_n(x) = f(x)$.

Множество функций $F \subset C(\mathscr{X})$ называется *слабо замкнутым*, если предел любой слабо сходящейся последовательности функций из F принадлежит F.

Теорема 3. *Последовательность мер μ_n слабо сходится к мере μ тогда и только тогда, когда она слабо компактна и для некоторого множества функций $F_0 \subset C(\mathscr{X})$, слабое замыкание которого совпадает со всем $C(\mathscr{X})$, выполняется соотношение*

$$\lim_{n \to \infty} \int f(x)\,\mu_n(dx) = \int f(x)\,\mu(dx), \quad f \in F_0. \tag{1.7}$$

Замечание. Множество функций $F \subset C(\mathscr{X})$ называется *тотальным*, если из равенства $\int f(x)\,\mu_1(dx) = \int f(x)\,\mu_2(dx)$ для всех $f \in F$ следует $\mu_1 = \mu_2$. Теорема 3 остается справедливой, если F_0 — тотальное множество.

При доказательстве предельных теорем для случайных процессов удобно применять условия сходимости частных распределений.

Теорема 4. *Пусть $\mathfrak{A}_0$ — класс открытых множеств в $\mathscr{X}$, содержащий вместе с двумя множествами их сумму и пересечение и удовлетворяющий условиям:* 1) *σ-алгебра, порожденная классом $\mathfrak{A}_0$, содержит все открытые множества;* 2) *все множества из $\mathfrak{A}_0$ являются множествами непрерывности данной меры μ.*

Если для слабо компактной последовательности мер μ_n выполняется условие

$$\lim_{n \to \infty} \mu_n(A) = \mu(A), \quad A \in \mathfrak{A}_0, \tag{1.8}$$

то μ_n слабо сходится к μ.

Замечание. В различных функциональных пространствах в качестве класса $\mathfrak{A}_0$ обычно рассматривается класс всех открытых цилиндрических множеств непрерывности предельной меры, а значит, применяются условия сходимости конечномерных распределений.

При слабой сходимости мер имеет место сходимость интегралов и для некоторых разрывных функций. При этом используется то обстоятельство, что множество точек разрыва $\mathfrak{B}$-измеримой функции является $\mathfrak{B}$-измеримым множеством.

Лемма. *Если последовательность мер μ_n слабо сходится к μ, то*

$$\lim_{n\to\infty}\int f(x)\,\mu_n(dx)=\int f(x)\,\mu(dx) \tag{1.9}$$

для всякой $\mathfrak{B}$-измеримой μ-почти всюду непрерывной и ограниченной функции f.

18.1.4. Сходимость мер в линейных нормированных пространствах. В линейных нормированных пространствах условия слабой сходимости мер можно сформулировать в виде условий сходимости характеристических функционалов.

Пусть $\{\mathscr{X}, \mathfrak{B}\}$ — сепарабельное банахово пространство с σ-алгеброй борелевских множеств, L — линейное множество линейных функционалов на $\mathscr{X}$ такое, что минимальная σ-алгебра, относительно которой измеримы все функционалы $l\in L$, совпадает с $\mathfrak{B}$.

Теорема 5. *Последовательность мер μ_n на $\{\mathscr{X}, \mathfrak{B}\}$ слабо сходится к мере μ тогда и только тогда, когда она слабо компактна и выполняется соотношение*

$$\lim_{n\to\infty}\int e^{il(x)}\mu_n(dx)=\int e^{il(x)}\mu(dx),\quad l\in L. \tag{1.10}$$

18.2. Слабая сходимость мер в гильбертовом пространстве

Случайные величины и вероятностные распределения в гильбертовом пространстве могут быть использованы при изучении случайных процессов и полей, интегрируемых с квадратом в своей области определения по некоторой мере. В частности, слабая сходимость мер в гильбертовом пространстве может быть использована для получения предельных теорем для интегрируемых с квадратом процессов и полей.

18.2.1. Случайные величины в гильбертовом пространстве и их распределения. В дальнейшем через X обозначается сепарабельное гильбертово пространство, а через $\mathfrak{B}$ — σ-алгебра его борелевских подмножеств. *Случайной величиной* со значениями в X называется измеримое отображение $\xi=\xi(\omega)$ некоторого вероятностного пространства $\{\Omega, \mathfrak{F}, \mathsf{P}\}$ в $(X, \mathfrak{B})$. Мера $\mu(A)=\mathsf{P}\{\xi^{-1}(A)\}=\mathsf{P}\{\omega;\ \xi(\omega)\in A\}$ называется *распределением величины* ξ. Обозначим через Q_L оператор проектирования на пространство $L\subset X$. С каждой мерой μ на $\mathfrak{B}$ можно связать семейство мер $\{\mu_L,\ L$ — конечномерное подпространство $X\}$, где $\mu_L(C)=\mu\left(Q_L^{-1}C\right)$, $C\subset L$, $C\in\mathfrak{B}$. Это семейство согласовано: при $L\subset N$ (L, N — конечномерные пространства)

$$\mu_N\left(Q_L^{-1}C\cap N\right)=\mu_L(C). \tag{2.1}$$

Совокупности семейств мер $\{\mu_L,\ L$ — конечномерное подпространство $X\}$, где μ_L — борелевская мера на L, удовлетворяющая (2.1), называются *конечномерными распределениями.* По конечномерным распределениям можно построить конечно-аддитивную функцию множеств $\tilde{\mu}$, заданную на алгебре $\mathfrak{B}_0$ множеств вида $Q_L^{-1}C$, где C — борелевское множество в L:

$$\tilde{\mu}\left(Q_L^{-1}C\right)=\mu_L(C).$$

Основной вопрос в теории меры в гильбертовых пространствах следующий: когда такая функция продолжается с $\mathfrak{B}_0$ на $\mathfrak{B}$ как счетно-аддитивная? Ниже дается ответ на этот вопрос. Конечномерные распределения могут быть заданы своим преобразованием Фурье или характеристическим функционалом

$$\chi(z) = \int e^{i|z|t}\mu_{L_z}(dt), \quad z \in X,$$

где L_z — одномерное пространство, порожденное вектором z; $\chi(z)$ — непрерывная и положительно определенная функция (для любых $z_1, \ldots, z_n \in X$ и комплексных $\alpha_1, \ldots, \alpha_n$

$$\sum \chi(z_k - z_l)\,\alpha_k\bar{\alpha}_l \geqslant 0).$$

Обозначим через $L_1^+(X)$ множество симметрических неотрицательных ядерных операторов (т. е. таких операторов V, для которых $\sum\limits_k (Ve_k, e_k) < \infty$ для некоторого ортонормированного базиса $\{e_k\}$ в X).

Теорема (Минлоса — Сазонова). *Для того чтобы непрерывный положительно определенный функционал $\chi(z)$ был характеристическим функционалом конечной меры μ на $\mathfrak{B}$ $\left(\text{т. е. } \chi(z) = \int e^{i(z,x)}\mu(dx)\right)$, необходимо и достаточно, чтобы для всякого $\varepsilon > 0$ можно было указать такой оператор $S \in L_1^+(X)$, что $\operatorname{Re}(\chi(0) - \chi(z)) \leqslant \varepsilon$, как только $(Sz, z) \leqslant 1$.*

Пусть ξ — величина в X. Тогда (ξ, z) — числовая случайная величина для всех z. Если μ — ее распределение, то $\mathsf{M}(\xi, z) = \int (x, z)\,\mu(dx)$. Если эта величина определена для всех $z \in X$, то она имеет вид (a, z), где $a \in X$ — элемент, называемый *средним значением* величины ξ (или распределения μ); дисперсия (ξ, z)

$$\mathsf{D}(\xi, z) = \mathsf{M}(\xi, z)^2 - (\mathsf{M}(\xi, z))^2 = \int (x - a, z)^2\,\mu(dx).$$

Если она конечна для всех z, то величина справа есть неотрицательный квадратичный функционал от z; он представим в виде (Bz, z), где B — симметрический неотрицательный оператор в X. Оператор B называется *корреляционным оператором* величины ξ (или распределения μ).

Случайная величина ξ в X называется *гауссовской*, если (ξ, z) имеет нормальное распределение для всех z. Характеристический функционал гауссовской величины ξ определяется равенством

$$\chi(z) = \mathsf{M}e^{i(\xi, z)} = \exp\left\{i(a, z) - \frac{1}{2}(Bz, z)\right\},$$

где a — среднее значение ξ, B — корреляционный оператор ξ; a — произвольный элемент X, $B \in L_1^+(X)$.

18.2.2. Условия компактности семейства мер в гильбертовом пространстве. Пусть T_C — множество всех симметрических неотрицательных вполне непрерывных (компактных) операторов в X, и пусть $S_1 = \left\{A \in L_1^+(X),\ \operatorname{Sp} A \leqslant 1\right\}$.

Слабая компактность семейства мер в гильбертовом пространстве эквивалентна равномерной непрерывности (в определенном смысле) семейства характеристических функционалов мер.

Теорема 1. *Пусть M — семейство конечных мер на $\mathfrak{B}$, а $\chi_\mu(z)$ ($z \in \mathscr{X}$) — характеристический функционал меры $\mu \in M$. Для слабой компактности множества M необходимо и достаточно, чтобы:*

а) *$\chi_\mu(0)$ были ограничены в совокупности для всех $\mu \in M$;*

б) *для всякого $\varepsilon > 0$ можно указать оператор $B \in T_C$ и для каждой меры $\mu \in M$ — оператор $A_\mu \in S_1$ такие, что $\operatorname{Re}[\chi_\mu(0) - \chi_\mu(z)] \leqslant \varepsilon$ при $(BA_\mu Bz, z) \leqslant 1$.*

Замечание. Существует пример, в котором для слабо компактной совокупности мер нельзя указать такой оператор $A \in S$ (общий для всех мер), что $\operatorname{Re}[\chi_\mu(0) - \chi_\mu(z)] \leqslant \varepsilon$ при $(Az, z) \leqslant 1$.

В условии б) теоремы 1 строятся операторы $C_\mu \in S$, представимые в виде $C_\mu = BA_\mu B$, где $B \in T_C$, $A_\mu \in S_1$. Ниже приводятся условия, когда такое представление возможно.

Лемма 1. *Для того чтобы семейство операторов $C_\mu \in S$ было представимо в виде $C_\mu = BA_\mu B$, где $B \in T_C$, $A_\mu \in S_1$, необходимо, чтобы в каждом ортонормированном базисе $\{e_k\}$ ряд*

$$\operatorname{Sp} C_\mu = \sum_{k=1}^{\infty} (C_\mu e_k, e_k)$$

сходился равномерно по μ, и достаточно, чтобы этот ряд сходился равномерно хотя бы в одном базисе.

Приведем результаты, позволяющие заменять условия леммы 1 условиями компактности в некоторых пространствах операторов. Обозначим через $L_1(X)$ пространство симметрических операторов C, для которых $\operatorname{Sp}|C| < \infty$ (модуль симметрического оператора определяется как положительный корень квадратный из его квадрата), и будем полагать $\|C\|_1 = \operatorname{Sp}|C|$. С такой нормой $L_1(X)$ является банаховым пространством.

Лемма 2. *Для того чтобы семейство операторов $\{C_\mu,\ \mu \in M\}$ было представимо в виде $C_\mu = BA_\mu B$, где $B \in T_C$, $A_\mu \in S_1$, необходимо и достаточно, чтобы множество $\{C_\mu,\ \mu \in M\}$ было вполне ограниченным в $L_1(X)$.*

(Подмножество метрического пространства вполне ограничено, если оно имеет для всех $\varepsilon > 0$ конечную ε-сеть. Для полных метрических пространств это означает, что замыкание этого подмножества есть компакт.)

Обозначим через $L_2(X)$ пространство симметрических операторов C, для которых $C^2 \in L_1(X)$: это — гильбертово пространство со скалярным произведением $\langle A, B\rangle = \operatorname{Sp} AB$.

Лемма 3. *Множество $\{C_\mu,\ \mu \in M\}$ операторов из S будет вполне ограниченным в $L_1(X)$ тогда и только тогда, когда множество $\left\{C_\mu^{1/2},\ \mu \in M\right\}$ будет вполне ограниченным в $L_2(X)$.*

Далее приводятся условия компактности семейства мер или сходимости последовательности мер, использующие операторы, аналогичные корреляционным. Пусть P_μ — оператор из S, определяемый равенством

$$(B_\mu z,\ z) = \int (z,\ x)^2 (1 + |x|^2)^{-1} \mu(dx).$$

Теорема 2. *Для слабой компактности семейства мер M на $\mathfrak{B}$ необходимо и достаточно, чтобы*

а) *для всякого $\varepsilon > 0$ существовало такое $c > 0$, что*

$$\sup_{\mu \in M} \mu\{x: |x| > c\} < \varepsilon;$$

б) *семейство операторов $\{B_\mu,\ \mu \in M\}$ было вполне ограниченным в $L_1(X)$ (или семейство операторов $\left\{B_\mu^{1/2},\ \mu \in M\right\}$ было вполне ограниченным в $L_2(X)$).*

Замечание. Для выполнения условия а) необходимо и достаточно выполнения одного из условий:

1) $\lim\limits_{\lambda \downarrow 0} \sup\limits_{\mu} \int \frac{1}{1 + \lambda |x|^2} \mu(dx) = 1;$

2) $\lim\limits_{\lambda \downarrow 0} \sup\limits_{\mu} \int e^{-\lambda |x|^2} \mu(dx) = 1;$

3) существует такая возрастающая функция $\varphi(t)$, $\varphi \uparrow +\infty$ при $t \uparrow +\infty$, что $\sup\limits_{\mu} \int \varphi(|x|)\, \mu(dx) < \infty$.

Следствие. *Пусть меры $\mu \in M$ имеют корреляционные операторы $A_\mu \in S$, где $(A_\mu z,\ z) = \int (z,\ x)^2 \mu(dx)$. Тогда для слабой компактности семейства M достаточно, чтобы множество $\{A_\mu,\ \mu \in M\}$ было вполне ограниченным в $L_1(X)$, или множество $\left\{A_\mu^{1/2},\ \mu \in M\right\}$ было вполне ограниченным в $L_2(X)$, или в некотором ортонормированном базисе $\{e_k\}$ ряд $\Sigma (A_\mu e_k,\ e_k)$ сходился бы равномерно по $\mu \in M$.*

Приведем условия слабой сходимости последовательности мер.

Теорема 3. *Для того чтобы последовательность мер μ_n слабо сходилась к мере μ, необходимо и достаточно, чтобы*

а) *множество операторов $\left\{B_{\mu_n},\ n = 1, 2, \ldots\right\}$ было компактным в $L_1(X)$;*

б) *для всех $z \in X$*

$$\lim_{n \to \infty} \int e^{i(z,\ x)} \mu_n(dx) = \int e^{i(z,\ x)} \mu(dx).$$

18.2.3. Суммирование независимых случайных величин со значениями в гильбертовом пространстве. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — последовательность независимых случайных величин со значениями в гильбертовом пространстве X. Будем обозначать через μ_n распределение величины ξ_n; $\chi_n(z)$ — ее характеристический функционал.

Теорема 4 (обобщение теоремы Колмогорова о трех рядах). *Для того чтобы ряд из случайных величин $\sum\limits_{n=1}^{\infty} \xi_n$ сходился в X*

(по норме) с вероятностью 1, необходимо, чтобы для всех $c > 0$ сходились ряды

а) $\sum_{n=1}^{\infty} \mathsf{P}\{|\xi_n| > c\}$;

б) $\sum_{n=1}^{\infty} a_n(c)$, $a_n(c) = \int_{|x| \leqslant c} x\mu_n(dx)$;

в) $\sum_{n=1}^{\infty} \int_{|x| \leqslant c} |x - a_n(c)|^2 \mu_n(dx)$,

и достаточно, чтобы эти ряды сходились хотя бы при одном $c > 0$.

Теорема 5. *Для сходимости ряда $\sum_{n=1}^{\infty} \xi_n$ необходимо и достаточно, чтобы произведение $\prod_{n=1}^{\infty} \chi_n(z)$ сходилось к характеристическому функционалу некоторой случайной величины равномерно в каждой области $\{|z| \leqslant c\}$.*

Приведем также простейший вариант центральной предельной теоремы в гильбертовом пространстве.

Теорема 6. *Пусть ξ_n одинаково распределены, $\int x\mu_n(dx) = 0$, $\int |x|^2 \mu_n(dx) < \infty$. Тогда последовательность распределений ν_n величин*

$$\zeta_n = n^{-1/2}(\xi_1 + \ldots + \xi_n)$$

слабо сходится к нормальному распределению в X, задаваемому характеристическим функционалом

$$\chi(z) = \exp\left\{-\frac{1}{2}\int (z, x)^2 \mu_1(dx)\right\}.$$

18.3. Предельные теоремы для непрерывных случайных процессов

18.3.1. Общие условия сходимости распределений функционалов. Здесь рассматриваются случайные процессы, непрерывные с вероятностью 1.

Пусть $C_{[a,b]}(\mathscr{X})$ — множество непрерывных функций $x(t)$, определенных на отрезке $[a, b]$ и принимающих значения в полном сепарабельном метрическом пространстве $\mathscr{X}$.

Введем в пространстве $C_{[a,b]}(\mathscr{X})$ метрику

$$r(x, y) = \sup_{a \leqslant t \leqslant b} \rho(x(t), y(t)), \tag{3.1}$$

где $\rho(x, y)$ — расстояние в $\mathscr{X}$. Метрика (3.1) превращает $C_{[a, b]}(\mathscr{X})$ в полное сепарабельное метрическое пространство.

Обозначим через $\mathfrak{B}_{[a, b]}(\mathscr{X})$ σ-алгебру всех борелевских множеств в $C_{[a, b]}(\mathscr{X})$. Эта σ-алгебра совпадает с минимальной σ-алгеброй, содержащей все цилиндрические множества из $C_{[a, b]}(\mathscr{X})$.

Пусть $\xi(t)$ — случайный процесс со значениями в $\mathscr{X}$, определенный для $t \in [a, b]$ и непрерывный с вероятностью 1. Тогда вероятностная мера μ, соответствующая случайному процессу $\xi(t)$, сосредоточена на измеримом пространстве $\{C_{[a, b]}(\mathscr{X}), \mathfrak{B}_{[a, b]}(\mathscr{X})\}$. При этом значения меры μ на цилиндрических множествах из $C_{[a, b]}(\mathscr{X})$ задаются конечномерными распределениями процесса $\xi(t)$.

Обычно в предельных теоремах для случайных процессов предполагается сходимость конечномерных распределений, т. е. сходимость мер $\mu_n(A)$ к мере $\mu(A)$ для всех цилиндрических множеств A, являющихся множествами непрерывности предельной меры μ.

Если предельная мера μ сосредоточена на пространстве непрерывных функций $C_{[a, b]}(\mathscr{X})$, то класс $\mathfrak{A}_0$ открытых цилиндрических множеств непрерывности меры μ удовлетворяет условиям теоремы 4 § 18.1. Поэтому для доказательства слабой сходимости мер μ_n к μ требуется установить условия слабой компактности мер μ_n $(n \geqslant 0)$, для чего достаточно указать общий вид компакта в пространстве $C_{[a, b]}(\mathscr{X})$ (см. теорему 2 § 18.1).

Пусть λ_δ — положительная монотонная непрерывная функция, определенная при $\delta > 0$ и удовлетворяющая условию $\lambda_{+0} = 0$. Пусть X_0 — компакт в $\mathscr{X}$.

Лемма 1. *Множество $K(X_0, \lambda_\delta)$ функций $x(t)$, удовлетворяющих условиям:* а) $x(t) \in X_0$, $a \leqslant t \leqslant b$; б) $\rho(x(t_1), x(t_2)) \leqslant \lambda_\delta$, $|t_1 - t_2| < \delta$ $\forall \delta > 0$, *компактно в* $C_{[a, b]}(\mathscr{X})$.

Для всякого компакта K_0 в $C_{[a, b]}(\mathscr{X})$ можно указать компакт X_0 в $\mathscr{X}$ и функцию λ_δ, положительную монотонную непрерывную при $\delta > 0$, с $\lambda_{+0} = 0$ такие, что $K_0 \subset K(X_0, \lambda_\delta)$.

Пусть $\xi_n(t)$ $(n \geqslant 0)$ — последовательность случайных процессов, выборочные функции которых принадлежат пространству $C_{[a, b]}(\mathscr{X})$ с вероятностью 1, и μ_n — вероятностные меры, соответствующие процессам $\xi_n(t)$.

Лемма 2. *Слабая сходимость мер $\mu_n \Rightarrow \mu_0$ при $n \to \infty$ эквивалентна сходимости распределений $f(\xi_n(\cdot))$ к распределению $f(\xi_0(\cdot))$ для всякого μ-почти всюду непрерывного $\mathfrak{B}_{[a, b]}(\mathscr{X})$-измеримого функционала $f(x)$.*

Приведенные ниже условия (3.2)—(3.4) являются условиями слабой компактности последовательности мер μ_n, соответствующих случайным процессам $\xi_n(t)$.

Теорема 1. *Пусть конечномерные распределения процессов $\xi_n(t)$ сходятся к конечномерным распределениям процесса $\xi_0(t)$. Для того чтобы для всех функционалов f, непрерывных на $C_{[a, b]}(\mathscr{X})$, распределения $f(\xi_n(\cdot))$ сходились к распределению $f(\xi_0(\cdot))$, необходимо и достаточно, чтобы для всякого $\lambda > 0$ выполнялось соотношение*

$$\lim_{\delta \to 0} \sup_n \mathbf{P}\left\{\sup_{|t_1 - t_2| \leqslant \delta} \rho(\xi_n(t_1), \xi_n(t_2)) > \lambda\right\} = 0. \qquad (3.2)$$

Замечания. 1. Вместо условия (3.2) для любого $\lambda > 0$ достаточно выполнения условия

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \mathsf{P}\left\{\sup_{|t_1-t_2|<\delta} \rho(\xi_n(t_1), \xi_n(t_2)) > \lambda\right\} = 0. \tag{3.3}$$

2. Вместо условия (3.2) достаточно выполнения следующего условия: существуют такие $\alpha > 0$, $\beta > 0$ и $H > 0$, что для всех $t_1, t_2 \in [a, b]$ и всех n

$$\mathsf{M}[\rho(\xi_n(t_1), \xi_n(t_2))]^\alpha \leqslant H|t_1 - t_2|^{1+\beta}. \tag{3.4}$$

Для различных типов непрерывных случайных процессов условия сходимости функционалов конкретизируются.

18.3.2. Процессы с независимыми приращениями. Для непрерывных процессов с независимыми приращениями $\xi_n(t)$ $(n \geqslant 0)$, определенных на отрезке $[a, b]$, со значениями в банаховом пространстве $\mathscr{X}$ условия сходимости функционалов устанавливаются с учетом следующего свойства выборочных функций:

$$\lim_{\delta\to 0} \sum_{k=0}^{n-1} \mathsf{P}\{|\xi(t_{k+1}) - \xi(t_k)| > \varepsilon\} = 0 \tag{3.5}$$

для всех $\varepsilon > 0$. Здесь $a = t_0 < \ldots < t_n = b$, $\delta = \max_k (t_{k+1} - t_k)$.

Теорема 2. *Для того чтобы для всякой функции $\varphi(x)$, непрерывной на $C_{[a,b]}(\mathscr{X})$, распределения случайных величин $\varphi(\xi_n(\cdot))$ сходились к распределению величины $\varphi(\xi_0(\cdot))$, необходимо и достаточно, чтобы выполнялись условия:*

1) *частные распределения процессов $\xi_n(t)$ сходятся к частным распределениям процесса $\xi_0(t)$;*

2) *для всякого $\varepsilon > 0$*

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \sup_{|t_2-t_1|\leqslant\delta} \mathsf{P}\{|\xi_n(t_2) - \xi_n(t_1)| > \varepsilon\} = 0. \tag{3.6}$$

18.3.3. Марковские процессы. Для непрерывных марковских процессов $\xi_n(t)$ $(n \geqslant 0)$, определенных на отрезке $[a, b]$, со значениями в полном метрическом пространстве $\mathscr{X}$ с метрикой ρ и вероятностями перехода $P_n(t, x, s, A)$ введем

$$\alpha_n(h, \varepsilon) = \sup\{\mathsf{P}_n(t_1, x, t_2, V_\varepsilon(x));\ x \in \mathscr{X},\ |t_2 - t_1| \leqslant h\}, \tag{3.7}$$

где $V_\varepsilon(x) = \{y\colon \rho(x, y) > \varepsilon\}$.

Теорема 3. *Пусть конечномерные распределения процессов $\xi_n(t)$ $(n \geqslant 1)$ сходятся при $n \to \infty$ к конечномерным распределениям процесса $\xi_0(t)$ и выполнены условия: для всякого $\varepsilon > 0$*

$$\lim_{h\to 0} \sup_n \alpha_n(h, \varepsilon) = 0, \quad \lim_{\delta\to 0} \sum_{k=0}^{n-1} \mathsf{P}\{\rho(\xi(t_{k+1}), \xi(t_k)) > \varepsilon\} = 0, \tag{3.8}$$

где $a = t_0 < t_1 < \ldots < t_n = b$, $\delta = \max_k (t_{k+1} - t_k)$.

Тогда для всякой функции φ, непрерывной на $C_{[a,c]}(\mathscr{X})$, распределения $\varphi(\xi_n(\cdot))$ сходятся к распределению $\varphi(\xi_0(\cdot))$.

18.3.4. Непрерывные процессы, построенные по суммам независимых случайных величин. Пусть $\{\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n}\}_{n\geq 1}$ — последовательность серий независимых в каждой серии числовых случайных величин, удовлетворяющих условиям

$$\mathsf{M}\xi_{ni}=0, \quad i=1, \ldots, k_n; \quad \mathsf{D}\xi_{ni}=b_{ni}, \quad \sum_{i=1}^{k_n} b_{ni}=1. \tag{3.9}$$

Определим случайные функции $\xi_n(t)$ для $t \in [0, 1]$ соотношениями

$$\begin{gathered} S_{nk}=\sum_{i=1}^{k}\xi_{ni}, \quad t_{nk}=\sum_{i=1}^{k} b_{ni}, \\ \xi_n(t)=S_{nk}+\frac{t-t_{nk}}{t_{nk+1}-t_{nk}}\left[S_{nk+1}-S_{nk}\right], \quad t \in [t_{nk}, t_{nk+1}). \end{gathered} \tag{3.10}$$

При этом $S_{n0}=0$, $t_{n0}=0$. Тогда $\xi_n(t)$ является случайной ломаной, соединяющей точки на плоскости с координатами (t_{nk}, S_{nk}) $(k=0, 1, \ldots, k_n)$.

Приведем условия, при которых частные распределения процессов $\xi_n(t)$ и распределения функционалов от этих процессов сходятся к частным распределениям и распределениям соответствующих функционалов от винеровского процесса $w(t)$ (см. п. 16.6.1).

Теорема 4. *Пусть независимые случайные величины ξ_{ni} с функциями распределения $F_{ni}(x)$ удовлетворяют условию* (3.9) *и условию Линдеберга*:

$$\lim_{n\to\infty}\sum_{i=1}^{k_n}\int_{|x|>\varepsilon} x^2\,dF_{ni}(x)=0 \tag{3.11}$$

для всякого $\varepsilon>0$.

Тогда конечномерные распределения процессов $\xi_n(t)$, определяемых соотношением (3.10), *сходятся к конечномерным распределениям винеровского процесса $w(t)$ и распределения $f(\xi_n(\cdot))$ сходятся к распределению $f(w(\cdot))$ для всякого непрерывного на $C_{[0,1]}(\mathbf{R})$ функционала f.*

18.3.5. Предельные теоремы для функционалов от сумм независимых одинаково распределенных величин. Пусть $\{\xi_n, n=1, 2, \ldots\}$ — последовательность независимых одинаково распределенных величин, $\mathsf{M}\xi_n=0$, $\mathsf{D}\xi_n=1$. Положим $S_k=\sum_{i=1}^{k}\xi_i$, $S_0=0$. Здесь приведены некоторые предельные теоремы для функционалов типа maximuma и аддитивных функционалов от последовательности сумм S_k.

Теорема 5. *Пусть $a(t)$ и $b(t)$ ($t \in [0, 1]$) — две непрерывные функции. Тогда*

$$\lim_{n\to\infty}\mathsf{P}\left\{\max_{1\leqslant k\leqslant n}\frac{|S_k-\sqrt{n}\,a(k/n)|}{b(k/n)}\leqslant\sqrt{n}\right\}= \\ =\mathsf{P}\{|w(t)-a(t)|\leqslant b(t), \quad t\in[0, 1]\}.$$

Следствие.

1) $\lim\limits_{n\to\infty} \mathsf{P}\left\{\max\limits_{k} |S_k| \leqslant a\sqrt{n}\right\} = \mathsf{P}\left\{\max\limits_{0\leqslant t\leqslant 1} |w(t)| < a\right\}$;

2) $\lim\limits_{n\to\infty} \mathsf{P}\left\{\max\limits_{k} S_k \leqslant a\sqrt{n}\right\} = \mathsf{P}\left\{\max\limits_{0\leqslant t\leqslant 1} w(t) < a\right\}$;

3) $\lim\limits_{n\to\infty} \mathsf{P}\left\{S_k \leqslant a(k/n)\sqrt{n},\quad k=1, \ldots, n\right\} =$

$$= \mathsf{P}\{w(t) \leqslant a(t),\quad t \in [0, 1]\}.$$

В следующей теореме устанавливаются условия сходимости распределений по вариации.

Теорема 6. *Пусть при некотором m величина S_m имеет ненулевую абсолютно непрерывную компоненту. Тогда распределения величин*

$$\max_{1\leqslant k\leqslant n}\left(\frac{1}{\sqrt{n}} S_k + h\left(\frac{k}{n}\right)\right), \quad \max_{1\leqslant k\leqslant n}\left|\frac{1}{\sqrt{n}} S_k + h\left(\frac{k}{n}\right)\right|,$$

где $h(t)$ — функция из $C_{[0,1]}$, сходятся при $n\to\infty$ по вариации соответственно к распределениям функционалов от винеровского процесса $\max\limits_{0\leqslant t\leqslant 1}(w(t)+h(t))$ и $\max\limits_{0\leqslant t\leqslant 1}|w(t)+h(t)|$.

Приведем результаты, относящиеся к аддитивным функционалам.

Теорема 7. *Пусть $h(x)$ — локально интегрируемая на прямой функция. Тогда для почти всех a*

$$\lim_{n\to\infty} \mathsf{P}\left\{\frac{1}{n}\sum_{k=1}^{n} h\left(\frac{1}{\sqrt{n}} S_k\right) < a\right\} = \mathsf{P}\left\{\int_0^1 h(w(t))\,dt < a\right\}.$$

Теорема 8. *Пусть величины ξ_k имеют абсолютно непрерывную плотность $p(x)$ относительно меры Лебега, и $\int |p'(x)|^2 p^{-1}(x)\,dx < \infty$. Если функция $h(x)$ дифференцируема в некоторой окрестности нуля и имеет там отличную от нуля производную, то распределение величины*

$$\frac{1}{n}\sum_{k=1}^{n} h\left(\frac{1}{\sqrt{n}} S_k\right)$$

сходится при $n\to\infty$ по вариации к распределению величины

$$\int_0^1 h(w(t))\,dt.$$

Пусть $\Phi_n(x_1, x_2, \ldots, x_m)$ — последовательность измеримых локально ограниченных функций в $\mathbf{R}^m$. Рассмотрим последовательность случайных величин

$$\eta_n = \sum_{k=0}^{n-m} \Phi_n(S_{k+1}, \ldots, S_{k+m}).$$

Теорема 9. *Пусть для каждого* $C > 0$

$$\lim_{n\to\infty} \sup_{|x_k|\leqslant C\sqrt{n},\ k=1,\ \dots,\ m} |\Phi_n(x_1, \dots, x_m)| = 0.$$

Тогда

1) *предельное распределение случайной величины* η_n *существует тогда и только тогда, когда существует предельное распределение величины*

$$\bar\eta_n = \sum_{k=1}^{n} \bar\Phi_n(S_k),$$

где $\bar\Phi_n(x) = \mathsf{M}\Phi_n(x, x+S_1, \dots, x+S_{m-1})$;

2) *предельные распределения величин* η_n *и* $\bar\eta_n$ *в этом случае совпадают;*

3) *предельное распределение величины* $\bar\eta_n$ *существует, если для всех* $C > 0$

$$\varlimsup_{n\to\infty} \frac{1}{\sqrt{n}} \int\limits_{|y|\leqslant C\sqrt{n}} \bar\Phi_n(y)\, dy < \infty$$

и существует для всех x *предел*

$$G(x) = \lim_{n\to\infty} \frac{2}{\sqrt{n}} \int\limits_0^{x\sqrt{n}} \bar\Phi_n(y)\, dy,$$

а предельное распределение величины $\bar\eta_n$ *совпадает с распределением величины*

$$\bar\eta = \int\limits_0^{w(1)} G(x)\, dx - \int\limits_0^1 G(w(s))\, dw(s).$$

18.3.6. Предельные теоремы для возрастающих процессов и мартингалов. В этом пункте рассматриваются непрерывные возрастающие процессы и числовые мартингалы. Рассмотрим сначала условия слабой компактности мер, соответствующих таким процессам в $C_{[0,1]}$.

Теорема 10. *Пусть* $\eta_n(t)$ $(t \in [0, 1])$ — *последовательность возрастающих процессов,* $\eta_n(0) = 0$. *Для того чтобы последовательность соответствующих им мер была слабо компактна, необходимо и достаточно, чтобы для каждого* $c > 0$ *и* $\varepsilon > 0$

$$\lim_{\max \Delta t_k \to 0} \sup_n \sum_{k=0}^{m-1} \mathsf{P}\{\eta_n(t_{k+1}) \wedge c - \eta_n(t_k) \wedge c > \varepsilon\} = 0,$$

где $0 = t_0 < t_1 < \dots < t_m = 1$, $\Delta t_k = t_{k+1} - t_k$.

Теорема 11. *Пусть* $\xi_n(t)$ $(t \in [0, 1])$ — *последовательность непрерывных мартингалов,* $\alpha_n(t)$ — *характеристики* $\xi_n(t)$. *Для того чтобы последовательность мер, соответствующих* $\xi_n(t)$, *была слабо компактна, необходимо и достаточно, чтобы*

а) *последовательность* $\xi_n(0)$ *была ограничена по вероятности;*
б) *меры, соответствующие* $\alpha_n(t)$, *были слабо компактны.*

Используя теоремы о компактности, можем получить теоремы о слабой сходимости мер.

Следствие. *Пусть* $\xi_n(t)$ $(t \in [0, 1])$ — *последовательность непрерывных мартингалов с характеристиками* $\alpha_n(t)$. *Для того чтобы меры, соответствующие* $\xi_n(t)$, *слабо сходились к мере* μ_w, *соответствующей винеровскому процессу, необходимо и достаточно выполнения условий:*

а) $\xi_n(0) \to 0$ *по вероятности;*
б) $\alpha_n(t) \to t$ *по вероятности для* $t \in [0, 1]$;
в) *последовательность мер, соответствующих* $\alpha_n(t)$, *слабо компактна.*

Приведем общую теорему о сходимости к винеровскому процессу.

Теорема 12. *Пусть* $\xi_n(t)$ $(t \in [0, 1])$ — *последовательность непрерывных процессов, для которой последовательность соответствующих им мер* μ_n *слабо компактна. Для того чтобы* $\mu_n \Rightarrow \mu_w$, *необходимо и достаточно, чтобы, каковы бы ни были натуральные* $l > k$ *и непрерывная ограниченная функция* $\varphi(x_1, \ldots, x_k)$ *на* $\mathbf{R}^k$,

$$\text{а)} \lim_{c\to\infty} \lim_{n\to\infty} M\varphi\left(\xi_n\left(\frac{1}{l}\right), \ldots, \xi_n\left(\frac{k}{l}\right)\right) \times$$
$$\times I_{\left\{\left|\xi_n\left(\frac{k+1}{l}\right)-\xi_n\left(\frac{k}{l}\right)\right| \leqslant c\right\}} \left(\xi_n\left(\frac{k+1}{l}\right) - \xi_n\left(\frac{k}{l}\right)\right) = 0;$$

$$\text{б)} \lim_{c\to\infty} \lim_{n\to\infty} M\varphi\left(\xi_n\left(\frac{1}{l}\right), \ldots, \xi_n\left(\frac{k}{l}\right)\right) \times$$
$$\times I_{\left\{\left|\xi_n\left(\frac{k+1}{l}\right)-\xi_n\left(\frac{k}{l}\right)\right| \leqslant c\right\}} \left[\left(\xi_n\left(\frac{k+1}{l}\right)-\xi_n\left(\frac{k}{l}\right)\right)^2 - \frac{1}{l}\right] = 0.$$

18.4. Предельные теоремы для процессов без разрывов второго рода

18.4.1. Метрика в пространстве функций без разрывов второго рода. Пусть $D_{[0, 1]}(\mathscr{X})$ — множество функций $x(t)$, определенных на отрезке $[0, 1]$, принимающих значения из полного сепарабельного метрического пространства $\mathscr{X}$ с метрикой ρ и имеющих предельные значения $x(t+0)$ для $0 \leqslant t < 1$ и $x(t-0)$ для $0 < t \leqslant 1$.

Поскольку функции, совпадающие во всех точках непрерывности, не различаются, то естественно фиксировать значения функций в точках разрыва:

$$x(t) = x(t+0), \quad x(0) = x(+0), \quad x(1) = x(1-0).$$

Величину $\rho(x(t-0), x(t))$ называют *величиной скачка* функции $x(t)$ в точке t.

Обозначим через Λ совокупность всех непрерывных монотонно возрастающих числовых функций на отрезке $[0, 1]$ с $\lambda(0) = 0$, $\lambda(1) = 1$, т. е. $\lambda(t)$ — взаимно однозначное и непрерывное отображение $[0, 1]$ на $[0, 1]$.

Метрика $r_D(x, y)$ в пространстве $D_{[0,1]}(\mathscr{X})$ определяется соотношением

$$r_D(x, y) = \inf\left\{\varepsilon > 0 \colon \exists \lambda \in \Lambda,\ \sup_{s \neq t}\left|\log \frac{\lambda(t) - \lambda(s)}{t - s}\right| \leqslant \varepsilon,\right.$$
$$\left.\sup_t \rho(x(t),\ y(\lambda(t))) \leqslant \varepsilon\right\}. \quad (4.1)$$

Метрика $r_D(x, y)$ (ее называют *метрикой Скорохода*) превращает $D_{[0,1]}(\mathscr{X})$ в полное сепарабельное метрическое пространство (пространство Скорохода).

Общий вид компактных множеств в $D_{[0,1]}(\mathscr{X})$ определяется с использованием критерия отсутствия разрывов второго рода. Определим для каждой $x(t) \in D_{[0,1]}(\mathscr{X})$ и $C > 0$ величину

$$\Delta_C(x) = \sup\{\min[\rho(x(t'),\ x(t)),\ \rho(x(t),\ x(t''))];$$
$$t - C < t' < t < t'' < t + C,\ t',\ t,\ t'' \in [0,\ 1]\} +$$
$$+ \sup\{\rho(x(0),\ x(t));\ 0 < t < C\} +$$
$$+ \sup\{\rho(x(t),\ x(1));\ 1 - C < t < 1\}. \quad (4.2)$$

Пусть X_0 — компакт в $\mathscr{X}$, а λ_C — положительная монотонная непрерывная функция, определенная при $C > 0$ и удовлетворяющая условию $\lambda_{+0} = 0$.

Теорема 1. *Множество функций $K_D(X_0, \lambda_\delta)$, удовлетворяющих условиям:*

1) $x(t) \in X_0$, $0 \leqslant t \leqslant 1$;

2) $\Delta_C(x) \leqslant \lambda_C \quad \forall C > 0$,

компактно в $D_{[0,1]}(\mathscr{X})$.

Для всякого компакта K_0 в $D_{[0,1]}(\mathscr{X})$ можно указать компакт $X_0 \subset \mathscr{X}$ и функцию λ_C, положительную монотонную и непрерывную при $C > 0$ с $\lambda_{+0} = 0$ такие, что $K_0 \subset K_D(X_0, \lambda_C)$.

18.4.2. Основная предельная теорема для процессов без разрывов второго рода.

Теорема 2. *Пусть частные распределения процессов $\xi_n(t)$ $(0 \leqslant t \leqslant 1,\ n \geqslant 0)$, выборочные функции которых принадлежат $D_{[0,1]}(\mathscr{X})$ с вероятностью 1, сходятся при $n \to \infty$ к частным распределениям процесса $\xi_0(t)$ без разрывов второго рода. Для того чтобы для всякого функционала f, определенного на $D_{[0,1]}(\mathscr{X})$ непрерывного в метрике r_D, распределения $f(\xi_n(\cdot))$ сходились к распределению $f(\xi_0(\cdot))$, необходимо и достаточно, чтобы для всех $\varepsilon > 0$ выполнялось условие*

$$\lim_{C \to 0} \overline{\lim_{n \to \infty}} \mathsf{P}\{\Delta_C(\xi_n(\cdot)) > \varepsilon\} = 0. \quad (4.3)$$

Замечание. Вместо условия (4.3) достаточно выполнения следующего условия: существуют такие $\alpha > 0$, $\beta > 0$ и $H > 0$, что для всех $0 \leqslant t_1 < t_2 < t_3 \leqslant 1$ и всех $n \geqslant 1$ выполнено неравенство

$$\mathsf{M}[\rho(\xi_n(t_1),\ \xi_n(t_2))\,\rho(\xi_n(t_2),\ \xi_n(t_3))]^\alpha \leqslant H(t_3 - t_1)^{1+\beta}. \quad (4.4)$$

18.4.3. Предельная теорема для марковских процессов. Пусть $\xi_n(t)$ $(0 \leqslant t \leqslant 1,\ n \geqslant 0)$ — последовательность марковских процессов, определенных на отрезке $[0, 1]$, выборочные функции которых принадлежат пространству $D_{[0,1]}(\mathscr{X})$ с вероятностью 1. Обозначим через $P_n(t, x, s, A)$ вероятности перехода для процесса $\xi_n(t)$ и введем $V_\varepsilon(x) = \{y: \rho(x, y) > \varepsilon\}$ для $\varepsilon > 0$.

Теорема 3. *Пусть частные распределения марковских процессов $\xi_n(t)$ сходятся при $n \to \infty$ к частным распределениям $\xi_0(t)$ и для каждого $\varepsilon > 0$ выполнено условие*

$$\lim_{h\to 0} \overline{\lim_{n\to\infty}} \sup \{P_n(t, x, s, V_\varepsilon(x));\ x \in \mathscr{X},\ 0 \leqslant s - t \leqslant h\} = 0. \tag{4.5}$$

Тогда для всякого непрерывного на $D_{[0,1]}(\mathscr{X})$ функционала f распределения $f(\xi_n(\cdot))$ сходятся к распределению $f(\xi_0(\cdot))$.

Процессы с независимыми приращениями в полном линейном нормированном пространстве $\mathscr{X}$ являются частным случаем марковских процессов. Следующая теорема является следствием теоремы 3.

Теорема 4. *Пусть $\xi_n(t)$ $(n \geqslant 0)$ — последовательность процессов с независимыми приращениями, определенных на $[0, 1]$, выборочные функции которых принадлежат $D_{[0,1]}(\mathscr{X})$ с вероятностью 1. Если частные распределения процесса $\xi_n(t)$ сходятся к частным распределениям процесса $\xi_0(t)$ и для всякого $\varepsilon > 0$ выполнено условие*

$$\lim_{h\to 0} \overline{\lim_{n\to\infty}} \sup_{|t-s|\leqslant h} \mathsf{P}\{|\xi_n(t) - \xi_n(s)| > \varepsilon\} = 0, \tag{4.6}$$

то для всякого функционала f, непрерывного на $D_{[0,1]}(\mathscr{X})$, распределения $f(\xi_n(\cdot))$ сходятся к распределению $f(\xi_0(\cdot))$.

Замечание. В теоремах 2—4 достаточно требовать, чтобы функционал f был измерим и μ_0-почти всюду непрерывен; μ_0 — мера, соответствующая предельному процессу $\xi_0(t)$.

18.4.4. Применение к исследованию асимптотического поведения эмпирической функции распределения. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые одинаково распределенные числовые случайные величины с непрерывной функцией распределения $F(t)$. Положим

$$F_n^*(t) = \frac{1}{n} \sum I_{\{\xi_k < t\}}, \quad \eta_n(t) = \sqrt{n}\left(F_n^*(t) - F(t)\right).$$

Лемма. *Конечномерные распределения процессов $\eta_n(t)$ сходятся к конечномерным распределениям гауссовского процесса $\eta(t)$, для которого* $\mathsf{M}\eta(t) = 0$, $\mathsf{M}\eta(t)\eta(s) = F(t \wedge s)(1 - F(t \vee s))$.

Положим $\xi_n(t) = \eta_n(F^{-1}(t))$, $\xi(t) = \eta(F^{-1}(t))$ (F^{-1} — обратная функция к F). Здесь $\xi(t)$ — гауссовский процесс, определенный на $[0, 1]$, $\mathsf{M}\xi(t) = 0$, $\mathsf{M}\xi(t)\xi(s) = t \wedge s(1 - t \vee s)$. Он называется *броуновским мостом*, и его распределения совпадают с условным распределением винеровского процесса $w(t)$ при условии $w(1) = 0$.

Теорема 5. *Меры, соответствующие процессам $\xi_n(t)$ в $D_{[0,1]}(\mathbf{R})$, слабо сходятся к мере, соответствующей процессу $\xi(t)$.*

Следствие. 1) *Для всех $a > 0$*

$$\lim_{n\to\infty} \mathsf{P}\left\{\sqrt{n} \sup_{-\infty<t<\infty} \left(F_n^*(t) - F(t)\right) < a\right\} = \mathsf{P}\left\{\sup_{0\leqslant t\leqslant 1} \xi(t) < a\right\} = 1 - e^{-2a^2}.$$

(Предельное распределение критерия Н. В. Смирнова.)

2) *Для всех* $a > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\sqrt{n}\sup_{-\infty<t<\infty}\left|F_n^*(t)-F(t)\right|<a\right\}=$$

$$=\mathsf{P}\left\{\sup_{0\leqslant t\leqslant 1}|\xi(t)|<a\right\}=\sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2a^2}.$$

(Предельное распределение критерия А. Н. Колмогорова.)

3) *Для всех* $a > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{n\int_{-\infty}^{\infty}\left[F_n^*(t)-F(t)\right]^2 dF(t)<a\right\}=G(a),$$

где $G(a)$ — *функция распределения, для которой*

$$\int e^{isa}\,dG(a)=\prod_{k=1}^{\infty}\left(1-\frac{2is}{k^2\pi^2}\right)^{-1/2}.$$

(Предельное распределение ω^2-критерия.)

18.4.5. Сходимость сумм одинаково распределенных независимых случайных величин к однородному процессу с независимыми приращениями. Пусть $\xi_{n1}, \ldots, \xi_{nn}$ при каждом n являются независимыми одинаково распределенными случайными величинами в $\mathbf{R}^m$. Положим

$$\xi_n(t)=\sum_{l=1}^{k}\xi_{nl}=S_{nk},\quad t\in\left[\frac{k-1}{n},\ \frac{k}{n}\right),\quad S_{n0}=0.$$

Рассмотрим также однородный процесс с независимыми приращениями $\xi(t)$ в $\mathbf{R}^m$.

Будем предполагать, что $\xi(t)$ — сепарабельный процесс. Тогда его можно считать непрерывным справа; при этом его выборочные функции на $[0, 1]$ с вероятностью 1 принадлежат $D_{[0,\,1]}(\mathbf{R}^m)$. Этому пространству принадлежат и выборочные функции $\xi_n(t)$.

Теорема 6. *Если распределение* $\xi_n(t)$ *сходится к распределению* $\xi(t)$, *то последовательность мер, соответствующих процессам* $\xi_n(t)$, *слабо сходится в пространстве* $D_{[0,\,1]}(\mathbf{R}^m)$ *к мере, соответствующей* $\xi(t)$.

Замечание. Распределение $\xi(t)$ является безгранично делимым. Условия сходимости распределения $\sum_{k=1}^{[nt]+1}\xi_{nk}=\xi_n(t)$ к безгранично делимым распределениям приведены в п. 5.3.1.

Следствие. 1) *Пусть* $a(t)$ *и* $b(t)>0$ $(t\in[0, 1])$ — *непрерывные функции,* $\alpha>0$, $\mathbf{R}^m=\mathbf{R}$. *Тогда в условиях теоремы* 6

$$\lim_{n\to\infty}\mathsf{P}\left\{a\left(\frac{k}{n}\right)-\alpha b\left(\frac{k}{n}\right)<S_{nk}<a\left(\frac{k}{n}\right)+\alpha b\left(\frac{k}{n}\right),\right.$$

$$\left.k=1,\ \ldots,\ n\right\}=\mathsf{P}\left\{\sup_t\frac{1}{b(t)}\,|\xi(t)-a(t)|<\alpha\right\}$$

в каждой точке непрерывности по α вероятности в правой части равенства.

2) *Пусть* $g(x)$ $(x \in \mathrm{R}^m)$ — *непрерывная функция. В условиях теоремы* 6

$$\lim_{n\to\infty} \mathrm{P}\left\{\frac{1}{n}\sum_{k=1}^{n} g(S_{nk}) < \alpha\right\} = \mathrm{P}\left\{\int_0^1 g(\xi(t))\,dt < \alpha\right\}$$

при всех α, *для которых вероятность справа непрерывна (по* α).

3) *Пусть в* $D_{[0,1]}(\mathrm{R})$ *определены функционалы* $\tau_a(x)$ *и* $\gamma_a(x)$ *на всех* $x = x(t)$, *для которых* $\sup\limits_{0\leqslant t\leqslant 1} x(t) > a$: $\tau_a(x) = \inf\{t\colon x(t) \geqslant a\}$, $\gamma_a(x) = x(\tau_a(x)) - a$. *Если* $\mathrm{R}^m = \mathrm{R}$ *и* a *таково, что для всех* $0 \leqslant t_1 < t_2 \leqslant 1$ $\mathrm{P}\left\{\sup\limits_{t_1\leqslant t\leqslant t_2} \xi(t) = a\right\} = 0$, *то совместное распределение величин* $\tau_a(\xi_n(\cdot))$ *и* $\gamma_a(\xi_n(\cdot))$ *сходится к совместному распределению величин* $\tau_a(\xi(\cdot))$ *и* $\gamma_a(\xi(\cdot))$. *Величина* $\tau_a(\xi_n)$ — *момент первого достижения процессом* ξ *уровня* a, $\gamma_a(\xi)$ — *величина перескока через этот уровень.*

Литература: [6, 9, 10, 17, 19, 20, 67, 68, 80, 82, 96, 98].

Глава 19. СТОХАСТИЧЕСКИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

19.1. Диффузионные процессы

19.1.1. Определение. Определение однородного диффузионного процесса, основанное на понятии характеристического оператора, было приведено в § 15.1. Приведем другое определение, пригодное в случае неоднородного диффузионного процесса и использующее лишь понятие вероятности перехода. О связи между этими определениями можно судить по теореме 1 п. 15.7.1.

Пусть $\mathfrak{B}$ — σ-алгебра борелевских подмножеств m-мерного евклидова пространства R^m. Функция $P(s, x, t, \Gamma)$ $(0 \leqslant s < t \leqslant T$, $x \in \mathrm{R}^m$, $\Gamma \in \mathfrak{B})$ называется *вероятностью перехода*, если выполнены условия:

а) $P(s, x, t, \Gamma)$ является $\mathfrak{B}$-измеримой функцией по x при фиксированных s, t, Γ;

б) $P(s, x, t, \Gamma)$ является вероятностной мерой на $\mathfrak{B}$ при фиксированных s, x, t (так что $P(s, x, t, \mathrm{R}^m) \equiv 1$);

в) при всех $0 \leqslant s < t_1 < t_2$, $x \in \mathrm{R}^m$ и $\Gamma \in \mathfrak{B}$ выполнено соотношение

$$P(s, x, t_2, \Gamma) = \int\limits_{\mathrm{R}^m} P(s, x, t_1, dy)\, P(t_1, y, t_2, \Gamma),$$

называемое *уравнением Колмогорова — Чепмена.*

Будем говорить, что задан марковский процесс в широком смысле со значениями в R^m, если задана вероятность перехода $P(s, x, t, \Gamma)$.

Марковский процесс в широком смысле со значениями в $\mathbf{R}^m$ на временно́м интервале $[0, T]$ называется *диффузионным*, если выполнены следующие условия:

1) при всех $\varepsilon > 0$, $t \in [0, T]$ и $x \in \mathbf{R}^m$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|y-x|>\varepsilon} P(t, x, t+\Delta t, dy) = 0;$$

2) существует функция $a(s, x)$ со значениями в $\mathbf{R}^m$ и линейный симметрический неотрицательно определенный оператор $b(s, x)$, отображающий $\mathbf{R}^m$ в $\mathbf{R}^m$, такие, что при всех $\varepsilon > 0$, $x \in \mathbf{R}^m$ и $t \in [0, T]$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|y-x|<\varepsilon} (y - x) P(t, x, t+\Delta t, dy) = a(t, x),$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|y-x|<\varepsilon} (y - x, \theta)^2 P(t, x, t+\Delta t, dy) = (b(t, x)\theta, \theta),$$

каково бы ни было $\theta \in \mathbf{R}^m$. Здесь (θ, y) — скалярное произведение в $\mathbf{R}^m$.

Если выполнено условие 1), а в условии 2) пределы существуют при некотором $\varepsilon > 0$, то они существуют и при всех $\varepsilon > 0$ и не зависят от ε.

Название «диффузионные процессы» связано с тем, что такие процессы достаточно точно описывают явление диффузии. Если в момент t диффундирующая частица находилась в точке x, то ее перемещение за время от t до $t + \Delta t$ можно записать с точностью до $o(\Delta t)$ в виде суммы $a(t, x)\Delta t + \delta(t, t+\Delta t, x)$, где $a(t, x)\Delta t$ — неслучайное смещение, связанное с макроскопическим движением среды, в которой происходит диффузия, а $\delta(t, t+\Delta t, x)$ — случайный вектор, связанный с хаотическим тепловым движением молекул рассматриваемой среды. При этом мы считаем, что математическое ожидание вектора $\delta(t, t+\Delta t, x)$ равно нулю, а среднее квадрата его проекции на произвольное направление $\theta \in \mathbf{R}^m$ имеет вид $|\theta|^{-2}(b(t, x)\theta, \theta)\Delta t$. Вектор $a(t, x)$ называется *вектором переноса*, а оператор $b(t, x)$ — *оператором диффузии*. В одномерном случае они называются соответственно *коэффициентом переноса* и *коэффициентом диффузии*.

19.1.2. Уравнения Колмогорова. Следующие две теоремы показывают, что диффузионные процессы тесно связаны с дифференциальными уравнениями в частных производных параболического типа.

Пусть в $\mathbf{R}^m$ выбран некоторый базис. Обозначим через $a^i(s, x)$ координаты вектора $a(s, x)$, а через $b^{ij}(s, x)$ элементы матрицы $b(s, x)$ в этом базисе.

Теорема 1. *Пусть задан диффузионный процесс, для которого функции $a(s, x)$ и $b(s, x)$ непрерывны, а ограниченная непрерывная функция $f(x)$ с вещественными значениями такова, что функция*

$$u(s, x) = \int\limits_{\mathbf{R}^m} P(s, x, t, dy) f(y), \quad s \in [0, t], \quad x \in \mathbf{R}^m,$$

дважды непрерывно дифференцируема по x. Тогда функция $u(s, x)$ дифференцируема по s и удовлетворяет уравнению

$$-\frac{\partial u}{\partial s} = \frac{1}{2} \sum_{i, j=1}^{m} b^{ij}(s, x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1}^{m} a^i(s, x) \frac{\partial u}{\partial x^i},$$

$$x \in \mathbf{R}^m, \quad 0 \leqslant s < t,$$

и начальному условию

$$\lim_{s \uparrow t} u(s, x) = f(x).$$

Это уравнение называется *обратным уравнением Колмогорова*.

Во многих случаях вероятность перехода имеет плотность $G(s, x, t, y)$ относительно лебеговой меры. Это означает, что при всех $0 \leqslant s < t$, $x \in \mathbf{R}^m$, $\Gamma \in \mathfrak{B}$

$$P(s, x, t, \Gamma) = \int_{\Gamma} G(s, x, t, y)\, dy.$$

Если функция $G(s, x, t, y)$ достаточно гладкая как функция от (t, y), то она удовлетворяет прямому уравнению Колмогорова, называемому также *уравнением Фоккера — Планка*.

Теорема 2. *Если для диффузионного процесса предельные соотношения в определении диффузионного процесса выполнены равномерно относительно $x \in \mathbf{R}^m$ и существуют непрерывные производные*

$$\frac{\partial G(s, x, t, y)}{\partial t}, \quad \frac{\partial}{\partial y^i}\left(a^i(t, y) G(s, x, t, y)\right),$$

$$\frac{\partial^2}{\partial y^i \partial y^j}\left(b^{ij}(t, y) G(s, x, t, y)\right),$$

то функция $G(s, x, t, y)$ при всех $(t, y) \in (s, T) \times \mathbf{R}^m$ удовлетворяет уравнению

$$\frac{\partial G(s, x, t, y)}{\partial t} = \frac{1}{2} \sum_{i, j=1}^{m} \frac{\partial^2}{\partial y^i \partial y^j}\left(b^{ij}(t, y) G(s, x, t, y)\right) - \sum_{i=1}^{m} \frac{\partial}{\partial y^i}\left(a^i(t, y) G(s, x, t, y)\right).$$

Теоремы 1 и 2 показывают, что при построении диффузионных процессов и при изучении их свойств можно использовать теорию дифференциальных уравнений в частных производных параболического типа (см. п. 15.7.2). Однако есть и другой метод построения диффузионных процессов, основанный на прямом построении траекторий таких процессов как решений (систем) стохастических дифференциальных уравнений. Для того чтобы понять, какой вид должны иметь эти уравнения, рассмотрим диффузионный процесс $\xi(t)$. Его приращение $\xi(t+\Delta t)-\xi(t)$ будет иметь такие же условные (при фиксированном $\xi(t)$) урезанные моменты первых двух порядков, как и вектор

$$a(t, \xi(t))\Delta t + \sigma(t, \xi(t))(w(t+\Delta t) - w(t)),$$

где симметрический оператор $\sigma(t, x)$ таков, что $\sigma^2(t, x) = b(t, x)$, а $w(t)$ — m-мерный винеровский процесс. С точностью до $o(\Delta t)$ мы можем записать приближенное равенство (имея в виду совпадение условных распределений левой и правой частей)

$$\xi(t+\Delta t) - \xi(t) \approx a(t, \xi(t))\Delta t + \sigma(t, \xi(t))(w(t+\Delta t) - w(t)).$$

Естественно ожидать, что при переходе к дифференциалам мы получим точное совпадение распределений. Само уравнение должно иметь вид

$$d\xi(t) = a(t, \xi(t))\,dt + \sigma(t, \xi(t))\,dw(t).$$

Чтобы придать смысл этому уравнению, запишем его в интегральной форме:

$$\xi(t) = \xi(0) + \int_0^t a(s, \xi(s))\,ds + \int_0^t \sigma(s, \xi(s))\,dw(s).$$

Теперь задача состоит в том, чтобы придать смысл второму интегралу в правой части этого уравнения. Интегралы такого вида, называемые *стохастическими интегралами*, рассмотрены в § 19.2. Заметим, что винеровский процесс с вероятностью 1 имеет неограниченную вариацию на каждом промежутке, и потому этот интеграл нельзя понимать в смысле Стилтьеса.

В заключение приведем достаточные условия, при выполнении которых процесс будет диффузионным.

Для того чтобы марковский процесс (в широком смысле) был диффузионным, достаточно, чтобы вероятность перехода процесса $P(s, x, t, \Gamma)$ удовлетворяла условиям:

1) при некотором $\delta > 0$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathbf{R}^m} |y - x|^{2+\delta} P(t, x, t+\Delta t, dy) = 0, \quad x \in \mathbf{R}^m,$$
$$t \in [0, T];$$

2) существуют вектор-функция $a(t, x) \in \mathbf{R}^m$ и операторная функция $b(t, x)$, представляющая собой при каждом t и x симметричный неотрицательно определенный оператор, действующий в $\mathbf{R}^m$, такие, что при всех t и x выполнены соотношения ($x \in \mathbf{R}^m$, $t \in [0, T]$)

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathbf{R}^m} (y - x) P(t, x, t+\Delta t, dy) = a(t, x),$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathbf{R}^m} (y - x, \theta)^2 P(t, x, t+\Delta t, dy) = (b(t, x)\theta, \theta), \quad \theta \in \mathbf{R}^m.$$

19.2. Стохастические интегралы по винеровскому процессу

19.2.1. Определение стохастического интеграла по одномерному винеровскому процессу. Приведенные в этом параграфе результаты являются частным случаем общей теории стохастического интеграла по локально квадратически интегрируемым мартингалам (см. п. 13.4.3). Определим сначала стохастический интеграл

$$\int_0^T f(t)\,dw(t)$$

в том случае, когда $w(t)$ — одномерный винеровский процесс.

Пусть на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ (σ-алгебра $\mathfrak{F}$ предполагается полной относительно меры P) задан поток σ-алгебр $\{\mathfrak{F}_t, t \in [0, T]\}$, т. е. семейство σ-алгебр, подчиненное условиям (см. п. 9.4.1):

а) при $s < t$ $\mathfrak{F}_s \subset \mathfrak{F}_t \subset \mathfrak{F}$;

б) σ-алгебра $\mathfrak{F}_0$ содержит все события вероятности 0,

и винеровский процесс $w(t)$, $t \in [0, T]$, со значениями в $\mathbf{R}$, такой, что $w(0) = 0$; при каждом $t \in [0, T]$ процесс $w(t)$ $\mathfrak{F}_t$-измерим, а приращения $w(t+s) - w(t)$ при $s > 0$ не зависят от σ-алгебры $\mathfrak{F}_t$.

Обозначим через $H_2[0, T]$ пространство всех прогрессивно измеримых (относительно потока $\{\mathfrak{F}_t, t \in [0, T]\}$) случайных функций $f(t) = f(t, \omega)$ со значениями в $\mathbf{R}$, заданных на $[0, T]$ и таких, что с вероятностью 1 конечен интеграл $\int_0^T f^2(t)\,dt$.

Теорема 1. *Каждому процессу $\{f(t), t \in [0, T]\}$ из пространства $H_2[0, T]$ можно поставить в соответствие некоторую случайную величину $I_T^0(f)$, определенную на пространстве $(\Omega, \mathfrak{F})$ и обладающую следующими свойствами:*

1) *если $f_1, f_2 \in H_2[0, T]$, а α_1 и α_2 — произвольные постоянные, то почти наверное*

$$I_T^0(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 I_T^0(f_1) + \alpha_2 I_T^0(f_2);$$

2) *если $\chi_{[t_1,\ t_2]}(t)$ — индикатор отрезка $[t_1, t_2]$, то почти наверное*

$$I_T^0(\chi_{[t_1,\ t_2]}) = w(t_2) - w(t_1);$$

3) *если $f \in H_2[0, T]$ и $\mathsf{M}\int_0^T f^2(t)\,dt < \infty$, то*

$$\mathsf{M} I_T^0(f) = 0, \quad \mathsf{M}\left(I_T^0(f)\right)^2 = \mathsf{M}\int_0^T f^2(t)\,dt;$$

4) *каковы бы ни были* $f \in H_2[0, T]$ *и постоянные* $C > 0$ *и* $N > 0$, *имеет место неравенство*

$$\mathsf{P}\{|I_T^0(f)| > C\} \leqslant \mathsf{P}\left\{\int_0^T f^2(t)\,dt > N\right\} + \frac{N}{C^2}.$$

Случайная величина $I_T^0(f)$ называется *стохастическим интегралом* от функции $f(t)$ по винеровскому процессу и обозначается

$$I_T^0(f) = \int_0^T f(t)\,dw(t).$$

Назовем функцию $f \in H_2[0, T]$ *ступенчатой*, если существует такое разбиение отрезка $[0, T]$ точками $0 = t_0 < t_1 < \ldots < t_n = T$, что $f(t) = f(t_k)$ при $t_k \leqslant t < t_{k+1}$ $(k = 0, 1, \ldots, n-1)$. Очевидно, для ступенчатых функций

$$\int_0^T f(t)\,dw(t) = \sum_{k=0}^{n-1} f(t_k)\,[w(t_{k+1}) - w(t_k)].$$

Если $\{f_n(t),\ t \in [0, T]\}$ $(n = 1, 2, \ldots)$ — последовательность ступенчатых функций, для которых при любом $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\int_0^T |f_n(t) - f(t)|^2\,dt > \varepsilon\right\} = 0,$$

где $f(t)$ — некоторая функция из $H_2[0, T]$, то, согласно условию 4),

$$\mathsf{P}\left\{\left|\int_0^T f_n(t)\,dw(t) - \int_0^T f_r(t)\,dw(t)\right| > \varepsilon\right\} \leqslant$$

$$\leqslant \rho + \mathsf{P}\left\{\int_0^T |f_n(t) - f_r(t)|^2\,dt > \rho\varepsilon^2\right\},$$

откуда следует фундаментальность последовательности случайных величин $\int_0^T f_n(t)\,dw(t)$ в смысле сходимости по вероятности. Предел этой последовательности и есть интеграл $\int_0^T f(t)\,dw(t)$.

В том случае, когда $f \in H_2[0, T]$ и $\mathsf{M}\int\limits_0^T f^2(t)\,dt < \infty$, существует последовательность ступенчатых функций $f_n(t) \in H_2[0, T]$ таких, что

$$\lim_{n\to\infty} \mathsf{M}\int\limits_0^T |f_n(t) - f(t)|^2\,dt = 0.$$

Из свойства 3) следует равенство

$$\mathsf{M}\left[\int\limits_0^T f_n(t)\,dw(t) - \int\limits_0^T f_r(t)\,dw(t)\right]^2 = \mathsf{M}\int\limits_0^T |f_n(t) - f_r(t)|^2\,dt,$$

которое означает, что последовательность случайных величин $\int\limits_0^T f_n(t)\,dw(t)$ фундаментальна в смысле сходимости в среднеквадратическом, и, стало быть, в этом случае

$$\int\limits_0^T f(t)\,dw(t) = \underset{n\to 0}{\mathrm{l.i.m.}}\int\limits_0^T f_n(t)\,dw(t).$$

Если процесс $f \in H_2[0, T]$ с вероятностью 1 непрерывен, то

$$\int\limits_0^T f(t)\,dw(t) = \lim_{\max\limits_k \Delta t_k \to 0} \sum_{k=0}^{n-1} f(t_k)\,[w(t_{k+1}) - w(t_k)],$$

где $0 = t_0 < t_1 < \ldots < t_n = T$, $\Delta t_k = t_{k+1} - t_k$.

19.2.2. Оценки моментов. Для оценок моментов стохастических интегралов полезна следующая теорема.

Теорема 2. *Если $f \in H_2[0, T]$ и при некотором $p > 0$*

$$\mathsf{M}\left(\int\limits_0^T |f(t)|^2\,dt\right)^{p/2} < \infty,$$

то имеют место неравенства:

$$\mathsf{M}\left|\int\limits_0^T f(t)\,dw(t)\right|^p \geqslant A_p\mathsf{M}\left\{\int\limits_0^T |f(t)|^2\,dt\right\}^{p/2},$$

если $p > 1$, и

$$\mathsf{M}\left|\int\limits_0^T f(t)\,dw(t)\right|^p \leqslant B_p\mathsf{M}\left(\int\limits_0^T |f(t)|^2\,dt\right)^{p/2},$$

если $p > 0$. Здесь A_p и B_p — постоянные, зависящие лишь от p.

При $p=2$ оба неравенства превращаются в равенства, причем $A_2 = B_2 = 1$.

Отметим еще одно равенство для стохастических интегралов, следующее из определения: если $f_1, f_2 \in H_2[0, T]$ и $\mathsf{M}\int\limits_0^T f_1^2(t)\,dt < \infty$, $\mathsf{M}\int\limits_0^T f_2^2(t)\,dt < \infty$, то

$$\mathsf{M}\int\limits_0^T f_1(t)\,dw(t)\int\limits_0^T f_2(t)\,dw(t) = \mathsf{M}\int\limits_0^T f_1(t) f_2(t)\,dt.$$

19.2.3. Стохастический интеграл как функция верхнего предела. Для $f \in H_2[0, T]$ и $0 \leqslant t_1 \leqslant t_2 \leqslant T$ положим

$$\int\limits_{t_1}^{t_2} f(t)\,dw(t) = \int\limits_0^T \chi_{[t_1,\ t_2]}(t) f(t)\,dw(t),$$

где $\chi_{[t_1,\ t_2]}(t)$ — индикатор отрезка $[t_1, t_2]$.

