Е. С. ВЕНТЦЕЛЬ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ



Е. С. ВЕНТЦЕЛЬ

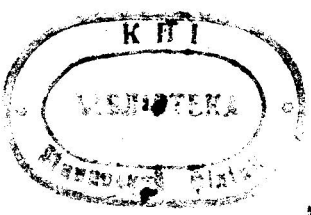
[illegible]

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

ИЗДАНИЕ ЧЕТВЕРТОЕ, СТЕРЕОТИПНОЕ

Допущено Министерством
высшего и среднего специального образования СССР
в качестве учебника
для высших технических учебных заведений

79871



АБОНЕМ[illegible]2

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1969

75

517.8
В 29
УДК 519.21 (075.8)

АННОТАЦИЯ

Книга представляет собой учебник, предназначенный для лиц, знакомых с математикой в объеме обычного втузовского курса и интересующихся техническими приложениями теории вероятностей, в частности теорией стрельбы. Книга представляет также интерес для инженеров других специальностей, которым приходится применять теорию вероятностей в их практической деятельности.

От других учебников, предназначенных для той же категории читателей, книга отличается большим вниманием к важным для приложений новым ветвям теории вероятностей (например, теории вероятностных процессов, теории информации, теории массового обслуживания и др.).

Елена Сергеевна Вентцель

Теория вероятностей

М., 1969 г., 576 стр. с илл.

Редакторы *А. П. Баева, В. В. Донченко*

Техн. редактор *Л. А. Пыжова* Корректор *К. В. Булатова*

Печать с матриц. Подписано к печати 4/VII 1969 г. Бумага 60×90/16. Физ. печ. л. 36. Условн. печ. л. 36. Уч.-изд. л. 33,80. Тираж 150 000 экз. Цена книги 1 р. 28 к. Заказ 196.

Издательство «Наука».
Главная редакция физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Ордена Трудового Красного Знамени Первая Образцовая типография имени А. А. Жданова Главполиграфпрома Комитета по печати при Совете Министров СССР
Москва, М-54, Валовая, 28

2-2-3
4-69

ОГЛАВЛЕНИЕ

Предисловие ко второму изданию 8
Предисловие к первому изданию 9

Глава 1. Введение 11

1.1. Предмет теории вероятностей 11
1.2. Краткие исторические сведения 17

Глава 2. Основные понятия теории вероятностей 23

2.1. Событие. Вероятность события 23
2.2. Непосредственный подсчет вероятностей 24
2.3. Частота, или статистическая вероятность, события 28
2.4. Случайная величина 32
2.5. Практически невозможные и практически достоверные события. Принцип практической уверенности 34

Глава 3. Основные теоремы теории вероятностей 37

3.1. Назначение основных теорем. Сумма и произведение событий 37
3.2. Теорема сложения вероятностей 40
3.3. Теорема умножения вероятностей 45
3.4. Формула полной вероятности 54
3.5. Теорема гипотез (формула Бейеса) 56

Глава 4. Повторение опытов 59

4.1. Частная теорема о повторении опытов 59
4.2. Общая теорема о повторении опытов 61

Глава 5. Случайные величины и их законы распределения 67

5.1. Ряд распределения. Многоугольник распределения 67
5.2. Функция распределения 72
5.3. Вероятность попадания случайной величины на заданный участок 78
5.4. Плотность распределения 80
5.5. Числовые характеристики случайных величин. Их роль и назначение 84
5.6. Характеристики положения (математическое ожидание, мода, медиана) 85
5.7. Моменты. Дисперсия. Среднее квадратическое отклонение . . 92
5.8. Закон равномерной плотности 103
5.9. Закон Пуассона 106

Глава 6. Нормальный закон распределения 115
6.1. Нормальный закон и его параметры 116
6.2. Моменты нормального распределения 120
6.3. Вероятность попадания случайной величины, подчиненной нормальному закону, на заданный участок. Нормальная функция распределения 122
6.4. Вероятное (срединное) отклонение 127

Глава 7. Определение законов распределения случайных величин на основе опытных данных 131
7.1. Основные задачи математической статистики 131
7.2. Простая статистическая совокупность. Статистическая функция распределения 133
7.3. Статистический ряд. Гистограмма 136
7.4. Числовые характеристики статистического распределения . . 139
7.5. Выравнивание статистических рядов 143
7.6. Критерии согласия 149

Глава 8. Системы случайных величин 159
8.1. Понятие о системе случайных величин 159
8.2. Функция распределения системы двух случайных величин . . 160
8.3. Плотность распределения системы двух случайных величин 163
8.4. Законы распределения отдельных величин, входящих в систему. Условные законы распределения 168
8.5. Зависимые и независимые случайные величины 171
8.6. Числовые характеристики системы двух случайных величин. Корреляционный момент. Коэффициент корреляции 175
8.7. Система произвольного числа случайных величин 182
8.8. Числовые характеристики системы нескольких случайных величин 184

Глава 9. Нормальный закон распределения для системы случайных величин 188
9.1. Нормальный закон на плоскости 188
9.2. Эллипсы рассеивания. Приведение нормального закона к каноническому виду 193
9.3. Вероятность попадания в прямоугольник со сторонами, параллельными главным осям рассеивания 196
9.4. Вероятность попадания в эллипс рассеивания 198
9.5. Вероятность попадания в область произвольной формы . . . 202
9.6. Нормальный закон в пространстве трех измерений. Общая запись нормального закона для системы произвольного числа случайных величин 205

Глава 10. Числовые характеристики функций случайных величин 210
10.1. Математическое ожидание функции. Дисперсия функции . . 210
10.2. Теоремы о числовых характеристиках 219
10.3. Применения теорем о числовых характеристиках 230

Глава 11. Линеаризация функций 252
11.1. Метод линеаризации функций случайных аргументов 252
11.2. Линеаризация функции одного случайного аргумента 253
11.3. Линеаризация функции нескольких случайных аргументов 255
11.4. Уточнение результатов, полученных методом линеаризации 259

Глава 12. Законы распределения функций случайных аргументов 263

12.1. Закон распределения монотонной функции одного случайного аргумента 263
12.2. Закон распределения линейной функции от аргумента, подчиненного нормальному закону 266
12.3. Закон распределения немонотонной функции одного случайного аргумента 267
12.4. Закон распределения функции двух случайных величин . . . 269
12.5. Закон распределения суммы двух случайных величин. Композиция законов распределения 271
12.6. Композиция нормальных законов 275
12.7. Линейные функции от нормально распределенных аргументов 279
12.8. Композиция нормальных законов на плоскости 280

Глава 13. Предельные теоремы теории вероятностей 286

13.1. Закон больших чисел и центральная предельная теорема . . 286
13.2. Неравенство Чебышева 287
13.3. Закон больших чисел (теорема Чебышева) 290
13.4. Обобщенная теорема Чебышева. Теорема Маркова 292
13.5. Следствия закона больших чисел: теоремы Бернулли и Пуассона 295
13.6. Массовые случайные явления и центральная предельная теорема 297
13.7. Характеристические функции 299
13.8. Центральная предельная теорема для одинаково распределенных слагаемых 302
13.9. Формулы, выражающие центральную предельную теорему и встречающиеся при ее практическом применении 306

Глава 14. Обработка опытов 312

14.1. Особенности обработки ограниченного числа опытов. Оценки для неизвестных параметров закона распределения 312
14.2. Оценки для математического ожидания и дисперсии 314
14.3. Доверительный интервал. Доверительная вероятность 317
14.4. Точные методы построения доверительных интервалов для параметров случайной величины, распределенной по нормальному закону 324
14.5. Оценка вероятности по частоте 330
14.6. Оценки для числовых характеристик системы случайных величин 339
14.7. Обработка стрельб 347
14.8. Сглаживание экспериментальных зависимостей по методу наименьших квадратов 351

Глава 15. Основные понятия теории случайных функций 370

15.1. Понятие о случайной функции 370
15.2. Понятие о случайной функции как расширение понятия о системе случайных величин. Закон распределения случайной функции 374
15.3. Характеристики случайных функций 377
15.4. Определение характеристик случайной функции из опыта . . 383
15.5. Методы определения характеристик преобразованных случайных функций по характеристикам исходных случайных функций 385

15.6. Линейные и нелинейные операторы. Оператор динамической системы 388
15.7. Линейные преобразования случайных функций 393
15.8. Сложение случайных функций 399
15.9. Комплексные случайные функции 402

Глава 16. Канонические разложения случайных функций 406

16.1. Идея метода канонических разложений. Представление случайной функции в виде суммы элементарных случайных функций 406
16.2. Каноническое разложение случайной функции 410
16.3. Линейные преобразования случайных функций, заданных каноническими разложениями 414

Глава 17. Стационарные случайные функции 419

17.1. Понятие о стационарном случайном процессе 419
17.2. Спектральное разложение стационарной случайной функции на конечном участке времени. Спектр дисперсий 427
17.3. Спектральное разложение стационарной случайной функции на бесконечном участке времени. Спектральная плотность стационарной случайной функции 431
17.4. Спектральное разложение случайной функции в комплексной форме 438
17.5. Преобразование стационарной случайной функции стационарной линейной системой 447
17.6. Применения теории стационарных случайных процессов к решению задач, связанных с анализом и синтезом динамических систем 454
17.7. Эргодическое свойство стационарных случайных функций 457
17.8. Определение характеристик эргодической стационарной случайной функции по одной реализации 462

Глава 18. Основные понятия теории информации 468

18.1. Предмет и задачи теории информации 468
18.2. Энтропия как мера степени неопределенности состояния физической системы 469
18.3. Энтропия сложной системы. Теорема сложения энтропий . . 475
18.4. Условная энтропия. Объединение зависимых систем 477
18.5. Энтропия и информация 481
18.6. Частная информация о системе, содержащаяся в сообщении о событии. Частная информация о событии, содержащаяся в сообщении о другом событии 489
18.7. Энтропия и информация для систем с непрерывным множеством состояний 493
18.8. Задачи кодирования сообщений. Код Шеннона — Фэно . . . 502
18.9. Передача информации с искажениями. Пропускная способность канала с помехами 509

Глава 19. Элементы теории массового обслуживания 515

19.1. Предмет теории массового обслуживания 515
19.2. Случайный процесс со счетным множеством состояний . . 517
19.3. Поток событий. Простейший поток и его свойства 520
19.4. Нестационарный пуассоновский поток 527

19. 5. Поток с ограниченным последействием (поток Пальма) . . 529
19. 6. Время обслуживания 534
19. 7. Марковский случайный процесс 537
19. 8. Система массового обслуживания с отказами. Уравнения Эрланга 540
19. 9. Установившийся режим обслуживания. Формулы Эрланга . 544
19.10. Система массового обслуживания с ожиданием 548
19.11. Система смешанного типа с ограничением по длине очереди 557

Приложение. Таблицы 561

Литература 573

Предметный указатель 574

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

В настоящее издание по сравнению с первым внесены следующие изменения и дополнения: материал, касающийся теоремы Ляпунова и связанного с ней метода характеристических функций (прежние nn° 13.7—13.11) сокращен и изложен в более простой форме в новых nn° 13.7—13.9; изменено изложение вопросов, связанных с обработкой опытов (глава 14), в связи с чем вместо прежних nn° 14.1 и 14.2 вставлены новые nn° 14.1—14.5, а прежние nn° 14.3—14.5 сохранены под номерами 14.6—14.8 соответственно. Несколько изменен n° 5.9, посвященный закону Пуассона. Добавлены две новые главы: глава 18 «Основные понятия теории информации» и глава 19 «Элементы теории массового обслуживания» — касающиеся двух важных ветвей теории вероятностей, до сих пор почти не освещавшихся в учебной литературе. Изменен состав таблиц в «Приложении». Изменены некоторые примеры.

Автор выражает глубокую благодарность академику АН УССР Б. В. Гнеденко за ряд ценных указаний.

Е. Вентцель

В настоящем, третьем издании существенных изменений нет. Произведены незначительные исправления и заменены некоторые примеры и задачи. Некоторые таблицы заменены другими. Слегка изменены в методическом отношении отдельные параграфы книги, например n°n° 5.9, 6.3, 9.4, 9.5 и 9.6. Существенно обновлено содержание n° 10.3.

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Настоящая книга написана на базе лекций по теории вероятностей, читанных автором в течение ряда лет слушателям Военно-воздушной инженерной академии им. Н. Е. Жуковского, а также учебника автора по тому же предмету, изданного Академией ограниченным тиражом в 1952 г.

В настоящем издании первоначальный текст учебника подвергся весьма значительной переработке.

Книга рассчитана, в основном, на инженера, имеющего математическую подготовку в объеме обычного курса высших технических учебных заведений. При составлении книги автор ставил себе задачу изложить предмет наиболее просто и наглядно, не связывая себя рамками полной математической строгости. В связи с этим отдельные положения приводятся без доказательства (раздел о доверительных границах и доверительных вероятностях; теорема А. Н. Колмогорова, относящаяся к критерию согласия, и некоторые другие); некоторые положения доказываются не вполне строго (теорема умножения законов распределения; правила преобразования математического ожидания и корреляционной функции при интегрировании и дифференцировании случайной функции и др.).

Применяемый математический аппарат, в основном, не выходит за рамки нормального курса высшей математики, излагаемого в высших технических учебных заведениях; там, где автору приходится пользоваться менее общеизвестными понятиями (например, понятием линейного оператора, матрицы, квадратичной формы и т. д.), эти понятия поясняются.

Книга снабжена большим количеством примеров, в ряде случаев — примерами расчетного характера, в которых применение излагаемых методов иллюстрируется на конкретном практическом материале и

доводится до численного результата. В связи с целевым назначением курса значительная часть примеров взята из области авиационной техники, воздушной стрельбы, бомбометания и теории боеприпасов. Ряд примеров относится к общей теории стрельбы. С этой точки зрения книга может быть особенно полезной для специалистов в области наземной, зенитной и морской артиллерии. Однако, несмотря на несколько специфический подбор примеров, иллюстративный материал, помещенный в книге, вполне понятен и для инженеров, работающих в других областях техники.

Автор выражает глубокую благодарность профессору Е. Б. Дынкину и профессору В. С. Пугачеву за ряд ценных указаний.

Е. Вентцель

ГЛАВА 1

ВВЕДЕНИЕ

1.1. Предмет теории вероятностей

Теория вероятностей есть математическая наука, изучающая закономерности в случайных явлениях.

Условимся, что́ мы будем понимать под «случайным явлением».

При научном исследовании различных физических и технических задач часто приходится встречаться с особого типа явлениями, которые принято называть случайными. *Случайное явление* — это такое явление, которое при неоднократном воспроизведении одного и того же опыта протекает каждый раз несколько по-иному.

Приведем примеры случайных явлений.

1. Производится стрельба из орудия, установленного под заданным углом к горизонту (рис. 1.1.1).

Пользуясь методами внешней баллистики (науки о движении снаряда в воздухе), можно найти теоретическую траекторию снаряда (кривая K на рис. 1.1.1). Эта траектория вполне определяется условиями стрельбы: начальной скоростью снаряда v_0, углом бросания θ_0 и баллистическим коэффициентом снаряда c. Фактическая траектория каждого отдельного снаряда неизбежно несколько отклоняется от теоретической за счет совокупного влияния многих факторов. Среди этих факторов можно, например, назвать: ошибки изготовления снаряда, отклонение веса заряда от номинала, неоднородность структуры заряда, ошибки установки ствола в заданное положение, метеорологические условия и т. д. Если произвести несколько выстрелов при неизменных основных условиях (v_0, θ_0, c), мы получим не одну теоретическую траекторию, а целый пучок или «сноп» траекторий, образующий так называемое «рассеивание снарядов».

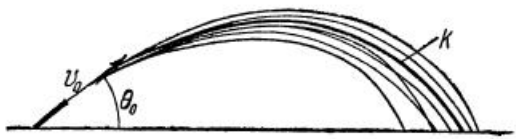


Рис. 1.1.1.

2. Одно и то же тело несколько раз взвешивается на аналитических весах; результаты повторных взвешиваний несколько отличаются друг от друга. Эти различия обусловлены влиянием многих

второстепенных факторов, сопровождающих операцию взвешивания, таких как положение тела на чашке весов, случайные вибрации аппаратуры, ошибки отсчета показаний прибора и т. д.

3. Самолет совершает полет на заданной высоте; теоретически он летит горизонтально, равномерно и прямолинейно. Фактически полет сопровождается отклонениями центра массы самолета от теоретической траектории и колебаниями самолета около центра массы. Эти отклонения и колебания являются случайными и связаны с турбулентностью атмосферы; от раза к разу они не повторяются.

4. Производится ряд подрывов осколочного снаряда в определенном положении относительно цели. Результаты отдельных подрывов несколько отличаются друг от друга: меняются общее число осколков, взаимное расположение их траекторий, вес, форма и скорость каждого отдельного осколка. Эти изменения являются случайными и связаны с влиянием таких факторов, как неоднородность металла корпуса снаряда, неоднородность взрывчатого вещества, непостоянство скорости детонации и т. п. В связи с этим различные подрывы, осуществленные, казалось бы, в одинаковых условиях, могут приводить к различным результатам: в одних подрывах цель будет поражена осколками, в других — нет.

Все приведенные примеры рассмотрены здесь под одним и тем же углом зрения: подчеркнуты случайные вариации, неодинаковые результаты ряда опытов, основные условия которых остаются неизменными. Эти вариации всегда связаны с наличием каких-то второстепенных факторов, влияющих на исход опыта, но не заданных в числе его основных условий. Основные условия опыта, определяющие в общих и грубых чертах его протекание, сохраняются неизменными; второстепенные — меняются от опыта к опыту и вносят случайные различия в их результаты.

Совершенно очевидно, что в природе нет ни одного физического явления, в котором не присутствовали бы в той или иной мере элементы случайности. Как бы точно и подробно ни были фиксированы условия опыта, невозможно достигнуть того, чтобы при повторении опыта результаты полностью и в точности совпадали.

Случайные отклонения неизбежно сопутствуют любому закономерному явлению. Тем не менее в ряде практических задач этими случайными элементами можно пренебречь, рассматривая вместо реального явления его упрощенную схему, «модель», и предполагая, что в данных условиях опыта явление протекает вполне определенным образом. При этом из бесчисленного множества факторов, влияющих на данное явление, выделяются самые главные, основные, решающие; влиянием остальных, второстепенных факторов просто пренебрегают. Такая схема изучения явлений постоянно применяется в физике, механике, технике. При пользовании этой схемой для решения любой задачи прежде всего выделяется основной круг учи-

тываемых условий и выясняется, на какие параметры задачи они влияют; затем применяется тот или иной математический аппарат (например, составляются и интегрируются дифференциальные уравнения, описывающие явление); таким образом выявляется основная закономерность, свойственная данному явлению и дающая возможность предсказать результат опыта по его заданным условиям. По мере развития науки число учитываемых факторов становится все больше; явление исследуется подробнее; научный прогноз становится точнее.

Однако для решения ряда вопросов описанная схема — классическая схема так называемых «точных наук» — оказывается плохо приспособленной. Существуют такие задачи, где интересующий нас исход опыта зависит от столь большого числа факторов, что практически невозможно зарегистрировать и учесть все эти факторы. Это — задачи, в которых многочисленные второстепенные, тесно переплетающиеся между собой случайные факторы играют заметную роль, а вместе с тем число их так велико и влияние столь сложно, что применение классических методов исследования себя не оправдывает.

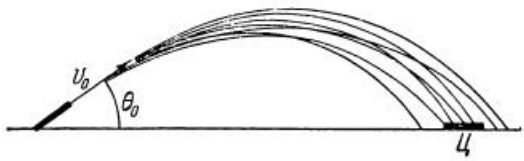


Рис. 1.1.2.

Рассмотрим пример. Производится стрельба по некоторой цели $Ц$ из орудия, установленного под углом θ_0 к горизонту (рис. 1.1.2). Траектории снарядов, как было указано выше, не совпадают между собой; в результате точки падения снарядов на земле рассеиваются. Если размеры цели велики по сравнению с областью рассеивания, то этим рассеиванием, очевидно, можно пренебречь: при правильной установке орудия любой выпущенный снаряд попадает в цель. Если же (как обычно и бывает на практике) область рассеивания снарядов превышает размеры цели, то некоторые из снарядов в связи с влиянием случайных факторов в цель не попадут. Возникает ряд вопросов, например: какой процент выпущенных снарядов в среднем попадает в цель? Сколько нужно потратить снарядов для того, чтобы достаточно надежно поразить цель? Какие следует принять меры для уменьшения расхода снарядов?

Чтобы ответить на подобные вопросы, обычная схема точных наук оказывается недостаточной. Эти вопросы органически связаны со случайной природой явления; для того чтобы на них ответить, очевидно, нельзя просто пренебречь случайностью, — надо изучить случайное явление рассеивания снарядов с точки зрения закономерностей, присущих ему именно как случайному явлению. Надо исследовать закон, по которому распределяются точки падения снарядов;

нужно выяснить случайные причины, вызывающие рассеивание, сравнить их между собой по степени важности и т. д.

Рассмотрим другой пример. Некоторое техническое устройство, например система автоматического управления, решает определенную задачу в условиях, когда на систему непрерывно воздействуют случайные помехи. Наличие помех приводит к тому, что система решает задачу с некоторой ошибкой, в ряде случаев выходящей за пределы допустимой. Возникают вопросы: как часто будут появляться такие ошибки? Какие следует принять меры для того, чтобы практически исключить их возможность?

Чтобы ответить на такие вопросы, необходимо исследовать природу и структуру случайных возмущений, воздействующих на систему, изучить реакцию системы на такие возмущения, выяснить влияние конструктивных параметров системы на вид этой реакции.

Все подобные задачи, число которых в физике и технике чрезвычайно велико, требуют изучения не только основных, главных закономерностей, определяющих явление в общих чертах, но и анализа случайных возмущений и искажений, связанных с наличием второстепенных факторов и придающих исходу опыта при заданных условиях элемент неопределенности.

Какие же существуют пути и методы для исследования случайных явлений?

С чисто теоретической точки зрения те факторы, которые мы условно назвали «случайными», в принципе ничем не отличаются от других, которые мы выделили в качестве «основных». Теоретически можно неограниченно повышать точность решения каждой задачи, учитывая все новые и новые группы факторов: от самых существенных до самых ничтожных. Однако практически такая попытка одинаково подробно и тщательно проанализировать влияние решительно всех факторов, от которых зависит явление, привела бы только к тому, что решение задачи, в силу непомерной громоздкости и сложности, оказалось бы практически неосуществимым и к тому же не имело бы никакой познавательной ценности.

Например, теоретически можно было бы поставить и решить задачу об определении траектории вполне определенного снаряда, с учетом всех конкретных погрешностей его изготовления, точного веса и конкретной структуры данного, вполне определенного порохового заряда при точно определенных метеорологических данных (температура, давление, влажность, ветер) в каждой точке траектории. Такое решение не только было бы необозримо сложным, но и не имело бы никакой практической ценности, так как относилось бы только к данному конкретному снаряду и заряду в данных конкретных условиях, которые практически больше не повторятся.

Очевидно, должна существовать принципиальная разница в методах учета основных, решающих факторов, определяющих в главных

чертах течение явления, и вторичных, второстепенных факторов, влияющих на течение явления в качестве «погрешностей» или «возмущений». Элемент неопределенности, сложности, многопричинности, присущий случайным явлениям, требует создания специальных методов для изучения этих явлений.

Такие методы и разрабатываются в теории вероятностей. Ее предметом являются специфические закономерности, наблюдаемые в случайных явлениях.

Практика показывает, что, наблюдая в совокупности массы однородных случайных явлений, мы обычно обнаруживаем в них вполне определенные закономерности, своего рода *устойчивости*, свойственные именно массовым случайным явлениям.

Например, если много раз подряд бросать монету, частота появления герба (отношение числа появившихся гербов к общему числу бросаний) постепенно стабилизируется, приближаясь к вполне определенному числу, именно к $\frac{1}{2}$. Такое же свойство «устойчивости частот» обнаруживается и при многократном повторении любого другого опыта, исход которого представляется заранее неопределенным, случайным. Так, при увеличении числа выстрелов частота попадания в некоторую цель тоже стабилизируется, приближаясь к некоторому постоянному числу.

Рассмотрим другой пример. В сосуде заключен какой-то объем газа, состоящий из весьма большого числа молекул. Каждая молекула за секунду испытывает множество столкновений с другими молекулами, многократно меняет скорость и направление движения; траектория каждой отдельной молекулы случайна. Известно, что давление газа на стенку сосуда обусловлено совокупностью ударов молекул об эту стенку. Казалось бы, если траектория каждой отдельной молекулы случайна, то и давление на стенку сосуда должно было бы изменяться случайным и неконтролируемым образом; однако это не так. Если число молекул достаточно велико, то давление газа практически не зависит от траекторий отдельных молекул и подчиняется вполне определенной и очень простой закономерности. Случайные особенности, свойственные движению каждой отдельной молекулы, в массе взаимно компенсируются; в результате, несмотря на сложность и запутанность отдельного случайного явления, мы получаем весьма простую закономерность, справедливую для массы случайных явлений. Отметим, что именно *массовость* случайных явлений обеспечивает выполнение этой закономерности; при ограниченном числе молекул начинают сказываться случайные отклонения от закономерности, так называемые флуктуации.

Рассмотрим еще один пример. По некоторой мишени производится один за другим ряд выстрелов; наблюдается распределение точек попадания на мишени. При ограниченном числе выстрелов точки

попадания распределяются по мишени в полном беспорядке, без какой-либо видимой закономерности. По мере увеличения числа выстрелов в расположении точек попадания начинает наблюдаться некоторая закономерность; эта закономерность проявляется тем отчетливее, чем большее количесто выстрелов произведено. Расположение точек попадания оказывается приблизительно симметричным относительно некоторой центральной точки: в центральной области группы пробоин они расположены гуще, чем по краям; при этом густота пробоин убывает по вполне определенному закону (так называемый «нормальный закон» или «закон Гаусса», которому будет уделено большое внимание в данном курсе).

Подобные специфические, так называемые «статистические», закономерности наблюдаются всегда, когда мы имеем дело с массой однородных случайных явлений. Закономерности, проявляющиеся в этой массе, оказываются практически независимыми от индивидуальных особенностей отдельных случайных явлений, входящих в массу. Эти отдельные особенности в массе как бы взаимно погашаются, нивелируются, и средний результат массы случайных явлений оказывается практически уже не случайным. Именно эта многократно подтвержденная опытом устойчивость массовых случайных явлений и служит базой для применения вероятностных (статистических) методов исследования. Методы теории вероятностей по природе приспособлены только для исследования массовых случайных явлений; они не дают возможности предсказать исход отдельного случайного явления, но дают возможность предсказать средний суммарный результат массы однородных случайных явлений, предсказать средний исход массы аналогичных опытов, конкретный исход каждого из которых остается неопределенным, случайным.

Чем большее количество однородных случайных явлений участвует в задаче, тем определеннее и отчетливее проявляются присущие им специфические законы, тем с большей уверенностью и точностью можно осуществлять научный прогноз.

Во всех случаях, когда применяются вероятностные методы исследования, цель их в том, чтобы, минуя слишком сложное (и зачастую практически невозможное) изучение отдельного явления, обусловленного слишком большим количеством факторов, обратиться непосредственно к законам, управляющим массами случайных явлений. Изучение этих законов позволяет не только осуществлять научный прогноз в своеобразной области случайных явлений, но в ряде случаев помогает целенаправленно влиять на ход случайных явлений, контролировать их, ограничивать сферу действия случайности, сужать ее влияние на практику.

Вероятностный, или статистический, метод в науке не противопоставляет себя классическому, обычному методу точных наук, а является его дополнением, позволяющим глубже анали-

зировать явление с учетом присущих ему элементов случайности.

Характерным для современного этапа развития естественных и технических наук является весьма широкое и плодотворное применение статистических методов во всех областях знания. Это вполне естественно, так как при углубленном изучении любого круга явлений неизбежно наступает этап, когда требуется не только выявление основных закономерностей, но и анализ возможных отклонений от них. В одних науках, в силу специфики предмета и исторических условий, внедрение статистических методов наблюдается раньше, в других — позже. В настоящее время нет почти ни одной естественной науки, в которой так или иначе не применялись бы вероятностные методы. Целые разделы современной физики (в частности, ядерная физика) базируются на методах теории вероятностей. Все шире применяются вероятностные методы в современной электротехнике и радиотехнике, метеорологии и астрономии, теории автоматического регулирования и машинной математике.

Обширное поле применения находит теория вероятностей в разнообразных областях военной техники: теория стрельбы и бомбометания, теория боеприпасов, теория прицелов и приборов управления огнем, аэронавигация, тактика и множество других разделов военной науки широко пользуются методами теории вероятностей и ее математическим аппаратом.

Математические законы теории вероятностей — отражение реальных статистических законов, объективно существующих в массовых случайных явлениях природы. К изучению этих явлений теория вероятностей применяет математический метод и по своему методу является одним из разделов математики, столь же логически точным и строгим, как другие математические науки.

79871

1.2. Краткие исторические сведения

Теория вероятностей, подобно другим математическим наукам, развилась из потребностей практики.

Начало систематического исследования задач, относящихся к массовым случайным явлениям, и появление соответствующего математического аппарата относятся к XVII веку. В начале XVII века знаменитый физик Галилей уже пытался подвергнуть научному исследованию ошибки физических измерений, рассматривая их как случайные и оценивая их вероятности. К этому же времени относятся первые попытки создания общей теории страхования, основанной на анализе закономерностей в таких массовых случайных явлениях, как заболеваемость, смертность, статистика несчастных случаев и т. д. Необходимость создания математического аппарата, специально приспособленного для анализа случайных явлений, вытекала из потребностей

БИБЛИОТЕКА

обработки и обобщения обширного статистического материала во всех областях науки.

Однако теория вероятностей как математическая наука сформировалась, в основном, не на материале указанных выше практических задач: эти задачи слишком сложны; в них законы, управляющие случайными явлениями, проступают недостаточно отчетливо и затушеваны многими осложняющими факторами. Необходимо было сначала изучить закономерности случайных явлений на более простом материале. Таким материалом исторически оказались так называемые «азартные игры». Эти игры с незапамятных времен создавались рядом поколений именно так, чтобы в них исход опыта был независим от поддающихся наблюдению условий опыта, был чисто случайным. Самое слово «азарт» (фр. «le hasard») означает «случай». Схемы азартных игр дают исключительные по простоте и прозрачности модели случайных явлений, позволяющие в наиболее отчетливой форме наблюдать и изучать управляющие ими специфические законы; а возможность неограниченно повторять один и тот же опыт обеспечивает экспериментальную проверку этих законов в условиях действительной массовости явлений. Вплоть до настоящего времени примеры из области азартных игр и аналогичные им задачи на «схему урн» широко употребляются при изучении теории вероятностей как упрощенные модели случайных явлений, иллюстрирующие в наиболее простом и наглядном виде основные законы и правила теории вероятностей.

Возникновение теории вероятностей в современном смысле слова относится к середине XVII века и связано с исследованиями Паскаля (1623—1662), Ферма (1601—1665) и Гюйгенса (1629—1695) в области теории азартных игр. В этих работах постепенно сформировались такие важные понятия, как вероятность и математическое ожидание; были установлены их основные свойства и приемы их вычисления. Непосредственное практическое применение вероятностные методы нашли прежде всего в задачах страхования. Уже с конца XVII века страхование стало производиться на научной математической основе. С тех пор теория вероятностей находит все более широкое применение в различных областях.

Крупный шаг вперед в развитии теории вероятностей связан с работами Якова Бернулли (1654—1705). Ему принадлежит первое доказательство одного из важнейших положений теории вероятностей — так называемого закона больших чисел.

Еще до Якова Бернулли многие отмечали как эмпирический факт ту особенность случайных явлений, которую можно назвать «свойством устойчивости частот при большом числе опытов». Было неоднократно отмечено, что при большом числе опытов, исход каждого из которых является случайным, относительная частота появления каждого данного исхода имеет тенденцию стабилизироваться, при-

ближаясь к некоторому определенному числу — вероятности этого исхода. Например, если много раз бросать монету, относительная частота появления герба приближается к $^1/_2$; при многократном бросании игральной кости частота появления грани с пятью очками приближается к $^1/_6$ и т. д. Яков Бернулли впервые дал теоретическое обоснование этому эмпирическому факту. Теорема Якова Бернулли — простейшая форма закона больших чисел — устанавливает связь между вероятностью события и частотой его появления; при достаточно большом числе опытов можно с практической достоверностью ожидать сколь угодно близкого совпадения частоты с вероятностью.

Другой важный этап в развитии теории вероятностей связан с именем Моавра (1667—1754). Этот ученый впервые ввел в рассмотрение и для простейшего случая обосновал своеобразный закон, очень часто наблюдаемый в случайных явлениях: так называемый нормальный закон (иначе — закон Гаусса). Нормальный закон, как мы увидим далее, играет исключительно важную роль в случайных явлениях. Теоремы, обосновывающие этот закон для тех или иных условий, носят в теории вероятностей общее название «центральной предельной теоремы».

Выдающаяся роль в развитии теории вероятностей принадлежит знаменитому математику Лапласу (1749—1827). Он впервые дал стройное и систематическое изложение основ теории вероятностей, дал доказательство одной из форм центральной предельной теоремы (теоремы Моавра — Лапласа) и развил ряд замечательных приложений теории вероятностей к вопросам практики, в частности к анализу ошибок наблюдений и измерений.

Значительный шаг вперед в развитии теории вероятностей связан с именем Гаусса (1777—1855), который дал еще более общее обоснование нормальному закону и разработал метод обработки экспериментальных данных, известный под названием «метода наименьших квадратов». Следует также отметить работы Пуассона (1781—1840), доказавшего более общую, чем у Якова Бернулли, форму закона больших чисел, а также впервые применившего теорию вероятностей к задачам стрельбы. С именем Пуассона связан один из законов распределения, играющий большую роль в теории вероятностей и ее приложениях.

Для всего XVIII и начала XIX века характерны бурное развитие теории вероятностей и повсеместное увлечение ею. Теория вероятностей становится «модной» наукой. Ее начинают применять не только там, где это применение правомерно, но и там, где оно ничем не оправдано. Для этого периода характерны многочисленные попытки применить теорию вероятностей к изучению общественных явлений, к так называемым «моральным» или «нравственным» наукам. Во множестве появились работы, посвященные вопросам судопроизводства, истории, политики, даже богословия, в которых применялся

аппарат теории вероятностей. Для всех этих псевдонаучных исследований характерен чрезвычайно упрощенный, механистический подход к рассматриваемым в них общественным явлениям. В основу рассуждения полагаются некоторые произвольно заданные вероятности (например, при рассмотрении вопросов судопроизводства склонность каждого человека к правде или лжи оценивается некоторой постоянной, одинаковой для всех людей вероятностью), и далее общественная проблема решается как простая арифметическая задача. Естественно, что все подобные попытки были обречены на неудачу и не могли сыграть положительной роли в развитии науки. Напротив, их косвенным результатом оказалось то, что примерно в 20-х — 30-х годах XIX века в Западной Европе повсеместное увлечение теорией вероятностей сменилось разочарованием и скептицизмом. На теорию вероятностей стали смотреть как на науку сомнительную, второсортную, род математического развлечения, вряд ли достойный серьезного изучения.

Замечательно, что именно в это время в России создается та знаменитая Петербургская математическая школа, трудами которой теория вероятностей была поставлена на прочную логическую и математическую основу и сделана надежным, точным и эффективным методом познания. Со времени появления этой школы развитие теории вероятностей уже теснейшим образом связано с работами русских, а в дальнейшем — советских ученых.

Среди ученых Петербургской математической школы следует назвать В. Я. Буняковского (1804—1889) — автора первого курса теории вероятностей на русском языке, создателя современной русской терминологии в теории вероятностей, автора оригинальных исследований в области статистики и демографии.

Учеником В. Я. Буняковского был великий русский математик П. Л. Чебышев (1821—1894). Среди обширных и разнообразных математических трудов П. Л. Чебышева заметное место занимают его труды по теории вероятностей. П. Л. Чебышеву принадлежит дальнейшее расширение и обобщение закона больших чисел. Кроме того, П. Л. Чебышев ввел в теорию вероятностей весьма мощный и плодотворный метод моментов.

Учеником П. Л. Чебышева был А. А. Марков (1856—1922), также обогативший теорию вероятностей открытиями и методами большой важности. А. А. Марков существенно расширил область применения закона больших чисел и центральной предельной теоремы, распространив их не только на независимые, но и на зависимые опыты. Важнейшей заслугой А. А. Маркова явилось то, что он заложил основы совершенно новой ветви теории вероятностей — теории случайных, или «стохастических», процессов. Развитие этой теории составляет основное содержание новейшей, современной теории вероятностей.

Учеником П. Л. Чебышева был и А. М. Ляпунов (1857 — 1918), с именем которого связано первое доказательство центральной предельной теоремы при чрезвычайно общих условиях. Для доказательства своей теоремы А. М. Ляпунов разработал специальный метод характеристических функций, широко применяемый в современной теории вероятностей.

Характерной особенностью работ Петербургской математической школы была исключительная четкость постановки задач, полная математическая строгость применяемых методов и наряду с этим тесная связь теории с непосредственными требованиями практики. Трудами ученых Петербургской математической школы теория вероятностей была выведена с задворков науки и поставлена как полноправный член в ряд точных математических наук. Условия применения ее методов были строго определены, а самые методы доведены до высокой степени совершенства.

Современное развитие теории вероятностей характерно всеобщим подъемом интереса к ней и резким расширением круга ее практических применений. За последние десятилетия теория вероятностей превратилась в одну из наиболее быстро развивающихся наук, теснейшим образом связанную с потребностями практики и техники. Советская школа теории вероятностей, унаследовав традиции Петербургской математической школы, занимает в мировой науке ведущее место.

Здесь мы назовем только некоторых крупнейших советских ученых, труды которых сыграли решающую роль в развитии современной теории вероятностей и ее практических приложений.

С. Н. Бернштейн разработал первую законченную аксиоматику теории вероятностей, а также существенно расширил область применения предельных теорем.

А. Я. Хинчин (1894 — 1959) известен своими исследованиями в области дальнейшего обобщения и усиления закона больших чисел, но главным образом своими исследованиями в области так называемых стационарных случайных процессов.

Ряд важнейших основополагающих работ в различных областях теории вероятностей и математической статистики принадлежит А. Н. Колмогорову. Он дал наиболее совершенное аксиоматическое построение теории вероятностей, связав ее с одним из важнейших разделов современной математики — метрической теорией функций. Особое значение имеют работы А. Н. Колмогорова в области теории случайных функций (стохастических процессов), которые в настоящее время являются основой всех исследований в данной области. Работы А. Н. Колмогорова, относящиеся к оценке эффективности легли в основу целого нового научного направления в теории стрельбы, переросшего затем в более широкую науку об эффективности боевых действий.

В. И. Романовский (1879 — 1954) и Н. В. Смирнов известны своими работами в области математической статистики, Е. Е. Слуцкий (1880 — 1948) — в теории случайных процессов, Б. В. Гнеденко — в области теории массового обслуживания, Е. Б. Дынкин — в области марковских случайных процессов, В. С. Пугачев — в области случайных процессов в применении к задачам автоматического управления.

Развитие зарубежной теории вероятностей в настоящее время также идет усиленными темпами в связи с настоятельными требованиями практики. Преимущественным вниманием пользуются, как и у нас, вопросы, относящиеся к случайным процессам. Значительные работы в этой области принадлежат, например, Н. Винеру, В. Феллеру, Д. Дубу. Важные работы по теории вероятностей и математической статистике принадлежат Р. Фишеру, Д. Нейману и Г. Крамеру.

За самые последние годы мы являемся свидетелями зарождения новых и своеобразных методов прикладной теории вероятностей, появление которых связано со спецификой исследуемых технических проблем. Речь идет, в частности, о таких новых дисциплинах, как «теория информации» и «теория массового обслуживания». Возникшие из непосредственных потребностей практики, эти разделы теории вероятностей приобретают общее теоретическое значение, а круг их приложений быстро увеличивается.

ГЛАВА 2

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

2.1. Событие. Вероятность события

Каждая наука, развивающая общую теорию какого-либо круга явлений, содержит ряд основных понятий, на которых она базируется. Таковы, например, в геометрии понятия точки, прямой, линии; в механике — понятия силы, массы, скорости, ускорения и т. д. Естественно, что не все основные понятия могут быть строго определены, так как определить понятие — это значит свести его к другим, более известным. Очевидно, процесс определения одних понятий через другие должен где-то заканчиваться, дойдя до самых первичных понятий, к которым сводятся все остальные и которые сами строго не определяются, а только поясняются.

Такие основные понятия существуют и в теории вероятностей. В качестве первого из них введем понятие *события*.

Под «событием» в теории вероятностей понимается всякий факт, который в результате опыта может произойти или не произойти.

Приведем несколько примеров событий:

A — появление герба при бросании монеты;

B — появление трех гербов при трехкратном бросании монеты;

C — попадание в цель при выстреле;

D — появление туза при вынимании карты из колоды;

E — обнаружение объекта при одном цикле обзора радиолокационной станции;

F — обрыв нити в течение часа работы ткацкого станка.

Рассматривая вышеперечисленные события, мы видим, что каждое из них обладает какой-то степенью возможности: одни — большей, другие — меньшей, причем для некоторых из этих событий мы сразу же можем решить, какое из них более, а какое менее возможно. Например, сразу видно, что событие A более возможно, чем B и D. Относительно событий C, E и F аналогичных выводов сразу сделать нельзя; для этого следовало бы несколько уточнить условия опыта. Так или иначе ясно, что каждое из таких событий обладает той или иной степенью возможности. Чтобы количественно сравнивать между собой события по степени их возможности, очевидно, нужно с каждым

событием связать определенное число, которое тем больше, чем более возможно событие. Такое число мы назовем *вероятностью события*.

Таким образом, мы ввели в рассмотрение второе основное понятие теории вероятностей — понятие вероятности события. Вероятность события есть численная мера степени объективной возможности этого события.

Заметим, что уже при самом введении понятия вероятности события мы связываем с этим понятием определенный практический смысл, а именно: на основании опыта мы считаем более вероятными те события, которые происходят чаще; менее вероятными — те события, которые происходят реже; мало вероятными — те, которые почти никогда не происходят. Таким образом, понятие вероятности события в самой своей основе связано с опытным, практическим понятием *частоты* события.

Сравнивая между собой различные события по степени их возможности, мы должны установить какую-то единицу измерения. В качестве такой единицы измерения естественно принять вероятность *достоверного события*, т. е. такого события, которое в результате опыта непременно должно произойти. Пример достоверного события — выпадение не более 6 очков при бросании одной игральной кости.

Если приписать достоверному событию вероятность, равную единице, то все другие события — возможные, но не достоверные — будут характеризоваться вероятностями, меньшими единицы, составляющими какую-то долю единицы.

Противоположностью по отношению к достоверному событию является *невозможное событие*, т. е. такое событие, которое в данном опыте не может произойти. Пример невозможного события — появление 12 очков при бросании одной игральной кости. Естественно приписать невозможному событию вероятность, равную нулю.

Таким образом, установлены единица измерения вероятностей — вероятность достоверного события — и диапазон изменения вероятностей любых событий — числа от 0 до 1.

2.2. Непосредственный подсчет вероятностей

Существует целый класс опытов, для которых вероятности их возможных исходов легко оценить непосредственно из условий самого опыта. Для этого нужно, чтобы различные исходы опыта обладали симметрией и в силу этого были объективно одинаково возможными.

Рассмотрим, например, опыт, состоящий в бросании игральной кости, т. е. симметричного кубика, на гранях которого нанесено различное число очков: от 1 до 6.

В силу симметрии кубика есть основания считать все шесть возможных исходов опыта одинаково возможными. Именно это дает нам

право предполагать, что при многократном бросании кости все шесть граней будут выпадать примерно одинаково часто. Это предположение для правильно выполненной кости действительно оправдывается на опыте; при многократном бросании кости каждая ее грань появляется примерно в одной шестой доле всех случаев бросания, причем отклонение этой доли от $\frac{1}{6}$ тем меньше, чем большее число опытов произведено. Имея в виду, что вероятность достоверного события принята равной единице, естественно приписать выпадению каждой отдельной грани вероятность, равную $\frac{1}{6}$. Это число характеризует некоторые объективные свойства данного случайного явления, а именно свойство симметрии шести возможных исходов опыта.

Для всякого опыта, в котором возможные исходы симметричны и одинаково возможны, можно применить аналогичный прием, который называется *непосредственным подсчетом вероятностей.*

Симметричность возможных исходов опыта обычно наблюдается только в искусственно организованных опытах, типа азартных игр. Так как первоначальное развитие теория вероятностей получила именно на схемах азартных игр, то прием непосредственного подсчета вероятностей, исторически возникший вместе с возникновением математической теории случайных явлений, долгое время считался основным и был положен в основу так называемой «классической» теории вероятностей. При этом опыты, не обладающие симметрией возможных исходов, искусственно сводились к «классической» схеме.

Несмотря на ограниченную сферу практических применений этой схемы, она все же представляет известный интерес, так как именно на опытах, обладающих симметрией возможных исходов, и на событиях, связанных с такими опытами, легче всего познакомиться с основными свойствами вероятностей. Такого рода событиями, допускающими непосредственный подсчет вероятностей, мы и займемся в первую очередь.

Предварительно введем некоторые вспомогательные понятия.

1. Полная группа событий.

Говорят, что несколько событий в данном опыте образуют *полную группу событий*, если в результате опыта непременно должно появиться хотя бы одно из них.

Примеры событий, образующих полную группу:

1) выпадение герба и выпадение цифры при бросании монеты;

2) попадание и промах при выстреле;

3) появление 1, 2, 3, 4, 5, 6 очков при бросании игральной кости;

4) появление белого шара и появление черного шара при вынимании одного шара из урны, в которой 2 белых и 3 черных шара;

5) ни одной опечатки, одна, две, три и более трех опечаток при проверке страницы напечатанного текста;

6) хотя бы одно попадание и хотя бы один промах при двух выстрелах.

2. Несовместные события.

Несколько событий называются *несовместными* в данном опыте, если никакие два из них не могут появиться вместе.

Примеры несовместных событий:

1) выпадение герба и выпадение цифры при бросании монеты;

2) попадание и промах при одном выстреле;

3) появление 1, 3, 4 очков при одном бросании игральной кости;

4) ровно один отказ, ровно два отказа, ровно три отказа технического устройства за десять часов работы.

3. Равновозможные события.

Несколько событий в данном опыте называются *равновозможными*, если по условиям симметрии есть основание считать, что ни одно из этих событий не является объективно более возможным, чем другое.

Примеры равновозможных событий:

1) выпадение герба и выпадение цифры при бросании монеты;

2) появление 1, 3, 4, 5 очков при бросании игральной кости;

3) появление карты бубновой, червонной, трефовой масти при вынимании карты из колоды;

4) появление шара с № 1, 2, 3 при вынимании одного шара из урны, содержащей 10 перенумерованных шаров.

Существуют группы событий, обладающие всеми тремя свойствами: они образуют полную группу, несовместны и равновозможны; например: появление герба и цифры при бросании монеты; появление 1, 2, 3, 4, 5, 6 очков при бросании игральной кости. События, образующие такую группу, называются *случаями* (иначе «шансами»).

Если какой-либо опыт по своей структуре обладает симметрией возможных исходов, то случаи представляют собой исчерпывающую систему равновозможных и исключающих друг друга исходов опыта. Про такой опыт говорят, что он «сводится к схеме случаев» (иначе — к «схеме урн»).

Схема случаев по преимуществу имеет место в искусственно организованных опытах, в которых заранее и сознательно обеспечена одинаковая возможность исходов опыта (как, например, в азартных играх). Для таких опытов возможен непосредственный подсчет вероятностей, основанный на оценке доли так называемых «благоприятных» случаев в общем числе случаев.

Случай называется *благоприятным* (или «благоприятствующим») некоторому событию, если появление этого случая влечет за собой появление данного события.

Например, при бросании игральной кости возможны шесть случаев: появление 1, 2, 3, 4, 5, 6 очков. Из них событию A — появ-

лению четного числа очков — благоприятны три случая: 2, 4, 6 и не благоприятны остальные три.

Если опыт сводится к схеме случаев, то вероятность события A в данном опыте можно оценить по относительной доле благоприятных случаев. Вероятность события A вычисляется как отношение числа благоприятных случаев к общему числу случаев:

$$P(A) = \frac{m}{n}, \tag{2.2.1}$$

где $P(A)$ — вероятность события A; n — общее число случаев; m — число случаев, благоприятных событию A.

Так как число благоприятных случаев всегда заключено между 0 и n (0 — для невозможного и n — для достоверного события), то вероятность события, вычисленная по формуле (2.2.1), всегда есть рациональная правильная дробь:

$$0 \leqslant P(A) \leqslant 1. \tag{2.2.2}$$

Формула (2.2.1), так называемая «классическая формула» для вычисления вероятностей, долгое время фигурировала в литературе как определение вероятности. В настоящее время при определении (пояснении) понятия вероятности обычно исходят из других принципов, непосредственно связывая понятие вероятности с эмпирическим понятием частоты; формула же (2.2.1) сохраняется лишь как формула для непосредственного подсчета вероятностей, пригодная тогда и только тогда, когда опыт сводится к схеме случаев, т. е. обладает симметрией возможных исходов.

Пример 1. В урне находится 2 белых и 3 черных шара. Из урны наугад вынимается один шар. Требуется найти вероятность того, что этот шар будет белым.

Решение. Обозначим A событие, состоящее в появлении белого шара. Общее число случаев $n = 5$; число случаев, благоприятных событию A, $m = 2$. Следовательно,

$$P(A) = \frac{2}{5}.$$

Пример 2. В урне a белых и b черных шаров. Из урны вынимаются два шара. Найти вероятность того, что оба шара будут белыми.

Решение. Обозначим B событие, состоящее в появлении двух белых шаров. Подсчитаем общее число возможных случаев n и число случаев m, благоприятных событию B:

$$n = C_{a+b}^2;$$

$$m = C_a^2\ ^1);$$

следовательно,

$$P(B) = \frac{C_a^2}{C_{a+b}^2}.$$

---

[1]) Знаком C_k^l обозначено число сочетаний из k элементов по l.

Пример 3. В партии из N изделий M бракованных. Из партии выбирается наугад n изделий. Определить вероятность того, что среди этих n изделий будет ровно m бракованных.