Можно показать, что если $f \in H_2[0, T]$ и $\mathsf{M}\int\limits_0^T f^2(t)\,dt < \infty$, то с вероятностью 1

$$\mathsf{M}\left\{\int\limits_{t_1}^{t_2} f(t)\,dw(t)\,|\,\mathfrak{F}_{t_1}\right\} = 0,$$

$$\mathsf{M}\left\{\left(\int\limits_{t_1}^{t_2} f(t)\,dw(t)\right)^2 \middle|\, \mathfrak{F}_{t_1}\right\} = \int\limits_{t_1}^{t_2} \mathsf{M}\left(f^2(t)\,|\,\mathfrak{F}_{t_1}\right)dt,$$

где $0 \leqslant t_1 \leqslant t_2 \leqslant T$.

Рассмотрим процесс

$$I_t(f) = \int\limits_0^t f(s)\,dw(s), \quad t \in [0, T],$$

где $f \in H_2[0, T]$. При каждом t этот процесс определен лишь при почти всех ω, т. е. с точностью до стохастической эквивалентности. Мы будем считать, что среди всех стохастически эквивалентных процессов в качестве $I_t(f)$ выбран сепарабельный процесс. Тогда можно показать, что процесс $\{I_t(f),\ t \in [0, T]\}$ непрерывен с вероятностью 1 и имеет место неравенство

$$\mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T}\left|\int\limits_0^t f(s)\,dw(s)\right| > C\right\} \leqslant \frac{N}{C^2} + \mathsf{P}\left\{\int\limits_0^T f^2(s)\,ds > N\right\}.$$

Если $f \in H_2[0, T]$ и $\mathsf{M}\int\limits_0^T f^2(t)\,dt < \infty$, то процесс $(I_t(f), \mathfrak{F}_t)$ $(t \in [0, T])$ представляет собой непрерывный мартингал с интегрируемым квадратом, характеристика которого определяется формулой

$$\langle I(f)\rangle_t = \int\limits_0^t f^2(s)\,ds.$$

При этом выполнены неравенства

$$\mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T}\left|\int\limits_0^t f(s)\,dw(s)\right| > C\right\} \leqslant \frac{1}{C^2}\mathsf{M}\int\limits_0^T f^2(s)\,ds,$$

$$\mathsf{M}\sup_{0 \leqslant t \leqslant T}\left|\int\limits_0^t f(s)\,dw(s)\right|^2 \leqslant 4\mathsf{M}\int\limits_0^T f^2(s)\,ds.$$

19.2.4. Формула Ито. Предположим, что некоторый процесс $\{\zeta(t),\ t \in [0, T]\}$, согласованный с потоком σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$, представим при всех $0 \leqslant t_1 < t_2 \leqslant T$ в виде

$$\zeta(t_2) - \zeta(t_1) = \int\limits_{t_1}^{t_2} a(t)\,dt + \int\limits_{t_1}^{t_2} b(t)\,dw(t),$$

где $b \in H_2[0, T]$, а прогрессивно измеримый относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$ процесс $a(t)$ таков, что

$$\mathsf{P}\left\{\int\limits_0^T |a(t)|\,dt < \infty\right\} = 1.$$

Тогда будем говорить, что процесс $\zeta(t)$ имеет *стохастический дифференциал* на $[0, T]$:

$$d\zeta(t) = a(t)\,dt + b(t)\,dw(t).$$

Очевидно, что если $\zeta_1(t)$ и $\zeta_2(t)$ — два процесса, имеющие стохастические дифференциалы, а α_1 и α_2 — произвольные постоянные, то

$$d(\alpha_1\zeta_1(t) + \alpha_2\zeta_2(t)) = \alpha_1\,d\zeta_1(t) + \alpha_2\,d\zeta_2(t),$$

т. е. операция дифференцирования линейна. Приведем теперь формулы дифференцирования произведения двух процессов, а также формулу дифференцирования сложной функции.

Теорема 3. *Если процессы $\zeta_1(t)$ и $\zeta_2(t)$ имеют стохастические дифференциалы*

$$d\zeta_1(t) = a_1(t)\,dt + b_1(t)\,dw(t), \quad d\zeta_2(t) = a_2(t)\,dt + b_2(t)\,dw(t),$$

то процесс $\zeta_1(t)\zeta_2(t)$ также имеет стохастический дифференциал и

$$d(\zeta_1(t)\,\zeta_2(t)) = \zeta_1(t)\,d\zeta_2(t) + \zeta_2(t)\,d\zeta_1(t) + b_1(t)\,b_2(t)\,dt.$$

Теорема 4. *Если процесс* $\xi(t)$ *имеет стохастический дифференциал*

$$d\xi(t) = a(t)\,dt + b(t)\,dw(t),$$

а непрерывная функция $f(t, x)$ $(t \in [0, T],\ x \in \mathbf{R})$ *с вещественными значениями имеет непрерывные производные* $f'_t(t, x)$, $f'_x(t, x)$ *и* $f''_{xx}(t, x)$, *то процесс* $f(t, \xi(t))$ *также имеет стохастический дифференциал*

$$df(t, \xi(t)) = \Big[f'_t(t, \xi(t)) + f'_x(t, \xi(t))\,a(t) + \frac{1}{2} f''_{xx}(t, \xi(t))\,b^2(t)\Big]dt + f'_x(t, \xi(t))\,dw(t).$$

Последняя формула называется *формулой Ито*. Предположим теперь, что процессы $\xi_1(t), \xi_2(t), \ldots, \xi_l(t)$ имеют стохастические дифференциалы

$$d\xi_i(t) = a_i(t)\,dt + b_i(t)\,dw(t), \quad i = 1, 2, \ldots, l,$$

а непрерывная вещественная функция $f(t, x_1, \ldots, x_l)$ $(t \in [0, T],$ $x_1, \ldots, x_l \in \mathbf{R})$ имеет непрерывные частные производные

$$f'_t,\ f'_{x_i},\ f''_{x_i x_k}, \quad k, i = 1, \ldots, l.$$

Тогда процесс $f(t, \xi_1(t), \xi_2(t), \ldots, \xi_l(t))$ также имеет стохастический дифференциал, причем

$$df(t, \xi_1(t), \xi_2(t), \ldots, \xi_l(t)) = \Big[\frac{\partial f}{\partial t}(t, \xi_1(t), \ldots, \xi_l(t)) + \sum_{i=1}^{l} f'_{x_i}(t, \xi_1(t), \xi_2(t), \ldots, \xi_l(t))\,a_i(t) + \frac{1}{2}\sum_{i,\,j=1}^{l} f''_{x_i x_j}(t, \xi_1(t), \ldots, \xi_l(t))\,b_i(t)\,b_j(t)\Big]dt + \sum_{i=1}^{l} f'_{x_i}(t, \xi_1(t), \ldots, \xi_l(t))\,b_i(t)\,dw(t).$$

Эта формула также называется формулой Ито.

Пример. Положив $f(x) = x^2$, из формулы Ито получаем соотношение

$$\int_{t_1}^{t_2} w(t)\,dw(t) = \frac{1}{2}(w(t_2))^2 - \frac{1}{2}(w(t_1))^2 - \frac{1}{2}(t_2 - t_1), \quad 0 \leqslant t_1 \leqslant t_2.$$

19.2.5. Моментные тождества. Определим функции

$$G_n(t, x) = t^{n/2}\,\mathrm{He}_n(t^{-1/2}x), \quad t \geqslant 0, \quad x \in \mathbf{R}, \quad n = 0, 1, 2, \ldots,$$

где $\mathrm{He}_n(z)$ — полиномы Эрмита:

$$\mathrm{He}_n(z) = (-1)^n \exp\left\{\frac{z^2}{2}\right\} \frac{d^n}{dz^n} \exp\left\{-\frac{z^2}{2}\right\}.$$

Выпишем для примера пять первых функций $G_k(t, x)$:

$$G_0(t, x) = 1, \quad G_1(t, x) = x, \quad G_2(t, x) = x^2 - t,$$
$$G_3(t, x) = x^3 - 3tx, \quad G_4(t, x) = x^4 - 6x^2t + 3t^2.$$

Следствием формулы Ито является следующая теорема.

Теорема 5. *Если $f \in H_2[0, T]$ и при некотором натуральном n выполнено условие*

$$\mathrm{M}\left(\int_0^T f^2(t)\,dt\right)^{n/2} < \infty,$$

то при любых вещественных α и β процесс

$$\left\{G_n\left(\alpha + \int_0^t f^2(s)\,ds,\ \beta + \int_0^t f(s)\,dw(s)\right),\ \mathfrak{F}_t,\quad t \in [0, T]\right\}$$

является мартингалом и, в частности,

$$\mathrm{M}G_n\left(\alpha + \int_0^T f^2(s)\,ds,\ \beta + \int_0^T f(s)\,dw(s)\right) = G_n(\alpha, \beta).$$

Из этой теоремы (при $n = 1$) следует, что при условии

$$\mathrm{M}\left(\int_0^T f^2(t)\,dt\right)^{1/2} < \infty$$

процесс

$$\left\{\int_0^t f(s)\,dw(s),\ \mathfrak{F}_t,\quad t \in [0, T]\right\}$$

является мартингалом (возможно, не имеющим второго момента), и, в частности, при этом условии имеет место равенство

$$\mathrm{M}\int_0^T f(s)\,dw(s) = 0.$$

В качестве следствия этого утверждения нетрудно получить следующее свойство винеровского процесса.

Если τ — марковский момент относительно винеровского процесса $\{w(t),\ t \geqslant 0\}$, то при условии $\mathrm{M}\tau^{1/2} < \infty$ справедливо равенство $\mathrm{M}w(\tau) = 0$. Что это не всегда так, показывает пример марковского момента $\tau_1 = \inf\{t\colon w(t) \geqslant 1\}$, для которого $w(\tau_1) = 1$ и потому $\mathrm{M}w(\tau_1) = 1$. Заметим, что $\mathrm{M}\tau_1^{1/2} = \infty$, хотя при любом $\varepsilon \in (0, 1/2)$ $\mathrm{M}\tau_1^{1/2-\varepsilon} < \infty$.

19.2.6. Стохастические интегралы по многомерному винеровскому процессу. Пусть на некотором вероятностном пространстве с потоком σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$, удовлетворяющим условиям п. 19.2.1, заданы m независимых между собой одномерных винеровских процессов $w^1(t), w^2(t), \ldots, w^m(t)$ таких, что $w^k(0) = 0$, все они согласованы с потоком σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$, а приращения $w^k(t+s) - w^k(t)$ $(s \geqslant 0,\ k = 1, \ldots, m)$ не зависят от σ-алгебры $\mathfrak{F}_t$. Через $w(t)$ будем обозначать m-мерный винеровский процесс $(w^1(t), w^2(t), \ldots, w^m(t))$.

Далее, пусть $f(t)$ $(t \in [0, T])$ — матричнозначный случайный процесс. Элементы матрицы $f(t)$ будем обозначать через $f^{ij}(t)$ $(i = 1, 2, \ldots, l,\ j = 1, 2, \ldots, m)$. Предположим, что при всех i и j $f^{ij} \in H_2[0, T]$. Это означает, что процессы $f^{ij}(t)$ прогрессивно измеримы относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$, и

$$\mathsf{P}\left\{\int_0^T (f^{ij}(t))^2\,dt < \infty\right\} = 1, \quad i = 1, \ldots, l,\ j = 1, \ldots, m.$$

Мы сохраним обозначение $H_2[0, T]$ для совокупности матриц $f(t)$, элементы которых удовлетворяют перечисленным условиям. Для следа матрицы $f(t)f^T(t)$ имеем формулу

$$\mathrm{Sp}\left(f(t)\,f^T(t)\right) = \sum_{i=1}^{l}\ \sum_{j=1}^{m} (f^{ij}(t))^2,$$

где $f^T(t)$ — транспонированная матрица. В частности, если $l = 1$, то $f(t)$ — вектор с координатами $(f^1(t), \ldots, f^m(t))$. В этом случае

$$\mathrm{Sp}\left(f(t)\,f^T(t)\right) = \sum_{j=1}^{m} (f^j(t))^2 = |f(t)|^2.$$

Определим теперь для $f \in H_2[0, T]$ стохастический интеграл

$$\int_0^T f(t)\,dw(t)$$

как случайный вектор с координатами

$$\sum_{j=1}^{m}\int_0^T f^{ij}(t)\,dw^j(t), \quad i = 1, 2, \ldots, l.$$

Если $l = 1$, то это — скалярная случайная величина $\sum_{j=1}^{m}\int_0^T f^j(t)\,dw^j(t)$, которую будем обозначать также как $\int_0^T (f(t), dw(t))$. Здесь $f^j(t)$ $(j = 1, 2, \ldots, m)$ — координаты вектора $f(t)$.

Отметим следующие свойства стохастических интегралов по многомерному винеровскому процессу.

1) Стохастический интеграл представляет собой линейную функцию от процесса $f(t)$.

2) Если $f \in H_2[0, T]$ и $\mathsf{M}\int_0^T \operatorname{Sp}(f(t) f^T(t))\,dt < \infty$,

то

$$\mathsf{M}\int_0^T f(t)\,dw(t) = 0, \quad \mathsf{M}\left|\int_0^T f(t)\,dw(t)\right|^2 = \mathsf{M}\int_0^T \operatorname{Sp}(f(t) f^T(t))\,dt.$$

3) Если $f(t)$ и $g(t)$ — два матричных процесса порядка $l \times m$ из пространства $H_2[0, T]$, для которых

$$\mathsf{M}\int_0^T \operatorname{Sp}(f(t) g^T(t))\,dt < \infty,$$

то

$$\mathsf{M}\left(\int_0^T f(t)\,dw(t), \quad \int_0^T g(t)\,dw(t)\right) = \mathsf{M}\int_0^T \operatorname{Sp}(f(t) g^T(t))\,dt.$$

В частности, при $l = 1$

$$\mathsf{M}\left[\int_0^T (f(t), dw(t)) \int_0^T (g(t), dw(t))\right] = \mathsf{M}\int_0^T (f(t), g(t))\,dt,$$

если только конечна величина, стоящая в правой части этого равенства. При $f = g$ отсюда получаем равенство

$$\mathsf{M}\left[\int_0^T (f(t), dw(t))\right]^2 = \mathsf{M}\int_0^T |f(t)|^2\,dt.$$

4) Сепарабельная модификация процесса

$$\int_0^t f(s)\,dw(s), \quad t \in [0, T],$$

при любом $f \in H_2[0, T]$ представляет собой непрерывный l-мерный процесс. Если θ — произвольный (неслучайный) элемент из пространства $\mathbf{R}^l$, то, положив

$$\xi_\theta(t) = \left(\theta, \int_0^t f(s)\,dw(s)\right),$$

будем иметь неравенство

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant T}|\xi_\theta(t)|>C\right\}\leqslant\frac{N}{C^2}+\mathsf{P}\left\{\int_0^T(f(t)f^T(t)\theta,\ \theta)\,dt>N\right\},$$

где N и C — произвольные положительные постоянные.

Если $l=1$, то отсюда получаем

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant T}\left|\int_0^t(f(s),\ dw(s))\right|>C\right\}\leqslant\frac{N}{C^2}+\mathsf{P}\left\{\int_0^T|f(t)|^2\,dt>N\right\}.$$

5) Если $f\in H_2[0,\ T]$ и $\mathsf{M}\int\limits_0^T\mathrm{Sp}\left(f(t)f^T(t)\right)dt<\infty$,

то процесс

$$\left(\int_0^t f(s)\,dw(s),\ \mathfrak{F}_t\right),\quad t\in[0,\ T],$$

является l-мерным непрерывным мартингалом с интегрируемым квадратом. Его характеристика, представляющая собой матричнозначный процесс порядка $l\times l$, определяется интегралом

$$\int_0^t f(s)f^T(s)\,ds,\quad t\in[0,\ T].$$

Кроме того, для произвольного $\theta\in\mathbf{R}^l$ выполнены неравенства

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant T}|\xi_\theta(t)|>C\right\}\leqslant\frac{1}{C^2}\mathsf{M}\int_0^T\left(f(t)f^T(t)\theta,\ \theta\right)dt,$$

$$\mathsf{M}\sup_{0\leqslant t\leqslant T}|\xi_\theta(t)|^2\leqslant 4\mathsf{M}\int_0^T\left(f(t)f^T(t)\theta,\ \theta\right)dt,$$

где $\xi_\theta(t)$ определено в свойстве 4).

В частности, при $l=1$ характеристика мартингала

$$\int_0^t(f(s),\ dw(s)),\quad t\in[0,\ T],$$

равна

$$\int_0^t|f(s)|^2\,ds,\quad t\in[0,\ T].$$

6) Если $f \in H_2[0, T]$ и при некотором $p > 0$

$$\mathsf{M}\left[\int_0^T \operatorname{Sp}(f(t) f^T(t))\,dt\right]^{p/2} < \infty,$$

то имеют место неравенства:

$$\mathsf{M}\left|\int_0^T f(t)\,dw(t)\right|^p \geqslant A_p \mathsf{M}\left(\int_0^T \operatorname{Sp}(f(t) f^T(t))\,dt\right)^{p/2},$$

если $p > 1$, и

$$\mathsf{M}\left|\int_0^T f(t)\,dw(t)\right|^p \leqslant B_p \mathsf{M}\left(\int_0^T \operatorname{Sp}(f(t) f^T(t))\,dt\right)^{p/2},$$

если $p > 0$. Здесь A_p и B_p — постоянные, зависящие лишь от p.

7) Предположим, что для некоторого l-мерного процесса $\zeta(t)$ при всех $0 \leqslant t_1 < t_2 \leqslant T$ имеет место представление

$$\zeta(t_2) - \zeta(t_1) = \int_{t_1}^{t_2} a(s)\,ds + \int_{t_1}^{t_2} b(s)\,dw(s),$$

где $b(t)$ — матричнозначный процесс порядка $l \times m$ из пространства $H_2[0, T]$, а l-мерный процесс $a(t)$ таков, что все его координаты прогрессивно измеримы относительно потока σ-алгебр $\{\mathfrak{F}_t, t \in [0, T]\}$ и интегрируемы с вероятностью 1 на отрезке $[0, T]$. Тогда будем говорить, что процесс $\zeta(t)$ имеет стохастический дифференциал

$$d\zeta(t) = a(t)\,dt + b(t)\,dw(t).$$

Если процесс $\zeta(t)$ имеет указанный стохастический дифференциал, а непрерывная вещественная функция $f(t, x)$ $(t \in [0, T],$ $x \in \mathbf{R}^l)$ имеет непрерывные частные производные

$$f'_t(t, x),\quad f'_{x^i}(t, x),\quad f''_{x^i x^j}(t, x),\qquad i, j = 1, \ldots, l,$$

то процесс $f(t, \zeta(t))$ $(t \in [0, T])$ также имеет стохастический дифференциал и

$$df(t, \zeta(t)) = \Big[f'_t(t, \zeta(t)) + (a(t),\ f'_x(t, \zeta(t))) + \frac{1}{2}\operatorname{Sp}\left(b(t)\,b^T(t)\,f''_{xx}(t, \zeta(t))\right)\Big]dt + \left(b^T(t)\,f'_x(t, \zeta(t)),\ dw(t)\right),$$

где f'_x — вектор с координатами $f'_{x^1}, f'_{x^2}, \ldots, f'_{x^l}$, а f''_{xx} — матрица с элементами $f''_{x^i x^j}$ $(i, j = 1, \ldots, l)$.

Приведенная формула также называется *формулой Ито*. Запишем ее в развернутом виде:

$$df(t, \zeta(t)) = \left[f'_t(t, \zeta(t)) + \sum_{i=1}^{l} a^i(t) f'_{x^i}(t, \zeta(t)) + \right.$$
$$\left. + \frac{1}{2} \sum_{k,\, i=1}^{l} \sum_{j=1}^{m} b^{ij}(t) b^{kj}(t) f''_{x^i x^k}(t, \zeta(t)) \right] dt +$$
$$+ \sum_{k=1}^{l} \sum_{j=1}^{m} f'_{x^k}(t, \zeta(t)) b^{kj}(t) dw^j(t).$$

19.3. Стохастические дифференциальные уравнения для непрерывных процессов

19.3.1. Теорема существования и единственности решения. Пусть заданы:

1) вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ с потоком σ-алгебр $\{\mathfrak{F}_t, t \in [0, T]\}$ (σ-алгебра $\mathfrak{F}$ предполагается полной относительно меры P, а σ-алгебра $\mathfrak{F}_0$ — содержащей все события вероятности 0);

2) согласованный с потоком σ-алгебр $\{\mathfrak{F}_t, t \in [0, T]\}$ m-мерный винеровский процесс $w(t) = \{w^1(t), \ldots, w^m(t)\}$ (это означает, что $w(0) = 0$; при всех $t \in [0, T]$ величина $w(t)$ $\mathfrak{F}_t$-измерима и приращения $w(t+s) - w(t)$ при $s \geqslant 0$ не зависят от σ-алгебры $\mathfrak{F}_t$);

3) $\mathfrak{F}_0$-измеримый случайный вектор ξ_0 (заметим, что в силу 2) σ-алгебра $\mathfrak{F}_0$, а стало быть, и вектор ξ_0 не зависят от процесса w_t);

4) функции $a(t, x)$ и $\sigma(t, x)$ ($t \in [0, T]$, $x \in \mathbf{R}^m$), принимающие значения в $\mathbf{R}^m$ и $L(\mathbf{R}^m)$ соответственно, где $L(\mathbf{R}^m)$ — совокупность всех линейных операторов, действующих из $\mathbf{R}^m$ в $\mathbf{R}^m$ (предполагается, что $a(t, x)$ и $\sigma(t, x)$ измеримы по совокупности переменных).

Будем рассматривать стохастическое дифференциальное уравнение

$$d\xi(t) = a(t, \xi(t))\, dt + \sigma(t, \xi(t))\, dw(t) \tag{3.1}$$

с начальным условием

$$\xi(0) = \xi_0.$$

Это уравнение может быть записано в интегральной форме:

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(s))\, ds + \int_0^t \sigma(s, \xi(s))\, dw(s), \quad t \in [0, T].$$

Здесь $\xi(t)$ — искомый процесс. Приведем точное определение того, что мы будем понимать под решением уравнения (3.1).

Решением уравнения (3.1) с начальным условием $\xi(0) = \xi_0$ называется m-мерный процесс $\{\xi(t), t \in [0, T]\}$ такой, что:

а) процесс $\xi(t)$ прогрессивно измерим относительно потока σ-алгебр $\{\mathfrak{F}_t, t \in [0, T]\}$;

б) все координаты векторного процесса $\{a(t, \xi(t)), t \in [0, T]\}$ абсолютно интегрируемы на отрезке $[0, T]$ с вероятностью 1;

в) все элементы матричного процесса $\{\sigma(t, \xi(t)), t \in [0, T]\}$ интегрируемы с квадратом на отрезке $[0, T]$ с вероятностью 1;

г) процесс $\xi(t)$ имеет стохастический дифференциал, причем $d\xi(t) = a(t, \xi(t))dt + \sigma(t, \xi(t))dw(t)$ и $\xi(0) = \xi_0$.

Заметим, что в случае, когда $\sigma(t, x) \equiv 0$, уравнение (3.1) превращается в обыкновенное дифференциальное уравнение, правда, со случайным начальным условием. Такое уравнение можно решать обычными методами при каждом ω.

Говорят, что уравнение (3.1) имеет единственное решение, если для любых двух его решений $\xi_1(t)$ и $\xi_2(t)$ выполнено условие

$$\mathsf{P}\left\{\sup_{0 \leqslant t \leqslant T} |\xi_1(t) - \xi_2(t)| > 0\right\} = 0.$$

Приведем теорему существования и единственности решения уравнения (3.1).

Теорема 1. *Предположим, что коэффициенты уравнения* (3.1) *удовлетворяют условиям:*

1) *при всех* $t \in [0, T]$, $x \in \mathbb{R}^m$ *выполнено неравенство*

$$|a(t, x)|^2 + |\sigma(t, x)|^2 \leqslant K(1 + |x|^2),$$

где K — *некоторая постоянная,* $|\sigma(t, x)|^2 = \sum_{j, k=1}^{m} (\sigma^{jk}(t, x))^2$, $\sigma^{jk}(t, x)$ — *элементы матрицы* $\sigma(t, x)$;

2) *для любого* $R > 0$ *найдется постоянная* C_R *такая, что при* $|x| \leqslant R$, $|y| \leqslant R$ *и* $t \in [0, T]$ *выполнено неравенство*

$$|a(t, x) - a(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leqslant C_R |x - y|^2.$$

Тогда существует единственное непрерывное решение $\xi(t)$ $(t \in [0, T])$ *уравнения* (3.1).

Замечание 1. Пусть заданы функции $a_i(t, x)$, $\sigma_i(t, x)$ $(i = 1, 2)$, удовлетворяющие условиям теоремы 1, такие, что для некоторого $N > 0$ при $|x| \leqslant N$ и $t \in [0, T]$ имеют место равенства $a_1(t, x) = a_2(t, x)$ и $\sigma_1(t, x) = \sigma_2(t, x)$. Обозначим через $\xi_i(t)$ $(i = 1, 2)$ решение уравнения

$$d\xi_i(t) = a_i(t, \xi_i(t))\,dt + \sigma_i(t, \xi_i(t))\,dw(t)$$

с одним и тем же начальным условием $\xi_i(0) = \xi_0$ $(i = 1, 2)$. Далее, положим

$$\tau_i = \inf\{t\colon |\xi_i(t)| \geqslant N\}, \quad i = 1, 2,$$

причем, если множество в фигурных скобках пусто, полагаем $\tau_i = T$. Тогда можно показать, что $\mathsf{P}\{\tau_1 = \tau_2\} = 1$ и

$$\mathsf{P}\left\{\sup_{0 \leqslant s \leqslant \tau_1} |\xi_1(s) - \xi_2(s)| > 0\right\} = 0.$$

Это свойство решений уравнения (3.1) характеризует так называемую локальную зависимость решения от коэффициентов уравнения.

Замечание 2. Решение уравнения (3.1) может быть построено следующим образом.

Пусть сначала коэффициенты $a(t, x)$ и $\sigma(t, x)$ удовлетворяют условию 1) теоремы 1, а также условию:

2′) существует постоянная $C > 0$ такая, что при всех $t \in [0, T]$, $x, y \in R^m$

$$|a(t, x) - a(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leqslant C|x - y|^2.$$

Предположим, кроме того, что $M|\xi_0|^2 < \infty$. Тогда, положив $\eta_0(t) = \xi_0$,

$$\eta_{n+1}(t) = \xi_0 + \int_0^t a(s, \eta_n(s))\,ds + \int_0^t \sigma(s, \eta_n(s))\,dw(s), \quad n = 0, 1, 2, \ldots,$$

с помощью несложных рассуждений придем к неравенствам

$$M \sup_{0 \leqslant t \leqslant T} |\eta_{n+1}(t) - \eta_n(t)|^2 \leqslant \frac{L^n T^n}{n!}, \quad n = 0, 1, 2, \ldots,$$

где L — некоторая постоянная. Эти неравенства позволяют заключить, что ряд

$$\eta_0(t) + \sum_{n=0}^{\infty} (\eta_{n+1}(t) - \eta_n(t))$$

сходится равномерно с вероятностью 1 и представляет собой решение уравнения (3.1) в том случае, когда его коэффициенты удовлетворяют условиям 1), 2), а начальное условие таково, что $M|\xi_0|^2 < \infty$.

В общем случае можно указать последовательность коэффициентов $a_n(t, x)$, $\sigma_n(t, x)$ $(n = 1, 2, \ldots)$ таких, что при $|x| \leqslant n$, $t \in [0, T]$ имеем $a_n(t, x) = a(t, x)$, $\sigma_n(t, x) = \sigma(t, x)$, причем функции $a_n(t, x)$ и $\sigma_n(t, x)$ удовлетворяют условию 2′) с некоторой постоянной C_n, возможно растущей с ростом n. Полагая еще $\xi_0^{(n)} = \xi_0$, если $|\xi_0| \leqslant n$, и $\xi_0^{(n)} = 0$, если $|\xi_0| > n$, получим решение $\xi^{(n)}(t)$ уравнения (3.1) с коэффициентами $a_n(t, x)$, $\sigma_n(t, x)$ и начальным условием $\xi_0^{(n)}$.

В силу замечания 1 при всех $k = 1, 2, \ldots$ $\xi^{(k+n)}(t) = \xi^{(n)}(t)$ для $t \leqslant \tau^n$, где $\tau^n = \inf\{t\colon |\xi^{(n)}(t)| \geqslant n\}$ (если множество в фигурных скобках пусто, то полагаем $\tau^n = T$).

Теперь уж совсем несложно доказать, что при $n \to \infty$ процессы $\xi^{(n)}(t)$ сходятся с вероятностью 1 равномерно к некоторому процессу $\xi(t)$, который и представляет собой решение уравнения (3.1).

Из этого построения и из единственности решения уравнения (3.1) следует, что в условиях теоремы 1 решение $\xi(t)$ измеримо относительно минимальной σ-алгебры событий, порожденной случайным вектором ξ_0 и значениями процесса $w(s)$ при $s \leqslant t$.

Замечание 3. Пусть коэффициенты $a(t, x)$ и $\sigma(t, x)$ заданы при $t \in [0, \infty)$, $x \in R^m$, и предположим, что для всякого $R > 0$ существует постоянная C_R такая, что при $t \leqslant R$, $|x| \leqslant R$, $|y| \leqslant R$

$$|a(t, x) - a(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leqslant C_R|x - y|^2.$$

Тогда при каждом $R > 0$ можно построить функции $a_R(t, x)$ и $\sigma_R(t, x)$, совпадающие при $t \leqslant R$, $|x| \leqslant R$ с функциями $a(t, x)$ и

$\sigma(t, x)$ соответственно и такие, что $a_R(t, x)$ и $\sigma_R(t, x)$ при всех $t \leqslant R$, $x \in \mathbf{R}^m$ удовлетворяют условию 1) теоремы 1 с некоторой постоянной K_R.

Обозначим через $\xi_R(t)$, $t \in [0, R]$ решение уравнения (3.1) с коэффициентами a_R, σ_R и заданным начальным условием ξ_0. Определим случайную величину τ_R, полагая

$$\tau_R = \inf\{t \leqslant R\colon |\xi_R(t)| \geqslant R\},$$

если множество в фигурных скобках непусто. В противном случае полагаем $\tau_R = R$. Как следует из замечания 1, при всех $R' > R$ с вероятностью 1 $\xi_R(t) = \xi_{R'}(t)$ для $t \leqslant \tau_R$.

Положим $\xi(t) = \lim\limits_{R \to \infty} \xi_R(t)$. Если коэффициенты $a(t, x)$ и $\sigma(t, x)$ удовлетворяют неравенству

$$|a(t, x)|^2 + |\sigma(t, x)|^2 \leqslant K(1 + |x|^2)$$

при всех $t \geqslant 0$, $x \in \mathbf{R}^m$ (K — некоторая постоянная), то можно показать, что при $R \to \infty$ с вероятностью 1 $\tau_R \to +\infty$, и, стало быть, решение $\xi(t)$ уравнения (3.1) определено при всех $t \geqslant 0$. Если же это условие не выполняется, то, вообще говоря, нельзя утверждать, что $\tau_R \to +\infty$ при $R \to \infty$ с вероятностью 1. В этом случае (единственное) решение $\xi(t)$ уравнения (3.1) существует лишь до момента времени

$$\zeta = \lim_{R \to \infty} \tau_R,$$

который называется моментом взрыва. Очевидно, что

$$\sup_{0 \leqslant t < \zeta} |\xi(t)| = \infty.$$

Сформулируем одно достаточное условие отсутствия взрыва (т. е. условие, при котором можно гарантировать, что $\zeta = +\infty$ п. н.).

Предположим, что коэффициенты $a(t, x)$ и $\sigma(t, x)$ удовлетворяют сформулированному в начале замечания 3 условию, и пусть для каждого $R > 0$ существует такая дважды непрерывно дифференцируемая функция $v_R(x)$, $x \in \mathbf{R}^m$, что:

1) $v_R(x) \to +\infty$ при $|x| \to \infty$;

2) найдется такое $\lambda_R > 0$, что при $s \leqslant R$, $x \in \mathbf{R}^m$

$$\frac{1}{2} \sum_{i, j, k=1}^{m} \sigma^{ij}(s, x)\, \sigma^{kj}(s, x) \frac{\partial^2 v_R}{\partial x^i \partial x^k} + \sum_{j=1}^{m} a^j(s, x) \frac{\partial v_R}{\partial x^j} \leqslant \lambda_R v_R(x).$$

Тогда уравнение (3.1) имеет единственное определенное при всех $t \geqslant 0$ решение $\xi(t)$, удовлетворяющее заданному начальному условию ξ_0.

4. Если выполнены условия теоремы 1 и $\mathsf{M}|\xi_0|^{2p} < \infty$ при некотором целом p, то решение уравнения (3.1) с начальным условием ξ_0 удовлетворяет условиям

$$\mathsf{M}|\xi(t)|^{2p} \leqslant K_p(1 + \mathsf{M}|\xi_0|^{2p}), \qquad t \in [0, T],$$

$$\mathsf{M}|\xi(t) - \xi_0|^{2p} \leqslant K_p'\left(1 + \mathsf{M}|\xi_0|^{2p}\right) t^p, \quad t \in [0, T],$$

где K_p, K_p' — постоянные, зависящие лишь от p, K и T.

19.3.2. Решение как диффузионный процесс. В условиях теоремы 1 решение $\xi(t)$ $(t \in [0, T])$ уравнения (3.1) обладает марковским свойством относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in [0, T]\}$. Это означает, что при любых $0 \leqslant s \leqslant t \leqslant T$ и $\Gamma \in \mathfrak{B}$, где $\mathfrak{B}$ — σ-алгебра борелевских подмножеств пространства $\mathbf{R}^m$, с вероятностью 1 выполнено соотношение

$$\mathsf{P}\{\xi(t) \in \Gamma \mid \mathfrak{F}_s\} = \mathsf{P}\{\xi(t) \in \Gamma \mid \xi(s)\}.$$

Таким образом, процесс $\xi(t)$ $(t \in [0, T])$ представляет собой марковскую случайную функцию с начальным распределением

$$\mu(\Gamma) = \mathsf{P}\{\xi_0 \in \Gamma\}, \quad \Gamma \in \mathfrak{B}.$$

Вероятность перехода $P(s, x, t, \Gamma)$ марковской случайной функции $\xi(t)$ $(t \in [0, T])$ определяется формулой

$$P(s, x, t, \Gamma) = \mathsf{P}\{\xi_{sx}(t) \in \Gamma\}, \quad 0 \leqslant s \leqslant t \leqslant T, \quad x \in \mathbf{R}^m, \quad \Gamma \in \mathfrak{B},$$

где $\xi_{sx}(t)$ — решение уравнения

$$\xi_{sx}(t) = x + \int_s^t a(\tau, \xi_{sx}(\tau))\, d\tau + \int_s^t \sigma(\tau, \xi_{sx}(\tau))\, dw(\tau). \tag{3.2}$$

Здесь x — неслучайный вектор из $\mathbf{R}^m$, $t \in [s, T]$.

Следующая теорема показывает, что решение уравнения (3.1) при некоторых условиях представляет собой диффузионный процесс (см. п. 19.1.1).

Теорема 2. *Предположим, что функции $a(t, x)$ и $\sigma(t, x)$ удовлетворяют условиям теоремы* 1 *и, кроме того, непрерывны по совокупности переменных. Тогда процесс $\xi(t)$ $(t \in [0, T])$, являющийся решением уравнения* (3.1), *представляет собой диффузионный процесс с вектором переноса $a(t, x)$ и матрицей диффузии $b(t, x) = \sigma(t, x)\sigma^T(t, x)$.*

Таким образом, теория стохастических дифференциальных уравнений дает возможность строить диффузионные процессы при весьма широких предположениях относительно коэффициентов $a(t, x)$ и $b(t, x)$. Более того, потребовав дополнительной гладкости функций $a(t, x)$ и $\sigma(t, x)$, можно доказать существование двух непрерывных производных от функции

$$u(s, x) = \mathsf{M} f(\xi_{sx}(t)), \quad 0 \leqslant s < t \leqslant T, \quad x \in \mathbf{R}^m,$$

по x и тем самым получить обратное уравнение Колмогорова.

Теорема 3. *Предположим, что функции $a(t, x)$ и $\sigma(t, x)$ удовлетворяют условиям теоремы* 1, *непрерывны и дважды непрерывно дифференцируемы по x. Пусть, кроме того, при некоторых $p > 0$ и $K > 0$ выполнено неравенство*

$$\sum_{i,k=1}^{m}\left|\frac{\partial a^i(t, x)}{\partial x^k}\right| + \sum_{i,l,k=1}^{m}\left|\frac{\partial^2 a^i(t, x)}{\partial x^l\, \partial x^k}\right| + \sum_{l,i,k=1}^{m}\left|\frac{\partial \sigma^{il}(t, x)}{\partial x^k}\right| + \sum_{i,j,k,l=1}^{m}\left|\frac{\partial^2 \sigma^{ij}(t, x)}{\partial x^k\, \partial x^l}\right| \leqslant K(1 + |x|^p).$$

Тогда если $f(x)$ $(x \in \mathrm{R}^m)$ — *дважды непрерывно дифференцируемая функция с вещественными значениями, для которой*

$$|f(x)| + \sum_{i=1}^{m} \left| \frac{\partial f(x)}{\partial x^i} \right| + \sum_{i,\,k=1}^{m} \left| \frac{\partial^2 f(x)}{\partial x^i \partial x^k} \right| \leqslant K(1 + |x|^p), \quad p > 0,$$

то функция

$$u(s, x) = \mathsf{M} f(\xi_{sx}(t)), \quad 0 \leqslant s < t \leqslant T, \quad x \in \mathrm{R}^m,$$

где $\xi_{sx}(t)$ — *решение уравнения* (3.2), *дважды непрерывно дифференцируема по* x, *непрерывно дифференцируема по* s *и удовлетворяет уравнению*

$$\frac{\partial u(s, x)}{\partial s} + \sum_{i=1}^{m} a^i(s, x) \frac{\partial u(s, x)}{\partial x^i} + \\ + \frac{1}{2} \sum_{i,\,k,\,j=1}^{m} \sigma^{ij}(s, x)\, \sigma^{kj}(s, x) \frac{\partial^2 u(s, x)}{\partial x^i \partial x^k} = 0$$

в области $s \in (0, t)$, $x \in \mathrm{R}^m$, *с граничным условием*

$$\lim_{s \uparrow t} u(s, x) = f(x).$$

Следствием этой теоремы является следующая теорема существования и единственности решения задачи Коши для уравнений в частных производных параболического типа.

Теорема 4. *Пусть в области* $0 \leqslant s < T$, $x \in \mathrm{R}^m$, *задан дифференциальный оператор*

$$\mathscr{L} u(s, x) = \frac{\partial u}{\partial s}(s, x) + \sum_{i=1}^{m} a^i(s, x) \frac{\partial u(s, x)}{\partial x^i} + \\ + \frac{1}{2} \sum_{i,\,j=1}^{m} b^{ij}(s, x) \frac{\partial^2 u(s, x)}{\partial x^i \partial x^j}$$

параболического типа (*это означает, что при всех* $s \in [0, T]$, $x \in \mathrm{R}^m$, *выполнено неравенство* $\sum_{i,\,j=1}^{m} b^{ij}(s, x)\, \theta^i \theta^j \geqslant 0$, *каковы бы ни были вещественные числа* $\theta^1, \theta^2, \ldots, \theta^m$). *Если матрица* $b(t, x)$ *с элементами* $b^{ij}(t, x)$ $(i, j = 1, 2, \ldots, m)$ *такова, что* $b(t, x) = \sigma(t, x)\sigma^T(t, x)$, *а функции* $a(t, x)$ *и* $\sigma(t, x)$ *удовлетворяют условиям теоремы* 3, *то задача Коши*

$$\mathscr{L} u(s, x) = 0, \quad \lim_{s \uparrow T} u(s, x) = f(x)$$

имеет единственное решение при любой дважды непрерывно дифференцируемой функции $f(x)$, такой, что она сама и все ее частные производные до второго порядка включительно растут на бесконечности не быстрее, чем некоторая степень $|x|$. При этом решение $u(s, x)$ указанной задачи Коши может быть записано в виде

$$u(s, x) = \mathsf{M} f(\xi_{sx}(T)), \quad 0 \leqslant s < T, \quad x \in \mathbf{R}^m,$$

где $\xi_{sx}(t)$ $(t \in [s, T])$ *— решение уравнения* (3.2).

Сформулированная теорема показывает, что при изучении уравнений в частных производных параболического типа с успехом может быть применена теория стохастических дифференциальных уравнений. Особо отметим то обстоятельство, что в теореме 4 не предполагается невырожденность матрицы $b(t, x)$, и это является существенным преимуществом методов, основанных на теории стохастических дифференциальных уравнений, перед аналитическими методами.

19.3.3. Уравнения для характеристических функций функционалов. Пусть $\{\xi(t),\ t \in [0, T]\}$ — решение уравнения (3.1). Рассмотрим функционалы

$$\int_0^T g(s, \xi(s))\, ds, \quad \int_0^T h(s, \xi(s))\, dw(s),$$

где $g(s, x)$ $(s \in [0, T],\ x \in \mathbf{R}^m)$ — функции со значениями в $\mathbf{R}^l$, а $h(s, x)$ $(s \in [0, T],\ x \in \mathbf{R}^m)$ — матричнозначная функция порядка $l \times m$. Распределения указанных функционалов будут определены, если будет найдена функция

$$u(s, x) = \mathsf{M} f(\xi_{sx}(T)) \exp\left\{i\left(\lambda, \int_s^T g(\tau, \xi_{sx}(\tau))\, d\tau\right) + i\left(\theta, \int_s^T h(\tau, \xi_{sx}(\tau))\, dw(\tau)\right)\right\},$$

где $0 \leqslant s < T$, $x \in \mathbf{R}^m$, $\lambda, \theta \in \mathbf{R}^l$, $\xi_{sx}(t)$ — решение уравнения (3.2), а $f(x)$ $(x \in \mathbf{R}^m)$ — некоторая вещественная функция.

Можно показать, что если коэффициенты уравнения (3.1) и функция $f(x)$ удовлетворяют условиям теоремы 3, а функции $g(t, x)$ и $h(t, x)$ дважды непрерывно дифференцируемы по x, причем при некоторых $p > 0$ и $K > 0$ $(t \in [0, T],\ x \in \mathbf{R}^m)$

$$\sum_{j=1}^m \sum_{k=1}^l \left|\frac{\partial g^k(t, x)}{\partial x^j}\right| + \sum_{k, j=1}^m \sum_{r=1}^l \left|\frac{\partial^2 g^r(t, x)}{\partial x^k \partial x^j}\right| + \sum_{r, k, j=1}^m \sum_{i=1}^l \left|\frac{\partial^2 h^{ir}(t, x)}{\partial x^k \partial x^j}\right| + \sum_{r, k=1}^m \sum_{i=1}^l \left|\frac{\partial h^{ir}(t, x)}{\partial x^k}\right| \leqslant K(1 + |x|^p),$$

то функция $u(s, x)$ в области $0 \leqslant s < T$, $x \in \mathbf{R}^m$, удовлетворяет уравнению

$$\frac{\partial u(s, x)}{\partial s} + \sum_{k=1}^{m} a^k(s, x) \frac{\partial u(s, x)}{\partial x^k} + \frac{1}{2} \sum_{j, k=1}^{m} b^{jk}(s, x) \frac{\partial^2 u(s, x)}{\partial x^j \partial x^k} + \\ + i \sum_{j=1}^{m} \frac{\partial u(s, x)}{\partial x^j} \sum_{k=1}^{l} d^{jk}(s, x) \theta^k + \\ + u(s, x) \left[i \sum_{k=1}^{l} \lambda^k g^k(s, x) - \frac{1}{2} \sum_{j, k}^{l} \theta^j \theta^k c^{jk}(s, x) \right] = 0,$$

где

$$b^{jk}(s, x) = \sum_{r=1}^{m} \sigma^{jr}(s, x) \sigma^{kr}(s, x), \quad j, k = 1, 2, \ldots, m,$$

$$c^{jk}(s, x) = \sum_{r=1}^{m} h^{jr}(s, x) h^{kr}(s, x), \quad j, k = 1, 2, \ldots, l,$$

$$d^{jk}(s, x) = \sum_{r=1}^{m} \sigma^{jr}(s, x) h^{kr}(s, x), \quad j = 1, 2, \ldots, m; \; k = 1, 2, \ldots, l.$$

К этому уравнению нужно присоединить граничное условие

$$\lim_{s \uparrow T} u(s, x) = f(x).$$

Уравнение для функции $u(s, x)$ может быть записано короче:

$$\frac{\partial u(s, x)}{\partial x} + \left(a(s, x), u'_x(s, x)\right) + \frac{1}{2} \operatorname{Sp}\left(\sigma(s, x) \sigma^T(s, x) u''_{xx}(s, x)\right) + \\ + i\left(\sigma(s, x) h^T(s, x) \theta, u'_x(s, x)\right) + u(s, x) \Big[i\left(g(s, x), \lambda\right) - \\ - \frac{1}{2} \left| h^T(s, x) \theta \right|^2 \Big] = 0,$$

где u'_x — вектор с координатами $\partial u / \partial x^k$ $(k = 1, 2, \ldots, m)$, u''_{xx} — матрица с элементами $\partial^2 u / \partial x^j \partial x^k$ $(j, k = 1, \ldots, m)$.

19.3.4. Мартингальная постановка задачи. Условие Липшица в теореме 1 является слишком жестким. Многие задачи приводят к необходимости рассматривать стохастические дифференциальные уравнения с коэффициентами, не удовлетворяющими этому условию. В связи с этим удобно несколько расширить и само понятие решения стохастического дифференциального уравнения, а именно можно считать, что заданы лишь коэффициенты уравнения. Требуется построить вероятностное пространство и определить на нем винеровский процесс $w(t)$ и процесс $\xi(t)$ так, чтобы они были связаны между собой уравнением (3.1) с заданными коэффициентами. Разумеется, при этом необходимо, чтобы минимальная σ-алгебра, порожденная значениями процесса $\xi(s)$ при $s \leqslant t$, и приращения $w(t+\tau) - w(t)$ при $\tau \geqslant 0$ были независимы.

Далее, если $\xi(t)$ — решение уравнения (3.1) с начальным условием ξ_0, то очевидно, что процесс $\xi(t) - \xi_0 - \int_0^t a(s, \xi(s))\,ds$ является непрерывным мартингалом с интегрируемым квадратом, характеристика которого равна $\int_0^t b(s, \xi(s))\,ds$, где $b(s, x) = \sigma(s, x)\sigma^T(s, x)$. Наоборот, если $\xi(t)$ обладает такими свойствами, то, несколько расширяя (если это нужно) вероятностное пространство, можно построить винеровский процесс $w(t)$ так, чтобы процессы $\xi(t)$ и $w(t)$ были связаны уравнением (3.1).

Таким образом, мы можем сформулировать задачу так: при каждых $s \in [0, T]$ и $x \in \mathbf{R}^m$ заданы вектор $a(s, x) \in \mathbf{R}^m$ и симметрическая неотрицательно определенная матрица $b(s, x)$ порядка $m \times m$; требуется на некотором вероятностном пространстве построить процесс $\xi(t)$ так, чтобы процесс

$$\xi(t) - \xi(0) - \int_0^t a(s, \xi(s))\,ds, \quad t \in [0, T],$$

был непрерывным мартингалом с интегрируемым квадратом и с характеристикой

$$\int_0^t b(s, \xi(s))\,ds.$$

Более того, так как искомый процесс $\xi(t)$ является непрерывным, то мы можем считать, что основное вероятностное пространство совпадает с пространством всех непрерывных функций $x(t)$, заданных на $[0, T]$ и принимающих значения в $\mathbf{R}^m$. При этом считаем, что $\xi(t) = x(t)$, и задача заключается в построении меры в пространстве непрерывных функций такой, чтобы $x(t)$ удовлетворяла указанным условиям. Сформулируем более точно задачу.

Задача. Пусть заданы: а) измеримая функция $a(s, x)$, $s \in [0, T]$, $x \in \mathbf{R}^m$, со значениями в $\mathbf{R}^m$; б) измеримая функция $b(s, x)$, $s \in [0, T]$, $x \in \mathbf{R}^m$, значениями которой служат линейные симметрические неотрицательно определенные операторы, действующие из $\mathbf{R}^m$ в $\mathbf{R}^m$.

Обозначим через Ω пространство всех непрерывных функций, заданных на $[0, T]$ со значениями в $\mathbf{R}^m$, и пусть $\mathfrak{F}_t^s$ — минимальная σ-алгебра подмножеств Ω, содержащая все множества вида $\{x(\tau) \in \Gamma\}$, $\tau \in [s, t]$, где $0 \leqslant s \leqslant t \leqslant T$, Γ — борелевское подмножество пространства $\mathbf{R}^m$. Здесь $0 \leqslant s \leqslant t \leqslant T$.

Для заданных $s \in [0, T]$ и $x \in \mathbf{R}^m$ требуется построить вероятностную меру P_{sx} на измеримом пространстве $(\Omega, \mathfrak{F}_T^s)$ так, чтобы:

1) $\mathsf{P}_{sx}\{x(s) = x\} = 1$;

2) процесс

$$x(t)-x(s)-\int_s^t a(\tau, x(\tau))\,d\tau, \quad t\in[s, T],$$

был мартингалом относительно $\left(\mathfrak{F}_t^s, \mathsf{P}_{sx}\right)$ с интегрируемым квадратом, характеристика которого определялась бы формулой

$$\int_s^t b(\tau, x(\tau))\,d\tau.$$

Такую задачу будем называть *проблемой мартингалов* на отрезке $[s, T]$ с начальным условием x для коэффициентов $a(t, y)$, $b(t, y)$, а всякую меру P_{sx}, удовлетворяющую условиям 1), 2), будем называть *мерой, решающей проблему мартингалов.*

Весьма широкие условия существования и единственности решения проблемы мартингалов приведены в следующей теореме.

Теорема 5. *Если в сформулированной выше задаче функции $a(t, x)$ и $b(t, x)$, $t\in[0, T]$, $x\in\mathbf{R}^m$, удовлетворяют условиям:*

а) *$b(t, x)$ непрерывна по совокупности переменных и для некоторой постоянной $C>0$ выполнено неравенство*

$$(b(t, x)\theta, \theta)\leqslant C|\theta|^2$$

при всех $t\in[0, T]$, $x, \theta\in\mathbf{R}^m$;

б) *при любых $t\in[0, T]$ и $x\in\mathbf{R}^m$ матрица $b(t, x)$ положительно определена, т. е. $(b(t, x)\theta, \theta)>0$ при всех $\theta\in\mathbf{R}^m$ и $\theta\neq 0$;*

в) *$a(t, x)$ измерима и ограничена,*

то, каковы бы ни были $s\in[0, T]$ и $x\in\mathbf{R}^m$, существует единственная вероятностная мера $\mathsf{P}_{sx}^{a, b}$ на пространстве $\left(\Omega, \mathfrak{F}_T^s\right)$, удовлетворяющая условиям 1) и 2). При этом процесс $\left(x(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}^{a, b}\right)$ является марковским.

Согласно этой теореме, процесс

$$\xi(t)=x(t)-x(s)-\int_s^t a(\tau, x(\tau))\,d\tau, \quad t\in[s, T],$$

представляет собой мартингал с интегрируемым квадратом относительно потока σ-алгебр $\mathfrak{F}_t^s$ $(t\in[s, T])$ и меры $\mathsf{P}_{sx}^{a, b}$. Если $\eta(t)$ $(t\in[s, T])$ — некоторый матричнозначный (порядка $l\times m$) процесс, прогрессивно измеримый относительно потока σ-алгебр $\mathfrak{F}_t^s$ $(t\in [s, T])$ и удовлетворяющий условию

$$\mathsf{P}_{sx}^{a, b}\left\{\int_s^T \operatorname{Sp}\left(\eta(t)\, b(t, x(t))\, \eta^T(t)\right) dt<\infty\right\}=1,$$

то аналогично тому, как были определены стохастические интегралы по винеровскому процессу, можно определить стохастический интеграл

$$\int_s^T \eta(t)\, d\xi(t).$$

Если, кроме того,

$$\mathsf{M}_{sx}^{a,\,b} \int_s^T \operatorname{Sp}\left(\eta(t)\, b(t, x(t))\, \eta^T(t)\right) dt < \infty,$$

то сепарабельный процесс

$$\int_s^t \eta(\tau)\, d\xi(\tau), \quad t \in [s, T],$$

будет непрерывным мартингалом с интегрируемым квадратом относительно $\left(\mathfrak{F}_t^s, \mathsf{P}_{sx}^{a,\,b}\right)$ с характеристикой

$$\int_s^t \eta(\tau)\, b(\tau, x(\tau))\, \eta^T(\tau)\, d\tau.$$

В частности, полагая $\eta(t) = b^{-1/2}(t, x(t))$, где $b^{-1/2}(t, x)$ — симметричный положительный корень из положительного оператора $b^{-1}(t, x)$, получим, что процесс

$$w(t) = \int_s^t b^{-1/2}(\tau, x(\tau))\, d\xi(\tau), \quad t \in [s, T],$$

является винеровским процессом относительно $\left(\mathfrak{F}_t^s, \mathsf{P}_{sx}^{a,\,b}\right)$, причем

$$\xi(t) = \int_s^t b^{1/2}(\tau, x(\tau))\, dw(\tau), \quad t \geqslant s,$$

где $b^{1/2}(t, x)$ — положительный корень из оператора $b(t, x)$.

Таким образом, в условиях теоремы 5 для каждых $s \in [0, T]$ и $x \in \mathsf{R}^m$ существует m-мерный процесс $w(t)$ $(t \geqslant s)$, заданный на Ω, такой, что процесс $\left(w(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}^{a,\,b}\right)$ винеровский и $\mathsf{P}_{sx}^{a,\,b}$-почти наверное при всех $t \in [s, T]$:

$$x(t) = x + \int_s^t a(\tau, x(\tau))\, d\tau + \int_s^t b^{1/2}(\tau, x(\tau))\, dw(\tau).$$

Пусть $D \subset \mathsf{R}^m$ и $p > 1$; обозначим через $W_p^{1,\,2}([0, T] \times D)$ пополнение семейства бесконечно дифференцируемых финитных функций, заданных на $[0, T] \times D$, по норме

$$\|u\|_{W_p^{1,\,2}} = \|u\|_{L_p} + \|u_t'\|_{L_p} + \|u_x'\|_{L_p} + \|u_{xx}''\|_{L_p},$$

где

$$\|u\|_{L_p} = \left(\int_0^T \int_D |u(t, x)|^p \, dt \, dx \right)^{1/p},$$

$u'_x = \left(\frac{\partial u}{\partial x^1}, \ldots, \frac{\partial u}{\partial x^m} \right)$, u''_{xx} — матрица с элементами $\frac{\partial^2 u}{\partial x^i \partial x^j}$ $(i, j = 1, 2, \ldots, m)$, $u'_t = \partial u/\partial t$. Аналогично пространство $W_p^2(D)$ — это пополнение семейства бесконечно дифференцируемых финитных функций, заданных на D, по норме

$$\|v\|_{W_p^2} = \|v\|_{L_p} + \|v'_x\|_{L_p} + \|v''_{xx}\|_{L_p},$$

где

$$\|v\|_{L_p} = \left(\int_D |v(x)|^p \, dx \right)^{1/p}, \quad D \subset \mathbf{R}^m, \quad p > 1.$$

Далее, назовем функцию $f(t, x)$ *медленно растущей на бесконечности*, если при любом $k > 0$

$$\lim_{|x| \to \infty} \sup_{0 \leqslant t \leqslant T} |f(t, x)| e^{-k|x|^2} = 0.$$

Рассмотрим задачу Коши в области $s \in [0, T)$, $x \in \mathbf{R}^m$:

$$\frac{\partial u}{\partial s} + \frac{1}{2} \sum_{j, k=1}^m b^{jk}(s, x) \frac{\partial^2 u}{\partial x^j \partial x^k} + \sum_{j=1}^m a^j(s, x) \frac{\partial u}{\partial x^j} + f(s, x) = 0,$$

$$u(T, x) = \varphi(x),$$

где функции $f(t, x)$, $\varphi(x)$ $(t \in [0, T],\ x \in \mathbf{R}^m)$ заданы, а $u(s, x)$ — неизвестная функция.

Дополнением к теореме 5 служит следующая теорема.

Теорема 5′. *Предположим, что выполнены условия теоремы 5 и матрица $b(t, x)$ равномерно положительно определена (это означает, что существует число $\varkappa > 0$, такое, что $(b(t, x)\theta, \theta) \geqslant \varkappa|\theta|^2$ при всех $t \in [0, T]$, $x \in \mathbf{R}^m$, $\theta \in \mathbf{R}^m$). Пусть заданы медленно растущие функции $f(t, x)$ и $\varphi(x)$, такие, что для некоторого $p > (m+2)/2$ и любой ограниченной области $D \subset \mathbf{R}^m$ справедливы включения $f \in L_p([0, T] \times D)$, $\varphi \in W_p^2(D)$.*

Тогда в классе медленно растущих функций решение $u(s, x)$ вышеприведенной задачи Коши существует и единственно, а $u \in W_p^{1,2}([0, T] \times D)$, какова бы ни была ограниченная область $D \subset \mathbf{R}^m$.

Если при этом $p \geqslant m + 1$, то решение $u(s, x)$ имеет вероятностное представление

$$u(s, x) = \mathsf{M}_{sx}^{a, b} \left\{ \int_s^T f(t, x(t)) \, dt + \varphi(x(T)) \right\},$$

где $\mathsf{M}_{sx}^{a,b}$ — *операция усреднения по мере* $\mathsf{P}_{sx}^{a,b}$, *существование которой утверждается в теореме* 5.

19.3.5. Дифференцируемость мер, соответствующих решениям стохастических дифференциальных уравнений. Следующая теорема показывает, что мера в пространстве непрерывных функций, соответствующая решению стохастического дифференциального уравнения с данной матрицей диффузии и ненулевым вектором переноса, эквивалентна мере, соответствующей решению такого уравнения с той же матрицей и нулевым вектором переноса.

Теорема 6. *В условиях теоремы* 5 *меры* $\mathsf{P}_{sx}^{0,b}$ *и* $\mathsf{P}_{sx}^{a,b}$ *эквивалентны, причем*

$$\frac{d\mathsf{P}_{sx}^{a,b}}{d\mathsf{P}_{sx}^{0,b}}=\exp\left\{\int_s^T\left(b^{-1}(t,x(t))\,a(t,x(t)),\,dx(t)\right)-\right.$$
$$\left.-\frac{1}{2}\int_s^T\left(b^{-1}(t,x(t))\,a(t,x(t)),\,a(t,x(t))\right)dt\right\}.\qquad(3.3)$$

Заметим, что первый интеграл в правой части последней формулы представляет собой стохастический интеграл по мартингалу.

Таким образом, если функции $a(t, y)$ и $b(t, y)$ удовлетворяют условиям теоремы 5′ и мера $\mathsf{P}_{sx}^{0,b}$ решает проблему мартингалов на отрезке $[s, T]$ с начальным условием x для коэффициентов 0 и $b(t, y)$, то меру $\mathsf{P}_{sx}^{a,b}$, решающую проблему мартингалов на том же отрезке с начальным условием x для коэффициентов $a(t, y)$ и $b(t, y)$, можно получить, полагая

$$\mathsf{P}_{sx}^{a,b}(A)=\int_A R_s(T)\,d\mathsf{P}_{sx}^{0,b},\quad A\in\mathfrak{F}_T^s,\qquad(3.4)$$

где $R_s(T)$ — функционал на $\left(\Omega,\,\mathfrak{F}_T^s\right)$, определяемый первой частью формулы (3.3).

Всякий раз, когда есть мера $\mathsf{P}_{sx}^{0,b}$, решающая проблему мартингалов на отрезке $[s, T]$ с начальным условием x для коэффициентов 0 и $b(t, y)$, и задана некоторая измеримая функция $a(t, y)$ со значениями в $\mathbf{R}^m$ ($t\in[s, T]$, $y\in\mathbf{R}^m$), такая, что интегралы в правой части равенства (3.3) существуют, мы можем составить функционал $R_s(T)$ по указанному правилу и попытаться произвести замену меры $\mathsf{P}_{sx}^{0,b}$ с помощью формулы (3.4) в надежде получить меру, решающую проблему мартингалов на том же отрезке с тем же начальным условием, но уже для коэффициентов $a(t, y)$ и $b(t, y)$.

Такая замена действительно приводит к цели, если только окажется, что функция $a(t, y)$, кроме условия существования интегралов правой части равенства (3.3), удовлетворяет условию

$$\int_\Omega R_s(T)\,d\mathsf{P}_{sx}^{0,b}=1.$$

На этом пути может быть доказана следующая теорема.

Теорема 7. *Предположим, что функция $b(t, x)$ такая же, как и в теореме 5, а измеримая функция $a(t,x)$, $t \in [0, T]$, $x \in \mathrm{R}^m$, со значениями в R^m обладает свойством: при некотором $p > m+2$*

$$\int_0^T \int_{\mathrm{R}^m} |a(t, x)|^p \, dt \, dx < \infty.$$

Тогда при любых $s \in [0, T]$, $x \in \mathrm{R}^m$, существует мера $\mathsf{P}_{sx}^{a,b}$ на пространстве $(\Omega, \mathfrak{F}_T^s)$, решающая проблему мартингалов на отрезке $[s, T]$ с начальным условием x для коэффициентов $a(t, y)$, $b(t, y)$ и такая, что процесс $(x(t), \mathfrak{F}_t^s, \mathsf{P}_{sx}^{a,b})$ является марковским.

Эта теорема показывает, что существуют решения стохастических дифференциальных уравнений с локально неограниченным вектором переноса (т. е. функция $a(t, x)$ может не быть локально ограниченной).

Теорема 8. *Пусть заданы функции $a(x)$ и $b(x)$ со значениями в R^m и $L^+(\mathrm{R}^m)$ соответственно, где $L^+(\mathrm{R}^m)$ — совокупность всех линейных симметрических положительных операторов, действующих в R^m. Предположим, что выполнены условия:*

1) *существуют такие положительные постоянные C_1 и C_2, что при всех x, $\theta \in \mathrm{R}^m$*

$$C_1 |\theta|^2 \leqslant (b(x)\theta, \theta) \leqslant C_2 |\theta|^2;$$

2) *при всех x, $y \in \mathrm{R}^m$*

$$\| b(x) - b(y) \| \leqslant K |x - y|^\alpha,$$

где K и α — положительные постоянные, $\alpha \leqslant 1$, $\|\cdot\|$ — норма оператора, $|\cdot|$ — норма вектора;

3) *при некотором $p > m$*

$$\int_{\mathrm{R}^m} |a(x)|^p \, dx < \infty.$$

Тогда при каждом $x \in \mathrm{R}^m$ на пространстве $(\Omega, \mathfrak{F})$ (здесь Ω — совокупность всех непрерывных функций, заданных на $[0, \infty)$, со значениями в R^m, а $\mathfrak{F}$ — минимальная σ-алгебра подмножеств Ω, содержащая все σ-алгебры $\mathfrak{F}_T^0$ при $T < \infty$) существует вероятностная мера $\mathsf{P}_x^{a,b}$ такая, что

а) $\mathsf{P}_x^{a,b}\{x(0) = x\} = 1$;

б) *процесс*

$$x(t) - x(0) - \int_0^t a(x(s)) \, ds, \quad t \geqslant 0,$$

является мартингалом с интегрируемым квадратом относительно $\left(\mathfrak{F}_t^0, \mathsf{P}_x^{a,b}\right)$, *характеристика которого определяется формулой*

$$\int_0^t b(x(s))\,ds.$$

Процесс $\left(x(t), \mathfrak{F}_t^0, \mathsf{P}_x^{a,b}\right)$ представляет собой однородный марковский процесс. В случае, когда $m \geqslant 2$ и $p > m$, а также $m = 1$ и $p \geqslant 2$, сужения мер $\mathsf{P}_x^{a,b}$ и $\mathsf{P}_x^{0,b}$ на σ-алгебры $\mathfrak{F}_T^0$ эквивалентны при любом $T > 0$. Если же $m = 1$ и $1 < p < 2$, то, вообще говоря, сужения мер $\mathsf{P}_x^{a,b}$ и $\mathsf{P}_x^{0,b}$ на σ-алгебры $\mathfrak{F}_T^0$ не эквивалентны ни при одном $T > 0$.

Заметим, что в случае эквивалентности сужений мер $\mathsf{P}_x^{a,b}$ и $\mathsf{P}_x^{0,b}$ на σ-алгебру $\mathfrak{F}_T^0$ плотность $d\mathsf{P}_x^{a,b}/d\mathsf{P}_x^{0,b}$ определяется формулой

$$\frac{d\mathsf{P}_x^{a,b}}{d\mathsf{P}_x^{0,b}} = \exp\left\{\int_0^T \left(b^{-1}(x(s))\,a(x(s)),\ dx(s)\right) - \frac{1}{2}\int_0^T \left(b^{-1}(x(t))\,a(x(t)),\ a(x(t))\right)dt\right\}.$$

19.3.6. Сильные и слабые решения. Приведенное в п. 19.3.1 понятие «решение стохастического дифференциального уравнения» подразумевало такой случайный процесс $\xi(t)$, который определен на заданном вероятностном пространстве и связан с заданным винеровским процессом $w(t)$ соотношением (3.1). В условиях теоремы 1 оказывалось, что именно такое решение $\xi(t)$ уравнения (3.1) существует, единственно и обладает тем свойством, что $\xi(t)$ измеримо относительно наименьшей σ-алгебры событий, порожденной начальным условием ξ_0 и значениями винеровского процесса $w(s)$ при $s \leqslant t$ (см. замечание 2 к теореме 1). Другими словами, решение в момент времени t полностью определяется значениями входящего в уравнение винеровского процесса во все предшествующие моменты времени, а также начальным условием.

В пп. 19.3.4, 19.3.5, говоря о решениях уравнения (3.1), мы считали заданными лишь его коэффициенты. В задачу входило построение вероятностного пространства и определенной на нем пары непрерывных с вероятностью 1 процессов $(\xi(t), w(t))$, таких, что $w(t)$ — винеровский процесс, связанный с процессом $\xi(t)$ соотношением (3.1) (для того чтобы это уравнение имело смысл, необходимо, чтобы σ-алгебра $\mathfrak{F}_t^\xi$, порожденная значениями процесса $\xi(s)$ при $s \leqslant t$, и приращения $w(t+\tau) - w(t)$ при $\tau \geqslant 0$ были независимы). При этом в теоремах 5, 7, 8 начальное условие считалось неслучайным, хотя можно было бы считать его и случайным с наперед заданным законом распределения $\mu(\Gamma)$, $\Gamma \in \mathfrak{B}$ (напоминаем, что $\mathfrak{B}$ —

это σ-алгебра борелевских подмножеств $\mathbf{R}^m$); в таком случае, например, в теореме 5 вместо меры $\mathsf{P}_{sx}^{a,b}$ решением служила бы мера

$$\mathsf{P}_s^{a,b}(A)=\int\limits_{\mathbf{R}^m}\mathsf{P}_{sx}^{a,b}(A)\,\mu(dx),\quad A\in\mathfrak{F}_T^s.$$

Важно подчеркнуть, что решение $x(t)$, о существовании которого идет речь в теоремах 5, 7, 8, вовсе не обязано полностью определяться начальным условием и тем винеровским процессом, с которым оно связано уравнением (3.1). В самом деле, построение этого винеровского процесса $w(t)$ (см. рассуждение после формулировки теоремы 5) показывает, что $w(t)$ измерим относительно σ-алгебры, порожденной значениями решения $x(\tau)$ при $\tau\leqslant t$, а не наоборот, как это было в условиях теоремы 1.

Условимся о некоторых обозначениях. Если $\xi(t)$ — некий случайный процесс на $[0, T]$, то $\mathfrak{F}_t^\xi$ обозначает наименьшую σ-алгебру событий, содержащую все события вида $\{\xi(s)\in\Gamma\}$, $s\in[0, t]$, $\Gamma\in\mathfrak{B}$. Если $\{\mathfrak{F}_t,\ t\in[0, T]\}$ — некоторый поток σ-алгебр, то говорят, что $\{w(t), \mathfrak{F}_t\}$ является винеровским процессом, если $w(t)$ — винеровский процесс, согласованный с потоком σ-алгебр $\mathfrak{F}_t$, такой, что приращения $w(t+h)-w(t)$ при $h\geqslant 0$ не зависят от σ-алгебры $\mathfrak{F}_t$, каково бы ни было $t\in[0, T]$. Начальное распределение будем считать неслучайным, так что мы рассматриваем уравнение

$$x(t)=x_0+\int\limits_0^t a(s, x(s))\,ds+\int\limits_0^t \sigma(s, x(s))\,dw(s), \tag{3.3}$$

где $x_0\in\mathbf{R}^m$, $t\in[0, T]$; $a(s, x)$, $\sigma(s, x)$ — заданные измеримые функции на $[0, T]\times\mathbf{R}^m$ (первая — со значениями в $\mathbf{R}^m$, а вторая принимает значения во множестве всех квадратных матриц порядка $m\times m$); $w(t)$ — m-мерный винеровский процесс.

Пусть задано вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ и на нем m-мерный винеровский процесс $\left\{w(t), \mathfrak{F}_t^w\right\}$, $t\in[0, T]$. Предположим, кроме того, что на $[0, T]\times\mathbf{R}^m$ заданы две измеримые функции $a(t, x)$, $\sigma(t, x)$ — векторнозначная (размерности m) и матричнозначная (порядка $m\times m$) соответственно.

Пару процессов $\left((x(t), w(t)), \mathfrak{F}_t^w\right)$ назовем *сильным решением* уравнения (3.5), если процесс $x(t)$ при каждом t $\mathfrak{F}_t^w$-измерим и с вероятностью 1 соотношение (3.3) выполняется для всех $t\in[0, T]$ одновременно.

Пусть заданы лишь функции $a(t, x)$, $\sigma(t, x)$, такие же, как и в предыдущем определении. Если найдется вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ с потоком σ-алгебр $\{\mathfrak{F}_t,\ t\in[0, T]\}$ и такая пара процессов $(x(t), w(t))$, согласованных с этим потоком, что процесс $\{w(t), \mathfrak{F}_t\}$ является винеровским, а процессы $x(t)$ и $w(t)$ с вероятностью 1 связаны соотношением (3.3) при всех $t\in[0, T]$ одновременно, то такую пару процессов $\{(x(t), w(t)), \mathfrak{F}_t\}$ назовем *слабым решением* уравнения (3.3).

Таким образом, всякое сильное решение уравнения (3.3) является в то же время и слабым, т. е. понятие слабого решения включает в себя понятие сильного решения. Если $(x(t), w(t))$ — слабое решение уравнения (3.3), то процесс $x(t)$ не обязан быть $\mathfrak{F}_t^w$-измеримым. Допуская некоторую вольность речи, иногда называют сильным или слабым решением уравнения (3.3) сам процесс $x(t)$, а не соответствующую пару.

Если для любых двух решений $(x'(t), w'(t))$ и $(x''(t), w''(t))$ уравнения (3.3), заданных на одном и том же вероятностном пространстве, из равенств $w'(t) \equiv w''(t)$, $x'(0) = x''(0)$ следует, что

$$\mathsf{P}\{\sup_{0 \leqslant t \leqslant T} |x'(t) - x''(t)| > 0\} = 0,$$

то говорят, что решение уравнения (3.3) *единственно в сильном смысле* (*сильно единственно, единственно по траекториям*).

Если любые два решения $(x'(t), w'(t))$ и $(x''(t), w''(t))$ уравнения (3.3) имеют одинаковые конечномерные распределения, то говорят, что решение уравнения (3.3) *единственно в слабом смысле* (*слабо единственно, единственно по мере*).

Существование сильного решения уравнения (3.3) и сильная единственность решения гарантируются, если выполнены условия теоремы 1. Сформулированная выше теорема 5 содержит весьма общие условия существования слабого решения уравнения (3.3) и единственности решения в слабом смысле. В теоремах 7, 8 даны условия существования слабого решения того же уравнения.

Следующий пример хорошо иллюстрирует различие между введенными понятиями.

Пример 1. Пусть $m = 1$; Ω — пространство непрерывных функций, заданных на $[0, T]$ и принимающих значения в $\mathbf{R}$; $\mathfrak{F}_t$ — минимальная σ-алгебра подмножеств Ω, содержащая все множества вида $\{x(\cdot): x(s) \in \Gamma\}$, $s \leqslant t$, $\Gamma \in \mathfrak{B}$; P — винеровская мера на пространстве $(\Omega, \mathfrak{F}_T)$.

Положим

$$\sigma(x) = \begin{cases} 1, & \text{если } x \geqslant 0; \\ -1, & \text{если } x < 0. \end{cases}$$

Определим процесс $w(t)$, $t \in [0, T]$, положив

$$w(t) = \int_0^t \sigma(x(s))\, dx(s).$$

Легко видеть, что процесс $(w(t), \mathfrak{F}_t, \mathsf{P})$ тоже винеровский, причем P-почти наверное

$$x(t) = \int_0^t \sigma(x(s))\, dw(s)$$

при всех $t \in [0, T]$ одновременно. Значит, процесс является решением уравнения

$$dx(t) = \sigma(x(t))\, dw(t)$$

с начальным условием $x(0) = 0$.

Далее можно доказать, что

$$w(t) = |x(t)| - \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t \chi_{[0,\varepsilon]}(|x(s)|)\, ds,$$

где $\chi_{[0,\varepsilon]}(x)$ — индикатор отрезка $[0, \varepsilon]$. Из этой формулы следует, что функция $w(t)$ измерима относительно минимальной σ-алгебры подмножеств Ω, порожденной значениями процесса $|x(s)|$ при $s \leqslant t$. Таким образом, σ-алгебра $\mathfrak{F}_t$ существенно богаче σ-алгебры $\mathfrak{F}_t^w$. Стало быть, решение рассмотренного уравнения нельзя сконструировать, исходя лишь из винеровского процесса $w(t)$. Значит, это решение слабое и не является сильным. Заметим еще, что в этом примере уравнение имеет неединственное решение: вместе с $(x(t), w(t))$ решением является также $(-x(t), w(t))$. С другой стороны, решение этого уравнения единственное в слабом смысле, поскольку конечномерные распределения каждого решения $(x(t), w(t))$ совпадают с конечномерными распределениями пары процессов $(\xi(t), \int_0^t \sigma(\xi(s))\, d\xi(s))$, где $\xi(t)$ — винеровский процесс.

Сформулируем ряд общих утверждений о сильных и слабых решениях.

1. *Если уравнение* (3.3) *обладает сильной единственностью, то всякое его решение является сильным.*

2. *Если уравнение* (3.3) *обладает слабой единственностью, то любые два его сильных решения* $(x'(t), w'(t))$ *и* $(x''(t), w''(t))$ *совпадают с вероятностью* 1 *при всех* t *одновременно.*

3. *Предположим, что функции* $a(t, x)$ *и* $b(t, x) = \sigma(t, x)\sigma^T(t, x)$ *удовлетворяют условиям теоремы* 5. *Тогда, если на некотором вероятностном пространстве уравнение* (3.3) *имеет сильное решение, то любое слабое решение является сильным.*

Таким образом, при указанных условиях имеет место альтернатива.

4. *Либо любое решение уравнения* (3.3), *заданное на любом вероятностном пространстве, является сильным, либо ни одно решение ни на одном вероятностном пространстве не является сильным, и эти решения обязательно неединственны по траекториям.*

5. Пусть $m = 1$. В пространстве $L_2[0, T]$ (функций $\varphi: [0, T] \to \mathbf{R}$ с интегрируемым квадратом), выберем ортонормированный базис $\{\varphi_k(t),\ k = 1, 2, \ldots\}$, состоящий из равномерно ограниченных функций. Для всякой числовой последовательности $z = \{z_1, z_2, \ldots\}$, для которой $\sum_{k=1}^{\infty} z_k^2 \leqslant C$, $\sum_{k=1}^{\infty} |z_k| \leqslant C$ (здесь $C > 0$ — фиксированная постоянная), положим

$$\varphi(t, z) = \sum_{k=1}^{\infty} z_k \varphi_k(t)$$

и обозначим через $u(s, x, z)$ решение задачи Коши (в области

$s \in [0, t)$, $x \in \mathbf{R}$):

$$\frac{\partial u}{\partial s} + \frac{1}{2} b(s, x) \frac{\partial^2 u}{\partial x^2} + (a(s, x) + \sigma(s, x) \varphi(s, z)) \frac{\partial u}{\partial x} = 0,$$
$$u(t, x, z) = x.$$

Через $v(s, x)$ обозначим решение задачи Коши (в той же области):

$$\frac{\partial v}{\partial s} + \frac{1}{2} b(s, x) \frac{\partial^2 v}{\partial x^2} + a(s, x) \frac{\partial v}{\partial x} = 0,$$
$$v(t, x) = x^2.$$

Если коэффициенты $b(t, x) = \sigma^2(t, x)$ и $a(t, x)$ удовлетворяют условиям теоремы 5 (в рассматриваемом случае это означает, что функция $b(t, x)$ непрерывна, $\inf\limits_{\substack{t \in [0, T] \\ x \in \mathbf{R}}} b(t, x) > 0$, функции $\sigma(t, x)$ и $a(t, x)$ ограничены и измеримы), то, согласно теореме 5, решения этих задач в классе медленно растущих функций, локально принадлежащих пространству $W_p^{1, 2}$ с любым $p \geqslant 2$, существуют и единственны.

Далее, если $I = (i_1, i_2, \ldots, i_n)$, где i_k — целые неотрицательные числа, то обозначим $|I| = i_1 + i_2 + \ldots + i_n$, $I! = i_1!\, i_2! \ldots i_n!$ и

$$D_z^I = \frac{\partial^{|I|}}{\partial z_1^{i_1} \ldots \partial z_n^{i_n}}$$

(если $I = (0, 0, \ldots, 0)$, то D_z^I — тождественный оператор).

Теперь мы можем сформулировать необходимое и достаточное условие существования сильного решения уравнения (3.3) (в одномерном случае) в терминах решений приведенных выше задач Коши.

В предположении, что коэффициенты $a(t, x)$ и $\sigma(t, x)$ удовлетворяют условиям теоремы 5′, уравнение (3.3) имеет сильное решение при $t \leqslant T$ тогда и только тогда, когда выполняется тождество

$$v(0, x) = \sum_I \frac{1}{I!} \left[D_z^I u(0, x, z) \right]^2 \Big|_{z=0}.$$

Хотя этот критерий и трудно проверяем, тем не менее он показывает, что наличие или отсутствие сильных решений уравнения (3.3) связано с устройством его коэффициентов, а не с удачным или неудачным выбором вероятностного пространства, винеровского процесса и т. п.

19.3.7. Теоремы существования сильных решений. Сильная единственность. В теореме 1 сформулированы условия, обеспечивающие существование сильного решения и его единственность по траекториям. Оказывается, что дополнительные предположения об одной группе коэффициентов позволяют ослабить условия на другие. Ниже сформулированы некоторые теоремы существования сильных решений, сильной единственности, а также результат, известный под названием «теоремы сравнения».

Теорема 9. *Предположим, что входящие в уравнение* (3.3) *функции* $a(t, x)$, $\sigma(t, x)$ *ограничены и измеримы; матрица* $b(t, x) = \sigma(t, x)\sigma^T(t, x)$ *непрерывна по совокупности переменных и равномерно положительно определена* (*см. теорему* 5′); *функция* $\sigma(t, x)$ *удовлетворяет условию Липшица по* x:

$$|\sigma(t, x) - \sigma(t, y)| \leqslant K|x - y|, \quad t \in [0, T], \quad x, y \in \mathbf{R}^m$$

(*здесь* K — *постоянная*); *функции* $a(t, x)$ *и* $\sigma(t, x)$ *удовлетворяют условию Дини: при некотором* $\varepsilon > 0$ $\int\limits_0^\varepsilon \rho(r) r^{-1} dr < \infty$, *где* $\rho(r)$ — *модуль непрерывности по переменным* t, x *любой из функций* $a(t, x)$, $\sigma(t, x)$.

Тогда уравнение (3.3) *имеет сильное решение, и оно единственное по траекториям.*

Замечание. Если функция $a(t, x)$ не удовлетворяет условию Липшица по x, то решение обыкновенного дифференциального уравнения $dx(t) = a(t, x(t))dt$ может оказаться неединственным (например, уравнение $dx(t) = (1-\alpha)^{-1}|x(t)|^\alpha dt$ с начальным условием $x(0) = 0$ при $0 < \alpha < 1$ имеет решения $x(t) \equiv 0$, $x(t) = t^{1/(1-\alpha)}$). В то же время прибавление к правой части этого уравнения сколь угодно малого стохастического возмущения превращает его в уравнение $dx(t) = a(t, x(t))dt + \varepsilon dw(t)$, для которого решение может оказаться единственным при почти всех траекториях винеровского процесса $w(t)$. Согласно теореме 9, это будет так, если функция $a(t, x)$ удовлетворяет условию Дини по переменным t, x.

В предыдущей теореме размерность пространства произвольна. В одномерном случае можно получить более сильные результаты.

Теорема 10. *Предположим, что функции* $a(t, x)$ *и* $\sigma(t, x)$, *заданные при* $t \in [0, T]$, $x \in \mathbf{R}$, *ограничены и измеримы*, $\inf\limits_{t, x} \sigma(t, x) > 0$ *и функция* $\sigma(t, x)$ *при всех* $t \in [0, T]$, $x \in \mathbf{R}$ *представима в виде* $\sigma(t, x) = \sigma_1(t, x)\sigma_2(x)\sigma_3(t, x)$, *где* σ_i — *ограниченные борелевские функции, причем:*

1) *при всех* $x, y \in \mathbf{R}$ *и почти всех* $t \in [0, T]$

$$|\sigma_1(t, x) - \sigma_1(t, y)| \leqslant c(t)\rho(|x - y|),$$

где $\rho(u)$, $u \geqslant 0$, — *возрастающая выпуклая вверх функция, такая, что при любом* $\varepsilon > 0$

$$\int\limits_0^\varepsilon \rho^{-2}(u)\, du = +\infty,$$

а $c(t)$ $(t \geqslant 0)$ — *измеримая по Лебегу функция, для которой*

$$\int\limits_0^T c^2(t)\, dt < \infty;$$

2) *функция* $\sigma_2(x)$ *имеет ограниченную вариацию на любом конечном промежутке изменения* x;

3) *функция* $\sigma_3(t, x)$ *абсолютно непрерывна по* x *при почти всех* $t \geqslant 0$, *и ее производная* $\frac{\partial \sigma_3(t, x)}{\partial x}$ *локально интегрируема с квадратом, т. е. при всех* $N > 0$

$$\int_0^T \int_{-N}^N \left| \frac{\partial \sigma_3(t, x)}{\partial x} \right|^2 dt\, dx < \infty.$$

Тогда уравнение (3.3) *имеет сильное решение, и оно единственно по траекториям.*

В частности, из этой теоремы следует результат: в одномерном случае, если $a(t, x)$ и $\sigma(t, x)$ — ограниченные измеримые функции, $\sigma(t, x)$ равномерно отделена от нуля и удовлетворяет условию Гёльдера по x:

$$|\sigma(t, x) - \sigma(t, y)| \leqslant K |x - y|^{\alpha}, \quad t \in [0, T], \quad x, y \in \mathbf{R},$$

при некотором $\alpha \geqslant 1/2$ (здесь K — постоянная), то уравнение (3.3) имеет единственное сильное решение.

В теоремах 9, 10 матрица диффузии $b(t, x) = \sigma(t, x)\sigma^T(t, x)$ предполагалась невырожденной. Как уже отмечалось, общие условия существования и единственности сильного решения уравнения (3.3), допускающие вырождение матрицы $\sigma(t, x)$, даются теоремой 1.

В одномерном случае следующая, более общая теорема содержит условия сильной единственности решения уравнения (3.3) без предположения о невырожденности коэффициента диффузии.

Теорема 11. *Пусть для* $t \in [0, T]$, $x \in \mathbf{R}$, *заданы измеримые функции* $a(t, x)$ *и* $\sigma(t, x)$, *такие, что*:

1) *при всех* $t \in [0, T]$, $x, y \in \mathbf{R}$, *выполнено неравенство*

$$|\sigma(t, x) - \sigma(t, y)| \leqslant \rho(|x - y|),$$

где положительная возрастающая функция $\rho(u)$ $(u > 0)$ *удовлетворяет условию*

$$\int_0^{\varepsilon} \rho^{-2}(u)\, du = +\infty$$

при любом $\varepsilon > 0$;

2) *при всех* $t \in [0, T]$, $x, y \in \mathbf{R}$, *справедливо неравенство*

$$|a(t, x) - a(t, y)| \leqslant k(|x - y|),$$

где положительная возрастающая выпуклая вверх функция $k(u)$ $(u > 0)$ *удовлетворяет условию*

$$\int_0^{\varepsilon} k^{-1}(u)\, du = +\infty$$

при любом $\varepsilon > 0$.

Тогда уравнение (3.3) *обладает потраекторной единственностью.*

Таким образом, если для некоторой пары функций $a(t, x)$, $\sigma(t, x)$, удовлетворяющих условиям теоремы 11, удается доказать существование слабого решения уравнения (3.3), то это решение автоматически оказывается сильным (см. утверждение 1 п. 19.3.6).

В частности, утверждение теоремы 11 справедливо, если коэффициент переноса $a(t, x)$ удовлетворяет условию Липшица по x, а коэффициент диффузии $\sigma(t, x)$ (вообще говоря, вырождающийся) удовлетворяет условию Гёльдера по x с показателем $\alpha \geqslant 1/2$ (оба коэффициента измеримы).

Следующий пример показывает, что при $a(t, x) \equiv 0$ условие 1) теоремы 11 нельзя ослабить.

Пример 2. Пусть $\rho(u)$ $(u \geqslant 0)$ — непрерывная возрастающая функция, для которой $\rho(0) = 0$ и

$$\int_0^1 \rho^{-2}(u)\, du < \infty.$$

Тогда с вероятностью 1 для всех t одновременно

$$\int_0^t \rho^{-2}(|w(s)|)\, ds < \infty,$$

где $w(s)$ — одномерный винеровский процесс. Положим $\xi(t) = w(\tau_t)$, где τ_t определяется из соотношения

$$\int_0^{\tau_t} \rho^{-2}(|w(s)|)\, ds = t, \quad t \geqslant 0.$$

Тогда нетрудно показать, что процесс $\xi(t)$ удовлетворяет уравнению

$$d\xi(t) = \rho(|\xi(t)|)\, d\tilde{w}(t),$$

где $\tilde{w}(t)$ — некоторый новый винеровский процесс. С начальным условием $\xi(0) = 0$ этому уравнению удовлетворяет, кроме указанного процесса, также процесс $\bar{\xi}(t) \equiv 0$, так что решение рассмотренного уравнения неединственно по траекториям.

Приведем еще один пример, показывающий, что при одних начальных условиях стохастическое дифференциальное уравнение обладает потраекторной единственностью, а при других такая единственность отсутствует.

Пример 3. Пусть $w(t)$ — двумерный винеровский процесс. Тогда, как легко подсчитать, при всех $0 < \alpha < 1$ и $t \geqslant 0$

$$\mathrm{M} \int_0^t |w(s)|^{-2\alpha}\, ds < \infty$$

и, значит, с вероятностью 1 при всех $t \geqslant 0$ конечен интеграл $\int_0^t |w(s)|^{-2\alpha}\, ds$. Определим τ_t из соотношения

$$\tau_t = \int_0^t |w(s)|^{-2\alpha}\, ds, \quad t \geqslant 0,$$

и положим $\xi(t) = w(\tau_t)$. Тогда нетрудно убедиться, что $\xi(t)$ является решением уравнения

$$d\xi(t) = |\xi(t)|^\alpha d\tilde{w}(t),$$

где $\tilde{w}(t)$ — некоторый новый (двумерный) винеровский процесс. Это уравнение, кроме указанного решения, выходящего из начала координат на плоскости, имеет еще решение $\bar{\xi}(t)$, не покидающее начала координат ни в один из моментов времени. Если же начальным условием служит точка, отличная от начала координат, то решение рассмотренного уравнения никогда не попадет в начало координат, и, стало быть, оно (решение) единственно в сильном смысле.

Следующий результат показывает, что существует достаточно широкий класс одномерных уравнений с невырожденным гёльдеровым коэффициентом диффузии, не обладающих потраекторной единственностью и, значит, не имеющих сильных решений.

Обозначим через $H^+_{\alpha,\beta}(I)$ множество всех функций $f(x)$, определенных на замкнутом ограниченном отрезке $I \subseteq \mathbf{R}$ и удовлетворяющих условиям:

а) $\inf\limits_x f(x) > 0$;

б) $\sup\limits_{x, y} |f(x) - f(y)|\,|x - y|^{-\alpha} < \infty$;

в) для всех $x < y$ $(x, y \in I)$ найдутся такие x', y', что $x < x' < y' < y$ и

$$|f(x') - f(y')| \geqslant C|x - y|^\beta;$$

здесь α, β — положительные числа, а C — положительная постоянная, зависящая от функции f.

При $\alpha \leqslant \beta \leqslant 1$ классы $H^+_{\alpha,\beta}(I)$ непусты. Примером функции f из класса $H^+_{\alpha,\alpha}([0, 1])$ может служить функция

$$f(x) = a + \sum_{k=0}^{\infty} b^{-k\alpha}\varphi(b^k x), \quad x \in [0, 1],$$

где $\varphi(x) = x\lambda(1 - x)$, $a > 0$, b — целое число, удовлетворяющее неравенству $b^{1-\alpha} > 2$.

Теорема 12. *Пусть $\sigma(x)$ — всюду положительная непрерывная ограниченная функция, заданная на $\mathbf{R}$. Предположим, что сужение этой функции на некоторый ограниченный замкнутый отрезок $I \subseteq \mathbf{R}$ принадлежит классу $H^+_{\alpha,\beta}(I)$, где $\alpha \leqslant \beta < \dfrac{2\alpha}{1 + 2\alpha} \vee \dfrac{2\alpha + 2\alpha^2}{2 - \alpha}$.*

Если x_0 — внутренняя точка отрезка I, то решение уравнения

$$d\xi(t) = \sigma(\xi(t))\,dw(t), \quad \xi(0) = x_0$$

неединственно.

Следующая теорема называется *теоремой сравнения.*

Теорема 13. *Рассмотрим два одномерных стохастических дифференциальных уравнения:*

$$dx_i(t) = a_i(t, x_i(t))\,dt + \sigma(t, x_i(t))\,dw(t), \quad i = 1, 2, \qquad (3.4)$$

где функции $a_i(t, x)$ $(i = 1, 2)$, $\sigma(t, x)$ *определены на* $[0, \infty) \times \mathbf{R}$, *непрерывны, ограничены и таковы, что функция* $\sigma(t, x)$ *удовлетворяет условию* 1) *теоремы* 11, *а функции* $a_1(t, x)$, $a_2(t, x)$ *связаны неравенством*

$$a_1(t, x) < a_2(t, x)$$

при всех $(t, x) \in [0, \infty) \times \mathbf{R}$.

Тогда, если $(x_1(t), w(t))$ *и* $(x_2(t), w(t))$ — *решения указанных уравнений при* $i = 1, 2$ *соответственно (эти решения заданы на одном вероятностном пространстве и с одним и тем же винеровским процессом* $w(t)$) *и* $x_1(0) = x_2(0)$, *то* $x_1(t) \leqslant x_2(t)$ *с вероятностью* 1 *для всех* t *одновременно.*

Следствие 1. *Пусть* $\sigma(t, x)$ *удовлетворяет условиям теоремы* 13, *а функция* $a(t, x)$ *непрерывна и ограничена. Тогда стохастическое дифференциальное уравнение*

$$dx(t) = a(t, x(t))\, dt + \sigma(t, x(t))\, dw(t) \tag{3.5}$$

с начальным условием $x(0) = x_0$ *имеет минимальное и максимальное решения:* $(\underline{x(t)}, w(t))$ *и* $(\overline{x(t)}, w(t))$ *соответственно.*

Это означает, что если $(x(t), w(t))$ — произвольное решение указанного уравнения с тем же начальным условием $x(0) = x_0$, то с вероятностью 1 $\underline{x(t)} \leqslant x(t) \leqslant \overline{x(t)}$ при всех t одновременно. При этом оба процесса $\underline{x(t)}$ и $\overline{x(t)}$ обладают марковским свойством и являются диффузионными в смысле п. 19.1 1. Если к тому же рассматриваемое уравнение обладает единственностью в слабом смысле, то с вероятностью 1 $\underline{x(t)} = \overline{x(t)}$ при всех t одновременно, и, стало быть, решение этого уравнения единственно по траекториям.

Следствие 2. *Пусть выполнены все условия теоремы* 13, *за исключением строгого неравенства:* $a_1(t, x) < a_2(t, x)$, *вместо которого теперь будем предполагать выполненным нестрогое неравенство:* $a_1(t, x) \leqslant a_2(t, x)$ *при всех* $(t, x) \in [0, \infty) \times \mathbf{R}$. *Предположим, кроме того, что оба уравнения* (3.6) *обладают потраекторной единственностью (в частности, это будет так, если* $a_i(t, x)$ $(i \in 1, 2)$ *удовлетворяет условию* 2) *теоремы* 11). *Тогда справедливо утверждение теоремы* 13.

Приведем два примера стохастических дифференциальных уравнений, обладающих потраекторной единственностью, что легко может быть выведено из теоремы сравнения и ее следствий.

Пример 4. Предположим, что в уравнении (3.5) коэффициент $\sigma(t, x)$ удовлетворяет условиям теоремы 13, а коэффициент $a(t, x)$ непрерывен, ограничен и при каждом $t \in [0, \infty)$ функция $a(t, x)$ не возрастает как функция x. Тогда решение уравнения (3.7) потраекторно единственно.

Пример 5. Рассмотрим уравнение

$$dx(t) = \sigma(x(t))\, dw(t) + a(x(t))\, dt,$$

где

$$\sigma(x) = \begin{cases} 0 & \text{при} \quad x < 0; \\ x & \text{при} \quad x \geqslant 0, \end{cases}$$

а непрерывная ограниченная функция $a(x)$ такова, что $a(x) \geqslant C > 0$ (здесь C — постоянная). Утверждается, что решение этого уравнения единственно по траекториям.

В заключение этого пункта приведем теорему существования и единственности решения уравнения (3.3), обобщающую теорему 1. Условия на коэффициенты $a(t, x)$ и $\sigma(t, x)$ в этой теореме носят название *условия монотонности.*

Теорема 14. *Пусть на* $[0, T]\times \mathbf{R}^m$ *заданы измеримые векторнозначная (размерности* m) *функция* $a(t, x)$ *и матричнозначная (порядка* $m\times m$) *функция* $\sigma(t, x)$, *и пусть для всякого* $x\in \mathbf{R}^m$

$$\int_0^T |a(t, x)|\,dt < \infty.$$

Предположим далее, что для всякого $N>0$ *существует неотрицательная измеримая функция* $K_t(N)$ *такая, что*

$$\int_0^T K_t(N)\,dt < \infty$$

и при почти всех $t\in[0, T]$ *и всех* $x, y, z\in \mathbf{R}^m$ *справедливы неравенство*

$$2(x-y,\ a(t, x)-a(t, y))+ \\ + \operatorname{Sp}[\sigma(t, x)-\sigma(t, y)][\sigma(t, x)-\sigma(t, y)]^T \leqslant K_t(N), \quad (3.6)$$

если только $|x|\leqslant N$, $|y|\leqslant N$, *и неравенство*

$$2(z,\ a(t, z))+\operatorname{Sp}[\sigma(t, z)\sigma^*(t, z)]\leqslant K_t(1)(1+|z|^2). \quad (3.7)$$

Тогда уравнение (3.3) *имеет сильное решение, и оно единственно по траекториям.*

Пример 6. Пусть заданы число $\gamma\in(1, 2)$ и вещественные измеримые функции $\alpha(t)$, $\beta(t)$, $t\in[0, T]$, связанные неравенством

$$-2(\gamma-1)\,\alpha(t)+\frac{\gamma^2}{4}\beta^2(t)\leqslant 0$$

при почти всех $t\in[0, \mathrm{T}]$, и пусть

$$\int_0^T \alpha(t)\,dt < \infty.$$

Положим для $t\in[0, T]$, $x\in \mathbf{R}$ $a(t, x)=-\alpha(t)|x|^{\gamma-1}\operatorname{sign} x$, $\sigma(t, x)= =\beta(t)|x|^{\gamma/2}$. Нетрудно проверить, что так определенные функции удовлетворяют условиям (3.6), (3.7) с $K_t(N)$, и потому (одномерное) уравнение (3.3) с такими коэффициентами имеет сильное решение и оно единственно по траекториям.

19.3.8. Уравнения с коэффициентами, зависящими от прошлого. В этом пункте рассмотрим более общие стохастические дифференциальные уравнения. Именно, будем предполагать, что коэффициенты уравнения в момент времени t являются функционалами, зависящими от траектории искомого процесса до момента времени t.

Обозначим через $C_{[0, T]}(\mathbf{R}^m)$ пространство всех непрерывных функций $x(t)$, заданных на $[0, T]$ и принимающих значение в $\mathbf{R}^m$.

Через $\mathfrak{N}_t$ обозначим минимальную σ-алгебру подмножеств $C_{[0,T]}(\mathbf{R}^m)$, содержащую все множества вида $\{x(\cdot): x(s) \in \Gamma\}$ при $s \leqslant t$, $\Gamma \in \mathfrak{B}$ ($\mathfrak{B}$ по-прежнему обозначает σ-алгебру борелевских подмножеств $\mathbf{R}^m$). Пусть для $t \in [0, T]$, $x(\cdot) \in C_{[0,T]}(\mathbf{R}^m)$ заданы:

1) функция $a(t, x(\cdot))$ со значениями в $\mathbf{R}^m$, измеримая по совокупности переменных (т. е. измеримо отображение

$$a\colon ([0, T] \times C_{[0,T]}(\mathbf{R}^m), \mathfrak{B}_{[0,T]} \times \mathfrak{N}_T) \to (\mathbf{R}^m, \mathfrak{B}),$$

где $\mathfrak{B}_{[0,T]}$ — σ-алгебра борелевских подмножеств отрезка $[0, T]$), при каждом $t \in [0, T]$ $\mathfrak{N}_t$-измеримая как функция $x(\cdot)$, при каждом $x(\cdot) \in C_{[0,T]}(\mathbf{R}^m)$ ограниченная как функция t;

2) матричнозначная (порядка $m \times m$) функция $\sigma(t, x)$, измеримая по совокупности переменных, при каждом $t \in [0, T]$ $\mathfrak{N}_t$-измеримая как функция $x(\cdot)$, при каждом $x(\cdot) \in C_{[0,T]}(\mathbf{R}^m)$ ограниченная как функция t.

Пусть также заданы вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ с потоком σ-алгебр $\mathfrak{F}_t$ ($t \in [0, T]$), m-мерный винеровский процесс $w(t)$ ($t \in [0, T]$) и случайный вектор $\xi_0 \in \mathbf{R}^m$, удовлетворяющие условиям 1)—3) п. 19.3.1.

Рассмотрим стохастическое дифференциальное уравнение

$$d\xi(t) = a(t, \xi(\cdot))\,dt + \sigma(t, \xi(\cdot))\,dw(t) \tag{3.8}$$

с начальным условием $\xi(0) = \xi_0$, которое можно записать в интегральной форме:

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(\cdot))\,ds + \int_0^t \sigma(s, \xi(\cdot))\,dw(s). \tag{3.9}$$

Такие уравнения представляют собой обобщение тех уравнений, которые рассматривались в предыдущих пунктах. Именно, мы получим рассмотренные ранее уравнения, если окажется, что функционалы $a(t, x(\cdot))$, $\sigma(t, x(\cdot))$ зависят не от всей траектории $x(s)$ при $s \leqslant t$, а лишь от значения функции $x(\cdot)$ в момент времени t.

Понятие решения уравнения (3.8) почти дословно совпадает с определением решения уравнения (3.1). То же относится и к понятию единственности решения. Аналогом теоремы 1 является следующая теорема.

Теорема 15. *Пусть заданы перечисленные объекты (т. е. коэффициенты): $a(t, x(\cdot))$, $b(t, x(\cdot))$, вероятностное пространство с потоком σ-алгебр $\mathfrak{F}_t$ ($t \in [0, T]$), m-мерный винеровский процесс $w(t)$, случайная величина ξ_0, удовлетворяющие указанным выше условиям, и предположим, что выполняются условия:*

1) *существует постоянная K, такая, что при $t \leqslant T$*

$$|a(t, x(\cdot))|^2 + |\sigma(t, x(\cdot))|^2 \leqslant K(1 + \|x(\cdot)\|_t^2),$$

где $\|x(\cdot)\|_t = \sup\limits_{s \leqslant t} |x(s)|$;

2) *для каждого $R > 0$ существует такая постоянная C_R, что при $t \in [0, T]$, $\|x(\cdot)\|_T \leqslant R$, $\|y(\cdot)\|_T \leqslant R$ выполнено неравенство*

$$|a(t, x(\cdot)) - a(t, y(\cdot))|^2 + |\sigma(t, x(\cdot)) - \sigma(t, y(\cdot))|^2 \leqslant$$
$$\leqslant C_R \|x(\cdot) - y(\cdot)\|_t.$$

Тогда существует единственное $\mathfrak{F}_t$-измеримое решение $\xi(t)$ уравнения (3.9), *причем $\xi(t)$ измеримо относительно наименьшей σ-алгебры событий, порожденной значениями винеровского процесса $w(s)$ при $s \leqslant t$ и начальным значением ξ_0.*

На рассматриваемые уравнения переносятся понятия сильного и слабого решений (см. п. 19.3.6). Приведенная выше теорема 15 дает условия существования сильного решения уравнения (3.8) и сильной единственности решения. Следующая теорема содержит условия существования слабого решения.

Теорема 16. *Предположим, что коэффициенты уравнения* (3.8) *удовлетворяют условиям:*

а) *функции $a(t, x(\cdot))$, $\sigma(t, x(\cdot))$ непрерывны по $x(\cdot)$ в метрике пространства $C_{[0,\,T]}(\mathrm{R}^m)$ при каждом $t \leqslant T$;*

б) *существует такая постоянная K, что*

$$|a(t, x(\cdot))|^2 + |\sigma(t, x(\cdot))|^2 \leqslant K\left(1 + \|x(\cdot)\|_t^2\right).$$

Тогда существует слабое решение уравнения (3.8). *Слабые решения уравнения* (3.8) *можно получать и с помощью абсолютно непрерывной замены меры.*

Рассмотрим для примера уравнение

$$dx(t) = a(t, x(\cdot))\,dt + dw(t) \tag{3.10}$$

с неслучайным начальным значением $x(0) = x_0$, где x_0 — фиксированный вектор из R^m.

Теорема 17. *Пусть на $(C_{[0,\,T]}(\mathrm{R}^m), \mathfrak{N}_t)$ задана мера P такая, что $\mathsf{P}\{x(0) = x_0\} = 1$, и процесс $x(t)$, $t \in [0, T]$, является квадратично интегрируемым мартингалом относительно $(\mathfrak{N}_t, \mathsf{P})$ с характеристикой $t \cdot I$ (другими словами, процесс $(x(t), \mathfrak{N}_t, \mathsf{P})$ является винеровским, выходящим из точки x_0 в начальный момент времени).*

Если функция $a(t, x(\cdot))$ удовлетворяет перечисленным в начале пункта условиям измеримости и такова, что

$$\mathsf{P}\left\{\int_0^T |a(t, x(\cdot))|^2\,dt < \infty\right\} = 1, \tag{3.11}$$

$$\mathsf{M}R_T(x(\cdot)) = 1,$$

где M — символ операции усреднения по мере P, а $R_T(x(\cdot))$ — функционал от $x(\cdot)$, определяемый формулой

$$R_T(x(\cdot)) = \exp\left\{\int_0^T (a(t, x(\cdot))\,dx(t)) - \frac{1}{2}\int_0^T |a(t, x(\cdot))|^2\,dt\right\},$$

то, полагая

$$\tilde{\mathsf{P}}(A) = \int_A R_T(x(\cdot))\,d\mathsf{P}, \quad A \in \mathfrak{N}_T,$$

получим меру $\tilde{\mathsf{P}}$ *на* $(C_{[0,T]}(\mathbf{R}^m), \mathfrak{N}_T)$, *обладающую тем свойством, что* $\tilde{\mathsf{P}}\{x(0) = x_0\} = 1$, *и процесс*

$$\tilde{w}(t) = x(t) - x_0 - \int_0^t a(s, x(\cdot))\, ds, \quad t \in [0, T]$$

является квадратично интегрируемым мартингалом относительно $(\mathfrak{N}_t, \tilde{\mathsf{P}})$, *характеристика которого равна* $t \cdot I$ (*здесь и выше* I — *единичная матрица*).

Иначе говоря, пара процессов $(x(t), w(t))$, $t \in [0, T]$, заданных на вероятностном пространстве $(C_{[0,T]}(\mathbf{R}^m), \mathfrak{N}_T, \tilde{\mathsf{P}})$ с потоком σ-алгебр $\{\mathfrak{N}_t, t \in [0, T]\}$, представляет собой слабое решение уравнения (3.10). Более того, можно показать, что в классе мер, удовлетворяющих условию

$$\tilde{\mathsf{P}}\left\{\int_0^T |a(t, x(\cdot))|^2\, dt < \infty\right\} = 1,$$

построенная в теореме 17 мера единственна. Этот результат может быть дополнен следующим образом.

Теорема 17'. *Предположим, что функция* $a(t, x(\cdot))$ *такова, что*

$$\int_0^T |a(t, x(\cdot))|^2\, dt < \infty$$

для всякой функции $x(\cdot) \in C_{[0,T]}(\mathbf{R}^m)$. *Тогда второе из условий* (3.11) *является необходимым и достаточным для существования и слабой единственности слабого решения уравнения* (3.10).

Для того чтобы функция $a(t, x(\cdot))$ удовлетворяла условиям (3.11), достаточно, чтобы

$$\mathsf{M} \exp\left\{\frac{1}{2}\int_0^T |a(t, x(\cdot))|^2\, dt\right\} < \infty,$$

где M имеет то же значение, что и выше. В частности, условия (3.11) будут выполнены, если функция $a(t, x(\cdot))$ ограничена, так что уравнение (3.10) имеет единственное слабое решение, какова бы ни была ограниченная функция $a(t, x(\cdot))$, удовлетворяющая условию 1), приведенному в начале этого пункта.

Приведем в заключение пример уравнения (3.10) с ограниченной функцией $a(t, x(\cdot))$, которое не имеет сильных решений.

Пример 7. Пусть $\{t_n, n = 0, 1, 2, \ldots\}$ — числовая последовательность, такая, что $1 = t_0 > t_1 > t_2 > \ldots$, $\lim_{n\to\infty} t_n = 0$. Положим для $s \in [t_k, t_{k-1})$, $k = 1, 2, \ldots$,

$$a(s, x(\cdot)) = \left\{\frac{x(t_k) - x(t_{k+1})}{t_k - t_{k+1}}\right\},$$

где $\{\gamma\}$ — дробная часть числа γ. Пусть $(x(t),\ w(t))$ — решение уравнения (3.10) с таким коэффициентом $a(t,\ x(\cdot))$ (рассматривается одномерный случай).

Обозначим

$$\eta_k = \left\{ \frac{x(t_{k-1}) - x(t_k)}{t_{k-1} - t_k} \right\}, \quad \xi_k = \frac{w(t_{k-1}) - w(t_k)}{t_{k-1} - t_k}, \qquad k = 1, 2, \ldots$$

Из уравнения (3.10) легко получить соотношение

$$\exp\{2\pi i \eta_k\} = \exp\{2\pi i \eta_{n+1}\} \exp\{2\pi i (\xi_k + \ldots + \xi_n)\},$$

справедливое при всех $n > k$. Отсюда

$$\left| \mathsf{M} e^{2\pi i \eta_k} \right| \leqslant \prod_{j=k}^{n} \left| \mathsf{M} e^{2\pi i \xi_j} \right| = \exp\left\{ -2\pi^2 \sum_{j=k}^{n} (t_{j-1} - t_j)^{-1} \right\} \to 0$$

при $n \to \infty$ и, значит, $\mathsf{M} e^{2\pi i \eta_k} = 0$ при всех k. Обозначим через $\mathfrak{F}_t^s$ σ-алгебру событий, порожденную величинами $w(\tau) - w(s)$ при $\tau \in [s,\ t]$. Тогда при $k < n$

$$\mathsf{M}\left\{ e^{2\pi i \eta_k} \middle| \mathfrak{F}_{t_{k-1}}^{t_n} \right\} = \exp\{2\pi i (\xi_k + \ldots + \xi_n)\} \times$$
$$\times \mathsf{M}\left\{ e^{2\pi i \eta_{n+1}} \middle| \mathfrak{F}_{t_{k-1}}^{t_n} \right\} = 0,$$

поскольку величина η_{n+1} не зависит от σ-алгебры $\mathfrak{F}_{t_{k-1}}^{t_n}$. Переходя к пределу в последнем соотношении при $n \to \infty$, получаем равенство

$$\mathsf{M}\left\{ e^{2\pi i \eta_k} \middle| \mathfrak{F}_{t_{k-1}}^{0} \right\} = 0,$$

которое означает, что η_k не может быть $\mathfrak{F}_{t_{k-1}}^{0}$-измеримой величиной. Значит, никакое решение рассматриваемого уравнения не может быть сильным.

19.4. Стохастические дифференциальные уравнения для процессов с разрывами

19.4.1. Пуассоновские меры. Пусть $(\Theta,\ \mathfrak{M})$ — некоторое измеримое пространство с σ-конечной мерой Π. Через $\mathfrak{M}_0$ будем обозначать подалгебру $\mathfrak{M}$, состоящую из тех множеств $A \in \mathfrak{M}$, для которых $\Pi(A) < \infty$, а через $\mathfrak{B}_+$ — σ-алгебру борелевских подмножеств полуоси $\mathbf{R}_+ = [0,\ \infty)$. Пусть, далее, задано вероятностное пространство $(\Omega,\ \mathfrak{F},\ \mathsf{P})$ с потоком σ-алгебр $\{\mathfrak{F}_t,\ t \in \mathbf{R}_+\}$, удовлетворяющим условиям п. 9.4.1.

Говорят, что на пространстве $\Theta \times \mathbf{R}_+$ определена *пуассоновская мера* ν, если каждому множеству $C \in \mathfrak{M} \times \mathfrak{B}_+$ поставлена в соответствие случайная величина $\nu(C)$ такая, что:

1) при каждом $A \in \mathfrak{M}_0$ процесс $\nu_t(A) = \nu(A + [0,\ t])$ представляет собой однородный пуассоновский процесс с параметром $\Pi(A)$, согласованный с потоком σ-алгебр $\mathfrak{F}_t$, т. е. процесс $\nu_t(A)$, $t \in \mathbf{R}_+$,

является возрастающим процессом с независимыми приращениями, принимающим целые неотрицательные значения, причем

$$\mathrm{P}\{\nu_t(A) - \nu_s(A) = k\} = \frac{[(t-s)\Pi(A)]^k}{k!}\, e^{-(t-s)\Pi(A)},$$

$0 \leqslant s < t$, $k = 0, 1, 2, \ldots$, и, кроме того, при всех $t \geqslant 0$ величина $\nu_t(A)$ измерима относительно σ-алгебры $\mathfrak{F}_t$, а приращения $\nu_{t+s}(A) - \nu_t(A)$ при $s \geqslant 0$ не зависят от σ-алгебры $\mathfrak{F}_t$;

2) для любых n и непересекающихся множеств $A_i \in \mathfrak{M}_0$, $i = 1, 2, \ldots, n$, процессы $\{\nu_t(A_i),\ t \in \mathrm{R}_+\}$ независимы в совокупности;

3) для всякой последовательности множеств $C_1, C_2, \ldots, C_i \in \mathfrak{M} \times \mathfrak{B}_+$ таких, что $C_i \cap C_j = \varnothing$ при $i \neq j$, с вероятностью 1 справедливо равенство

$$\nu\left(\bigcup_{i=1}^{\infty} C_i\right) = \sum_{i=1}^{\infty} \nu(C_i).$$

Если ν — пуассоновская мера на $\Theta \times \mathrm{R}_+$, то ее компенсатором служит мера на σ-алгебре $\mathfrak{M} \times \mathfrak{B}_+$, являющаяся произведением меры Π и лебеговой меры на $\mathfrak{B}_+$, так что мера $\tilde{\nu}(d\theta \times dt) = \nu(d\theta \times dt) - dt\Pi(d\theta)$ является мартингальной мерой (см. п. 13.4.4). При любом $A \in \mathfrak{M}_0$ процесс $\tilde{\nu}_t(A) = \tilde{\nu}(A \times [0, t]) = \nu(A \times [0, t]) - t\Pi(A)$, $t \in \mathrm{R}_+$, является квадратично интегрируемым мартингалом относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in \mathrm{R}_+\}$ с характеристикой $\langle\tilde{\nu}\rangle_t(A) = t\Pi(A)$.

19.4.2. Стохастические интегралы по пуассоновским мерам. Обозначим через $H_2(\Pi)$ совокупность случайных функций $\varphi(\theta, t) = \varphi(\theta, t, \omega)$, $\theta \in \Theta$, $t \in \mathrm{R}_+$, $\omega \in \Omega$, измеримых относительно $\mathfrak{M} \times \mathscr{P}$ ($\mathscr{P}$ — предсказуемая σ-алгебра), для которых с вероятностью 1

$$\int_\Theta \int_0^T \varphi^2(\theta, t)\, \Pi(d\theta)\, dt < \infty$$

при любом $T < \infty$. В соответствии с п. 13.4.4 для $\varphi \in H_2(\Pi)$ определен стохастический интеграл

$$I_t(\varphi) = \int_0^t \int_\Theta \varphi(\theta, s)\, \tilde{\nu}(d\theta \times ds), \quad t \in \mathrm{R}_+,$$

обладающий свойствами:

а) если $\varphi_1, \varphi_2 \in H_2(\Pi)$ и α_1, α_2 — произвольные постоянные, то с вероятностью 1 для всех t одновременно

$$I_t(\alpha_1\varphi_1 + \alpha_2\varphi_2) = \alpha_1 I_t(\varphi_1) + \alpha_2 I_t(\varphi_2);$$

б) при $\varphi \in H_2(\Pi)$ процесс $I_t(\varphi)$, $t \in \mathrm{R}_+$, непрерывен справа, имеет пределы слева и представляет собой локальный квадратично интегрируемый мартингал относительно потока $\mathfrak{F}_t$ (см. п. 13.3.2) с характеристикой

$$\langle I(\varphi)\rangle_t = \int_0^t \int_\Theta \varphi^2(\theta, s)\,\Pi(d\theta)\,ds;$$

если $\varphi_1, \varphi_2 \in H_2(\Pi)$, то взаимная характеристика стохастических интегралов $I_t(\varphi_1)$ и $I_t(\varphi_2)$ выражается формулой

$$\langle I(\varphi_1), I(\varphi_2)\rangle = \int_0^t \int_\Theta \varphi_1(\theta, s)\,\varphi_2(\theta, s)\,\Pi(d\theta)\,ds;$$

в) если $\varphi(\theta, s) = \chi_A(\theta)\,\chi_{[t_1, t_2]}(s)\,\xi$, где $A \in \mathfrak{M}_0$, ξ — $\mathfrak{F}_{t_1-}$-измеримая величина ($\chi_\Gamma(\cdot)$ — индикатор множества Γ), то

$$I_t(\varphi) = \xi\tilde{\nu}(A \times \{[t_1, t_2] \cap [0, t]\});$$

г) если $\varphi \in H_2(\Pi)$ и $\int_0^T \int_\Theta \mathsf{M}\varphi^2(\theta, t)\,\Pi(d\theta)\,dt < \infty$ при некотором $T > 0$, то при всех $t \leqslant T$

$$\mathsf{M}I_t(\varphi) = 0, \quad \mathsf{M}(I_t(\varphi))^2 = \int_0^t \int_\Theta \mathsf{M}\varphi^2(\theta, t)\,\Pi(d\theta)\,dt,$$

кроме того,

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant T}\left|\int_0^t \int_\Theta \varphi(\theta, s)\,\tilde{\nu}(d\theta \times ds)\right| > C\right\} \leqslant \frac{1}{C^2}\int_0^T \int_\Theta \mathsf{M}\varphi^2(\theta, t)\,\Pi(d\theta)\,dt,$$

$$\mathsf{M}\sup_{0\leqslant t\leqslant T}\left|\int_0^t \int_\Theta \varphi(\theta, t)\,\tilde{\nu}(d\theta \times dt)\right|^2 \leqslant 4\int_0^T \int_\Theta \mathsf{M}\varphi^2(\theta, t)\,\Pi(d\theta)\,dt;$$

д) если $\varphi \in H_2(\Pi)$, то при любых $N > 0$, $C > 0$, $T > 0$

$$\mathsf{P}\left\{\sup_{0\leqslant t\leqslant T}\left|\int_0^t \int_\Theta \varphi(\theta, s)\,\tilde{\nu}(d\theta \times ds)\right| > C\right\} \leqslant \frac{N}{C^2} + \mathsf{P}\left\{\int_0^T \int_\Theta \varphi^2(\theta, s)\,\Pi(d\theta)\,ds > N\right\}.$$

Обозначим через $H_1(\Pi)$ совокупность всех случайных функций $f(\theta, t) = f(\theta, t, \omega)$, $\theta \in \Theta$, $t \in \mathbf{R}_+$, $\omega \in \Omega$, измеримых относительно $\mathfrak{M} \times \mathfrak{B}_+ \times \mathfrak{F}$, для которых при всех $T < \infty$ с вероятностью 1

$$\int_0^T \int_\Theta |f(\theta, t)| \, \Pi(d\theta) \, dt < \infty.$$

Для $f \in H_1(\Pi) \cap H_2(\Pi)$ определен стохастический интеграл

$$\int_0^t \int_\Theta f(\theta, s) \, \nu(d\theta \times ds) = \int_0^t \int_\Theta f(\theta, s) \, \tilde{\nu}(d\theta \times ds) + \int_0^t \int_\Theta f(\theta, s) \, \Pi(d\theta) ds.$$

Интеграл по мере ν можно определить и для более широкого класса функций (см. п. 13.4.4).

19.4.3. Формула Ито. Пусть на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ с потоком σ-алгебр $\{\mathfrak{F}_t, t \in \mathbf{R}_+\}$, удовлетворяющим условиям п. 9.4.1, заданы l-мерный винеровский процесс $(w(t), \mathfrak{F}_t)$, $t \in R_+$, и независящая от процесса $w(t)$ пуассоновская мера ν на $(\Theta \times \mathbf{R}_+, \mathfrak{M} \times \mathfrak{B}_+)$, для которой $\mathsf{M}\nu(d\theta \times dt) = \Pi(d\theta) \, dt$ (здесь Π — некоторая σ-конечная мера на $\mathfrak{M}$).

Через $\tilde{\nu}(d\theta + dt)$ обозначим соответствующую мартингальную меру: $\tilde{\nu}(d\theta \times dt) = \nu(d\theta \times dt) - \Pi(d\theta) dt$. Предположим далее, что $\Theta = \Theta_1 \cup \Theta_2$, где $\Theta_1, \Theta_2 \in \mathfrak{M}$, $\Theta_1 \cap \Theta_2 = \varnothing$ и $\Pi(\Theta_2) < \infty$.

Будем говорить, что некий m-мерный процесс $\zeta(t)$, $t \in \mathbf{R}_+$, согласованный с потоком σ-алгебр $\{\mathfrak{F}_t, t \in \mathbf{R}_+\}$, допускает стохастический дифференциал

$$d\zeta(t) = \alpha(t) \, dt + \beta(t) \, dw(t) + \\ + \int_{\Theta_1} \gamma(\theta, t) \, \tilde{\nu}(d\theta \times dt) + \int_{\Theta_2} \gamma(\theta, t) \, \nu(d\theta \times dt),$$

если он с вероятностью 1 при всех t одновременно представим в виде

$$\zeta(t) = \zeta(0) + \int_0^t \alpha(s) \, ds + \int_0^t \beta(s) \, dw(s) + \\ + \int_0^t \int_{\Theta_1} \gamma(\theta, s) \, \tilde{\nu}(d\theta \times ds) + \int_0^t \int_{\Theta_2} \gamma(\theta, s) \, \nu(d\theta \times ds),$$

где $\alpha(t) = (\alpha^1(t), \ldots, \alpha^m(t))$ — m-мерный прогрессивно измеримый относительно потока σ-алгебр $\{\mathfrak{F}_t, t \in \mathbf{R}_+\}$ процесс, для которого с вероятностью 1 при всех $T < \infty$

$$\int_0^T |\alpha(t)| \, dt < \infty;$$

$\beta(t) = (\beta_{kj}(t))$ $(k = 1, 2, \ldots, m;\ j = 1, 2, \ldots, l)$ — матричнозначный (порядка $m \times l$) прогрессивно измеримый относительно потока σ-алгебр $\{\mathfrak{F}_t,\ t \in \mathbf{R}_+\}$ процесс, для которого с вероятностью 1 при всех $T < \infty$

$$\sum_{i=1}^{m} \sum_{j=1}^{l} \int_0^T (\beta_{ij}(t))^2\, dt < \infty;$$

$\gamma(\theta, t) = (\gamma^1(\theta, t), \ldots, \gamma^m(\theta, t))$ $(t \in \mathbf{R}_+,\ \theta \in \Theta)$ — m-мерный случайный процесс, измеримый относительно σ-алгебры $\mathfrak{M} \times \mathscr{P}$ ($\mathscr{P}$ — предсказуемая σ-алгебра) и такой, что с вероятностью 1 при всех $T < \infty$

$$\int_0^T \int_\Theta |\gamma(\theta, t)|^2\, \Pi(d\theta)\, dt < \infty.$$

Предположим, что задана функция $f(x)$, $x \in \mathbf{R}^m$, с вещественными значениями, дважды непрерывно дифференцируемая по x, и пусть m-мерный процесс $\zeta(t)$, $t \in \mathbf{R}_+$, допускает стохастический дифференциал в указанном смысле. Тогда процесс $f(\zeta(t))$, $t \in \mathbf{R}_+$, также допускает стохастический дифференциал, причем

$$df(\zeta(t)) = \left\{ \sum_{k=1}^{m} \frac{\partial f(\zeta(t))}{\partial x^k} \alpha^k(t) + \frac{1}{2} \sum_{j,\, k=1}^{\infty} \sum_{i=1}^{l} \left\{ \frac{\partial^2 f(\zeta(t))}{\partial x^j \partial x^k} \beta_{ji}(t) \beta_{ki}(t) + \right.\right.$$

$$\left.\left. + \int_\Theta \left[f(\zeta(t) + \gamma(\theta, t)) - f(\zeta(t)) - \sum_{k=1}^{m} \frac{\partial f(\zeta(t))}{\partial x^k} \gamma^k(\theta, t) \right] \Pi(d\theta) \right\} dt + \right.$$

$$+ \sum_{k=1}^{m} \sum_{j=1}^{l} \frac{\partial f(\zeta(t))}{\partial x^k} \beta_{kj}(t)\, dw^j(t) +$$

$$+ \int_{\Theta_1} [f(\zeta(t-) + \gamma(\theta, t)) - f(\zeta(t-))]\, \tilde{\nu}(d\theta \times dt) +$$

$$+ \int_{\Theta_2} [f(\zeta(t-) + \gamma(\theta, t)) - f(\zeta(t-))]\, \nu(d\theta \times dt).$$

Эта формула называется *формулой Ито* (см. п. 13.4.6) или *формулой дифференцирования сложной функции* (употребляется также термин *формула замены переменной*).

Обозначим через $f'(x)$ вектор с координатами $\dfrac{\partial f(x)}{\partial x^k}$ $(k = 1, 2, \ldots, m)$, а через $f''_{xx}(x)$ матрицу с элементами $\dfrac{\partial^2 f(x)}{\partial x^j \partial x^k}$ $(j,\ k = 1,$

2, ..., m). Тогда формула Ито может быть записана в виде

$$f(\zeta(t))-f(\zeta(0))=\int_0^t\Big\{\big(f'_x(\zeta(s)),\alpha(s)\big)+\frac{1}{2}\operatorname{Sp}\big(\beta^T(s)f''_{xx}(\zeta(s))\beta(s)\big)+$$

$$+\int_{\Theta_1}\big[f(\zeta(s)+\gamma(\theta,s))-f(\zeta(s))-\big(f'_x(\zeta(s)),\gamma(\theta,s)\big)\big]\Pi(d\theta)\Big\}ds+$$

$$+\int_0^t\big(\beta^T(s)f'_x(\zeta(s)),dw(s)\big)+$$

$$+\int_0^t\int_{\Theta_1}[f(\zeta(s-)+\gamma(0,s))-f(\zeta(s-))]\tilde{\nu}(d\theta\times ds)+$$

$$+\int_0^t\int_{\Theta_2}[f(\zeta(s-)+\gamma(0,s))-f(\zeta(s-)))]\nu(d\theta\times ds).$$

19.4.4. Стохастическое дифференциальное уравнение. Марковский случай. Пусть на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ с потоком σ-алгебр $(\mathfrak{F}_t,\ t\in \mathbf{R}_+)$ (удовлетворяющим условиям п. 9.4.1) заданы l-мерный винеровский процесс $(w(t), \mathfrak{F}_t)$, $t\in\mathbf{R}_+$, и не зависящая от него согласованная с потоком $\mathfrak{F}_t$ пуассоновская мера ν на $(\Theta\times\mathbf{R}_+, \mathfrak{M}\times\mathfrak{B}_+)$, удовлетворяющая всем условиям, сформулированным в начале п. 19.4.3. Пусть, кроме того, заданы:

1) векторнозначная (размерности m) локально ограниченная функция $a(t, x)$, $t\in\mathbf{R}_+$, $x\in\mathbf{R}^m$, измеримая по совокупности переменных;

2) матричнозначная (порядка $m\times l$) локально ограниченная функция $\sigma(t, x)$, $t\in\mathbf{R}_+$, $x\in\mathbf{R}^m$, измеримая по совокупности переменных;

3) векторнозначная (размерности m) функция $f(\theta, t, x)$ на $\Theta\times\mathbf{R}_+\times\mathbf{R}^m$, измеримая по совокупности переменных и такая, что функция

$$\int_\Theta |f(\theta, t, x)|^2\,\Pi(d\theta)$$

является локально ограниченной;

4) $\mathfrak{F}_0$-измеримый случайный вектор $\xi_0\in\mathbf{R}^m$.

Рассмотрим стохастическое дифференциальное уравнение

$$d\xi(t)=a(t,\xi(t))\,dt+\sigma(t,\xi(t))\,dw(t)+$$

$$+\int_{\Theta_1}f(\theta,t,\xi(t-))\,\tilde{\nu}(d\theta\times dt)+\int_{\Theta_2}f(\theta,t,\xi(t-))\,\nu(d\theta\times dt)\tag{4.1}$$

с начальным условием $\xi(0) = \xi_0$. Его можно записать в интегральной форме:

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(s))\,ds + \int_0^t \sigma(s, \xi(s))\,dw(s) +$$

$$+ \int_0^t \int_{\Theta_1} f(\theta, s, \xi(s-))\,\tilde{\nu}(d\theta \times ds) + \int_0^t \int_{\Theta_2} f(\theta, s, \xi(s-))\,\nu(d\theta \times ds). \tag{4.2}$$

Назовем *решением уравнения* (4.1) с начальным условием $\xi(0) = \xi_0$ всякий непрерывный справа и имеющий пределы слева прогрессивно измеримый относительно потока σ-алгебр $\{\mathfrak{F}_t, t \in R_+\}$ процесс $\xi(t)$, $t \in R_+$, со значениями в R^m, такой, что все интегралы, входящие в (4.2), существуют при всех $t < \infty$, а само соотношение (4.2) выполняется с вероятностью 1 для всех $t \in R_+$ одновременно. Скажем, что уравнение (4.1) с заданным начальным условием имеет единственное согласованное с потоком σ-алгебр $\{\mathfrak{F}_t, t \in R_+\}$ решение, если из того, что $(\xi(t), \mathfrak{F}_t)$ и $(\tilde{\xi}(t), \mathfrak{F}_t)$ — решения уравнения (4.1) с начальным условием ξ_0, следует соотношение

$$\mathsf{P}\left\{\sup_{t \in R_+} |\xi(t) - \tilde{\xi}(t)| > 0\right\} = 0.$$

Прежде чем сформулировать теорему существования и единственности решения уравнения (4.1), сделаем следующее замечание.

Так как по предположению $\Pi(\Theta_2) < \infty$, то процесс $\nu(\Theta_2 \times [0, T])$ на каждом конечном промежутке времени имеет лишь конечное число скачков. Обозначим моменты скачков в порядке возрастания через $\tau_1 < \tau_2 < \ldots$. Через $\theta_1, \theta_2, \ldots$ обозначим Θ_2-значные $\mathfrak{F}_t$-измеримые случайные элементы, для которых $\nu(\{\theta_k\} \times \{\tau_k\}) = 1$. Положим еще $\tau_0 = 0$. Уравнение (4.1) можно решать последовательно на каждом из полуинтервалов $[\tau_k, \tau_{k+1})$, $k = 0, 1, 2, \ldots$ На полуинтервале $[\tau_k, \tau_{k+1})$ уравнение (4.1) будет эквивалентно уравнению

$$\xi(t) = \xi(\tau_k) + \int_{\tau_k}^t a(s, \xi(s))\,ds + \int_{\tau_k}^t \sigma(s, \xi(s))\,dw(s) +$$

$$+ \int_{\tau_k}^t \int_{\Theta_1} f(\theta, s, \xi(s-))\,\tilde{\nu}(d\theta \times ds).$$

Пусть $\tilde{\xi}(t)$ $(t \geqslant \tau_k)$ — решение этого уравнения. Тогда при построении решения $\xi(t)$ уравнения (4.1) мы можем положить $\xi(t) = \tilde{\xi}(t)$ для $t \in [\tau_k, \tau_{k+1})$. Чтобы получить значение $\xi(\tau_{k+1})$, нужно к вектору $\xi(\tau_{k+1}-) = \lim_{t \uparrow \tau_{k+1}} \tilde{\xi}(t)$ прибавить вектор $f(\theta_{k+1}, \tau_{k+1}, \xi(\tau_{k+1}-))$. Затем точно так же уравнение (4.1) решается на полуинтервале $[\tau_{k+1}, \tau_{k+2})$ и так далее.

Теорема 1. *Предположим, что входящие в уравнение* (4.2) *объекты удовлетворяют следующим условиям:*

1) *для всякого* $R > 0$ *существует такая постоянная* C_R, *что*

$$|a(t, x) - a(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 + \int_{\Theta_1} |f(\theta, t, x) - f(\theta, t, y)|^2 \Pi(d\theta) \leqslant C_R |x - y|^2$$

при $t \leqslant R$, $|x| \leqslant R$, $|y| \leqslant R$ (*если* B — *матрица, то* $|B|^2$ *означает* $\operatorname{Sp} BB^T$);

2) *для всякого* $T > 0$ *существует постоянная* K_T, *такая, что при* $t \in [0, T]$, $x \in \mathbf{R}^m$, $y \in \mathbf{R}^m$

$$|a(t, x)|^2 + |\sigma(t, x)|^2 + \int_{\Theta} |f(\theta, t, x)|^2 \Pi(d\theta) \leqslant K_T (1 + |x|^2);$$

3) *каковы бы ни были* $t_0 \in \mathbf{R}_+$, $x_0 \in \mathbf{R}^m$ *и* $\varepsilon > 0$, *справедливо соотношение*

$$\lim_{\substack{t \to t_0 \\ x \to x_0}} \Pi\{\theta \in \Theta_2 \colon |f(\theta, t, x) - f(\theta, t_0, x_0)| > \varepsilon\} = 0$$

(*другими словами, сужение функции* f *на множество* $\Theta_2 \times \mathbf{R}_+ \times \mathbf{R}^m$ *непрерывно по совокупности переменных* t, x *в смысле сходимости по сужению меры* Π *на множество* Θ_2);

4) *случайный вектор* ξ_0 *не зависит от совокупности величин* $\{w(t), \nu(d\theta \times dt)\}$.

Пусть, далее, $\mathfrak{F}_t^*$ *обозначает* σ-*алгебру событий, порожденную величинами* ξ_0 *и* $\{w(s), \nu(d\theta \times [0, s]), s \leqslant t\}$. *Тогда существует единственное* $\mathfrak{F}_t^*$-*измеримое решение уравнения* (4.2), *не имеющее разрывов второго рода и непрерывное справа.*

Для уравнений (4.1) справедлив следующий результат о локальной зависимости решений от коэффициентов уравнения (ср. замечание 1 к теореме 1 п. 19.3.1).

Пусть ξ_t $(t \in \mathbf{R}_+)$ — решение уравнения (4.2) (предполагается, что выполнены условия теоремы 1), а $\bar\xi(t)$ $(t \in \mathbf{R}_+)$ — решение уравнения

$$\bar\xi(t) = \xi_0 + \int_0^t \bar a(s, \bar\xi(s))\, ds + \int_0^t \bar\sigma(s, \bar\xi(s))\, dw(s) +$$

$$+ \int_0^t \int_{\Theta_1} \bar f(\theta, s, \bar\xi(s-))\, \tilde\nu(d\theta \times ds) + \int_0^t \int_{\Theta_2} \bar f(\theta, s, \bar\xi(s-))\, \nu(d\theta \times ds),$$

относительно которого также предполагаются выполненными условия теоремы 1.

Пусть, кроме того, $a(s, x) = \bar a(s, x)$, $\sigma(s, x) = \bar\sigma(s, x)$, $f(\theta, s, x) = \bar f(\theta, s, x)$ при $|x| \leqslant N$. Тогда с вероятностью 1 $\xi(s) = \bar\xi(s)$ для $s \leqslant \tau$, где $\tau = \inf\{s \in \mathbf{R}_+ \colon |\xi(s)| \geqslant N\}$.

Можно показать, что в условиях теоремы 1 решение уравнения (4.2) обладает марковским свойством, т. е. представляет собой марковскую случайную функцию. Ее вероятность перехода определяется формулой

$$P(s, x, t, \Gamma) = \mathsf{P}\{\xi_{sx}(t) \in \Gamma\}, \quad 0 \leqslant s \leqslant t, \quad x \in \mathbf{R}^m, \ \Gamma \in \mathfrak{B},$$

где $\mathfrak{B}$ — σ-алгебра борелевских подмножеств $\mathbf{R}^m$, а $\xi_{sx}(t)$ ($t \in [s, \infty)$) — решение уравнения

$$\xi_{sx}(t) = x + \int_s^t a(\tau, \xi_{sx}(\tau))\, d\tau + \int_s^t \sigma(\tau, \xi_{sx}(\tau))\, dw(\tau) +$$
$$+ \int_s^t \int_{\Theta_1} f(\theta, \tau, \xi_{sx}(\tau-))\, \tilde{\nu}(d\theta \times d\tau) + \int_s^t \int_{\Theta_2} f(\theta, \tau, \xi_{sx}(\tau-))\, \nu(d\theta \times d\tau). \tag{4.3}$$

На уравнения вида (4.1) очевидным образом переносятся понятия сильных и слабых решений, сильной и слабой единственности решения (см. п. 19.3.6). Теорема 1, таким образом, содержит условия, при которых существует сильное решение уравнения (4.1), и оно единственно по траекториям. Весьма широкие условия существования слабого решения уравнения (4.1) приведены в следующей теореме.

Теорема 2. *Пусть входящие в уравнение* (4.1) *коэффициенты удовлетворяют следующим условиям*:

а) *функции* $a(t, x)$ *и* $\sigma(t, x)$ *измеримы, локально ограничены и для всех* $t \in \mathbf{R}_+$ *непрерывны по* x;

б) *при всех* $t \in \mathbf{R}_+$, $x_0 \in \mathbf{R}^m$

$$\lim_{x \to x_0} \int_{\Theta_1} |f(\theta, t, x) - f(\theta, t, x_0)|^2\, \Pi(d\theta) = 0;$$

в) *для каждого* $T > 0$ *существует такая постоянная* K_T, *что*

$$|a(t, x)|^2 + |\sigma(t, x)|^2 + \int_{\Theta_1} |f(\theta, t, x)|^2\, \Pi(d\theta) \leqslant K_T(1 + |x|^2),$$

каковы бы ни были $t \in [0, T]$ *и* $x \in \mathbf{R}^m$;

г) *при всех* $t \in \mathbf{R}_+$, $x \in \mathbf{R}^m$ *и* $\varepsilon > 0$

$$\lim_{x \to x_0} \Pi\{\theta \in \Theta_2 : |f(\theta, t, x) - f(\theta, t, x_0)| > \varepsilon\} = 0.$$

Тогда уравнение (4.1) *имеет слабое решение.*

Условия слабой единственности, а значит, и марковости решения уравнения (4.1) содержатся в следующей теореме.