Решение. Общее число случаев, очевидно, равно C_N^n; число благоприятных случаев $C_M^m C_{N-M}^{n-m}$, откуда вероятность интересующего нас события

$$P(A) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}.$$

2.3. Частота, или статистическая вероятность, события

Формула (2.2.1) для непосредственного подсчета вероятностей применима только, когда опыт, в результате которого может появиться интересующее нас событие, обладает симметрией возможных исходов (сводится к схеме случаев). Очевидно, что далеко не всякий опыт может быть сведен к схеме случаев, и существует обширный класс событий, вероятности которых нельзя вычислить по формуле (2.2.1). Рассмотрим, например, неправильно выполненную, несимметричную игральную кость. Выпадение определенной грани уже не будет характеризоваться вероятностью $\frac{1}{6}$; вместе с тем ясно, что для данной конкретной несимметричной кости выпадение этой грани обладает некоторой вероятностью, указывающей, насколько часто в среднем должна появляться данная грань при многократном бросании. Очевидно, что вероятности таких событий, как «попадание в цель при выстреле», «выход из строя радиолампы в течение одного часа работы» или «пробивание брони осколком снаряда», также не могут быть вычислены по формуле (2.2.1), так как соответствующие опыты к схеме случаев не сводятся. Вместе с тем ясно, что каждое из перечисленных событий обладает определенной степенью объективной возможности, которую в принципе можно измерить численно и которая при повторении подобных опытов будет отражаться в относительной частоте соответствующих событий. Поэтому мы будем считать, что каждое событие, связанное с массой однородных опытов, — сводящееся к схеме случаев или нет, — имеет определенную вероятность, заключенную между нулем и единицей. Для событий, сводящихся к схеме случаев, эта вероятность может быть вычислена непосредственно по формуле (2.2.1). Для событий, не сводящихся к схеме случаев, применяются другие способы определения вероятностей. Все эти способы корнями своими уходят в опыт, в эксперимент, и для того чтобы составить представление об этих способах, необходимо уяснить себе понятие частоты события и специфику той органической связи, которая существует между вероятностью и частотой.

Если произведена серия из n опытов, в каждом из которых могло появиться или не появиться некоторое событие A, то *частотой*

события A в данной серии опытов называется отношение числа опытов, в которых появилось событие A, к общему числу произведенных опытов.

Частоту события часто называют его *статистической вероятностью* (в отличие от ранее введенной «математической» вероятности).

Условимся обозначать частоту (статистическую вероятность) события A знаком $P^*(A)$. Частота события вычисляется на основании результатов опыта по формуле

$$P^*(A) = \frac{m}{n}, \tag{2.3.1}$$

где m — число появлений события A; n — общее число произведенных опытов.

При небольшом числе опытов частота события носит в значительной мере случайный характер и может заметно изменяться от одной группы опытов к другой. Например, при каких-то десяти бросаниях монеты вполне возможно, что герб появится только два раза (частота появления герба будет равна 0,2); при других десяти бросаниях мы вполне можем получить 8 гербов (частота 0,8). Однако при увеличении числа опытов частота события все более теряет свой случайный характер; случайные обстоятельства, свойственные каждому отдельному опыту, в массе взаимно погашаются, и частота проявляет тенденцию стабилизироваться, приближаясь с незначительными колебаниями к некоторой средней, постоянной величине. Например, при многократном бросании монеты частота появления герба будет лишь незначительно уклоняться от $\frac{1}{2}$.

Это свойство «устойчивости частот», многократно проверенное экспериментально и подтверждающееся всем опытом практической деятельности человечества, есть одна из наиболее характерных закономерностей, наблюдаемых в случайных явлениях. Математическую формулировку этой закономерности впервые дал Я. Бернулли в своей теореме, которая представляет собой простейшую форму закона больших чисел. Я. Бернулли доказал, что при неограниченном увеличении числа однородных независимых опытов с практической достоверностью можно утверждать, что частота события будет сколь угодно мало отличаться от его вероятности в отдельном опыте.

Связь между частотой события и его вероятностью — глубокая, органическая связь. Эти два понятия по существу неразделимы. Действительно, когда мы оцениваем степень возможности какого-либо события, мы неизбежно связываем эту оценку с большей или меньшей частотой появления аналогичных событий на практике. Характеризуя вероятность события каким-то числом, мы не можем придать этому числу иного реального значения и иного практического смысла, чем относительная частота появления данного события при большом числе опытов. Численная оценка степени возможности события

посредством вероятности имеет практический смысл именно потому, что более вероятные события происходят в среднем чаще, чем менее вероятные. И если практика определенно указывает на то, что при увеличении числа опытов частота события имеет тенденцию выравниваться, приближаясь сквозь ряд случайных уклонений к некоторому постоянному числу, естественно предположить, что это число и есть вероятность события.

Проверить такое предположение мы, естественно, можем только для таких событий, вероятности которых могут быть вычислены непосредственно, т. е. для событий, сводящихся к схеме случаев, так как только для этих событий существует точный способ вычисления математической вероятности. Многочисленные опыты, производившиеся со времен возникновения теории вероятностей, действительно подтверждают это предположение. Они показывают, что для события, сводящегося к схеме случаев, частота события при увеличении числа опытов всегда приближается к его вероятности. Вполне естественно допустить, что и для события, не сводящегося к схеме случаев, тот же закон остается в силе и что постоянное значение, к которому при увеличении числа опытов приближается частота события, представляет собой не что иное, как вероятность события. Тогда частоту события при достаточно большом числе опытов можно принять за приближенное значение вероятности. Так и поступают на практике, определяя из опыта вероятности событий, не сводящихся к схеме случаев.

Следует отметить, что характер приближения частоты к вероятности при увеличении числа опытов несколько отличается от «стремления к пределу» в математическом смысле слова.

Когда в математике мы говорим, что переменная x_n с возрастанием n стремится к постоянному пределу a, то это означает, что разность $|x_n - a|$ становится меньше любого положительного числа ε для всех значений n, начиная с некоторого достаточно большого числа.

Относительно частоты события и его вероятности такого категорического утверждения сделать нельзя. Действительно, нет ничего физически невозможного в том, что при большом числе опытов частота события будет значительно уклоняться от его вероятности; но такое значительное уклонение является весьма маловероятным, тем менее вероятным, чем большее число опытов произведено. Например, при бросании монеты 10 раз физически возможно (хотя и маловероятно), что все 10 раз появится герб, и частота появления герба будет равна 1; при 1000 бросаниях такое событие все еще остается физически возможным, но приобретает настолько малую вероятность, что его смело можно считать практически неосуществимым. Таким образом, при возрастании числа опытов частота приближается к вероятности, но не с полной достоверностью, а с большой

вероятностью, которая при достаточно большом числе опытов может рассматриваться как практическая достоверность.

В теории вероятностей чрезвычайно часто встречается такой характер приближения одних величин к другим, и для его описания введен специальный термин: «сходимость по вероятности».

Говорят, что величина X_n *сходится по вероятности* к величине a, если при сколь угодно малом ε вероятность неравенства $|X_n - a| < \varepsilon$ с увеличением n неограниченно приближается к единице.

Применяя этот термин, можно сказать, что при увеличении числа опытов частота события не «стремится» к вероятности события, а «сходится к ней по вероятности».

Это свойство частоты и вероятности, изложенное здесь пока без достаточных математических оснований, просто на основании практики и здравого смысла, составляет содержание теоремы Бернулли, которая будет доказана нами в дальнейшем (см. гл. 13).

Таким образом, вводя понятие частоты события и пользуясь связью между частотой и вероятностью, мы получаем возможность приписать определенные вероятности, заключенные между нулем и единицей, не только событиям, которые сводятся к схеме случаев, но и тем событиям, которые к этой схеме не сводятся; в последнем случае вероятность события может быть приближенно определена по частоте события при большом числе опытов.

В дальнейшем мы увидим, что для определения вероятности события, не сводящегося к схеме случаев, далеко не всегда необходимо непосредственно определять из опыта его частоту. Теория вероятностей располагает многими способами, позволяющими определять вероятности событий косвенно, через вероятности других событий, с ними связанных. В сущности, такие косвенные способы и составляют основное содержание теории вероятностей. Однако и при таких косвенных методах исследования в конечном счете все же приходится обращаться к экспериментальным данным. Надежность и объективная ценность всех практических расчетов, выполненных с применением аппарата теории вероятностей, определяется качеством и количеством экспериментальных данных, на базе которых этот расчет выполняется.

Кроме того, при практическом применении вероятностных методов исследования всегда необходимо отдавать себе отчет в том, действительно ли исследуемое случайное явление принадлежит к категории массовых явлений, для которых, по крайней мере на некотором участке времени, выполняется свойство устойчивости частот. Только в этом случае имеет смысл говорить о вероятностях событий, имея в виду не математические фикции, а реальные характеристики случайных явлений.

Например, выражение «вероятность поражения самолета в воздушном бою для данных условий равна 0,7» имеет определенный

конкретный смысл, потому что воздушные бои мыслятся как массовые операции, которые будут неоднократно повторяться в приблизительно аналогичных условиях.

Напротив, выражение «вероятность того, что данная научная проблема решена правильно, равна 0,7» лишено конкретного смысла, и было бы методологически неправильно оценивать правдоподобие научных положений методами теории вероятностей.

2.4. Случайная величина

Одним из важнейших основных понятий теории вероятностей является понятие о случайной величине.

Случайной величиной называется величина, которая в результате опыта может принять то или иное значение, причем неизвестно заранее, какое именно.

Примеры случайных величин:

1) число попаданий при трех выстрелах;

2) число вызовов, поступавших на телефонную станцию за сутки;

3) частота попадания при 10 выстрелах.

Во всех трех приведенных примерах случайные величины могут принимать отдельные, изолированные значения, которые можно заранее перечислить.

Так, в примере 1) эти значения:

$$0,\ 1,\ 2,\ 3;$$

в примере 2):

$$1,\ 2,\ 3,\ 4,\ \ldots;$$

в примере 3):

$$0;\ 0{,}1;\ 0{,}2;\ \ldots;\ 1.0.$$

Такие случайные величины, принимающие только отделенные друг от друга значения, которые можно заранее перечислить, называются *прерывными* или *дискретными случайными величинами*.

Существуют случайные величины другого типа, например:

1) абсцисса точки попадания при выстреле;

2) ошибка взвешивания тела на аналитических весах;

3) скорость летательного аппарата в момент выхода на заданную высоту;

4) вес наугад взятого зерна пшеницы.

Возможные значения таких случайных величин не отделены друг от друга; они непрерывно заполняют некоторый промежуток, который иногда имеет резко выраженные границы, а чаще — границы неопределенные, расплывчатые.

Такие случайные величины, возможные значения которых непрерывно заполняют некоторый промежуток, называются *непрерывными случайными величинами*.

Понятие случайной величины играет весьма важную роль в теории вероятностей. Если «классическая» теория вероятностей оперировала по преимуществу с событиями, то современная теория вероятностей предпочитает, где только возможно, оперировать со случайными величинами.

Приведем примеры типичных для теории вероятностей приемов перехода от событий к случайным величинам.

Производится опыт, в результате которого может появиться или не появиться некоторое событие A. Вместо события A можно рассмотреть случайную величину X, которая равна 1, если событие A происходит, и равна 0, если событие A не происходит. Случайная величина X, очевидно, является прерывной; она имеет два возможных значения: 0 и 1. Эта случайная величина называется *характеристической случайной величиной* события A. На практике часто вместо событий оказывается удобнее оперировать их характеристическими случайными величинами. Например, если производится ряд опытов, в каждом из которых возможно появление события A, то общее число появлений события равно сумме характеристических случайных величин события A во всех опытах. При решении многих практических задач пользование таким приемом оказывается очень удобным.

С другой стороны, очень часто для вычисления вероятности события оказывается удобно связать это событие с какой-то непрерывной случайной величиной (или системой непрерывных величин).

Пусть, например, измеряются координаты какого-то объекта O для того, чтобы построить точку M, изображающую этот объект на панораме (развертке) местности. Нас интересует событие A, состоящее в том, что ошибка R в положении точки M не превзойдет заданного значения r_0 (рис. 2.4.1). Обозначим X, Y случайные ошибки в измерении координат объекта. Очевидно, событие A равносильно попаданию случайной точки M с координатами X, Y в пределы круга радиуса r_0 с центром в точке O. Другими словами, для выполнения события A случайные величины X и Y должны удовлетворять неравенству

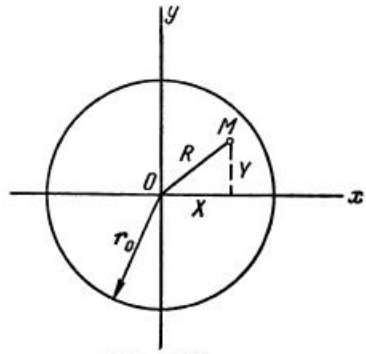


Рис. 2.4.1.

$$X^2+Y^2<r_0^2. \qquad (2.4.1)$$

Вероятность события A есть не что иное, как вероятность выполнения неравенства (2.4.1). Эта вероятность может быть определена, если известны свойства случайных величин X, Y.

Такая органическая связь между событиями и случайными величинами весьма характерна для современной теории вероятностей, которая, где только возможно, переходит от «схемы событий» к «схеме случайных величин». Последняя схема сравнительно с первой представляет собой гораздо более гибкий и универсальный аппарат для решения задач, относящихся к случайным явлениям.

2.5. Практически невозможные и практически достоверные события. Принцип практической уверенности

В п° 2.2 мы познакомились с понятиями невозможного и достоверного события. Вероятность невозможного события, равная нулю, и вероятность достоверного события, равная единице, занимают крайние положения на шкале вероятностей.

На практике часто приходится иметь дело не с невозможными и достоверными событиями, а с так называемыми «практически невозможными» и «практически достоверными» событиями.

Практически невозможным событием называется событие, вероятность которого не в точности равна нулю, но весьма близка к нулю.

Рассмотрим, например, следующий опыт: 32 буквы разрезной азбуки смешаны между собой; вынимается одна карточка, изображенная на ней буква записывается, после чего вынутая карточка возвращается обратно, и карточки перемешиваются. Такой опыт производится 25 раз. Рассмотрим событие A, заключающееся в том, что после 25 выниманий мы запишем первую строку «Евгения Онегина»:

«Мой дядя самых честных правил».

Такое событие не является логически невозможным; можно подсчитать его вероятность, которая равна $\left(\frac{1}{32}\right)^{25}$; но ввиду того что вероятность события A ничтожно мала, можно считать его практически невозможным.

Практически достоверным событием называется событие, вероятность которого не в точности равна единице, но весьма близка к единице.

Если какое-либо событие A в данном опыте практически невозможно, то противоположное ему событие $\overline{A}$, состоящее в невыполнении события A, будет практически достоверным. Таким образом, с точки зрения теории вероятностей все равно, о каких событиях говорить: о практически невозможных или о практически достоверных, так как они всегда сопутствуют друг другу.

Практически невозможные и практически достоверные события играют большую роль в теории вероятностей; на них основывается все практическое применение этой науки.

В самом деле, если нам известно, что вероятность события в данном опыте равна 0,3, это еще не дает нам возможности предсказать результат опыта. Но если вероятность события в данном опыте ничтожно мала или, наоборот, весьма близка к единице, это дает нам возможность предсказать результат опыта; в первом случае мы не будем ожидать появления события A; во втором случае будем ожидать его с достаточным основанием. При таком предсказании мы руководствуемся так называемым *принципом практической уверенности*, который можно сформулировать следующим образом.

Если вероятность некоторого события A в данном опыте E весьма мала, то можно быть практически уверенным в том, что при однократном выполнении опыта E событие A не произойдет.

Иными словами, если вероятность события A в данном опыте весьма мала, то, приступая к выполнению опыта, можно организовать свое поведение так, как будто это событие вообще невозможно, т. е. не рассчитывая совсем на его появление.

В повседневной жизни мы непрерывно бессознательно пользуемся принципом практической уверенности. Например, выезжая в путешествие по железной дороге, мы все свое поведение организуем, не считаясь с возможностью железнодорожной катастрофы, хотя некоторая, весьма малая, вероятность такого события все же имеется.

Принцип практической уверенности не может быть доказан математическими средствами; он подтверждается всем практическим опытом человечества.

Вопрос о том, насколько мала должна быть вероятность события, чтобы его можно было считать практически невозможным, выходит за рамки математической теории и в каждом отдельном случае решается из практических соображений в соответствии с той важностью, которую имеет для нас желаемый результат опыта.

Например, если вероятность отказа взрывателя при выстреле равна 0,01, мы еще можем помириться с этим и считать отказ взрывателя практически невозможным событием. Напротив, если вероятность отказа парашюта при прыжке также равна 0,01, мы, очевидно, не можем считать этот отказ практически невозможным событием и должны добиваться большей надежности работы парашюта.

Одной из важнейших задач теории вероятностей является выявление практически невозможных (или практически достоверных) событий, дающих возможность предсказывать результат опыта, и выявление условий, при которых те или иные события становятся практически невозможными (достоверными). Существует ряд теорем теории вероятностей — так называемых предельных теорем, в которых устанавливается существование событий, становящихся практически невозможными (достоверными) при увеличении числа опытов или при увеличении числа случайных величин, участвующих в задаче.

Примером такой предельной теоремы является уже сформулированная выше теорема Бернулли (простейшая форма закона больших чисел). Согласно теореме Бернулли при большом числе опытов событие, заключающееся в том, что разность между частотой события и его вероятностью сколь угодно мала, становится практически достоверным.

Наряду с практически невозможными (достоверными) событиями, которые позволяют с уверенностью предсказывать исход опыта, несмотря на наличие случайности, в теории вероятностей большую роль играют особого типа случайные величины, которые, хотя и являются случайными, но имеют такие незначительные колебания, что практически могут рассматриваться как не случайные. Примером такой «почти не случайной» величины может служить частота события при большом числе опытов. Эта величина, хотя и является случайной, но при большом числе опытов практически может колебаться только в очень узких пределах вблизи вероятности события.

Такие «почти не случайные» величины дают возможность предсказывать численный результат опыта, несмотря на наличие в нем элементов случайности, оперируя с этим результатом столь же уверенно, как мы оперируем с данными, которые доставляются обычными методами точных наук.

ГЛАВА 3

ОСНОВНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

3.1. Назначение основных теорем. Сумма и произведение событий

В предыдущей главе мы познакомились со способами непосредственного определения вероятностей, а именно: с классической формулой для вероятности события, сводящегося к схеме случаев, и со способом приближенного определения вероятности по частоте для события, которое к схеме случаев не сводится. Однако не эти непосредственные способы являются основными в теории вероятностей: их применение не всегда удобно и не всегда возможно. Даже когда событие сводится к схеме случаев, зачастую эта схема бывает слишком сложна, и непосредственный подсчет вероятности по формуле (2.2.1) становится чрезмерно громоздким. Что касается событий, не сводящихся к схеме случаев, то и их вероятности лишь в редких случаях определяются непосредственно по частотам. На практике обычно требуется определять вероятности событий, непосредственное экспериментальное воспроизведение которых затруднено. Например, если требуется определить вероятность поражения самолета в воздушном бою, ясно, что определение этой вероятности по частоте практически невозможно. И не только потому, что такие опыты оказались бы непомерно сложными и дорогостоящими, а еще и потому, что часто нам требуется оценить вероятность того или иного исхода боя не для существующих образцов техники, а для перспективных, проектируемых. Обычно такая оценка и производится для того, чтобы выявить наиболее рациональные конструктивные параметры элементов перспективной техники.

Поэтому, как правило, для определения вероятностей событий применяются не непосредственные прямые методы, а косвенные, позволяющие по известным вероятностям одних событий определять вероятности других событий, с ними связанных. Вся теория вероятностей, в основном, и представляет собой систему таких косвенных методов, пользование которыми позволяет свести необходимый эксперимент к минимуму.

Применяя эти косвенные методы, мы всегда в той или иной форме пользуемся *основными теоремами* теории вероятностей. Этих теорем две: теорема сложения вероятностей и теорема умножения вероятностей. Строго говоря, оба эти положения являются теоремами и могут быть доказаны только для событий, сводящихся к схеме случаев. Для событий, не сводящихся к схеме случаев, они принимаются аксиоматически, как принципы или постулаты.

Перед тем как формулировать и доказывать основные теоремы, введем некоторые вспомогательные понятия, а именно понятия о *сумме событий* и *произведении событий*.

Во многих областях точных наук применяются символические операции над различными объектами, которые получают свои названия по аналогии с арифметическими действиями, рядом свойств которых они обладают. Таковы, например, операции сложения и умножения векторов в механике, операции сложения и умножения матриц в алгебре и т. д. Эти операции, подчиненные известным правилам, позволяют не только упростить форму записей, но в ряде случаев существенно облегчают логическое построение научных выводов. Введение таких символических операций над событиями оказывается плодотворным и в теории вероятностей.

Суммой двух событий A и B называется событие C, состоящее в выполнении события A или события B, или обоих вместе.

Например, если событие A — попадание в цель при первом выстреле, событие B — попадание в цель при втором выстреле, то событие $C = A + B$ есть попадание в цель вообще, безразлично при каком выстреле — при первом, при втором или при обоих вместе.

Если события A и B несовместны, то естественно, что появление обоих этих событий вместе отпадает, и сумма событий A и B сводится к появлению или события A, или события B. Например, если событие A — появление карты червонной масти при вынимании карты из колоды, событие B — появление карты бубновой масти, то $C = A + B$ есть появление карты красной масти, безразлично — червонной или бубновой.

Короче, *суммой двух событий* A и B называется событие C, состоящее в появлении хотя бы одного из событий A и B.

Суммой нескольких событий называется событие, состоящее в появлении хотя бы одного из этих событий.

Например, если опыт состоит в пяти выстрелах по мишени и даны события:

A_0 — ни одного попадания,
A_1 — ровно одно попадание,
A_2 — ровно два попадания,
A_3 — ровно три попадания,
A_4 — ровно четыре попадания,
A_5 — ровно пять попаданий,

то

$$A = A_0 + A_1 + A_2$$

есть событие «не более двух попаданий», а

$$B = A_3 + A_4 + A_5$$

есть событие «не менее трех попаданий».

Произведением двух событий A и B называется событие C, состоящее в совместном выполнении события A и события B.

Например, если событие A — появление туза при вынимании карты из колоды, событие B — появление карты бубновой масти, то событие $C = AB$ есть появление бубнового туза. Если производится два выстрела по мишени и событие A — попадание при первом выстреле, событие B — попадание при втором выстреле, то $C = AB$ есть попадание при обоих выстрелах.

Произведением нескольких событий называется событие, состоящее в совместном появлении всех этих событий.

Например, если по мишени производится три выстрела и рассматриваются события

B_1 — промах при первом выстреле,
B_2 — промах при втором выстреле,
B_3 — промах при третьем выстреле,

то событие

$$B = B_1 B_2 B_3$$

состоит в том, что в мишень не будет ни одного попадания.

При определении вероятностей часто приходится представлять сложные события в виде комбинаций более простых событий, применяя и операцию сложения, и операцию умножения событий.

Например, пусть по мишени производится три выстрела и рассматриваются следующие элементарные события:

A_1 — попадание при первом выстреле,
$\bar{A}_1$ — промах при первом выстреле,
A_2 — попадание при втором выстреле,
$\bar{A}_2$ — промах при втором выстреле,
A_3 — попадание при третьем выстреле,
$\bar{A}_3$ — промах при третьем выстреле.

Рассмотрим более сложное событие B, состоящее в том, что в результате данных трех выстрелов будет ровно одно попадание в мишень. Событие B можно представить в виде следующей комбинации элементарных событий:

$$B = A_1 \bar{A}_2 \bar{A}_3 + \bar{A}_1 A_2 \bar{A}_3 + \bar{A}_1 \bar{A}_2 A_3.$$

Событие C, состоящее в том, что в мишень будет не менее двух попаданий, может быть представлено в виде:

$$C = A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3 + A_1 A_2 A_3.$$

Такие приемы представления сложных событий часто применяются в теории вероятностей.

Непосредственно из определения суммы и произведения событий следует, что

$$A + A = A,$$
$$AA = A.$$

Если событие B есть частный случай события A, то

$$A + B = A;$$
$$AB = B.$$

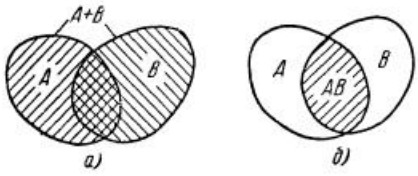


Рис. 3.1.1.

При пользовании понятиями суммы и произведения событий часто оказывается полезной наглядная геометрическая интерпретация этих понятий.

На рис. 3.1.1 наглядно иллюстрированы понятия суммы и произведения двух событий. Если событие A есть попадание точки в область A, соответственно событие B — попадание в область B, то событие $A + B$ есть попадание в область, заштрихованную на рис. 3.1.1, *а*), а событие AB — в область, заштрихованную на рис. 3.1.1, *б*).

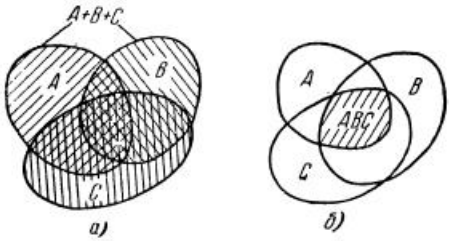


Рис. 3.1.2.

На рис. 3.1.2 аналогично показаны сумма и произведение трех событий.

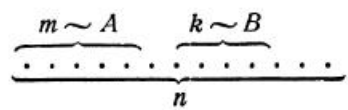
3.2. Теорема сложения вероятностей

Теорема сложения вероятностей формулируется следующим образом.

Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B). \qquad (3.2.1)$$

Докажем теорему сложения вероятностей для схемы случаев. Пусть возможные исходы опыта сводятся к совокупности случаев, которые мы для наглядности изобразим в виде n точек:



Предположим, что из этих случаев m благоприятны событию A, а k — событию B. Тогда

$$P(A)=\frac{m}{n};\quad P(B)=\frac{k}{n}.$$

Так как события A и B несовместны, то нет таких случаев, которые благоприятны и A, и B вместе. Следовательно, событию $A+B$ благоприятны $m+k$ случаев и

$$P(A+B)=\frac{m+k}{n}.$$

Подставляя полученные выражения в формулу (3.2.1), получим тождество. Теорема доказана.