Теорема 3. *Предположим, что выполнены условия*:

1) *функция* $a(t, x)$ *ограничена*;

2) *существует постоянная* q $(0 < q < 1)$ *такая, что при всех* $t \in \mathbf{R}_+$, $x \in \mathbf{R}^m$

$$\operatorname{Sp}(I - \sigma(t, x)\sigma^*(t, x))^2 \leqslant q;$$

3) *существуют меры* $\pi_1(dz)$ *и* $\pi_2(dz)$ *на* $(\mathbf{R}^m, \mathfrak{B})$ *такие, что*

$$\pi_2(\mathbf{R}^m) + \int_{\mathbf{R}^m} |z|^2 \pi_1(dz) < \infty,$$

а мера $n_k(t, x, dz)$ $(k = 1, 2)$, *определяемая формулой*

$$n_k(t, x, \Gamma) = \int_{\Theta_k} \chi_\Gamma(f(\theta, t, x))\, \Pi(d\theta), \quad t \in \mathbf{R}_+, \quad x \in \mathbf{R}^m, \quad \Gamma \in \mathfrak{B},$$

абсолютно непрерывна относительно меры $\pi_k(dz)$, *и соответствующая плотность* $g_k(t, x, z)$ $(k = 1, 2)$ *обладает тем свойством, что интегралы*

$$\int_{\mathbf{R}^m} (g_1(t, x, z))^2 |z|^2 \pi_1(dz), \qquad \int_{\mathbf{R}^m} (g_2(t, x, z))^2 \pi_2(dz)$$

представляют собой ограниченные функции переменных t, x.

Тогда решение уравнения (4.1) *слабо единственно.*

19.4.5. Уравнение Колмогорова. В этом пункте будем предполагать выполненными условия теоремы 1. Кроме того, предполагаем, что непрерывны по совокупности переменных (t, x), $t \in \mathbf{R}_+$, $x \in \mathbf{R}^m$, следующие функции:

$$a(t, x), \ \sigma(t, x), \ \int_{\Theta_1} |f(\theta, t, x)|^2 \Pi(d\theta),$$

$$\int_{\Theta_2} |f(\theta, t, x)|\, \Pi(d\theta).$$

Как уже отмечалось, решение уравнения (4.2) представляет собой марковскую случайную функцию с вероятностью перехода

$$P(s, x, t, \Gamma) = \mathrm{M}\chi_\Gamma(\xi_{sx}(t)), \quad 0 \leqslant s \leqslant t, \quad x \in \mathbf{R}^m, \quad \Gamma \in \mathfrak{B},$$

где $\xi_{sx}(t)$ — решение уравнения (4.3), а $\chi_\Gamma(x)$ — индикатор множества $\Gamma \in \mathfrak{B}$.

Пусть $\varphi(x)$ $(x \in \mathbf{R}^m)$ — дважды непрерывно дифференцируемая функция, ограниченная вместе со своими производными. Тогда с использованием формулы Ито нетрудно получить соотношение

$$\lim_{t \to s} \frac{1}{t-s} [\mathrm{M}\varphi(\xi_{sx}(t)) - \varphi(x)] =$$
$$= \left(\varphi'_x(x), a(s, x)\right) + \frac{1}{2} \mathrm{Sp}\left(\sigma^T(s, x) \varphi''_{xx} \sigma(s, x)\right) +$$
$$+ \int_{\Theta_1} \left[\varphi(x + f(\theta, s, x)) - \varphi(x) - \left(\varphi'_x(x), f(\theta, s, x)\right)\right] \Pi(d\theta) +$$
$$+ \int_{\Theta_2} [\varphi(x + f(\theta, s, x)) - \varphi(x)]\, \Pi(d\theta). \qquad (4.4)$$

Следствием соотношения (4.4) является следующее утверждение. Предположим, что непрерывная ограниченная функция $\varphi(x)$, $x \in \mathbf{R}^m$, такова, что функция

$$u(t, x) = \mathrm{M}\varphi(\xi_{tx}(T)) = \int_{\mathbf{R}^m} \varphi(y) P(t, x, T, dy)$$

при всех $t \in [0, T]$, $x \in \mathbf{R}^m$ дважды непрерывно дифференцируема по x и производные u'_x и u''_{xx} ограничены. Тогда она удовлетворяет в области $t \in [0, T]$, $x \in \mathbf{R}^m$ следующему интегро-дифференциальному уравнению

$$\frac{\partial u(t, x)}{\partial t} + \left(u'_x(t, x), a(t, x)\right) + \frac{1}{2} \operatorname{Sp}\left(\sigma^T(t, x) u''_{xx}(t, x) \sigma(t, x)\right) +$$
$$+ \int_{\Theta_1} \left[u(t, x + f(\theta, t, x)) - u(t, x) - \left(u'_x(t, x), f(\theta, t, x)\right)\right] \Pi(d\theta) +$$
$$+ \int_{\Theta_2} [u(t, x + f(\theta, t, x)) - u(t, x)] \Pi(d\theta) = 0 \quad (4.5)$$

с граничным условием

$$\lim_{t \uparrow T} u(t, x) = \varphi(x). \quad (4.6)$$

Возникает вопрос: при каких условиях на коэффициенты уравнения определенная выше функция $u(t, x)$ обладает свойствами, позволяющими записать для нее уравнение (4.5)? Ответ на этот вопрос в случае $f(\theta, t, x) \equiv 0$ при $\theta \in \Theta_2$ дает следующая теорема.

Теорема 4. *Предположим, что для всех* $\theta \in \Theta_2$ *функция* $f(\theta, t, x) \equiv 0$, *и пусть выполнены условия теоремы* 1. *Пусть, кроме того, существуют ограниченные и непрерывные по совокупности переменных производные*

$$a'_x(t, x), \quad a''_{xx}(t, x), \quad \sigma'_x(t, x), \quad \sigma''_{xx}(t, x).$$

Относительно функции $f(\theta, t, x)$ *предположим, что существуют непрерывные по* t, x *производные* $f'_x(\theta, t, x)$ *и* $f''_{xx}(\theta, t, x)$, *причем* $f'_x(\theta, t, x)$ *ограничена и интегралы*

$$\int_{\Theta_1} \left| f'_x(\theta, t, x)\right|^2 \Pi(d\theta), \qquad \int_{\Theta_1} \left| f''_{xx}(\theta, t, x)\right|^2 \Pi(d\theta)$$

также представляют собой ограниченные функции переменных t, x.

Если $\hat{\xi}_{sx}(t)$ — *решение уравнения*

$$\hat{\xi}_{sx}(t) = x + \int_s^t a(\tau, \hat{\xi}_{sx}(\tau))\, d\tau + \int_s^t \sigma(\tau, \hat{\xi}_{sx}(\tau))\, dw(\tau) +$$
$$+ \int_s^t \int_\Theta f(\theta, \tau, \hat{\xi}_{sx}(\tau-))\, \tilde{\nu}(d\theta \times d\tau), \quad t \geqslant s, \quad x \in \mathbf{R}^m,$$

а функция $\varphi(x)$, $x \in \mathbf{R}^m$, *дважды непрерывно дифференцируема и ограничена вместе со своими производными, то функция* $\mathrm{M}\varphi(\xi_{sx}(t))$ *также дважды непрерывно дифференцируема и ее производные ограничены.*

Таким образом, если $f(\theta, t, x) = 0$ для $\theta \in \Theta_2$, то соответствующая задача Коши (т. е. уравнение (4.5) без последнего интеграла в левой части с граничным условием (4.6)) имеет единственное решение. Эта задача позволяет определить вероятность перехода

$$\hat{P}(s, x, t, \Gamma) = \mathrm{M}\chi_{\Gamma}(\xi_{sx}(t)), \quad 0 \leqslant s \leqslant t, \quad x \in \mathbf{R}^m, \quad \Gamma \in \mathfrak{B},$$

марковской случайной функции, являющейся решением уравнения (4.2) (без интеграла по мере $\nu(d\theta \times dt)$). Для того чтобы определить вероятность перехода $P(s, x, t, \Gamma)$ в общем случае, можно воспользоваться следующим интегральным уравнением, связывающим функции P и $\hat{P}$:

$$P(s, x, t, \Gamma) =$$

$$= \exp\{-(t-s)\Pi(\Theta_2)\}\hat{P}(s, x, t, \Gamma) + \int_s^t \exp\{-(\tau - s)\Pi(\Theta_2)\}\,d\tau \times$$

$$\times \int_{\Theta_2} \Pi(d\theta) \int_{\mathbf{R}^m} \hat{P}(s, x, \tau, dz)\, P(\tau, z + f(\theta, \tau, z), t, \Gamma),$$

где $0 \leqslant s \leqslant t$, $x \in \mathbf{R}^m$, $\Gamma \in \mathfrak{B}$.

19.4.6. Уравнения с коэффициентами, зависящими от прошлого. В этом пункте рассмотрим стохастические дифференциальные уравнения, подобные уравнению (4.1), в котором, однако, коэффициенты представляют собой функционалы, зависящие от всех значений искомого процесса в «прошлом».

Обозначим через $D_{[0,T]}(\mathbf{R}^m)$ пространство функций $x(t)$, заданных на $[0, T]$, принимающих значения в $\mathbf{R}^m$, не имеющих разрывов второго рода, непрерывных справа в каждой точке $t \in [0, T)$, а в точке T непрерывных слева. Через $\mathfrak{N}_t$ $(t \in [0, T])$ обозначим наименьшую σ-алгебру подмножеств $D_{[0,T]}(\mathbf{R}^m)$, содержащую все множества вида $\{x(\cdot): x(s) \in \Gamma\}$ при $s \leqslant t$, $\Gamma \in \mathfrak{B}$. В пространстве $D_{[0,T]}(\mathbf{R}^m)$ можно ввести метрику

$$\rho(x(\cdot), y(\cdot)) = \inf_g \Big(\sup_{0 \leqslant t \leqslant T} [\,|x(t) - y(g(t))| + |t - g(t)|\,]\Big),$$

где inf берется по всем непрерывным монотонным функциям $g(t)$, $t \in [0, T]$, для которых $g(0) = 0$, $g(T) = T$.

Тогда σ-алгебра $\mathfrak{N}_T$ совпадает с σ-алгеброй борелевских подмножеств $D_{[0,T]}(\mathbf{R}^m)$ в метрике ρ. Символом $D(\mathbf{R}^m)$ обозначаем пространство всех непрерывных справа функций $x(t)$ со значениями в $\mathbf{R}^m$, заданных на $[0, \infty)$ и не имеющих разрывов второго рода; $\mathfrak{N}$ будет означать минимальную σ-алгебру подмножеств $D(\mathbf{R}^m)$, содержащую все σ-алгебры $\mathfrak{N}_t$ при $t \to \infty$, а $\mathfrak{N}_t$ — наименьшую σ-алгебру подмножеств $D(\mathbf{R}^m)$, содержащую все σ-алгебры $\mathfrak{N}_s$ при $s < t$.

Винеровский процесс $w(t)$ и пуассоновская мера $\nu(d\theta \times dt)$ считаются заданными точно так же, как это было в пп. 19.4.3, 19.4.4. Пусть заданы коэффициенты:

1) функция $a(t, x(\cdot))$ на $\mathbf{R}_+ \times D(\mathbf{R}^m)$ со значениями в $\mathbf{R}^m$, измеримая относительно $\mathfrak{B}_+ \times \mathfrak{N}$, при каждом $t \in \mathbf{R}_+$ измеримая относительно $\mathfrak{N}_{t-}$, при каждом $x(\cdot) \in D(\mathbf{R}^m)$ локально ограниченная по t;

2) матричнозначная (порядка $m \times l$) функция $\sigma(t, x(\cdot))$, определенная на $\mathbf{R}_+ \times D(\mathbf{R}^m)$, измеримая относительно $\mathfrak{B}_+ \times \mathfrak{N}$, при каждом $t \in \mathbf{R}_+$ измеримая относительно $\mathfrak{N}_{t-}$, при каждом $x(\cdot) \in D(\mathbf{R}^m)$ локально ограниченная по t;

3) векторнозначная (размерности m) функция $f(\theta, t, x(\cdot))$, определенная на $\Theta \times \mathbf{R}_+ \times D(\mathbf{R}^m)$, измеримая относительно $\mathfrak{M} \times \mathfrak{B}_+ \times \mathfrak{N}$, при каждом t измеримая относительно $\mathfrak{M} \times \mathfrak{N}_{t-}$, такая, что интеграл

$$\int_{\Theta} |f(\theta, t, x(\cdot))|^2 \Pi(d\theta)$$

представляет собой при любом $x(\cdot) \in D(\mathbf{R}^m)$ локально ограниченную функцию переменной t.

Рассмотрим уравнение (ξ_0 — $\mathfrak{F}_0$-измеримый вектор в $\mathbf{R}^m$)

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(\cdot))\, ds + \int_0^t \sigma(s, \xi(\cdot))\, dw(s) + $$
$$+ \int_0^t \int_{\Theta_1} f(\theta, s, \xi(\cdot))\, \tilde{\nu}(d\theta \times ds) + \int_0^t \int_{\Theta_2} f(\theta, s, \xi(\cdot))\, \nu(d\theta \times ds). \tag{4.7}$$

Заметим, что интегралы в правой части этого уравнения определены, если процесс ξ таков, что его траектории с вероятностью 1 принадлежат пространству $D(\mathbf{R}^m)$, а сам он согласован с потоком σ-алгебр $\{\mathfrak{F}_t, t \in \mathbf{R}_+\}$ (согласно предположению, винеровский процесс $w(t)$ и пуассоновская мера $\nu(d\theta \times dt)$ согласованы с потоком σ-алгебр $\{\mathfrak{F}_t, t \in \mathbf{R}^m\}$). С другой стороны, сами интегралы правой части (4.7) как функции переменной t представляют собой функции из $D(\mathbf{R}^m)$. Поэтому естественно назвать решением уравнения (4.7) всякий такой процесс $\xi(t)$, согласованный с потоком σ-алгебр $\mathfrak{F}_t$, траектории которого с вероятностью 1 принадлежат пространству $D(\mathbf{R}^m)$ и для которого соотношение (4.7) выполняется с вероятностью 1 при всех t одновременно. Если окажется, что на исходном вероятностном пространстве такого процесса не существует, то тогда можно рассматривать слабые решения. Очевидным образом вводится понятие потраекторной и слабой единственности.

Теорема 5. *Пусть объекты, определяющие уравнение* (4.7), *удовлетворяют условиям:*

1) *для всякого* $R > 0$ *существует такая постоянная* C_R, *что*

$$|a(t, x(\cdot)) - a(t, y(\cdot))|^2 + |\sigma(t, x(\cdot)) - \sigma(t, y(\cdot))|^2 +$$
$$+ \int_{\Theta_1} |f(\theta, t, x(\cdot)) - f(\theta, t, y(\cdot))|^2 \Pi(d\theta) \leqslant C_R \|x(\cdot) - y(\cdot)\|_t^2$$

при $t \leqslant R$, $\|x(\cdot)\|_R \leqslant R$, $\|y(\cdot)\|_R \leqslant R$, *где* $\|x(\cdot)\|_t = \sup_{s \leqslant t} |x(s)|$;

2) *для всякого* $T>0$ *существует постоянная* K_T, *такая, что при* $t\in[0, T]$

$$|a(t, x(\cdot))|^2+|\sigma(t, x(\cdot))|^2+\int_{\Theta_1}|f(\theta, t, x(\cdot))|^2\,\Pi(d\theta)\leqslant$$
$$\leqslant K_T\left(1+\|x(\cdot)\|_t^2\right);$$

3) *каковы бы ни были* $t_0\in\mathbf{R}_+$, $x_0(\cdot)\in D(\mathbf{R}^m)$, $\varepsilon>0$, *справедливо соотношение*

$$\Pi\{\theta\in\Theta_2:\ |f(\theta, t, x(\cdot))-f(\theta, t_0, x_0(\cdot))|>\varepsilon\}\to 0$$

при $t\to t_0$, $\rho(x(\cdot), x_0(\cdot))\to 0$;

4) *случайный вектор* ξ_0 *не зависит от совокупности величин* $\{w(t), \nu(d\theta\times dt)\}$.

Тогда существует единственное $\mathfrak{F}_t^*$-*измеримое решение уравнения* (4.7), *траектории которого с вероятностью* 1 *принадлежат пространству* $D(\mathbf{R}^m)$.

(Как и в теореме 1, σ-алгебра $\mathfrak{F}_t^*$ — это наименьшая σ-алгебра событий, относительно которой измеримы величины ξ_0 и $\{w(s), \nu(d\theta\times[0, s]),\ s\leqslant t\}$.)

Таким образом, сформулированная теорема содержит условия существования сильного решения уравнения (4.7) и его потраекторной единственности. Следующая теорема дает условия существования слабого решения.

Теорема 6. *Предположим, что коэффициенты уравнения* (4.7) *удовлетворяют условиям:*

а) *функции* $a(t, x(\cdot))$ *и* $\sigma(t, x(\cdot))$ *непрерывны по* $x(\cdot)$ *в метрике пространства* $D_{[0, T]}(\mathbf{R}^m)$ *при каждом* $t\in[0, T]$;

б) *для всех* $t\in[0, T]$, $x_0(\cdot)\in D_{[0, T]}(\mathbf{R}^m)$

$$\int_{\Theta_1}|f(\theta, t, x(\cdot))-f(\theta, t, x_0(\cdot))|^2\,\Pi(d\theta)\to 0$$

при $\rho(x(\cdot), x_0(\cdot))\to 0$;

в) *для всех* $t\in[0, T]$, $x_0(\cdot)\in D_{[0, T]}(\mathbf{R}^m)$ *и* $\varepsilon>0$

$$\Pi\{\theta\in\Theta_2:\ |f(\theta, t, x(\cdot))-f(\theta, t, x_0(\cdot))|>\varepsilon\}$$

при $\rho(x(\cdot), x_0(\cdot))\to 0$;

г) *существует постоянная* K, *такая, что*

$$|a(t, x(\cdot))|^2+|\sigma(t, x(\cdot))|^2+\int_{\Theta_1}|f(\theta, t, x(\cdot))|^2\,\Pi(d\theta)\leqslant$$
$$\leqslant K\left(1+\|x(\cdot)\|_t^2\right)$$

при всех $t\in[0, T]$, $x(\cdot)\in D_{[0, T]}(\mathbf{R}^m)$.

Тогда уравнение (4.7) *имеет на отрезке* $[0, T]$ *слабое решение.*

Литература: [21, 22, 32, 52, 59, 61, 66, 71, 79, 82, 108, 109, 111, 118].

Часть третья

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Глава 20. ОСНОВНЫЕ ПОНЯТИЯ И ЗАДАЧИ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

20.1. Статистическая структура

Если $x_1, \ldots, x_n$ — результаты наблюдений (измерений), полученных в ходе выполнения n независимых повторений случайного эксперимента, связанного со случайной величиной ξ с неизвестным распределением P, то вектор $y_n = (x_1, \ldots, x_n)$ называется *выборкой* (*простой*) объема n из генеральной совокупности с распределением P. Символически модель простой выборки записывается в виде тройки $(\mathbf{R}^n, \mathfrak{S}^n, P^n)$, называемой *выборочным пространством.*

Задача математической статистики состоит в том, чтобы на основе анализа выборки сделать научно обоснованное заключение о распределении P.

Обычно постулируется, что неизвестное распределение принадлежит известному семейству (классу) распределений $\mathcal{P}$. Тройка $(\mathbf{R}^n, \mathfrak{S}^n, P^n)$ дает пример одного из фундаментальных понятий математической статистики — *статистической структуры.*

Пусть $(X, \mathfrak{X})$ — измеримое пространство, $\mathcal{P}$ — семейство вероятностных мер на $\mathfrak{X}$. Тройка $(X, \mathfrak{X}, \mathcal{P})$ называется *статистической структурой.*

В большинстве конкретных рассмотрений семейство вероятностных распределений $\mathcal{P}$ параметризовано:

$$\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}, \tag{1.1}$$

где Θ — произвольное множество, P_θ — однозначно определяемое распределение при известном значении параметра θ. Основная статистическая структура классической математической статистики — это структура, отвечающая модели простой выборки вида

$$(X^n, \mathfrak{X}^n, \{P_\theta,\ \theta \in \Theta\}),$$

где $X \subset \mathbf{R}$, Θ — область в $\mathbf{R}^k$, $k < n$, $P_\theta = \Pi_\theta^n$ — n-кратное произведение распределений Π_θ на X.

Пусть $(X, \mathfrak{X}, \mathcal{P})$ — статистическая структура. Если существует положительная σ-конечная мера ν на $\mathfrak{X}$ такая, что каждое распределение P на $\mathcal{P}$ абсолютно непреравно относительно меры ν ($P \ll \mu$, см. п. 9.4.2), то структуру $(X, \mathfrak{X}, \mathcal{P})$ называют *доминируемой.*

Всякая статистическая структура, у которой X или (и) семейство $\mathcal{P}$ не более чем счетны, доминируема.

В том случае, когда семейство $\mathscr{P}$ параметризовано (см. (1.1)) и структура $(X, \mathfrak{X}, \mathscr{P})$ доминируема, чрезвычайно полезным инструментом для получения статистических выводов является понятие *функции правдоподобия* $L_\theta(x)$:

$$L_\theta(x) = \frac{dP_\theta}{d\nu}(x).$$

В частности, оказывается, что для многих дискретных и непрерывных распределений (см. § 6.1, 6.2) функция правдоподобия допускает представление

$$L_\theta(x) = C(\theta)\exp\left\{\sum_{j=1}^{s} Q_j(\theta)\,T_j(x)\right\}h(x), \quad \theta \in \Theta, \qquad (1.2)$$

где функции $C(\theta)$ и $Q_j(\theta)$ $(j = 1, \ldots, s)$ зависят от θ и не зависят от x, а $T_j(x)$ $(j = 1, \ldots, s)$ и $h(x)$ зависят от x и не зависят от θ.

Доминируемая статистическая структура $(\mathbf{R}^n, \mathfrak{S}^n, \{P_\theta, \theta \in \Theta\})$, у которой функция правдоподобия $L_\theta(x)$ допускает представление (1.2) и множество $\{x \in \mathbf{R}^n\colon L_\theta(x) > 0\}$ не зависит от $\theta \in \Theta$, называется *экспоненциальной структурой*.

Примеры экспоненциальных структур.

1. Проводятся n наблюдений $x_1, x_2, \ldots, x_n$, результаты которых представимы в виде

$$x_i = \sum_{j=1}^{m} \varphi_j(z_i)\,\beta_j + \varepsilon_i,$$

где переменные z_i $(j = 1, \ldots, n)$ и функции $\varphi_j(z)$ $(j = 1, \ldots, m)$ известны, ε_i — реализации независимых одинаково распределенных случайных величин, имеющих нормальное распределение $N(0, \sigma^2)$ с неизвестной дисперсией, β_j $(j = 1, \ldots, m)$ — неизвестные параметры.

Статистическая структура $(X^n, \mathfrak{S}^n, \{P_\theta, \theta \in \Theta\})$, где $\Theta = \mathbf{R}^m \times \mathbf{R}_+$ и P_θ — нормальное распределение с плотностью

$$p_\theta(s) = \frac{1}{\sigma^n\sqrt{(2\pi)^n}}\exp\left\{\frac{(s - X\beta)^T(s - X\beta)}{\sigma^2}\right\},$$

здесь $s = (s_1 \ldots s_n)^T$, $\theta = (\beta_1 \ldots \beta_m, \sigma^2)$, X — $n \times m$-матрица с элементами $x_{ij} = \varphi_j(z_i)$, $\beta = (\beta_1 \ldots \beta_m)^T$, соответствует невырожденной модели линейной нормальной регрессии (см. § 23.2)

$$x = X\beta + \varepsilon.$$

Если rang $X = m$, то, взяв в качестве доминирующей меры меру Лебега, имеем

$$L_\theta(x) = \frac{1}{\sigma^n\sqrt{(2\pi)^n}}\exp\left\{\frac{s^2_{\text{ост}}}{2\sigma^2} - \frac{(\hat\beta - \beta)^T X^T X(\hat\beta - \beta)}{2\sigma^2}\right\},$$

где $\hat\beta = (x^Tx)^{-1}\beta$ ($\hat\beta$ называется *оценкой метода наименьших квадратов*), $s^2_{\text{ост}} = x^Tx - \hat\beta^T X^T X\hat\beta$ (*остаточная сумма квадратов*). Здесь (ср. (1.2))

$$C(\theta)=\frac{1}{\sigma^n\sqrt{(2\pi)^n}},\quad s=2,$$

$$T_1(x)=x^Tx=\|x\|^2,\quad Q_1(\theta)=\frac{1}{2\sigma^2},$$

$$T_2(x)\equiv 1,\quad Q_2(\theta)=-\frac{\hat\beta^TX^TX\hat\beta}{2\sigma^2}-\frac{(\hat\beta-\beta)^TX^TX(\hat\beta-\beta)}{2\sigma^2}.$$

2. Пусть P_θ — биномиальное распределение с параметрами (n, p) (см. п. 6.2.3), ν — мера на множестве целых чисел, принимающая единичное значение на любом из них.

Тогда для $x\in\{0, 1, \ldots, n\}$

$$L_\theta(x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}=(1-\theta)^n\exp\left\{x\log\frac{\theta}{1-\theta}\right\}\binom{n}{x}.$$

Здесь (ср. (1.2)) $C(\theta)=(1-\theta)^n$, $s=1$, $T_1(x)=x$,

$$Q_1(\theta)=\frac{\theta}{1-\theta},\quad h(x)=\binom{n}{x}.$$

Кроме приведенных выше, экспоненциальными являются структуры, у которых семейства $\{P_\theta, \theta\in\Theta\}$ имеют распределения (см. § 6.1, 6.2):

а) Пуассона с неизвестным средним;

б) нормальные с неизвестным средним и известной либо неизвестной дисперсией;

в) k-мерные нормальные с неизвестным вектором средних и неизвестной ковариационной матрицей;

г) гамма-распределение с неизвестными параметрами α и λ.

Примером структуры, не являющейся экспоненциальной, является структура $(\mathbf{R}^n, \mathfrak{S}^n, \{P_\theta, \theta\in\Theta\})$, где P_θ — бета-распределение с неизвестными параметрами (α, β).

20.2. Статистики

20.2.1. Определение статистики и оценки. Пусть $(X, \mathfrak{X}, \mathcal{P})$ — статистическая структура и $(Y, \mathfrak{U})$ — произвольное измеримое пространство. Любое $(\mathfrak{X}, \mathfrak{U})$-измеримое отображение T пространства $(X, \mathfrak{X})$ в $(Y, \mathfrak{U})$ называется *статистикой* на статистической структуре $(X, \mathfrak{X}, \mathcal{P})$.

Статистику T, принимающей значения в пространстве параметров Θ, называют *оценкой*.

Пусть T — вещественная статистика, заданная на статистической структуре $(X, \mathfrak{X}, \mathcal{P})$. Для всякого фиксированного распределения P из $\mathcal{P}$ статистика T как отображение вероятностного пространства $(X, \mathfrak{X}, P)$ в $(\mathbf{R}, \mathfrak{S})$ является случайной величиной.

Если статистика T имеет математическое ожидание для любого распределения P из $\mathcal{P}$, то ее называют *интегрируемой*.

20.2.2. Достаточные статистики. Пусть $(X, \mathfrak{X}, \{P_\theta, \theta\in\Theta\})$ — статистическая структура, Y — полное метрическое сепарабельное пространство, $\mathfrak{U}$ — σ-алгебра борелевских множеств, T: $(X, \mathfrak{X})\to(Y, \mathfrak{U})$ —

статистика; $\mathfrak{X}_T = T^{-1}(\mathfrak{U})$ — прообраз σ-алгебры $\mathfrak{U}$ при отображении T.

Статистика T называется *достаточной* (для параметра θ), если условная вероятность

$$P_\theta(A \mid \mathfrak{X}_T), \quad \theta \in \Theta, \quad A \in \mathfrak{X},$$

не зависит от θ.

Таким образом, если статистика T достаточна, то условная вероятность $P_\theta(A \mid \mathfrak{X}_T)$ не содержит никакой информации о параметре θ.

Пример 1. Пусть в статистической структуре $(\mathbf{R}^n, \mathfrak{S}^n, \mathscr{P}^n)$ $\mathscr{P} = \mathscr{F}$, где $\mathscr{F}$ — класс абсолютно непрерывных распределений, и статистика $T(x)$ — статистика вида

$$y_n = (x_1, \ldots, x_n) \xrightarrow{T(x)} \left(x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}\right),$$

где $\left(x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}\right)$ — упорядоченная в порядке возрастания выборка $(x_1, \ldots, x_n)$ (ее называют *вариационным рядом*, см. п. 20.4.1).

Поскольку $y_n = (x_1, \ldots, x_n)$ — выборка, то ее совместная плотность распределения имеет вид

$$p_{y_n}(x) = \prod_{j=1}^{n} f(x_j), \tag{2.1}$$

где $f(z)$ — (неизвестная) плотность.

Совместная плотность распределения вариационного ряда равна

$$p_{T(y_n)}(t) = \begin{cases} n! \prod\limits_{j=1}^{n} f(t_j), & \text{если } t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n; \\ 0 & \text{в противном случае.} \end{cases}$$

Отсюда для совместной условной плотности распределения выборки следует

$$p_{y_n}\left(x \mid T(y_n) = t\right) = \begin{cases} 1/n!, & \text{если } (x_1, \ldots, x_n) \text{ является результатом некоторой перестановки } (t_1 \ldots t_n); \\ 0 & \text{в противном случае.} \end{cases}$$

Поскольку $p_{y_n}(x \mid T(y_n) = t)$ не зависит от распределения выборки, статистика $T(x)$ достаточна.

Эффективный способ проверки достаточности статистики, не требующий вычисления условных распределений, дает следующий факторизационный критерий Неймана — Фишера.

Теорема 1. *Пусть $(X, \mathfrak{X}, \{P_\theta, \theta \in \Theta\})$ — доминируемая статистическая структура. Статистика T со значениями в $(Y, \mathfrak{U})$ является достаточной тогда и только тогда, когда существует $\mathfrak{X}$-измеримая неотрицательная функция $h(x)$ на X и $\mathfrak{U}$-измеримая неотрицательная функция $g_\theta(y)$ на Y такие, что*

$$L_\theta(x) = g_\theta(T(x))\, h(x) \tag{2.2}$$

для всех $\theta \in \Theta$, $x \in X$.

Примеры 2. Пусть $(X, \mathfrak{X}, \{P_\theta, \theta \in \Theta\})$ — доминируемая структура, $T_\theta(x)$ — статистика вида

$$T_\theta(x) = \frac{L_\theta(x)}{L_{\theta_0}(x)}, \tag{2.3}$$

называемая *отношением правдоподобия*. В (2.3) θ_0 — произвольный параметр из Θ, для которого $L_{\theta_0}(x) > 0$, если такой существует.

Статистика $T_\theta(x)$ достаточна для параметра θ, поскольку

$$L_\theta(x) = \frac{L_\theta(x)}{L_{\theta_0}(x)} L_{\theta_0}(x) = T_\theta(x) L_{\theta_0}(x),$$

что совпадает с правой частью (2.3) при $g(y) = y$, $h(x) = L_{\theta_0}(x)$.

3. Для статистической структуры модели линейной нормальной регрессии, рассмотренной в примере 1 § 20.1, статистика $\left(\hat{\beta}, \bar{s}^2_{\text{ост}}\right)$ достаточна для параметра (β, σ^2).

При известной дисперсии статистика $\hat{\beta}$ достаточна для β.

4. Пусть $(\mathbf{R}^n, \mathfrak{S}^n, \{P_\theta, \theta \in \Theta\})$ — экспоненциальная статистическая структура (см. § 20.1) с функцией правдоподобия

$$L_\theta(x) = C(\theta) \exp\left\{\sum_{j=1}^{s} Q_j(\theta) T_j(x)\right\} h(x).$$

Если область Θ содержит нетривиальный s-мерный куб и $C(\theta) > 0$ для $\theta \in \Theta$, то вектор-функция

$$T(x) = \left(\sum_{j=1}^{n} T_1(x_j), \sum_{j=1}^{n} T_2(x_j), \ldots, \sum_{j=1}^{n} T_s(x_j)\right) \tag{2.4}$$

является достаточной статистикой для параметра $\theta \in \Theta$.

20.2.3. Минимальные достаточные статистики. Если T_1 — достаточная статистика со значениями в Y, а $g(y)$ — взаимно-однозначная функция со значениями в Y, то, согласно факторизационному критерию Неймана — Фишера, $T_2 = g(T_1)$ также является достаточной статистикой.

Среди двух достаточных статистик T_1 и T_2 лучшей, естественно, считать ту из них, при значении которой вторая не добавляет новой информации о параметре.

Пусть $(X, \mathfrak{X}, \{P_\theta, \theta \in \Theta\})$ — статистическая структура, $\mathcal{T}$ — класс достаточных статистик T со значениями в $(Y, \mathfrak{U})$ (см. п. 20.2.1), $\mathfrak{X}_T = T^{-1}(\mathfrak{U})$. Достаточная статистика T_0 называется ***минимальной***, если

$$\mathfrak{X}_{T_0} \subseteq \mathfrak{X}_T \quad \text{для} \quad \forall T \in \mathcal{T}. \tag{2.5}$$

В рассмотренных выше примерах достаточные статистики являются одновременно минимальными.

Пусть $L_\theta(x)$ — функция правдоподобия и $L_\theta(x) > 0$, $x \in X$, $\theta \in \Theta$.

Положим

$$C(x) = \left\{z \in X: \frac{L_\theta(z)}{L_\theta(x)} = \varphi(z, x) \quad \forall \theta \in \Theta\right\}.$$

Теорема 2. *Если для $T \in \mathcal{T}$*

$$T(x_1) = T(x_2) \quad \text{для} \quad x_1, x_2 \in C(x), \quad x \in X, \tag{2.6}$$

то T — минимальная достаточная статистика.

Пример 5. Пусть $X = \{0, 1\}$, $(X^n, \mathfrak{X}^n, \{P_\theta, \theta \in (0, 1)\})$ — статистическая структура, P_θ — биномиальное распределение с неизвестным параметром. Тривиальная статистика $T_1(x) = (x_1, x_2, \ldots, x_n)$, совпадающая с выборкой $y_n = (x_1, \ldots, x_n)$, является достаточной статистикой.

Достаточными являются также, например, статистики

$$\begin{aligned} T_2(x) &= (x_1 + x_2, x_3, \ldots, x_n), \\ T_3(x) &= (x_1 + x_2 + x_3, x_4, \ldots, x_n), \\ T_4(x) &= (x_1 + x_2 + x_3 + x_4, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ T_0(x) &= x_1 + x_2 + \ldots + x_n. \end{aligned}$$

Среди алгебр $\mathfrak{X}_{T_k} = T_k^{-1}(\mathfrak{X}^n)$ статистике T_1 отвечает максимальная алгебра $\mathfrak{X}^n$, статистике T_0 — минимальная, порожденная векторами $(x_1, x_2, \ldots, x_n)$, $x_i = 0, 1$, с фиксированной суммой компонент.

Статистика T_0 — минимальная согласно (2.5).

С другой стороны, отношение

$$\frac{L_\theta(y)}{L_\theta(x)} = \left(\frac{\theta}{1-\theta}\right)^{\sum y_i - \sum x_i}$$

не зависит от θ тогда и только тогда, когда $\sum\limits_{i=1}^{n} y_i = \sum\limits_{i=1}^{n} x_i = k$ $(k = 0, \ldots, n)$. Следовательно,

$$C(x) = \left\{y \in X^n: \frac{L_\theta(y)}{L_\theta(x)} = \varphi(x, y), \ \theta \in \Theta\right\}$$

для всякого $x = (x_1, \ldots, x_n)$ совпадает с $\left\{y \in X^n: \sum y_i = \sum x_i\right\}$. Статистика, удовлетворяющая (2.6), совпадает с T_0.

Аналогично проверяется, что статистика примера 4 минимальна.

20.2.4. Полные статистики. Статистика T на статистической структуре $(X, \mathfrak{X}, \{P_\theta, \theta \in \Theta\})$ называется (*ограниченно*) *полной*, если для всякой (ограниченной) интегрируемой вещественной статистики $h(T)$ из того, что

$$\int\limits_X h(T(x))\, P_\theta(dx) = 0$$

для $\theta \in \Theta$, следует, что

$$h(T(x)) = 0 \ (\text{mod}\, P_\theta), \quad \theta \in \Theta.$$

Теорема 3. *Полная достаточная статистика является минимальной достаточной статистикой.*

Обратное утверждение, вообще говоря, неверно.

Для экспоненциальной статистической структуры статистика T из примера 4 п. 20.2.2 является не только минимальной достаточной, но и полной.

Пусть T — достаточная статистика со значениями в $(Y, \mathfrak{U})$ и оценка θ^* задает отображение Y в $\Theta \subset \mathbb{R}^k$, измеримое относительно σ-алгебры $\mathfrak{U}_T$, порожденной статистикой T.

Теорема 4. *Статистика T является полной тогда и только тогда, когда найдется измеримая функция $b(\theta)$ со значениями в Θ такая, что оценка θ^* единственна в классе всех $\mathfrak{U}_T$-измеримых интегрируемых оценок $\widehat{\theta}$, для которых*

$$\int \hat{\theta}(T(y))\, P_\theta(dx) = \theta + b(\theta).$$

Свойство полноты выделяет в классе оценок, имеющих конечную дисперсию, оценки, обладающие минимальной дисперсией (см. п. 21.1.4).

20.3. Основные задачи математической статистики

20.3.1. Проверка простой гипотезы. Неизвестная функция распределения F принадлежит к некоторому классу распределений $\mathfrak{F}$. Из априорных соображений можно сделать вывод, что $F = F_0 \in \mathfrak{F}$. Нужно на основании сделанных наблюдений подтвердить или опровергнуть эту гипотезу. Например, $\mathfrak{F}$ — множество нормальных распределений, F_0 — нормальное распределение со средним 0 и дисперсией 1.

20.3.2. Проверка сложной гипотезы. Неизвестная функция распределения F принадлежит $\mathfrak{F}$, $\mathfrak{F}_0 \subset \mathfrak{F}$. Нужно проверить гипотезу: $F \in \mathfrak{F}_0$. Например, проверка гипотезы о том, что величина, имеющая нормальное распределение, имеет среднее 0. В этом случае $\mathfrak{F}$ совпадает с классом всех нормальных распределений, а $\mathfrak{F}_0$ — с классом нормальных распределений, имеющих среднее 0.

Задачи проверки простой и сложной гипотез сводятся к задаче проверки статистической гипотезы, для решения которой используются критерии согласия. Критерий согласия определяется заданием критической области G в выборочном пространстве. Если выборка попадает в критическую область, то гипотеза отвергается. Качество критерия определяется вероятностью отвергнуть истинную гипотезу. Чем меньше эта вероятность, тем критерий лучше. С другой стороны, критерий характеризуется вероятностями не отвергнуть (принять) ложную гипотезу (эта вероятность зависит, естественно, от того, каким является истинное распределение). Эти вероятности также желательно сделать как можно меньшими.

20.3.3. Оценка параметра распределения. Предполагается, что неизвестная функция распределения принадлежит некоторому семейству распределений $F(\theta, x)$, зависящему от некоторого параметра $\theta \in \Theta$, где Θ — множество на прямой либо в конечномерном евклидовом пространстве. Это значит, что распределение зависит от одного или нескольких вещественных параметров. Так, например, семейство нормальных распределений на прямой зависит от двух вещественных параметров — среднего значения и дисперсии. Нужно по наблюдениям оценить параметр (или несколько вещественных параметров). Для построения оценок используются статистики — функции от выборочных значений. Примерами статистик являются:

выборочное среднее

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k;$$

выборочная дисперсия

$$\bar{s}_n^2 = \frac{1}{n}\sum_{k=1}^{n} (x_k - \bar{x})^2$$

(здесь $x_1, \ldots, x_n$ — выборка объема n). В качестве оценки вещественного параметра θ используется некоторая статистика $\hat{\theta}_n(x_1, x_2, \ldots, x_n)$, которая рассматривается как приближенное значение неизвестного параметра. Качество оценки определяется распределением величины

$$\hat{\theta}_n(x_1, x_2, \ldots, x_n) - \theta$$

(это распределение, очевидно, зависит от значения искомого параметра). Оценка будет хорошей, если это распределение достаточно сосредоточено возле нуля. В практике обычно ограничиваются лишь двумя первыми моментами оценки. В этом случае оценка характеризуется смещением

$$\mathrm{M}_\theta\hat{\theta}_n(x_1, x_2, \ldots, x_n) - \theta$$

и дисперсией

$$\mathrm{M}_\theta[\hat{\theta}_n(x_1, x_2, \ldots, x_n) - \mathrm{M}_\theta\hat{\theta}_n(x_1, x_2, \ldots, x_n)]^2$$

(M_θ обозначает математическое ожидание в предположении, что истинное распределение совпадает с $F(\theta, x)$). Оценки, для которых смещение равно нулю, называются *несмещенными*. Для несмещенных оценок качество определяется величиной дисперсии: чем она меньше, тем оценка лучше.

20.3.4. Доверительные области для параметров. Условия задачи такие же, как и в предыдущем пункте. Однако вместо оценки параметра строится доверительная область для параметра, т. е. такая область S в Θ, для которой вероятность того, что S будет содержать истинное значение параметра, не меньше чем α (это число называется уровнем доверия, оно должно быть достаточно близким к 1).

Доверительная область строится по выборочным значениям. Она задается функцией, определенной на выборочном пространстве, значениями которой являются области в множестве Θ. В случае оценки одного вещественного параметра используются доверительные интервалы, задаваемые двумя статистиками, определяющими концы интервала. Пусть $S(x_1, x_2, \ldots, x_n) \subseteq \Theta$ — доверительная область. Качество доверительной области характеризуется уровнем доверия, а также формой и размерами области.

20.3.5. Общие задачи статистических решений. Как правило, определение функции распределения неизвестного параметра и принятие той или иной гипотезы являются составной частью некоторой более общей задачи, состоящей в принятии некоторого решения. Например, таким решением может быть принятие некоторой гипотезы. Более сложный пример: при выведении нового сорта культуры нужно на каждой стадии эксперимента принимать определённое решение, как производить отбор семян, а затем принять окончательное решение, что выведенный сорт удовлетворяет необходимым требованиям. Основанием для принятия того или иного решения служит наблюдаемый статистический материал (в приведенном примере — данные о тех или иных свойствах семян, полученных на экспериментальных участках). В каждой такой задаче имеется множество возможных решений D. Правило принятия решения задается функцией $d(x_1, x_2, \ldots, x_n)$ на выборном пространстве, принимающей значения из D, которая называется *решающей функцией*. Предполагается, что возможные распределения выборки принадлежат некоторому множеству распределений $\mathcal{P}$. Для оценки качества правила принятия решения используется некоторая функция потерь $W(d, P)$, определяющая тот убыток, который мы получим, приняв решение d, если истинное распределение выборки было $P \in \mathcal{P}$ (убыток может быть и отрицательным). Естественно искать такие правила, при которых средний убыток является минимальным.

20.3.6. Последовательный анализ. Среди правил приема решений особо важную роль играют последовательные правила. В этом случае удобно рассматривать бесконечномерное выборочное пространство, так как при решении используются выборки сколь угодно большого объема. Последовательное правило указывает, при каком n в зависимости от выборки $(x_1, x_2, \ldots, x_n)$ следует прекратить наблюдение и какое в этом случае принимается решение. Так, при последовательном различении двух гипотез величины $x_1, x_2, \ldots$ наблюдаются последовательно и при каждом $n = 1, 2, \ldots$ принимается одно из решений: d_1 — принята первая гипотеза, d_2 — принята вторая гипотеза, d_3 — необходимо произвести еще одно наблюдение (т. е. прибавить к выборке $x_1, x_2, \ldots, x_n$ наблюдение x_{n+1}).

Последовательное правило принятия решения может быть описано двумя последовательностями функций. Пусть $\varepsilon_n(x_1, x_2, \ldots, x_n) = 1$, если наблюдение прекращается на n-м шаге, и $\varepsilon_n(x_1, x_2, \ldots, x_n) = 0$ в противном случае; $d_n(x_1, x_2, \ldots, x_n)$ — решение, которое принимается на n-м шаге, это функция, определенная на подмножестве выборочного пространства, на котором $\varepsilon_n = 1$, и принимающая значения из D. Среди правил выбирается такое, которое минимизирует некоторый убыток, учитывающий и число необходимых по правилу наблюдений.

20.3.7. Статистика случайных процессов. Для случайных процессов решаются те же задачи, что и для независимых наблюдений: проверка гипотез и оценка параметров распределений. Особенностью статистических задач для случайных процессов является то, что обычно наблюдается единственная траектория случайного процесса и на основании этого наблюдения принимаются статистические решения. Таким образом, статистика случайных процессов — это статистика одного (независимого) наблюдения. Однако это наблюдения не одной, а бесконечного числа случайных величин (значений процесса в различные моменты времени), зависимость между которыми определяется структурой процесса. Отметим некоторые особенности статистики случайных процессов. Во-первых, параметры, от которых зависят распределения, сами часто являются бесконечномерными (так, семейство распределений гауссовских процессов зависит от двух функциональных параметров — среднего значения и корреляционной функции). Во-вторых, несмотря на то, что имеется лишь одно наблюдение, иногда можно достоверно выбрать одну из гипотез или абсолютно точно определить значение параметра. В классических задачах статистики для регулярных распределений такого эффекта нет.

Поскольку, вообще говоря, распределения случайных процессов не могут быть заданы эффективно, не всегда возможно конструктивно решать статистические задачи, используя все конечномерные распределения. В простейших случаях используются лишь моментные фукции до определенного порядка. Особенно широкое развитие получила статистика стационарных в широком смысле процессов (гл. 25).

20.4. Распределение выборки

20.4.1. Вариационный ряд. Исходным материалом для статистического анализа, полученным в результате простого случайного выбора из генеральной совокупности, определяемой случайной величиной ξ с функцией распределений $P(x)$, служит выборка конечного объема

$$x_1, x_2, \ldots, x_n, \tag{4.1}$$

т. е. последовательность независимых одинаково распределенных случайных величин x_k $(1 \leqslant k \leqslant n)$ с общей функцией распределения $P(x)$.

Упорядоченная по величине последовательность выборочных значений

$$x_1^{(n)} \leqslant x_2^{(n)} \leqslant \ldots \leqslant x_n^{(n)} \tag{4.2}$$

называется *вариационным рядом.* Равные между собой члены выборки нумеруются в произвольном порядке.

Члены вариационного ряда $x_m^{(n)}$ $(m = 1, 2, \ldots, n)$ называются *порядковыми* (*ранговыми*) *статистиками.* Число $\lambda_m = m/n$ называется *рангом* члена $x_m^{(n)}$.

Стастика $\nu_n(x)$, равная числу значений выборки, меньших x:

$$\left\{\nu_n(x) = m\right\} = \left\{x_m^{(n)} < x \leqslant x_{m+1}^{(n)}\right\}, \quad m = 0, 1, \ldots, n, \tag{4.3}$$

называется *эмпирической частотой.* Случайная величина $\nu_n(x)$ равна числу наступлений события $\{\xi < x\}$ в n независимых испытаниях, так что эмпирическая частота $\nu_n(x)$ имеет биномиальное распределение с параметром $p = \mathsf{P}\{\xi < x\} = P(x)$:

$$\mathsf{P}\left\{\nu_n(x) = m\right\} = C_n^m P^m(x)\,(1 - P(x))^{n-m}, \quad m = 0, 1, \ldots, n. \tag{4.4}$$

Распределение членов вариационного ряда (порядковых статистик) просто определяется по распределению эмпирической частоты:

$$\mathsf{P}\left\{x_m^{(n)} < x\right\} = \mathsf{P}\left\{\nu_n(x) \geqslant m\right\} = \sum_{k=m}^{n} C_n^k P^k(x)\,(1 - P(x))^{n-k}. \tag{4.5}$$

В частности, очень простой вид имеют распределения *крайних членов вариационного ряда* $x_1^{(n)}$ и $x_n^{(n)}$:

$$\begin{aligned} \mathsf{P}\left\{x_1^{(n)} < x\right\} &= 1 - (1 - P(x))^n, \\ \mathsf{P}\left\{x_n^{(n)} < x\right\} &= P^n(x). \end{aligned} \tag{4.6}$$

Распределение членов вариационного ряда можно представить в другой, более удобной для анализа форме:

$$\mathsf{P}\left\{x_m^{(n)} < x\right\} = \frac{n!}{(m-1)!\,(n-m)!} \int\limits_0^{P(x)} y^{m-1}(1-y)^{n-m}\,dy. \tag{4.7}$$

Распределение (4.7) принадлежит к типу бета-распределений.

Если исходное распределение генеральной совокупности $P(x)$ имеет плотность $p(x) = dP/dx$, то и распределение порядковых статистик имеет плотность в виде

$$\frac{d}{dx}\mathsf{P}\left\{x_m^{(n)} < x\right\} = \frac{n!}{(m-1)!\,(n-m)!} P^{m-1}(x)\,(1-P(x))^{n-m}\,p(x). \tag{4.8}$$

В задачах статистического контроля качества продукции часто используется статистика $R_n = x_n^{(n)} - x_1^{(n)}$, называемая *размахом* или *широтой выборки*. Распределение размаха имеет вид

$$\mathsf{P}\left\{x_n^{(n)} - x_1^{(n)} < t\right\} = n \int\limits_{-\infty}^{\infty} [P(x+t) - P(x)]^{n-1}\,dP(x). \tag{4.9}$$

Точное распределение порядковых статистик трудно использовать в статистическом анализе, так как они существенно зависят от исходного распределения. Естественно ожидать, что при неограниченном возрастании объема выборки n эта зависимость должна ослабевать. Приближенное представление порядковых статистик при больших n называют *асимптотическим представлением*. Преобразование членов вариационного ряда по формуле

$$z_m^{(n)} = nP\left(x_m^{(n)}\right) \tag{4.10}$$

дает плотность распределения $g_m^{(n)}(z)$ случайных величин $z_m^{(n)}$ в виде

$$g_m^{(n)}(z) = C_{n-1}^{m-1}\left(\frac{z}{n}\right)^{m-1}\left(1-\frac{z}{n}\right)^{n-m}. \tag{4.11}$$

При фиксированных z это распределение Бернулли (по m) при $n \to \infty$ аппроксимируется распределением Пуассона

$$g_m^{(n)}(z) \simeq \frac{z^{m-1}e^{-z}}{(m-1)!}, \quad z > 0, \tag{4.12}$$

которое по z является плотностью гамма-распределения с параметром $m-1$.

Аналогично преобразование крайних членов вариационного ряда по формуле $z_{n-m}^{(n)}=n\left(1-P\left(x_{n-m}^{(n)}\right)\right)$ дает плотность распределения $g_{n-m}^{(n)}(z)$ случайных величин $z_{n-m}^{(n)}$ в виде

$$g_{n-m}^{(n)}(z)=C_{n-1}^{m}\left(\frac{z}{n}\right)^{m}\left(1-\frac{z}{n}\right)^{n-m-1}. \tag{4.13}$$

При $n\to\infty$ и фиксированных m и z плотность распределения $g_{n-m}^{(n)}(z)$ аппроксимируется плотностью гамма-распределения с параметром m:

$$g_{n-m}^{(n)}(z)\simeq\frac{z^{m}e^{-z}}{m!},\quad z>0. \tag{4.14}$$

Распределения $z_{m}^{(n)}$ и $z_{n-m}^{(n)}$ можно использовать при фиксированных m для нахождения асимптотического представления для крайних членов вариационного ряда.

Примеры 1. Исходное распределение $P(x)$ равномерно на отрезке $[-a, a]$. Тогда для минимального $x_{1}^{(n)}$ и максимального $x_{n}^{(n)}$ членов вариационного ряда имеют место асимптотические представления

$$x_{1}^{(n)}\simeq -a+\frac{2a}{n}z,\quad x_{n}^{(n)}\simeq a-\frac{2a}{n}z, \tag{4.15}$$

где z — случайная величина с гамма-распределением и параметром $m=0$, т. е. с плотностью $g(z)=e^{-z}$, $z>0$.

2. Исходное распределение $P(x)$ имеет плотность вида $p(x)=e^{-|x|}/2$. Тогда минимальный и максимальный члены вариационного ряда имеют асимптотические представления вида

$$x_{1}^{(n)}\simeq v-\ln\frac{n}{2},\quad x_{n}^{(n)}\simeq -v+\ln\frac{n}{2}, \tag{4.16}$$

где v — случайная величина с плотностью распределения $g(v)=\exp\{v-e^{v}\}$.

В случае конечного ранга, когда $m=[nq]$, т. е. для целых m, удовлетворяющих неравенствам $0<m/n\leqslant q<(m+1)/n\leqslant 1$, преобразование членов вариационного ряда с конечным рангом $z_{q}^{(n)}=P\left(x_{[nq]}^{n}\right)$ дает плотность $g_{q}(z)$ распределения $z_{q}^{(n)}$ в виде

$$g_{q}(z)=C_{n-1}^{m-1}z^{m-1}(1-z)^{n-m},$$

которая в окрестности точки $z=q$ аппроксимируется плотностью нормального распределения с параметрами $\left(q,\sqrt{q(1-q)/n}\right)$, q — среднее значение, $\sqrt{q(1-q)/n}$ — среднеквадратическое отклонение. Это позволяет случайную величину $z_{q}^{(n)}$ представить асимптотически при $n\to\infty$ в виде

$$z_{q}^{(n)}\simeq q+u\sqrt{\frac{q(1-q)}{n}},$$

где u имеет нормальное распределение с параметрами (0, 1). При этом средние члены вариационного ряда $x_{[nq]}^{(n)}$ асимптотически представимы в виде

$$x_{[nq]}^{(n)} \simeq \varkappa_q + \frac{u}{p(\varkappa_q)} \sqrt{\frac{q(1-q)}{n}},$$

где $\varkappa_q$ — квантиль порядка q распределения $P(x)$, т. е. $P(\varkappa_q) = q$, $p(x)$ — плотность исходного распределения.

20.4.2. Эмпирическая функция распределения. Выше была введена эмпирическая частота $\nu_n(x)$, равная числу выборочных значений, меньших x, и имеющая биномиальное распределение с параметром $P(x)$.

Эмпирической функцией распределения называется функция $F_n(x)$, определяемая соотношением

$$\overline{P}_n(x) = \frac{\nu_n(x)}{n}. \tag{4.17}$$

Иначе, эмпирическая функция распределения определяется следующим соотношением:

$$\overline{P}_n(x) = \begin{cases} 0, & x \leqslant x_1^{(n)}, \\ m/n, & x_m^{(n)} < x \leqslant x_{m+1}^{(n)}, \\ 1, & x > x_n^{(n)}. \end{cases} \tag{4.18}$$

График эмпирической функции распределения представляет собой ступенчатую линию со скачками, кратными величине $1/n$ в точках, определяемых членами вариационного ряда $x_1^{(n)} \leqslant x_2^{(n)} \leqslant \ldots \leqslant x_n^{(n)}$. Эмпирическая функция распределения обладает всеми свойствами распределения вероятностей. Эмпирическую функцию распределения называют еще *распределением выборки.*

При фиксированном x $\mathsf{M}\,\overline{P}_n(x) = P(x)$. Следовательно, по закону больших чисел при $n \to \infty$ для каждого x эмпирическая функция распределения сходится по вероятности к исходному теоретическому распределению $P(x)$.

Теорема Гливенко. *Эмпирическая функция распределения $P_n(x)$ равномерно по x с вероятностью 1 сходится при $n \to \infty$ к теоретическому распределению $P(x)$*

$$\mathsf{P}\left\{\lim_{n\to\infty} \sup_{-\infty<x<+\infty} |F_n(x) - F(x)| = 0\right\} = 1.$$

В качестве возможной меры уклонения эмпирической функции распределения $\overline{P}_n(x)$ от теоретической $P(x)$ при фиксированном x можно воспользоваться тем, что разность $\overline{P}_n(x) - P(x)$ асимптотически нормальна с нулевым средним и дисперсией $\dfrac{P(x)(1-P(x))}{n}$. Однако такая мера уклонения неравномерна по x. Важную роль в математической статистике сыграло изучение статистики, введенной А. Н. Колмогоровым:

$$D_n = \sup_{-\infty<x<+\infty} |P_n(x) - P(x)|.$$

Теорема Колмогорова. *Если функция распределения $P(x)$ непрерывна, то*

$$\lim_{n\to\infty} \mathsf{P}\left\{\sqrt{n} \sup_{-\infty<x<+\infty} |\overline{P}_n(x) - P(x)| < z\right\} =$$

$$= K(z) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2z^2}, \quad z > 0. \quad (4.19)$$

Функция $K(z)$ табулирована. Статистика D_n используется в непараметрическом критерии согласия (см. п. 20.3.4) эмпирических результатов с гипотезой об исходном теоретическом распределении. Заметим, что распределение статистики D_n не зависит от вида исходного непрерывного распределения $P(x)$.

Практическое вычисление статистики D_n не представляет большого труда, так как максимальное уклонение эмпирической функции распределения от непрерывной теоретической функции распределения достигается в точках скачков $\overline{P}_n(x)$, так что

$$D_n = \max_{1\leq m\leq n} \left\{\left|\frac{m-1}{n} - P\left(x_m^{(n)}\right)\right|, \left|\frac{m}{n} - P\left(x_m^{(n)}\right)\right|\right\}, \quad (4.20)$$

или иначе

$$D_n = \max_{1\leq m\leq n} \left\{\left|\frac{2m-1}{2n} - P\left(x_m^{(n)}\right)\right| + \frac{1}{2n}\right\}. \quad (4.21)$$

В односторонних критериях согласия может быть использовано распределение статистик Смирнова:

$$D_n^+ = \sup_{-\infty<x<+\infty} \left[\overline{P}_n(x) - P(x)\right] \quad (4.22)$$

или

$$D_n^- = -\inf_{-\infty<x<+\infty} \left[\overline{P}_n(x) - P(x)\right]. \quad (4.22')$$

Эти статистики имеют одинаковые распределения:

$$\mathsf{P}\left\{D_n^+ \geqslant x\right\} = \mathsf{P}\left\{D_n^- \geqslant x\right\} =$$

$$= \sum_{k=0}^{[n(1-x)]} C_n^k x \left(x + \frac{k}{n}\right)^{k-1} \left(1 - x - \frac{k}{n}\right)^{n-k}, \quad 0 < x < 1. \quad (4.23)$$

Теорема Смирнова. *Если распределение $P(x)$ непрерывно, то*

$$\lim_{n\to\infty} \mathsf{P}\left\{\sqrt{n}\, D_n^+ < z\right\} = 1 - e^{-2z^2}, \quad z > 0. \quad (4.24)$$

Известно также совместное предельное распределение статистик D_n^+ и D_n^-.

Теорема. *Если распределение $P(x)$ непрерывно, то*

$$\lim_{n\to\infty} \mathsf{P}\left\{\sqrt{\frac{n}{2}}\, D_n^+ < z,\quad \sqrt{\frac{n}{2}}\, D_n^- < v\right\} =$$

$$= 1 + 2\sum_{k=1}^{\infty} \exp\left\{-2k^2(z+v)^2\right\} - \sum_{k=1}^{\infty}\left[\exp\left\{-2(kv+(k-1)z)^2\right\} + \right.$$
$$\left. + \exp\left\{-2((k-1)v+kz)^2\right\}\right].$$

Имеются также асимптотические разложения для распределений статистик D_n, D_n^+ и D_n^-.

20.4.3. Распределение выборочных характеристик. *Выборочными* (или *эмпирическими*) *характеристиками* называют характеристики эмпирической функции распределения.

Выборочные характеристики являются случайными величинами, функциями от выборочных значений, т. е. статистиками, представимыми в виде

$$\bar{t}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{\infty} t(x)\, d\bar{P}(x). \tag{4.25}$$

Так как эмпирическая функция распределения $\bar{P}(x)$ служит оценкой исходного распределения $P(x)$ (см. п. 20.4.5), то следует ожидать, что и выборочные характеристики могут служить оценками соответствующих характеристик исходного распределения. Этим объясняется важность изучения распределений выборочных статистик и их числовых характеристик.

Условимся в дальнейшем числовые характеристики выборочных статистик (т. е. эмпирической функции $\bar{P}(x)$) обозначать той же буквой, что и соответствующие числовые характеристики генеральной совокупности (т. е. исходного распределения $P(x)$), только с чертой сверху.

20.4.4. Выборочные (статистические) моменты. *Среднее значение выборки* определяется соотношением

$$\bar{x} = \int_{-\infty}^{\infty} x\, d\bar{P}_n(x) = \frac{1}{n}\sum_{k=1}^{n} x_k. \tag{4.26}$$

Числовые характеристики среднего значения легко вычисляются с учетом того, что $\bar{x}$ есть сумма независимых одинаково распределенных случайных величин. Например,

$$\mathsf{M}\bar{x} = m, \quad \mathsf{D}\bar{x} = \frac{\sigma^2}{n}, \tag{4.27}$$

где m — среднее значение, σ^2 — дисперсия генеральной совокупности

$$m = \int_{-\infty}^{\infty} x\, dP(x), \quad \sigma^2 = \int_{-\infty}^{\infty} (x-m)^2\, dP(x).$$

Из формул (4.27) немедленно следует, что выборочное среднее $\bar{x}$ сходится по вероятности к среднему значению m генеральной совокупности при $n \to \infty$. Более того, величина отклонения выборочного среднего от его математического ожидания $\sqrt{n}(\bar{x} - m)$ *асимптотически нормальна* с параметрами $(0, \sigma^2)$.

Выборочная (*статистическая*) *дисперсия* определяется (обычно) соотношением

$$\bar{s}^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2. \tag{4.28}$$

Основные числовые характеристики выборочной дисперсии имеют вид

$$\mathrm{M}\bar{s}^2 = \left(1 - \frac{1}{n}\right)\sigma^2,$$
$$\mathrm{D}\bar{s}^2 = \frac{m_4 - m_2^2}{n} - \frac{2\left(m_4 - 2m_2^2\right)}{n^2} + \frac{m_4 - 3m_2^2}{n^3}. \tag{4.29}$$

Здесь m_2 и m_4 — соответственно второй и четвертый центральные моменты генеральной совокупности, т. е.

$$m_2 = \sigma^2, \quad m_4 = \int_{-\infty}^{\infty} (x - m)^4 \, dP(x).$$

Из первой формулы (4.29) следует, что несмещенной оценкой дисперсии является статистика

$$\frac{n}{n-1} \bar{s}^2 = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \bar{x})^2.$$

Вместе с тем величина отклонения выборочной дисперсии от дисперсии генеральной совокупности $\sqrt{n}(\bar{s} - \sigma^2)$ *асимптотически нормальна* с параметрами $\left(0, m_4 - m_2^2\right)$.

Старшие *выборочные* (*центральные*) *моменты* определяются соотношением

$$\bar{m}_r = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^r, \quad r \geqslant 2. \tag{4.30}$$

Вычисление числовых характеристик старших выборочных моментов не представляет труда. Нужно воспользоваться очевидным соотношением $\mathrm{M}x_k^r = a_r = \int_{-\infty}^{\infty} x^r \, dP(x)$ и независимостью случайных величин x_k. Точные выражения числовых характеристик выборочных моментов $\bar{m}_r$ при $r \geqslant 3$ громоздки, однако при больших n асимптотические выражения значительно упрощаются:

$$\mathsf{M}\bar{m}_r = m_r + O\left(\frac{1}{n}\right), \tag{4.31}$$

$$\mathsf{D}\bar{m}_r = \frac{1}{n}\left[m_{2r} - 2rm_{r-1}m_{r+1} - m_r^2 + r^2 m_2 m_{r-1}^2\right] + O\left(\frac{1}{n^2}\right). \tag{4.32}$$

Как и в случае первых двух выборочных моментов, старшие выборочные моменты $\bar{m}_r$ асимптотически нормальны (при $n \to \infty$) со средним и дисперсией, определяемыми главными членами формул (4.31) и (4.32).

20.4.5. Функции от выборочных моментов. При определении числовых характеристик функции от выборочных моментов полезной является следующая теорема.

Теорема. *Пусть задана функция $H(\bar{m}_r, \bar{m}_s)$ от выборочных моментов $\bar{m}_r$ и $\bar{m}_s$, не зависящая явно от n и удовлетворяющая условиям:*

1) *функция $H(u, v)$ дважды непрерывно дифференцируема в окрестности точек m_r, m_s;*

2) *при всех значениях x_k ($k = 1, \ldots, n$) функция $H(\bar{m}_r, \bar{m}_s) = H(x_1, x_2, \ldots, x_k, n)$ удовлетворяет оценке $|H| < Cn^p$, где C и p — неотрицательные постоянные.*

Тогда среднее значение и дисперсия случайной величины $H(\bar{m}_r, \bar{m}_s)$ представимы асимптотическими формулами

$$\mathsf{M}H = H(m_r, m_s) + O\left(\frac{1}{n}\right), \tag{4.33}$$

$$\mathsf{D}H = \mathsf{D}\bar{m}_r \frac{\partial H}{\partial m_r}(m_r, m_s) + 2\mathsf{M}\left[(\bar{m}_r - m_r)(\bar{m}_s - m_s)\right] \times$$
$$\times \frac{\partial H}{\partial m_r}(m_r, m_s)\frac{\partial H}{\partial m_s}(m_r, m_s) + \mathsf{D}\bar{m}_s \frac{\partial H}{\partial m_s}(m_r, m_s) + O\left(\frac{1}{n^{3/2}}\right). \tag{4.34}$$

Приведенная теорема имеет место также и для функций любого числа аргументов в случае многомерных выборок.

При выполнении только первого условия теоремы статистика $H(\bar{m}_r, \bar{m}_s)$ *асимптотически нормальна* (при $n \to \infty$) со средним и дисперсией, задаваемыми главными членами формул (4.33) и (4.34). Заметим, что главный член формулы (4.34) может оказаться равным нулю. В этом случае $\sqrt{n}(H - \mathsf{M}H)$ асимптотически нормально с нулевой дисперсией, т. е. сходится по вероятности к нулю. Не исключена возможность, что при некотором $p > \frac{1}{2}$ $n^p(H - \mathsf{M}H)$ имеет нетривиальное предельное распределение, не обязательно нормальное.

20.4.6. Точные распределения выборочных характеристик. Явный вид распределения выборочных характеристик, представляющих, как правило, функционалы от эмпирической функции распределения, может быть получен в принципе для любых распределений. Наиболее просты и обозримы соответствующие формулы в случае нормальной генеральной совокупности.

Если исходная функция распределения $P(x)$ выборочных значений x_k $(k=1, \ldots, n)$ нормальна с параметрами (m, σ^2), то *выборочное среднее* $\bar{x}=\frac{1}{n}\sum_{k=1}^{n} x_k$ *и выборочная дисперсия* $\bar{s}^2$ *независимы*, причем $\bar{x}$ нормально распределено с параметрами $\left(m, \frac{\sigma^2}{n}\right)$, а $\frac{n}{\sigma^2}\bar{s}^2$ имеет χ^2-распределение с $n-1$ степенями свободы, или, иначе, $\frac{n}{n-1}\bar{s}^2$ имеет такое же распределение, как и среднее арифметическое $n-1$ квадратов независимых нормальных величин с параметрами $(0, \sigma^2)$. Кроме того, отношение

$$t=\sqrt{n-1}\frac{\bar{x}-m}{\bar{s}} \tag{4.35}$$

имеет распределение Стьюдента (см. гл. 6) с $n-1$ степенями свободы.

20.5. Процедуры проверки гипотез

20.5.1. Общая схема построения статистических (нерандомизированных) критериев. Пусть $(X, \mathfrak{X}, \mathcal{P})$ — статистическая структура, связанная со случайным элементом ξ с неизвестным распределением $P_0 \in \mathcal{P}$. Любое предположение, выделяющее подсемейство $\mathcal{P}_0 \subset \mathcal{P}$, которому может принадлежать P_0: $P_0 \in \mathcal{P}_0$ (в частности, $\mathcal{P}_0=\{P_0\}$), называют в математической статистике *гипотезой*.

Так, если $(X^n, \mathfrak{X}^n, \mathcal{P}^n)$ — статистическая структура модели простой выборки, связанной со случайным элементом ξ с распределением $P_\theta \in \mathcal{P}=\{P_\theta\colon \theta \in \Theta\}$, статистическая гипотеза утверждает: неизвестный параметр θ исходного распределения вероятностей P_θ принадлежит заданному подмножеству $H \subset \Theta$ множества возможных значений параметра θ. Дополнительное подмножество $K=\Theta\setminus H$ называется *альтернативой к гипотезе* H.

Гипотеза H называется *простой*, если множество H состоит из одного-единственного значения параметра θ; в противном случае H — сложная гипотеза.

Практическое применение математической статистики состоит в проверке фактического соответствия реальных результатов экспериментов предполагаемой гипотезе. С этой целью строится процедура *проверки гипотезы* (*критерий согласия*), позволяющая по результатам наблюдений принимать или отвергать данную гипотезу.

Выборочное пространство X разбивается на два непересекающихся подмножества: X_0 и X_1. *Правило проверки гипотезы* формулируется так. Если результаты наблюдений $x \in X_0$, то считается, что данная гипотеза H подтверждается эмпирическими данными, т. е. гипотеза H принимается. Если же выборочное значение $x \in X_1$, то утверждается, что данная гипотеза H не согласуется с результатами наблюдений, т. е. H отвергается. Множество X_0 называется *областью принятия гипотезы*, множество X_1 — *критической областью*. Ради краткости множество X_1 иногда называют также *критерием* *гипотезы* H. Поскольку событие $x \in X_1$ является случайным, то данная гипотеза принимается или отвергается в результате наблюдения случайного события, имеющего определенную вероятность $P_\theta\{x \in X_1\}$ при каждом заданном θ. Применение процедуры проверки гипотезы сопряжено с ошибками двух родов: *отвергнуть гипотезу, когда она верна* (*ошибка первого рода*); *принять гипотезу, когда она неверна* (*ошибка второго рода*).

При построении процедур проверки гипотез желательно добиваться минимизации значений ошибок обоих родов. В большинстве практически важных ситуаций невозможно построение критериев согласия со сколь угодно малыми ошибками первого и второго рода. Правила проверки гипотез имеют статистический смысл, т. е. при многократном применении определенного правила проценты числа неверных решений выражаются вероятностями ошибок первого и второго рода.

Если выборочные данные $x \in X_1$, тогда с вероятностью ошибки первого рода наблюдается случайное событие, которое противоречит гипотезе. Если вероятность такого случайного события мала, значит, наблюдается практически невозможное событие. В этом случае данная гипотеза должна быть отвергнута с практической достоверностью.

Когда экспериментальные данные согласуются с предполагаемой гипотезой, это еще не означает, что невозможно согласование этих же данных с другой гипотезой. При применении статистических критериев на основании наблюдений невозможно доказательство той или иной гипотезы. Можно лишь утверждать, что результаты наблюдений не противоречат принятой гипотезе.

Таким образом, выводы, принимаемые на основании статистических данных, формулируются в следующем виде: экспериментальные данные согласуются с данной гипотезой (противоречат ей).

20.5.2. Функция мощности критерия. Вероятность $\beta(\theta) = P_\theta(X_1)$, рассматриваемая как функция параметра $\theta \in \Theta$, называется *функцией мощности критерия*.

Допустимое максимальное значение α ошибки первого рода при данном критерии называется *уровнем значимости критерия*:

$$\sup_{\theta \in H} P_\theta(X_1) \leqslant \alpha.$$

Если для критерия выполняется условие $P_\theta(X_1) = \alpha$ для $\theta \in H$, то критическая область X_1 называется *подобной* выборочному пространству.

Обычно уровень значимости α выбирается из практических соображений в зависимости от отношения к гипотезе. Если предполагаемая гипотеза весьма правдоподобна, то уровень значимости выбирается довольно малым. Тогда гипотеза будет отвергаться с малой вероятностью. При выборе уровня значимости следует также учитывать поведение функции мощности критерия $\beta(\theta) = P_\theta(X_1)$ при альтернативных значениях параметра $\theta \in K$. Желательным свойством критерия является его *несмещенность*, что характеризуется условиями

$$P_\theta(X_1) \leqslant \alpha, \quad \theta \in H; \quad P_\theta(X_1) \geqslant \alpha, \quad \theta \in K.$$

Может оказаться, что при данном уровне значимости мощность критерия $\beta(\theta)$ слишком мала при $\theta \in K$, значит, велика ошибка второго рода критерия неверной гипотезы.

Естественно считать *оптимальным критерием* такой, у которого при заданном уровне значимости достигается максимальное значение функции мощности критерия (задача Неймана — Пирсона).

Как правило, критерий, оптимальный при фиксированном значении параметра $\theta \in K$, зависит от параметра $\theta \in K$.

Критерий X_1^* называется *равномерно наиболее мощным* (р. н. м.), если для любого другого критерия X_1 выполняются условия:

$$P_\theta(X_1^*) \leqslant P_\theta(X_1), \quad \theta \in H; \quad P_\theta(X_1^*) \geqslant P_\theta(X_1), \ \theta \in K.$$

Рандомизированный критерий определяется *критической функцией* $\varphi(x)$ на выборочном пространстве X, значение которой есть вероятность отклонения гипотезы H при данном значении x результатов наблюдений. Мощность рандомизированного критерия с критической функцией $\varphi(x)$ определяется соотношением

$$\beta(\theta) = \mathrm{M}_\theta \varphi(x) = \int_X \varphi(x) P_\theta(dx).$$

В частности, когда $\varphi(x) = 1$ для $x \in X_1$ и $\varphi(x) = 0$ для $x \in X \setminus X_1$, т. е. когда критическая функция является характеристической функцией множества X_1, критерий, определяемый функцией φ, становится нерандомизированным с критической областью X_1.

В простом случайном выборе статистические данные представляют собой результаты наблюдений значений случайной величины в последовательности независимых испытаний. В этом случае выборочное пространство X является n-мерным евклидовым пространством: $x = (x_1, x_2, \ldots, x_n)$, где x_k $(k = 1, \ldots, n)$ — независимые одинаково распределенные случайные величины. Число n элементов выборочной последовательности называют *объемом выборки*.

Критерий согласия с критической областью X_1 называется *состоятельным*, если имеет место соотношение

$$\lim_{n \to \infty} P_\theta(X_1) = 1, \quad \theta \in K.$$

Состоятельность критерия означает, что ошибка второго рода — вероятность принятия неверной гипотезы — стремится к нулю при $n \to \infty$.

Для сравнения различных критериев между собой используются *меры асимптотической эффективности критериев*, которые основаны на изучении скорости сходимости функции мощности в окрестности параметра $\theta \in H$.

20.5.3. Критерии проверки статистических гипотез. Для проверки *простой гипотезы* H_0, которой соответствует распределение P_0, против *простой альтернативы* K, которой соответствует распределение P_1, критическая функция $\varphi(x)$ оптимального *критерия Неймана — Пирсона* при заданном уровне значимости α определяется условиями:

$$\int_X \varphi(x) p_0(x)\, dx = \alpha,$$

$$\varphi(x) = \begin{cases} 1, & p_1(x) \geqslant C_\alpha p_0(x), \\ 0, & p_1(x) < C_\alpha p_0(x), \end{cases}$$

где $p_0(x)$ и $p_1(x)$ — плотности распределений P_0 и P_1 по доминирующей мере μ. Критерий Неймана — Пирсона является наиболее мощным критерием уровня α.

Практическое применение критерия Неймана — Пирсона состоит в проверке неравенства $p_1(x) \geqslant C_\alpha p_0(x)$ для данных результатов наблюдений x. В случае выполнения этого неравенства гипотеза H_0 отклоняется; в противном случае гипотеза H_0 принимается.

Константа C_α в критерии Неймана — Пирсона определяется случайной величиной $T(\xi) = p_1(\xi)/p_0(\xi)$ при гипотезе H (ξ имеет распределение P_0) следующим образом:

$$P_0(T(\xi) \geqslant C_\alpha) = \alpha.$$

Если вероятности $P_\theta(T(\xi) \geqslant C_\alpha)$ не зависят от альтернативных значений параметра θ, то критерий Неймана — Пирсона проверки простой гипотезы H_θ является равномерно наиболее мощным по отношению ко всем таким альтернативам.

Критерии, основанные на использовании распределения статистик $T(\xi) = p_1(\xi)/p_0(\xi)$, называются *критериями отношения правдоподобия*; они обладают многими полезными свойствами.

Примеры. 1. Пусть $p(x;\ a,\ \sigma)$ — нормальные плотности распределения со средними значениями a и дисперсией σ^2. Рассмотрим простую гипотезу H_0: $a = a_0 \geqslant 0$ при известном значении параметра σ против класса альтернатив K: $a > a_0$ с тем же значением параметра σ. В простом случайном выборе $x = (x_1, x_2, \ldots, x_n)$ — n-мерный вектор независимых одинаково распределенных случайных величин $x_1, x_2, \ldots, x_n$.

В этом случае критерий Неймана — Пирсона определяется неравенством

$$\exp\left\{-\sum_{k=1}^{n} \frac{(x_k - a)^2}{2\sigma^2}\right\} \geqslant l_\alpha \exp\left\{-\sum_{k=1}^{n} \frac{(x_k - a_0)^2}{2\sigma^2}\right\}.$$

После логарифмирования это неравенство приводится к следующему виду:

$$(a - a_0) \sum_{k=1}^{n} x_k \geqslant l_\alpha.$$

Для класса альтернатив $a > a_0$ критическая область задается неравенством

$$\sum_{k=1}^{n} x_k \geqslant C_\alpha,$$

где константа C_α определяется из условия

$$P_{a_0}\left(\sum_{k=1}^{n} x_k \geqslant C_\alpha\right) = \alpha.$$

Критерий $\sum_{k=1}^{n} x_k \geqslant C_\alpha$ простой гипотезы H: $a = a_0$ является равномерно наиболее мощным для класса альтернатив $a > a_0$. Аналогично для класса альтернатив $a < a_0$ равномерно наиболее мощный критерий простой гипотезы H_0; $a = a_0$ определяется неравенством $\sum_{k=1}^{n} x_k \leqslant C_\alpha$.

Если множество альтернативных значений параметра a содержит точки как слева, таки справа от a_0, то р. н. м. критерия не существует.

2. Для простой гипотезы H_0: $a = a_0$ при неизвестном значении параметра σ против альтернативы K: $a > a_0$ критерий Неймана — Пирсона определяется критической областью

$$\frac{\bar{x} - a_0}{\bar{s}} \geqslant C_\alpha; \quad \bar{x} = \frac{1}{n} \sum_{k=1}^{n} x_k, \quad \bar{s}^2 = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \bar{x})^2.$$

3. Для простой гипотезы H_0: $\sigma = \sigma_0$ против альтернативы K: $\sigma < \sigma_0$ р. н. м. определяется соотношением $\bar{s}^2 \geqslant C_\alpha$.

20.5.4. Критерии отношения правдоподобия. Критерии отношения правдоподобия строятся с использованием свойств функции отношения правдоподобия $L(x) = p_\theta(x)/p_{\theta_0}(x)$.

Естественным представляется выбор критической области таким образом, чтобы при $x \in X_1$ отношение правдоподобия принимало возможно бо́льшие значения при всех альтернативных значениях параметра $\theta \in K$.

Критерий отношения правдоподобия простой гипотезы H_0: $\theta = \theta_0$ против сложной альтернативы $\theta \in K$ определяется статистикой

$$l(x) = \frac{1}{p_{\theta_0}(x)} \sup_{\theta \in K} p_\theta(x).$$

Критическая область имеет вид $X_1 = \{x: l(x) \geqslant C_\alpha\}$, где константа C_α определяется условием $p_{\theta_0}(X_1) = \alpha$. Например, для выборки $x = (x_1, x_2, \ldots, x_n)$ из нормальной совокупности $P(x; a, \sigma)$ с неизвестными параметрами a и σ критерий отношения правдоподобия для гипотезы H_0: $a = a_0$ определяется статистикой

$$l(x) = \left(1 + \frac{t^2(x)}{n-1}\right)^{-n/2},$$

где $t(x) = \sqrt{n}\,(\bar{x} - a_0)/\bar{s}$ имеет распределение Стьюдента с $n-1$ степенями свободы. $\left(\text{Здесь } \bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k,\ \bar{s}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2.\right)$

Критерий отношения правдоподобия сложной гипотезы $\theta \in H$ против сложной альтернативы $\theta \in K$ определяется статистикой

$$l(x) = \sup_{\theta \in K} p_\theta(x) \Big/ \sup_{\theta_0 \in H} p_{\theta_0}(x).$$

20.5.5. Критерий χ^2. Для проверки простой гипотезы H_0, которой соответствует дискретное распределение $(p_1, p_2, \ldots, p_m)$, $\sum_{k=1}^{m} p_k = 1$, используется статистика

$$\chi^2 = \sum_{k=1}^{m} \frac{(\nu_k - np_k)^2}{np_k},$$

где $\nu_1, \nu_2, \ldots, \nu_n$ — частоты результатов наблюдений в выборке объема $n = \sum_{k=1}^{m} \nu_k$. При $n \to \infty$ распределение статистики χ^2 стремится к χ^2-распределению с $m-1$ степенями свободы и плотностью вероятности

$$K_{m-1}(x) = \left(2^{(m-1)/2}\,\Gamma\left(\frac{m-1}{2}\right)\right)^{-1} x^{(m-2)/2} e^{-x/2}, \quad x > 0.$$

Предельное распределение не зависит от вида исходного дискретного распределения. Критерий χ^2 определяется следующим образом. Пусть χ_α^2 таково, что $\alpha = \mathsf{P}\left\{\chi^2 > \chi_\alpha^2\right\}$. Тогда критическая область критерия χ^2 определяется неравенством для статистики χ^2: $\chi^2 > \chi_\alpha^2$.

Критерий χ^2 используется также для проверки простой гипотезы о виде исходного распределения при группировке значений наблюдаемой случайной величины ξ. В этом случае $p_k = \mathsf{P}\left\{\xi \in X_k\right\}$, где X_k — группы, на которые разбито множество возможных значений случайной величины ξ. На практике критерий χ^2 оказывается достаточно эффективным, когда все ожидаемые частоты $np_k \geqslant 10$.

Примеры. 1. В последовательности независимых испытаний наблюдается случайное событие, вероятность которого p $(0 < p < 1)$ неизвестна. Пусть ν_n — частота наблюдений события в выборке объема n. Тогда статистика $\chi^2 = (\nu - np)^2/(np(1-p))$ при больших n распределена приближенно с плотностью $K_1(x) = \left(\sqrt{2\pi x}\right)^{-1} e^{-x/2}$ $(x > 0)$.

Имеющиеся таблицы квантилей χ^2-распределения дают возможность проверить гипотезу о том, что в данной серии наблюдений вероятность случайного события равна заданному числу p.

Критерий χ^2 применим также в случае, когда имеется несколько независимых серий наблюдений.

2. Пусть $\nu_1, \nu_2, \ldots, \nu_m$ — частоты наблюдений в выборках объема $n_1, n_2, \ldots, n_m$ случайного события, вероятность которого предполагается равной p.

Для проверки предполагаемой гипотезы можно воспользоваться статистикой $\chi^2 = \sum_{k=1}^{m} \frac{(\nu_k - n_k p)^2}{n_k\, p\,(1-p)}$, распределение которой при больших $n = \sum_{k=1}^{m} n_k$ близко к χ^2-распределению с m степенями свободы.

Если исходное распределение зависит от неизвестных параметров $\theta_1, \theta_2, \ldots, \theta_r$, то в качестве критерия согласия используется статистика

$$\chi^2 = \sum_{k=1}^{m} \frac{\left[\nu_k - np_k\left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r\right)\right]^2}{np_k\left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r\right)},$$

в которой $\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_r$ — оценки неизвестных параметров, построенные по результатам наблюдений. При определенных условиях статистика $\widehat{\chi}^2$ в пределе при $n \to \infty$ имеет χ^2-распределение с $m - r - 1$ степенями свободы.

3. В условиях примера 2 воспользуемся оценкой неизвестной вероятности $\hat{p} = \frac{1}{n}\sum_{k=1}^{m} \nu_k$. Статистика $\chi^2 = \sum_{k=1}^{m} \frac{(\nu_k - n_k\hat{p})^2}{n_k\hat{p}(1-\hat{p})}$ (имеющая в пределе χ^2-распределение с $m - 1$ степенями свободы) используется для проверки гипотезы об однородности выборок: предположения, что во всех m сериях независимых наблюдений вероятность наблюдаемого события оставалась неизменной.

20.5.6. Непараметрические критерии. Одной из основных непараметрических статистических задач является задача проверки согласования выборочных данных с гипотезой о том, что исходная функция распределения $F(x)$ является заданной. Если исходная функция распределения $F(x)$ непрерывна, то в качестве непараметрического критерия используется статистика Колмогорова

$$\sqrt{n}\, D_n = \sqrt{n} \sup_{-\infty < x < +\infty} |F_n(x) - F(x)|,$$

распределение которой не зависит от вида $F(x)$ и для которой известно предельное распределение (см. § 20.4).

Критическая область критерия при данном уровне значимости определяется неравенством

$$\sqrt{n}\, D_n \geqslant d_\alpha,$$

где d_α — квантиль предельного распределения Колмогорова, $1 - K(d_\alpha) = \alpha$.

Задача проверки однородности статистических данных ставится следующим образом. На основании двух серий независимых наблюдений $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$ проверить гипотезу о том, что результаты наблюдений в обеих сериях получены в результате испытаний над случайными величинами с одинаковой функцией распределения $F(x)$. Если исходная функция распределения $F(x)$ непрерывна, то в качестве критерия однородности выборочных данных используется статистика Смирнова, распределение которой не зависит от вида функции $F(x)$:

$$\sqrt{\frac{nm}{n+m}}\, D_{nm} = \sqrt{\frac{nm}{n+m}} \max_{-\infty < x < +\infty} |F_n(x) - F_m(x)|,$$

где $F_n(x)$ и $F_m(x)$ — эмпирические функции распределения первой и второй серий наблюдений соответственно. Критическая область критерия однородности определяется неравенством

$$\sqrt{\frac{nm}{n+m}}\, D_{nm} \geqslant d_\alpha,$$

в котором d_α — квантиль распределения статистики Смирнова при малых n и m, а при больших n и m d_α — квантиль предельного распределения статистики Смирнова: $1-K(d_\alpha)=\alpha$. В частном случае, когда объемы выборок равны ($n=m$), точное распределение статистики Смирнова имеет довольно простое аналитическое выражение.

Порядковые непараметрические критерии строятся по статистикам вариационного ряда, которые не зависят от конкретных значений членов вариационного ряда.

Критерий серий Вальда — Вольфовица основан на статистике U_{nm} — числа серий наблюденных значений первой и второй выборок в общем вариационном ряду $z_1 \leqslant z_2 \leqslant \ldots \leqslant z_{n+m}$, где каждое z_k есть либо x_{ik}, либо y_{ik}.

Распределение статистики U_{nm} не зависит от вида $F(x)$ исходного распределения:

$$U_{nm} = \sum_{k=1}^{n+m} v_k,$$

где случайные величины v_k независимы и принимают значения 1 или 0 с вероятностями, равными 1/2. При больших n и m распределение статистики U_{nm} асимптотически нормально со средним $nm/2$ и дисперсией $\frac{nm}{2}(n+m+1)$. При фиксированном m и $n \to \infty$ статистика U_{nm}/n в пределе распределена как $U_m = \sum_{i=1}^{m} W_i$, где W_i независимы и равномерно распределены на интервале (0, 1).

Имеются и другие непараметрические критерии.

20.5.7. Последовательный критерий отношения правдоподобия. На практике эксперименты осуществляются последовательно. На каждом этапе имеется возможность принимать решение о необходимости продолжения или прекращения испытаний. Объем выборки в последовательном статистическом анализе заранее не фиксируется и является случайной величиной. Пусть H_0 и H_1 — две альтернативные гипотезы о виде плотности распределения $p_0(x)$ или $p_1(x)$ наблюдаемой случайной величины. Последовательный критерий отношения правдоподобия (Вальда) строится по результатам независимых наблюдений $x_1, x_2, \ldots, x_n, \ldots$ следующим образом. Задаются две константы A и B, которые определяют разбиение выборочного пространства по отношению правдоподобия

$$\prod_{k=1}^{n} \frac{p_1(x_k)}{p_0(x_k)}$$

при каждом n на три области

$$X_1:\ \prod_{k=1}^{n}\frac{p_1(x_k)}{p_0(x_k)}\geqslant B \text{ — область принятия гипотезы } H_1;$$

$$X_0:\ \prod_{k=1}^{n}\frac{p_1(x_k)}{p_0(x_k)}\leqslant A \text{ — область принятия гипотезы } H_0;$$

$$X_{01}:\ A<\prod_{k=1}^{n}\frac{p_1(x_k)}{p_0(x_k)}< B \text{ — область продолжения испытаний.}$$

Качество последовательного критерия отношения правдоподобия определяется ошибками первого рода $\alpha = P_0(X_1)$ и второго рода $\beta = P_1(X_0)$, а также средним числом наблюдений $M_0\nu$ и $M_1\nu$; случайное число наблюдений ν определяется условиями

$$\prod_{k=1}^{\nu-1}\frac{p_1(x_k)}{p_0(x_k)}\in(A,\ B),\qquad \prod_{k=1}^{\nu}\frac{p_1(x_k)}{p_0(x_k)}\notin(A,\ B).$$

Последовательный критерий отношения правдоподобия требует в среднем меньшего числа наблюдений, чем критерий с фиксированным объемом выборки при тех же ошибках первого и второго рода.

Последовательный критерий заканчивается с вероятностью 1 как при гипотезе H_0, так и при гипотезе H_1, т. е. $P_0(\nu<\infty) = P_1(\nu<\infty)=1$.

Точное определение границ A и B и среднего объема выборки $M_0\nu$ в последовательном критерии связано с большими трудностями. Однако имеются полезные для приложений соотношения:

$$A\geqslant\frac{\beta}{1-\alpha},\quad B\leqslant\frac{1-\beta}{\alpha},$$

$$M_0\nu\simeq\frac{(1-\alpha)\log B+\alpha\log A}{M_0z},$$

$$M_0\nu\geqslant\frac{(1-\alpha)\log(B/(1-\alpha))+\alpha\log((1-\beta)/\alpha)}{M_0z}.$$

Здесь $z=\log\dfrac{p_1(\xi)}{p_0(\xi)}$.

Литература: [1, 4, 13, 42, 43, 56, 58, 72, 85, 99].

Глава 21. ТЕОРИЯ ОЦЕНИВАНИЯ ПАРАМЕТРОВ

21.1. Задача оценивания и свойства оценок

21.1.1. Постановка задачи. Пусть $\{X^n, \mathfrak{X}^n, P^n\}$, где $P=\{P_\theta,\ \theta\in\Theta\}$, — (параметризованная) статистическая структура модели простого выбора, связанная с наблюдением над случайным элементом ξ, принимающим значения в X.

Считается, что в множестве Θ допустимых значений параметра θ существует такое θ_0, что распределение элемента ξ совпадает с P_{θ_0}, т. е. $\mathsf{P}\{\xi \in \Gamma\} = P_{\theta_0}(\Gamma)$, $\Gamma \in \mathfrak{X}$. Значение θ_0 называется *истинным значением* параметра.

Предположим, что истинное значение θ_0 неизвестно, и задача заключается в том, чтобы на основании эксперимента над ξ, т. е. по выборке $(x_1, \ldots, x_n) \in X^n$, оценить θ_0. Процедура оценивания состоит в том, что, во-первых, строится статистика $\theta^* = \theta^*(x)$: $X^n \to \Theta$ (оценка, см. § 20.2), обладающая рядом предпочтительных свойств, и, во-вторых, вместо аргументов в $\theta^*(x) = \theta^*(x_1, \ldots, x_n)$ подставляется выборка — результаты наблюдений. Полученное значение (реализация случайной величины $\theta^*(\xi_1, \xi_2, \ldots, \xi_n)$, где ξ_i независимы и одинаково с ξ распределены) и принимается в качестве оценки истинного значения параметра θ_0. Таких оценок можно построить бесконечно много, и возникает вопрос, какие из них предпочесть. Ответ неоднозначен, так как можно вводить различные критерии качества оценок.

21.1.2. Принципы постоения оценок. Естественно считать, что качество оценки θ^* зависит от близости θ^* к истинному значению параметра. Термин «близость» нуждается в уточнении.

Во-первых, можно по-разному вводить понятие близости в множестве Θ. Например, если Θ — метрическое пространство, то можно считать мерой близости между элементами $\theta_1, \theta_2 \in \Theta$ расстояние между ними. Более общим образом на множестве Θ можно ввести функцию потерь $r(\theta_1, \theta_2)$ $(\theta_1, \theta_2 \in \Theta)$, т. е. неотрицательную функцию, интерпретируемую как потери, которые мы несем, если принимаем в качестве оценки истинного значения параметра значение θ_2, в то время как истинное значение равно θ_1. В этом случае оценка θ^* тем ближе к θ, чем меньше потери $r(\theta, \theta^*)$.

Во-вторых, так как результаты наблюдений $x \in X^n$ случайны, то оценка $\theta^*(x_1, x_2, \ldots, x_n)$ представляет собой случайный элемент, и поэтому близость θ^* к θ должна пониматься в некотором усредненном смысле. Например, можно считать, что θ^* близко к θ, если малы средние потери

$$\mathsf{M}_\theta r(\theta, \theta^*(x_1, \ldots, x_n)) = \int_X \cdots \int_X r(\theta, \theta^*(x_1, \ldots, x_n))\, P_\theta(dx_1) \ldots P_\theta(dx_n).$$

Однако для данной оценки $\theta^*(x_1, \ldots, x_n)$ эти потери могут быть малыми при одних θ и достаточно большими при других. Разумеется, если бы существовала такая оценка $\theta^*(x_1, x_2, \ldots, x_n)$, что для любой другой $\theta_1^*(x_1, \ldots, x_n)$ при всех $\theta \in \Theta$ выполнялось неравенство

$$\mathsf{M}_\theta r\left(\theta, \theta^*(x_1, x_2, \ldots, x_n)\right) \leqslant \mathsf{M}_\theta r\left(\theta, \theta_1^*(x_1, x_2, \ldots, x_n)\right),$$

то с точки зрения выбранной меры близости ее следовало бы предпочесть любой другой оценке. Однако, вообще говоря, такой оценки не существует. Поэтому для выбора оценки неизвестного параметра принимают во внимание некоторые дополнительные соображения.

Можно, например, выбрать оценку $\theta^*(x_1, x_2, \ldots, x_n)$ так, чтобы значение функции

$$\sup_{\theta \in \Theta} \mathsf{M}_\theta r(\theta, \theta^*(x_1, x_2, \ldots, x_n))$$

было минимальным. Этот принцип выбора оценок носит название *принципа минимакса*, а соответствующие оценки, если они существуют, называются *минимаксными*. При таком подходе мы стараемся минимизировать максимальный убыток, связанный с выбором той или иной оценки.

Другим подходом к выбору оценки является так называемый *байесовский* подход: считается, что имеются некоторые априорные соображения о предпочтительности тех или иных значений параметра θ. Другими словами, на пространстве Θ считается заданным некоторое априорное распределение $\mu(d\theta)$. Оценка $\theta^*(x_1, \ldots, x_n)$ выбирается так, чтобы значение интеграла

$$\int_\Theta \mathsf{M}_\theta r(\theta, \theta^*(x_1, x_2, \ldots, x_n))\,\mu(d\theta)$$

было минимальным. Возможны и другие принципы построения оценок.

21.1.3. Неравенство Крамера — Рао. В дальнейшем будем предполагать, что множество Θ является некоторым интервалом в $\mathbf{R}$ или же некоторой областью в евклидовом пространстве $\mathbf{R}^d$. Множество X, как правило, будет совпадать с $\mathbf{R}$.

Назовем оценку $\theta^*(x_1, x_2, \ldots, x_n)$ *несмещенной*, если при всех $\theta \in \Theta$

$$\mathsf{M}_\theta \theta^*(x_1, x_2, \ldots, x_n) = \theta.$$

В классе несмещенных оценок естественно считать лучшей ту оценку, распределение которой при всех θ концентрируется «теснее» вокруг среднего значения, т. е. вокруг θ. В случае, когда θ — одномерный параметр, мерой такой концентрации распределения может служить дисперсия распределения. Таким образом, мы приходим к задаче о нахождении в классе всех несмещенных оценок такой оценки $\theta^*(x_1, x_2, \ldots, x_n)$, для которой при каждом $\theta \in \Theta$ значение функции

$$\mathsf{M}_\theta (\theta^*(x_1, x_2, \ldots, x_n) - \theta)^2$$

минимально. Оказывается, что это выражение ограничено снизу некоторой функцией от θ, так что если для некоторой оценки $\theta^*(x_1, x_2, \ldots, x_n)$ эта нижняя граница для дисперсии достигается при всех θ, то это и будет искомая оценка.

Предположим, что распределение $P_\theta(dx)$ $(\theta \in \Theta)$ (здесь θ — одномерный параметр) имеет плотность $p(\theta, x)$ относительно некоторой σ-конечной меры $\nu(dx)$ на $(X, \mathfrak{X})$. В частности, $P_\theta(dx)$ может быть дискретным распределением, сосредоточенным в точках $z_1, z_2, \ldots \in X$. При этом точки $z_1, z_2, \ldots$ не зависят от θ и $p(\theta, z_k)$ — масса (вероятность), соответствующая точке z_k, так что $\sum_k p(\theta, z_k) = 1$.

Предположим далее, что плотности $p(\theta, x)$ дифференцируемы по θ, причем для любого измеримого множества Γ из X

$$\frac{\partial}{\partial\theta}\int\limits_{\Gamma} p(\theta, x)\,\nu(dx) = \int\limits_{\Gamma}\frac{\partial p(\theta, x)}{\partial\theta}\,\nu(dx).$$

Положим

$$I(\theta) = \int\limits_{X}\left[\frac{\partial \log p(\theta, x)}{\partial\theta}\right]^2 p(\theta, x)\,\nu(dx), \quad \theta \in \Theta.$$

Величина $I(\theta)$ называется *количеством информации* о параметре θ, содержащемся в одном наблюдении. В дискретном случае $I(\theta)$ запишется в виде

$$I(\theta) = \sum_{k=1}^{\infty}\left[\frac{\partial \log p(\theta, z_k)}{\partial\theta}\right]^2 p(\theta, z_k).$$

Количество информации о θ, содержащееся в независимых наблюдениях $x_1, x_2, \ldots, x_n$, равно $nI(\theta)$.

Пусть $\theta^*(x_1, x_2, \ldots, x_n)$ — произвольная несмещенная оценка параметра θ. При некоторых условиях регулярности имеет место неравенство

$$\sigma_\theta^2(\theta^*) = \mathrm{M}_\theta\left(\theta^*(x_1, x_2, \ldots, x_n) - \theta\right)^2 \geqslant \frac{1}{nI(\theta)}, \tag{1.1}$$

причем равенство достигается тогда и только тогда, когда

$$\sum_{k=1}^{n}\frac{\partial \log p(\theta, x_k)}{\partial\theta} = \lambda\,[\theta^*(x_1, \ldots, x_n) - \theta]$$

для почти всех $x \in X^n$ относительно меры $p(\theta, x_1)\ldots p(\theta, x_n)\nu(dx_1)\ldots\nu(dx_n)$. Здесь λ не зависит от $x_1, x_2, \ldots, x_n$, однако может зависеть от θ.

Упомянутые условия регулярности состоят в выполнении равенства

$$\mathrm{M}_\theta\frac{\partial \log L(\theta, x_1, \ldots, x_n)}{\partial\theta} = \int\limits_{X}\ldots\int\limits_{X}\frac{\partial \log L}{\partial\theta}L\nu(dx_1)\ldots\nu(dx_n) = 0,$$

где $L(\theta, x_1, \ldots, x_n) = p(\theta, x_1)p(\theta, x_2)\ldots p(\theta, x_n)$, а также в возможности продифференцировать по θ равенство

$$\int\limits_{X}\ldots\int\limits_{X}\theta^*(x_1, \ldots, x_n)\,L(\theta, x_1, \ldots, x_n)\,\nu(dx_1)\ldots(dx_n) = 0.$$

Неравенство (1.1) называется *неравенством Крамера — Рао* и дает нижнюю границу для дисперсии несмещенной оценки. Если для оценки $\theta^*(x_1, x_2, \ldots, x_n)$ в неравенстве (1.1) достигается равенство, то такая оценка называется *эффективной*. Таким образом, среди несмещенных регулярных оценок эффективные оценки имеют минимальную дисперсию. Назовем *эффективностью оценки* $\theta^*(x_1, x_2, \ldots, x_n)$

отношение нижней границы для дисперсии оценки к фактической дисперсии оценки:

$$\mathrm{eff}(\theta^*) = \frac{1}{nI(\theta)} \cdot \frac{1}{\sigma_\theta^2(\theta^*)}.$$

Очевидно, что $0 \leqslant \mathrm{eff}(\theta^*) \leqslant 1$. Две эффективные оценки одного и того же параметра почти наверное совпадают при каждом θ.

21.1.4. Достаточные оценки. Пусть структура $(X^n, \mathscr{X}^n, P^n)$, где $P = \{P_\theta, \theta \in \Theta\}$, доминируема мерой Лебега (см. § 20.1) и $\theta^* = \theta^*(x_1, \ldots, x_n)$ — оценка параметра θ. Согласно п. 20.2.2, оценка θ^* является достаточной для параметра θ, если условное распределение

$$Q_\theta\{A \mid \theta^*(x_1, x_2, \ldots, x_n) = t\}, \quad A \in \mathscr{X}^n, \ Q_\theta = P_\theta^n,$$

не зависит от θ. Оценка $\theta^*(x_1, x_2, \ldots, x_n)$ достаточна тогда и только тогда, когда условное распределение любой другой оценки при условии $\theta^*(x_1, x_2, \ldots, x_n) = t$ не зависит от θ. Следующее утверждение является следствием факторизационного критерия Неймана — Фишера.

Теорема 1. *Предположим, что меры $P_\theta(dx)$ $(\theta \in \Theta)$ абсолютно непрерывны относительно некоторой σ-конечной меры $\nu(dx)$, заданной на $(X, \mathfrak{B})$, и пусть $p(\theta, x) = dP_\theta/d\nu$. Оценка $\theta^*(x_1, x_2, \ldots, x_n)$ достаточна для параметра θ тогда и только тогда, когда имеет место представление*

$$p(\theta, x_1)\, p(\theta, x_2) \ldots p(\theta, x_n) = g_\theta(\theta^*(x_1, x_2, \ldots, x_n))\, h(x_1, x_2, \ldots, x_n)$$

где $g_\theta(\theta^*)$ и $h(x_1, x_2, \ldots, x_n)$ — неотрицательные функции, причем g_θ зависит от $x_1, x_2, \ldots, x_n$ только через оценку $\theta^*(x_1, x_2, \ldots, x_n)$, а $h(x_1, x_2, \ldots, x_n)$ не зависит от θ.

Важность понятия достаточной оценки в теории оценивания подчеркивает следующее утверждение.

Теорема 2. *Предположим, что $\theta^*(x_1, x_2, \ldots, x_n)$ — достаточная оценка для параметра θ, а $\theta_1^*(x_1, x_2, \ldots, x_n)$ — произвольная несмещенная оценка параметра θ. Тогда при всех $\theta \in \Theta$*

$$\mathsf{M}_\theta\left[f\left(\theta^*(x_1, x_2, \ldots, x_n)\right) - \theta\right]^2 \leqslant \mathsf{M}_\theta\left[\theta_1^*(x_1, x_2, \ldots, x_n) - \theta\right]^2,$$

где $f(t) = \mathsf{M}_\theta\left\{\theta_1^*(x_1, x_2, \ldots, x_n) \mid \theta^*(x_1, x_2, \ldots, x_n) = t\right\}$ (как следует из вышесказанного, функция $f(t)$ не зависит от θ). При этом функция $f(\theta^*(x_1, x_2, \ldots, x_n))$ также является несмещенной оценкой параметра θ.

Таким образом, имея для параметра θ произвольную несмещенную оценку $\theta_1^*(x_1, x_2, \ldots, x_n)$ и достаточную оценку $\theta^*(x_1, x_2, \ldots, x_n)$, мы можем построить новую несмещенную оценку параметра θ, которая при всех θ будет иметь дисперсию, меньшую, чем исходная оценка.

Достаточная оценка $\theta^*(x_1, x_2, \ldots, x_n)$ называется *полной*, если из того, что для некоторой функции $\varphi(t)$, $t \in \mathbf{R}$, выполнено соотношение $\mathsf{M}_\theta \varphi(\theta^*(x_1, x_2, \ldots, x_n)) = 0$, следует, что $\varphi(\theta^*(x_1, x_2, \ldots, x_n)) = 0$ почти наверное относительно меры Q_θ при любом $\theta \in \Theta$. Если существует полная несмещенная оценка

$\theta^*(x_1, x_2, \ldots, x_n)$ параметра θ, то для любой другой оценки $\theta_1^*(x_1, x_2, \ldots, x_n)$ параметра θ выполнено неравенство

$$\mathsf{M}_\theta\left[\theta_1^*(x_1, x_2, \ldots, x_n) - \theta\right]^2 \geqslant \mathsf{M}_\theta\left[\theta^*(x_1, x_2, \ldots, x_n) - \theta\right]^2,$$

т. е. в этом случае оценка $\theta^*(x_1, x_2, \ldots, x_n)$ имеет минимальную дисперсию.

21.1.5. Оценки многомерных параметров. Предположим, что множество Θ допустимых значений параметра является (открытой) областью в евклидовом пространстве $\mathbf{R}^d$. В этом случае оценка $\theta^*(x_1, x_2, \ldots, x_n)$ представляет собой также d-мерный вектор. Как и в одномерном случае, поставим задачу о нахождении среди всех несмещенных оценок таких из них, распределения которых обладали бы наибольшей степенью концентрации вокруг среднего значения. В многомерном случае удобной мерой такой концентрации является эллипсоид рассеивания.

Пусть ξ — d-мерный случайный вектор со средним a и матрицей вторых моментов $V = \|v_{ij}\|$: $a = \mathsf{M}\xi$, $v_{ij} = \mathsf{M}\xi^i\xi^j$ $(i, j = 1, 2, \ldots, d)$, где ξ^i — i-я компонента вектора ξ. Пусть S — эллипсоид в $\mathbf{R}^d$ с центром в точке a такой, что если сосредоточить внутри него равномерное распределение (т. е. распределение с постоянной плотностью), то это распределение будет иметь своим средним вектор a, а матрицей вторых моментов — матрицу V. Такой эллипсоид называется эллипсоидом рассеивания. В одномерном случае он превращается в интервал $(a - \sigma\sqrt{3},\ a + \sigma\sqrt{3})$, где a и σ^2 — математическое ожидание и дисперсия величины ξ соответственно. В общем случае уравнение эллипсоида рассеивания имеет вид

$$\left(V^{-1}(x - a),\ x - a\right) = d + 2,$$

где V^{-1} — матрица, обратная матрице V, x — текущая точка эллипсоида, (y, x) — скалярное произведение в $\mathbf{R}^d$.

Предположим теперь, что $\{P_\theta, \theta \in \Theta\}$ — семейство распределений в $(X, \mathfrak{X})$, имеющих плотность $p(\theta, x)$ относительно некоторой σ-конечной меры $\nu(dx)$, заданной на $\mathfrak{X}$. Пусть θ — d-мерный параметр. Положим

$$i_{kr}(\theta) = \int\limits_X \frac{\partial \log p(\theta, x)}{\partial \theta^k} \cdot \frac{\partial \log p(\theta, x)}{\partial \theta^r}\, p(\theta, x)\, \nu(dx),$$

где $k, r = 1, 2, \ldots, d$, и обозначим через $I(\theta)$ матрицу с элементами $i_{kr}(\theta)$. Матрица $I(\theta)$ называется *информационной матрицей*. Пусть далее $\theta^*(x_1, x_2, \ldots, x_n)$ — некоторая несмещенная оценка параметра θ. Обозначим через $S_{\theta^*}(\theta)$ эллипсоид рассеивания для распределения вектора $\theta^*(x_1, x_2, \ldots, x_n)$ по мере $Q_\theta(dx_1, dx_2, \ldots, dx_n) = p(\theta, x_1)\, p(\theta, x_2) \ldots p(\theta, x_n)\, \nu(dx_1) \ldots \nu(dx_n)$.

Многомерный аналог неравенства Крамера — Рао гласит: для любой несмещенной оценки $\theta^*(x_1, x_2, \ldots, x_n)$ параметра θ при некоторых условиях регулярности эллипсоид $S_{\theta^*}(\theta)$ содержит в себе эллипсоид, уравнение которого имеет вид

$$n\,(I(\theta)(u - \theta),\ u - \theta) = d + 2, \tag{1.2}$$

где n — число наблюдений, u — текущая точка эллипсоида ($u \in \mathbf{R}^d$). В предельном случае оба эллипсоида совпадают. В этом случае оценка $\theta^*(x_1, x_2, \ldots, x_n)$ называется *совместно-эффективной*. Матрица вторых моментов для совместно-эффективной оценки совпадает с $n^{-1}I^{-1}(\theta)$. *Эффективностью* оценки $\theta^*(x_1, x_2, \ldots, x_n)$ называется отношение объема эллипсоида (1.2) к объему эллипсоида $S_{\theta^*}(\theta)$. Очевидным образом переносятся на многомерный случай понятие достаточной оценки, а также свойства достаточных оценок, выраженные теоремами 1, 2.

21.1.6. Асимптотические свойства оценок. Предположим, что объем выборки n растет, т. е. $n \to \infty$, будем интересоваться асимптотическими свойствами оценок.

Оценка $\theta^*(x_1, x_2, \ldots, x_n)$ параметра θ называется *состоятельной*, если $\theta^*(x_1, x_2, \ldots, x_n) \to \theta$ по вероятности при $n \to \infty$. Иногда такие оценки называют слабо состоятельными в отличие от сильно состоятельных оценок, для которых соответствующая сходимость имеет место с вероятностью 1.

Заметим далее, что эффективные оценки существуют далеко не всегда. Однако во многих случаях существуют асимптотически эффективные оценки.

Назовем *асимптотической эффективностью* оценки $\theta^*(x_1, x_2, \ldots, x_n)$ предел

$$e_0(\theta) = \lim_{n\to\infty} \operatorname{eff}(\theta^*) = \lim_{n\to\infty} \frac{1}{nI(\theta)\,\sigma_\theta^2(\theta^*)},$$

если он существует. Во многих случаях $\sigma_\theta^2(\theta^*) \sim c/n$ при $n \to \infty$, и тогда $e_0(\theta^*) = [cI(\theta)]^{-1}$ существует. Очевидно, что $0 \leqslant e_0(\theta^*) \leqslant 1$. Если $e_0(\theta^*) = 1$, то оценка $\theta^*(x_1, x_2, \ldots, x_n)$ называется *асимптотически эффективной*.

21.2. Методы построения оценок

21.2.1. Метод моментов. Пусть случайная величина имеет распределение, принадлежащее семейству $\mathscr{P} = \{P_\theta, \theta \in \Theta\}$, где Θ — некоторая область в $\mathbf{R}^d$. Предположим, что существуют первые d моментов распределения P_θ, и положим

$$m_r(\theta) = \int_X x^r P_\theta(dx), \quad r = 1, 2, \ldots, d.$$

Здесь X совпадает с $\mathbf{R}$, или X — счетное множество. Имея n независимых наблюдений $x_1, x_2, \ldots, x_n$ над случайной величиной ξ, построим выборочные моменты

$$\bar{m}_r = \frac{1}{n}\sum_{k=1}^{n} x_k^r, \quad r = 1, 2, \ldots, d.$$

Метод моментов состоит в приравнивании выборочных моментов теоретическим. Получаем систему уравнений

$$m_r(\theta) = \bar{m}_r, \quad r = 1, 2, \ldots, d,$$

относительно d неизвестных $\theta^1, \theta^2, \ldots, \theta^d$. Если существует единственное решение $\theta_r^*(x_1, x_2, \ldots, x_n) = f_r(\bar{m}_1, \bar{m}_2, \ldots, \bar{m}_d)$ $(r = 1, \ldots, d)$ этой системы и функции f_r непрерывны, то получающаяся оценка $\{\theta_r^*, r = 1, 2, \ldots, d\}$ является состоятельной оценкой параметра θ. Однако, вообще говоря, оценки по методу моментов неэффективны.

21.2.2. Метод максимума правдоподобия. Предположим, что семейство распределений $\mathscr{P} = \{P_\theta, \theta \in \Theta\}$ на измеримом пространстве $(X, \mathfrak{X})$ имеет плотность распределения $p(\theta, x)$ относительно некоторой σ-конечной меры $\nu(dx)$, заданной на $\mathfrak{X}$. Если X дискретно, $p(\theta, x)$ есть вероятность того, что $\xi = x$, при условии что истинное значение параметра равно θ. Пусть $x_1, \ldots, x_n$ — результаты n независимых наблюдений над случайной величиной ξ. Согласно методу максимального правдоподобия, в качестве оценки $\theta^*(x_1, x_2, \ldots, x_n)$ параметра θ выбирается такая функция от наблюдений, которая доставляет максимум функции

$$p(\theta, x_1)\, p(\theta, x_2) \ldots p(\theta, x_n) = L(\theta, x_1, x_2, \ldots, x_n),$$

называемой *функцией правдоподобия*. Если при $\theta = \theta^*(x_1, x_2, \ldots, x_n)$ функция правдоподобия достигает наибольшего значения, то при этом же θ достигает наибольшего значения и функция $\log L(\theta, x_1, x_2, \ldots, x_n)$. Использование функции $\log L$ часто бывает более удобным. Оценки, получаемые с помощью метода максимального правдоподобия, называются *оценками максимального правдоподобия.*

Для отыскания оценок максимального правдоподобия нужно решить уравнения $(k = 1, 2, \ldots, d)$

$$\frac{\partial \log L(\theta, x_1, x_2, \ldots, x_n)}{\partial \theta^k} = 0,$$

если $\theta = (\theta^1, \theta^2, \ldots, \theta^d)$. Уравнения такого типа называются *уравнениями правдоподобия.* При решении уравнений правдоподобия следует отбросить решения вида $\theta = \mathrm{const}$ и рассматривать лишь те решения, которые зависят от $x_1, x_2, \ldots, x_n$ и попадают в область допустимых значений параметра Θ. Следует также иметь в виду, что наибольшее значение функция правдоподобия может принимать на границе области Θ.

Оценки максимального правдоподобия обладают следующими двумя важными свойствами:

а) если существует достаточная оценка $\theta^*(x_1, x_2, \ldots, x_n)$ для параметра θ, то каждое решение уравнения правдоподобия является функцией от $\theta^*(x_1, x_2, \ldots, x_n)$;

б) если для параметра θ существует эффективная оценка (в многомерном случае — совместно-эффективная) $\theta^*(x_1, x_2, \ldots, x_n)$, то уравнение правдоподобия имеет единственное решение $\theta^*(x_1, x_2, \ldots, x_n)$.

21.2.3. Асимптотическое поведение оценок максимального правдоподобия. Пусть Θ — интервал в $\mathbf{R}$, $X = \mathbf{R}$ и $\nu(dx) = dx$, где dx — мера Лебега на $\mathbf{R}$. Сформулируем теорему, показывающую, что оценки максимального правдоподобия имеют целый ряд хороших свойств при $n \to \infty$.

Теорема 1. *Предположим, что плотность $p(\theta, x)$ удовлетворяет условиям:*

1) *при каждом $\theta \in \Theta$ для почти всех x существуют производные*

$$\frac{\partial^k \log p(\theta, x)}{\partial \theta^k}, \quad k = 1, 2, 3;$$

2) *при каждом $\theta \in \Theta$ выполнены неравенства*

$$\left|\frac{\partial p(\theta, x)}{\partial \theta}\right| \leqslant G_1(x), \quad \left|\frac{\partial^2 p(\theta, x)}{\partial \theta^2}\right| \leqslant G_2(x), \left|\frac{\partial^3 p(\theta, x)}{\partial \theta^3}\right| \leqslant G_3(x),$$

где функции $G_1(x)$ и $G_2(x)$ интегрируемы на $\mathbf{R}$ по мере Лебега, а

$$\sup_{\theta \in \Theta} \int_{\mathbf{R}} G_3(x) p(\theta, x)\, dx < \infty;$$

3) *при каждом $\theta \in \Theta$ интеграл*

$$I(\theta) = \int_{\mathbf{R}} \left[\frac{\partial \log p(\theta, x)}{\partial \theta}\right]^2 p(\theta, x)\, dx$$

конечен и положителен.

Тогда уравнение правдоподобия имеет решение $\theta^*(x_1, x_2, \ldots, x_n)$, представляющее собой состоятельную, асимптотически эффективную и асимптотически нормальную оценку параметра θ.

Последнее означает, что величина

$$I(\theta) \sqrt{n}\, (\theta^*(x_1, x_2, \ldots, x_n) - \theta)$$

асимптотически нормальна с параметрами (0, 1), если истинное значение параметра равно θ.

Эта теорема обобщается на случай дискретной случайной величины, а также на случай многомерного параметра θ.

21.2.4. Метод минимума χ^2. Пусть $x_1, x_2, \ldots, x_n$ — независимые наблюдения над случайной величиной ξ со значениями в $(X, \mathfrak{X})$, распределение которой принадлежит семейству $\mathscr{P} = \{P_\theta, \theta \in \Theta\}$. Предположим, что пространство X разбито на r непересекающихся измеримых множеств $X_1, X_2, \ldots, X_r$. Обозначим через n_i число наблюдений в выборке $x_1, x_2, \ldots, x_n$, попавших в множество X_i. Если множество X конечно, т. е. случайная величина ξ принимает лишь конечное число значений, то можно считать, что X_i — одноточечное множество. Таким образом, произведена группировка результатов наблюдений.

Составим величину

$$\chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i(\theta))^2}{np_i(\theta)},$$

где $p_i(\theta) = P_\theta(X_i)$ $(i = 1, 2, \ldots, r,\ \theta \in \Theta)$. Оценка $\theta^*(x_1, x_2, \ldots, x_n)$ называется *оценкой по методу минимума* χ^2, если она получена минимизацией по θ величины χ^2. Если θ — d-мерный параметр, то для нахождения оценки по методу минимума χ^2 получаем систему уравнений

$$\sum_{i=1}^{n}\left[\frac{n_i - np_i(\theta)}{p_i(\theta)} + \frac{(n_i - np_i(\theta))^2}{2np_i^2(\theta)}\right]\frac{\partial p_i(\theta)}{\partial\theta^k} = 0, \quad k = 1, 2, \ldots, d.$$

По своим асимптотическим свойствам оценки, получаемые по методу минимума χ^2, весьма близки к оценкам максимального правдоподобия. Например, при некоторых условиях с вероятностью 1 имеется лишь один состоятельный корень соответствующих уравнений и он доставляет абсолютный минимум величине χ^2.

21.3. Доверительные области

21.3.1. Понятие доверительной области. В некоторых случаях важно не только дать оценку неизвестному параметру распределения, но и указать область, в которой предположительно должно находиться истинное значение параметра. Будучи построенной по результатам наблюдений, такая область может изменяться от выборки к выборке и потому является случайной областью. Следовательно, можно говорить о вероятности того, что такая область накрывает истинное значение параметра. Выбрав некоторое достаточно малое число $\varepsilon > 0$, мы можем задаться целью построить правило, позволяющее поставить в соответствие результатам наблюдений такую область в параметрическом множестве, что с вероятностью $1-\varepsilon$ истинное значение параметра будет содержаться в этой области. Это означает, что в длинном ряду выборок мы ошибаемся лишь в 100ε % случаев.

Области, о которых шла речь, называются *доверительными областями*, а число $1-\varepsilon$ — *коэффициентом доверия.*

21.3.2. Построение доверительных областей. Пусть $x_1, x_2, \ldots$ $\ldots, x_n$ — независимые наблюдения над случайной величиной ξ, распределение которой содержит неизвестный параметр, изменяющийся в пространстве $\mathbf{R}$, и пусть $\theta^*(x_1, x_2, \ldots, x_n)$ — какая-либо оценка параметра θ. Обозначим через $q_\theta(dt)$ распределение оценки θ^* в предположении, что истинное значение параметра совпадает с θ, т. е.

$$q_\theta(dt) = Q_\theta\{\theta^*(x_1, x_2, \ldots, x_n) \in dt\},$$

где $Q_\theta(dx_1, dx_2, \ldots, dx_n)$ — распределение в выборочном пространстве $\mathbf{R}^n$, определяемое формулой

$$Q_\theta(dx_1, dx_2, \ldots, dx_n) = P_\theta(dx_1)\,P_\theta(dx_2)\ldots P_\theta(dx_n).$$

Если распределение $q_\theta(dt)$ не имеет атомов, то по заданному $\varepsilon > 0$ всегда можно выбрать числа $a_1(\theta, \varepsilon)$ и $a_2(\theta, \varepsilon)$ таким образом, чтобы $a_1(\theta, \varepsilon) < a_2(\theta, \varepsilon)$ и

$$\int\limits_{\{t < a_1(\theta,\varepsilon)\}} q_\theta(dt) + \int\limits_{\{t > a_2(\theta,\varepsilon)\}} q_\theta(dt) = \varepsilon.$$

Такой выбор неоднозначен. Предположим, что эти функции можно выбрать так, чтобы они были непрерывны по θ и чтобы каждое из уравнений $a_i(\theta, \varepsilon) = t$ $(i = 1, 2, t \in \Theta)$, имело единственное решение $c_i(t, \varepsilon)$ $(i = 1, 2)$. Тогда соотношения

$$Q_\theta\left\{a_1(\theta, \varepsilon) < \theta^* < a_2(\theta, \varepsilon)\right\} = 1 - \varepsilon$$

и

$$Q_\theta\left\{c_1(\theta^*, \varepsilon) < \theta < c_2(\theta^*, \varepsilon)\right\} = 1 - \varepsilon$$

эквивалентны.

Таким образом, зная распределение оценки $\theta^*(x_1, x_2, \ldots, x_n)$, по заданному $\varepsilon > 0$ мы можем построить доверительный интервал $(c_1(\theta^*, \varepsilon), c_2(\theta^*, \varepsilon))$ с коэффициентом доверия $1 - \varepsilon$.

Заметим, что если распределение $q_\theta(dt)$ имеет атомы, то числа $a_1(\theta, \varepsilon)$ и $a_2(\theta, \varepsilon)$ нужно выбирать, исходя из неравенства

$$\int\limits_{\{t < a_1(\theta, \varepsilon)\}} q_\theta(dt) + \int\limits_{\{t > a_2(\theta, \varepsilon)\}} q_\theta(dt) \leqslant \varepsilon,$$

поскольку в этом случае может не существовать таких a_1 и a_2, для которых выполнялось бы соответствующее равенство. Продолжая построение так же, как и выше, получим доверительный интервал с коэффициентом доверия, не меньшим, чем $1 - \varepsilon$.

Очевидно, отправляясь от различных оценок θ^* параметра θ, мы будем получать различные доверительные интервалы. Желательно, чтобы длина доверительного интервала была возможно меньшей. Поэтому при построении доверительных интервалов естественно исходить из эффективных или асимптотически эффективных оценок, получаемых, например, из метода максимального правдоподобия.

21.3.3. Один метод построения доверительных интервалов. Предположим, что распределение $P_\theta(dx)$ не имеет атомов, и пусть существует функция $g(\theta, x_1, x_2, \ldots, x_n)$, $\theta \in \Theta$, $x_1, x_2, \ldots, x_n \in \mathbf{R}$, обладающая свойствами:

1) $g(\theta, x_1, x_2, \ldots, x_n)$ непрерывна и монотонна по θ;

2) функция распределения $Q_\theta\left\{g(\theta, x_1, x_2, \ldots, x_n) < \alpha\right\}$ не зависит от θ при всех $\alpha \in \mathbf{R}$ (определение Q_θ см. в п. 21.3.2).

По заданному $\varepsilon > 0$ выберем числа $a_1(\varepsilon)$ и $a_2(\varepsilon)$ так, чтобы

$$Q_\theta\left\{a_1(\varepsilon) < g(\theta, x_1, x_2, \ldots, x_n) < a_2(\varepsilon)\right\} = 1 - \varepsilon.$$

В силу условия 2) числа $a_1(\varepsilon)$ и $a_2(\varepsilon)$ не зависят от θ.

Обозначим через c_i $(i = 1, 2)$ числа, удовлетворяющие соотношениям $g(c_i, x_1, x_2, \ldots, x_n) = a_i(\varepsilon)$ $(i = 1, 2)$. Величины c_i зависят лишь от $x_1, x_2, \ldots, x_n$ и ε. Нетрудно видеть, что

$$Q_\theta\left(c_1 < \theta < c_2\right\} = 1 - \varepsilon$$

и, стало быть, (c_1, c_2) — доверительный интервал с коэффициентом доверия $1 - \varepsilon$.

Положим

$$F_\theta(x) = \int\limits_{-\infty}^{x} P_\theta(dy), \quad x \in \mathbf{R},$$

и предположим, что F_θ непрерывна и монотонна по θ. Тогда, как нетрудно проверить, функция $\prod_{k=1}^{n} F_\theta(x_k)$ удовлетворяет условиям 1), 2), и, значит, ее можно использовать для построения доверительных интервалов. В силу непрерывности $F_\theta(x)$ по x,

$$Q_\theta\{F_\theta(x_k) < \alpha\} = \begin{cases} 0 & \text{при } \alpha \leqslant 0, \\ \alpha & \text{при } 0 \leqslant \alpha \leqslant 1, \\ 1 & \text{при } \alpha > 1, \end{cases}$$

сумма $\sum_{k=1}^{n} \log F_\theta(x_k)$ имеет гамма-распределение

$$Q_\theta\left\{-\log a_2 < -\sum_{k=1}^{n} \log F_\theta(x_k) < -\log a_1\right\} = \\ = \int_{-\log a_2}^{-\log a_1} \frac{1}{\Gamma(n)} t^{n-1} e^{-t}\, dt.$$

По заданному $\varepsilon > 0$ можно выбрать числа a_1 и a_2 ($a_1 < a_2$) так, чтобы интеграл справа был равен $1-\varepsilon$. Отсюда получаем

$$Q_\theta\left\{a_1 < \prod_{k=1}^{n} F_\theta(x_k) < a_2\right\} = 1 - \varepsilon.$$

Так как функция $\prod_{k=1}^{n} F_\theta(x_k)$ непрерывна и монотонна по θ, то существуют такие c_1 и c_2, зависящие лишь от ε и $x_1, x_2, \ldots, x_n$, что (c_1, c_2) — доверительный интервал для θ с коэффициентом доверия $1-\varepsilon$.

Если функция $g(\theta, x_1, x_2, \ldots, x_n)$ удовлетворяет условию 2), непрерывна по θ, но не обязательно монотонна, то вместо доверительного интервала получится некоторая доверительная область. Этот же принцип можно использовать и для построения доверительных областей в том случае, когда θ — многомерный параметр.

21.3.4. Метод Байеса. Применяя метод построения доверительных интервалов, основанный на формуле Байеса, исходят из предположения, что параметр θ сам случаен. Предполагается также, что известно априорное распределение параметра. Обозначим плотность этого распределения относительно лебеговой меры (напомним, что Θ — интервал в $\mathbf{R}$) через $\varphi(\theta)$.

Далее, пусть $\theta^*(x_1, x_2, \ldots, x_n)$ — некоторая оценка параметра θ, построенная по независимым наблюдениям $x_1, x_2, \ldots, x_n$ над случайной величиной ξ. Предположим, что распределение $Q_\theta(dt)$ оценки θ^* (см. п. 21.3.2) также абсолютно непрерывно относительно лебеговой меры с плотностью $g(\theta, t)$. Согласно формуле Байеса, плотность распределения величины θ при фиксированном θ^* равна

$$\psi(\theta \mid \theta^*) = g(\theta, \theta^*)\,\varphi(\theta)\left(\int_{\Theta} g(\theta, \theta^*)\,\varphi(\theta)\,d\theta\right)^{-1}.$$

Поэтому условная вероятность того, что параметр θ лежит в пределах между c_1 и c_2 при условии, что фиксировано θ^*, выразится формулой

$$\mathsf{P}\{c_1 < \theta < c_2 \mid \theta^*\} = \int_{c_1}^{c_2} \psi(t \mid \theta^*)\,dt.$$

Теперь по заданному $\varepsilon > 0$ можем определить числа $c_1(\theta^*, \varepsilon)$ и $c_2(\theta^*, \varepsilon)$ так, чтобы

$$\mathsf{P}\{c_1(\theta^*, \varepsilon) < \theta < c_2(\theta^*, \varepsilon)\} = 1 - \varepsilon.$$

Тем самым для θ получен доверительный интервал с коэффициентом доверия $1-\varepsilon$.

Этот метод удобен не всегда, так как не всегда есть основания считать парметр случайным, или даже если он случаен, то не всегда известно его априорное распределение.

Литература: [14, 42, 49, 69, 72, 88].

Глава 22. ОЦЕНКИ ПАРАМЕТРОВ НЕКОТОРЫХ РАСПРЕДЕЛЕНИЙ

22.1. Оценки параметров нормального распределения

22.1.1. Оценка среднего при известной дисперсии. Пусть $x_1, x_2, \ldots, x_n$ — n независимых наблюденных значений нормальной случайной величины ξ с плотностью распределения

$$p(\theta, x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-\theta)^2}{2\sigma^2}\right\}, \quad x \in \mathbf{R},$$

где θ — неизвестный параметр, а параметр σ известен. Положим

$$\theta^*(x_1, x_2, \ldots, x_n) = \bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

Очевидно, что θ^* — нормально распределенная случайная величина с параметрами $\mathsf{M}_\theta\theta^* = \theta$ и $\mathsf{M}_\theta[\theta^* - \theta]^2 = \sigma^2/n$. Таким образом, оценка θ^* является несмещенной и состоятельной.

Далее, так как $I(\theta) = 1/\sigma^2$, то правая часть неравенства Крамера — Рао в рассматриваемом случае равна σ^2/n. Это означает, что оценка θ^* эффективна.

Случайная величина $\sqrt{n}\,(\theta^*-\theta)/\sigma$ имеет нормальное распределение с параметрами (0, 1). Найдя по заданному $\varepsilon>0$ (например, по таблицам) такое число c_ε, что

$$\frac{1}{\sqrt{2\pi}}\int_{-c_\varepsilon}^{c_\varepsilon} e^{-x^2/2}\,dx=1-\varepsilon,$$

получим

$$Q_\theta\left\{-c_\varepsilon<\frac{\sqrt{n}\,(\theta^*-\theta)}{\sigma}<c_\varepsilon\right\}=1-\varepsilon,$$

откуда

$$Q_\theta\left\{\theta^*-\frac{c_\varepsilon}{\sqrt{n}}\sigma<\theta<\theta^*+\frac{c_\varepsilon}{\sqrt{n}}\sigma\right\}=1-\varepsilon.$$

Здесь

$$Q_\theta(dx_1,\ldots,dx_n)=$$
$$=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\theta)^2\right\}dx_1\ldots dx_n.$$

Таким образом, $\left(\theta^*-\frac{c_\varepsilon}{\sqrt{n}}\sigma,\ \theta^*+\frac{c_\varepsilon}{\sqrt{n}}\sigma\right)$ — доверительный интервал с коэффициентом доверия $1-\varepsilon$.

22.1.2. Оценка дисперсии при известном среднем. В этом случае

$$p(\theta,x)=\frac{1}{\sqrt{2\pi\theta}}\exp\left\{-\frac{(x-a)^2}{2\theta}\right\},\quad x\in\mathbf{R},$$

где $\theta\in(0,\infty)$ — неизвестный параметр, а параметр a известен. Эффективной несмещенной оценкой для θ будет

$$\theta^*(x_1,x_2,\ldots,x_n)=\frac{1}{n}\sum_{k=1}^{n}(x_k-a)^2,$$

где $x_1, x_2, \ldots, x_n$ — результаты независимых наблюдений. Дисперсия оценки θ^* равна

$$\mathrm{M}_\theta(\theta^*-\theta)^2=2\theta^2/n.$$

Величина $n\theta^*/\theta$ имеет распределение χ^2 с n степенями свободы. Для построения доверительного интервала с коэффициентом доверия $1-\varepsilon$ выберем числа a_1 и a_2 (по таблицам) так, чтобы

$$Q_\theta\left\{a_1<\frac{n\theta^*}{\theta}<a_2\right\}=1-\varepsilon,$$

где

$$Q_\theta(dx_1, dx_2, \ldots, dx_n) =$$
$$= (2\pi\theta)^{-n/2} \exp\left\{-\frac{1}{2\theta}\sum_{k=1}^{n}(x_k - a)^2\right\} dx_1\, dx_2 \ldots dx_n.$$

Тогда

$$Q_\theta\left\{\frac{n\theta^*}{a_2} < \theta < \frac{n\theta^*}{a_1}\right\} = 1 - \varepsilon,$$

так что $(n\theta^*/a_2,\ n\theta^*/a_1)$ — искомый интервал.

22.1.3. Оценка дисперсии при неизвестном среднем. Задача такая же, как и в п. 22.1.2, однако параметр a считается неизвестным. Несмещенной состоятельной оценкой для параметра θ будет оценка

$$\theta^*(x_1, x_2, \ldots, x_n) = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2, \quad n > 1,$$

где $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$. Для нее

$$\mathrm{M}_\theta(\theta^* - \theta)^2 = 2\theta^2/(n-1).$$

Величина $(n-1)\theta^*/\theta$ имеет χ^2-распределение с $n-1$-й степенью свободы, и доверительный интервал строится так же, как и в п. 22.1.2.

22.1.4. Оценка среднего при неизвестной дисперсии. Задача такая же, как и в п. 22.1.1, однако параметр σ неизвестен. По-прежнему оценка $\theta^* = \bar{x}$ не смещена и состоятельна. Для построения доверительного интервала воспользуемся тем фактом, что величина $(\bar{x} - \theta)/s$, где $s^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2$, имеет распределение Стьюдента с $n-1$-й степенью свободы. Плотность этого распределения имеет вид

$$S_{n-1}(x) = \frac{\Gamma(n/2)}{\sqrt{\pi}\,\Gamma((n-1)/2)}(1 + x^2)^{-n/2}.$$

Определяя число c_ε из соотношения

$$2\int_0^{c_\varepsilon} S_{n-1}(x)\,dx = 1 - \varepsilon,$$

получаем

$$Q_\theta\left\{-c_\varepsilon < \frac{\bar{x} - \theta}{s} < c_\varepsilon\right\} = 1 - \varepsilon,$$

где

$$Q_\theta(dx_1, dx_2, \ldots, dx_n) =$$
$$= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\theta)^2\right\} dx_1\, dx_2 \ldots dx_n.$$

Таким образом, доверительным интервалом для параметра θ с коэффициентом доверия $1-\varepsilon$ является интервал $(\bar{x}-sc_\varepsilon,\ \bar{x}+sc_\varepsilon)$.

22.1.5. Совместное оценивание параметров среднего и дисперсии. Пусть в нормальном распределении неизвестны параметры a и σ. Решая уравнения правдоподобия, находим

$$a^* = \frac{1}{n}\sum_{k=1}^{n} x_k = \bar{x}, \quad (\sigma^*)^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\bar{x})^2 = s^2.$$

Оценка $(\sigma^*)^2$, однако, смещена. Положим $b^* = \frac{n}{n-1}s^2$. Тогда (a^*, b^*) — несмещенная оценка для параметров (a, σ^2).

Оптимальный эллипс (см. п. 21.1.5) имеет уравнение

$$\frac{(u-a)^2}{\sigma^2} + \frac{(v-\sigma^2)^2}{2\sigma^4} = \frac{4}{n}.$$

Эллипс рассеивания для оценки (a^*, b^*) задается уравнением

$$\frac{(u-a)^2}{\sigma^2} + \frac{n-1}{n}\,\frac{(v-\sigma^2)^2}{2\sigma^4} = \frac{4}{n}.$$

Совместная эффективность оценки (a^*, b^*) равна $(n-1)/n$, так что оценка (a^*, b^*) асимптотически совместно эффективна. Величины a^* и b^* независимы.