Обобщим теорему сложения на случай трех событий. Обозначая событие $A+B$ буквой D и присоединяя к сумме еще одно событие C, легко доказать, что

$$P(A+B+C)=P(D+C)=P(D)+P(C)=$$
$$=P(A+B)+P(C)=P(A)+P(B)+P(C).$$

Очевидно, методом полной индукции можно обобщить теорему сложения на произвольное число несовместных событий. Действительно, предположим, что она справедлива для n событий:

$$A_1,\ A_2,\ \ldots,\ A_n,$$

и докажем, что она будет справедлива для $n+1$ событий:

$$A_1,\ A_2,\ \ldots,\ A_n,\ A_{n+1}.$$

Обозначим:

$$A_1+A_2+\ \ldots\ +A_n=C.$$

Имеем:

$$P(A_1+A_2+\ \ldots\ +A_n+A_{n+1})=P(C+A_{n+1})=P(C)+P(A_{n+1}).$$

Но так как для n событий мы считаем теорему уже доказанной, то

$$P(C)=P(A_1)+P(A_2)+\ \ldots\ +P(A_n),$$

откуда

$$P(A_1+A_2+\ \ldots\ +A_n+A_{n+1})=$$
$$=P(A_1)+P(A_2)+\ \ldots\ +P(A_n)+P(A_{n+1}),$$

что и требовалось доказать.

Таким образом, теорема сложения вероятностей применима к любому числу несовместных событий. Ее удобно записать в виде:

$$P\left(\sum_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i). \qquad (3.2.2)$$

Отметим следствия, вытекающие из теоремы сложения вероятностей.

Следствие 1. *Если события $A_1, A_2, \ldots, A_n$ образуют полную группу несовместных событий, то сумма их вероятностей равна единице*:

$$\sum_{i=1}^{n} P(A_i) = 1.$$

Доказательство. Так как события $A_1, A_2, \ldots, A_n$ образуют полную группу, то появление хотя бы одного из них — достоверное событие:

$$P(A_1 + A_2 + \ldots + A_n) = 1.$$

Так как $A_1, A_2, \ldots, A_n$ — несовместные события, то к ним применима теорема сложения вероятностей

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = \sum_{i=1}^{n} P(A_i),$$

откуда

$$\sum_{i=1}^{n} P(A_i) = 1,$$

что и требовалось доказать.

Перед тем как вывести второе следствие теоремы сложения, определим понятие о «противоположных событиях».

Противоположными событиями называются два несовместных события, образующих полную группу.

Событие, противоположное событию A, принято обозначать $\bar{A}$.

Примеры противоположных событий.

1) A — попадание при выстреле,
   $\bar{A}$ — промах при выстреле;
2) B — выпадение герба при бросании монеты,
   $\bar{B}$ — выпадение цифры при бросании монеты;
3) C — безотказная работа всех элементов технической системы,
   $\bar{C}$ — отказ хотя бы одного элемента;
4) D — обнаружение не менее двух бракованных изделий в контрольной партии,
   $\bar{D}$ — обнаружение не более одного бракованного изделия.

Следствие 2. *Сумма вероятностей противоположных событий равна единице*:

$$P(A) + P(\bar{A}) = 1.$$

Это следствие есть частный случай следствия 1. Оно выделено особо ввиду его большой важности в практическом применении теории вероятностей. На практике весьма часто оказывается легче вычислить

вероятность противоположного события $\bar{A}$, чем вероятность прямого события A. В этих случаях вычисляют $P(\bar{A})$ и находят $P(A) = 1 - P(\bar{A})$.

Рассмотрим несколько примеров на применение теоремы сложения и ее следствий.

Пример 1. В лотерее 1000 билетов; из них на один билет падает выигрыш 500 руб., на 10 билетов — выигрыши по 100 руб., на 50 билетов — выигрыши по 20 руб., на 100 билетов — выигрыши по 5 руб., остальные билеты невыигрышные. Некто покупает один билет. Найти вероятность выиграть не менее 20 руб.

Решение. Рассмотрим события:

A — выиграть не менее 20 руб.,
A_1 — выиграть 20 руб.,
A_2 — выиграть 100 руб.,
A_3 — выиграть 500 руб.

Очевидно,

$$A = A_1 + A_2 + A_3.$$

По теореме сложения вероятностей

$$P(A) = P(A_1) + P(A_2) + P(A_3) = 0{,}050 + 0{,}010 + 0{,}001 = 0{,}061.$$

Пример 2. Производится бомбометание по трем складам боеприпасов, причем сбрасывается одна бомба. Вероятность попадания в первый склад 0,01; во второй 0,008; в третий 0,025. При попадании в один из складов взрываются все три. Найти вероятность того, что склады будут взорваны.

Решение. Рассмотрим события:

A — взрыв складов,
A_1 — попадание в первый склад,
A_2 — попадание во второй склад,
A_3 — попадание в третий склад.

Очевидно,

$$A = A_1 + A_2 + A_3.$$

Так как при сбрасывании одной бомбы события A_1, A_2, A_3 несовместны, то

$$P(A) = P(A_1) + P(A_2) + P(A_3) = 0{,}01 + 0{,}008 + 0{,}025 = 0{,}043.$$

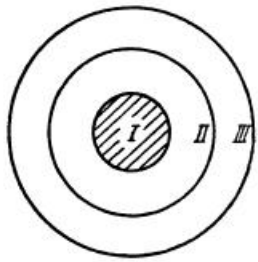


Рис. 3.2.1.

Пример 3. Круговая мишень (рис. 3.2.1) состоит из трех зон: *I*, *II* и *III*. Вероятность попадания в первую зону при одном выстреле 0,15, во вторую 0,23, в третью 0,17. Найти вероятность промаха.

Решение. Обозначим A — промах,

$\bar{A}$ — попадание.

Тогда

$$\bar{A} = \bar{A}_1 + \bar{A}_2 + \bar{A}_3,$$

где $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ — попадание соответственно в первую, вторую и третью зоны

$$P(\bar{A}) = P(\bar{A}_1) + P(\bar{A}_2) + P(\bar{A}_3) = 0{,}15 + 0{,}23 + 0{,}17 = 0{,}55,$$

откуда

$$P(A) = 1 - P(\bar{A}) = 0{,}45.$$

Как уже указывалось, теорема сложения вероятностей (3.2.1) справедлива только для несовместных событий. В случае, когда события A и B совместны, вероятность суммы этих событий выражается формулой

$$P(A+B)=P(A)+P(B)-P(AB). \qquad (3.2.3)$$

В справедливости формулы (3.2.3) можно наглядно убедиться, рассматривая рис. 3.2.2.

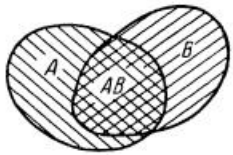


Рис. 3.2.2.

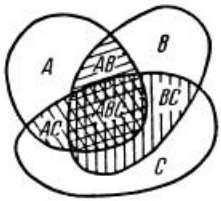


Рис. 3.2.3.

Аналогично вероятность суммы трех совместных событий вычисляется по формуле

$$P(A+B+C)=$$
$$=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC).$$

Справедливость этой формулы также наглядно следует из геометрической интерпретации (рис. 3.2.3).

Методом полной индукции можно доказать общую формулу для вероятности суммы любого числа совместных событий:

$$P\left(\sum_{i=1}^{n} A_i\right)=\sum_i P(A_i)-\sum_{i,\,j} P(A_iA_j)+\sum_{i,\,j,\,k} P(A_iA_jA_k)-\ldots$$
$$\ldots+(-1)^{n-1}P(A_1A_2\ldots A_n), \qquad (3.2.4)$$

где суммы распространяются на различные значения индексов i; i, j; i, j, k, и т. д.

Формула (3.2.4) выражает вероятность суммы любого числа событий через вероятности произведений этих событий, взятых по одному, по два, по три и т. д.

Аналогичную формулу можно написать для произведения событий. Действительно, из рис. 3.2.2 непосредственно ясно, что

$$P(AB)=P(A)+P(B)-P(A+B). \qquad (3.2.5)$$

Из рис. 3.2.3 видно, что

$$P(ABC)=P(A)+P(B)+P(C)-P(A+B)-$$
$$-P(A+C)-P(B+C)+P(A+B+C). \qquad (3.2.6)$$

Общая формула, выражающая вероятность произведения произвольного числа событий через вероятности сумм этих событий, взятых по одному, по два, по три и т. д., имеет вид:

$$P(A_1A_2 \dots A_n) = \sum_i P(A_i) - \sum_{i,j} P(A_i + A_j) +$$

$$+ \sum_{i,j,k} P(A_i + A_j + A_k) + \dots + (-1)^{n-1} P(A_1 + \dots + A_n). \quad (3.2.7)$$

Формулы типа (3.2.4) и (3.2.7) находят практическое применение при преобразовании различных выражений, содержащих вероятности сумм и произведений событий. В зависимости от специфики задачи в некоторых случаях удобнее бывает пользоваться только суммами, а в других только произведениями событий: для преобразования одних в другие и служат подобные формулы.

Пример. Техническое устройство состоит из трех агрегатов: двух агрегатов первого типа — A_1 и A_2 — и одного агрегата второго типа — B. Агрегаты A_1 и A_2 дублируют друг друга: при отказе одного из них происходит автоматическое переключение на второй. Агрегат B не дублирован. Для того чтобы устройство прекратило работу (отказало), нужно, чтобы одновременно отказали оба агрегата A_1 и A_2 или же агрегат B. Таким образом, отказ устройства — событие C — представляется в виде:

$$C = A_1A_2 + B,$$

где A_1 — отказ агрегата A_1, A_2 — отказ агрегата A_2, B — отказ агрегата B.

Требуется выразить вероятность события C через вероятности событий, содержащих только суммы, а не произведения элементарных событий A_1, A_2 и B.

Решение. По формуле (3.2.3) имеем:

$$P(C) = P(A_1A_2) + P(B) - P(A_1A_2B); \quad (3.2.8)$$

по формуле (3.2.5)

$$P(A_1A_2) = P(A_1) + P(A_2) - P(A_1 + A_2);$$

по формуле (3.2.6)

$$P(A_1A_2B) = P(A_1) + P(A_2) + P(B) - P(A_1 + A_2) - P(A_1 + B) - P(A_2 + B) + P(A_1 + A_2 + B).$$

Подставляя эти выражения в (3.2.8) и производя сокращения, получим:

$$P(C) = P(A_1 + B) + P(A_2 + B) - P(A_1 + A_2 + B).$$

3.3. Теорема умножения вероятностей

Перед тем как излагать теорему умножения вероятностей, введем еще одно важное понятие: понятие о независимых и зависимых событиях.

Событие A называется *независимым от события* B, если вероятность события A не зависит от того, произошло событие B или нет.

Событие A называется *зависимым от события* B, если вероятность события A меняется в зависимости от того, произошло событие B или нет.

Рассмотрим примеры.

1) Опыт состоит в бросании двух монет; рассматриваются события:

A — появление герба на первой монете,

B — появление герба на второй монете.

В данном случае вероятность события A не зависит от того, произошло событие B или нет; событие A независимо от события B.

2) В урне два белых шара и один черный; два лица вынимают из урны по одному шару; рассматриваются события:

A — появление белого шара у 1-го лица,

B — появление белого шара у 2-го лица.

Вероятность события A до того, как известно что-либо о событии B, равна $\frac{2}{3}$. Если стало известно, что событие B произошло, то вероятность события A становится равной $\frac{1}{2}$, из чего заключаем, что событие A зависит от события B.

Вероятность события A, вычисленная при условии, что имело место другое событие B, называется *условной вероятностью события* A и обозначается

$$P(A|B).$$

Для условий последнего примера

$$P(A)=\frac{2}{3}; \quad P(A|B)=\frac{1}{2}.$$

Условие независимости события A от события B можно записать в виде:

$$P(A|B)=P(A),$$

а условие зависимости — в виде:

$$P(A|B)\neq P(A).$$

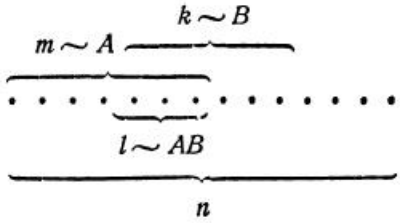
Перейдем к формулировке и доказательству теоремы умножения вероятностей.

Теорема умножения вероятностей формулируется следующим образом.

Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое имело место:

$$P(AB)=P(A)P(B|A). \tag{3.3.1}$$

Докажем теорему умножения для схемы случаев. Пусть возможные исходы опыта сводятся к n случаям, которые мы снова для наглядности изобразим в виде n точек:



Предположим, что событию A благоприятны m случаев, а событию B благоприятны k случаев. Так как мы не предполагали события A и B несовместными, то вообще существуют случаи, благоприятные и событию A, и событию B одновременно. Пусть число таких случаев l. Тогда

$$P(AB) = \frac{l}{n}; \qquad P(A) = \frac{m}{n}.$$

Вычислим $P(B \mid A)$, т. е. условную вероятность события B в предположении, что A имело место. Если известно, что событие A произошло, то из ранее возможных n случаев остаются возможными только те m, которые благоприятствовали событию A. Из них l случаев благоприятны событию B. Следовательно,

$$P(B \mid A) = \frac{l}{m}.$$

Подставляя выражения $P(AB)$, $P(A)$ и $P(B \mid A)$ в формулу (3.3.1), получим тождество. Теорема доказана.

Очевидно, при применении теоремы умножения вполне безразлично, какое из событий A и B считать первым, а какое вторым, и теорему умножения можно записать и в таком виде:

$$P(AB) = P(B)\,P(A \mid B).$$

Отметим следствия, вытекающие из теоремы умножения.

Следствие 1. *Если событие A не зависит от события B, то и событие B не зависит от события A.*

Доказательство. Дано, что событие A не зависит от B, т. е.

$$P(A) = P(A \mid B). \tag{3.3.2}$$

Требуется доказать, что и событие B не зависит от A, т. е.

$$P(B) = P(B \mid A).$$

При доказательстве будем предполагать, что $P(A) \neq 0$.

Напишем теорему умножения вероятностей в двух формах:

$$P(AB) = P(A)\,P(B \mid A),$$

$$P(AB) = P(B)\,P(A \mid B),$$

откуда

$$P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$$

или, согласно условию (3.3.2),

$$P(A)\,P(B \mid A) = P(B)\,P(A). \tag{3.3.3}$$

Разделим обе части равенства (3.3.3) на $P(A)$. Получим:

$$P(B \mid A) = P(B),$$

что и требовалось доказать.

Из следствия 1 вытекает, что зависимость или независимость событий всегда взаимны. В связи с этим можно дать следующее новое определение независимых событий.

Два события называются независимыми, если появление одного из них не изменяет вероятности появления другого.

Понятие независимости событий может быть распространено на случай произвольного числа событий. Несколько событий называются независимыми, если любое из них не зависит от любой совокупности остальных.

Следствие 2. *Вероятность произведения двух независимых событий равна произведению вероятностей этих событий.*

Следствие непосредственно вытекает из определения независимых событий.

Теорема умножения вероятностей может быть обобщена на случай произвольного числа событий. В общем виде она формулируется так.

Вероятность произведения нескольких событий равна произведению вероятностей этих событий, причем вероятность каждого следующего по порядку события вычисляется при условии, что все предыдущие имели место:

$$P(A_1A_2 \ldots A_n) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1A_2) \ldots \\ \ldots P(A_n \mid A_1A_2 \ldots A_{n-1}). \tag{3.3.4}$$

Доказательство может быть дано тем же методом полной индукции.

В случае независимых событий теорема упрощается и принимает вид:

$$P(A_1A_2 \ldots A_n) = P(A_1)\,P(A_2) \ldots P(A_n), \tag{3.3.5}$$

т. е. *вероятность произведения независимых событий равна произведению вероятностей этих событий.*

Применяя знак произведения, теорему можно записать в виде:

$$P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i). \tag{3.3.6}$$

Рассмотрим примеры на применение теоремы умножения вероятностей.

Пример 1. В урне 2 белых и 3 черных шара. Из урны вынимают подряд два шара. Найти вероятность того, что оба шара белые.

Решение. Обозначим:

A — появление двух белых шаров.

Событие A представляет собой произведение двух событий:

$$A = A_1 A_2,$$

где A_1 — появление белого шара при первом вынимании,
A_2 — появление белого шара при втором вынимании.
По теореме умножения вероятностей

$$P(A) = P(A_1)\, P(A_2|A_1) = \frac{2}{5} \cdot \frac{1}{4} = 0{,}1.$$

Пример 2. Те же условия, но после первого вынимания шар возвращается в урну, и шары в урне перемешиваются.

Решение. В данном случае события A_1 и A_2 независимы и

$$P(A) = P(A_1)\, P(A_2) = \frac{2}{5} \cdot \frac{2}{5} = 0{,}16.$$

Пример 3. Прибор, работающий в течение времени t, состоит из трех узлов, каждый из которых, независимо от других, может в течение времени t отказать (выйти из строя). Отказ хотя бы одного узла приводит к отказу прибора в целом. За время t надежность (вероятность безотказной работы) первого узла равна $p_1 = 0{,}8$; второго $p_2 = 0{,}9$; третьего $p_3 = 0{,}7$. Найти надежность прибора в целом.

Решение. Обозначая:

A — безотказная работа прибора,
A_1 — безотказная работа первого узла,
A_2 — безотказная работа второго узла,
A_3 — безотказная работа третьего узла,

имеем:

$$A = A_1 A_2 A_3,$$

откуда по теореме умножения для независимых событий

$$P(A) = P(A_1)\, P(A_2)\, P(A_3) = p_1 p_2 p_3 = 0{,}504.$$

На практике сравнительно редко встречаются задачи, в которых нужно применять только теорему сложения или только теорему умножения вероятностей. Обычно обе теоремы приходится применять совместно. При этом, как правило, событие, вероятность которого требуется определить, представляется в виде суммы нескольких несовместных событий (вариантов данного события), каждое из которых в свою очередь является произведением событий.

Пример 4. Производится три выстрела по одной и той же мишени. Вероятности попадания при первом, втором и третьем выстрелах равны соответственно

$$p_1 = 0{,}4; \qquad p_2 = 0{,}5; \qquad p_3 = 0{,}7.$$

Найти вероятность того, что в результате этих трех выстрелов в мишени будет ровно одна пробоина.

Решение. Рассмотрим событие A — ровно одно попадание в мишень. Это событие может осуществиться несколькими способами, т. е. распадается на несколько несовместных вариантов: может быть попадание при первом выстреле, промахи при втором и третьем; или же попадание при втором выстреле, промахи при первом и третьем; или, наконец, промахи при первом и втором выстрелах и попадание при третьем. Следовательно,

$$A = A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3,$$

где A_1, A_2, A_3 — попадание при первом, втором, третьем выстрелах,

$\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ — промах при первом, втором, третьем выстрелах.

Применяя теоремы сложения и умножения вероятностей и пользуясь свойством противоположных событий, находим:

$$P(A) = P(A_1\bar{A}_2\bar{A}_3) + P(\bar{A}_1A_2\bar{A}_3) + P(\bar{A}_1\bar{A}_2A_3) =$$
$$= 0{,}4 \cdot 0{,}5 \cdot 0{,}3 + 0{,}6 \cdot 0{,}5 \cdot 0{,}3 + 0{,}6 \cdot 0{,}5 \cdot 0{,}7 = 0{,}36.$$

Пример 5. В условиях предыдущего примера найти вероятность того, что в мишени будет хотя бы одна пробоина.

Решение. Рассмотрим событие B — хотя бы одно попадание в мишень. Пользуясь тем же приемом, который был применен в предыдущем примере, и теми же обозначениями, можно представить событие B в виде суммы несовместных вариантов:

$$B = A_1A_2A_3 + A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3 + A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3,$$

найти вероятность каждого варианта по теореме умножения и все эти вероятности сложить. Однако такой путь решения задачи слишком сложен; здесь целесообразно от прямого события B перейти к противоположному:

$\bar{B}$ — ни одного попадания в мишень.

Очевидно,

$$\bar{B} = \bar{A}_1\bar{A}_2\bar{A}_3.$$

По теореме умножения

$$P(\bar{B}) = P(\bar{A}_1\bar{A}_2\bar{A}_3) = 0{,}6 \cdot 0{,}5 \cdot 03 = 0{,}09,$$

откуда

$$P(B) = 1 - P(\bar{B}) = 1 - 0{,}09 = 0{,}91.$$

На последнем примере проиллюстрирован принцип целесообразности применения противоположных событий в теории вероятностей. Его можно сформулировать следующим образом.

Если противоположное событие распадается на меньшее число вариантов, чем прямое событие, то имеет смысл при вычислении вероятностей переходить к противоположному событию.

Пример 6. Происходит бой («дуэль») между двумя участниками (летательными аппаратами, танками, кораблями) A и B. У стороны A в запасе два выстрела, у стороны B — один. Начинает стрельбу A: он делает по B один выстрел и поражает его с вероятностью 0,2. Если B не поражен, он отвечает противнику выстрелом и поражает его с вероятностью 0,3. Если A этим выстрелом не поражен, то он делает по B свой последний выстрел, которым поражает его с вероятностью 0,4. Найти вероятность того, что в бою будет поражен: а) участник A, б) участник B.

Решение. Рассмотрим события:

A — поражение участника A,

B — поражение участника B.

Для выполнения события A необходимо совмещение (произведение) двух событий: 1) A не поразил B первым выстрелом и 2) B поразил A своим ответным выстрелом. По теореме умножения вероятностей получим

$$P(A) = 0{,}8 \cdot 0{,}3 = 0{,}24.$$

Перейдем к событию B. Оно, очевидно, состоит из двух несовместных вариантов:

$$B = B_1 + B_2,$$

где B_1 — поражение участника B первым выстрелом A,

B_2 — поражение участника B вторым выстрелом A.

По теореме сложения вероятностей

$$P(B) = P(B_1) + P(B_2).$$

По условию $P(B_1) = 0{,}2$. Что касается события B_2, то оно представляет собой совмещение (произведение) трех событий, а именно:

1) первый выстрел стороны A не должен поразить B;
2) ответный выстрел стороны B не должен поразить A;
3) последний (второй) выстрел стороны A должен поразить B.

По теореме умножения вероятностей

$$P(B_2) = 0{,}8 \cdot 0{,}7 \cdot 0{,}4 = 0{,}224,$$

откуда

$$P(B) = 0{,}2 + 0{,}224 = 0{,}424.$$

Пример 7. Цель, по которой ведется стрельба, состоит из трех различных по уязвимости частей. Для поражения цели достаточно одного попадания в первую часть, или двух попаданий во вторую, или трех попаданий в третью. Если снаряд попал в цель, то вероятность ему попасть в ту или другую часть пропорциональна площади этой части. На проекции цели на плоскость, перпендикулярную направлению стрельбы, первая, вторая и третья части занимают относительные площади 0,1, 0,2 и 0,7. Известно, что в цель попало ровно два снаряда. Найти вероятность того, что цель будет поражена.

Решение. Обозначим A — поражение цели; $P(A|2)$ — условную вероятность поражения цели при условии, что в нее попало ровно два снаряда. Два снаряда, попавшие в цель, могут поразить ее двумя способами: или хотя бы один из них попадает в первую часть, или же оба снаряда попадут во вторую. Эти варианты несовместны, так как в цель попало всего два снаряда; поэтому можно применить теорему сложения. Вероятность того, что хотя бы один снаряд попадет в первую часть, может быть вычислена через вероятность противоположного события (ни один из двух снарядов не попадет в первую часть) и равна $1 - 0{,}9^2$. Вероятность того, что оба снаряда попадут во вторую часть, равна $0{,}2^2$. Следовательно,

$$P(A|2) = 1 - 0{,}9^2 + 0{,}2^2 = 0{,}23.$$

Пример 8. Для условий предыдущего примера найти вероятность поражения цели, если известно, что в нее попало три снаряда.

Решение. Решим задачу двумя способами: через прямое и противоположное событие.

Прямое событие — поражение цели при трех попаданиях — распадается на четыре несовместных варианта:

A_1 — хотя бы одно попадание в первую часть,

A_2 — два попадания во вторую часть и одно — в третью,

A_3 — три попадания во вторую часть,

A_4 — три попадания в третью часть.

Вероятность первого варианта находим аналогично предыдущему примеру:

$$P(A_1) = 1 - 0{,}9^3 = 0{,}271.$$

Найдем вероятность второго варианта. Три попавших снаряда могут распределиться по второй и третьей частям нужным образом (два во вторую и один — в третью) тремя способами $\left(C_3^2 = 3\right)$. Следовательно,

$$P(A_2) = 3 \cdot 0{,}2^2 \cdot 0{,}7 = 0{,}084.$$

Далее находим:

$$P(A_3) = 0{,}2^3 = 0{,}008,$$

$$P(A_4) = 0{,}7^3 = 0{,}343.$$

Отсюда

$$P(A \mid 3) = 0{,}271 + 0{,}084 + 0{,}008 + 0{,}343 = 0{,}706.$$

Однако проще решается задача, если перейти к противоположному событию — непоражению цели при трех попаданиях. Это событие может осуществиться только одним способом: если два снаряда из трех попадут в третью часть, а один — во вторую. Таких комбинаций может быть три $\left(C_3^2 = 3\right)$, следовательно,

$$P(\bar{A} \mid 3) = 3 \cdot 0{,}7^2 \cdot 0{,}2 = 0{,}294,$$

откуда

$$P(A \mid 3) = 1 - 0{,}294 = 0{,}706.$$

Пример 9. Монета бросается 6 раз. Найти вероятность того, что выпадет больше гербов, чем цифр.

Решение. Для нахождения вероятности интересующего нас события A (выпадет больше гербов, чем цифр) можно было бы перечислить все возможные его варианты, например:

A_1 — выпадет шесть гербов и ни одной цифры,

A_2 — выпадет пять гербов и одна цифра

и т. д.

Однако проще будет применить другой прием. Перечислим все возможные исходы опыта:

A — выпадет больше гербов, чем цифр,

B — выпадет больше цифр, чем гербов,

C — выпадет одинаковое число цифр и гербов.