22.2. Оценки параметров биномиального и пуассоновского распределений

22.2.1. Биномиальное распределение. Пусть величина ξ принимает значения $0, 1, 2, \ldots, N$ с вероятностями $p_k(\theta) = C_N^k\theta^k(1-\theta)^{N-k}$ $(k = 0, 1, 2, \ldots, N)$ соответственно, причем параметр θ, $0 < \theta < 1$, неизвестен. Пусть $k_1, k_2, \ldots, k_n$ — результаты n независимых наблюдений случайной величины ξ. Положим

$$\theta^*(k_1, k_2, \ldots, k_n) = \frac{1}{nN}\sum_{i=1}^{n} k_i.$$

Тогда θ^* — несмещенная оценка параметра θ, для которой

$$\mathrm{M}_\theta(\theta^*-\theta)^2 = \frac{\theta(1-\theta)}{nN}.$$

С другой стороны, как нетрудно подсчитать,

$$I(\theta) = \frac{N}{\theta(1-\theta)}.$$

Отсюда следует, что θ^* — эффективная оценка параметра θ.

Для построения доверительного интервала воспользуемся тем фактом, что величина $\sqrt{n\theta}\,(\theta^*-\theta)/\sqrt{\theta(1-\theta)}$ асимптотически нормальна с параметрами (0, 1). Допуская

$$Q_\theta\left\{-a_\varepsilon < \frac{\sqrt{nN}\,(\theta^*-\theta)}{\sqrt{\theta(1-\theta)}} < a_\varepsilon\right\} \approx \frac{2}{\sqrt{2\pi}}\int_0^{a_\varepsilon} e^{-t^2/2}\,dt$$

и найдя a_ε из условия

$$\frac{2}{\sqrt{2\pi}}\int_0^{a_\varepsilon} e^{-t^2/2}\,dt = 1-\varepsilon,$$

получим

$$Q_\theta\left\{-a_\varepsilon < \frac{\sqrt{Nn}\,(\theta^*-\theta)}{\sqrt{\theta(1-\theta)}} < a_\varepsilon\right\} \approx 1-\varepsilon,$$

где Q_θ определяется равенством

$$Q_\theta(k_1, k_2, \ldots, k_n) = p_{k_1}(\theta)\,p_{k_2}(\theta)\ldots p_{k_n}(\theta), \qquad k_i = 0, 1, 2, \ldots, N$$

Значит, границами доверительного интервала с коэффициентом доверия $1-\varepsilon$ являются корни квадратного уравнения

$$\left(Nn + a_\varepsilon^2\right)x^2 - \left(2Nn\theta^* + a_\varepsilon^2\right)x + Nn\theta^{*2} = 0.$$

В частности, если в рассмотренной схеме положить $N = 1$, то мы получим, что результат i-го наблюдения случайной величины ξ есть появление или непоявление события $A = \{\xi = 1\}$. Вероятность этого события равна θ. Оценкой для нее служит $\theta^* = \nu/n$, где ν — число тех наблюдений, в которых событие A произошло.

22.2.2. Пуассоновское распределение. Пусть ξ имеет пуассоновское распределение. Это означает, что ξ может принимать значения 0, 1, 2, ... с вероятностями $p_k(\theta) = \frac{\theta^k}{k!}e^{-\theta}$ $(0 < \theta < \infty)$, причем θ — неизвестный параметр.

Несмещенной оценкой параметра θ является оценка

$$\theta^* = \frac{1}{n}\sum_{i=1}^{n} k_i,$$

где $k_1, k_2, \ldots, k_n$ — результаты независимых наблюдений случайной величины ξ. При этом

$$\mathsf{M}_\theta(\theta^*-\theta)^2 = \theta/n.$$

Поскольку $I(\theta)=\theta^{-1}$, то θ^* — эффективная оценка. Далее, величина

$$\sqrt{n/\theta}\,(\theta^*-\theta)$$

асимптотически нормальна с параметрами (0, 1). Поэтому при больших n имеет место приближенное равенство

$$Q_\theta\left\{-a_\varepsilon<\sqrt{n/\theta}\,(\theta^*-\theta)<a_\varepsilon\right\}\approx 1-\varepsilon,$$

где a_ε выбрано так же, как и в § 22.1, а мера Q_θ определяется формулой

$$Q_\theta(k_1, k_2, \ldots, k_n)=\frac{\theta^{k_1+k_2+\ldots+k_n}}{k_1!\,k_2!\ldots k_n!}e^{-n\theta},\qquad k_i=0, 1, 2, \ldots$$

Отсюда находим, что границы доверительного интервала с коэффициентом доверия $1-\varepsilon$ являются корнями квадратного уравнения

$$x^2-\left(2\theta^*+\frac{a_\varepsilon^2}{n}\right)x+\theta^{*2}=0.$$

Подчеркнем еще раз, что здесь, как и в п. 22.2.1, доверительный интервал построен приближенно, однако ошибка с ростом n стремится к нулю.

22.3. Оценки параметров равномерного распределения и Г-распределения

22.3.1. Равномерное распределение на отрезке с фиксированным концом. Пусть $x_1, x_2, \ldots, x_n$ — независимые наблюдения случайной величины ξ, распределенной равномерно на отрезке $[0, \theta]$ с неизвестным параметром θ. Положим $x_n^*=\max\limits_{1\leqslant k\leqslant n} x_k$. Оценка x_n^* является достаточной для параметра θ, и функция правдоподобия

$$L(\theta, x_1, x_2, \ldots, x_n)=\begin{cases} 1/\theta^n, & \text{если } \max\limits_{1\leqslant i\leqslant n} x_i\leqslant\theta,\ x_i\geqslant 0;\\ 0 & \text{в остальных случаях,}\end{cases}$$

принимает максимальное значение при $\theta=x_n^*$. Однако оценка x_n^* смещена. Несмещенной оценкой параметра θ будет

$$\theta^*=\frac{n+1}{n}x_n^*.$$

При этом θ^* будет иметь наименьшую среди всех несмещенных оценок дисперсию, равную

$$\mathsf{M}_\theta(\theta^*-\theta)^2=\frac{\theta^2}{n(n+2)}.$$

Обозначим через $Q_\theta(dx_1, dx_2, \ldots, dx_n)$ меру в $\mathbf{R}^n$, плотность которой относительно лебеговой меры равна $L(\theta, x_1, x_2, \ldots, x_n)$. Заметим, что распределение величины θ^*/θ не зависит от θ:

$$Q_\theta\left\{\frac{\theta^*}{\theta}<\alpha\right\}=\begin{cases} 0, & \text{если } \alpha\leqslant 0;\\ \left(\dfrac{\alpha n}{n+1}\right)^n, & \text{если } \alpha\in[0, 1+1/n];\\ 1, & \text{если } \alpha>1+1/n.\end{cases}$$

По заданному $\varepsilon > 0$ выберем a_ε так, чтобы

$$1-\varepsilon = Q_\theta\left\{\frac{\theta^*}{\theta} > a_\varepsilon\right\} = 1-\left(\frac{a_\varepsilon n}{n+1}\right)^n,$$

откуда найдем $a_\varepsilon = \frac{n+1}{n}\sqrt[n]{\varepsilon}$. Значит,

$$Q_\theta\left\{x_n^* < \theta < \frac{x_n^*}{\sqrt[n]{\varepsilon}}\right\} = 1-\varepsilon,$$

т. е. интервал $\left(x_n^*,\ x_n^*/\sqrt[n]{\varepsilon}\right)$ является доверительным с коэффициентом доверия $1-\varepsilon$.

22.3.2. Равномерное распределение на отрезке с неизвестными концами. Если случайная величина ξ имеет равномерное распределение на отрезке $[\theta_1, \theta_2]$ с неизвестными параметрами θ_1, θ_2 $(0 < \theta_1 < \theta_2)$, то несмещенными оценками с наименьшей дисперсией для параметров θ_1 и θ_2 будут соответственно оценки

$$\theta_1^* = \frac{nx_1^*}{n-1} - \frac{x_n^*}{n-1},$$

$$\theta_2^* = \frac{nx_n^*}{n-1} - \frac{x_1^*}{n-1},$$

где $x_1^* = \min\limits_{1\leqslant k\leqslant n} x_k$, $x_n^* = \max\limits_{1\leqslant k\leqslant n} x_k$. Несмещенными оценками с минимальной дисперсией для средней точки $(\theta_1+\theta_2)/2$ и размаха $\theta_2-\theta_1$ будут оценки соответственно

$$\ddot{\theta}_1^* = \frac{x_n^*+x_1^*}{2}, \quad \bar{\theta}_2^* = \frac{n+1}{n-1}\left(x_n^* - x_1^*\right).$$

Совместное распределение величин x_1^* и x_2^* определяется плотностью

$$f(x_1, x_2) = n(n-1)(x_2-x_1)^{n-2}(\theta_2-\theta_1)^{-n},$$

где $\theta_1 \leqslant x_1 \leqslant x_2 \leqslant \theta_2$.

22.3.3. Оценка масштабного параметра в Г-распределении. Пусть случайная величина ξ имеет Г-распределение

$$P_\theta(dx) = \begin{cases} \dfrac{x^{\lambda-1}e^{-x/\theta}}{\Gamma(\lambda)\,\theta^\lambda}, & \text{если } x > 0; \\ 0, & \text{если } x \leqslant 0, \end{cases}$$

с неизвестным параметром $\theta \in (0, \infty)$ и известным параметром $\lambda \in (0, \infty)$. Заметим, что при $\lambda = 1$ это распределение превращается в показательное распределение

$$P_\theta^0(dx) = \begin{cases} \theta^{-1}e^{-x/\theta}\,dx, & \text{если } x > 0; \\ 0, & \text{если } x \leqslant 0. \end{cases}$$

Пусть $x_1, x_2, \ldots, x_n$ — независимые наблюдения случайной величины ξ. Эффективной оценкой параметра θ является

$$\theta^* = \frac{1}{n\lambda} \sum_{k=1}^{n} x_k,$$

причем

$$\mathsf{M}_\theta (\theta^* - \theta)^2 = \theta^2/(n\lambda).$$

Заметим далее, что распределение величины $n\lambda\theta^*/\theta$ также является Г-распределением, не зависящим от θ:

$$Q_\theta \left\{ \frac{n\lambda\theta^*}{\theta} < x \right\} = \begin{cases} \displaystyle\int_0^x \frac{y^{n\lambda-1}e^{-y}}{\Gamma(n\lambda)}\,dy, & \text{если } x \geqslant 0; \\ 0, & \text{если } x < 0, \end{cases}$$

где мера $Q_\theta(dx_1, dx_2, \ldots, dx_n)$ в $\mathbf{R}^n$ определяется формулой

$$Q_\theta(dx_1, dx_2, \ldots, dx_n) = \begin{cases} \displaystyle\prod_{k=1}^{n} \frac{x_k^{\lambda-1}e^{-x_k/\theta}}{\theta^\lambda \Gamma(\lambda)}\,dx_k, & \text{если } x_k \geqslant 0,\quad k = 1, 2, \ldots, n; \\ 0 & \text{в противном случае.} \end{cases}$$

Таким образом, найдя числа a_1 и a_2 из соотношения

$$Q_\theta \left\{ a_1 + n\lambda < \frac{n\lambda\theta^*}{\theta} < a_2 + n\lambda \right\} = 1 - \varepsilon,$$

определим доверительный интервал $\left(\frac{n\lambda\theta^*}{n\lambda + a_2}, \frac{n\lambda\theta^*}{n\lambda + a_1} \right)$ с коэффициентом доверия $1 - \varepsilon$.

22.3.4. Оценка среднего значения Г-распределения. Пусть случайная величина ξ имеет Г-распределение

$$P_\theta(dx) = \begin{cases} \dfrac{x^{\theta-1}e^{-x}}{\Gamma(\theta)}\,dx, & \text{если } x > 0; \\ 0, & \text{если } x \leqslant 0, \end{cases}$$

с неизвестным параметром θ, $0 < \theta < \infty$. Нетрудно подсчитать, что в этом случае

$$I(\theta) = \frac{d^2 \log \Gamma(\theta)}{d\theta^2}.$$

Метод моментов приводит к оценке $\theta_1^* = \frac{1}{n} \sum_{k=1}^{n} x_k$, где $x_1, x_2, \ldots, x_n$ — независимые наблюдения над случайной величиной ξ. Для нее $M_\theta\theta_1^* = \theta$ и $M_\theta(\theta_1^* - \theta)^2 = \theta/n$. Эффективность оценки θ_1^* определяется формулой

$$\mathrm{eff}\left(\theta_1^*\right) = \left(\theta \frac{d^2 \log \Gamma(\theta)}{d\theta^2}\right)^{-1},$$

так что она не зависит от n и всегда меньше единицы. Уравнение правдоподобия имеет вид

$$\frac{1}{n}\sum_{k=1}^{n} \log x_i - \frac{d \log \Gamma(\theta)}{d\theta} = 0.$$

Единственный положительный корень этого уравнения определяет оценку максимального правдоподобия θ^*. Она является асимптотически эффективной и асимптотически нормальной с параметрами $\left(\theta, \left[n \frac{d^2 \log \Gamma(\theta)}{d\theta^2}\right]^{-1/2}\right)$.

Литература: [14, 42, 49, 69, 72, 88].

Глава 23. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

23.1. Линейные модели регрессии

23.1.1. Модели регрессии. Рассмотрим следующую задачу, часто возникающую при обработке результатов наблюдений, экспериментов, измерений и т. д.

Пусть $\xi \in \mathbf{R}^p$ и $\eta \in \mathbf{R}^q$ — зависимые случайные векторы. Требуется по результатам наблюдений $(x_1, \ldots, x_n)$ за вектором ξ и $(y_1, \ldots, y_n)$ за вектором η сделать обоснованное заключение о виде (характере) зависимости η от ξ.

Характерными примерами могут быть, например, исследование зависимости выхода готовой продукции от параметров технологических процессов, светимости звезд от их размеров, урожайности данного участка от количества выпавших на него осадков и т. п.

Теоретико-вероятностная формулировка подобного рода задач выглядит следующим образом. Пусть $\rho(y', y'')$ — некоторая метрика в $\mathbf{R}^q$. Требуется найти (борелевскую) функцию $\varphi(\cdot)$: $\mathbf{R}^p \to \mathbf{R}^q$, для которой $M\rho(\eta, \varphi(\xi))$ принимает минимальное значение. В том случае, когда $\rho(y', y'') = \|y' - y''\|^2$ и $\|\cdot\|$ — евклидова норма в $\mathbf{R}^q$, решение сформулированной задачи дается *функцией теоретической регрессии* $y = \varphi_{\eta|\xi}(x)$, где

$$\varphi_{\eta|\xi}(x) = M[\eta \mid \xi = x], \tag{1.1}$$

при этом переменную x называют *регрессионной переменной* или *регрессором*, а y — *откликом*. Решение задачи регрессии в форме (1.1) требует знания (или по крайней мере оценки) совместной функции распределения векоторов ξ и η — информации, которой исследователь в большинстве случаев не располагает. В этих условиях разумным представляется следующий подход.

Во-первых, в качестве оценки теоретической функции регрессии $y = \varphi_{\eta|\xi}(x)$ выбирается функция $y = \varphi(x)$ из определенного, обладающего хорошими аппроксимирующими возможностями класса функций $\Phi = \{\varphi: \mathbf{R}^p \to \mathbf{R}^q\}$.

Во-вторых, предполагается, что функция $\varphi(x) = \varphi(x, \theta)$ однозначно определяется некоторым параметром $\theta \in \Theta$, где Θ, например, некоторая область в $\mathbf{R}^m$ (m фиксировано).

В-третьих, функция $\varphi(x, \theta)$ зависит от θ линейно.

Таким образом, приходят к следующей линейной модели регрессии:

$$y = \theta_0\varphi_0(x) + \theta_1\varphi_1(x) + \ldots + \theta_{m-1}\varphi_{m-1}(x), \tag{1.2}$$

где $\varphi_i(x)$ — известные функции, определяющие план эксперимента, а m (порядок модели) и параметр $\theta = (\theta_0, \ldots, \theta_{m-1})$ подлежат оценке на основе результатов наблюдений $(x_1, \ldots, x_n)$ и $(y_1, \ldots, y_n)$ с очевидным требованием $m < n$ (цель эксперимента).

К классу линейных моделей принадлежат и более общие модели регрессии вида

$$p(y) = \theta_0\varphi_0(x) + \theta_1\varphi_1(x) + \ldots + \theta_{m-1}\varphi_{m-1}(x) \tag{1.3}$$

в случае, когда $p(y)$ имеет обратную функцию $p^{-1}(y)$.

Наиболее употребительные модели в статистическом регрессионном анализе (для простоты положим $p = q = 1$) имеют вид

а) $y = \theta_0 + \theta_1 x$, $y = \theta_0 + \theta_1 x + \ldots + \theta_k x^k$, $y = \theta_0 + \theta_1 \ln x$,

$$y = \theta_0 + \theta_1 x + \theta_2 \frac{1}{x}, \quad \ln y = \theta_0 + \theta_1 x + \ldots + \theta_k x^k,$$

$$\ln y = \theta_0 + \theta_1 \ln x, \quad 1/y = \theta_0 + \theta_1 x + \ldots + \theta_k x^k,$$

б) $\varphi_i(x)$ $(i = 0, \ldots, m-1)$ в (1.2) и (1.3) — ортогональные многочлены (например, многочлены Чебышева),

в) $\varphi_i(x)$ $(i = 0, \ldots, m-1)$ принадлежат классу сплайн-функций.

Наконец, в применении к задачам обработки результатов наблюдений (обработки данных) естественной, эффективно исследуемой является следующая общая линейная модель структуры наблюдений:

$$\begin{aligned} y_1 &= \theta_0\varphi_0(x_1) + \theta_1\varphi_1(x_1) + \ldots + \theta_{m-1}\varphi_{m-1}(x_1) + \varepsilon_1, \\ y_2 &= \theta_0\varphi_0(x_2) + \theta_1\varphi_1(x_2) + \ldots + \theta_{m-1}\varphi_{m-1}(x_2) + \varepsilon_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= \theta_0\varphi_0(x_n) + \theta_1\varphi_1(x_n) + \ldots + \theta_{m-1}\varphi_{m-1}(x_n) + \varepsilon_n, \end{aligned} \tag{1.4}$$

где ε_i $(i = 1, \ldots, n)$ интерпретируется как ошибка (шум) i-го измерения. Исследование модели вида (1.4), в которой вместо y_i стоят $p(y_i)$ с известной функцией $p(y)$, лишь очевидными деталями отличается от (1.4).

Приведенную общую линейную модель удобнее записать в матричной форме:

$$y = X\theta + \varepsilon, \tag{1.5}$$

где y — вектор-столбец результатов наблюдений: $X = \{X_{ij},\ i = 1, \ldots, n,\ j = 0, \ldots, m-1\}$ — $n \times m$-матрица известных коэффициентов ($X_{ij} = \varphi_j(x_i)$), которую называют *регрессионной матрицей*; $\theta = (\theta_0, \ldots, \theta_{m-1})^T$ — вектор-столбец неизвестных и подлежащих оценке параметров; $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T$ — вектор-столбец случайных ошибок. Обычно предполагается, имея в виду качество эксперимента, что $\mathrm{M}\,\varepsilon_i = 0$ $(i = 1, \ldots, n)$ (отсутствие систематической ошибки) $\mathrm{M}\varepsilon_i\varepsilon_j = \sigma^2\delta_{ij}$, σ^2 — дисперсия (неизвестная) ошибки наблюдения.

В том случае, когда ошибки измерений коррелированы (например, $\mathrm{M}\varepsilon\varepsilon^T = \sigma^2 V$, где V — известная матрица), замена $y_0 = V^{-1/2}y$, $x_0 = V^{-1/2}x$, $\varepsilon_0 = V^{-1/2}\varepsilon$ приводит к модели с $\mathrm{M}\varepsilon_0 = 0$, $\mathrm{M}\varepsilon_0\varepsilon_0^T = \sigma^2 I$, где I — единичная $n \times n$-матрица.

23.1.2. Принцип наименьших квадратов. Пусть в модели (1.4) $q = 1$ и m известно. В соответствии с приведенным в предыдущем пункте экстремальным свойством теоретической регрессии в качестве точечной оценки неизвестного параметра θ в (1.4) принимается оценка $\hat\theta$, реализующая *принцип наименьших квадратов*:

$$\min_{\theta \in \Theta} \sum_{i=1}^{n}\left[y_i - \sum_{j=0}^{m-1}\varphi_j(x_i)\,\theta_j\right]^2 = \sum_{i=1}^{n}\left[y_i - \sum_{j=0}^{m-1}\varphi_j(x_i)\,\hat\theta_j\right]^2. \tag{1.6}$$

Решение $\hat\theta = (\hat\theta_0, \ldots, \hat\theta_{m-1})$ задачи (1.6) называется *оценкой метода наименьших квадратов* или *МНК-оценкой* параметра θ.

Пусть область значений Θ параметра θ является линейным подпространством в $\mathbf{R}^m$ (в частности, $\Theta = \mathbf{R}^m$). В геометрической интерпретации, если θ — МНК-оценка параметра $\hat\theta$, то:

а) в случае, когда $\Theta = \mathbf{R}^m$ и $\operatorname{rang} X = m$, вектор $X\hat\theta$ является проекцией вектора наблюдений y (как вектора в $\mathbf{R}^n$, $m < n$) на подпространство $\Theta = \mathbf{R}^m$: $X\hat\theta = \mathrm{Pr}_\Theta y$ (оценка $\hat\theta$ единственна);

б) в случае, когда $\Theta = \mathbf{R}^m$, $\operatorname{rang} X < m$ и $\mathrm{Pr}_\Theta y$ не принадлежит области значений $R[X]$ линейного преобразования X, вектор $X\hat\theta$ является проекцией вектора $\mathrm{Pr}_\Theta y$ (как вектора в $\Theta = \mathbf{R}^m$) на подпространство $R[X]$ (оценка θ единственна);

в) в случае, когда $\Theta \subset \mathbf{R}^m$, $\operatorname{rang} X < m$ и $\mathrm{Pr}_\Theta y \in R[X]$, минимум в (1.6) достигается на всяком векторе $\hat\theta = \hat\theta_0 + h$, где $\hat\theta_0$ — любой вектор, для которого $X\hat\theta_0 = \mathrm{Pr}_\Theta y$, и h — любой вектор из подпространства пространства $\Theta \subset \mathbf{R}^m$, ортогонального $R[X]$ (МНК-оценка неединственна).

Использование принципа наименьших квадратов для получения точечных оценок неизвестных параметров (линейных) моделей в отличие от других методов точечного оценивания (см. § 21.2) не апеллирует к априорной информации о виде распределений, которой исследователь может не располагать на начальных этапах обработки данных.

23.2. Свойства МНК-оценок

23.2.1. Невырожденная линейная модель. Пусть в линейной модели (1.5) m известно, $\Theta \subset \mathbf{R}^m$, $\operatorname{rang} X = m$, или, что эквивалентно, $\det X^T X \neq 0$. Такую модель называют *невырожденной*. МНК-оценкой

$\hat{\theta}$ параметра θ является решение системы нормальных уравнений

$$X^T X\hat{\theta} = X^T y, \tag{2.1}$$

т. е.

$$\hat{\theta} = (X^T X)^{-1} X^T y,$$

причем $\mathsf{M}\hat{\theta} = \theta$, $\mathsf{M}(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T = \sigma^2 (X^T X)^{-1}$,

$$s^2_{\text{ост}} = y^T y - \hat{\theta}^T X^T X\hat{\theta}.$$

Несмещенной оценкой для σ^2 является

$$s^2 = \frac{1}{n-m}(y - X\hat{\theta})^T (y - X\hat{\theta}) = \frac{1}{n-m}\sum_{i=1}^{n}\left[y_i - \sum_{j=0}^{m-1}\varphi_i(x_j)\,\tilde{\theta}_j\right]^2.$$

Положим $P = X(X^T X)^{-1} X^T$. Матрица P, определяемая невырожденной линейной моделью, обладает следующими свойствами:

1) P — ортогональная проекционная матрица, т. е. $P = P^T$; $P^2 = P$;

2) $\operatorname{rang} P = \operatorname{Sp} P = m$;

3) $(I - P)X = 0$;

4) если столбцы матрицы X взаимно ортогональны, т. е. $\sum_{k=1}^{n}\varphi_i(x_k)\varphi_j(x_k) = 0$, $i \neq j$, $i, j = 0, \ldots, m-1$, то $X^T X$, $(X^T, X)^{-1}$ и P — диагональные матрицы.

Если в невырожденной модели (1.5) вектор ε имеет нормальное распределение с ковариационной матрицей $\sigma^2 I$ (I — $n \times n$-матрица), то

а) МНК-оценка $\hat{\theta}$ имеет m-мерное нормальное распределение

$$N_m\left(\theta,\ \sigma^2\left(X^T X\right)^{-1}\right);$$

б) квадратичная форма $(\hat{\theta} - \theta)^T X^T X(\hat{\theta} - \theta)/\sigma^2$ имеет χ^2-распределение с m степенями свободы;

в) МНК-оценка $\hat{\theta}$ не зависит от остаточной суммы квадратов $s^2_{\text{ост}}$;

г) $s^2_{\text{ост}}/\sigma^2$ имеет χ^2-распределение с $n - m$ степенями свободы.

23.2.2. Прямые измерения. Результаты независимых повторных измерений представляются в виде

$$y_i = \theta + \varepsilon_i, \quad i = 1, \ldots, n,$$

где θ — измеряемая величина, ε_i — ошибки измерений. Матрица X в линейной модели (1.5) для данного случая представляет вектор-столбец из единиц. Если $\mathsf{D}\varepsilon_k = \sigma^2$, то говорят о *прямых равноточных измерениях.*

МНК-оценка $\hat{\theta}$ для прямых равноточных измерений имеет вид

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \mathsf{D}\hat{\theta} = \frac{\sigma^2}{n}.$$

Несмещенная оценка для дисперсии σ^2 имеет вид

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat\theta)^2.$$

Если $\mathsf{D}e_k = \sigma^2/p_k$, где p_k известны и, вообще говоря, различны, то говорят о *прямых неравноточных измерениях.*

МНК-оценка $\hat\theta$ для прямых неравноточных измерений имеет вид

$$\hat\theta = \sum_{i=1}^{n} p_i y_i \left(\sum_{i=1}^{n} p_i\right)^{-1},$$

при этом

$$\mathsf{D}\hat\theta = \sigma^2\left(\sum_{k=1}^{n} p_k\right)^{-1}.$$

Несмещенная оценка для σ^2 имеет вид

$$s^2 = (n-1)^{-1}\sum_{k=1}^{n} p_k\left(y_k - \hat\theta\right)^2.$$

23.2.3. Невырожденная линейная модель с ограничениями. Рассмотрим невырожденную линейную модель из п. 23.2.1 при следующем дополнительном линейном ограничении: $C\theta = c$, где C — известная $r \times m$-матрица ранга r, c — известный вектор размерности r. Приведенное ограничение эквивалентно тому, что Θ является линейным подпространством в $\mathbf{R}^m$ размерности $r \leqslant m$.

Пусть $\hat\theta_C$ — МНК-оценка модели (1.5) с указанным ограничением. Тогда

$$\hat\theta_C = (I-P)\,\hat\theta + Ac = (I-P)\,\hat\theta + P\theta = \hat\theta + A\left(c - C\hat\theta\right),$$

$$\mathsf{M}\hat\theta_C = \theta, \quad \mathsf{M}\left(\hat\theta_C - \theta\right)^T\left(\hat\theta_C - \theta\right) = \\ = \mathsf{M}\left(\hat\theta - \theta\right)^T(I-P)\left(\hat\theta - \theta\right) + \theta^T P\theta,$$

где $A = (X^TX)^{-1}C^T[C(X^TX)^{-1}C^T]^{-1}$, $P = AC$. Несмещенная оценка для σ^2 имеет вид

$$s^2 = \frac{1}{n-m}\left(y^Ty - \hat\theta\,(I-P)\,y - [c^TA^TX^Ty + y^TXAc] + \right. \\ \left. + c^TC\,(X^TX)^{-1}\,C^Tc\right).$$

Здесь P — ортогональная проекционная $m \times m$-матрица, $\operatorname{rang} P = r$ (и, следовательно, $P = I$, если $r = m$), $C(I-P) = 0$.

23.2.4. Вырожденная линейная модель. Если линейная модель (1.5) вырождена, т. е. $\det X^TX = 0$, то МНК-оценки удобно определять в терминах так называемых псевдообратных или обобщенных обратных матриц.

Пусть Π — проекционная $m\times m$-матрица, однозначно определяемая как матрица максимального ранга, удовлетворяющая равенствам

$$\Pi X^TX = X^TX\Pi = 0, \quad \Pi^2 = \Pi.$$

Псевдообратной (*обобощенной обратной*) к матрице X^TX называется матрица $(X^TX)^{(-1)}$, определяемая равенством

$$(X^TX)^{(-1)} = (X^TX + \Pi)^{-1} - \Pi.$$

Если $\det X^TX \neq 0$, то $\Pi = 0$ и, следовательно, $(X^TX)^{(-1)} = (X^TX)^{-1}$, т. е. в этом случае псевдообратная матрица совпадает с обратной.

В случае вырожденной линейной модели метод наименьших квадратов определяет подпространство векторов, на котором достигается минимум суммы квадратов s^2 в (1.6), имеющее вид

$$(X^TX)^{(-1)} X^Ty + \Pi h,$$

где h — произвольный вектор из $\mathbf{R}^m$. По этой причине в вырожденной линейной модели оценивается не сам параметр 0, а некоторая линейная функция $B\theta$ этого параметра, где $B = \{b_{ij}, i = 1, \ldots, k, j = 1, \ldots, m\}$ — определенная $k \times m$-матрица, k — произвольное, но фиксированное число.

Линейная функция $B\theta$ параметра θ называется *оцениваемой*, если $B\Pi = 0$. МНК-оценка $\widehat{B\theta}$ оцениваемой функции $B\theta$ имеет вид

$$B\hat{\theta} = B(X^TX)^{(-1)} X^Ty,$$

причем $\mathsf{M}B\hat{\theta} = B\theta$, $\mathsf{M}[B\hat{\theta} - B\theta][B\hat{\theta} - B\theta]^T = \sigma^2 B(X^TX)^{(-1)} B^T$.

Несмещенной оценкой для σ^2 является

$$s^2 = \frac{1}{n-r}(y - X\hat{\theta})^T(y - X\hat{\theta}) = \frac{1}{n-r}\sum_{i=1}^{n}\left[y_i - \sum_{j=1}^{m} x_{ij}\hat{\theta}_j\right]^2,$$

где $\hat{\theta} = (X^TX)^{(-1)}X^Ty$, r — ранг матрицы X^TX.

Если $B_1\theta$ и $B_2\theta$ — две оцениваемые функции, то

$$\operatorname{Cov}\left(B_1\hat{\theta}, B_2\hat{\theta}\right) = \sigma^2 B_1 \left(X^TX\right)^{(-1)} B_2^T.$$

23.2.5. Линейные модели со случайными параметрами. Рассмотрим линейную модель вида

$$y = X\theta + \varepsilon,$$

где $y = (y_1, y_2, \ldots, y_n)^T$ — вектор наблюдаемых значений, $X = \{x_{ij}, i = 1, \ldots, n, j = 1, \ldots, m\}$ — матрица известных коэффициентов, $\theta = (\theta_1, \theta_2, \ldots, \theta_m)^T$ — неизвестный случайный вектор с математическим ожиданием $\mathsf{M}\theta = a$ и ковариационной матрицей $\mathsf{M}(\theta - a)(\theta - a)^T = C$, $\varepsilon = \left(\varepsilon_1, \ldots, \varepsilon_n\right)^T$ — не зависящий от θ вектор ошибок, $\mathsf{M}\varepsilon = 0$, $\mathsf{M}\varepsilon\varepsilon^T = \sigma^2 I$.

Задача оценки неизвестного вектора θ методом наименьших квадратов состоит в определении линейной функции $\hat{\theta}$ от вектора y наблюденных значений, на которой случайная величина

$$s(\theta) = (y - X\theta)^T(y - X\theta) = \sum_{i=1}^{n}\left[y_i - \sum_{j=1}^{m} x_{ij}\theta_j\right]^2$$

имеет минимальное математическое ожидание. Если вектор a математических ожиданий параметра θ известен, то

$$\hat{\theta} = (I - LX)\, a + Ly,$$

где $L = CX^T(\sigma^2 I + XCX^T)^{-1}$, причем $\mathsf{M}\hat{\theta} = a$, $\mathsf{M}(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T = C - CX^T(\sigma^2 + XCX^T)^{-1} XC$.

Если вектор a неизвестен, то

$$\hat{\theta} = (X^T X)^{-1} X^T y,$$

причем $\mathsf{M}[\hat{\theta} - \theta][\hat{\theta} - \theta]^T = \sigma^2 (X^T X)^{-1}$ и МНК-оценка случайного параметра θ та же, что и в случае детерминированного параметра.

Это верно, даже если $\det X^T X = 0$. Разумеется, здесь, как и в детерминированном случае, следует рассматривать оцениваемые линейные функции параметра θ и вместо $(X^T X)^{-1}$ использовать псевдообратную матрицу $(X^T X)^{(-1)}$.

23.2.6. **Общие свойства МНК-оценок.** МНК-оценки в общей линейной модели

$$y = X\theta + \varepsilon, \quad \mathsf{M}\varepsilon = 0, \quad \operatorname{Cov}\varepsilon = \sigma^2 I, \quad \det X^T X \neq 0$$

обладают следующими свойствами:

1) МНК-оценка $\hat{\theta}$ является несмещенной;

2) если $\lim\limits_{n\to\infty} \frac{1}{n}(X^T X)$ является невырожденной матрицей, то МНК-оценка $\hat{\theta}$ является состоятельной, а вектор $\sqrt{n}(\hat{\theta} - \theta)$ асимптотически нормален;

3) теорема Гаусса—Маркова. *МНК-оценка $\hat{\theta}$ обладает минимальной дисперсией в классе всех несмещенных линейных оценок параметра θ, причем $\hat{\theta}$ и $y - X\hat{\theta}$ некоррелированы.*

В классе несмещенных оценок МНК-оценка θ в линейной модели не обладает, вообще говоря, минимальной дисперсией.

Если вектор ε в (1.22) нормально распределен, то МНК-оценка $\hat{\theta}$ параметра θ наряду с 1)—3) обладает следующими свойствами:

4) $\hat{\theta}$ является оценкой метода максимального правдоподобия;

5) $\hat{\theta}$ обладает минимальной дисперсией в классе всех несмещенных оценок параметра θ;

6) $\hat{\theta}$ имеет нормальное распределение с математическим ожиданием θ и ковариационной матрицей $\sigma^2 (X^T X)^{-1}$;

7) случайная величина

$$\frac{(y - X\hat{\theta})^T (y - X\hat{\theta})}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left[y_i - \sum_{j=1}^{m} x_{ij}\hat{\theta}_j \right]^2$$

имеет χ^2-распределение с $n - m$ степенями свободы;

8) оценки $\hat{\theta}$ и $s^2 = \frac{1}{n-m}(y - X\hat{\theta})^T (y - X\hat{\theta})$ являются достаточными для θ и σ^2;

9) $\hat{\theta}$ и $y - X\hat{\theta}$ независимы.

В случае вырожденной линейной модели утверждения 1)—9) имеют место для любой оцениваемой функции параметра θ. Наконец, утверждения 1)—9) применимы и к линейной модели с ограничениями (п. 23.2.3), если заметить, что последняя эквивалентна следующей линейной модели без ограничений:

$$\tilde{y} = X\tilde{\theta} + \varepsilon,$$

где $\quad \tilde{y} = y - X(X^TX)^{-1}[C(X^TX)^{-1}C^T]^{-1}c, \qquad \tilde{\theta} = (I - P)\theta,$
$P = (X^TX)^{-1}C^T(C(X^TX)^{-1}C^T)^{-1}C.$

Для линейной модели с ограничениями имеем, кроме того, следующее: если вектор ошибок наблюдений имеет распределение $N(0, \sigma^2 I)$, то

10) $s^2_{\text{ост}}$ и $s^2_{C,\,\text{ост}} - s^2_{\text{ост}}$ независимы, где $s^2_{\text{ост}}$ — остаточная сумма квадратов для модели без ограничений, $s^2_{C,\,\text{ост}}$ — для модели с ограничениями, $s^2_{C,\,\text{ост}} - s^2_{\text{ост}} = (C\hat{\theta} - c)^T[C(X^TX)^{-1}C^T]^{-1}(C\hat{\theta} - c)$;

11) $s^2_{\text{ост}}/\sigma^2$ имеет χ^2-распределение с $n - s$ степенями свободы, где $s = \operatorname{rang} X$; для $R[C^T] \subset R[X^T]$ $\left(s^2_{C,\,\text{ост}} - s^2_{\text{ост}}\right)/\sigma^2$ имеет χ^2-распределение с $r = \operatorname{rang} C$ степенями свободы в предположении, что $C\theta = c$ верно; $F = \dfrac{s^2_{C,\,\text{ост}} - s^2_{\text{ост}}}{r}\left(\dfrac{s^2_{\text{ост}}}{n - s}\right)^{-1}$ имеет F-распределение Фишера — Снедекора с $(r, n - s)$ степенями свободы и служит основанием для использования F-критерия при проверке гипотезы о наличии или отсутствии линейных ограничений.

Пусть $y = X\theta + \varepsilon$ — невырожденная линейная модель,

$$X = X_0 + X_1, \quad \frac{\|X_1\|}{\|X_0\|} \ll 1,$$
$$\det X_0^T X_0 = 0, \quad \operatorname{rang} X_0^T X_0 = m - 1,$$

так что модель оказывается близкой к вырожденной. В этом случае ($\approx$ означает равенство с точностью до величин, определяемых порядком $\det X^TX$)

$$\tilde{\theta} \approx \frac{y^T X_1 \rho_0}{\rho_0^T Z_1 \rho_0}\rho_0, \quad s^2_{\text{ост}} \approx y^T y - \frac{(y^T X_1 \rho_0)^2}{\|X_1 \rho_0\|^2},$$

где вектор ρ_0 — собственный вектор матрицы $Z_0 = X_0^T X_0$.

23.3. Оценка параметров линейной регрессии

23.3.1. Простая линейная регрессия. Пусть модель линейной регрессии имеет вид

$$y = X\theta + \varepsilon,$$

где $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdot & \cdot \\ 1 & x_n \end{bmatrix}$, $\theta = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$. Здесь

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \end{bmatrix} = \begin{bmatrix} \dfrac{\sum x_i^2 \sum y_k - \sum x_i \sum x_k y_k}{n \sum x_i^2 - \left(\sum x_i\right)^2} \\ \dfrac{n \sum x_i y_i - \sum x_i \sum y_k}{n \sum x_i^2 - \left(\sum x_i\right)^2} \end{bmatrix},$$

$$s^2_{\text{ост}} = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2, \quad \hat{y}_i = \hat{\theta}_0 + \hat{\theta}_1 x_i.$$

Выборочная линия регрессии: $y = \hat{\theta}_0 + \hat{\theta}_1 x$.

23.3.2. Экспоненциальная регрессия. Пусть в модели (1.3) $p(y) = \ln y$, $\varphi_0(x) \equiv 1$, $\varphi_1(x) = X$. Тогда

$$y = \begin{bmatrix} \ln y_1 \\ \ln y_2 \\ \cdots \\ \ln y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdot & \cdot \\ 1 & x_n \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix},$$

$$\hat{\theta} = \begin{bmatrix} \dfrac{n \sum x_i \ln y_i - \sum x_i \sum \ln y_k}{n \sum x_i^2 - \left(\sum x_i\right)^2} \\ \dfrac{\sum x_i^2 \sum \ln y_k - \sum x_i \sum x_k \ln y_k}{n \sum x_i^2 - \left(\sum x_i\right)^2} \end{bmatrix}.$$

Выборочная линия регрессии: $y = \exp\left\{\hat{\theta}_0 + \hat{\theta}_1 x\right\}$.

23.3.3. Полиноминальная регрессия. Пусть в модели (1.2) $y = \theta_0 + \theta_1 x + \theta_2 x^2 + \ldots + \theta_{m-1} x^{m-1}$. Тогда

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{m-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 & \ldots & x_n^{m-1} \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \cdots \\ \theta_{m-1} \end{bmatrix}.$$

МНК-оценка $\hat{\theta}$ является решением системы нормальных уравнений

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \ldots & \sum x_i^{m-1} \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \ldots & \sum x_i^m \\ \cdot & \cdot & \cdot & \cdot \\ \sum x_i^{m-1} & \sum x_i^m & \sum x_i^{m+1} \ldots & \sum x_i^{2(m-1)} \end{bmatrix} \begin{bmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \\ \cdots \\ \hat{\theta}_{m-1} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \cdots \\ \sum x_i^{m-1} y_i \end{bmatrix}.$$

Выборочная линия регрессии:

$$y=\hat\theta_0+\hat\theta_1 x+\hat\theta_2 x^2+\ldots+\hat\theta_{m-1}x^{m-1}.$$

Практическое применение полиноминальной регрессии обладает существенным недостатком: повышение степени сглаживающего полинома требует проведения всех вычислений заново.

Полезным в этом случае является использование ортогональных полиномов Чебышева. Вместо оценивания параметра $\theta=(\theta_0,\ldots,\theta_{m-1})^T$ в модели

$$y_i=\theta_0+\theta_1 x_i+\theta_2 x_i^2+\ldots+\theta_{m-1}x_i^{m-1}+\varepsilon_i$$

оценивается параметр $\alpha=(\alpha_0,\alpha_1,\ldots,\alpha_{m-1})^T$ в модели

$$y_i=\alpha_0\varphi_0(x_i)+\alpha_1\varphi_1(x_i)+\ldots+\alpha_{m-1}\varphi_{m-1}(x_i)+\varepsilon_i, \tag{3.1}$$

где $\varphi_k(x)$ — полиномы степени k, обладающие следующим свойством ортогональности:

$$\sum_{i=1}^{n}\varphi_k(x_i)\,\varphi_l(x_i)=0,\quad k\neq l.$$

Модель (3.1) в матричной записи имеет вид

$$y=\Phi\alpha+\varepsilon,$$

где

$$\Phi=\{\varphi_{ij},\ i=1,\ldots,n,\ j=0,\ldots,m-1\},\quad \varphi_{ij}=\varphi_j(x_i).$$

Для параметра α МНК-оценка имеет вид

$$\hat\alpha=(\Phi^T\Phi)^{-1}\Phi^T y,$$

т. е.

$$\hat\alpha_k=\sum_{i=1}^{n}\varphi_k(x_i)\,y_i\left(\sum_{i=1}^{n}\varphi_k^2(x_i)\right)^{-1}.$$

Ортогональные полиномы Чебышева $\varphi_k(x)$ вычисляются по формулам

$$\varphi_0(x)\equiv 1,\quad \varphi_{k+1}(x)=\frac{\det M_{k+1}(x)}{\det M_k},\qquad k=0,\ldots,m-2,$$

где

$$M_k=\begin{bmatrix} n & \sum\limits_{i=1}^{n}x_i & \ldots & \sum\limits_{i=1}^{n}x_i^k\\ \sum\limits_{i=1}^{n}x_i & \sum\limits_{i=1}^{n}x_i^2 & \ldots & \sum\limits_{i=1}^{n}x_i^{k+1}\\ \cdots & \cdots & \cdots & \cdots\\ \sum\limits_{i=1}^{n}x_i^k & \sum\limits_{i=1}^{n}x_i^{k+1} & \ldots & \sum\limits_{i=1}^{n}x_i^{2k}\end{bmatrix};$$

$$M_{k+1}(x) = \begin{bmatrix} n & \sum\limits_{i=1}^{n} x_i & \dots & \sum\limits_{i=1}^{n} x_i^{k+1} \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2 & \dots & \sum\limits_{i=1}^{n} x_i^{k+2} \\ \dots & \dots & \dots & \dots \\ \sum\limits_{i=1}^{n} x_i^k & \sum\limits_{i=1}^{n} x_i^{k+1} & \dots & \sum\limits_{i=1}^{n} x_i^{2k+1} \\ 1 & x & \dots & x^{k+1} \end{bmatrix},$$

либо из рекуррентных соотношений

$$\varphi_0(x) \equiv 1,$$

$$\varphi_k(x) = x^k - \frac{\sum\limits_{i=1}^{n} x_i^k \varphi_{k-1}(x_i)}{\sum\limits_{i=1}^{n} \varphi_{k-1}^2(x_i)} \varphi_{k-1}(x) - \dots - \frac{\sum\limits_{i=1}^{n} x_i^k \varphi_0(x_i)}{\sum\limits_{i=1}^{n} \varphi_0^2(x_i)} \varphi_0(x).$$

Пример.

$$\varphi_1(x) = \frac{\begin{vmatrix} n & \sum\limits_{i=1}^{n} x_i \\ 1 & x \end{vmatrix}}{|n|} = x - \bar{x},$$

$$\varphi_2(x) = \frac{\begin{vmatrix} n & \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2 \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2 & \sum\limits_{i=1}^{n} x_i^3 \\ 1 & x & x^2 \end{vmatrix}}{\begin{vmatrix} n & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2 \end{vmatrix}} =$$

$$= x^2 - \frac{\sum\limits_{i=1}^{n} (x_i - \bar{x}) x_i^2}{\sum\limits_{i=1}^{n} (x_i - \bar{x})^2} (x - \bar{x}) - \frac{\sum\limits_{i=1}^{n} x_i^2}{n},$$

где

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

23.3.4. Оценка порядка линейной модели. Предположим, что в линейной модели (1.5) истинное значение порядка равно m_0 и нам не известно; в качестве предполагаемой модели выбирается модель $L_{(m)}$:

$$y_{(m)} = X_{(m)}\theta_{(m)} + \varepsilon$$

порядка m.

Такая модель $L_{(m)}$ является смещенной для $m < m_0$ и несмещенной для $m \leqslant m_0$.

В качестве статистики для проверки гипотезы о порядке модели естественно взять статистику

$$T_{(m)} = \frac{s^2_{(m),\,\text{ост}}}{s^2_{(m+1),\,\text{ост}}},$$

где $s^2_{(m),\,\text{ост}}$ — остаточная сумма квадратов для модели $L_{(m)}$. Статистика $T_{(m)} - \frac{n-m-1}{n-m}$ имеет F-распределение Фишера — Снедекора с $(1, n-m-1)$ степенями свободы, а статистика $\frac{n-m-1}{n-m} \times \frac{1}{T_{(m)}}$ имеет бета-распределение с параметрами $\left(\frac{n-m-1}{2}, \frac{1}{2}\right)$.

Критерий для выбора порядка модели с уровнем значимости α имеет вид

$$\frac{n-m-1}{n-m} < T_{(m)} < T_\alpha, \tag{3.2}$$

где $T_\alpha = \frac{n-m-1}{n-m} + \frac{1}{n-m} F_\alpha(1, n-m-1)$, $F_\alpha(1, n-m-1)$ — α-квантиль F-распределения Фишера — Снедекора с $(1, n-m-1)$ степенями свободы. Критерий (3.2) является одновременно критерием значимости параметра $\theta_{(m)}$ в модели $L_{(m+1)}$.

Литература: [3, 35, 57, 76].

Глава 24. СТАТИСТИКА СЛУЧАЙНЫХ ПРОЦЕССОВ

24.1. Различение гипотез

При различении гипотез для случайных процессов наблюдают траекторию случайного процесса $x(t)$ $(t \in [0, T])$; относительно конечномерных распределений имеются некоторые гипотезы, одну из которых следует выбрать на основании наблюдаемой траектории. В настоящее время разработано решение простейшей задачи — выбора одной из двух гипотез. Важным применением этой задачи является обнаружение сигнала на фоне шума. Например, на приемное устройство поступает некоторое сообщение. Следует установить наличие или отсутствие в поступившем сообщении полезного сигнала на фоне случайного шума.

24.1.1. Общая формулировка задачи о различении двух гипотез о распределении случайного процесса. Наблюдается траектория случайного процесса $x(t)$ $(t \in [0, T])$, о которой заранее известно, что она обязательно принадлежит некоторому пространству функций $F_{[0, T]}$, заданных на $[0, T]$ (например, пространству непрерывных функций или пространству функций без разрывов второго рода и т. п.). Относительно случайного процесса имеются две гипотезы H_i $(i = 1, 2)$. По гипотезе H_i мера, соответствующая процессу в пространстве $F_{[0, T]}$, есть мера μ_i (эта мера задана на σ-алгебре, порожденной цилиндрическими множествами, см. гл. 9). Нужно решить на основании наблюдения, какую из гипотез предпочесть.

Пусть R — некоторое правило, по которому принимается та либо другая гипотеза. Наиболее общий вид правила такой: для каждой возможной траектории $x(\cdot)$ должна быть задана вероятность $p(x(\cdot))$ (p — измеримый функционал от траектории) принять гипотезу H_1, если наблюдается траектория $x(\cdot)$, $1 - p(x(\cdot))$ — вероятность принять гипотезу H_2. Такое правило характеризуется вероятностями ошибок: $\alpha_{12} = \mathsf{P}(H_1/H_2)$ — принять гипотезу H_1, если верна H_2; $\alpha_{21} = \mathsf{P}(H_2|H_1)$ — принять гипотезу H_2, если верна H_1. Выражения для α_{12} и α_{21} задаются формулами

$$\alpha_{12} = \int p(x(\cdot))\,\mu_2(dx), \quad \alpha_{21} = \int (1 - p(x(\cdot)))\,\mu_1(dx).$$

Естественно искать правила, которые делают вероятности ошибок минимальными. Как мы увидим дальше, во многих случаях можно ограничиться нерандомизированными правилами, для которых $p(x(\cdot))$ принимает лишь значения 0 или 1. Тогда $F_{[0, T]}$ разбивается на два подмножества: G_1 и $G_2 = F_{[0, T]} - G_1$ (если $x(\cdot) \in G_1$, то принимается гипотеза H_1, если $x(\cdot) \in G_2$, то принимается гипотеза H_2). Вероятности ошибок задаются выражением

$$\alpha_{ij} = \mu_j(G_i), \quad i \neq j.$$

24.1.2. Абсолютная непрерывность мер в функциональных пространствах. При решении задач статистики случайных процессов очень важную роль играет абсолютная непрерывность или сингулярность мер, соответствующих рассматриваемым процессам. Пусть $F_{[0, T]}$ — некоторое фиксированное пространство, $\mathfrak{F}[0, T]$ — σ-алгебра подмножеств этого пространства, порожденная цилиндрическими множествами, μ_1 и μ_2 — две меры на $\mathfrak{F}[0, T]$, отвечающие случайным процессам (или одному и тому же процессу при разных гипотезах). Меры μ_1 и μ_2 взаимно *сингулярны* (ортогональны), если существуют такие непересекающиеся множества $G_i \in \mathfrak{F}[0, T]$, что $\mu_1(G_2) = \mu_2(G_1) = 0$. Мера μ_2 абсолютно непрерывна относительно μ_1, если $\mu_2(C) = 0$ для всех $C \in \mathfrak{F}[0, T]$, для которых $\mu_1(C) = 0$. В этом случае существует $\mathfrak{F}[0, T]$-измеримая функция $\rho(x(\cdot))$ такая, что для всех $C \in \mathfrak{F}[0, T]$

$$\mu_2(C) = \int_C \rho(x)\,\mu_1(dx). \tag{1.1}$$

Эта функция $\rho(x(\cdot))$ называется *плотностью меры* μ_2 *относительно* μ_1. Если μ_1 также абсолютно непрерывна относительно μ_2, то меры μ_1 и μ_2 называются *эквивалентными*; это будет тогда и только тогда, когда функция $\rho(x(\cdot))$, входящая в (1.1), будет положительна почти всюду по мере μ_1. В этом случае

$$\mu_1(C) = \int_C \frac{1}{\rho(x)} \mu_2(x). \tag{1.2}$$

Для любых двух (конечных) мер μ_1 и μ_2 можно указать такие попарно непересекающиеся множества Δ_1, Δ_2 и Δ, что $\mu_1(\Delta_2) = \mu_2(\Delta_1) = 0$, а на Δ эти меры эквивалентны, т. е. существует такая измеримая функция $\rho(x(\cdot))$, определенная на Δ, что для всех C

$$\mu_2(C \cap \Delta) = \int_{C \cap \Delta} \rho(x)\, \mu_1(dx), \quad \mu_1(C \cap \Delta) = \int_{C \cap \Delta} \rho^{-1}(x)\, \mu_2(dx). \tag{1.3}$$

24.1.3. Критерий Неймана — Пирсона дает правило, различающее гипотезы, такое, что при $\alpha_{12} \leqslant \varepsilon$ $(0 < \varepsilon < 1)$ α_{21} минимально. Предлагаемое правило в определенном смысле оптимально. Действительно, пусть имеется некоторое правило с вероятностями ошибок $\bar{\alpha}_{12}$ и $\bar{\alpha}_{21}$. С помощью критерия Неймана — Пирсона можно построить правило, у которого $\alpha_{12} \leqslant \bar{\alpha}_{12}$ и α_{21} — минимально возможное (значит, $\alpha_{21} \leqslant \bar{\alpha}_{21}$).

Опишем критерий в зависимости от свойств взаимной абсолютной непрерывности мер μ_1 и μ_2, соответствующих наблюдаемому процессу по гипотезам H_1 и H_2.

1. Пусть μ_1 и μ_2 ортогональны. Тогда можно указать такое множество G_1, что $\mu_1(G_1) = 1$, $\mu_2(G_1) = 0$. Если $x(\cdot) \in G_1$, принимаем гипотезу H_1; если $x(\cdot) \in G_2$, принимаем гипотезу H_2. Вероятности ошибок

$$\alpha_{12} = \mu_2(G_1) = 0, \quad \alpha_{21} = \mu_1(F_{[0,T]} - G_1) = 1 - \mu_1(G_1) = 0.$$

Таким образом, в рассматриваемом случае возможно безошибочное правило различения гипотез, т. е. такое правило, при котором $\alpha_{12} = \alpha_{21} = 0$.

2. Пусть μ_1 и μ_2 — эквиваленты и $\rho(x(\cdot))$ — плотность μ_2 относительно μ_1; $\rho(x(\cdot)) > 0$ для почти всех $x(\cdot)$ по мере μ_1. Обозначим

$$R_\lambda = \{x(\cdot) \colon \rho(x(\cdot)) < \lambda\}; \quad \Gamma_\lambda = \{x(\cdot) \colon \rho(x(\cdot)) = \lambda\}.$$

Пусть $\bar{\lambda}$ удовлетворяет соотношению

$$\mu_2(R_{\bar{\lambda}}) \leqslant \varepsilon, \quad \mu_2(R_{\bar{\lambda}} \cup \Gamma_{\bar{\lambda}}) \geqslant \varepsilon.$$

При $\varepsilon \in (0, 1)$ такое $\bar{\lambda}$ существует, так как при λ, возрастающем от 0 до ∞, $\mu_1(R_\lambda)$ возрастает от 0 до 1. Рассмотрим три подслучая:

а) $\mu_2\left(R_{\bar\lambda}\right)=\varepsilon$; тогда возьмем $G_1=R_{\bar\lambda}$, $G_2=F_{[0,\,T]}-G_1$, при этом $\alpha_{12}=\mu_2\left(R_{\bar\lambda}\right)=\varepsilon$,

$$\alpha_{21}=\mu_1\left(F_{[0,\,T]}-G_1\right)=1-\mu_1\left(R_{\bar\lambda}\right);$$

б) $\mu_2\left(R_{\bar\lambda}\right)<\varepsilon$, $\mu_2\left(R_{\bar\lambda}\cup\Gamma_{\bar\lambda}\right)=\varepsilon$; в этом случае $G_1=R_{\bar\lambda}\cup\Gamma_{\bar\lambda}$, $G_2=F_{[0,\,T]}-G_1$, $\alpha_{12}=\varepsilon$, $\alpha_{21}=1-\mu_1\left(R_{\bar\lambda}\right)-\mu_1\left(\Gamma_{\bar\lambda}\right)$;

в) $\mu_2\left(R_{\bar\lambda}\right)<\varepsilon$, $\mu_2\left(R_{\bar\lambda}\cup\Gamma_{\bar\lambda}\right)>\varepsilon$; тогда можно построить следующий рандомизированный критерий, задаваемый функционалом вероятности $p(x(\cdot))$: $p(x(\cdot))=1$, если $x(\cdot)\in R_{\bar\lambda}$; $p(x(\cdot))=0$, если $x(\cdot)\in F_{[0,\,T]}-R_{\bar\lambda}-\Gamma_{\bar\lambda}$; $p(x(\cdot))=\dfrac{\varepsilon-\mu_2\left(R_{\bar\lambda}\right)}{\mu_2\left(\Gamma_{\bar\lambda}\right)}$, если $x(\cdot)\in\Gamma_{\bar\lambda}$.

Заметим, что когда мера μ_2 (а значит, и μ_1) непрерывна, т. е. не существует траекторий, имеющих положительную вероятность, то и в случае в) можно построить нерандомизированный критерий: существует такое $D\subset\Gamma_{\bar\lambda}$, что $\mu_2(D)=\varepsilon-\mu_2\left(R_{\bar\lambda}\right)$. Полагая $G_1=R_{\bar\lambda}\cup D$, $G_2=F_{[0,\,T]}-R_{\bar\lambda}-D$, получим правило с $\alpha_{12}=\varepsilon$ и минимальным α_{21}.

3. В общем случае можно указать такие Δ_1, Δ_2, Δ, чтобы выполнялись равенства (1.3) и $\mu_2(\Delta_1)=\mu_1(\Delta_2)=0$. Пусть

$$R_\lambda=\{x(\cdot)\in\Delta\colon\ \rho(x(\cdot))<\lambda\}\cup\Delta_1;\quad \Gamma_\lambda=\{x(\cdot)\in\Delta\colon\ \rho(x(\cdot))=\lambda\}.$$

Если $\varepsilon\geqslant1-\mu_2(\Delta_2)$, то, выбирая $G_1=\Delta_1\cup\Delta$, $G_2=\Delta_2$, будем иметь

$$\alpha_{12}=\mu_2(\Delta_1\cup\Delta)=1-\mu_2(\Delta_2)\leqslant\varepsilon,\quad \alpha_{21}=\mu_1(\Delta_2)=0.$$

Если $\varepsilon<1-\mu_2(\Delta_2)$, выбираем $\bar\lambda$ так, чтобы $\mu_2\left(R_{\bar\lambda}\right)\leqslant\varepsilon$, $\mu_2\left(R_{\bar\lambda}\cup\Gamma_{\bar\lambda}\right)\geqslant\varepsilon$, и поступаем точно так же, как в случае 2.

Таким образом, для построения наилучшего критерия по Нейману — Пирсону нужно уметь строить множества Δ_i, на которых сосредоточены взаимно сингулярные меры, а в случае абсолютно непрерывных мер — плотность одной меры относительно другой, а также нужно знать распределение $\rho(x(\cdot))$ при одной и другой гипотезе.

24.2. Различение гипотез для процессов с независимыми приращениями

24.2.1. Основные обозначения. Пусть $x(t)$ — наблюдаемая траектория процесса, определенная на $[0,\,T]$, $x(t)\in\mathbf{R}$, $x(\cdot)\in D_{[0,\,T]}$, где $D_{[0,\,T]}$ — пространство функций без разрывов второго рода. По гипотезе H_k ($k=1,\ 2$) $x(t)$ является стохастически непрерывным процессом с независимыми приращениями, характеристическая функция которого задается формулой

$$\mathsf{M}\exp\{izx(t)\} = \\ = \exp\left\{iz\gamma_k(t) - \frac{1}{2}b_k(t)z^2 + \int\left(e^{izx} - 1 - \frac{izx}{1+x^2}\right)\Pi_k(t, dx)\right\}, \tag{2.1}$$

где функции $\Pi_k(t, A)$ для всех борелевских A, лежащих на положительном расстоянии от точки 0, непрерывны и не убывают, $\gamma_k(t)$ непрерывны, $\gamma_k(0) = b_k(0) = \Pi_k(0, A) = 0$ (т. е. предполагается, что $x(0) = 0$; если это не так, можно рассмотреть функцию $x(t) - x(0)$). Введем меры на борелевских подмножествах $[0, T]\times\mathbf{R}$:

$$\pi_k(B) = \iint\limits_{(t,\,x)\in B} d_t\Pi_k(t, dx).$$

Пусть далее Λ_1, Λ_2 и Λ — такие непересекающиеся множества в $[0, T]\times\mathbf{R}$, что $\Lambda_1\cup\Lambda_2\cup\Lambda = [0, T]\times\mathbf{R}$, $\pi_1(\Lambda_2) = \pi_2(\Lambda_1) = 0$, а на Λ меры π_1 и π_2 эквивалентны; тогда существует такая положительная измеримая функция $f(t, x)$, что

$$\pi_2(\Lambda\cap B) = \iint\limits_{\Lambda\cap B} f(t, x)\,\pi_1(dt\times dx),$$

$$\pi_1(\Lambda\cap B) = \iint\limits_{\Lambda\cap B} f^{-1}(t, x)\,\pi_2(dt\times dx).$$

Определим эмпирическую меру скачков:

$$\nu_{x(\cdot)}(B) = \sum_t \chi_B(t, x(t+0) - x(t-0)) \tag{2.2}$$

(так как процесс не имеет разрывов второго рода, то эта величина конечна для всех B, для которых $B\cap\{[0, T]\times(-\varepsilon, \varepsilon)\} = \varnothing$). Будем обозначать через μ_k меру, соответствующую процессу $x(t)$ в $D_{[0,\,T]}$ при гипотезе H_k.

24.2.2. Условия ортогональности. Безошибочное различение гипотез. Меры μ_1 и μ_2 ортогональны, если выполнено хотя бы одно из условий:

1) $\pi_1(\Lambda_1) + \pi_2(\Lambda_2) = +\infty$;

2) для некоторого $c\in(0, 1)$

$$\pi_1\{(s, x)\colon\ |1 - f(s, x)| > c\} = \infty;$$

3) $\displaystyle\iint\limits_{\Lambda} \frac{(1 - f(s, x))^2}{1 + f^2(s, x)}\,\pi_1(ds\times dx) = +\infty$;

4) $b_1(t) \not\equiv b_2(t)$ (при некотором $t\in[0, T]$);

5) пусть $\displaystyle\iint\limits_{\Lambda} \frac{(1 - f(s, x))^2}{1 + f^2(s, x)}\,\pi_1(ds\times dx) < \infty$ и $b_1(t) = b_2(t) = b(t)$ для всех $t\in[0, T]$, тогда определена функция

$$a_2(t) = \int \frac{y}{1+y^2}\,[\Pi_1(t, dy) - \Pi_2(t, dy)], \tag{2.3}$$

меры μ_k будут ортогональны, если либо функция $\gamma(t) = \gamma_2(t) - \gamma_1(t) + a_2(t)$ не абсолютно непрерывна относительно $b(t)$, либо

$$\int_0^T \left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t) = \infty. \tag{2.4}$$

Укажем безошибочные правила выбора гипотезы в каждом из этих случаев.

1а) Пусть $\pi_1(\Lambda_1) = +\infty$. Тогда обозначим через G_1 множество тех $x(\cdot) \in D_{[0,T]}$, для которых $\nu_{x(\cdot)}(\Lambda_1) > 0$. Если $x(\cdot) \in G_1$, принимаем гипотезу H_1; если $x(\cdot) \notin G_1$, принимаем гипотезу H_2.

1б) Пусть $\pi_2(\Lambda_2) = +\infty$. Обозначим через G_2 множество тех $x(\cdot) \in D_{[0,T]}$, для которых $\nu_{x(\cdot)}(\Lambda_2) > 0$. Если $x(\cdot) \in G_2$, принимаем гипотезу H_2; если $x(\cdot) \notin G_2$, принимаем гипотезу H_1.

2а) Пусть $\pi_1\{(s, x): f(s, x) > 1 + c\} = +\infty$. Выберем возрастающую последовательность измеримых множеств B_n так, чтобы $\pi_1(B_n) > n$ и $f(s, x) > 1 + c$ при $(s, x) \in B_n$. Обозначим через G_1 множество тех $x(\cdot)$, для которых

$$\lim_{n\to\infty} \frac{1}{\pi_1(B_n)} \nu_{x(\cdot)}(B_n) = 1.$$

Если $x(\cdot) \in G_1$, принимаем гипотезу H_1; если $x(\cdot) \notin G_1$, принимаем гипотезу H_2.

2б) Пусть $\pi_1\{(s, x): f(s, x) < 1 - c\} = +\infty$. Выберем возрастающую последовательность измеримых множеств B_n так, чтобы $\pi_1(B_n) > n$ и $f(s, x) < 1 - c$ при $(s, x) \in B_n$. Множество G_1 строится точно так же, как и в 2а).

3) Выберем такую последовательность разбиений $\Lambda = \bigcup_{k=1}^{m_n} V_{nk}$, чтобы для величин

$$\theta_n = \sum_{k=1}^{m_n} \pi_1(V_{nk}) \frac{(\pi_2(V_{nk}) - \pi_1(V_{nk}))^2}{\pi_1^2(V_{nk}) + \pi_2^2(V_{nk})}$$

выполнялось соотношение $\sum_{n=1}^{\infty} \theta_n^{-1} < \infty$. Тогда гипотеза H_1 принимается, если

$$\lim_{n\to\infty} \frac{1}{\theta_n} \sum_{k=1}^{m_n} [\nu_{x(\cdot)}(V_{nk}) - \pi_1(V_{nk})] \frac{\pi_2(V_{nk}) - \pi_1(V_{nk})}{\sqrt{\pi_1^2(V_{nk}) + \pi_2^2(V_{nk})}} = 0;$$

если это равенство не выполнено, принимается гипотеза H_2.

4) Предположим, что

$$\iint_\Lambda \frac{(1 - f(s, x))^2}{1 + f^2(s, x)} \pi_1(ds \times dx) < \infty,$$

$a_2(t)$ определено соотношением (2.3). Введем процесс

$$x^0(t) = x(t) - \iint y\left[\nu_{x(\cdot)}(ds \times dy) - \frac{1}{1-y^2}\pi_1(ds \times dy)\right]$$

(интеграл понимается как предел интеграла по области $[0, T] \times \{y: |y| > \varepsilon\}$ при $\varepsilon \downarrow 0$). При каждой гипотезе H_k $x^0(t)$ является непрерывным процессом с независимыми приращениями со средним $\gamma_1(t)$ при гипотезе H_1, $\gamma_2(t) + a_2(t)$ при гипотезе H_2 и дисперсией $b_k(t)$ при гипотезе H_k. Пусть t_0 таково, что $b_1(t_0) \neq b_2(t_0)$. Будем считать, что функции $b_k(t)$ строго возрастают. Рассмотрим два случая.

4а) Пусть при некотором $\delta > 0$ $b_1(t_0) < (1-\delta)b_2(t_0)$. Тогда для каждого n можно выбрать такие точки $t'_{n1} < t''_{n1} \leqslant t'_{n2} < t''_{n2} \leqslant \ldots$, чтобы $(1-\delta)\left[b_2(t''_{nk}) - b_2(t'_{nk})\right] \geqslant b_1(t''_{nk}) - b_1(t'_{nk}) > 0$. Примем G_1 равным множеству тех $x(\cdot)$, для которых

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{1}{b_1(t''_{nk}) - b_1(t'_{nk})}\left[x^0(t_{nk}) - x^0(t'_{nk}) - \gamma_1(t''_{nk}) + \gamma_1(t'_{nk})\right]^2 = 1.$$

4б) Пусть при некотором $\delta > 0$ $b_2(t_0) < (1-\delta)b_1(t_0)$. Выбираем точки $t'_{n1} < t''_{n1} \leqslant t'_{n2} < t''_{n2} \leqslant \ldots$ так, чтобы $(1-\delta)\left[b_1(t''_{nk}) - b_1(t'_{nk})\right] \geqslant b_2(t''_{nk}) - b_2(t'_{nk}) > 0$. Тогда обозначим через G_2 множество тех $x(\cdot)$, для которых

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{1}{b_2(t''_{nk}) - b_2(t'_{nk})}\left[x^0(t''_{nk}) - x^0(t'_{nk}) - \gamma_2(t''_{nk}) + \gamma_2(t'_{nk}) - a_2(t''_{nk}) + a_2(t'_{nk})\right]^2 = 1$$

В каждом из этих случаев гипотеза H_k принимается, если $x(\cdot) \in G_k$; если же $x(\cdot) \notin G_k$, то принимается другая гипотеза.

5) Если $\gamma(t)$ не абсолютно непрерывна относительно $b(t)$, то можно выбрать последовательность разбиений отрезка $[0, T]$: $0 = t_{n0} < t_{n1} < \ldots < t_{nm_n} = T$ так, чтобы величины

$$\theta_n = \sum_{k=1}^{m_n} \frac{(\gamma(t_{nk}) - \gamma(t_{nk-1}))^2}{b(t_{nk}) - b(t_{nk-1})}$$

при некотором $\alpha > 0$ удовлетворяли неравенству $\theta_n > n^\alpha$. Обозначим через G_1 множество тех $x(\cdot)$, для которых

$$\lim_{n\to 0} \frac{1}{\theta_n}\sum_{k=1}^{m_n}\left[\frac{x^0(t_{nk}) - x^0(t_{nk-1}) - \gamma_1(t_{nk}) + \gamma_1(t_{nk-1})}{b(t_{nk}) - b(t_{nk-1})}\right] \times \times \left[b(t_{nk}) - b(t_{nk-1})\right] = 0,$$

Принимаем гипотезу H_1, если $x(\cdot) \in G_1$, и гипотезу H_2, если $x(\cdot) \notin G_1$.

24.2.3. Условия эквивалентности мер. Пусть

1) $\Lambda_1 \cup \Lambda_2 = \varnothing$;

2) $\iint\limits_{\Lambda} \frac{(1-f(s,x))^2}{1+f^2(s,x)} \pi_1(ds \times dx) < \infty$;

3) $b_1(t) = b_2(t) = b(t)$;

4) функция $\gamma(t)$ абсолютно непрерывна относительно $b(t)$ и

$$\int_0^T \left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t) < \infty.$$

Тогда меры μ_1 и μ_2 эквивалентны и плотность μ_2 относительно μ_1 задается формулой

$$\rho(x(\cdot)) = \exp\left\{\int_0^T \frac{d\gamma(t)}{db(t)}\, d[x^0(t) - \gamma_1(t)] - \frac{1}{2}\int_0^T \left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t) + \right.$$

$$+ \int\limits_{|1-f(s,y)|\leqslant c} \ln f(s,y)\,[\nu_{x(\cdot)}(ds \times dy) - \pi_1(ds \times dy)] +$$

$$+ \int\limits_{|1-f(s,y)|>c} \ln f(s,y)\,\nu_{x(\cdot)}(ds \times dy) +$$

$$+ \int\limits_{|1-f(s,y)|\leqslant c} [\ln f(s,y) - f(s,y) + 1]\,\pi_1(ds \times dy) +$$

$$\left. + \int\limits_{|1-f(s,y)|>c} (1-f(s,y))\,\pi_1(ds \times dy)\right\} \quad (2.5)$$

(интегралы в формуле (2.5) следует понимать как стохастические).

Распределение $\rho(x(\cdot))$ зададим с помощью характеристической функции величины $\ln \rho(x(\cdot))$. При гипотезе H_1

$$\mathrm{M} \exp\{iz \ln \rho(x(\cdot))\} = \exp\left\{-\frac{1}{2} z^2 \int_0^T \left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t) + \right.$$

$$\left. + \int [e^{iz \ln f(s,y)} - 1 - iz(f(s,y) - 1)]\,\pi_1(ds \times dy) + iz\alpha_1\right\},$$

где

$$\alpha_1 = -\frac{1}{2}\int_0^T \left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t).$$

При гипотезе H_2

$$\mathsf{M}\exp\{iz\ln\rho(x(\cdot))\}=\exp\Big\{-\frac{1}{2}z^2\int_0^T\left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t)+$$
$$+\int\Big[e^{iz\ln f(s,\,y)}-1-iz\frac{f(s,\,y)-1}{2-2f(s,\,y)+f^2(s,\,y)}\Big]\times$$
$$\times\pi_2(ds\times dy)+iz\alpha_2\Big\},$$

где

$$\alpha_2=+\frac{1}{2}\int_0^T\left(\frac{d\gamma(t)}{db(t)}\right)^2 db(t)+\int\left(\frac{(1-f(s,\,y))^3}{2-2f(s,\,y)+f^2(s,\,y)}\right)\times$$
$$\times\pi_1(ds\times dy).$$

24.2.4. Общий случай. Очевидно, что при $\nu_{x(\cdot)}(\Lambda_1\cup\Lambda_2)>0$ можно достоверно выбрать H_i для того i, для которого $\nu_{x(\cdot)}(\Lambda_i)>0$ (одновременно $\nu_{x(\cdot)}(\Lambda_1)$ и $\nu_{x(\cdot)}(\Lambda_2)$ положительными быть не могут). Если $\nu_{x(\cdot)}(\Lambda_1\cup\Lambda_2)=0$, то распределение $x(\cdot)$ совпадает с распределением процесса с независимыми приращениями $x^*(t)$, характеристическая функция которого при гипотезе H_i имеет вид

$$\exp\Big\{i\gamma_k^*(t)\,z-\frac{1}{2}b_k(t)\,z^2+\iint\limits_{\{s<t\}\cap\Lambda}\left(e^{izy}-1-\frac{izy}{1+y^2}\right)\times$$
$$\times\pi_k(ds\times dy)\Big\},$$

где

$$\gamma_k^*(t)=\gamma_k(t)+\iint\limits_{\{s<t\}\cap\Lambda_k}\frac{y}{1+y^2}\pi_k(ds\times dy).$$

Меры μ_k^*, отвечающие процессу $x^*(\cdot)$ в $D_{[0,\,T]}$ при гипотезе H_k, будут эквивалентны, при этом плотность $\rho(x(\cdot))$ меры μ_2^* относительно μ_1^* определяется формулой (2.5), если в ней заменить $\pi_1(B)$ на $\pi_1^*(B)=\pi_1(B\cap\Lambda)$. Таким образом, если на основании наблюдения нельзя достоверно различить гипотезы, то для построения наилучшего правила можно использовать результаты предыдущего пункта.

24.2.5. Определение параметра однородного процесса Пуассона. Пусть $x(t)$ — ступенчатая функция, все скачки которой равны 1, $x(0)=0$. Гипотеза H_k заключается в том, что $x(t)$ — однородный процесс Пуассона с параметром λ_k ($k=1,\,2$). Таким образом, при гипотезе H_k

$$\mathsf{P}\{x(t)=n\}=\frac{(t\lambda_k)^n}{n!}e^{-t\lambda_k}.$$

Мера $\pi_k(ds \times dx)$ в этом случае при $B \in \mathbf{R}$ имеет вид

$$\pi_k(ds \times B) = \lambda_k \, ds \chi_B(1).$$

Значит, меры μ_k эквивалентны и

$$\rho(x(\cdot)) = (\lambda_2/\lambda_1)^{x(T)} e^{(\lambda_1 - \lambda_2) x(T)}.$$

Очевидно, что область $\{\rho(x(\cdot)) < \lambda\}$ совпадает с областью $\left\{x(T) < \dfrac{\ln \lambda}{\lambda_1 - \lambda_2 + \ln(\lambda_2/\lambda_1)}\right\}$. Поэтому правило, для которого $\alpha_{12} = \varepsilon$ и α_{21} минимально, строится следующим образом.

Обозначим через n_ε такое число, что

$$\sum_{k<n_\varepsilon} \frac{(T\lambda_2)^k}{k!} e^{-T\lambda_2} \leqslant \varepsilon \leqslant \sum_{k \leqslant n_\varepsilon} \frac{(T\lambda_2)^k}{k!} e^{-T\lambda_2}$$

$\left(\sum_{k<0} = 0\right)$. Тогда при $x(T) < n_\varepsilon$ принимается гипотеза H_1 и при $x(T) > n_\varepsilon$ — гипотеза H_2. Если $x(T) = n_\varepsilon$, то с вероятностью

$$\left[\varepsilon - \sum_{k<n_\varepsilon} \frac{(T\lambda_2)^k}{k!} e^{-T\lambda_2}\right] \left(\frac{(T\lambda_2)^{n_\varepsilon}}{n_\varepsilon!} e^{-T\lambda_2}\right)^{-1}$$

принимается гипотеза H_2, а с вероятностью

$$\left[\sum_{k \leqslant n_\varepsilon} \frac{(T\lambda_2)^k}{k!} e^{-T\lambda_2} - \varepsilon\right] \left(\frac{(T\lambda_2)^{n_\varepsilon}}{n_\varepsilon!} e^{-T\lambda_2}\right)^{-1}$$

— принимается гипотеза H_1.

24.2.6. Определение среднего однородного винеровского процесса. Пусть $x(t)$ — непрерывный процесс; при гипотезе H_1 — это винеровский процесс со средним 0 и дисперсией t, при гипотезе H_2 — со средним γt и дисперсией t. Меры μ_1 и μ_2 эквивалентны, и

$$\rho(x(\cdot)) = \exp\left\{\gamma x(T) - \frac{\gamma^2}{2} T\right\}.$$

Очевидно, что область $\{\rho(x(\cdot)) < \lambda\}$ имеет вид

$$x(T) < \frac{\ln \lambda + \gamma^2 T/2}{\gamma}$$

(считаем $\gamma > 0$). При гипотезе H_1 процесс $x(T)$ имеет нормальное распределение со средним 0 и дисперсией T, а при гипотезе H_2 — со средним γT и дисперсией T. Пусть $c(\varepsilon)$ удовлетворяет соотношению

$$\varepsilon = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c(\varepsilon)} e^{-u^2/2} \, du,$$

и пусть $\lambda = c(\varepsilon)\sqrt{T} + \gamma T$. Будем принимать гипотезу H_1, если $x(T) < \lambda$, и гипотезу H_2, если $x(T) \geqslant 1$. При этом

$$\alpha_{12} = \varepsilon, \quad \alpha_{21} = \frac{1}{\sqrt{2\pi}} \int\limits_{c(\varepsilon)+\gamma\sqrt{T}}^{\infty} e^{-u^2/2}\, du.$$

24.3. Различение гипотез для диффузионных процессов

24.3.1. Определение оператора диффузии. Пусть $x(t)$ $(t \in [0, T])$ — траектория диффузионного процесса в $\mathbf{R}^m$; при гипотезе H_k вектор переноса диффузионного процесса будет $a_k(t, x)$, а оператор диффузии — $B_k(t, x)$. Будем предполагать, что функции $a_k(t, x)$, $B_k(t, x)$ определены и непрерывны при $t \in [0, T]$, $x \in \mathbf{R}^m$. Заметим, что, зная траекторию процесса $x(t)$, можно определить $B_k(t, x(t))$ $(t \in [0, T])$, если верна гипотеза H_k. Это можно сделать следующим образом. Для $z \in \mathbf{R}^m$ положим

$$\lambda(t, z) = \lim_{n\to\infty} \sum_{k=0}^{2^n-1} \left(x\left(\frac{k+1}{2^n} t\right) - x\left(\frac{k}{2^n} t\right),\ z \right)^2 \tag{3.1}$$

($(\cdot, \cdot)$ — скалярное произведение в $\mathbf{R}^m$); предел (3.1) существует с вероятностью 1 при любой гипотезе. Тогда

$$\lambda(t, z) = \int\limits_0^t (B_k(s, x(s))\, z, z)\, ds, \tag{3.2}$$

если справедлива гипотеза H_k.

Если на наблюдаемой траектории при некоторых $t \in [0, T]$ и $z \in \mathbf{R}^m$

$$\int\limits_0^t (B_1(s, x(s))\, z, z)\, ds \neq \int\limits_0^t (B_2(s, x(s))\, z, z)\, ds$$

(в силу непрерывности подынтегральных функций это будет в том случае, когда при некоторых $t \in [0, T]$ и $z \in \mathbf{R}^m$ $(B_1(t, x(t))z, z) \neq$ $\neq (B_2(t, x(t))z, z)$), то равенство (3.2) может выполняться лишь при одном k. Значит, выбирая гипотезу H_k, если (3.2) выполняется при заданном k, получим безошибочное правило различения гипотез.

Пусть теперь вдоль наблюдаемой траектории

$$(B_1(t, x(t))\, z, z) = (B_2(t, x(t))\, z, z)$$

для всех $z \in \mathbf{R}^m$. Тогда $B_1(t, x(t)) = B_2(t, x(t))$. Поэтому можем считать, что $B_1(t, x) = B_2(t, x)$ для всех $t \in [0, T]$ и $x \in \mathbf{R}^m$.

24.3.2. Условие эквивалентности мер. Пусть μ_k — мера, соответствующая наблюдаемому процессу при гипотезе H_k (распределение $x(0)$ считаем заданным и не зависящим от выбора гипотезы). Обозначим $B(t, x) = B_1(t, x) = B_2(t, x)$, $a(t, x) = a_2(t, x) - a_1(t, x)$. Для того чтобы меры μ_1 и μ_2 были эквивалентны, достаточно, чтобы для всех t, x существовал вектор $b(t, x) \in \mathbb{R}^m$ такой, что

$$a(t, x) = B(t, x) b(t, x)$$

и с вероятностью 1

$$\int_0^T (a(t, x(t)), b(t, x(t)))\, dt < \infty.$$

При выполнении этих условий плотность меры μ_2 относительно μ_1 имеет вид

$$\rho(x(\cdot)) = \exp\left\{\int_0^T (b(t, x(t)), dx(t)) - \right.$$
$$\left. - \frac{1}{2}\int_0^T (b(t, x(t)), a_1(t, x(t)) + a_2(t, x(t)))\, dt\right\}. \quad (3.3)$$

Пусть $c(t, x) = (b(t, x), a_1(t, x) + a_2(t, x))$. Для нахождения распределения $\rho(x(\cdot))$ при гипотезе H_k введем величину

$$\xi_t = \int_t^T (b(s, x(s)), \quad dx(s)) - \frac{1}{2}\int_t^T c(s, x(s))\, ds.$$

Обозначим через $\mathsf{M}_{t,x}$ условное математическое ожидание при условии $x(t) = x$. Положим

$$u_\lambda(t, x) = \mathsf{M}_{t,x} e^{i\lambda\xi_t},$$

тогда функция $u_\lambda(t, x)$ удовлетворяет следующему уравнению при $t \in [0, T]$:

$$\frac{\partial u_\lambda(t, x)}{\partial t} + \frac{1}{2}[(B(t, x)\nabla, \nabla) + (a_1(t, x) + i\lambda b(t, x), \nabla)]\, u_\lambda -$$
$$- \left[i\lambda c(t, x) + \frac{\lambda^2}{2}(a(t, x), b(t, x))\right] u_\lambda = 0 \quad (3.4)$$

(здесь ∇ — вектор с компонентами $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$, где $x^1, \ldots, x^m$ — координаты точки x). При $t = T$ функция $u_\lambda(t, x)$ удовлетворяет граничному условию $u_\lambda(T, x) = 1$.

24.3.3. Процессы, однородные по пространству. Пусть $a_k(t, x) = a_k(t)$, $B_k(t, x) = B_k(t)$, т. е. коэффициенты диффузионного процесса не зависят от пространственной координаты. В этом случае процесс $x(t)$ будет процессом с независимыми приращениями. Из (3.2) следует, что

$$\lambda(t, z) = \int_0^t (B_k(s) z, z)\, ds,$$

если справедлива гипотеза H_k. Поэтому гипотезы достоверно различаются, если $B_1(t) \neq B_2(t)$ при некотором t.

Пусть $B_1(t) = B_2(t) = B(t)$, $a_2(t) - a_1(t) = a(t)$. Обозначим через E множество тех $t \in [0, T]$, для которых $a(t)$ не принадлежит области значений оператора $B(t)$. Пусть $P(t)$ — оператор проектирования на область значений оператора $B(t)$. Если мера Лебега множества E положительна, то гипотезы достоверно различаются: если справедлива гипотеза H_1, то

$$\int_0^T |P(t)(x(t) - a_1(t))|^2\, dt = 0;$$

если же справедлива гипотеза H_2, то

$$\int_0^T |P(t)(x(t) - a_1(t))|^2\, dt > 0.$$

Предположим, что почти для всех t вектор $a(t)$ принадлежит области значений оператора $B(t)$, т. е. существует такой вектор $b(t)$, что

$$a(t) = B(t)b(t)$$

(чтобы вектор $b(t)$ определялся однозначно, будем выбирать его в области значений $B(t)$; это возможно, так как $B(t)$ симметричен). Заметим, что $(a(t), b(t)) > 0$. Необходимым и достаточным условием абсолютной непрерывности мер является условие

$$\int_0^T (a(t), b(t))\, dt < \infty. \tag{3.5}$$

Если (3.5) выполнено, то меры μ_1 и μ_2 эквивалентны и плотность меры μ_2 относительно μ_1 имеет вид

$$\rho(x(\cdot)) = \exp\left\{\int_0^T (b(t), dx(t)) - \frac{1}{2}\int_0^T (b(t), a_1(t) + a_2(t))\, dt\right\}.$$

При любой из гипотез $\ln \rho(x(\cdot))$ имеет нормальное распределение с дисперсией

$$\int_0^T (a(t), b(t))\, dt$$

и средним значением

$$-\frac{1}{2}\int_0^T (a(t), b(t))\, dt \quad \text{при гипотезе } H_1;$$

$$\frac{1}{2}\int_0^T (a(t), b(t))\, dt \quad \text{при гипотезе } H_2.$$

Укажем, кроме того, правило безошибочного различения гипотез при условии

$$\int_0^T (a(t),\ b(t))\, dt = +\infty. \tag{3.6}$$

Выберем последовательность функций $b_n(t)$ такую, чтобы

$$\int_0^T (B(t)\, b_n(t),\ b_n(t))\, dt = \int_0^T (a(t),\ b_n(t))\, dt \geqslant n$$

(это возможно в силу (3.6)). Тогда выбираем гипотезу H_1, если

$$\lim_{n\to\infty} \left(\int_0^T (B(t)\, b_n(t),\ b_n(t))\, dt \right)^{-1} \int_0^T b_n(t)\, d\,[x(t) - a_1(t)] = 0,$$

и гипотезу H_2, если это условие не выполнено.

24.4. Различение гипотез о среднем значении гауссовского процесса

Пусть $x(t)$ $(t \in [0,\ T])$ — траектория одномерного гауссовского процесса с заданной непрерывной корреляционной функцией $R(t, s)$. Среднее значение процесса при гипотезе H_1 равно 0, а при гипотезе H_2 — заданной непрерывной функции $a(t)$ (если бы среднее значение при гипотезе H_1 было $a_1(t)$, можно было бы вместо $x(t)$ рассматривать $x(t) - a_1(t)$).

24.4.1. Условия ортогональности. Безошибочное различение гипотез. Обозначим через $L_2[0,\ T]$ пространство функций $g(t)$ на $[0,\ T]$, интегрируемых с квадратом; это — гильбертово пространство со скалярным произведением $(g_1, g_2) = \int_0^T g_1(t)\, g_2(t)\, dt.$

Пусть $Rg(t) = \int_0^T R(t,\ s)\, g(s)\, ds$ — линейный оператор в $L_2[0,\ T]$.

Этот линейный оператор вполне непрерывен и имеет последовательности собственных значений и собственных функций $\{\lambda_k,\ \varphi_k\}$

$$\int_0^T R(t,\ s)\, \varphi_k(s)\, ds = \lambda_k \varphi_k(t),$$

функции φ_k попарно ортогональны и полны в области значений оператора R, $\lambda_k > 0$ и $\sum_k \lambda_k < \infty$.

1) Если $a(t)$ не разлагается в ряд по функциям $\varphi_k(t)$, то меры μ_1 и μ_2, отвечающие процессу при гипотезах H_1 и H_2, ортогональны.

Правило, различающее гипотезы, заключается в следующем. Пусть

$$\hat{a}(t)=a(t)-\sum_{k=1}^{\infty}\varphi_k(t)\int_0^T\varphi_k(s)\,a(s)\,ds,\quad (\varphi_k,\varphi_k)=1.$$

Если $\int_0^T x(t)\,\hat{a}(t)\,dt=0$, принимаем гипотезу H_1;

если $\int_0^T x(t)\,\hat{a}(t)\,dt\neq 0$, принимаем гипотезу H_2.

2) Пусть $a(t)$ разлагается по собственным функциям φ_k:

$$a(t)=\sum_{k=1}^{\infty}\alpha_k\varphi_k(t),\quad \alpha_k=\int_0^T a(t)\,\varphi_k(t)\,dt.$$

Если

$$\sum_{k=1}^{\infty}\frac{\alpha_k^2}{\lambda_k}=\infty,\tag{4.1}$$

то меры μ_1 и μ_2 ортогональны. Различающее правило строится следующим образом. Выберем подпоследовательность m_n так, чтобы выполнялось неравенство

$$\sum_{k=1}^{m_n}\frac{\alpha_k^2}{\lambda_k}\geqslant n.$$

Если

$$\lim_{n\to\infty}\left(\sum_{k=1}^{m_n}\frac{\alpha_k^2}{\lambda_k}\right)^{-1}\sum_{k=1}^{m_n}\frac{\alpha_k}{\lambda_k}\int_0^T x(t)\,\varphi_k(t)\,dt=0,$$

то принимается гипотеза H_1; если это условие не выполняется, то принимается гипотеза H_2.

Приведенное выше правило различения гипотез использует собственные значения и собственные функции интегрального оператора R. Можно предложить и другие правила, не использующие этих данных.

Пусть $\{\psi_k(t),\ k=1,2,\ldots\}$ — произвольная ортонормированная система функций в $L_2[0,T]$. Примем обозначения

$$x_k=\int_0^T x(t)\,\psi_k(t)\,dt,\quad a_k=\int_0^T a(t)\,\psi_k(t)\,dt,$$

$$r_{jk}=\mathrm{M}\int_0^T x(t)\,\psi_k(t)\,dt\int_0^T x(t)\,\psi_j(t)\,dt$$

(математическое ожидание берется при гипотезе H_1). Пусть $b_k^{(n)}$ — решение системы линейных уравнений

$$a_k = \sum_{j=1}^{n} r_{jk} b_j^{(n)}.$$

Если выполнено (4.1), то

$$\lim_{n\to\infty} \sum_{k=1}^{n} a_k b_k^{(n)} = +\infty.$$

Выберем m_n так, чтобы выполнялось неравенство

$$\sum_{k=1}^{m_n} a_k b_k^{(m_n)} \geqslant n.$$

Тогда правило выбора гипотезы заключается в следующем: если

$$\lim_{n\to\infty} \left(\sum_{k=1}^{m_n} a_k b_k^{(m_n)} \right)^{-1} \sum_{k=1}^{m_n} x_k b_k^{(m_n)} = 0,$$

выберем гипотезу H_1; если последнее соотношение не выполнено, выбираем гипотезу H_2.

24.4.2. Условия эквивалентности мер. В обозначениях предыдущего пункта необходимым и достаточным условием эквивалентности мер μ_1 и μ_2 является условие

$$\sum_{k=1}^{\infty} \frac{\alpha_k^2}{\lambda_k} < \infty. \tag{4.2}$$

Если (4.2) выполнено, то плотность меры μ_2 относительно меры μ_1 имеет вид

$$\rho(x(\cdot)) = \exp\left\{ \sum_{k=1}^{\infty} \frac{x_k \alpha_k}{\lambda_k} - \frac{1}{2} \sum_{k=1}^{\infty} \frac{\alpha_k^2}{\lambda_k} \right\}. \tag{4.3}$$

В том случае, когда существует решение $b(t)$ уравнения Фредгольма первого рода

$$a(t) = \int_0^T R(T(t, s)\, b(s))\, ds, \tag{4.4}$$

плотность $\rho(x(\cdot))$ может быть записана в более обозримом виде:

$$\rho(x(\cdot)) = \exp\left\{ \int_0^T x(s)\, b(s)\, ds - \frac{1}{2} \int_0^T a(s)\, b(s)\, ds \right\} \tag{4.5}$$

(заметим, что во многих случаях решение уравнения (4.4) существует как обобщенная функция, а интегралы в (4.5) можно превратить в обычные интегрированием по частям); $\ln \rho(x(\cdot))$ имеет нормальное распределение с дисперсией

$$d = \sum_{k=1}^{\infty} \frac{\alpha_k^2}{\lambda_k}$$

при обеих гипотезах; среднее значение равно $-d/2$ при гипотезе H_1 и $d/2$ при гипотезе H_2. В том случае, когда уравнение (4.4) решить трудно, можно использовать следующий приближенный критерий различения гипотез.

Пусть $a_k^{(n)}$ такие же, как в предыдущем пункте. Будем принимать гипотезу H_1, если $\sum_{k=1}^{n} b_k^{(n)} x_k < \lambda$, и гипотезу H_2, если $\sum_{k=1}^{n} b_k^{(n)} x_k \geqslant \lambda$. При этом

$$\lambda = c(\varepsilon)\sqrt{d_n} + \frac{1}{2} d_n,$$

где

$$\varepsilon = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c(\varepsilon)} e^{-u^2/2}\, du, \quad d_n = \sum_{k=1}^{n} a_k b_k^{(n)}.$$

Здесь $\alpha_{12} = \varepsilon$, а

$$\alpha_{21} = \frac{1}{\sqrt{2\pi}} \int_{c(\varepsilon)+\sqrt{d_n}}^{\infty} e^{-u^2/2}\, du.$$

Поскольку $d_n \to d = \sum_{k=1}^{\infty} \frac{\alpha_k^2}{\lambda_k}$, а при оптимальном правиле различения с $\alpha_{12} = \varepsilon$

$$\alpha_{21} = \frac{1}{\sqrt{2\pi}} \int_{c(\varepsilon)+\sqrt{d}}^{\infty} e^{-u^2/2}\, du,$$

то $\alpha_{21}^{(n)} \to \alpha_{21}$.

24.4.3. Случай стационарных процессов. Пусть $R(t, s) = r(t-s)$, т. е. $x(t)$ является стационарным процессом при гипотезе H_1. Обозначим через $F(\lambda)$ спектральную функцию процесса. Пусть W_T — пространство функций $g(\lambda)$, представимых в виде

$$g(\lambda) = \int_{-T}^{T} e^{i\lambda t} \varphi(t)\, dt,$$

где φ — комплекснозначные функции на $[-T, T]$, для которых $\int\limits_{-T}^{T} |\varphi(t)|^2 dt < \infty$. Обозначим далее через $W_T(F)$ замыкание W_T в метрике, определяемой равенством

$$\|g\|_F^2 = \int |g(\lambda)|^2 dF(\lambda).$$

Для эквивалентности мер μ_1 и μ_2 необходимо и достаточно, чтобы функция $a(t)$ при $t \in [-T, T]$ имела представление

$$a(t) = \int e^{-i\lambda t} b(\lambda)\, dF(\lambda),$$

где $b(\lambda) \in W_T(F)$. Если это условие выполнено, то плотность $\rho(x(\cdot))$ имеет вид

$$\rho(x(\cdot)) = \exp\left\{\int b(\lambda)\, dy(\lambda) - \frac{1}{2}\int |b(\lambda)|^2 dF(\lambda)\right\}, \tag{4.6}$$

где $y(\lambda)$ — спектральная мера процесса $x(t)$:

$$x(t) = \int e^{i\lambda t} dy(\lambda).$$

Заметим, что стохастический интеграл в (4.6) может быть вычислен как

$$\lim_{n\to\infty} \int\limits_{-T}^{T} \varphi_n(t)\, x(t)\, dt,$$

где φ_n — такая последовательность функций, что $\|b - b_n\|_F \to 0$, если только $b_n(\lambda) = \int\limits_{-T}^{T} e^{i\lambda t} \varphi_n(t)\, dt$.

24.5. Различение гипотез о корреляционной функции гауссовского процесса

Наблюдается траектория гауссовского одномерного процесса $x(t)$ ($t \in [0, T]$), среднее значение процесса равно нулю, относительно корреляционной функции имеется две гипотезы: по гипотезе H_1 она равна $R_1(t, s)$, а по гипотезе H_2 равна $R_2(t, s)$; $R_k(t, s)$ $(k = 1, 2)$ предполагаются непрерывными функциями.

24.5.1. Условия ортогональности. Будем обозначать, как и ранее, через μ_k меру, соответствующую процессу при гипотезе H_k $(k = 1, 2)$. Эти меры можно всегда считать сосредоточенными на пространстве $L_2[0, T]$.

1) Если существует такая последовательность функций $g_n(t) \in L_2[0, T]$, что

$$\lim_{n\to\infty} \int_0^T \int_0^T R_2(t, s) g_n(t) g_n(s)\, dt\, ds \times$$
$$\times \left(\int_0^T \int_0^T R_1(t, s) g_n(t) g_n(s)\, dt\, ds \right)^{-1} = +\infty, \qquad (5.1)$$

то меры μ_1 и μ_2 ортогональны. Безошибочное правило различения гипотез получим, выбрав такую последовательность функций $\psi_n(t)$, что

$$\int_0^T \int_0^T R_1(t, s) \psi_n(t) \psi_n(s)\, dt\, ds = 1,$$
$$\int_0^T \int_0^T R_2(t, s) \psi_n(t) \psi_n(s)\, dt\, ds \geqslant n.$$

Принимаем гипотезу H_1, если

$$\left(\int_0^T \int_0^T R_2(t, s) \psi_n(t) \psi_n(s)\, dt\, ds \right)^{-1/2} \int_0^T x(t) \psi_n(t)\, dt \to 0,$$

и гипотезу H_2, если это условие не выполняется. Аналогично строится правило выбора, если соотношение (5.1) выполняется при перестановке индексов 1 и 2.

2) Предположим, что существует такое $\delta > 0$, что для любой функции $g(t) \not\equiv 0$

$$\delta \int_0^T \int_0^T R_1(t, s) g(t) g(s)\, dt\, ds \leqslant$$
$$\leqslant \int_0^T \int_0^T R_2(t, s) g(t) g(s)\, dt\, ds \leqslant \frac{1}{\delta} \int_0^T \int_0^T R_1(t, s) g(t) g(s)\, dt\, ds.$$

Построим некоторую последовательность функций $\psi_k(t)$ таких, что

$$\int_0^T \int_0^T R_i(t, s) \psi_k(t) \psi_j(s)\, dt\, ds = 0 \quad \text{при} \quad k \neq j.$$

Если

$$\sum_{k=1}^{\infty}\left[\int_0^T\int_0^T R_2(t, s)\,\psi_k(t)\,\psi_k(s)\,dt\,ds \times \right.$$

$$\left.\times\left(\int_0^T\int_0^T R_1(t, s)\,\psi_k(t)\,\psi_k(s)\,dt\,ds\right)^{-1} - 1\right]^2 = +\infty, \qquad (5.2)$$

то меры μ_1 и μ_2 ортогональны. Пусть

$$\int_0^T\int_0^T R_1(t, s)\,\psi_k(t)\,\psi_k(s)\,dt\,ds = 1,$$

$$\int_0^T\int_0^T R_2(t, s)\,\psi_k(t)\,\psi_k(s)\,dt\,ds = 1 + \delta_k.$$

Безошибочное правило различения гипотез получим, выбрав m_n так, чтобы выполнялось неравенство

$$\sum_{k=1}^{m_n}\delta_k^2 > n,$$

и принимая гипотезу H_1, если

$$\lim_{n\to\infty}\left(\sum_{k=1}^{m_n}\delta_k^2\right)^{-1}\sum_{k=1}^{m_n}\left[\left(\int_0^T\psi_k(t)\,x(t)\,dt\right)^2 - 1\right]\delta_k = 0,$$

и гипотезу H_2, если это условие не выполнено.

24.5.2. Условия эквивалентности. Введем вместе с функцией $R_k(t, s)$ функцию $R_k^{1/2}(t, s)$ — это симметрическая функция, интегрируемая с квадратом на $[0, T]\times[0, T]$, удовлетворяющая соотношению

$$R_k(t, s) = \int_0^T R_k^{1/2}(t, u)\,R_k^{1/2}(u, s)\,du.$$

Если $\varphi_n^{(k)}(t)$ и $\lambda_n^{(k)}$ $(k = 1, 2, n = 1, 2, \ldots)$ — соответственно собственные функции и собственные значения интегрального оператора

$$R_k\varphi(t) = \int_0^T R_k(t, s)\,\varphi(s)\,ds,$$

то функция $R_k^{1/2}(t, s)$ может быть построена следующим образом:

$$R_k^{1/2}(t, s) = \sum_{n=1}^{\infty}\sqrt{\lambda_n^{(k)}}\,\varphi_n^{(k)}(t)\,\varphi_n^{(k)}(s) \qquad (5.3)$$

(ряд справа сходится в среднеквадратическом на $[0, T]\times[0, T]$). Для эквивалентности мер μ_1 и μ_2 необходимо и достаточно, чтобы существовала такая интегрируемая с квадратом функция $D(t, s)$, что

$$R_1(t, s) - R_2(t, s) = \int_0^T \int_0^T R_2^{1/2}(t, u)\, D(u, v)\, R_2^{1/2}(v, s)\, du\, dv. \qquad (5.4)$$

Пусть $D(t, s)$ — такая функция. Обозначим через $\theta_k(t)$ собственные функции и через δ_k собственные значения симметричного интегрального оператора с ядром $D(t, s)$

$$\delta_k \theta_k(t) = \int_0^T D(t, s)\, \theta_k(s)\, ds.$$

(Из соотношения (5.4) следует, что $\delta_k > -1$.) Тогда плотность меры μ_2 относительно μ_1 имеет вид

$$\rho(x(\cdot)) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{\infty}\left[\xi_k^2 \frac{\delta_k}{1+\delta_k} - \ln(1+\delta_k)\right]\right\}, \qquad (5.5)$$

где ξ_k определяются по наблюдаемой функции $x(t)$ с помощью равенства

$$\xi_k = \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n^{(2)}}} \int_0^T \varphi_n^{(2)}(s)\, x(s)\, ds \int_0^T \theta_k(t)\, \varphi_n^{(2)}(t)\, dt. \qquad (5.6)$$

Для нахождения распределения величины $\rho(x(\cdot))$ следует учесть, что величины ξ_k при каждой из гипотез являются последовательностью независимых гауссовских величин со средним 0 и дисперсией 1 при гипотезе H_2 и дисперсией $1+\delta_k$ при гипотезе H_1. Поэтому

$$\mathsf{M} \exp\{is \ln \rho(x(\cdot))\} = \prod_{k=1}^{\infty} \frac{(1+\delta_k)^{is/2}}{(1+is\delta_k)^{1/2}}$$

при гипотезе H_1 и

$$\mathsf{M} \exp\{is \ln \rho(x(\cdot))\} = \prod_{k=1}^{\infty} \frac{(1+\delta_k)^{is/2}}{(1+is\delta_k/(1+\delta_k))^{1/2}}$$

при гипотезе H_2.

Как видим, нахождение плотности одной меры относительно другой и распределения этой плотности в случае различных дисперсий приводит к необходимости находить собственные функции и собственные значения интегральных операторов. Иногда можно при вычислении плотности и ее распределения обойтись без вычисления оператора D и величин ξ_k. Рассмотрим уравнение

$$\lambda \int_0^T R_2(t, s)\, \psi(s)\, ds = \int_0^T R_1(t, s)\, \psi(s)\, ds. \qquad (5.7)$$

Это уравнение имеет в случае эквивалентности мер (возможно, обобщенные) решения при не более чем счетном множестве λ. Обозначим эти значения λ через λ_k, а соответствующие решения через ψ_k. Тогда $\lambda_k = 1 + \delta_k$ и плотность $\rho(x(\cdot))$ имеет вид

$$\rho(x(\cdot)) = \exp\left\{-\frac{1}{2}\left[\sum_{k=1}^{\infty}\int_0^T\left(x(t)\,\psi_k(t)\,dt\,\frac{\lambda_k - 1}{\lambda_k} - \ln\lambda_k\right)\right]\right\} \tag{5.8}$$

(предполагается, что ψ_k нормированы таким образом, что $\int_0^T\int_0^T R_2(t, s)\,\psi_k(t)\,\psi_k(s)\,dt\,ds = 1$). Обобщенное решение уравнения

$$\lambda_k\int_0^T R_2(t, s)\,\psi_k(s)\,ds = \int_0^T R_1(t, s)\,\psi_k(s)\,ds$$

определяется последовательностью функции $\psi_k^{(n)}(t)$, для которых

$$\int_0^T\int_0^T R_2(t, s)\,\psi_k^{(n)}(t)\,\psi_k^{(n)}(s)\,dt\,ds = 1,$$

$$\lim_{n\to\infty}\left[\int_0^T\left(\lambda_k\int_0^T R_2(t, s)\,\psi_k^{(n)}(s)\,ds - \int_0^T R_1(t, s)\,\psi_k^{(n)}(s)\,ds\right)^2 dt\right] = 0.$$

24.5.3. Различение гипотез для стационарных процессов. Пусть $x(t)$ является стационарным процессом со средним 0 и корреляционной функцией $R_k(t)$ при гипотезе H_k ($k = 1, 2$). Пусть $F_k(\lambda)$ — спектральная функция процесса при гипотезе H_k:

$$R_k(t) = \int e^{i\lambda t}\,dF_k(\lambda).$$

Обозначим через W_T^2 множество функций вида

$$\psi(\alpha, \beta) = \int_{-T}^{T}\int_{-T}^{T} e^{i\alpha s + i\beta t}\,g(s, t)\,ds\,dt,$$

где $g(s, t)$ ограничена и измерима на $[-T, T]\times[-T, T]$. Пусть далее $W_T^2(F_1)$ — замыкание W_T^2 в норме

$$\|\psi\|_{F_1}^2 = \iint |\psi(\alpha, \beta)|^2\,dF_1(\alpha)\,dF_1(\beta).$$

Для эквивалентности мер μ_1 и μ_2 необходимо и достаточно, чтобы существовала такая функция $b(\alpha, \beta)$ из $W_T^2(F_1)$, чтобы имело место равенство

$$R_2(t-s) - R_1(t-s) = \iint e^{-i\alpha t + i\beta s}\, b(\alpha, \beta)\, dF_1(\alpha)\, dF_2(\beta),$$

при этом плотность имеет вид

$$\rho(x(\cdot)) = \exp\left\{\frac{1}{2}\iint \Phi(\alpha, \beta)\, dy(\alpha)\, \overline{dy(\beta)} + c\right\}, \tag{5.9}$$

где $y(\alpha)$ — спектральная мера, соответствующая процессу $x(t)$:

$$x(t) = \int e^{i\lambda t}\, dy(\lambda),$$

функция $\Phi(\alpha, \beta)$ связана с $b(\alpha, \beta)$ соотношением $b(\alpha, \beta) = \Phi(\alpha, \beta) + \int \Phi(\alpha, \gamma)\, \overline{b(\gamma, \beta)}\, dF_1(\gamma)$, а кратный стохастический интеграл в (5.9) определяется как интеграл по мере ν на $(-\infty, \infty) \times (-\infty, \infty)$, для которой

$$\nu([\alpha, \beta) \times [\gamma, \delta)) = [y(\delta) - y(\gamma)] \times [y(\beta) - y(\alpha)] - \\ - F_1([\alpha, \beta) \cap [\gamma, \delta))$$

$\left(F_1(\Delta) = \int\limits_\Delta dF_1(\lambda)\right)$. Постоянная c в формуле (5.9) определяется из равенства

$$c = -\frac{1}{2}\sum_{k=1}^{\infty} \ln(1 + \lambda_k),$$

где λ_k — собственные значения оператора

$$Vg(\beta) = \int b(\alpha, \beta)\, g(\alpha)\, dF_1(\alpha)$$

в $W_T(F_1)$. Отметим также, что интеграл в (5.9) можно записать в виде

$$\sum_{k,j} c_{kj}\left[\int g_k(\alpha)\, dy(\alpha)\, \overline{\int f_j(\beta)\, dy(\beta)} - \int g_k(\alpha)\, \overline{f_j(\alpha)}\, dF_1(\alpha)\right],$$

если

$$b(\alpha, \beta) = \sum_{k,j} c_{kj} g_k(\alpha)\, \overline{f_j(\beta)}.$$

24.5.4. Различение гипотез о спектральной плотности. Предположим, что спектральные функции $F_k(\lambda)$ имеют спектральные плотности $f_k(\lambda)$. Приведем некоторые достаточные условия абсолютной непрерывности и ортогональности мер в терминах спектральных плотностей.

1. Предположим, что выполнены условия:

а) при некоторых c_1 и c_2

$$c_1 |\varphi_0(\lambda)|^2 \leqslant f_1(\lambda) \leqslant c_2 |\varphi_0(\lambda)|^2,$$

где $\varphi_0(\lambda)$ при некотором $s > 0$ имеет вид

$$\varphi_0(\lambda) = \int_{-s}^{s} e^{i\lambda t} g(t)\, dt \tag{5.10}$$

с некоторой интегрируемой с квадратом функцией $g(t)$;

б) $\displaystyle\int \left[\frac{f_2(\lambda) - f_1(\lambda)}{f_1(\lambda)}\right]^2 d\lambda < \infty$;

тогда, каково бы ни было $T > 0$, меры μ_1 и μ_2, отвечающие $x(t)$ ($t \in [0, T]$) при гипотезах H_1 и H_2, эквивалентны.

2. Предположим, что для достаточно больших $|\lambda|$ при некоторых положительных c_1 и c_2 выполнено соотношение

$$c_1 \leqslant |\lambda|^\alpha f_1(\lambda) \leqslant c_2,$$

где $\alpha > 1$. Если

$$\int \frac{[f_2(\lambda) - f_1(\lambda)]^2}{1 + \lambda^{2\alpha}}\, d\lambda < \infty,$$

то меры μ_1 и μ_2, отвечающие $x(t)$ ($t \in [0, T]$) при гипотезах H_1, H_2, эквивалентны при всех $T > 0$.

3. Пусть существует целая аналитическая функция $\varphi_0(\lambda)$ экспоненциального типа не выше $s < T$, для которой при некоторых $0 < c_1 < c_2$ выполнено неравенство

$$c_1 \leqslant |\varphi_0(\lambda)|^2 f_1(\lambda) \leqslant c_2.$$

Тогда если

$$\iint \frac{\sin^2 (T - s)(\alpha - \beta)}{(\alpha - \beta)^2} \cdot \frac{f_2(\alpha) - f_1(\alpha)}{f_1(\alpha)} \cdot \frac{f_2(\beta) - f_1(\beta)}{f_1(\beta)}\, d\alpha\, d\beta = +\infty, \tag{5.11}$$

то меры μ_1 и μ_2 ортогональны.

4. Если $f_1(\lambda)$ и $f_2(\lambda)$ — дробно рациональные плотности, то необходимым и достаточным условием эквивалентности является условие

$$\lim_{\lambda \to \infty} \frac{f_2(\lambda)}{f_1(\lambda)} = 1.$$

Если $\displaystyle\lim_{\lambda \to \infty} \frac{f_2(\lambda)}{f_1(\lambda)} \neq 1$, то правило, достоверно различающее гипотезы, заключается в следующем. Пусть m таково, что существует конечный отличный от нуля предел

$$\lim_{\lambda \to \infty} \lambda^{2m} [f_1(\lambda) + f_2(\lambda)] = a.$$

Тогда существуют также пределы

$$\lim_{\lambda \to \infty} \lambda^{2m} f_1(\lambda) = a_1, \quad \lim_{\lambda \to \infty} \lambda^{2m} f_2(\lambda) = a_2, \tag{5.12}$$

при этом $a_1 + a_2 = a \neq 0$, $a_1 \neq a_2$. Из (5.12) следует существование $m-1$-й производной у наблюдаемой траектории $\tilde{x}(t) = x^{(m-1)}(t)$. С вероятностью 1 существует предел

$$\lim_{n\to\infty} \sum_{k=1}^{2^n} \left[\tilde{x}\left(\frac{k}{2^n}T\right) - \tilde{x}\left(\frac{k-1}{2^n}T\right)\right]^2 = a_i, \tag{5.13}$$

если верна гипотеза H_i. Вычисляя левую часть (5.13) и сравнивая ее со значениями a_i, можем безошибочно выбрать истинную гипотезу.

24.6. Оценки параметров распределений для случайных процессов

24.6.1. Постановка задачи. Предположим, что наблюдаемой траектории процесса $x(t)$ $(t \in [0, T])$ отвечает мера μ_θ на пространстве функций $F_{[0,\,T]}$. Параметр θ, который нужно оценить по наблюдению, меняется в некотором параметрическом множестве Θ. Особенность задач статистики случайных процессов заключается в том, что этот параметр, как правило, меняется в бесконечномерном пространстве (например, в качестве параметра может выступать неизвестное среднее значение, которое в принципе может быть любой функцией из $F_{[0,\,T]}$). Предположим, что существует мера ν на $F_{[0,\,T]}$, относительно которой все меры μ_θ абсолютно непрерывны и

$$\frac{d\mu_\theta}{d\nu}(x) = \rho(\theta, x). \tag{6.1}$$

В этом случае семейство мер $\{\mu_\theta, \theta \in \Theta\}$ называется *регулярным*. Предполагают, что Θ — линейное многообразие, $\rho(\theta, x)$ — достаточно регулярная функция. Тогда оценку параметра можно искать, например, методом максимального правдоподобия. Заметим, что для процессов с независимыми приращениями, диффузионных, а также гауссовских процессов условия регулярности и вид плотности можно извлечь из результатов § 24.2—24.5.

Более интересным (и специфичным именно для статистики случайных процессов) представляется случай попарно сингулярного семейства мер $\{\mu_\theta, \theta \in \Theta\}$. Тогда при $\theta_1 \neq \theta_2$ меры μ_{θ_1} и μ_{θ_2} ортогональны. Поэтому естественно ожидать, что параметр θ можно определить достоверно по единственному наблюдению. Под оценкой параметра θ будем понимать функцию $\hat{\theta}(x)$, определенную на $F_{[0,T]}$ со значениями в Θ. Пусть Θ — полное сепарабельное метрическое пространство (или борелевское множество в таком пространстве). Предположим, что $\hat{\theta}(x)$ измерима относительно σ-алгебры $\mathfrak{F}_{[0,\,T]}$, порожденной в $F_{[0,T]}$ цилиндрическими множествами, и относительно $\mathfrak{B}$ — σ-алгебры борелевских множеств в Θ, т. е. что прообраз всякого борелевского множества в Θ измерим в $F_{[0,\,T]}$. Оценка параметра θ называется *состоятельной*, если

$$\mu_\theta \{x\colon\ \hat{\theta}(x) = \theta\} = 1 \quad \forall \theta \in \Theta. \tag{6.2}$$

Существование состоятельной оценки параметра θ дает возможность безошибочного определения параметра. Примеры показывают, что бывают такие семейства попарно ортогональных мер, для которых не существует состоятельной оценки. Поэтому вопрос о существовании состоятельных оценок, как и способы их построения в случае существования, представляет определенный интерес. Ниже приведены некоторые методы построения состоятельных оценок.

24.6.2. Проекционные методы. Предположим, что меры μ_θ таковы, что существуют среднее значение процесса $a_\theta(t)$ и корреляционная функция $R_\theta(t, s)$. Пусть далее существуют два линейных многообразия L_1 и L_2 в $F_{[0, T]}$, имеющих нулевое пересечение, такие, что $a_\theta(\cdot) \in L_1$ для всех θ, и если θ — истинное значение параметра, то $x(t) - a_\theta(t)$ с вероятностью 1 принадлежит L_2. Последнее означает, что если $\varphi_k(\theta, t)$ — собственные функции интегрального оператора с ядром $R_\theta(t, s)$, а $\lambda_k(\theta)$ — соответствующие собственные значения, т. е.

$$\lambda_k(\theta)\varphi_k(\theta, t) = \int_0^T R_\theta(t, s)\varphi_k(\theta, s)\,ds,$$

то

$$\sum_{k=1}^{\infty}\left[\int_0^T (x(s) - a_\theta(s))\varphi_k(\theta, s)\,ds\right]\varphi_k(\theta, t) \in L_2.$$

Пусть $a_\theta(\cdot)$ — взаимно однозначное отображение Θ в L_1, $L = L_1 + L_2$ и P — оператор проектирования L на L_1: если $z(t) = y_1(t) + y_2(t)$, где $y_1(t) \in L_1$ (такое представление единственно для всех $z(t) \in L$), то $Pz(t) = y_1(t)$. Тогда при наших предположениях

$$\mu_\theta(\{x(\cdot):\ Px(t) = a_\theta(t)\}) = 1.$$

Зная $a_\theta(t)$, в силу взаимной однозначности отображения $a_\theta(t)$ можем определить по $a_\theta(t)$ параметр θ. Таким образом, основная задача при использовании проекционного метода — построение проекционного оператора P.

Пример. Семейство мер μ_θ — семейство гауссовских мер со средним значением

$$a_\theta(t) = \int_0^T B(t, s)\,\theta(s)\,ds, \tag{6.3}$$

$\theta(s) \in L_2[0, T]$, $\int_0^T\int_0^T B^2(t, s)\,dt\,ds < \infty$, корреляционный оператор $R_\theta(t, s) = R(t, s)$ от θ не зависит.

Пусть $\varphi_k(t)$ и λ_k — соответственно собственные функции и значения интегрального оператора с ядром $R(t, s)$. Обозначим через $R^{1/2}$

линейное многообразие функций $z(t)$ из $L_2[0, T]$ вида $z(t) = \sum_{k=1}^{\infty} \alpha_k \sqrt{\lambda_k}\, \varphi_k(t)$, где $\sum_k \alpha_k^2 < \infty$, а через B линейное многообразие функций, представимых в виде правой части (6.3)

Предположим, что $B \cap R^{1/2} = \{0\}$. Это означает, что для всех $\theta(s)$, для которых $\int_0^T \theta^2(s)\, ds > 0$,

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(\int_0^T \int_0^T B(t, s)\, \theta(s)\, ds\, \varphi_k(t)\, dt \right)^2 = +\infty.$$

Введем в $B + R^{1/2}$ скалярное произведение: если

$$y(t) = \sum_{k=1}^{\infty} \alpha_k \sqrt{\lambda_k}\, \varphi_k(t) + \int_0^T B(t, s)\, \theta(s)\, ds,$$

то

$$(y(t), y(t))_1 = \sum_{k=1}^{\infty} \alpha_k^2 + \int_0^T \theta^2(s)\, ds.$$

Предположим, что $B(t, s)$ — симметрическая функция и $\{\psi_k(t), \mu_k\}$ — соответствующая интегральному оператору с ядром $B(t, s)$ последовательность собственных функций и собственных значений. Пусть

$$y(t) = \sum_{k=1}^{\infty} \alpha_k \sqrt{\lambda_k}\, \varphi_k(t) + \sum_{k=1}^{\infty} \beta_k \mu_k \psi_k(t), \tag{6.4}$$

где $\sum \alpha_k^2 + \sum \beta_k^2 < \infty$. Покажем, как можно построить оператор P, ставящий в соответствие функции $y(t)$ вида (6.4) выражение

$$\sum_{k=1}^{\infty} \beta_k \mu_k \psi_k(t).$$

Обозначим через $\beta_k^{(n)}$ такие числа, что выражение

$$\sum_{k=1}^{n} \frac{1}{\lambda_k} \left\{ \int_0^T \left[y(t) - \sum_{m=1}^{\infty} \gamma_m \mu_m \psi_m(t) \right] \varphi_k(t)\, dt \right\}^2$$

принимает минимальное значение при $\gamma_k = \beta_k^{(n)}$. Тогда $\beta_k = \lim_{n\to\infty} \beta_k^{(n)}$ и

$$Py(t) = \sum_{k=1} \lim_{n\to\infty} \beta_k^{(n)} \psi_k(t).$$

Таким образом,

$$a_\theta(t) = Px(t),$$

где $x(t)$ — наблюденная траектория.

24.6.3. Метод минимизации квадратического функционала. Пусть семейство $\{\mu_\theta, \theta \in \Theta\}$ такое же, как и в предыдущем пункте. Отметим, что в рассмотренном выше примере оценка среднего строилась путем минимизации некоторой последовательности квадратических функционалов. В общем случае можно рассмотреть некоторую последовательность квадратических функционалов

$$K_n(x(\cdot)) = \int_0^T\int_0^T K_n(t, s)\, x(t)\, x(s)\, dt\, ds$$

и оценку среднего $a_\theta(t)$ искать в виде предела последовательности функций $a_\theta^{(n)}(t)$, дающих минимум выражению

$$\int_0^T\int_0^T K_n(t, s)\,[x(t) - z(t)]\,[x(s) - z(s)]\, dt\, ds$$

при $z(\cdot)$, меняющемся во множестве M возможных средних $\{a_\theta(\cdot), \theta \in \Theta\}$. Для состоятельности этой оценки (в частности, для существования предела $a_\theta^{(n)}(t)$) достаточно выполнения следующих условий:

1) при $\theta_1 \neq \theta_2$

$$\lim_{n\to\infty}\int_0^T\int_0^T K_n(t, s)\,[a_{\theta_1}(t) - a_{\theta_2}(t)]\,[a_{\theta_1}(s) - a_{\theta_2}(s)]\, dt\, ds = +\infty;$$

2) существует такая метрика ρ в Θ, что если $\xi(t) = x(t) - a_\theta(t)$ (θ — истинное значение параметра), то с вероятностью 1

$$\lim_{n\to\infty}\frac{\displaystyle\int_0^T\int_0^T K_n(t, s)\,\xi(t)\,[a_{\theta_1}(s) - a_\theta(s)]\, dt\, ds}{\displaystyle\int_0^T\int_0^T K_n(t, s)\,[a_{\theta_1}(t) - a_\theta(t)]\,[a_{\theta_1}(s) - a_\theta(s)]\, dt\, ds} = 0$$

равномерно по θ_1, для которых $\rho(\theta, \theta_1) \geqslant \varepsilon$, каково бы ни было $\varepsilon > 0$. При выполнении этих условий, если θ_n определяется из равенства $a_\theta^{(n)}(t) = a_{\theta_n}(t)$, то θ_n сходится по вероятности (в метрике ρ) к θ.

24.6.4. Метод максимального правдоподобия. Рассмотрим два варианта этого метода.

1. Пусть для любых $t_1, \ldots, t_n$ из $[0, T]$ существует совместная всюду положительная плотность распределения величин $x(t_1), \ldots, x(t_n)$

$$p_\theta(t_1, \ldots, t_n, x_1, \ldots, x_n),$$

если θ — истинное значение параметра. Выбрав некоторое значение $\theta_0 \in \Theta$, введем функции

$$f_n(\theta, x(\cdot)) = \frac{p_\theta(t_1^{(n)}, \ldots, t_n^{(n)}, x(t_1^{(n)}), \ldots, x(t_n^{(n)}))}{p_{\theta_0}(t_1^{(n)}, \ldots, t_n^{(n)}, x(t_1^{(n)}), \ldots, x(t_n^{(n)}))} \tag{6.5}$$

(f_n — некоторый функционал от наблюдаемой траектории), где $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = T$, $\max\limits_k (t_k^{(n)} - t_{k-1}^{(n)}) \to 0$ при $n \to \infty$. Если процесс $x(t)$ стохастически непрерывен каково бы ни было истинное значение параметра, то в том случае, когда θ_0 — истинное значение параметра, $\lim\limits_{n \to \infty} f_n(\theta, x(\cdot)) = 0$ с вероятностью 1 для всех $\theta \neq \theta_0$, а при $\theta = \theta_0$ этот предел равен 1.

Пусть $\widehat{\theta}_n$ — значение, при котором $f_n(\theta, x(\cdot))$ достигает максимума (предполагается, что $f_n(\theta, x(\cdot))$ непрерывна по θ, а само θ меняется в некотором компакте). Естественно выбирать в качестве оценки величину $\widehat{\theta}_n$. Состоятельность этой оценки нужно исследовать в каждом отдельном случае.

2. Предположим, что $x(t) \in L_2[0, T]$. Выберем некоторую ортонормированную систему функций $\{\varphi_k(t),\ k = 1, 2, \ldots\}$ в $L_2[0, T]$ и положим

$$x_k = \int_0^T \varphi_k(t)\, x(t)\, dt.$$

Пусть $p_\theta^{(n)}(y_1, \ldots, y_n)$ — совместная плотность распределения величин $x_1, \ldots, x_n$, если θ — истинное значение параметра,

$$f_n(\theta, x(\cdot)) = \frac{p_\theta^{(n)}(x_1, \ldots, x_n)}{p_{\theta_0}^{(n)}(x_1, \ldots, x_n)}. \tag{6.6}$$

Оценка $\widehat{\theta}_n$ ищется как точка, где последняя функция достигает максимума.

24.6.5. Метод Байеса. Пусть выполнены условия предыдущего пункта. Обозначим через $f_n(\theta, x(\cdot))$ функцию, определенную равенством (6.5) или (6.6). Будем предполагать, что Θ — выпуклое открытое множество в линейном нормированном пространстве, зададим на Θ борелевскую меру ν такую, что мера всякого открытого множества положительна. В качестве оценки Байеса параметра θ берут последовательность оценок

$$\hat{\theta}_n = \int \theta f_n(\theta, x(\cdot))\, \nu(d\theta) \left(\int f_n(\theta, x(\cdot))\, \nu(d\theta) \right)^{-1}.$$

Состоятельность этой оценки нужно исследовать в каждом конкретном случае.

Литература: [17, 25, 37, 52, 59].

Глава 25. СТАТИСТИКА СТАЦИОНАРНЫХ В ШИРОКОМ СМЫСЛЕ СЛУЧАЙНЫХ ПРОЦЕССОВ

25.1. Свойства статистических оценок характеристик стационарных процессов

25.1.1. Задачи статистики стационарных процессов. Предположим, что наблюдается некоторый случайный процесс $\{\xi(t),\ t \in T\}$. Априорные соображения либо предварительно проведенный статистический тест дают основания считать, что процесс $\{\xi(t),\ t \in T\}$ имеет вид:

1) $\xi(t) = \xi_0(t), \qquad t \in T,$

2) $\xi(t) = m + \xi_0(t), \quad t \in T,$

3) $\xi(t) = \sum_{k=1}^{r} \theta_k a_k(t) + \xi_0(t),$

где $\xi_0(t)$ — стационарный в широком смысле случайный процесс с нулевым математическим ожиданием.

Пусть $x(t)$ $(t \in T_0)$ — траектория процесса $\xi(t)$, наблюдавшаяся в течение времени T_0, где T_0 может быть отрезком на временной оси $T_0 = [a, b]$ в случае непрерывного времени либо последовательностью моментов наблюдения $T_0 = \{t_k,\ k = 1, \ldots, n\}$.

В случае 1) требуется на основе наблюдения $x(t)$ $(t \in T_0)$ оценить спектральную функцию или спектральную плотность процесса $\xi(t)$. В случае 2) спектральная функция процесса $\xi_0(t)$ предполагается известной и требуется на основе наблюдений $x(t)$ $(t \in T_0)$ оценить неизвестное среднее m. В случае 3) спектральная функция процесса $\xi_0(t)$ также предполагается известной и требуется на основе наблюдений $x(t)$ $(t \in T_0)$ оценить неизвестные параметры θ_1, $\theta_2, \ldots, \theta_r$ *регрессии* $A(t) = \sum_{k=1}^{r} \theta_k a_k(t)$, где функции $a_k(t)$ $(k = 1, \ldots, n)$ предполагаются известными.

Это основные задачи статистики стационарных процессов. Возможны варианты, например предварительная оценка спектральной функции для последующей оценки среднего или параметров регрессии.

25.1.2. Свойства оценок. Пусть $\hat{\mu}$ — статистика, предназначенная для решения какой-либо из перечисленных задач, представляющая собой функционал от наблюдаемой траектории $x(t)$, $t \in T_0$: $\hat{\mu} = g(x(t),\ t \in T_0)$.

Среди множества статистик $\hat{\mu}$ естественно выбрать те, которые обладают наиболее желательными свойствами:

1) *линейность* (функционал $g(\cdot)$ должен быть линейным);

2) *несмещенность* (если h — оцениваемая характеристика процесса $\{\xi(t),\ t \in T\}$ и $\hat{\mu}$ — статистика, предназначенная для оценки h, требуется, чтобы $M\hat{\mu} = h$):

3) *состоятельность* (статистика $\hat{\mu}$ должна сходиться по вероятности к h при увеличении интервала наблюдения);

4) *эффективность* (статистика $\hat{\mu}$ должна обладать минимальной дисперсией среди статистик заданного класса).

Иногда приходится ограничиваться более слабыми требованиями, чем требования несмещенности и эффективности:

2′) *асимптотической несмещенностью* ($M\hat{\mu} \to h$ при неограниченном возрастании интервала наблюдения);

4′) *асимптотической эффективностью* (статистика $\hat{\mu}$ должна обладать асимптотически минимальной дисперсией среди статистик заданного класса при неограниченном возрастании интервала наблюдения).

Для проверки состоятельности оценки достаточно убедиться в том, что ее дисперсия стремится к нулю.

25.2. Оценки неизвестного среднего

25.2.1. Временно́е среднее (среднеарифметическая оценка). Пусть $x(t)$ $(t \in T_0)$ — траектория процесса $\xi(t) = m + \xi_0(t)$ $(t \in T_0)$, где $\xi_0(t)$ — стационарный в широком смысле процесс с нулевым средним и ковариационной функцией $B(t)$. Предполагается, что время t непрерывно и $T_0 = [a, b]$.

Среди линейных несмещенных оценок среднего стационарного процесса наиболее простой вид имеет статистика $\overline{m}$, называемая временны́м средним или среднеарифметической оценкой:

$$\overline{m} = \frac{1}{b-a} \int_a^b x(t)\, dt. \tag{2.1}$$

Если рассматриваемый стационарный процесс эргодичен, то среднеарифметическая оценка $\overline{m}$ является состоятельной. Оценкой среднего m по методу наименьших квадратов является среднеарифметическая оценка $\overline{m}$.

Определим класс M_g линейных несмещенных оценок неизвестного среднего:

$$M_g = \left\{ \hat{\mu}: \hat{\mu} = \int_a^b g(t)\, x(t)\, dt \right\},$$

где функции $g(t)$ принадлежат классу равностепенно непрерывных и равномерно ограниченных функций на $[a, b]$ таких, что $\int_a^b g(t)\, dt = 1$, а $x(t)$ $(t \in [a, b])$ — траектория процесса $\{\xi(t),\ t \in T\}$, имеющего непрерывную в нуле спектральную плотность.

Теорема 1. *Среднеарифметическая оценка $\overline{m}$ обладает асимптотически минимальной дисперсией в классе M_g.*

Таким образом, среди оценок $\hat{\mu} \in M_g$ при $b - a \to \infty$ не существует оценок более эффективных, чем среднеарифметическая.

Значительно более широкий, чем M_g, класс линейных несмещенных оценок можно получить, рассматривая «взвешенные» оценки вида

$$\hat{\mu} = \sum c_k^{(n)} x(t_k), \quad a = t_1 < t_2 < \ldots < t_n = b, \tag{2.2}$$

где $\sum_{k=1}^{n} c_k^{(n)} = 1$ для любого $n \geqslant 1$.

Пусть M_0 — замыкание класса оценок вида (2.2) в среднеквадратичном.

Теорема 2. *В M_0 существует единственная с точностью до эквивалентности оценка $\hat{m}$ неизвестного среднего, имеющая минимальную дисперсию, причем*

$$\mathsf{M}\hat{m}x(t) \equiv C, \quad t \in [a, b], \tag{2.3}$$

где $C = \inf\limits_{\hat{\mu} \in M_0} \mathsf{D}\hat{\mu}$ — *дисперсия оценки* $\hat{m}$.

25.2.2. Вычисление оценок среднего на основе прогноза. Один из методов построения эффективных оценок среднего для эргодических стационарных процессов основан на анализе прогноза, построенного по наблюдениям $x(t)$ $(t \in T_0 = [a, b])$.

Пусть $\hat{x}(t)$ для $t \notin [a, b]$ есть наилучший линейный несмещенный прогноз значений $x(t)$ $(t \in [a, b])$, т. е. $\mathsf{M}\hat{x}(t) \equiv m$, $\mathsf{M}\,|\hat{x}(t) - \xi(t)|^2 = \min$ для всех $t \notin [a, b]$.

Введем статистику $\hat{\mu}_A$, определяемую как временно́е среднее, построенное по прогнозу $\hat{x}(t)$: для $[a, b] \subset [-A, A]$

$$\hat{\mu}_A = \frac{1}{2A}\left[\int_{-A}^{a} \hat{x}(t)\,dt + \int_{a}^{b} x(t)\,dt + \int_{b}^{A} \hat{x}(t)\,dt\right]. \tag{2.4}$$

Для линейной несмещенной оценки $\hat{m}$ с минимальной дисперсией имеет место следующая теорема.

Теорема 3.

$$\hat{m} = \lim_{A \to \infty} \hat{\mu}_A.$$

Если процесс $\xi_0(t)$ таков, что $\int\limits_{-\infty}^{\infty} B(t)\,dt < \infty$, то результат теоремы 3 может быть записан в более удобном для вычислений виде.

Пусть $\hat{x}_0(t)$ — наилучший линейный прогноз значений $x(t)$ $(t \in [a, b])$, сделанный в предположении, что $m = 0$ (если $t' \in [a, b]$, то $\hat{x}_0(t) = x(t)$).

Теорема 4. *Если* $\int\limits_{-\infty}^{\infty} B(t)\,dt < \infty$, *то существует константа* d *такая, что*

$$\hat{m} = d\int_{-\infty}^{\infty} \hat{x}_0(t)\,dt,$$

и d определяется единственным образом из условия несмещенности: $\mathsf{M}\hat{m} = m$.

Пример 1. Пусть ковариационная функция процесса $\xi(t) = m + \xi_0(t)$ равна $B(t) = e^{-\alpha|t|}$ и m неизвестно.

Наблюдается траектория $x(t)$ $(t \in [a, b])$. В предположении, что $m = 0$, находим наилучший линейный прогноз:

$$\hat{x}_0(b+\tau) = e^{-\alpha\tau}x(b), \quad \tau > 0,$$
$$\hat{x}_0(t) = x(t), \quad t \in [a, b],$$
$$\hat{x}_0(a-\tau) = e^{-\alpha\tau}x(a), \quad \tau > 0.$$

Следовательно, $\hat{m} = d\int_{-\infty}^{\infty} \hat{x}_0(t)\,dt = d\left[\int_{-\infty}^{a} \hat{x}_0(t)\,dt + \int_a^b x(t)\,dt + \right.$

$$\left. + \int_b^{\infty} \hat{x}_0(t)\,dt\right] = d\left[\frac{x(a)}{\alpha} + \int_a^b x(t)\,dt + \frac{x(b)}{\alpha}\right].$$

Условие несмещенности дает

$$m = \mathsf{M}\hat{m} = d\mathsf{M}\left[\frac{x(a)}{\alpha} + \int_a^b x(t)\,dt + \frac{x(b)}{\alpha}\right] =$$
$$= d\left[\frac{m}{\alpha} + m(b-a) + \frac{m}{\alpha}\right],$$

откуда

$$d = \frac{\alpha}{2 + \alpha(b-a)}.$$

Таким образом,

$$\hat{m} = \left(x(a) + \int_a^b x(t)\,dt + x(b)\right)\Big/(2 + \alpha(b-a)),$$
$$\mathsf{D}\hat{m} = \frac{2}{2 + \alpha(b-a)}. \tag{2.5}$$

Для сравнения укажем, что дисперсия среднеарифметической оценки $\bar{m}$

$$\mathsf{D}\bar{m} = \frac{2\left[e^{-\alpha(b-a)} - 1 + \alpha(b-a)\right]}{\alpha^2(b-a)^2}, \tag{2.6}$$

причем $\mathsf{D}\bar{m} > \mathsf{D}\hat{m}$, $\mathsf{D}\hat{m}/\mathsf{D}\bar{m} \to 1$, $b - a \to \infty$.

25.2.3. Уравнения типа Винера — Хопфа. В ряде случаев явную формулу для несмещенной линейной оценки $\hat{m}$ неизвестного среднего m можно получить, исходя из формального представления

$$\hat{m} = \int_a^b x(t)\,dG(t). \tag{2.7}$$

Функция $G(t)$ должна удовлетворять условию несмещенности $\int\limits_a^b dG(t) = 1$ и быть решением *интегрального уравнения типа Винера — Хопфа*

$$\int\limits_a^b B(t-s)\,dG(s) \equiv C, \quad t \in [a, b], \tag{2.8}$$

получающегося из (2.3) для данного вида оценок.

Для процессов с дробно-рациональной спектральной плотностью уравнение (2.8) всегда имеет решение, которое содержит линейные комбинации дельта-функций Дирака и их производных. Такое решение может быть найдено явно (см. п. 25.3.3).

Пример 2. В случае марковского стационарного процесса уравнение

$$\int\limits_a^b e^{-\alpha(t-s)}\,dG(s) \equiv C, \quad t \in [a, b],$$

имеет решение

$$G(t) = \frac{C}{2}\left[u(t-a) + \alpha t + u(b-t)\right]$$

где

$$u(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geqslant 1. \end{cases}$$

Оценка $\hat{m} = \int\limits_a^b x(t)\,dG(t)$ естественно совпадает с приведенной в примере 1.

Пример 3. Пусть процесс $\xi_0(t)$ имеет дробно-рациональную спектральную плотность вида

$$f(\lambda) = \frac{1}{|Q(i\lambda)|^2},$$

где $Q(z) = \sum\limits_{k=0}^{q} q_k z^k$, q_k — действительные числа (процесс авторегрессии порядка q).