События A, B, C несовместны и образуют полную группу. Следовательно,

$$P(A) + P(B) + P(C) = 1.$$

Так как задача симметрична относительно «герба» и «цифры»,

$$P(A) = P(B),$$

откуда

$$2P(A) + P(C) = 1$$

и

$$P(A) = \frac{1 - P(C)}{2}.$$

Найдем вероятность события C, состоящего в том, что при шести бросаниях монеты появится ровно три герба (а значит, ровно три цифры). Вероятность любого из вариантов события C (например, последовательности г, ц, г, г, ц, ц при шести бросаниях) одна и та же и равна $\left(\frac{1}{2}\right)^6$. Число таких комбинаций равно $C_6^3 = 20$ (числу способов, какими можно из шести бросаний выбрать три, в которых появился герб). Следовательно,

$$P(C) = \frac{20}{64} = \frac{5}{16};$$

отсюда

$$P(A) = \frac{1}{2}\left(1 - \frac{5}{16}\right) = \frac{11}{32}.$$

Пример 10. Прибор состоит из четырех узлов: A_1, A_2, A_3, A_4, причем узел A_2 дублирует A_1, а узел A_4 дублирует узел A_3. При отказе (выходе из строя) любого из основных узлов (A_1 или A_3) происходит автоматическое переключение на дублирующий узел. Надежность (вероятность безотказной работы) в течение заданного времени каждого из узлов равна соответственно p_1, p_2, p_3, p_4. Надежность каждого из переключающих устройств равна p. Все элементы выходят из строя независимо друг от друга. Определить надежность прибора.

Решение. Рассмотрим совокупность узлов A_1, A_2 и соответствующего переключающего устройства как один «обобщенный узел» B, а совокупность узлов A_3, A_4 и соответствующего переключающего устройства — как обобщенный узел C. Рассмотрим события:

A — безотказная работа прибора,
B — безотказная работа обобщенного узла B,
C — безотказная работа обобщенного узла C.

Очевидно,

$$A = BC,$$

откуда

$$P(A) = P(B)\,P(C).$$

Найдем вероятность события B. Оно распадается на два варианта:

A_1 — исправно работал узел A_1

и

A_2' — узел A_1 отказал, но оказались исправными переключающее устройство и узел A_2.

Имеем:

$$P(B) = P(A_1) + P(A_2') = p_1 + (1 - p_1)\,pp_2,$$

аналогично

$$P(C) = p_3 + (1 - p_3)\,pp_4,$$

откуда

$$P(A) = [p_1 + (1 - p_1)\,pp_2]\,[p_3 + (1 - p_3)\,pp_4].$$

3.4. Формула полной вероятности

Следствием обеих основных теорем — теоремы сложения вероятностей и теоремы умножения вероятностей — является так называемая *формула полной вероятности.*

Пусть требуется определить вероятность некоторого события A, которое может произойти вместе с одним из событий:

$$H_1, H_2, \ldots, H_n,$$

образующих полную группу несовместных событий. Будем эти события называть *гипотезами.*

Докажем, что в этом случае

$$P(A) = \sum_{i=1}^{n} P(H_i) P(A|H_i), \tag{3.4.1}$$

т. е. вероятность события A вычисляется как сумма произведений вероятности каждой гипотезы на вероятность события при этой гипотезе.

Формула (3.4.1) носит название *формулы полной вероятности.*

Доказательство. Так как гипотезы $H_1, H_2, \ldots, H_n$ образуют полную группу, то событие A может появиться только в комбинации с какой-либо из этих гипотез:

$$A = H_1A + H_2A + \ldots + H_nA.$$

Так как гипотезы $H_1, H_2, \ldots, H_n$ несовместны, то и комбинации $H_1A, H_2A, \ldots, H_nA$ также несовместны; применяя к ним теорему сложения, получим:

$$P(A) = P(H_1A) + P(H_2A) + \ldots + P(H_nA) = \sum_{i=1}^{n} P(H_iA).$$

Применяя к событию H_iA теорему умножения, получим:

$$P(A) = \sum_{i=1}^{n} P(H_i) P(A|H_i),$$

что и требовалось доказать.

Пример 1. Имеются три одинаковые на вид урны; в первой урне два белых и один черный шар; во второй — три белых и один черный; в третьей — два белых и два черных шара. Некто выбирает наугад одну из урн и вынимает из нее шар. Найти вероятность того, что этот шар белый.

Решение. Рассмотрим три гипотезы:

H_1 — выбор первой урны,
H_2 — выбор второй урны,
H_3 — выбор третьей урны

и событие A — появление белого шара.

Так как гипотезы, по условию задачи, равновозможны, то

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}.$$

Условные вероятности события A при этих гипотезах соответственно равны:

$$P(A \mid H_1) = \frac{2}{3}; \qquad P(A \mid H_2) = \frac{3}{4}; \qquad P(A \mid H_3) = \frac{1}{2}.$$

По формуле полной вероятности

$$P(A) = \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{2} = \frac{23}{36}.$$

Пример 2. По самолету производится три одиночных выстрела. Вероятность попадания при первом выстреле равна 0,4, при втором — 0,5, при третьем — 0,7. Для вывода самолета из строя заведомо достаточно трех попаданий; при одном попадании самолет выходит из строя с вероятностью 0,2, при двух попаданиях с вероятностью 0,6. Найти вероятность того, что в результате трех выстрелов самолет будет выведен из строя.

Решение. Рассмотрим четыре гипотезы:

H_0 — в самолет не попало ни одного снаряда,
H_1 — в самолет попал один снаряд,
H_2 — в самолет попало два снаряда,
H_3 — в самолет попало три снаряда.

Пользуясь теоремами сложения и умножения, найдем вероятности этих гипотез:

$$P(H_0) = 0{,}6 \cdot 0{,}5 \cdot 0{,}3 = 0{,}09;$$
$$P(H_1) = 0{,}4 \cdot 0{,}5 \cdot 0{,}3 + 0{,}6 \cdot 0{,}5 \cdot 0{,}3 + 0{,}6 \cdot 0{,}5 \cdot 0{,}7 = 0{,}36;$$
$$P(H_2) = 0{,}6 \cdot 0{,}5 \cdot 0{,}7 + 0{,}4 \cdot 0{,}5 \cdot 0{,}7 + 0{,}4 \cdot 0{,}5 \cdot 0{,}3 = 0{,}41;$$
$$P(H_3) = 0{,}4 \cdot 0{,}5 \cdot 0{,}7 = 0{,}14.$$

Условные вероятности события A (выход самолета из строя) при этих гипотезах равны:

$$P(A \mid H_0) = 0; \quad P(A \mid H_1) = 0{,}2; \quad P(A \mid H_2) = 0{,}6; \quad P(A \mid H_3) = 1{,}0.$$

Применяя формулу полной вероятности, получим:

$$P(A) = P(H_0)\,P(A \mid H_0) + P(H_1)\,P(A \mid H_1) + P(H_2)\,P(A \mid H_2) + P(H_3)\,P(A \mid H_3) = 0{,}36 \cdot 0{,}2 + 0{,}41 \cdot 0{,}6 + 0{,}14 \cdot 1{,}0 = 0{,}458.$$

Заметим, что первую гипотезу H_0 можно было бы не вводить в рассмотрение, так как соответствующий член в формуле полной вероятности обращается в нуль. Так обычно и поступают при применении формулы полной вероятности, рассматривая не полную группу несовместных гипотез, а только те из них, при которых данное событие возможно.

Пример 3. Работа двигателя контролируется двумя регуляторами. Рассматривается определенный период времени t, в течение которого желательно обеспечить безотказную работу двигателя. При наличии обоих регуляторов двигатель отказывает с вероятностью $q_{1,2}$, при работе только первого из них — с вероятностью q_1, при работе только второго — с вероятностью q_2, при отказе обоих регуляторов — с вероятностью q_0. Первый из регуляторов имеет надежность P_1, второй — P_2. Все элементы выходят из строя независимо друг от друга. Найти полную надежность (вероятность безотказной работы) двигателя.

Р е ш е н и е. Рассмотрим гипотезы:

$H_{1,2}$ — работают оба регулятора,
H_1 — работает только первый регулятор (второй вышел из строя),
H_2 — работает только второй регулятор (первый вышел из строя),
H_0 — оба регулятора вышли из строя

и событие

A — безотказная работа двигателя.

Вероятности гипотез равны:

$$P(H_{1,2}) = P_1P_2;\ P(H_1) = P_1(1-P_2);$$
$$P(H_2) = P_2(1-P_1);\ P(H_0) = (1-P_1)(1-P_2).$$

Условные вероятности события A при этих гипотезах заданы:

$$P(A \mid H_{1,2}) = 1 - q_{1,2};\ P(A \mid H_1) = 1 - q_1;\ P(A \mid H_2) = 1 - q_2;$$
$$P(A \mid H_0) = 1 - q_0.$$

По формуле полной вероятности получим:

$$P(A) = P_1P_2(1-q_{1,2}) + P_1(1-P_2)(1-q_1) + P_2(1-P_1)(1-q_2) + (1-P_1)(1-P_2)(1-q_0).$$

3.5. Теорема гипотез (формула Бейеса)

Следствием теоремы умножения и формулы полной вероятности является так называемая *теорема гипотез*, или *формула Бейеса*.

Поставим следующую задачу.

Имеется полная группа несовместных гипотез $H_1, H_2, \ldots, H_n$. Вероятности этих гипотез до опыта известны и равны соответственно $P(H_1), P(H_2), \ldots, P(H_n)$. Произведен опыт, в результате которого наблюдено появление некоторого события A. Спрашивается, как следует изменить вероятности гипотез в связи с появлением этого события?

Здесь, по существу, речь идет о том, чтобы найти условную вероятность $P(H_i \mid A)$ для каждой гипотезы.

Из теоремы умножения имеем:

$$P(AH_i) = P(A)P(H_i \mid A) = P(H_i)P(A \mid H_i) \quad (i = 1, 2, \ldots, n),$$

или, отбрасывая левую часть,

$$P(A)P(H_i \mid A) = P(H_i)P(A \mid H_i) \quad (i = 1, 2, \ldots, n),$$

откуда

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{P(A)} \quad (i = 1, 2, \ldots, n).$$

Выражая $P(A)$ с помощью формулы полной вероятности (3.4.1), имеем:

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{\sum_{i=1}^{n} P(H_i)P(A \mid H_i)} \quad (i = 1, 2, \ldots, n). \qquad (3.5.1)$$

Формула (3.5.1) и носит название *формулы Бейеса* или *теоремы гипотез.*

Пример 1. Прибор может собираться из высококачественных деталей и из деталей обычного качества; вообще около 40% приборов собирается из высококачественных деталей. Если прибор собран из высококачественных деталей, его надежность (вероятность безотказной работы) за время t равна 0,95; если из деталей обычного качества — его надежность равна 0,7. Прибор испытывался в течение времени t и работал безотказно. Найти вероятность того, что он собран из высококачественных деталей.

Решение. Возможны две гипотезы:

H_1 — прибор собран из высококачественных деталей,

H_2 — прибор собран из деталей обычного качества.

Вероятности этих гипотез до опыта:

$$P(H_1) = 0{,}4; \quad P(H_2) = 0{,}6.$$

В результате опыта наблюдено событие A — прибор безотказно работал время t.

Условные вероятности этого события при гипотезах H_1 и H_2 равны:

$$P(A \mid H_1) = 0{,}95; \quad P(A \mid H_2) = 0{,}7.$$

По формуле (3.5.1) находим вероятность гипотезы H_1 после опыта:

$$P(H_1) = \frac{0{,}4 \cdot 0{,}95}{0{,}4 \cdot 0{,}95 + 0{,}6 \cdot 0{,}6} = 0{,}475.$$

Пример 2. Два стрелка независимо один от другого стреляют по одной мишени, делая каждый по одному выстрелу. Вероятность попадания в мишень для первого стрелка 0,8, для второго 0,4. После стрельбы в мишени обнаружена одна пробоина. Найти вероятность того, что эта пробоина принадлежит первому стрелку.

Решение. До опыта возможны следующие гипотезы:

H_1 — ни первый, ни второй стрелок не попадет,

H_2 — оба стрелка попадут,

H_3 — первый стрелок попадет, а второй не попадет,

H_4 — первый стрелок не попадет, а второй попадет.

Вероятности этих гипотез:

$$P(H_1) = 0{,}2 \cdot 0{,}6 = 0{,}12;$$
$$P(H_2) = 0{,}8 \cdot 0{,}4 = 0{,}32;$$
$$P(H_3) = 0{,}8 \cdot 0{,}6 = 0{,}48;$$
$$P(H_4) = 0{,}2 \cdot 0{,}4 = 0{,}08.$$

Условные вероятности наблюденного события A при этих гипотезах равны:

$$P(A \mid H_1) = 0; \quad P(A \mid H_2) = 0; \quad P(A \mid H_3) = 1; \quad P(A \mid H_4) = 1.$$

После опыта гипотезы H_1 и H_2 становятся невозможными, а вероятности гипотез H_3 и H_4 будут равны:

$$P(H_3 \mid A) = \frac{0{,}48 \cdot 1}{0{,}48 \cdot 1 + 0{,}08 \cdot 1} = \frac{6}{7}; \quad P(H_4 \mid A) = \frac{0{,}08 \cdot 1}{0{,}48 \cdot 1 + 0{,}08 \cdot 1} = \frac{1}{7}.$$

Следовательно, вероятность того, что пробоина принадлежит первому стрелку, равна $\frac{6}{7}$.

Пример 3. Производится наблюдение за некоторым объектом с помощью двух наблюдательных станций. Объект может находиться в двух различных состояниях S_1 и S_2, случайно переходя из одного в другое. Долговременной практикой установлено, что примерно 30% времени объект находится в состоянии S_1, а 70% — в состоянии S_2. Наблюдательная станция № 1 передает ошибочные сведения приблизительно в 2% всех случаев, а наблюдательная станция № 2 — в 8%. В какой-то момент времени наблюдательная станция № 1 сообщила: объект находится в состоянии S_1, а наблюдательная станция № 2: объект находится в состоянии S_2.

Спрашивается: какому из сообщений следует верить?

Решение. Естественно верить тому из сообщений, для которого больше вероятность того, что оно соответствует истине. Применим формулу Бейеса. Для этого сделаем гипотезы о состоянии объекта:

H_1 — объект находится в состоянии S_1,

H_2 — объект находится в состоянии S_2.

Наблюденное событие A состоит в следующем: станция № 1 сообщила, что объект находится в состоянии S_1, а станция № 2 — что он находится в состоянии S_2. Вероятности гипотез до опыта

$$P(H_1) = 0{,}3; \quad P(H_2) = 0{,}7.$$

Найдем условные вероятности наблюденного события A при этих гипотезах. При гипотезе H_1, чтобы произошло событие A, нужно, чтобы первая станция передала верное сообщение, а вторая — ошибочное:

$$P(A \mid H_1) = 0{,}98 \cdot 0{,}08 = 0{,}0784.$$

Аналогично

$$P(A \mid H_2) = 0{,}92 \cdot 0{,}02 = 0{,}0184.$$

Применяя формулу Бейеса, найдем вероятность того, что истинное состояние объекта — S_1:

$$P(H_1 \mid A) = \frac{0.3 \cdot 0{,}0784}{0{,}3 \cdot 0{,}0784 + 0{,}7 \cdot 0{,}0184} \approx 0{,}645,$$

т. е. из двух переданных сообщений более правдоподобным является сообщение первой станции.

ГЛАВА 4

ПОВТОРЕНИЕ ОПЫТОВ

4.1. Частная теорема о повторении опытов

При практическом применении теории вероятностей часто приходится встречаться с задачами, в которых один и тот же опыт или аналогичные опыты повторяются неоднократно. В результате каждого опыта может появиться или не появиться некоторое событие A, причем нас интересует не результат каждого отдельного опыта, а общее число появлений события A в результате серии опытов. Например, если производится группа выстрелов по одной и той же цели, нас, как правило, интересует не результат каждого выстрела, а общее число попаданий. В подобных задачах требуется уметь определять вероятность любого заданного числа появлений события в результате серии опытов. Такие задачи и будут рассмотрены в данной главе. Они решаются весьма просто в случае, когда опыты являются независимыми.

Несколько опытов называются *независимыми*, если вероятность того или иного исхода каждого из опытов не зависит от того, какие исходы имели другие опыты. Например, несколько последовательных бросаний монеты представляют собой независимые опыты. Несколько последовательных вынимаций карты из колоды представляют собой независимые опыты при условии, что вынутая карта каждый раз возвращается в колоду и карты перемешиваются; в противном случае это — зависимые опыты. Несколько выстрелов представляют собой независимые опыты только в случае, если прицеливание производится заново перед каждым выстрелом; в случае, когда прицеливание производится один раз перед всей стрельбой или непрерывно осуществляется в процессе стрельбы (стрельба очередью, бомбометание серией) выстрелы представляют собой зависимые опыты.

Независимые опыты могут производиться в одинаковых или различных условиях. В первом случае вероятность события A во всех опытах одна и та же. Во втором случае вероятность события A от опыта к опыту меняется. К первому случаю относится *частная теорема*, а ко второму — *общая теорема о повторении опытов*.

Мы начнем с частной теоремы, как более элементарной. Прежде всего рассмотрим конкретный пример.

Пример. Производится три независимых выстрела по мишени, вероятность попадания в которую при каждом выстреле равна p. Найти вероятность того, что при этих трех выстрелах мы получим ровно два попадания.

Решение. Обозначим B_2 событие, состоящее в том, что в мишень попадет ровно два снаряда. Это событие может произойти тремя способами:

1) попадание при первом выстреле, попадание при втором, промах при третьем;

2) попадание при первом выстреле, промах при втором, попадание при третьем;

3) промах при первом выстреле, попадание при втором, попадание при третьем.

Следовательно, событие B_2 можно представить как сумму произведений событий:

$$B_2 = A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3,$$

где A_1, A_2, A_3 — попадание при первом, втором, третьем выстрелах соответственно, $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ — промах при первом, втором, третьем выстрелах.

Учитывая, что три перечисленных варианта события B_2 несовместны, а события, входящие в произведения, независимы, по теоремам сложения и умножения получим:

$$P(B_2) = pp(1-p) + p(1-p)p + (1-p)pp,$$

или, обозначая $1-p=q$,

$$P(B_2) = 3p^2q.$$

Аналогичным образом, перечисляя все возможные варианты, в которых интересующее нас событие может появиться заданное число раз, можно решить и следующую общую задачу.

Производится n независимых опытов, в каждом из которых может появиться или не появиться некоторое событие A; вероятность появления события A в каждом опыте равна p, а вероятность непоявления $q=1-p$. Требуется найти вероятность $P_{m,n}$ того, что событие A в этих n опытах появится ровно m раз.

Рассмотрим событие B_m, состоящее в том, что событие A появится в n опытах ровно m раз. Это событие может осуществиться различными способами. Разложим событие B_m на сумму произведений событий, состоящих в появлении или непоявлении события A в отдельном опыте. Будем обозначать A_i появление события A в i-м опыте; $\bar{A}_i$ — непоявление события A в i-м опыте.

Очевидно, каждый вариант появления события B_m (каждый член суммы) должен состоять из m появлений события A и $n-m$ непоявлений, т. е. из m событий A и $n-m$ событий $\bar{A}$ с различными индексами. Таким образом,

$$B_m = A_1A_2 \dots A_m\bar{A}_{m+1} \dots \bar{A}_n + \dots$$
$$\dots + A_1\bar{A}_2A_3 \dots \bar{A}_{n-1}A_n + \dots$$
$$\dots + \bar{A}_1\bar{A}_2 \dots \bar{A}_{n-m}A_{n-m+1} \dots A_n,$$

причем в каждое произведение событие A должно входить m раз, а $\bar{A}$ должно входить $n-m$ раз.

Число всех комбинаций такого рода равно C_n^m, т. е. числу способов, какими можно из n опытов выбрать m, в которых произошло событие A. Вероятность каждой такой комбинации, по теореме умножения для независимых событий, равна $p^m q^{n-m}$. Так как комбинации между собой несовместны, то, по теореме сложения, вероятность события B_m равна

$$P_{m,n} = \underbrace{p^m q^{n-m} + \ldots + p^m q^{n-m}}_{C_n^m \text{ раз}} = C_n^m p^m q^{n-m}.$$

Таким образом, мы можем дать следующую формулировку частной теоремы о повторении опытов.

Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p, то вероятность того, что событие A появится ровно m раз, выражается формулой

$$P_{m,n} = C_n^m p^m q^{n-m}, \tag{4.1.1}$$

где $q = 1 - p$.

Формула (4.1.1) описывает, как распределяются вероятности между возможными значениями некоторой случайной величины — числа появлений события A при n опытах.

В связи с тем, что вероятности $P_{m,n}$ по форме представляют собой члены разложения бинома $(q+p)^n$, распределение вероятностей вида (4.1.1) называется *биномиальным распределением*.

4.2. Общая теорема о повторении опытов

Частная теорема о повторении опытов касается того случая, когда вероятность события A во всех опытах одна и та же. На практике часто приходится встречаться с более сложным случаем, когда опыты производятся в неодинаковых условиях, и вероятность события от опыта к опыту меняется. Например, если производится ряд выстрелов в переменных условиях (скажем, при изменяющейся дальности), то вероятность попадания от выстрела к выстрелу может заметно меняться.

Способ вычисления вероятности заданного числа появлений события в таких условиях дает общая теорема о повторении опытов.

Пусть производится n независимых опытов, в каждом из которых может появиться или не появиться некоторое событие A, причем вероятность появления события A в i-м опыте равна p_i, а вероятность непоявления $q_i = 1 - p_i$ ($i = 1, \ldots, n$). Требуется найти вероятность $P_{m,n}$ того, что в результате n опытов событие A появится ровно m раз.

Обозначим по-прежнему B_m событие, состоящее в том, что событие A появится m раз в n опытах. По-прежнему представим B_m как сумму произведений элементарных событий:

$$B_m = A_1A_2 \dots A_m\bar{A}_{m+1} \dots \bar{A}_n + \dots$$
$$\dots + A_1\bar{A}_2A_3 \dots \bar{A}_{n-1}A_n + \dots$$
$$\dots + \bar{A}_1\bar{A}_2 \dots \bar{A}_{n-m}A_{n-m+1} \dots A_n,$$

причем в каждое из произведений событие A входит m раз, событие $\bar{A}$ $n-m$ раз. Число таких комбинаций по-прежнему будет C_n^m, но сами комбинации между собой будут уже неравновероятны.

Применяя теорему сложения и теорему умножения для независимых событий, получим:

$$P_{m,n} = p_1p_2 \dots p_mq_{m+1} \dots q_n + \dots$$
$$\dots + p_1q_2p_3 \dots q_{n-1}p_n + \dots$$
$$\dots + q_1q_2 \dots q_{n-m}p_{n-m+1} \dots p_n,$$

т. е. искомая вероятность равна сумме всех возможных произведений, в которые буквы p с разными индексами входят m раз, а буквы q с разными индексами $n-m$ раз.

Для того чтобы чисто механически составлять все возможные произведения из m букв p и $n-m$ букв q с разными индексами, применим следующий формальный прием. Составим произведение n биномов:

$$\varphi_n(z) = (q_1 + p_1z)(q_2 + p_2z) \dots (q_n + p_nz)$$

или короче

$$\varphi_n(z) = \prod_{l=1}^{n}(q_l + p_lz),$$

где z — произвольный параметр.

Зададимся целью найти в этом произведении биномов коэффициент при z^m. Для этого перемножим биномы и произведем приведение подобных членов. Очевидно, каждый член, содержащий z^m, будет иметь в качестве коэффициента произведение m букв p с какими-то индексами и $n-m$ букв q, а после приведения подобных членов коэффициент при z^m будет представлять собой сумму всех возможных произведений такого типа. Следовательно, способ составления этого коэффициента полностью совпадает со способом вычисления вероятности $P_{m,n}$ в задаче о повторении опытов.

Функция $\varphi_n(z)$, разложение которой по степеням параметра z дает в качестве коэффициентов вероятности $P_{m,n}$, называется *производящей функцией вероятностей* $P_{m,n}$ или просто *производящей функцией*.

Пользуясь понятием производящей функции, можно сформулировать общую теорему о повторении опытов в следующем виде.

Вероятность того, что событие A в n независимых опытах появится ровно m раз, равна коэффициенту при z^m в выражении производящей функции:

$$\varphi_n(z) = \prod_{l=1}^{n} (q_l + p_l z),$$

где p_l — вероятность появления события A в l-м опыте, $q_l = 1 - p_l$.

Вышеприведенная формулировка общей теоремы о повторении опытов в отличие от частной теоремы не дает явного выражения для вероятности $P_{m,n}$. Такое выражение в принципе написать можно, но оно является слишком сложным, и мы не будем его приводить. Однако, не прибегая к такому явному выражению, все же можно записать общую теорему о повторении опытов в виде одной формулы:

$$\prod_{l=1}^{n} (q_l + p_l z) = \sum_{m=0}^{n} P_{m,n} z^m. \tag{4.2.1}$$

Левая и правая части равенства (4.2.1) представляют собою одну и ту же производящую функцию $\varphi_n(z)$, только слева она написана в виде одночлена, а справа — в виде многочлена. Раскрывая скобки в левой части и выполняя приведение подобных членов, получим все вероятности:

$$P_{0,n},\ P_{1,n},\ \ldots,\ P_{n,n}$$

как коэффициенты соответственно при нулевой, первой и т. д. степенях z.

Очевидно, частная теорема о повторении опытов вытекает из общей при

$$p_1 = p_2 = \ldots = p_n = p,$$

$$q_1 = q_2 = \ldots = q_n = q.$$

В этом случае производящая функция обращается в n-ю степень бинома $(q + pz)$:

$$\varphi_n(z) = (q + pz)^n.$$

Раскрывая это выражение по формуле бинома, имеем:

$$(q + pz)^n = \sum_{m=0}^{n} C_n^m p^m q^{n-m} z^m,$$

откуда следует формула (4.1.1).

Отметим, что как в общем, так и в частном случае сумма всех вероятностей $P_{m,n}$ равна единице:

$$\sum_{m=0}^{n} P_{m,n} = 1. \tag{4.2.2}$$

Это следует прежде всего из того, что события $B_0, B_1, \ldots, B_n$ образуют полную группу несовместных событий. Формально к равенству (4.2.2) можно придти, полагая в общей формуле (4.2.1) $z = 1$.

Во многих случаях практики, кроме вероятности $P_{m,n}$ ровно m появлений события A, приходится рассматривать вероятность не менее m появлений события A.

Обозначим C_m событие, состоящее в том, что событие A появится не менее m раз, а вероятность события C_m обозначим $R_{m,n}$. Очевидно,

$$C_m = B_m + B_{m+1} + \ldots + B_n,$$

откуда, по теореме сложения,

$$R_{m,n} = P_{m,n} + P_{m+1,n} + \ldots + P_{n,n},$$

или короче

$$R_{m,n} = \sum_{l=m}^{n} P_{l,n}. \tag{4.2.3}$$

При вычислении $R_{m,n}$ часто бывает удобнее не пользоваться непосредственно формулой (4.2.3), а переходить к противоположному событию и вычислять вероятность $R_{m,n}$ по формуле

$$R_{m,n} = 1 - \sum_{l=0}^{m-1} P_{l,n}. \tag{4.2.4}$$

Пример 1. Производится 4 независимых выстрела по одной и той же цели с различных расстояний; вероятности попадания при этих выстрелах равны соответственно

$$p_1 = 0{,}1; \quad p_2 = 0{,}2; \quad p_3 = 0{,}3; \quad p_4 = 0{,}4.$$

Найти вероятности ни одного, одного, двух, трех и четырех попаданий:

$$P_{0,4}; \quad P_{1,4}; \quad P_{2,4}; \quad P_{3,4}; \quad P_{4,4}.$$

Решение. Составляем производящую функцию:

$$\varphi_4(z) = \prod_{l=1}^{4} (q_l + p_l z) =$$
$$= (0{,}9 + 0{,}1z)(0{,}8 + 0{,}2z)(0{,}7 + 0{,}3z)(0{,}6 + 0{,}4z) =$$
$$= 0{,}302 + 0{,}440z + 0{,}215z^2 + 0{,}040z^3 + 0{,}002z^4,$$

откуда

$$P_{0,4} = 0{,}302; \quad P_{1,4} = 0{,}440; \quad P_{2,4} = 0{,}215; \quad P_{3,4} = 0{,}040; \quad P_{4,4} = 0{,}002.$$

Пример 2. Производится 4 независимых выстрела в одинаковых условиях, причем вероятность попадания p есть средняя из вероятностей p_1, p_2, p_3, p_4 предыдущего примера:

$$p = \frac{1}{4}(p_1 + p_2 + p_3 + p_4) = 0{,}25.$$

Найти вероятности

$$P_{0,4};\quad P_{1,4};\quad P_{2,4};\quad P_{3,4};\quad P_{4,4}.$$

Решение. По формуле (4.1.1) имеем:

$$P_{0,4} = q^4 = 0{,}316;$$

$$P_{1,4} = C_4^1 p q^3 = 0{,}421;$$

$$P_{2,4} = C_4^2 p^2 q^2 = 0{,}211;$$

$$P_{3,4} = C_4^3 p^3 q = 0{,}047;$$

$$P_{4,4} = p^4 = 0{,}004.$$

Пример 3. Имеется 5 станций, с которыми поддерживается связь. Время от времени связь прерывается из-за атмосферных помех. Вследствие удаленности станций перерыв друг от друга связи с каждой из них происходит независимо от остальных с вероятностью $p = 0{,}2$. Найти вероятность того, что в данный момент времени будет иметься связь не более чем с двумя станциями.

Решение. Событие, о котором идет речь, сводится к тому, что будет нарушена связь не менее чем с тремя станциями. По формуле (4.2.3) получим:

$$R_{3,5} = P_{3,5} + P_{4,5} + P_{5,5} = C_5^3 \cdot 0{,}2^3 \cdot 0{,}8^2 + C_5^4 \cdot 0{,}2^4 \cdot 0{,}8 + 0{,}2^5 =$$

$$= 0{,}0512 + 0{,}0064 + 0{,}0003 = 0{,}0579.$$

Пример 4. Система радиолокационных станций ведет наблюдение за группой объектов, состоящей из 10 единиц. Каждый из объектов может быть (независимо от других) потерян с вероятностью 0,1. Найти вероятность того, что хотя бы один из объектов будет потерян.

Решение. Вероятность потери хотя бы одного объекта $R_{1,10}$ можно было бы найти по формуле

$$R_{1,10} = P_{1,10} + P_{2,10} + \ldots + P_{10,10},$$

но несравненно проще воспользоваться вероятностью противоположного события — ни один объект не потерян — и вычесть ее из единицы:

$$R_{1,10} = 1 - P_{0,10} = 1 - 0{,}9^{10} \approx 0{,}65.$$

Пример 5. Прибор состоит из 8 однородных элементов, но может работать при наличии в исправном состоянии не менее 6 из них. Каждый из элементов за время работы прибора t выходит из строя независимо от других с вероятностью 0,2. Найти вероятность того, что прибор откажет за время t.

Решение. Для отказа прибора требуется выход из строя не менее двух из восьми элементов. По формуле (4.2.4) имеем:

$$R_{2,8} = 1 - (P_{0,8} + P_{1,8}) = 1 - \left(0{,}8^8 + C_8^1 \cdot 0{,}2 \cdot 0{,}8^7\right) \approx 0{,}497.$$

Пример 6. Производится 4 независимых выстрела с самолета по самолету. Вероятность попадания при каждом выстреле равна 0,3. Для

поражения (выхода из строя) самолета заведомо достаточно двух попаданий; при одном попадании самолет поражается с вероятностью 0.6. Найти вероятность того, что самолет будет поражен.

Решение. Задача решается по формуле полной вероятности. Можно было бы рассмотреть гипотезы

H_1 — в самолет попал 1 снаряд,

H_2 — в самолет попало 2 снаряда,

H_3 — в самолет попало 3 снаряда,

H_4 — в самолет попало 4 снаряда

и находить вероятность события A — поражения самолета — с помощью этих четырех гипотез. Однако значительно проще рассмотреть всего две гипотезы:

H_0 — в самолет не попало ни одного снаряда,

H_1 — в самолет попал 1 снаряд,

и вычислять вероятность события $\bar{A}$ — непоражения самолета:

$$P(\bar{A}) = P(H_0)\,P(\bar{A} \mid H_0) + P(H_1)\,P(\bar{A} \mid H_1).$$

Имеем:

$$P(H_0) = P_{0,4} = 0{,}7^4 = 0{,}240;$$

$$P(H_1) = P_{1,4} = C_4^1 \cdot 0{,}3 \cdot 0{,}7^3 = 0{,}412;$$

$$P(\bar{A} \mid H_0) = 1;$$

$$P(\bar{A} \mid H_1) = 1 - 0{,}6 = 0{,}4.$$

Следовательно,

$$P(\bar{A}) = 0{,}240 + 0{,}412 \cdot 0{,}4 \approx 0{,}405,$$

откуда

$$P(A) = 1 - 0{,}405 = 0{,}595.$$

ГЛАВА 5

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ

5.1. Ряд распределения. Многоугольник распределения

В разделе курса, посвященном основным понятиям теории вероятностей, мы уже ввели в рассмотрение чрезвычайно важное понятие *случайной величины*. Здесь мы дадим дальнейшее развитие этого понятия и укажем способы, с помощью которых случайные величины могут быть описаны и характеризованы.

Как уже было сказано, случайной величиной называется величина, которая в результате опыта может принять то или иное значение, неизвестно заранее — какое именно. Мы условились также различать случайные величины прерывного (дискретного) и непрерывного типа. Возможные значения прерывных величин могут быть заранее перечислены. Возможные значения непрерывных величин не могут быть заранее перечислены и непрерывно заполняют некоторый промежуток.

Примеры прерывных случайных величин:

1) число появлений герба при трех бросаниях монеты (возможные значения 0, 1, 2, 3);

2) частота появления герба в том же опыте (возможные значения $0, \frac{1}{3}, \frac{2}{3}, 1$);

3) число отказавших элементов в приборе, состоящем из пяти элементов (возможные значения 0, 1, 2, 3, 4, 5);

4) число попаданий в самолет, достаточное для вывода его из строя (возможные значения $1, 2, 3, \ldots, n, \ldots$);

5) число самолетов, сбитых в воздушном бою (возможные значения $0, 1, 2, \ldots, N$, где N — общее число самолетов, участвующих в бою).

Примеры непрерывных случайных величин:

1) абсцисса (ордината) точки попадания при выстреле;

2) расстояние от точки попадания до центра мишени;

3) ошибка измерителя высоты;

4) время безотказной работы радиолампы.

Условимся в дальнейшем случайные величины обозначать большими буквами, а их возможные значения — соответствующими малыми буквами. Например, X — число попаданий при трех выстрелах; возможные значения: $x_1 = 0$; $x_2 = 1$; $x_3 = 2$; $x_4 = 3$.

Рассмотрим прерывную случайную величину X с возможными значениями $x_1, x_2, \ldots, x_n$. Каждое из этих значений возможно, но не достоверно, и величина X может принять каждое из них с некоторой вероятностью. В результате опыта величина X примет одно из этих значений, т. е. произойдет одно из полной группы несовместных событий:

$$\left.\begin{aligned} X &= x_1, \\ X &= x_2, \\ &\ldots \\ X &= x_n. \end{aligned}\right\} \qquad (5.1.1)$$

Обозначим вероятности этих событий буквами p с соответствующими индексами:

$$P(X = x_1) = p_1; \quad P(X = x_2) = p_2; \ \ldots; \ P(X = x_n) = p_n.$$

Так как несовместные события (5.1.1) образуют полную группу, то

$$\sum_{i=1}^{n} p_i = 1,$$

т. е. сумма вероятностей всех возможных значений случайной величины равна единице. Эта суммарная вероятность каким-то образом р а с п р е д е л е н а между отдельными значениями. Случайная величина будет полностью описана с вероятностной точки зрения, если мы зададим это распределение, т. е. в точности укажем, какой вероятностью обладает каждое из событий (5.1.1). Этим мы установим так называемый *закон распределения* случайной величины.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину мы будем говорить, что она *подчинена* данному закону распределения.

Установим форму, в которой может быть задан закон распределения прерывной случайной величины X. Простейшей формой задания этого закона является таблица, в которой перечислены возможные значения случайной величины и соответствующие им вероятности:

x_i	x_1	x_2	. . .	x_n
p_i	p_1	p_2	. . .	p_n

Такую таблицу мы будем называть *рядом распределения* случайной величины X.

Чтобы придать ряду распределения более наглядный вид, часто прибегают к его графическому изображению: по оси абсцисс откладываются возможные значения случайной величины, а по оси ординат — вероятности этих значений. Для наглядности полученные точки соединяются отрезками прямых. Такая фигура называется *многоугольником распределения* (рис. 5.1.1). Многоугольник распределения,

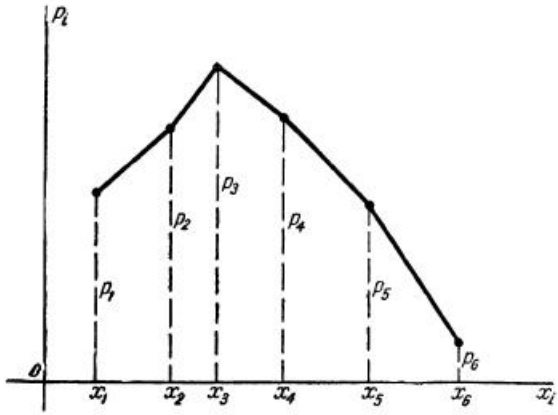


Рис. 5.1.1.

так же как и ряд распределения, полностью характеризует случайную величину; он является одной из форм закона распределения.

Иногда удобной оказывается так называемая «механическая» интерпретация ряда распределения. Представим себе, что некоторая масса, равная единице, распределена по оси абсцисс так, что в n отдельных точках $x_1, x_2, \ldots, x_n$ сосредоточены соответственно массы $p_1, p_2, \ldots, p_n$. Тогда ряд распределения интерпретируется как система материальных точек с какими-то массами, расположенных на оси абсцисс.

Рассмотрим несколько примеров прерывных случайных величин с их законами распределения.

Пример 1. Производится один опыт, в котором может появиться или не появиться событие A. Вероятность события A равна 0,3. Рассматривается случайная величина X — число появлений события A в данном опыте (т. е. характеристическая случайная величина события A, принимающая значение 1, если оно появится, и 0, если не появится). Построить ряд распределения и многоугольник распределения величины X.

Р е ш е н и е. Величина X имеет всего два значения: 0 и 1. Ряд распределения величины X имеет вид:

x_i	0	1
p_i	0,7	0,3

Многоугольник распределения изображен на рис. 5.1.2.

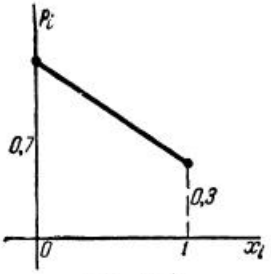


Рис. 5.1.2.

Пример 2. Стрелок производит три выстрела по мишени. Вероятность попадания в мишень при каждом выстреле равна 0,4. За каждое попадание стрелку засчитывается 5 очков. Построить ряд распределения числа выбитых очков.

Р е ш е н и е. Обозначим X число выбитых очков. Возможные значения величины X: $x_1 = 0$; $x_2 = 5$; $x_3 = 10$; $x_4 = 15$.

Вероятности этих значений находим по теореме о повторении опытов:

$$p_1 = 0{,}6^3 = 0{,}216; \quad p_2 = C_3^1 \cdot 0{,}4 \cdot 0{,}6^2 = 0{,}432;$$

$$p_3 = C_3^2 \cdot 0{,}4^2 \cdot 0{,}6 = 0{,}288; \quad p_4 = 0{,}4^3 = 0{,}064.$$

Ряд распределения величины X имеет вид:

x_i	0	5	10	15
p_i	0,216	0,432	0,288	0,064

Многоугольник распределения изображен на рис. 5.1.3.

Пример 3. Вероятность появления события A в одном опыте равна p. Производится ряд независимых опытов, которые продолжаются до первого появления события A, после чего опыты прекращаются. Случайная величина X — число произведенных опытов. Построить ряд распределения величины X.

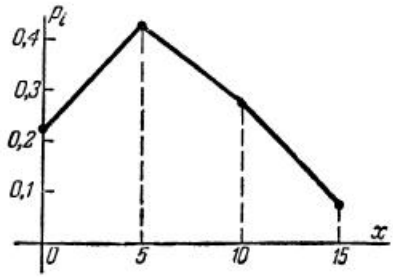


Рис. 5.1.3.

Р е ш е н и е. Возможные значения величины X: 1, 2, 3, ... (теоретически они ничем не ограничены). Для того чтобы величина X приняла значение 1, необходимо, чтобы событие A произошло в первом же опыте; вероятность этого равна p. Для того чтобы величина X приняла значение 2, нужно, чтобы в первом опыте событие A не появилось, а во втором — появилось; вероятность этого равна qp, где $q = 1 - p$, и т. д. Ряд

распределения величины X имеет вид:

x_i	1	2	3	. . .	i	. . .
p_i	p	pq	pq^2	. . .	pq^{i-1}	. . .

Первые пять ординат многоугольника распределения для случая $p = q = 0{,}5$ показаны на рис. 5.1.4.

Пример 4. Стрелок ведет стрельбу по мишени до первого попадания, имея боезапас 4 патрона. Вероятность попадания при каждом выстреле равна 0,6. Построить ряд распределения боезапаса, оставшегося неизрасходованным.

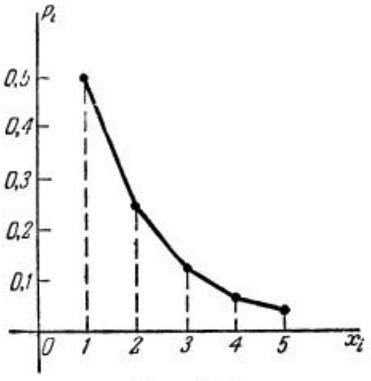


Рис. 5.1.4.

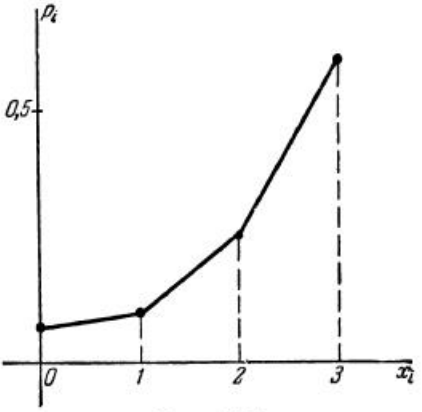


Рис. 5.1.5.

Решение. Случайная величина X — число неизрасходованных патронов — имеет четыре возможных значения: 0, 1, 2 и 3. Вероятности этих значений равны соответственно:

$$p_0 = 0{,}4^3 = 0{,}064;$$
$$p_1 = 0.4^2 \cdot 0{,}6 = 0{,}096;$$
$$p_2 = 0{,}4 \cdot 0{,}6 = 0{,}240;$$
$$p_3 = 0{,}600.$$

Ряд распределения величины X имеет вид:

x_i	0	1	2	3
p_i	0,064	0,096	0,240	0,600

Многоугольник распределения показан на рис. 5.1.5.

Пример 5. Техническое устройство может применяться в различных условиях и в зависимости от этого время от времени требует регулировки. При однократном применении устройства оно может случайным образом попасть в благоприятный или неблагоприятный режим. В благоприятном режиме устройство выдерживает три применения без регулировки; перед четвертым его приходится регулировать. В неблагоприятном режиме устройство приходится регулировать после первого же применения. Вероятность того, что устройство попадает в благоприятный режим, — 0,7, что в неблагоприятный, — 0,3. Рассматривается случайная величина X — число применений устройства до регулировки. Построить ее ряд распределения.

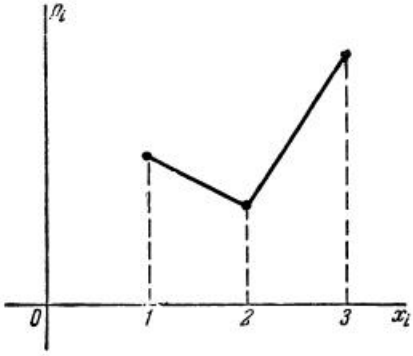


Рис. 5.1.6.

Решение. Случайная величина X имеет три возможных значения: 1, 2 и 3. Вероятность того, что $X = 1$, равна вероятности того, что при первом же применении устройство попадет в неблагоприятный режим, т. е. $p_1 = 0{,}3$. Для того чтобы величина X приняла значение 2, нужно, чтобы при первом применении устройство попало в благоприятный режим, а при втором — в неблагоприятный; вероятность этого $p_2 = 0{,}7 \cdot 0{,}3 = 0{,}21$. Чтобы величина X приняла значение 3, нужно, чтобы два первых раза устройство попало в благоприятный режим (после третьего раза его все равно придется регулировать). Вероятность этого равна $p_3 = 0{,}7^2 = 0{,}49$.

Ряд распределения величины X имеет вид:

x_i	1	2	3
p_i	0,30	0,21	0,49

Многоугольник распределения показан на рис. 5.1.6.

5.2. Функция распределения

В предыдущем п° мы ввели в рассмотрение ряд распределения как исчерпывающую характеристику (закон распределения) прерывной случайной величины. Однако эта характеристика не является универсальной; она существует только для прерывных случайных величин. Нетрудно убедиться, что для непрерывной случайной величины такой характеристики построить нельзя. Действительно, непрерывная случайная величина имеет бесчисленное множество возможных значений, сплошь заполняющих некоторый промежуток (так называемое «несчетное множество»). Составить таблицу, в которой были бы перечислены все возможные значения такой случайной величины, не-

возможно. Кроме того, как мы увидим в дальнейшем, каждое отдельное значение непрерывной случайной величины обычно не обладает никакой отличной от нуля вероятностью. Следовательно, для непрерывной случайной величины не существует ряда распределения в том смысле, в каком он существует для прерывной величины. Однако различные области возможных значений случайной величины все же не являются одинаково вероятными, и для непрерывной величины существует «распределение вероятностей», хотя и не в том смысле, как для прерывной.

Для количественной характеристики этого распределения вероятностей удобно воспользоваться не вероятностью события $X = x$, а вероятностью события $X < x$, где x — некоторая текущая переменная. Вероятность этого события, очевидно, зависит от x, есть некоторая функция от x. Эта функция называется *функцией распределения* случайной величины X и обозначается $F(x)$:

$$F(x) = P(X < x). \tag{5.2.1}$$

Функцию распределения $F(x)$ иногда называют также *интегральной функцией распределения или интегральным законом распределения*.

Функция распределения — самая универсальная характеристика случайной величины. Она существует для всех случайных величин: как прерывных, так и непрерывных. Функция распределения полностью характеризует случайную величину с вероятностной точки зрения, т. е. является одной из форм закона распределения.

Сформулируем некоторые общие свойства функции распределения.

1. Функция распределения $F(x)$ есть неубывающая функция своего аргумента, т. е. при $x_2 > x_1$ $F(x_2) \geqslant F(x_1)$.

2. На минус бесконечности функция распределения равна нулю:

$$F(-\infty) = 0.$$

3. На плюс бесконечности функция распределения равна единице:

$$F(+\infty) = 1.$$

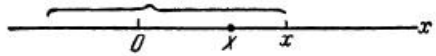


Рис. 5.2.1.

Не давая строгого доказательства этих свойств, проиллюстрируем их с помощью наглядной геометрической интерпретации. Для этого будем рассматривать случайную величину X как случайную точку X на оси Ox (рис. 5.2.1), которая в результате опыта может занять то или иное положение. Тогда функция распределения $F(x)$ есть вероятность того, что случайная точка X в результате опыта попадет левее точки x.

Будем увеличивать x, т. е. перемещать точку x вправо по оси абсцисс. Очевидно, при этом вероятность того, что случайная точка X

попадет левее x, не может уменьшиться; следовательно, функция распределения $F(x)$ с возрастанием x убывать не может.

Чтобы убедиться в том, что $F(-\infty)=0$, будем неограниченно перемещать точку x влево по оси абсцисс. При этом попадание случайной точки X левее x в пределе становится невозможным событием; естественно полагать, что вероятность этого события стремится к нулю, т. е. $F(-\infty)=0$.

Аналогичным образом, неограниченно перемещая точку x вправо, убеждаемся, что $F(+\infty)=1$, так как событие $X<x$ становится в пределе достоверным.

График функции распределения $F(x)$ в общем случае представляет собой график неубывающей функции (рис. 5.2.2), значения которой начинаются от 0 и доходят до 1, причем в отдельных точках функция может иметь скачки (разрывы).

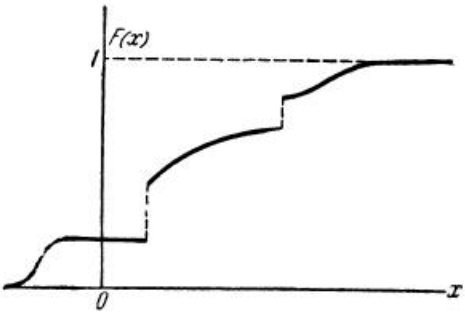


Рис. 5.2.2.

Зная ряд распределения прерывной случайной величины, можно легко построить функцию распределения этой величины. Действительно,

$$F(x)=P(X<x)=\sum_{x_i<x}P(X=x_i),$$

где неравенство $x_i<x$ под знаком суммы указывает, что суммирование распространяется на все те значения x_i, которые меньше x.

Когда текущая переменная x проходит через какое-нибудь из возможных значений прерывной величины X, функция распределения меняется скачкообразно, причем величина скачка равна вероятности этого значения.

Пример 1. Производится один опыт, в котором может появиться или не появиться событие A. Вероятность события A равна 0,3. Случайная величина X — число появлений события A в опыте (характеристическая случайная величина события A). Построить ее функцию распределения.

Решение. Ряд распределения величины X имеет вид:

x_i	0	1
p_i	0,7	0,3

Построим функцию распределения величины X:

1) при $x \leqslant 0$

$$F(x) = P(X < x) = 0;$$

2) при $0 < x \leqslant 1$

$$F(x) = P(X < x) = P(X = 0) = 0{,}7;$$

3) при $x > 1$

$$F(x) = P(X < x) = P(X = 0) + P(X = 1) = 1.$$

График функции распределения представлен на рис. 5.2.3. В точках разрыва функция $F(x)$ принимает значения, отмеченные на чертеже точками (функция непрерывна слева).

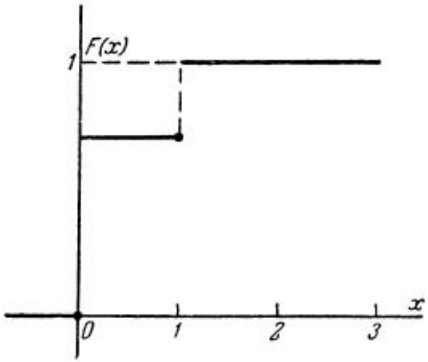


Рис. 5.2.3.

Пример. 2. В условиях предыдущего примера производится 4 независимых опыта. Построить функцию распределения числа появлений события A.