В этом случае решение уравнения (2.8) имеет вид

$$\frac{dG(t)}{dt} = \frac{Cq_0}{2\pi}\{q_0 + q_1[\delta(b-t) + \delta(t-a)] + q_2[\delta'(b-t) - \\ - \delta'(t-a)] + \ldots + q_q[\delta^{(q-1)}(b-t) + (-1)^{q-1}\delta^{(q-1)}(t-a)]\},$$

где $\delta(t)$ — дельта-функция Дирака, $\delta^{(k)}(t)$ — ее k-я производная. Следовательно,

$$\hat{m} = \frac{Cq_0}{2\pi}\left\{q_0 \int_a^b x(t)\,dt + q_1[x(b) + x(a)] + \right.$$

$$\left. + q_2[x'(b) - x'(a)] + \ldots + q_n\left[x^{(q-1)}(b) + (-1)^{q-1}x^{(q-1)}(a)\right]\right\},$$

где

$$C = \mathrm{D}\hat{m} = \frac{2\pi}{q_0[2q_1 + q_0(b-a)]}.$$

25.2.4. Метод Яглома. Пусть процесс $\xi_0(t)$ имеет дробно-рациональную спектральную плотность

$$f(\lambda) = \left|\frac{P(i\lambda)}{Q(i\lambda)}\right|^2,$$

где $P(z)$ — многочлен степени p, $Q(z)$ — многочлен степени q, $p < q$, и пусть $x(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\,d\zeta(\lambda)$ — спектральное представление траектории $x(t)$ $(t \in [a, b])$. Метод Яглома состоит в представлении наилучшей линейной несмещенной оценки $\hat{m}$ в виде

$$\hat{m} = \int_{-\infty}^{\infty} \varphi_{(a,b)}(\lambda)\,d\zeta(\lambda) \tag{2.9}$$

и указании условий, однозначно определяющих *спектральную характеристику* $\varphi_{(a,b)}(\lambda)$ оценки $\hat{m}$ и позволяющих эффективно ее вычислять.

Теорема Яглома. *Для процессов с дробно-рациональными спектральными плотностями спектральная характеристика* $\varphi_{(a,b)}(\lambda)$ *в* (2.9) *однозначно определяется условиями:*

а) $\varphi_{(a,b)}(\lambda)$ — *целая функция такая, что*

$$\int_{-\infty}^{\infty} |\varphi_{(a,b)}(\lambda)|^2 f(\lambda)\,d\lambda < \infty;$$

б) $\varphi_{(a,b)}(\lambda)$ *представима в виде*

$$\varphi_{(a,b)}(\lambda) = e^{i\lambda a}\frac{w_a(\lambda)\overline{Q(i\lambda)}}{\lambda|P(i\lambda)|^2} + e^{i\lambda b}\frac{w_b(\lambda)Q(i\lambda)}{\lambda|P(i\lambda)|^2}, \tag{2.10}$$

где $\varphi_a(\lambda) = \dfrac{w_a(\lambda)\overline{Q(i\lambda)}}{\lambda|P(i\lambda)|^2}$ — *функция, аналитическая в верхней полуплоскости,* $w_a(0) \neq 0$; $\varphi_b(\lambda) = \dfrac{w_b(\lambda)Q(i\lambda)}{\lambda|P(i\lambda)|^2}$ — *функция, аналитическая в нижней полуплоскости,* $w_b(0) \neq 0$;

в) $\lim\limits_{\lambda \to 0} \varphi_{(a,b)}(\lambda) = 1$ (*условие несмещенности*).

Функции $w_a(\lambda)$ и $w_b(\lambda)$ в силу второй части условия а) могут быть лишь многочленами степени не выше p. Для дисперсии оценки $\hat{m}$ метод Яглома дает формулу

$$\mathrm{D}\hat{m} = \int_{-\infty}^{\infty} |\varphi_{(a,\ b)}(\lambda)|^2 f(\lambda)d\lambda = 2\pi\left|\frac{w_a(0)}{q_0}\right| = 2\pi\left|\frac{w_b(0)}{q_0}\right|. \quad (2.11)$$

Пример 4. Пусть процесс $\xi_0(t)$ имеет спектральную плотность $f(\lambda) = B\dfrac{\lambda^2+\alpha^2}{\lambda^4+\alpha^4}$ (смешанная модель авторегрессии и скользящего суммирования). По наблюдению $x(t)$ $(t \in [a, b])$ требуется дать наилучшую несмещенную оценку среднего m процесса $\xi(t) = m + \xi_0(t)$.

В данном случае условие б) теоремы Яглома дает

$$\varphi_{a,\ b}(\lambda) = e^{ia\lambda}\frac{w_a(\lambda)\,[\lambda^2 + i\sqrt{2}\,\alpha\lambda - \alpha^2]}{\lambda(\lambda^2+\alpha^2)} + e^{ib\lambda}\frac{w_b(\lambda)\,[\lambda^2 - i\sqrt{2}\,\alpha\lambda - \alpha^2]}{\lambda(\lambda^2+\alpha^2)}, \quad (2.12)$$

где $w_a(\lambda)$ и $w_b(\lambda)$ — некоторые многочлены не выше первого порядка, коэффициенты которых должны быть подобраны так, чтобы удовлетворялись условия а)—в).

Правую часть (2.12) удобнее представить в виде

$$\varphi_{a,\ b}(\lambda) = e^{ia\lambda}\left[c_a^0 + \frac{c_a^1}{\lambda} + \frac{c_a^2}{\lambda - i\alpha} + \frac{c_a^3}{\lambda + i\alpha}\right] + e^{ib\lambda}\left[c_b^0 + \frac{c_b^1}{\lambda} + \frac{c_b^2}{\lambda - i\alpha} + \frac{c_b^3}{\lambda + i\alpha}\right], \quad (2.13)$$

где коэффициенты c_a^k, c_b^k $(k = 0, \ldots, 3)$ подлежат определению.

Из условия а) следует, что

$$c_a^1 + c_b^1 = 0, \quad e^{-a\alpha}c_a^2 + e^{-b\alpha}c_b^2 = 0, \quad e^{a\alpha}c_a^3 + e^{b\alpha}c_b^3 = 0. \quad (2.14)$$

Условие несмещенности в) дает:

$$c_a^0 + c_b^0 + i\left(ac_a^1 + bc_b^1\right) + i\frac{c_a^2 + c_b^2}{\alpha} - i\,\frac{c_a^3 + c_b^3}{\alpha} = 1. \quad (2.15)$$

Из условия б) следует, что коэффициент при $e^{ia\lambda}$ в правой части (2.13) обращается в нуль при $\lambda = \dfrac{\pm 1 - i}{\sqrt{2}}\alpha$, а коэффициент при $e^{ib\lambda}$ обращается в нуль при $\lambda = \dfrac{\pm 1 + i}{\sqrt{2}}\alpha$, что дает недостающие к (2.14), (2.15) четыре уравнения для определения c_a^k, c_b^k $(k = 0, \ldots$

..., 3). Решение соответствующих уравнений дает

$$\hat{m} = d\left\{\left(\sqrt{2}\,\mathrm{sh}\,\frac{\alpha T_0}{2} + \mathrm{ch}\,\frac{\alpha T_0}{2}\right)[x(a) + x(b)] + \right.$$

$$+ \alpha\left(\mathrm{sh}\,\frac{\alpha T_0}{2} + \sqrt{2}\,\mathrm{ch}\,\frac{\alpha T_0}{2}\right)\int_a^b x(t)\,dt -$$

$$\left. - \sqrt{2}\,\alpha\int_a^b \mathrm{ch}\left(\frac{a+b}{2} - t\right)x(t)\,dt\right\},$$

где $T_0 = b - a$,

$$d = \left[(2 + \sqrt{2}\,\alpha T_0)\,\mathrm{ch}\,\frac{\alpha T_0}{2} + \alpha T_0\,\mathrm{sh}\,\frac{\alpha T_0}{2}\right]^{-1},$$

$$\mathsf{D}\hat{m} = 2\pi B\left(\mathrm{sh}\,\frac{\alpha T_0}{2} + \sqrt{2}\,\mathrm{ch}\,\frac{\alpha T_0}{2}\right)\times$$

$$\times\left(\alpha(2 + \sqrt{2}\,\alpha T_0)\,\mathrm{ch}\,\frac{\alpha T_0}{2} + \alpha^2 T_0\,\mathrm{sh}\,\frac{\alpha T_0}{2}\right)^{-1}.$$

25.3. Оценки параметров регрессии

25.3.1. Оценки параметров регрессии по методу наименьших квадратов. Предположим, что наблюдается процесс вида

$$\xi(t) = \sum_{k=1} \theta_k a_k(t) + \xi_0(t), \tag{3.1}$$

где $\xi_0(t)$ — стационарный процесс с нулевым средним, $a_k(t)$ ($k = 1, \ldots, r$) — известные неслучайные функции, которые предполагаются линейно независимыми, θ_k ($k = 1, \ldots, r$) — неизвестные параметры. Задачу нахождения оценок параметров θ_k по реализации $x(t)$ ($t \in [a, b]$) процесса $\xi(t)$ называют *задачей оценки параметров регрессии*

$$A(t, \theta_1, \theta_2, \ldots, \theta_r) = \sum_{k=1}^{r} \theta_k a_k(t).$$

В радиотехнических приложениях стационарных процессов $A(t) = A(t, \theta_1, \ldots, \theta_r)$ называют *сигналом* (*полезным*), $\xi_0(t)$ — *стационарным шумом.* В экономических, биологических, социологических приложениях $A(t)$ называют *трендом.*

Наиболее простой вид имеют линейные несмещенные оценки параметров регрессии, вычисляемые по методу наименьших квадратов, — оценки $\tilde{\theta}_k$, минимизирующие квадратичный функционал

$$\int_a^b \left| x(t) - \sum_{k=1}^{r} \theta_k a_k(t) \right|^2 dt.$$

Если $a_k(t) \in L_2[a, b]$ $(k = 1, \ldots, r)$, то

$$\tilde{\theta}_k = \sum_{j=1}^{r} c_{kj}^{-1} \int_a^b \overline{a_j(t)}\, x(t)\, dt, \tag{3.2}$$

где c_{kj}^{-1} — (k, j)-й элемент матрицы, обратной к матрице

$$c_{kj} = \int_a^b \overline{a_k(t)}\, a_j(t)\, dt.$$

Заметим, что вычисление параметров регрессии по методу наименьших квадратов не предполагает знания корреляционных и спектральных свойств процесса $\xi(t)$.

Если же предположить, что спектральная функция $F_0(\lambda)$ процесса $\xi_0(t)$ абсолютно непрерывна и спектральная плотность $f_0(\lambda)$ ограничена и почти всюду положительна, то можно утверждать большее:

Теорема 1 *Для того чтобы оценки $\tilde{\theta}_k$ параметров θ_k по методу наименьших квадратов были состоятельными, необходимо и достаточно, чтобы для любых $\rho_1, \rho_2, \ldots, \rho_r$ функция $a(t) = \sum_{k=1}^{r} \rho_k a_k(t)$ удовлетворяла условию*

$$\int_{-\infty}^{\infty} |a(t)|^2\, dt = \infty.$$

В предположении, что процесс $\xi_0(t)$ имеет дробно-рациональную спектральную плотность, а базисные функции $a_k(t)$ имеют вид

$$a_k(t) = e^{it\omega_k} t^{m_k}, \tag{3.3}$$

где m_k — целые неотрицательные числа, ω_k — действительные числа (в этом случае $A(t)$ называют *полиномиально-тригонометрической регрессией*), оценки $\tilde{\theta}_k$ параметров θ_k регрессии являются асимптотически эффективными в следующем смысле.

Пусть матрица $G(a, b)$ — ковариационная матрица наилучших линейных несмещенных оценок параметров регрессии (ее явный вид приводится в следующем пункте), $\tilde{G}(a, b)$ — ковариационная матрица оценок $\tilde{\theta}_k$ параметров θ_k, полученных по методу наименьших квадратов. Тогда $\tilde{G}(a, b) \geqslant G(a, b)$ (т. е. $\tilde{G}(a, b) - G(a, b)$ — неотрицательно определенная матрица) и существует неотрицательная неубывающая функция $g(t)$ такая, что

$$\lim_{b-a\to\infty} g(b-a)\, \tilde{G}(a, b) = \lim_{b-a\to\infty} g(b-a)\, G(a, b) \neq 0.$$

25.3.2. Наилучшие линейные несмещенные оценки параметров регрессии. Если спектральная функция $F(\lambda)$, а следовательно, и корреляционная функция $B(t)$ процесса $\xi_0(t)$ известны, то обычно предполагается, что функции $a_k(t)$, образующие базис регрессии $A(t)$, таковы, что процесс $\xi(t)$ допускает спектральное представление $\xi(t)=\int\limits_{-\infty}^{\infty} e^{it\lambda}\,d\zeta(\lambda)$, где $d\zeta(\lambda)=\sum\limits_{k=1}^{r}\theta_k\overline{\alpha_k(\lambda)}\,dF(\lambda)+d\zeta_0(\lambda)$, $\zeta_0(\lambda)$ — спектральный процесс, соответствующий процессу $\xi_0(t)$: $\xi_0(t)=\int\limits_{-\infty}^{\infty} e^{it\lambda}\,d\zeta_0(\lambda)$, а $\alpha_k(\lambda)$ — функции, интегрируемые с квадратом по спектральной мере $F(\cdot)$:

$$\int\limits_{-\infty}^{\infty} |\alpha_k(\lambda)|^2\,dF(\lambda)<\infty$$

(или кратко $\alpha_k(\lambda)\in L_2(F)$), которые являются решениями интегрального уравнения типа Винера — Хопфа:

$$\int\limits_{-\infty}^{\infty} e^{it\lambda}\alpha_k(\lambda)\,dF(\lambda)=a_k(t). \tag{3.4}$$

Пример 1. Пусть $\xi(t)=\theta a(t)+\xi_0(t)$ и почти все траектории процесса $\xi_0(t)$ непрерывны. Если функция $a(t)$ имеет разрыв в точке $t_0\in(a,\ b)$, то по единственной реализации $x(t)$ $(t\in[a,\ b])$ можно точно определить значения параметра θ, а именно оценка

$$\theta_h=\frac{1}{a(t_0+)-a(t_0-)}\,[x(t_0+h)-x(t_0-h)]$$

с вероятностью 1 при $h\to 0$ сходится к точному значению θ. Для функции $a(t)$ из данного примера уравнение (3.4) не имеет решений в $L_2(F)$.

Теорема 2. *Для того чтобы уравнение* (3.4) *имело решение в* $L_2(F)$, *необходимо и достаточно, чтобы*

$$\inf\int\limits_{-\infty}^{\infty} |\psi(\lambda)|^2\,dF(\lambda)\neq 0,$$

либо, что эквивалентно,

$$\inf\sum_{j,\,l} c_j c_l b(t_j-t_l)>0,$$

где inf *берется по* $\psi(\lambda)$, *являющимся конечными суммами вида*

$$\sum_j c_j e^{i\lambda t_j},\quad t_j\in T,$$

и такими, что $\sum\limits_j c_j a_k(t_j)=1$ $(k=1,\ \ldots,\ r)$. *Если решение уравнения* (3.4) *существует, то оно единственно в* $L_2(F)$.

Если решения уравнения (3.4) найдены, то задача оценки параметров регрессии сводится к решению линейной системы алгебраических уравнений.

Теорема 3. *Если спектральная функция $F(\lambda)$ и базисные функции $a_k(t)$ $(k=1, \ldots, r)$ регрессии $A(t)=\sum_k \theta_k a_k(t)$ удовлетворяют условия теоремы 1 и $\alpha_k(\lambda)$ — решения уравнений (3.4), то линейные несмещенные оценки $\widehat{\theta}_k$ параметров θ_k, имеющие минимальную дисперсию, определяются формулами*

$$\hat{\theta}_k=\sum_{j=1}^{r} c_{kj} \int_{-\infty}^{\infty} \overline{\alpha_j(\lambda)}\, d\zeta(\lambda), \qquad (3.5)$$

где $\zeta(\lambda)$ — спектральное представление траектории $x(t)$ $(t \in [a, b])$, $c_{kj} = \operatorname{cov}(\widehat{\theta}_k, \widehat{\theta}_j)$ — (k, j)-й элемент матрицы C, обратной к матрице $D = \{d_{kj}, k, j = 1, \ldots, r\}$, где

$$d_{kj}=\int_{-\infty}^{\infty} \overline{\alpha_k(\lambda)}\, f(\lambda)\, \alpha_j(\lambda)\, d\lambda.$$

Если C_1 — ковариационная матрица линейных несмещенных оценок, отличных от $\widehat{\theta}_k$ $(k = 1, \ldots, r)$, определяемых равенством (3.5), то $C \leqslant C_1$ в том смысле, что матрица $C_1 - C$ неотрицательно определена.

25.3.3. Решение уравнений типа Винера — Хопфа для дробно-рациональных спектральных плотностей. В том практически наиболее важном случае, когда спектральная плотность $f(\lambda)$ процесса $\xi_0(t)$ дробно-рациональна, т. е. когда

$$f(\lambda)=\left|\frac{P(i\lambda)}{Q(i\lambda)}\right|^2,$$

где $P(z)=\sum_{k=1}^{p} p_k z^k$, $Q(z)=\sum_{k=1}^{q} q_k z^k$ $(p<q)$, решение уравнений (3.4), а вместе с ним и решение задачи оценки параметров регрессии могут быть найдены в явном виде.

Теорема 4. *Если многочлены $P(z)$ и $Q(z)$ имеют нули лишь в левой полуплоскости, ни один из корней многочлена $P(z)$ не является чисто мнимым и функции $a_k(t)$ $(k = 1, \ldots, r)$ удовлетворяют условиям теоремы 1, то решение $\alpha_k(\lambda)$ уравнения*

$$\int_{-\infty}^{\infty} e^{ib\lambda} \alpha_k(\lambda) \left|\frac{P(i\lambda)}{Q(i\lambda)}\right|^2 d\lambda = a_k(t)$$

имеет вид

$$\alpha_k(\lambda) = e^{ia\lambda} \sum_{j=0}^{n-m-1} c_{jk}^{(a)} (i\lambda)^j + e^{ib\lambda} \sum_{j=0}^{n-m-1} c_{jk}^{(b)} (i\lambda)^j + \int_a^b e^{i\lambda t} c_k(t)\, dt,$$

где $c_k(t) = \frac{1}{2\pi} Q\left(\frac{d}{dt}\right) Q\left(-\frac{d}{dt}\right) v_k(t)$, $v_k(t)$ *— решение дифференциального уравнения*

$$P\left(\frac{d}{dt}\right) P\left(-\frac{d}{dt}\right) v_k(t) = a_k(t), \quad t \in (a, b),$$

с граничными условиями

$$\lim_{t \downarrow a} \frac{d^l}{dt^l} Q\left(-\frac{d}{dt}\right) v_k(t) = \lim_{t \uparrow b} \frac{d^l}{dt^l} Q\left(\frac{d}{dt}\right) v_k(t) = 0,$$

$$l = 0, \ldots, m-1,$$

$$c_{jk}^{(a)} = \lim_{t \downarrow a} \frac{1}{2\pi} \sum_{l=j+m+1}^{n} q_l Q\left(-\frac{d}{dt}\right) v_k(t),$$

$$c_{jk}^{(b)} = \lim_{t \uparrow b} \frac{1}{2\pi} \sum_{l=j+m+1}^{n} (-1)^l q_l Q\left(\frac{d}{dt}\right) v_k(t).$$

25.4. Оценки спектральной плотности и спектральной функции стационарных последовательностей

25.4.1. Периодограмма. Пусть $\{\xi(t), t \in T\}$ — стационарная в широком смысле случайная последовательность (*временной ряд*) и $x(t)$ $(t \in T_0)$ — траектория процесса $\xi(t)$, где $T_0 = \{t_k, k = 1, \ldots, n\}$ $(t_k \in T)$.

В основе большинства статистик, предназначенных для оценки спектральной функции и спектральной плотности процесса $\{\xi(t), t \in T\}$, лежит статистика, называемая *периодограммой* и определяемая равенством

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t \in T_0} x(t) e^{it\lambda} \right|^2. \tag{4.1}$$

Замечание. Для процессов с непрерывным временем периодограмма определяется как

$$I_{(a,b)}(\lambda) = \frac{1}{2\pi(b-a)} \left| \int_a^b x(t) e^{it\lambda} dt \right|^2, \tag{4.2}$$

где $x(t)$ $(t \in [a, b])$ — траектория изучаемого стационарного процесса. Формулируемые ниже результаты имеют соответствующие непрерывные аналоги.

Если $f(\lambda)$ — спектральная плотность процесса $\xi(t)$ $(t \in T)$, то

$$\lim_{n \to \infty} MI_n(\lambda) = f(\lambda), \tag{4.3}$$

т. е. периодограмма является асимптотически несмещенной оценкой спектральной плотности. Однако

$$\lim_{n\to\infty} \operatorname{cov}\left[I_n(\lambda_1), I_n(\lambda_2)\right] = \begin{cases} 2f^2(0), & \lambda_1 = \lambda_2 = 0, \\ f^2(\lambda), & \lambda_1 = \lambda_2 = \lambda, \\ 0, & \lambda_1 \neq \lambda_2, \end{cases} \tag{4.4}$$

т. е. периодограмма не является состоятельной оценкой для спектральной плотности.

Периодограмма $I_n(\lambda)$, рассматриваемая как случайный процесс по λ, имеет в силу (4.4) при больших n сильно флуктуирующие траектории.

25.4.2. Оценки спектральной функции. Пусть

$$\widehat{F}_n(\lambda) = \int_{-\pi}^{\lambda} I_n(\mu)\, d\mu \tag{4.5}$$

и $F(\lambda)$ — спектральная функция процесса $\xi(t)$ $(t \in T)$. Статистика $\widehat{F}_n(\lambda)$ является асимптотически несмещенной оценкой спектральной функции $F(\lambda)$:

$$\lim_{n\to\infty} \mathsf{M}\widehat{F}_n(\lambda) = F(\lambda). \tag{4.6}$$

Если $\xi(t)$ $(t \in T)$ — эргодический процесс, то статистика $\widehat{F}_n(\lambda)$ является состоятельной оценкой спектральной функции: это оправдывает определение $\widehat{F}_n(\lambda)$ как *эмпирической спектральной функции.*

Если $F(\lambda)$ — абсолютно непрерывна, то

$$\lim_{n\to\infty} \sup_{\lambda} \left|\widehat{F}_n(\lambda) - F(\lambda)\right| = 0. \tag{4.7}$$

Пусть $\xi(t)$ $(t \in T)$ — регулярная стационарная последовательность и $\xi(t) = \sum_{k\in T} c_k \zeta(t-k)$ — ее представление в виде скользящего суммирования, $\mathsf{M}\zeta(t) = 0$, $\mathsf{M}\zeta^2(t) = 1$.

Положим $\widehat{F}_n^{(0)}(\lambda) = \int_0^{\lambda} I_n(\mu)\, d\mu$, $F^{(0)}(\lambda) = \int_0^{\lambda} f(\mu)\, d\mu$, где $f(\lambda)$ — спектральная плотность процесса $\xi(t)$ $(t \in T)$.

Теорема 1. *Если выполнены условия:* а) $\xi(t)$ *имеет конечный четвертый момент* μ_4; б) $f(\lambda)$ — *абсолютно непрерывна*; в) $c_k = O(k^{\beta})$, $\beta < -3/2$, *то*

$$\lim \mathsf{P}\left\{\max_{0\leqslant\lambda\leqslant\pi} \sqrt{n}\left|\widehat{F}_n^{(0)}(\lambda) - F^{(0)}(\lambda)\right| \leqslant z\right\} = \\ = \mathsf{P}\left\{\max_{0\leqslant\lambda\leqslant\pi} |\eta(\lambda)| \leqslant z\right\}, \tag{4.8}$$

где $\eta(\lambda)$ — *гауссовский процесс с* $\mathsf{M}\eta(\lambda) = 0$, $\mathsf{M}\eta(\lambda)\eta(\mu) =$

$$= (\mu_4 - 3) F^{(0)}(\lambda) F^{(0)}(\mu) + 2\pi \int_0^{\min(\lambda,\mu)} f^2(s)\, ds.$$

Равенство (4.8) напоминает соответствующий результат для статистики Колмогорова, однако в отличие от последней в данном случае предельное распределение зависит от оцениваемой функции.

25.4.3. Оценки спектральной плотности. В качестве точечной оценки спектральной плотности $f(\lambda)$ стационарной последовательности $\{\xi(t),\ t \in T\}$ по наблюдению $x(t),\ t \in T_0 = \{t_k,\ k = 1, \ldots, n\}$, выбирают статистики $\hat{f}_n(\lambda)$ вида

$$\hat{f}_n(\lambda) = \int_{-\pi}^{\pi} W_n(\lambda - \mu)\, I_n(\mu)\, d\mu, \tag{4.9}$$

где $I_n(\lambda)$ — периодограмма, а «весовые» функции $W_n(\lambda)$, называемые *спектральными окнами*, подбираются так, чтобы

1) $W_n(\lambda)$ имела резко выраженный максимум в нуле;

2) $\int_{-\pi}^{\pi} W_n(\lambda)\, d\lambda = 1$;

3) $\mathsf{D}\hat{f}_n(\lambda) \to 0$ при $n \to \infty$.

Условие 1) «вырезает» оценку требуемой частоты, которая в силу 2) оказывается асимптотически несмещенной и в силу 3) — состоятельной.

Если выполнены условия:

1) $\{\xi(t),\ t \in T\}$ — регулярный процесс и $\xi(t) = \sum_{k=0}^{\infty} c_k \zeta(t-k)$ — его представление в виде скользящего суммирования;

2) $\mathsf{M}\xi^2(t) = 1$, $\mathsf{M}\xi^4(t) < \infty$;

3) $c_k = O\left(|k|^{-(2+\delta)}\right)$ для некоторого $\delta > 0$;

4) $\left|\dfrac{W_n^{2*}(\mu)}{W_n^{2*}(0)} - 1\right| \to 0$ при $n \to \infty$ для $|\mu| \leqslant c/n$,

где $W_n^{2*}(\mu) = W_n(\mu) * W_n(\mu)$ — свертка $W_n(\mu)$ с собой, c — некоторая положительная константа, то статистика $\hat{f}_n(\lambda)$ асимптотически нормальна со средним

$$\mathsf{M}\hat{f}_n(\lambda) = \int_{-\pi}^{\pi} W_n(\lambda - \mu)\, f(\lambda)\, d\mu + O\left(\frac{\ln n}{n}\right) \tag{4.10}$$

и дисперсией

$$\mathsf{D}\hat{f}_n(\lambda) \sim \frac{2\pi}{n} \int_{-\pi}^{\pi} W_n^2(\lambda - \mu)\, f^2(\mu)\, d\mu \sim \frac{2\pi f^2(\lambda)}{n} \int_{-\pi}^{\pi} W_n^2(\mu)\, d\mu. \tag{4.11}$$

Статистика $\hat{f}_n(\lambda)$ является состоятельной оценкой для $f(\lambda)$, если

$$\lim_{n \to \infty} \frac{1}{n} \int_{-\pi}^{\pi} W_n^2(\lambda)\, d\lambda = 0.$$

Среди множества спектральных окон имеется класс, чаще всего используемый в статистической практике, — класс спектральных окон, допускающих представление вида

$$W_n(\lambda) = 2 \sum_{-n+1}^{n-1} k_n(l)\, e^{-il\lambda}, \tag{4.12}$$

где $k_n(l) = k(l/m_n)$, $\{m_n\}$ — некоторая неограниченно возрастающая последовательность целых положительных чисел такая, что $m_n/n \to 0$, $n \to \infty$, $k(x)$ — ограниченная четная функция, удовлетворяющая условиям: $k(0) = 1$, $|k(x)| < 1$ при $x \neq 0$, $\int\limits_{-\infty}^{\infty} k^2(x)\, dx < \infty$.

Предположим, что стационарная последовательность $\xi(t)$ является регулярной, $\xi(t) = \sum\limits_{k=0}^{\infty} c_k \zeta(t-k)$ — ее представление в виде скользящего суммирования, $\sum\limits_{k=0}^{\infty} |c_k| < \infty$, независимые случайные величины $\zeta(t)$ имеют нулевые средние и конечный четвертый момент. Пусть также $B(t)$ — корреляционная функция $\xi(t)$.

Представление об асимптотическом характере смещения $f(\lambda) - \mathsf{M}f_n(\lambda)$ дает следующая теорема.

Теорема 2. *Если*

1) *при некотором* $q > 1$ *существует и конечен предел*

$$k_q = \lim_{x \to 0} \frac{1 - k(x)}{|x|^q};$$

2) $n/m_n^q \to \infty$, $n \to \infty$;

3) $\sum\limits_{t \in T} |t^q B(t)| < \infty$,

то

$$\lim_{n \to \infty} m_n^q [f(\lambda) - \mathsf{M}\hat{f}_n(\lambda)] = \frac{k_q}{2\pi} \sum_{t \in T} |t|^q B(t)\, e^{it\lambda} < \infty. \tag{4.13}$$

Асимптотические ковариационные свойства статистик $\hat{f}_n(\lambda)$ описывает следующая теорема.

Теорема 3.

$$\lim_{n \to \infty} \frac{n}{m_n} \operatorname{cov}[\hat{f}_n(\lambda_1), \hat{f}_n(\lambda_2)] = \begin{cases} \lambda_1 \neq \pm\lambda_2, \\ f^2(\lambda) \int\limits_{-\infty}^{\infty} k^2(t)\, dt, \ \lambda_1 = \pm\lambda_2 = \lambda \neq 0, \\ 2f^2(\lambda) \int\limits_{-\infty}^{\infty} k^2(t)\, dt, \ \lambda_1 = \lambda_2 = 0, \pm\pi. \end{cases}$$

В условиях теоремы 2 наиболее быстрая сходимость дисперсии $\mathsf{D}\hat{f}_n(x)$ к нулю достигается при

$$m_n = O\left(n^{1/(1+2q)}\right).$$

25.4.4. Примеры спектральных окон. Обозначим через $B_n(t) = \frac{1}{n-t}\sum_{j=1}^{n-t} x_{j+t}\bar{x}_j$ оценку корреляционной функции $B(t)$ (эмпирическую корреляционную функцию).

1. *Финитное преобразование Фурье (оценка Даниэля)*

$$W_n(\lambda) = \begin{cases} m_n/2, & |\lambda| \leqslant \pi/m_n, \\ 0, & |\lambda| > \pi/m_n, \end{cases}$$

$$\hat{f}_n(\lambda) = \frac{m_n}{2\pi}\sum_{t=-n+1}^{n-1}\left(1-\frac{|t|}{n}\right)B_n(t)\frac{1}{\pi t}\sin\frac{\pi t}{m_n}e^{it\lambda},$$

$$k(x) = \frac{\sin \pi x}{\pi x}.$$

2. *Усеченная оценка*

$$W_n(\lambda) = 2\sin\left(\frac{2m_n+1}{2}\lambda\right)\left(\sin\frac{\lambda}{2}\right)^{-1},$$

$$\hat{f}_n(\lambda) = \frac{1}{2\pi}\sum_{t=-m_n}^{m_n}\left(1-\frac{|t|}{n}\right)B_n(t)e^{it\lambda},$$

$$k(x) = \begin{cases} 1, & |x| \leqslant 1, \\ 0, & |x| > 1. \end{cases}$$

3. *Оценка Бартлетта*

$$W_n(\lambda) = \frac{\sin^2\frac{m_n\lambda}{2}}{m_n\sin^2\frac{\lambda}{2}},$$

$$\hat{f}_n(\lambda) = \frac{1}{2\pi}\sum_{t=-m_n}^{m_n}\left(1-\frac{|t|}{n}\right)\left(1-\frac{|t|}{m_n}\right)B_n(t)e^{it\lambda},$$

$$k(x) = \begin{cases} 1-|x|, & |x| \leqslant 1, \\ 0, & |x| > 0. \end{cases}$$

4. *Оценка Тьюки — Хэннинга*

$$\hat{f}_n(\lambda) = \frac{1}{2}\hat{f}'_n(\lambda) + \frac{1}{n}\hat{f}'_n\left(\lambda - \frac{\pi}{m_n}\right) + \frac{1}{n}\hat{f}'_n\left(\lambda + \frac{\pi}{m_n}\right),$$

где $\hat{f}'_n(\lambda)$ — усеченная оценка из примера 2;

$$k(x) = \begin{cases} (1 + \cos \pi x)/2, & |x| \leqslant 1, \\ 0, & |x| > 1. \end{cases}$$

25.5. Оценки параметров спектральной плотности

Предположим, что наблюдается траектория $x(t)$ $(t \in T_0)$ стационарного процесса $\{\xi(t),\ t \in T\}$, $\mathsf{M}\xi(t) = 0$, $\mathsf{D}\xi(t) = \sigma^2$, спектральная плотность которого известна с точностью до одного или нескольких параметров, принадлежащих некоторому параметрическому множеству Θ:

$$f(\lambda) = f(\lambda, \theta), \quad \theta \in \Theta.$$

Оценке подлежат σ^2 и $\theta \in \Theta$.

В предположении, что процесс $\xi(t)$ $(t \in T)$ регулярен для любых $\theta \in \Theta$ и $\xi(t) = \sum_{k=0}^{\infty} c_k(\theta)\zeta(t-k)$ — его представление в виде скользящего суммирования, в качестве оценок $\hat{\sigma}_n^2$ и $\hat{\theta}_n$ неизвестных σ^2 и θ можно выбрать те значения $\hat{\sigma}_n^2$ и $\hat{\theta}_n$, для которых достигается минимум выражения

$$n \ln \sigma^2 + \frac{W_n(\theta, x(t), t \in T_0)}{\sigma^2}, \tag{4.14}$$

где $W_n(\theta, x(t), t \in T_0) = n \sum_{-n+1}^{n-1} W(t, \theta)\left(1 - \frac{|t|}{n}\right) B_n(t)$; $B_n(t) = \frac{1}{n-t}\sum_{j=1}^{n-t} x(t_{j+t})\, x(t_j)$ — эмпирическая корреляционная функция; $W(t, \theta)$ — коэффициент при z^t в лорановском разложении функции $v(z, \theta) = \frac{1}{u(z, \theta)\, u(1/z, \theta)}$, $u(z, \theta) = \sum_{k=0}^{\infty} c_k(\theta) z^k$ (предполагается, что такое разложение возможно в кольце, содержащем единичную окружность).

Теорема 1. *Если выполнены условия:*

1) *случайные величины $\zeta(t)$ в представлении процесса $\xi(t)$ в виде скользящего суммирования независимы, одинаково распределены и имеют конечные четвертые моменты;*

2) *параметрическое множество Θ является компактом в конечномерном евклидовом пространстве $\mathbb{R}^k$, $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$;*

3) $|v(z, \theta^1)|^2 \neq |v(z, \theta^2)|^2$ *для почти всех* $|z| = 1$, $\theta^1, \theta^2 \in \Theta$, $\theta^1 \neq \theta^2$;

4) $f(\lambda, \theta)$ *и* $f^{-1}(\lambda, \theta)$ *непрерывны на* $[-\pi, \pi]$,

то оценки $\hat{\sigma}_n^2$ *и* $\hat{\theta}_n$ *являются состоятельными оценками* σ^2 *и* θ.

Если, кроме того, выполнено условие:

5) *функция* $(|v(z, \theta)|^2)^{-1}$ *имеет непрерывные производные по* θ_i *до третьего порядка включительно в окрестности истинного значения* θ_0 *параметра* θ *и*

$$\sum_{k=1}^{\infty} k\,|c_k(\theta_0)| < \infty,$$

то распределение вектора $\frac{1}{\sqrt{n}}(\hat{\theta}_n - \theta_0)$ *сходится при* $n \to \infty$ *к нормальному с нулевым средним и ковариационной матрицей* G_0^{-1}, *где* G_0 *имеет элементы*

$$g_{lm}^0 = \lim_{\theta \to \theta_0} \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial \ln |u(e^{i\lambda}, \theta)|^2}{\partial \theta_l} \frac{\partial \ln |u(e^{i\lambda}, \theta)|^2}{\partial \theta_m} \, d\lambda$$

$(l, m = 1, \ldots, k)$ *(в предположении, что* G_0 *невырождена)*.

Литература: [2, 7, 25, 65, 92, 101, 110].

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика. Основы моделирования и первичная обработка данных.— М.: Финансы и статистика, 1983.
2. Андерсон Т. Статистический анализ временных рядов/Пер. с англ. — М.: Мир, 1976.
3. Андерсон Т. Введение в многомерный статистический анализ/Пер. с англ. — М.: Физматгиз, 1963.
4. Барра Ж.-Р. Основные понятия математической статистики. — М.: Мир, 1974.
5. Бикялис А. О центральной предельной теореме в $\mathbf{R}^k$ — Литов. мат. сб., 1971, № 1, с. 27—58.
6. Биллингсли П. Сходимость вероятностных мер/Пер. с англ. — М.: Наука, 1977.
7. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление/Пер. с англ. — М.: Мир, 1974.
8. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. — 3-е изд. — М.: Наука, 1983.
9. Боровков А. А. Сходимость мер и случайных процессов. — УМН, 1976, 31, № 2, с. 5—68.
10. Боровков А. А. Асимптотические методы в теории массового обслуживания. — М.: Наука, 1980.
11. Боровков А. А. Вероятностные процессы в теории массового обслуживания. — М.: Наука, 1972.
12. Боровков А. А. Курс теории вероятностей. — М.: Наука, 1972.
13. Боровков А. А. Математическая статистика. Оценка параметров, проверка гипотез. — М.: Наука, 1984.
14. Ван дер Варден Б. А. Математическая статистика/Пер. с нем. — М.: ИЛ, 1960.
15. Вентцель А. Д. Курс теории случайных процессов. — М.: Наука, 1975.
16. Гельфанд И. М., Виленкин Н. Я. Некоторые применения гармонического анализа. — М.: Физматгиз, 1961 г.
17. Гихман И. И. Предельные теоремы для последовательностей серий случайных величин. — Теория случайных процессов. Респ. межвед. сб. Киев: Наук. думка, 1974, вып. 2, 37—47 с.
18. Гихман И. И., Скороход А. В. Стохастические дифференциальные уравнения и их приложения. Киев: Наук. думка, 1982.
19. Гихман И. И., Скороход А. В. Введение в теорию случайных процессов. — 2-е изд. — М.: Наука, 1977.
20. Гихман И. И., Скороход А. В. Теория случайных процессов. Т. 1. — М.: Наука, 1971.

21. Гихман И. И., Скороход А. В. Стохастические дифференциальные уравнения. — Киев: Наук. думка, 1968.
22. Гихман И. И., Скороход А. В. Стохастические дифференциальные уравнения и их приложения. — Киев: Наук. думка, 1982.
23. Гнеденко Б. В. Курс теории вероятностей. — 4-е изд. — М.: Наука, 1965.
24. Гнеденко Б. В., Колмогоров А. Н. Предельные распределения для сумм независимых случайных величин. — Л. — М.: Гостехиздат, 1949.
25. Гренандер У. Случайные процессы и статистические выводы/Пер. с англ. — М.: ИЛ, 1961.
26. Григелионис Б. О мартингальной характеризации случайных процессов с независимыми приращениями. — Литов. мат. сб., 1974, 14, № 4, с. 45—61.
27. Деллашери К. Емкости и случайные процессы. — М.: Мир 1975.
28. Дуб Дж. Вероятностные процессы/Пер. с англ. — М.: ИЛ, 1956.
29. Дынкин Е. Б. Основания теории марковских процессов. — М.: Физматгиз, 1959.
30. Дынкин Е. Б. Марковские процессы. — М.: Физматгиз, 1963.
31. Дынкин Е. Б., Юшкевич А. А. Теоремы и задачи о процессах Маркова. — М.: Наука, 1967.
32. Звонкин А. К., Крылов Н. В. О сильных решениях стохастических дифференциальных уравнений.— В кн.: Труды школы-семинара по теории случайных процессов (Друскининкай, 1974 г.). Вильнюс: Изд-во Ин-та физики и математики АН ЛитовССР, 1975, с. 9—88.
33. Золотарев В. М. Одномерные устойчивые распределения. — М.: Наука, 1983.
34. Ибрагимов И. А., Линник Ю. В. Независимые и стационарно связанные величины. — М.: Наука, 1965.
35. Иванов Г. А., Чешкин Ю. Р. Исследование статистических критериев, используемых для построения математической модели при аппроксимации опытных данных. — В сб.: Некоторые вопросы теории случайных процессов. Киев: ИМ АН УССР, 1984, с. 133—140.
36. Ито К. Вероятностные процессы/Пер. с япон. — М.: ИЛ, 1963.
37. Ито К., Маккин Г. Диффузионные процессы и их траектории/Пер. с англ. — М.: Мир, 1968.
38. Кабанов Ю. М., Липцер Р. Ш., Ширяев А. Н. Мартингальные методы в теории точечных процессов. — В кн.: Труды школы-семинара по теории случайных процессов (Друскининкай, 1974). Вильнюс: Изд-во ин-та физики и математики АН ЛитовССР, 1975, с. 269—354.
39. Карлин С. Основы теории случайных процессов, М.: Мир, 1970.
40. Кемени Дж., Снелл Дж. Конечные цепи Маркова/Пер. с англ. — М.: Наука, 1970.
41. Кендалл М., Стьюарт А. Теория распределений/Пер. с англ. — М.: Наука, 1973.
42. Кендалл М., Стьюарт А. Статистические выводы и связи/Пер. с англ. — М.: Наука, 1975.

43. Кокс Д., Хинкли Д. Теоретическая статистика/Пер. с англ.— М.: Мир, 1978.
44. Колмогоров А. Н. Основные понятия теории вероятностей. — 2-е изд. — М.: Наука, 1974.
45. Колмогоров А. Н. Стационарные последовательности в гильбертовом пространстве. — Бюллетень МГУ, 1941, 2, № 6, 1—40.
46. Колмогоров А. Н. Интегрирование и экстраполирование случайных последовательностей. — Изв. АН СССР, сер. мат., 1941, 5, № 5, 3—14.
47. Королюк В. С. Граничные задачи для сложных пуассоновских процессов. — Киев: Наук. думка, 1975.
48. Королюк В. С., Турбин А. Ф. Полумарковские процессы и их применение. — Киев: Наук. думка, 1976.
49. Крамер Г. Математические методы статистики/Пер. с англ.— 2-е изд. — М.: Мир, 1975.
50. Крамер Г., Лидбеттер М. Стационарные случайные процессы/ Пер. с англ. — М.: Мир, 1969.
51. Круглов В. М. Дополнительные главы теории вероятностей.— М.: Высш. шк., 1984.
52. Крылов Н. В. Управляемые процессы диффузионного типа. — М.: Наука, 1977.
53. Кузнецов С. Е. Неоднородные марковские процессы. — В кн.: Современные проблемы математики. — ВИНИТИ АН СССР, Итоги науки и техн., 1972, 20, с. 37—178.
54. Ламперти Дж. Случайные процессы.— Киев: Вища шк., 1983.
55. Леви П. Стохастические процессы и броуновское движение/Пер. с франц. — М.: Наука, 1972.
56. Леман Е. Л. Проверка статистических гипотез/Пер. с англ. — М.: Наука, 1964.
57. Линник Ю. В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. — М.: Физматгиз, 1962.
58. Линник Ю. В. Статистические задачи с мешающими параметрами. — М.: Наука, 1966.
59. Липцер Р. М., Ширяев А. Н. Статистика случайных процессов. — М.: Наука, 1974.
60. Лоэв М. Теория вероятностей/Пер. с англ. — М.: ИЛ, 1962.
61. Маккин Г. Стохастические интегралы/Пер. с англ. — М.: Мир, 1972.
62. Мейер П. Вероятность и потенциалы/Пер. с англ. — М.: Мир, 1973.
63. Невё Ж. Математические основы теории вероятностей/Пер. с франц. — М.: Мир, 1969.
64. Петров В. В. Суммы независимых случайных величин. — М.: Наука, 1972.
65. Писаренко В. Ф., Розанов Ю. А. О некоторых задачах для стационарных процессов, приводящих к интегральным уравнениям, родственным уравнению Винера — Хопфа. — Проблемы передачи информации, 1963, вып. 14.
66. Портенко Н. И. Обобщенные диффузионные процессы. — Киев: Наук. думка, 1982.
67. Прохоров Ю. В. Сходимость случайных процессов и предельные теоремы теории вероятностей. — Теория вероятностей и ее применения, 1956, 1, № 2, с. 127—238.

68. Прохоров Ю. В., Розанов Ю. А. Теория вероятностей: основные понятия, предельные теоремы, случайные процессы. — М.: Наука, 1973.
69. Рао С. Р. Линейные статистические методы и их применения/Пер. с англ. — М.: Наука, 1968.
70. Розанов Ю. А. Стационарные случайные процессы. — М.: Физматгиз, 1963.
71. Розовский Б. Л. Эволюционные стохастические системы. — М.: Наука, 1983.
72. Романовский В. И. Математическая статистика. — М.: ГОНТИ, 1938.
73. Романовский В. И. Дискретные цепи Маркова. — Л. — М.: Гостехиздат, 1949.
74. Сазонов В. В. О скорости сходимости в многомерной центральной предельной теореме. — Теория вероятностей и ее применение, 1968, 1, с. 191—194.
75. Сарымсаков Т. А. Основы теории процессов Маркова. — М.: Гостехиздат, 1954.
76. Себер Дж. Линейный регрессионный анализ. — М.: Мир, 1980.
77. Севастьянов Б. А. Ветвящиеся процессы. — М.: Наука, 1971.
78. Сираждинов С. Х. Предельные теоремы для однородных цепей Маркова. — Ташкент: Изд-во АН УзССР, 1955.
79. Скороход А. В. Стохастические уравнения для сложных систем. — М.: Наука, 1983.
80. Скороход А. В. Предельные теоремы для случайных процессов. — Теория вероятностей и ее применение, 1956, 1, № 3, с. 289—319.
81. Скороход А. В. Случайные процессы с независимыми приращениями. — М.: Наука, 1964.
82. Скороход А. В. Исследования по теории случайных процессов. — Киев: Изд-во Киев. ун-та, 1961.
83. Скороход А. В. Елементи теорії ймовірностей та випадкових процесів. — Київ: Вища шк., 1975.
84. Скороход А. В., Слободенюк Н. П. Предельные теоремы для случайных блужданий. — Киев: Наук. думка, 1970.
85. Соле Дж. Л. Основные структуры математической статистики/Пер. с франц. — М.: Мир, 1972.
86. Спицер Ф. Принципы случайного блуждания/Пер. с англ. — М.: Мир, 1969.
87. Такач Л. Комбинаторные методы в теории случайных процессов/Пер. с англ. — М.: Мир, 1971.
88. Уилкс С. Математическая статистика/Пер. с англ. — М.: Наука, 1967.
89. Феллер В. Введение в теорию вероятностей и ее приложения/Пер. с англ. Т. 1, 2. — М.: Мир, 1984.
90. Хант Дж. Марковские процессы и потенциалы/Пер. с англ. — М.: ИЛ, 1962.
91. Харрис Т. Е. Теория ветвящихся процессов/Пер. с англ. — М.: Мир, 1966.
92. Хеннан Э. Анализ временных рядов/Пер. с англ. — М.: Наука, 1964.
93. Хеннан Э. Многомерные временные ряды/Пер. с англ. — М.: Мир, 1974.

94. Хинчин А. Я. Предельные законы для сумм независимых случайных величин. — М.: ГОНТИ, 1938.
95. Хинчин А. Я. Теория корреляции стационарных случайных процессов. — УМН, 1938, вып. 5, с. 42—51.
96. Хинчин А. Я. Асимптотические законы теории вероятностей. — М. — Л.: ОНТИ, 1936.
97. Чжун Кай-лай. Однородные цепи Маркова/Пер. с англ. — М.: Мир, 1964.
98. Ширяев А. Н. Вероятность. — М.: Наука, 1980.
99. Ширяев А. Н. Статистический последовательный анализ. — М.: Наука, 1969.
100. Шуренков В. М. Эргодические теоремы и смежные вопросы теории случайных процессов. — Киев: Наук. думка, 1981.
101. Яглом А. М. Экстраполирование, интерполирование и фильтрация стационарных процессов с рациональной спектральной плотностью. — Тр. Моск. мат. об-ва, 1955, **4**, с. 333—374.
102. Яглом А. М. Введение в теорию стационарных случайных функций. — УМН, 1952, УП, вып. 5, 3—168.
103. Яглом А. М. Корреляционная теория стационарных случайных функций. — Л.: Гидрометеоиздат, 1981.
104. Яглом А. М. Некоторые классы случайных полей в n-мерном пространстве, родственные стационарным случайным процессам. — Теория вероятностей и ее применения, 1957, **2**, № 3, с. 293—333.
105. Ядренко М. И. Спектральная теория случайных полей. — Киев: Изд-во Киев. ун-та, 1980.
106. Athreya K. B., Ney P. Branching processes. Berlin-Heidelberg, № 4, 1972.
107. Blumenthal R. M., Getoor R. K. Markov Processes and Potential Theory.— N. Y.: Academic Press, 1968.
108. Barlow M. T. One dimensional stochastic differential equations without strong solution. — Journal London Math. Soc., **26**, 1982, p. 335—347.
109. Friedman A., Stochastic differential equations and applications, v. 1, 2. — N. Y.: Academic Press, 1976.
110. Grenander U., Rosenblatt M. Statistical analysis of time series. — N. Y., Wiley, 1957.
111. Jacod J. Calcul stochastique et problemes de martingales. Lecture Notes in Mathematics, v. 714. — Berlin: Springer Verlag, 1979.
112. Kemeny J. G., Snell J. L., Knapp A. W. Denumerable Markov chains. — N. Y. — L.: Van Nostrand, 1966.
113. Kalaith T. Lectures on linear least-squares estimation. — Springer Verlag Wien, 1976.
114. Meyer P. A. Processus de Markov. Lecture Notes in Mathematics, v. 26. — Berlin: Springer Verlag, 1967.
115. Moyal I. E. The general theory of stochastic population processes. — Acta math., 1962, **108**, № 1, p. 1—31.
116. Neveu J. Martingales a temps discret. — Paris, 1972.
117. Revuz D. Markov chains. — Amsterdam: Elsevier, 1975.
118. Stroock D. W., Varadhan S. R. S. Multidimensional diffusion processes. — Berlin: Springer Verlag, 1979.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Абсолютная непрерывность мер 225, 582
Автокорреляционная функция 242
Алгебра событий 12
Альтернатива 542
Асимптотические разложения 84, 142
Асимметрии коэффициент 108
Атом 20

Байеса метод 561, 609
— формула 32
Бартлетта оценка 625
Бернулли распределение 109
— схема 40
Берри — Эссеена неравенство 82
Биркгофа теорема 280
Благоприятствующий исход 14
Бляшке — Привалова оператор 391
Большие уклонения 87
Больших чисел закон 41, 46, 47, 293, 295
— — — усиленный 48, 281, 412
Борелевские множества 17
Бореля — Кантелли теорема 39
Бореля — Таннера распределение 116
Бохнера — Хинчина теорема 65, 245
Буркхольдера неравенство 317

Вальда — Вольфовица критерий серий 549
— тождество 154
Вариационный ряд 528, 534
Вероятностное пространство 17
Вероятность 13, 14
— перехода 169, 172, 340, 343, 468
— — стохастически непрерывная 368
— — феллеровская 351, 364
— условная 32, 35

Винера метод 266
Винеровский процесс 211, 422
Вольда разложение 269
Восстановления процесс 145
— — линейчатый 149
— — стационарный 147
— уравнение 146
— функция 146, 151
Вполне монотонная функция 62
Временной ряд 621
Выборка 524
Выборки размах (широта) 535
Выборочная дисперсия 532, 540, 542
— функция 206, 285
Выборочное пространство 524
— среднее 532, 539, 542
Выборочные моменты 539

Гальтона — Ватсона процесс 427
Гармоническая функция 179, 181, 393
Гаусса — Маркова теорема 576
Гауссовская случайная величина 455
Гипотеза простая 542, 544
— сложная 542
Гливенко теорема 537
Граничные функционалы 161

Даниэля оценка 625
Деблина условие 183
Дебют множества 222
Дини условие 502
Дирихле распределение 139
Дисперсия 29
Доверительная область 532, 559
Доверительный интервал 532, 560
Дуба неравенство 318
— разложение 320, 327
Дубинса неравенство 318

Единственность решения сильная 499
— — слабая 499

Закон больших чисел 41, 46, 47, 293, 295
— — — усиленный 48, 281, 412
— нуля и единицы 39, 354
— повторного логарифма 50, 410, 413, 422
Значимости уровень 543

Измеримая функция 19
Измеримое пространство 17, 225
Инвариантное множество 184
Инвариантный заряд 179
Индикатор события 17
Интерполяция линейная 274
Инфинитезимальный оператор 362, 363
Информации количество 553
Информационная матрица 555
Ито формула 336, 477, 515

Квазимартингал 324
Квантиль 108
Класс апериодический 185, 194
— верхний 50
— нижний 50
— периодический 194
— эргодический 185
Ковариационная матрица 30
— функция 241
Ковариация 29, 229
Колмогорова неравенство 47, 406
— статистика 537, 548
— теорема 206, 538
— уравнения 372, 373, 470, 520
— формула 94
Колмогорова — Розанова теорема 284
Колмогорова — Чепмена уравнение 169, 173, 340, 359, 468
Компактность слабая 453
— — семейства распределений 54
Корреляции коэффициент 29
Корреляционная матрица 30, 306
— функция 207, 208, 241, 286
— — взаимная 208
— — эмпирическая 625
Корреляционный оператор 455
Коши распределение 123
Коэффициент асимметрии 108
— диффузии 469
— доверия 559
— корреляции 29
— обрыва 329
— переноса 469
— сильного перемешивания 283
— усиления фильтра 256
— эксцесса 108
Крамера ряд 88
— условие 87
Крамера — Рао неравенство 553, 555
Крикиберга разложение 319
Критерий 542
— Неймана — Пирсона 544, 583
— Неймана — Фишера факторизационный 528
— непараметрический 548
— — порядковый 549
— несмещенный 543
— отношения правдоподобия 545, 546
— последовательный отношения правдоподобия 549
— равномерно наиболее мощный 544
— рандомизированный 544
— серий Вальда — Вольфовица 549
— согласия 542
— состоятельный 544
— χ^2 (хи-квадрат) 547
Критическая область 542
— — подобная 543
— функция 544
Кумулянта 68, 410
Кунита — Ватанабе неравенство 331

Лапласа преобразование 60
— распределение 124
Леви — Линдеберга теорема 71, 78
Леви представление 93
— разложение 405
— расстояние 55
— теорема 332
Леви — Хинчина представление 93
Леви — Хинчина формула 407
Лестничные величины 152, 157, 159

Линдеберга — Феллера теорема 74, 78
Линейная интерполяция 274
— фильтрация 239, 264, 276
— экстраполяция (прогноз) 238, 264, 271
Локальное время 383
Ляпунова неравенство 106
— теорема 73

Марковский момент 176, 220, 348
Мартингал 311
— локально квадратически интегрируемый 330
— локальный 323
Мартингалов проблема 492
Математическое ожидание 27
— — условное 33, 34, 36
Матрица информационная 355
— ковариационная 30
— корреляционная 30, 306
— псевдообратная 575
— регрессионная 572
— стохастическая 192
— структурная 236
Медиана 107
Медленно меняющаяся функция 59, 63, 74, 103
— растущая на бесконечности функция 494
Мера 17
— пуассоновская 511
— спектральная 246, 292, 296
— стохастическая ортогональная 233, 236
Метод Байеса 561, 609
— Винера 266
— наименьших квадратов 570
— Яглома 267, 615
Миллса отношение 120
Минимаксная оценка 552
Минлоса — Сазонова теорема 455
Множество инвариантное 184
— непрерывности 452
— цилиндрическое 206
Мода 107
Момент 28
— абсолютный 28
— выборочный 539
— марковский 176, 220, 348
— обрыва 343
— остановки 220, 314
Момент остановки непредсказуемый 221
— — предсказуемый 221
— первого выхода 366
— — достижения 177, 415
— смешанный 29, 69
— факториальный 28
— центральный 28
Моментные функции 207, 285
Моментов проблема 106
Монотонный класс множеств 16
Муавра — Лапласа теорема 42, 71, 77

Наименьших квадратов метод 570
— — принцип 572
Независимые случайные величины 37
— события 37, 39
Неймана — Пирсона критерий 544, 583
Неймана — Фишера критерий факторизованный 528
Непрерывности аксиома 16
Неравенство Берри — Эссеена 82
— Буркхольдера 317
— Дуба 318
— Дубинса 318
— Колмогорова 47, 406
— Крамера — Рао 553, 555
— Кунита — Ватанабе 331
— Ляпунова 106
— Чебышева 44
Несмещенная оценка 532, 552, 610
Несовместимые события 12
Нуля и единицы закон 39, 354

Оператор Бляшке — Привалова 391
— инфинитезимальный 362, 363
— корреляционный 455
— характеристический 368
Отклик 570
Оцениваемая функция параметра 575
Оценка 527, 551
— асимптотически эффективная 556, 611
— Бартлетта 625
— Даниэля 625
— достаточная 554
— — полная 554

Оценка максимального правдоподобия 557, 609
— минимаксная 552
— МНК (метода наименьших квадратов) 572
— несмещенная 532, 552, 610
— по методу минимума χ^2 558
— — — моментов 556
— совместно-эффективная 556
— состоятельная 556, 605, 610
— среднеарифметическая 611
— Тьюки — Хэннинга 626
— эффективная 553, 610
Ошибка второго рода 543
— первого рода 543

Парето распределение 130
Парсеваля равенство 66
Паскаля распределение 111
Пересечение событий 12
Переходная функция импульсная 255, 265
Периодограмма 621
Пирсона распределение 133
Плотное семейство распределений (мер) 55, 452
Плотность 225, 583
— распределения 20, 23, 24, 106
— спектральная 247, 258, 260, 292, 296
Повторного логарифма закон 50, 410, 413, 422
Пойа распределение 113
Полиномы Чебышева — Эрмита 84
— Эрмита 478
Поллачека — Спицера тождество 159
Полная группа событий 12
Полной вероятности формула 32
Полугруппа операторов сжимающая 361
Потенциал 179, 319, 326
Поток σ-алгебр 220
Правило трех сигм 121
Правильно (регулярно) меняющаяся функция 103, 132
Правдоподобия отношение 529
— уравнения 557
— функция 526, 557
Преобразование, сохраняющее меру 279
— — — метрически транзитивное 281
Прогнозирование 263
Произведение событий 12
Производящая функция 56
Пространство вероятностное 17
— выборочное 524
— измеримое 17, 225
Процесс авторегрессии 262
— белого шума 258
— броуновского движения 211
— ветвящийся 427, 432
— — вырождающийся 430
— — докритический 430, 436, 441, 446
— — критический 430, 436, 441, 446
— — надкритический 430, 436, 441, 446
— — общий марковский 448
— — регулярный 433
— — с конечным числом типов частиц 438, 444
— — — — — — — — (не)периодический 441
— — — — — — — — (не)разложимый 440, 446
— винеровский 211, 422
— восстановления 145
— вполне измеримый (опциональный) 222
— Гальтона — Ватсона 427
— гибели 429
— детерминированный 269
— дифференцируемый в среднем квадратическом 230
— диффузионный 391, 469
— — канонический 391
— измеримый 212
— локально интегрируемый 326
— — ограниченной вариации 326
— марковский 339, 344
— — нормальный 352, 359
— — обрывающийся 351
— — однородный 358
— — прогрессивно измеримый 348
— — регулярный 378
— — стандартный 350, 367
— недетерминированный 269
— непрерывный в среднем квадратическом 230, 251
— Пуассона 210, 420
— предсказуемый 223
— прогрессивно измеримый 221
— размножения и гибели 378
— регулярный 269

Процесс с независимыми приращениями 218, 403
— — — — однородный 409
— — ортогональными приращениями 235, 244
— сепарабельный 215
— сингулярный 269
— скользящего среднего 262
— стационарный в узком смысле 277
— — — широком смысле 241
— — гауссовский 243
— стохастически непрерывный 208
— строго марковский 349, 366
— счетно порожденный 223
— точечный 335
— — маркированный 335
— чистого роста 211, 377
— эргодический 400
— — непериодический 402
— — периодический 401
Процессы, стационарно связанные 242
Прямые измерения неравноточны 574
— — равноточные 573
Пуассона процесс 210, 420
— распределение 21, 114
— — сложное 115
— теорема 43
Пуассоновская мера 511

Равномерная интегрируемость 45, 54, 315
— малость 72
Равных вероятностей эллипсы 138
Размах (широта) выборки 535
Распределение 19, 22, 454
— абсолютно непрерывное 20, 23, 105
— безгранично делимое 91, 96
— Бернулли 109
— бета 122
— биномиальное 21, 109
— Бореля — Таннера 116
— вырожденное 108
— гамма 121
— геометрическое 21, 112
— гипергеометрическое 112
— гиперэкспоненциальное 118
— Дирихле 139
— дискретное 19, 22, 105
Распределение Коши 123
— Лапласа 124
— логарифмическое 115
— логистическое 130
— логнормальное 129
— маргинальное 22
— непрерывное 20
— нормальное 24, 119, 138
— отрицательное биномиальное 111
— Парето 130
— Паскаля 111
— Пирсона 133
— Пойа 113
— показательное (экспоненциальное) 20, 118
— полиномиальное 135
— Пуассона 21, 114
— — сложное 115
— равномерное 20, 24, 116, 136
— решетчатое 20
— Салема 144
— сингулярное 105, 143
— совместное 22
— Стьюдента 127, 140
— треугольное 117
— Уишарта 140
— унимодальное 107
— устойчивое 95, 103, 141
— Фишера — Снедекора 127
— экспоненциальное 20, 118
— Эрланга 121
— χ (хи) 126
— χ^2 (хи-квадрат) 124
Распределения конечномерные 205, 454
Расстояние Леви 55
— по вариации 55
Регрессии теоретической функция 570
Регрессионная матрица 572
— переменная 570
Регрессия полиномиально-тригонометрическая 618
Регрессор 570
Резольвента 362
Решающая функция 533
Решение стохастического дифференциального уравнения 483, 517
— — — — сильное 498
— — — — слабое 498

Салема распределение 144

Свертка мер 89
Семиинварианты 68
Семимартингал 329
— специальный 329
Сингулярные меры 225, 582
Скорохода метрика 465
Случайная величина 19, 454
— — гильбертова 229
— — несобственная 105
— замена времени 388
— функция гильбертова 229
— — — дифференцируемая 230
— — — непрерывная 230
Случайное блуждание 150, 164
— — (не)возвратное 151
— — полунепрерывное 156
— множество 221
— поле 285
— — гауссовское 287
— — измеримое 289
— — изотропное 301, 302, 309
— — обобщенное 304
— — — однородное 305
— — — с однородными приращениями 308
— — однородное 291
— — с независимыми приращениями 288
— — сепарабельное 290
— — стохастически непрерывное 289
Случайный процесс 205
— эксперимент 11
— элемент 163
Смирнова статистика 538, 548
— теорема 538
Событие 11, 12
— достоверное 12
— невозможное 12
— противоположное 12
— элементарное 12
Состояние возвратное 194
— — нулевое 197
— — положительное 197
— невозвратное 194
— несущественное 193
— существенное 193
Состояния сообщающиеся 193
Сохраняющее меру преобразование 279
Спектральная мера 246, 292, 296
— плотность 247, 258, 260, 292, 296
— функция 246, 301
Спектральная функция эмпирическая 622
Спектральное окно 623
— представление 248
Среднеквадратическое отклонение 29
Статистика 527
— достаточная 528
— — минимальная 529
— Колмогорова 537, 548
— полная 530
— порядковая (ранговая) 534
— Смирнова 538, 548
Статистическая структура 524
— — доминируемая 524
Стационарная мера 179
Стационарные вероятности 378
Стохастическая матрица 192
— мера ортогональная 233
— — — векторная 236
Стохастический интеграл 233, 237, 332, 472
Стохастическое дифференциальное уравнение 483, 516, 522
Строгая марковость 177
Строго марковский процесс 349, 366
Структурная матрица 236
— функция 233
Стьюдента распределение 127, 140
Субгармоническая функция 179
Субмартингал 311
Сумма событий 12
Супергармоническая функция 179
Супермартингал 311
Сходимость в основном 53
— в среднем 45
— — — квадратическом 45, 229
— по вероятности 41
— почти наверное (с вероятностью 1) 48
— слабая 54, 451

Табу-вероятности 199
Теорема Берри — Эссеена 82
— Биркгофа 280
— Бореля — Кантелли 39
— Бохнера — Хинчина 65, 245
— восстановления 146
— — узловая 147
— Гаусса — Маркова 576
— Гливенко 537
— Колмогорова 206, 538

Теорема Колмогорова—Розанова 284
— Леви 332
— Леви — Линдеберга 71, 78
— Линдеберга — Феллера 74, 78
— локальная предельная 80, 81
— Ляпунова 73
— Минлоса — Сазонова 455
— Муавра — Лапласа 42, 71, 72
— непрерывности 59, 62, 66
— о представлении 62
— — трех рядах 52, 457
— обращения 60, 65
— Пуассона 43
— Смирнова 538
— сравнения 505
— тауберова 59, 63
— центральная предельная 70, 76, 79, 191, 203, 284, 458
— Ченцова 290
— эргодическая 202, 253
Тождество Вальда 154
— Поллачека — Спицера 159
— факторизационное 156
Траектории процесса 206
Тренд 617
Тьюки — Хэннинга 626

Уишарта распределение 140
Умножения вероятностей формула 32
Уравнение восстановления 146
— Колмогорова 372, 373, 470, 520
— Колмогорова — Чепмена 169, 173, 340, 359, 468
— Фоккера — Планка 470
Уровень значимости 543
Условие Деблина 183
— Дини 502
Условная вероятность 32, 35
Условное математическое ожидание 33, 34, 36
— распределение 33

Фаза фильтра 256
Фазовое пространство 163, 205, 344
Факторизации компоненты 156, 157
Факторизационные тождества 156
Фильтр линейный 255
Фильтр полосовой 256
— физически осуществимый 257
Фильтрация 263
— линейная 239, 264, 276
Фишера — Снедекора распределение 127
Фоккера — Планка уравнение 470
Формула Байеса 32
— Ито 336, 477, 515
— Колмогорова 94
— Леви — Хинчина 407
— умножения вероятностей 32
— Хинчина — Поллачека 159
Функционал аддитивный 379
— граничный 161
— мультипликативный 353, 378
Функция автокорреляционная 242
— восстановления 146, 151
— вполне монотонная 62
— выборочная 206, 285
— гармоническая 179, 181, 393
— измеримая 19
— ковариационная 241
— корреляционная 207, 208, 241, 286
— — взаимная 208
— медленно меняющаяся 59, 63, 74, 103
— — растущая на бесконечности 494
— моментная 207, 285
— параметра оцениваемая 575
— положительно определенная 64, 229
— правдоподобия 526, 577
— правильно (регулярно) меняющаяся 103, 132
— производящая 56
— распределения 18, 19, 22, 105
— — эмпирическая 537
— решающая 533
— спектральная 246, 301
— структурная 233
— субгармоническая 179
— супергармоническая 179
— характеристическая 64, 69
— центрирующая 405
— эксцессивная 179, 195, 381

Характеристика взаимная 330
— квадратическая 331

Характеристика фильтра частотная 255, 266
Характеристическая функция 64, 69
Характеристический оператор 368
— функционал 207, 455
Хинчина — Поллачека формула 159

Центрирующая функция 405
Цепь Маркова 163, 171
— — возвратная по Харрису 188
— — — нулевая 190
— — — положительная 190
— — неприводимая 193
— — однородная 172, 174
— — стационарная 179
Цилиндрическое множество 206

Частота 13, 534
Частóтная характеристика фильтра 255, 266
Чебышева неравенство 44
Чебышева — Эрмита полиномы 84
Ченцова теорема 290

Шаг распределения 21

Эквивалентность мер 583
Эскпоненциальная структура 526
Эксцесса коэффициент 108
Эксцессивная функция 179, 195, 381
Эллипсоид рассеивания 555
Эллипсы равных вероятностей 138
Эмпирическая функция корреляционная 625
— — распределения 537
— — спектральная 622
Эрланга распределение 121
Эрмита полиномы 478
Эффективная оценка 553
Эффективность 553, 556
— асимптотическая 556

Яглома метод 267, 615
σ-алгебра событий 16