Решение. Обозначим X — число появлений события A в четырех опытах. Эта величина имеет ряд распределения

x_i	0	1	2	3	4
p_i	0,2401	0,4116	0,2646	0,0756	0,0081

Построим функцию распределения случайной величины X:

1) при $x \leqslant 0$ $F(x) = 0$;
2) при $0 < x \leqslant 1$ $F(x) = 0{,}2401$;
3) при $1 < x \leqslant 2$ $F(x) = 0{,}6517$;
4) при $2 < x \leqslant 3$ $F(x) = 0{,}9163$;
5) при $3 < x \leqslant 4$ $F(x) = 0{,}9919$;
6) при $x > 4$ $F(x) = 1$.

График функции распределения представлен на рис. 5.2.4.

Функция распределения любой прерывной случайной величины всегда есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной

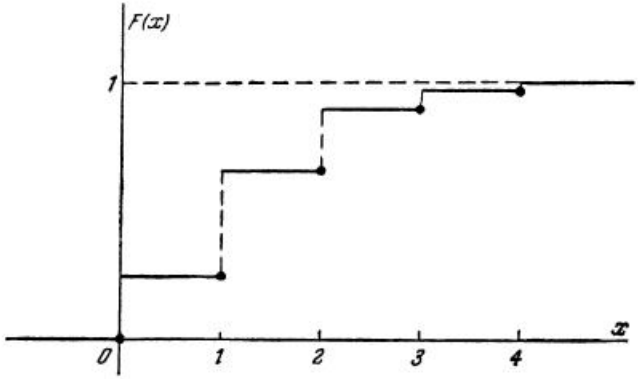


Рис. 5.2.4.

величины, и равны вероятностям этих значений. Сумма всех скачков функции $F(x)$ равна единице.

По мере увеличения числа возможных значений случайной величины и уменьшения интервалов между ними число скачков становится больше, а сами скачки — меньше; ступенчатая кривая становится

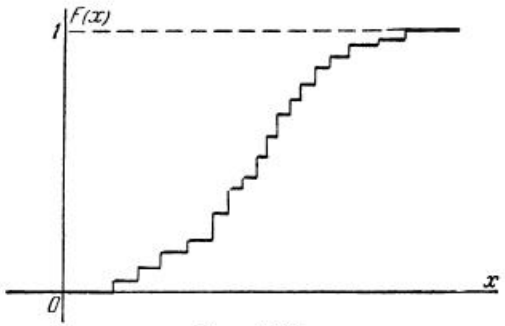


Рис. 5.2.5.

более плавной (рис. 5.2.5); случайная величина постепенно приближается к непрерывной величине, а ее функция распределения — к непрерывной функции (рис. 5.2.6).

На практике обычно функция распределения непрерывной случайной величины представляет собой функцию, непрерывную во всех

точках, как это показано на рис. 5.2.6. Однако можно построить примеры случайных величин, возможные значения которых непрерывно заполняют некоторый промежуток, но для которых функция распределения не везде является непрерывной, а в отдельных точках терпит

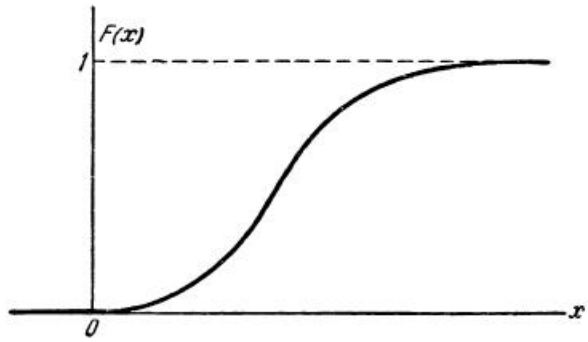


Рис. 5.2.6.

разрывы (рис. 5.2.7). Такие случайные величины называются *смешанными*. В качестве примера смешанной величины можно привести площадь разрушений, наносимых цели бомбой, радиус разрушительного действия которой равен R (рис. 5.2.8). Значения этой случайной величины непрерывно заполняют промежуток от 0 до πR^2, но при

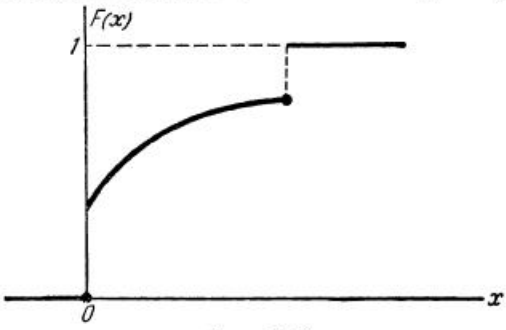


Рис. 5.2.7.

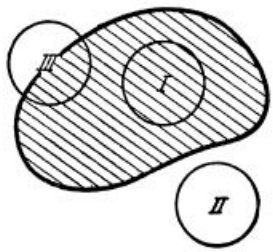


Рис. 5.2.8.

этом крайние значения промежутка 0 и πR^2, осуществляющиеся при положениях бомбы типа *I* и *II*, обладают определенной конечной вероятностью, и этим значениям соответствуют скачки функции распределения, тогда как в промежуточных значениях (положение типа *III*) функция распределения непрерывна. Другой пример смешанной случайной величины — время T безотказной работы прибора, испытываемого в течение времени t. Функция распределения этой случайной величины непрерывна всюду, кроме точки t.

5.3. Вероятность попадания случайной величины на заданный участок

При решении практических задач, связанных со случайными величинами, часто оказывается необходимым вычислять вероятность того, что случайная величина примет значение, заключенное в некоторых пределах, например от α до β. Это событие мы будем называть «попаданием случайной величины X на участок от α до β».

Условимся для определенности левый конец α включать в участок (α, β), а правый — не включать. Тогда попадание случайной величины X на участок (α, β) равносильно выполнению неравенства:

$$\alpha \leqslant X < \beta.$$

Выразим вероятность этого события через функцию распределения величины X. Для этого рассмотрим три события:

событие A, состоящее в том, что $X < \beta$;
событие B, состоящее в том, что $X < \alpha$;
событие C, состоящее в том, что $\alpha \leqslant X < \beta$.

Учитывая, что $A = B + C$, по теореме сложения вероятностей имеем:

$$P(X < \beta) = P(X < \alpha) + P(\alpha \leqslant X < \beta),$$

или

$$F(\beta) = F(\alpha) + P(\alpha \leqslant X < \beta),$$

откуда

$$P(\alpha \leqslant X < \beta) = F(\beta) - F(\alpha), \qquad (5.3.1)$$

т. е. *вероятность попадания случайной величины на заданный участок равна приращению функции распределения на этом участке.*

Будем неограниченно уменьшать участок (α, β), полагая, что $\beta \to \alpha$. В пределе вместо вероятности попадания на участок получим вероятность того, что величина примет отдельно взятое значение α:

$$P(X = \alpha) = \lim_{\beta \to \alpha} P(\alpha \leqslant X < \beta) = \lim_{\beta \to \alpha} [F(\beta) - F(\alpha)]. \qquad (5.3.2)$$

Значение этого предела зависит от того, непрерывна ли функция $F(x)$ в точке $x = \alpha$ или же терпит разрыв. Если в точке α функция $F(x)$ имеет разрыв, то предел (5.3.2) равен значению скачка функции $F(x)$ в точке α. Если же функция $F(x)$ в точке α непрерывна, то этот предел равен нулю.

В дальнейшем изложении мы условимся называть «непрерывными» только те случайные величины, функция распределения которых везде непрерывна. Имея это в виду, можно сформулировать следующее положение:

Вероятность любого отдельного значения непрерывной случайной величины равна нулю.

Остановимся на этом положении несколько подробнее. В данном курсе мы уже встречались с событиями, вероятности которых были равны нулю: это были невозможные события. Теперь мы видим, что обладать нулевой вероятностью могут не только невозможные, но и возможные события. Действительно, событие $X = a$, состоящее в том, что непрерывная случайная величина X примет значение a, возможно; однако вероятность его равна нулю. Такие события — возможные, но с нулевой вероятностью — появляются только при рассмотрении опытов, не сводящихся к схеме случаев.

Понятие о событии «возможном, но обладающем нулевой вероятностью» кажется на первый взгляд парадоксальным. В действительности оно не более парадоксально, чем представление о теле, имеющем определенную массу, тогда как ни одна из точек внутри тела определенной конечной массой не обладает. Сколь угодно малый объем, выделенный из тела, обладает определенной конечной массой; эта масса приближается к нулю по мере уменьшения объема и в пределе равна нулю для точки. Аналогично при непрерывном распределении вероятностей вероятность попадания на сколь угодно малый участок может быть отлична от нуля, тогда как вероятность попадания в строго определенную точку в точности равна нулю.

Если производится опыт, в котором непрерывная случайная величина X должна принять одно из своих возможных значений, то до опыта вероятность каждого из таких значений равна нулю; однако, в исходе опыта случайная величина X непременно примет одно из своих возможных значений, т. е. заведомо произойдет одно из событий, вероятности которых были равны нулю.

Из того, что событие $X = a$ имеет вероятность, равную нулю, вовсе не следует, что это событие не будет появляться, т. е. что частота этого события равна нулю. Мы знаем, что частота события при большом числе опытов не равна, а только приближается к вероятности. Из того, что вероятность события $X = a$ равна нулю, следует только, что при неограниченном повторении опыта это событие будет появляться сколь угодно редко.

Если событие A в данном опыте возможно, но имеет вероятность, равную нулю, то противоположное ему событие $\bar{A}$ имеет вероятность, равную единице, но не достоверно. Для непрерывной случайной величины X при любом a событие $X \neq a$ имеет вероятность, равную единице, однако это событие не достоверно. Такое событие при неограниченном повторении опыта будет происходить п о ч т и всегда, но не всегда.

В п° 5.1 мы познакомились с «механической» интерпретацией прерывной случайной величины как распределения единичной массы, сосредоточенной в нескольких изолированных точках на оси абсцисс.

В случае непрерывной случайной величины механическая интерпретация сводится к распределению единичной массы не по отдельным точкам, а непрерывно по оси абсцисс, причем ни одна точка не обладает конечной массой.

5.4. Плотность распределения

Пусть имеется непрерывная случайная величина X с функцией распределения $F(x)$, которую мы предположим непрерывной и дифференцируемой. Вычислим вероятность попадания этой случайной величины на участок от x до $x+\Delta x$:

$$P(x < X < x+\Delta x) = F(x+\Delta x) - F(x),$$

т. е. приращение функции распределения на этом участке. Рассмотрим отношение этой вероятности к длине участка, т. е. среднюю вероятность, приходящуюся на единицу длины на этом участке, и будем приближать Δx к нулю. В пределе получим производную от функции распределения:

$$\lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = F'(x). \qquad (5.4.1)$$

Введем обозначение:

$$f(x) = F'(x). \qquad (5.4.2)$$

Функция $f(x)$ — производная функции распределения — характеризует как бы плотность, с которой распределяются значения случайной величины в данной точке. Эта функция называется *плотностью распределения* (иначе — «плотностью вероятности») непрерывной случайной величины X. Иногда функцию $f(x)$ называют также «дифференциальной функцией распределения» или «дифференциальным законом распределения» величины X.

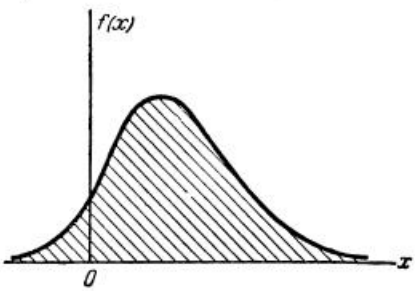


Рис. 5.4.1.

Термины «плотность распределения», «плотность вероятности» становятся особенно наглядными при пользовании механической интерпретацией распределения; в этой интерпретации функция $f(x)$ буквально характеризует плотность распределения масс по оси абсцисс (так называемую «линейную плотность»). Кривая, изображающая плотность распределения случайной величины, называется *кривой распределения* (рис. 5.4.1).

Плотность распределения, так же как и функция распределения, есть одна из форм закона распределения. В противоположность функции распределения эта форма не является универсальной: она существует только для непрерывных случайных величин.

Рассмотрим непрерывную случайную величину X с плотностью распределения $f(x)$ и элементарный участок dx, примыкающий к точке x (рис. 5.4.2). Вероятность попадания случайной величины X на этот элементарный участок (с точностью до бесконечно малых высшего порядка) равна $f(x)\,dx$. Величина $f(x)\,dx$ называется *элементом вероятности*. Геометрически это есть площадь элементарного прямоугольника, опирающегося на отрезок dx (рис. 5.4.2).

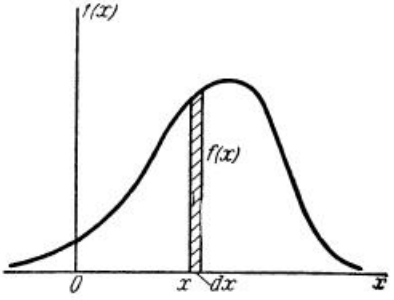


Рис. 5.4.2.

Выразим вероятность попадания величины X на отрезок от α до β (рис. 5.4.3) через плотность распределения. Очевидно, она равна сумме элементов вероятности на всем этом участке, т. е. интегралу:

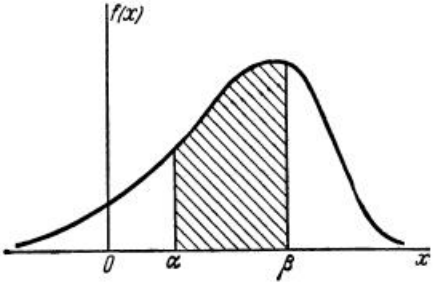


Рис. 5.4.3.

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)\,dx\ ^{1}). \tag{5.4.3}$$

Геометрически вероятность попадания величины X на участок (α, β) равна площади кривой распределения, опирающейся на этот участок (рис. 5.4.3).

Формула (5.4.2) выражает плотность распределения через функцию распределения. Зададимся обратной задачей: выразить функцию распределения через плотность. По определению

$$F(x) = P(X < x) = P(-\infty < X < x),$$

откуда по формуле (5.4.3) имеем:

$$F(x) = \int_{-\infty}^{x} f(x)\,dx. \tag{5.4.4}$$

[1]) Так как вероятность любого отдельного значения непрерывной случайной величины равна нулю, то можно рассматривать здесь отрезок (α, β), не включая в него левый конец, т. е. отбрасывая знак равенства в $\alpha \leqslant X < \beta$.

Геометрически $F(x)$ есть не что иное, как площадь кривой распределения, лежащая левее точки x (рис. 5.4.4).

Укажем основные свойства плотности распределения.

1. Плотность распределения есть неотрицательная функция:

$$f(x) \geqslant 0.$$

Это свойство непосредственно вытекает из того, что функция распределения $F(x)$ есть неубывающая функция.

2. Интеграл в бесконечных пределах от плотности распределения равен единице:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

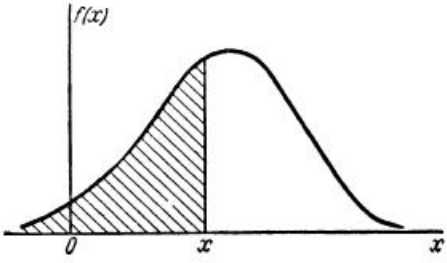


Рис. 5.4.4.

Это следует из формулы (5.4.4) и из того, что $F(+\infty) = 1$.

Геометрически основные свойства плотности распределения означают, что:

1) вся кривая распределения лежит не ниже оси абсцисс;

2) полная площадь, ограниченная кривой распределения и осью абсцисс, равна единице.

Выясним размерности основных характеристик случайной величины — функции распределения и плотности распределения. Функция распределения $F(x)$, как всякая вероятность, есть величина безразмерная. Размерность плотности распределения $f(x)$, как видно из формулы (5.4.1), обратна размерности случайной величины.

Пример 1. Функция распределения непрерывной случайной величины X задана выражением

$$F(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ ax^2 & \text{при } 0 < x \leqslant 1, \\ 1 & \text{при } x > 1. \end{cases}$$

а) Найти коэффициент a.

б) Найти плотность распределения $f(x)$.

в) Найти вероятность попадания величины X на участок от 0,25 до 0,5.

Решение. а) Так как функция распределения величины X непрерывна, то при $x = 1$ $ax^2 = 1$, откуда $a = 1$.

б) Плотность распределения величины X выражается формулой

$$f(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ 2x & \text{при } 0 < x \leqslant 1, \\ 0 & \text{при } x > 1. \end{cases}$$

в) По формуле (5.3.1) имеем:

$$P(0{,}25 < X < 0{,}5) = F(0{,}5) - F(0{,}25) = 0{,}5^2 - 0{,}25^2 = 0{,}1875.$$

Пример 2. Случайная величина X подчинена закону распределения с плотностью:

$$f(x) = a\cos x \quad \text{при} \quad -\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2};$$

$$f(x) = 0 \qquad \text{при} \quad x < -\frac{\pi}{2} \quad \text{или} \quad x > \frac{\pi}{2}.$$

а) Найти коэффициент a.
б) Построить график плотности распределения $f(x)$.
в) Найти функцию распределения $F(x)$ и построить ее график.
г) Найти вероятность попадания величины X на участок от 0 до $\frac{\pi}{4}$.

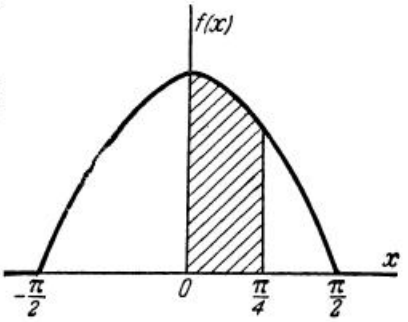


Рис. 5.4.5.

Решение. а) Для определения коэффициента a воспользуемся свойством плотности распределения:

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a\cos x\,dx = 2a = 1,$$

откуда $a = \frac{1}{2}$.

б) График плотности $f(x)$ представлен на рис. 5.4.5.

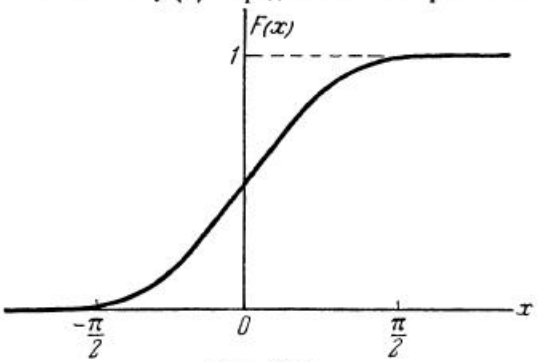


Рис. 5.4.6.

в) По формуле (5.4.4) получаем выражение функции распределения:

$$F(x) = \begin{cases} 0 & \text{при } x < -\frac{\pi}{2}; \\ \frac{1}{2}(\sin x + 1) & \text{при } -\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}; \\ 1 & \text{при } x > \frac{\pi}{2}. \end{cases}$$

График функции $F(x)$ изображен на рис. 5.4.6.

г) По формуле (5.3.1) имеем:

$$P\left(0 < X < \frac{\pi}{4}\right) = \frac{1}{2}\left(\sin\frac{\pi}{4} + 1\right) - \frac{1}{2}(\sin 0 + 1) = \frac{\sqrt{2}}{4}.$$

Тот же результат, но несколько более сложным путем можно получить по формуле (5.4.3).

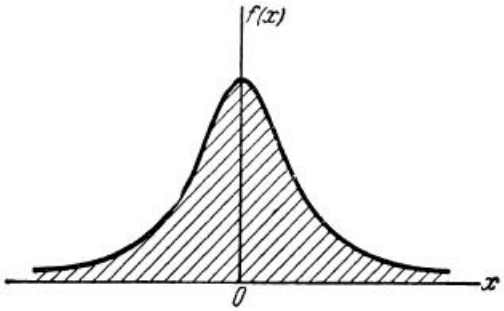


Рис. 5.4.7.

Пример 3. Плотность распределения случайной величины X задана формулой:

$$f(x) = \frac{1}{\pi(1 + x^2)}\ ^{1}).$$

а) Построить график плотности $f(x)$.

б) Найти вероятность того, что величина X попадет на участок $(-1, +1)$.

Решение. а) График плотности дан на рис. 5.4.7.

б) По формуле (5.4.3) имеем:

$$P(-1 < X < 1) = \int_{-1}^{1} \frac{dx}{\pi(1 + x^2)} = \frac{1}{\pi}\operatorname{arctg} x\Big|_{-1}^{1} = \frac{1}{2}.$$

5.5. Числовые характеристики случайных величин. Их роль и назначение

В данной главе мы познакомились с рядом полных, исчерпывающих характеристик случайных величин — так называемых законов распределения. Такими характеристиками были:

для дискретной случайной величины

а) функция распределения;

б) ряд распределения (графически — многоугольник распределения);

для непрерывной величины

а) функция распределения;

б) плотность распределения (графически — кривая распределения).

---

[1]) Так называемый закон Коши.

Каждый закон распределения представляет собой некоторую функцию, и указание этой функции полностью описывает случайную величину с вероятностной точки зрения.

Однако во многих вопросах практики нет необходимости характеризовать случайную величину полностью, исчерпывающим образом. Зачастую достаточно бывает указать только отдельные числовые параметры, до некоторой степени характеризующие существенные черты распределения случайной величины: например, какое-то среднее значение, около которого группируются возможные значения случайной величины; какое-либо число, характеризующее степень разбросанности этих значений относительно среднего, и т. д. Пользуясь такими характеристиками, мы хотим все существенные сведения относительно случайной величины, которыми мы располагаем, выразить наиболее компактно с помощью минимального числа числовых параметров. Такие характеристики, назначение которых — выразить в сжатой форме наиболее существенные особенности распределения, называются *числовыми характеристиками* случайной величины.

В теории вероятностей числовые характеристики и операции с ними играют огромную роль. С помощью числовых характеристик существенно облегчается решение многих вероятностных задач. Очень часто удается решить задачу до конца, оставляя в стороне законы распределения и оперируя одними числовыми характеристиками. При этом весьма важную роль играет то обстоятельство, что когда в задаче фигурирует большое количество случайных величин, каждая из которых оказывает известное влияние на численный результат опыта, то закон распределения этого результата в значительной мере можно считать независимым от законов распределения отдельных случайных величин (возникает так называемый нормальный закон распределения). В этих случаях по существу задачи для исчерпывающего суждения о результирующем законе распределения не требуется знать законов распределения отдельных случайных величин, фигурирующих в задаче; достаточно знать лишь некоторые числовые характеристики этих величин.

В теории вероятностей и математической статистике применяется большое количество различных числовых характеристик, имеющих различное назначение и различные области применения. Из них в настоящем курсе мы введем только некоторые, наиболее часто применяемые.

5.6. Характеристики положения (математическое ожидание, мода, медиана)

Среди числовых характеристик случайных величин нужно прежде всего отметить те, которые характеризуют положение случайной величины на числовой оси, т. е. указывают некоторое среднее, ориентировочное значение, около которого группируются все возможные значения случайной величины.

Среднее значение случайной величины есть некоторое число, являющееся как бы ее «представителем» и заменяющее ее при грубо ориентировочных расчетах. Когда мы говорим: «среднее время работы лампы равно 100 часам» или «средняя точка попадания смещена относительно цели на 2 *м* вправо», мы этим указываем определенную числовую характеристику случайной величины, описывающую ее местоположение на числовой оси, т. е. «характеристику положения».

Из характеристик положения в теории вероятностей важнейшую роль играет *математическое ожидание* случайной величины, которое иногда называют просто *средним значением* случайной величины.

Рассмотрим дискретную случайную величину X, имеющую возможные значения $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$. Нам требуется охарактеризовать каким-то числом положение значений случайной величины на оси абсцисс с учетом того, что эти значения имеют различные вероятности. Для этой цели естественно воспользоваться так называемым «средним взвешенным» из значений x_i, причем каждое значение x_i при осреднении должно учитываться с «весом», пропорциональным вероятности этого значения. Таким образом, мы вычислим среднее значение случайной величины X, которое мы обозначим $M[X]$:

$$M[X] = \frac{x_1 p_1 + x_2 p_2 + \ldots + x_n p_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i},$$

или, учитывая, что $\sum_{i=1}^{n} p_i = 1$,

$$M[X] = \sum_{i=1}^{n} x_i p_i. \qquad (5.6.1)$$

Это среднее взвешенное значение и называется математическим ожиданием случайной величины. Таким образом, мы ввели в рассмотрение одно из важнейших понятий теории вероятностей — понятие математического ожидания.

Математическим ожиданием случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений.

Заметим, что в вышеприведенной формулировке определение математического ожидания справедливо, строго говоря, только для дискретных случайных величин; ниже будет дано обобщение этого понятия на случай непрерывных величин.

Для того чтобы сделать понятие математического ожидания более наглядным, обратимся к механической интерпретации распределения

дискретной случайной величины. Пусть на оси абсцисс расположены точки с абсциссами $x_1, x_2, \ldots, x_n$, в которых сосредоточены соответственно массы $p_1, p_2, \ldots, p_n$, причем $\sum_{i=1}^{n} p_i = 1$. Тогда, очевидно, математическое ожидание $M[X]$, определяемое формулой (5.6.1), есть не что иное, как *абсцисса центра тяжести* данной системы материальных точек.

Математическое ожидание случайной величины X связано своеобразной зависимостью со средним арифметическим наблюденных значений случайной величины при большом числе опытов. Эта зависимость того же типа, как зависимость между частотой и вероятностью, а именно: при большом числе опытов среднее арифметическое наблюденных значений случайной величины приближается (сходится по вероятности) к ее математическому ожиданию. Из наличия связи между частотой и вероятностью можно вывести как следствие наличие подобной же связи между средним арифметическим и математическим ожиданием.

Действительно, рассмотрим дискретную случайную величину X, характеризуемую рядом распределения:

x_i	x_1	x_2	. . .	x_n
p_i	p_1	p_2	. . .	p_n

где $p_i = P(X = x_i)$.

Пусть производится N независимых опытов, в каждом из которых величина X принимает определенное значение. Предположим, что значение x_1 появилось m_1 раз, значение x_2 появилось m_2 раз, вообще значение x_i появилось m_i раз. Очевидно,

$$\sum_{i=1}^{n} m_i = N.$$

Вычислим среднее арифметическое наблюденных значений величины X, которое в отличие от математического ожидания $M[X]$ мы обозначим $M^*[X]$:

$$M^*[X] = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_n m_n}{N} =$$

$$= x_1 \frac{m_1}{N} + x_2 \frac{m_2}{N} + \ldots + x_n \frac{m_n}{N} = \sum_{i=1}^{n} x_i \frac{m_i}{N}.$$

Но $\frac{m_i}{N}$ есть не что иное, как частота (или статистическая вероятность) события $X = x_i$; эту частоту можно обозначить p_i^*. Тогда

$$M^*[X] = \sum_{i=1}^{n} x_i p_i^*,$$

т. е. среднее арифметическое наблюденных значений случайной величины равно сумме произведений всех возможных значений случайной величины на частоты этих значений.

При увеличении числа опытов N частоты p_i^* будут приближаться (сходиться по вероятности) к соответствующим вероятностям p_i. Следовательно, и среднее арифметическое наблюденных значений случайной величины $M^*[X]$ при увеличении числа опытов будет приближаться (сходиться по вероятности) к ее математическому ожиданию $M[X]$.

Сформулированная выше связь между средним арифметическим и математическим ожиданием составляет содержание одной из форм закона больших чисел. Строгое доказательство этого закона будет дано нами в главе 13.

Мы уже знаем, что все формы закона больших чисел констатируют факт устойчивости некоторых средних при большом числе опытов. Здесь речь идет об устойчивости среднего арифметического из ряда наблюдений одной и той же величины. При небольшом числе опытов среднее арифметическое их результатов случайно; при достаточном увеличении числа опытов оно становится «почти не случайным» и, стабилизируясь, приближается к постоянной величине — математическому ожиданию.

Свойство устойчивости средних при большом числе опытов легко проверить экспериментально. Например, взвешивая какое-либо тело в лаборатории на точных весах, мы в результате взвешивания получаем каждый раз новое значение; чтобы уменьшить ошибку наблюдения, мы взвешиваем тело несколько раз и пользуемся средним арифметическим полученных значений. Легко убедиться, что при дальнейшем увеличении числа опытов (взвешиваний) среднее арифметическое реагирует на это увеличение все меньше и меньше и при достаточно большом числе опытов практически перестает меняться.

Формула (5.6.1) для математического ожидания соответствует случаю дискретной случайной величины. Для непрерывной величины X математическое ожидание, естественно, выражается уже не суммой, а интегралом:

$$M[X] = \int_{-\infty}^{\infty} x f(x)\, dx, \qquad (5.6.2)$$

где $f(x)$ — плотность распределения величины X.

Формула (5.6.2) получается из формулы (5.6.1), если в ней заменить отдельные значения x_i непрерывно изменяющимся параметром x, соответствующие вероятности p_i — элементом вероятности $f(x)\,dx$, конечную сумму — интегралом. В дальнейшем мы часто будем пользоваться таким способом распространения формул, выведенных для прерывных величин, на случай непрерывных величин.

В механической интерпретации математическое ожидание непрерывной случайной величины сохраняет тот же смысл — абсциссы центра тяжести в случае, когда масса распределена по оси абсцисс непрерывно, с плотностью $f(x)$. Эта интерпретация часто позволяет найти математическое ожидание без вычисления интеграла (5.6.2), из простых механических соображений.

Выше мы ввели обозначение $M[X]$ для математического ожидания величины X. В ряде случаев, когда величина $M[X]$ входит в формулы как определенное число, ее удобнее обозначать одной буквой. В этих случаях мы будем обозначать математическое ожидание величины X через m_x:

$$m_x = M[X].$$

Обозначения m_x и $M[X]$ для математического ожидания будут в дальнейшем применяться параллельно в зависимости от удобства той или иной записи формул. Условимся также в случае надобности сокращать слова «математическое ожидание» буквами м. о.

Следует заметить, что важнейшая характеристика положения — математическое ожидание — существует не для всех случайных величин. Можно составить примеры таких случайных величин, для которых математического ожидания не существует, так как соответствующая сумма или интеграл расходятся.

Рассмотрим, например, прерывную случайную величину X с рядом распределения:

x_i	2	2^2	$\ldots$	2^i	$\ldots$
p_i	$\frac{1}{2}$	$\frac{1}{2^2}$	$\ldots$	$\frac{1}{2^i}$	$\ldots$

Нетрудно убедиться в том, что $\sum_{i=1}^{n} p_i = 1$, т. е. ряд распределения имеет смысл; однако сумма $\sum_{i=1}^{\infty} x_i p_i$ в данном случае расходится и, следовательно, математического ожидания величины X не существует. Однако для практики такие случаи существенного интереса не представляют. Обычно случайные величины, с которыми мы имеем дело, имеют ограниченную область возможных значений и безусловно обладают математическим ожиданием.

Выше мы дали формулы (5.6.1) и (5.6.2), выражающие математическое ожидание соответственно для прерывной и непрерывной случайной величины X.

Если величина X принадлежит к величинам смешанного типа, то ее математическое ожидание выражается формулой вида:

$$M[X] = \sum_i x_i p_i + \int x F'(x)\, dx, \quad (5.6.3)$$

где сумма распространяется на все точки x_i, в которых функция распределения терпит разрыв, а интеграл — на все участки, на которых функция распределения непрерывна.

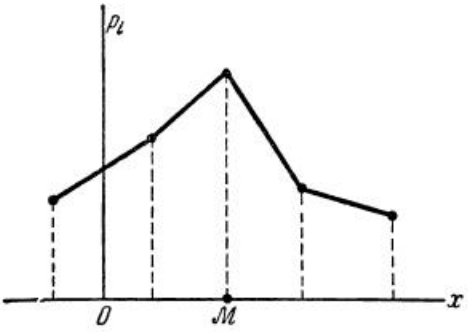


Рис. 5.6.1.

Кроме важнейшей из характеристик положения — математического ожидания, — на практике иногда применяются и другие характеристики положения, в частности *мода* и *медиана* случайной величины.

Модой случайной величины называется ее наиболее вероятное значение. Термин «наиболее вероятное значение», строго говоря, применим только к прерывным величинам; для непрерывной величины

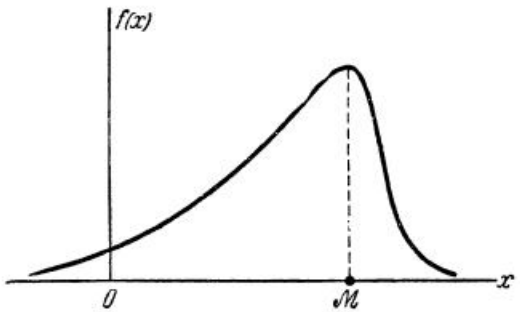


Рис. 5.6.2.

модой является то значение, в котором плотность вероятности максимальна. Условимся обозначать моду буквой M. На рис. 5.6.1 и 5.6.2 показана мода соответственно для прерывной и непрерывной случайных величин.

Если многоугольник распределения (кривая распределения) имеет более одного максимума, распределение называется «полимодальным» (рис. 5.6.3 и 5.6.4).

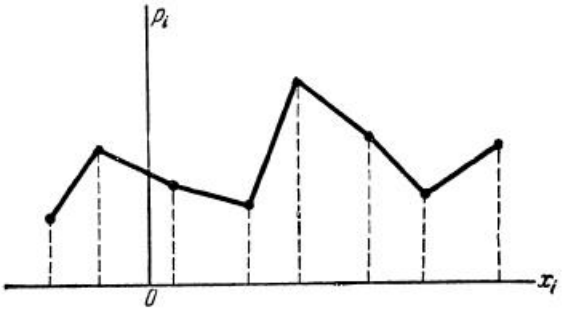


Рис. 5.6.3.

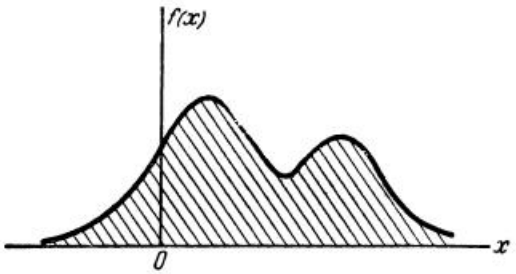


Рис. 5.6.4.

Иногда встречаются распределения, обладающие посередине не максимумом, а минимумом (рис. 5.6.5 и 5.6.6). Такие распределения

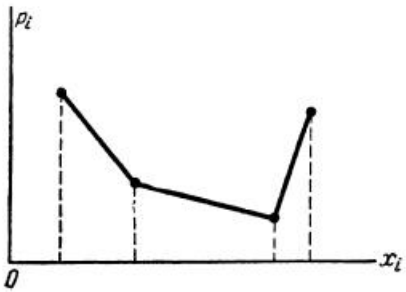


Рис. 5.6.5.

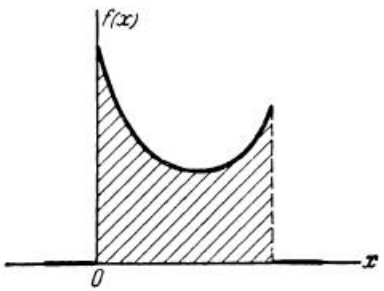


Рис. 5.6.6.

называются «антимодальными». Примером антимодального распределения может служить распределение, полученное в примере 5, п° 5.1.

В общем случае мода и математическое ожидание случайной величины не совпадают. В частном случае, когда распределение является симметричным и модальным (т. е. имеет моду) и существует математическое ожидание, то оно совпадает с модой и центром симметрии распределения.

Часто применяется еще одна характеристика положения — так называемая *медиана* случайной величины. Этой характеристикой пользуются обычно только для непрерывных случайных величин, хотя формально можно ее определить и для прерывной величины.

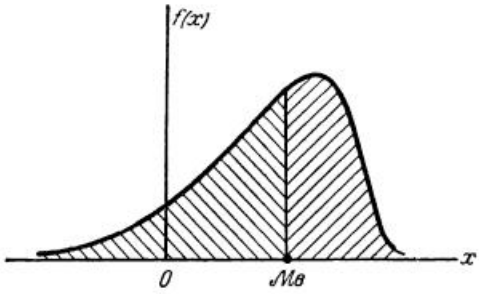


Рис. 5.6.7.

Медианой случайной величины X называется такое ее значение Me, для которого

$$P(X < Me) = P(X > Me),$$

т. е. одинаково вероятно, окажется ли случайная величина меньше или больше Me. Геометрическая медиана — это абсцисса точки, в которой площадь, ограниченная кривой распределения, делится пополам (рис. 5.6.7).

В случае симметричного модального распределения медиана совпадает с математическим ожиданием и модой.

5.7. Моменты. Дисперсия. Среднее квадратическое отклонение

Кроме характеристик положения — средних, типичных значений случайной величины, — употребляется еще ряд характеристик, каждая из которых описывает то или иное свойство распределения. В качестве таких характеристик чаще всего применяются так называемые *моменты*.

Понятие момента широко применяется в механике для описания распределения масс (статические моменты, моменты инерции и т. д.). Совершенно теми же приемами пользуются в теории вероятностей для описания основных свойств распределения случайной величины. Чаще всего применяются на практике моменты двух видов: начальные и центральные.

Начальным моментом s-го порядка прерывной случайной величины X называется сумма вида:

$$\alpha_s[X] = \sum_{i=1}^{n} x_i^s p_i. \qquad (5.7.1)$$

Очевидно, это определение совпадает с определением начального момента порядка s в механике, если на оси абсцисс в точках $x_1, x_2, \ldots, x_n$ сосредоточены массы $p_1, p_2, \ldots, p_n$.

Для непрерывной случайной величины X начальным моментом s-го порядка называется интеграл

$$\alpha_s[X] = \int_{-\infty}^{\infty} x^s f(x)\, dx. \tag{5.7.2}$$

Нетрудно убедиться, что введенная в предыдущем п° основная характеристика положения — математическое ожидание — представляет собой не что иное, как *первый начальный момент* случайной величины X.

Пользуясь знаком математического ожидания, можно объединить две формулы (5.7.1) и (5.7.2) в одну. Действительно, формулы (5.7.1) и (5.7.2) по структуре полностью аналогичны формулам (5.6.1) и (5.6.2), с той разницей, что в них вместо x_i и x стоят, соответственно, x_i^s и x^s. Поэтому можно написать общее определение начального момента s-го порядка, справедливое как для прерывных, так и для непрерывных величин:

$$\alpha_s[X] = M[X^s], \tag{5.7.3}$$

т. е. *начальным моментом s-го порядка случайной величины X называется математическое ожидание s-й степени этой случайной величины* [1]).

Перед тем как дать определение центрального момента, введем новое понятие «центрированной случайной величины».

Пусть имеется случайная величина X с математическим ожиданием m_x. *Центрированной случайной величиной*, соответствующей величине X, называется отклонение случайной величины X от ее математического ожидания:

$$\mathring{X} = X - m_x. \tag{5.7.4}$$

Условимся в дальнейшем везде обозначать центрированную случайную величину, соответствующую данной случайной величине, той же буквой со значком ° наверху.

Нетрудно убедиться, что математическое ожидание центрированной случайной величины равно нулю. Действительно, для прерывной

---

[1]) Понятие математического ожидания функции от случайной величины будет уточнено далее (см. главу 10).

величины

$$M[\mathring{X}] = M[X - m_x] = \sum_{i=1}^{n} (x_i - m_x) p_i =$$

$$= \sum_{i=1}^{n} x_i p_i - m_x \sum_{i=1}^{n} p_i = m_x - m_x = 0; \quad (5.7.5)$$

аналогично и для непрерывной величины.

Центрирование случайной величины, очевидно, равносильно переносу начала координат в среднюю, «центральную» точку, абсцисса которой равна математическому ожиданию.

Моменты центрированной случайной величины носят название *центральных моментов*. Они аналогичны моментам относительно центра тяжести в механике.

Таким образом, *центральным моментом порядка s* случайной величины X называется математическое ожидание s-й степени соответствующей центрированной случайной величины:

$$\mu_s[X] = M[\mathring{X}^s] = M[(X - m_x)^s]. \quad (5.7.6)$$

Для прерывной случайной величины s-й центральный момент выражается суммой

$$\mu_s = \sum_{i=1}^{n} (x_i - m_x)^s p_i, \quad (5.7.7)$$

а для непрерывной — интегралом

$$\mu_s = \int_{-\infty}^{\infty} (x - m_x)^s f(x)\, dx. \quad (5.7.8)$$

В дальнейшем в тех случаях, когда не возникает сомнений, к какой случайной величине относится данный момент, мы будем для краткости вместо $\alpha_s[X]$ и $\mu_s[X]$ писать просто α_s и μ_s.

Очевидно, для любой случайной величины *центральный момент первого порядка равен нулю*:

$$\mu_1 = M[\mathring{X}] = M[X - m_x] = 0, \quad (5.7.9)$$

так как математическое ожидание центрированной случайной величины всегда равно нулю.

Выведем соотношения, связывающие центральные и начальные моменты различных порядков. Вывод мы проведем только для прерывных величин; легко убедиться, что точно те же соотношения справедливы и для непрерывных величин, если заменить конечные суммы интегралами, а вероятности — элементами вероятности.

Рассмотрим второй центральный момент:

$$\mu_2 = M[\mathring{X}^2] = \sum_{i=1}^{n} (x_i - m_x)^2 p_i =$$

$$= \sum_{i=1}^{n} x_i^2 p_i - 2m_x \sum_{i=1}^{n} x_i p_i + m_x^2 \sum_{i=1}^{n} p_i = \alpha_2 - 2m_x^2 + m_x^2 = \alpha_2 - m_x^2.$$

Аналогично для третьего центрального момента получим:

$$\mu_3 = M[\mathring{X}^3] = \sum_{i=1}^{n} (x_i - m_x)^3 p_i = \sum_{i=1}^{n} x_i^3 p_i - 3m_x \sum_{i=1}^{n} x_i^2 p_i +$$

$$+ 3m_x^2 \sum_{i=1}^{n} x_i p_i - m_x^3 \sum_{i=1}^{n} p_i = \alpha_3 - 3\alpha_2 m_x + 2m_x^3.$$

Выражения для μ_4, μ_5 и т. д. могут быть получены аналогичным путем.

Таким образом, для центральных моментов любой случайной величины X справедливы формулы:

$$\left.\begin{aligned} &\mu_1 = 0, \\ &\mu_2 = \alpha_2 - m_x^2, \\ &\mu_3 = \alpha_3 - 3m_x\alpha_2 + 2m_x^3, \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \end{aligned}\right\} \qquad (5.7.10)$$

Вообще говоря, моменты могут рассматриваться не только относительно начала координат (начальные моменты) или математического ожидания (центральные моменты), но и относительно произвольной точки a:

$$\gamma_s = M[(X - a)^s]. \qquad (5.7.11)$$

Однако центральные моменты имеют перед всеми другими преимущество: первый центральный момент, как мы видели, всегда равен нулю, а следующий за ним, второй центральный момент при этой системе отсчета имеет минимальное значение. Докажем это. Для прерывной случайной величины X при $s = 2$ формула (5.7.11) имеет вид:

$$\gamma_2 = \sum_{i=1}^{n} (x_i - a)^2 p_i. \qquad (5.7.12)$$

Преобразуем это выражение:

$$\gamma_2 = \sum_{i=1}^{n} (x_i - m_x + m_x - a)^2 p_i =$$

$$= \sum_{i=1}^{n} (x_i - m_x)^2 p_i - 2(m_x - a) \sum_{i=1}^{n} (x_i - m_x) p_i +$$

$$+ (m_x - a)^2 = \mu_2 + (m_x - a)^2.$$

Очевидно, эта величина достигает своего минимума, когда $m_x = a$, т. е. когда момент берется относительно точки m_x.

Из всех моментов в качестве характеристик случайной величины чаще всего применяются первый начальный момент (математическое ожидание) $m_x = \alpha_1$ и второй центральный момент μ_2.

Второй центральный момент называется *дисперсией* случайной величины. Ввиду крайней важности этой характеристики среди других моментов введем для нее специальное обозначение $D[X]$:

$$\mu_2 = D[X].$$

Согласно определению центрального момента

$$D[X] = M[\mathring{X}^2], \tag{5.7.13}$$

т. е. *дисперсией случайной величины X называется математическое ожидание квадрата соответствующей центрированной величины.*

Заменяя в выражении (5.7.13) величину $\mathring{X}$ ее выражением, имеем также:

$$D[X] = M[(X - m_x)^2]. \tag{5.7.14}$$

Для непосредственного вычисления дисперсии служат формулы:

$$D[X] = \sum_{i=1}^{n} (x_i - m_x)^2 p_i, \tag{5.7.15}$$

$$D[X] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x)\, dx \tag{5.7.16}$$

— соответственно для прерывных и непрерывных величин.

Дисперсия случайной величины есть характеристика *рассеивания*, разбросанности значений случайной величины около ее математического ожидания. Само слово «дисперсия» означает «рассеивание».

Если обратиться к механической интерпретации распределения, то дисперсия представляет собой не что иное, как *момент инерции* заданного распределения масс относительно центра тяжести (математического ожидания).

Дисперсия случайной величины имеет размерность квадрата случайной величины; для наглядной характеристики рассеивания удобнее пользоваться величиной, размерность которой совпадает с размерностью случайной величины. Для этого из дисперсии извлекают квадратный корень. Полученная величина называется *средним квадратическим отклонением* (иначе — «стандартом») случайной величины X. Среднее квадратическое отклонение будем обозначать $\sigma[X]$:

$$\sigma[X] = \sqrt{D[X]}. \tag{5.7.17}$$

Для упрощения записей мы часто будем пользоваться сокращенными обозначениями среднего квадратического отклонения и дисперсии: σ_x и D_x. В случае, когда не возникает сомнения, к какой случайной величине относятся эти характеристики, мы будем иногда опускать значок x у σ_x и D_x и писать просто σ и D. Слова «среднее квадратическое отклонение» иногда будем сокращенно заменять буквами с. к. о.

На практике часто применяется формула, выражающая дисперсию случайной величины через ее второй начальный момент (вторая из формул (5.7.10)). В новых обозначениях она будет иметь вид:

$$D_x = \alpha_2 - m_x^2. \tag{5.7.18}$$

Математическое ожидание m_x и дисперсия D_x (или среднее квадратическое отклонение σ_x) — наиболее часто применяемые характеристики случайной величины. Они характеризуют наиболее важные черты распределения: его положение и степень разбросанности. Для более подробного описания распределения применяются моменты высших порядков.

Третий центральный момент служит для характеристики *асимметрии* (или «скошенности») распределения. Если распределение симметрично относительно математического ожидания (или, в механической интерпретации, масса распределена симметрично относительно центра тяжести), то все моменты нечетного порядка (если они существуют) равны нулю. Действительно, в сумме

$$\mu_s = \sum_{i=1}^{n} (x_i - m_x)^s p_i$$

при симметричном относительно m_x законе распределения и нечетном s каждому положительному слагаемому соответствует равное ему по абсолютной величине отрицательное слагаемое, так что вся сумма равна нулю. То же, очевидно, справедливо и для интеграла

$$\mu_s = \int_{-\infty}^{\infty} (x - m_x)^s f(x)\, dx,$$

который равен нулю, как интеграл в симметричных пределах от нечетной функции.

Естественно поэтому в качестве характеристики асимметрии распределения выбрать какой-либо из нечетных моментов. Простейший из них есть третий центральный момент. Он имеет размерность куба случайной величины; чтобы получить безразмерную характеристику, третий момент μ_3 делят на куб среднего квадратического отклонения.

Полученная величина носит название «коэффициента асимметрии» или просто «асимметрии»; мы обозначим ее Sk:

$$Sk = \frac{\mu_3}{\sigma^3}. \qquad (5.7.19)$$

На рис. 5.7.1 показано два асимметричных распределения; одно из них (кривая I) имеет положительную асимметрию ($Sk > 0$); другое (кривая II) — отрицательную ($Sk < 0$).

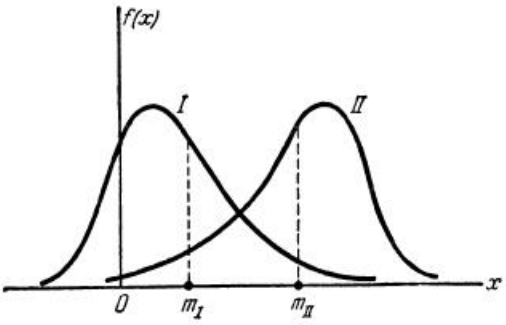


Рис. 5.7.1.

Четвертый центральный момент служит для характеристики так называемой «крутости», т. е. островершинности или плосковершинности распределения. Эти свойства распределения описываются с помощью так называемого *эксцесса*. Эксцессом случайной величины X называется величина

$$Ex = \frac{\mu_4}{\sigma^4} - 3. \qquad (5.7.20)$$

Число 3 вычитается из отношения $\frac{\mu_4}{\sigma^4}$ потому, что для весьма важного и широко распространенного в природе нормального закона распределения (с которым мы подробно познакомимся в дальнейшем) $\frac{\mu_4}{\sigma^4} = 3$. Таким образом, для нормального распределения эксцесс равен нулю; кривые, более островершинные по сравнению с нормальной, обладают положительным эксцессом; кривые более плосковершинные — отрицательным эксцессом.

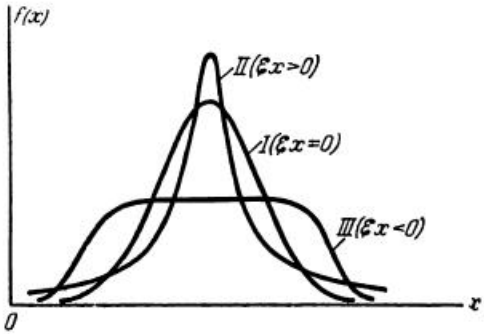


Рис. 5.7.2.

На рис. 5.7.2 представлены: нормальное распределение (кривая I), распределение с положительным эксцессом (кривая II) и распределение с отрицательным эксцессом (кривая III).

Кроме рассмотренных выше начальных и центральных моментов, на практике иногда применяются так называемые *абсолютные* моменты (начальные и центральные), определяемые формулами

$$\beta_s = M[|X|^s]; \quad \nu_s = M[|\mathring{X}|^s].$$

Очевидно, абсолютные моменты четных порядков совпадают с обычными моментами.

Из абсолютных моментов наиболее часто применяется первый абсолютный центральный момент

$$\nu_1 = M[|\mathring{X}|] = M[|X - m_x|], \tag{5.7.21}$$

называемый *средним арифметическим отклонением*. Наряду с дисперсией и средним квадратическим отклонением среднее арифметическое отклонение иногда применяется как характеристика рассеивания.

Математическое ожидание, мода, медиана, начальные и центральные моменты и, в частности, дисперсия, среднее квадратическое отклонение, асимметрия и эксцесс представляют собой наиболее употребительные числовые характеристики случайных величин. Во многих задачах практики полная характеристика случайной величины — закон распределения — или не нужна, или не может быть получена. В этих случаях ограничиваются приблизительным описанием случайной величины с помощью числовых характеристик, каждая из которых выражает какое-либо характерное свойство распределения.

Очень часто числовыми характеристиками пользуются для приближенной замены одного распределения другим, причем обычно стремятся произвести эту замену так, чтобы сохранились неизменными несколько важнейших моментов.

Пример 1. Производится один опыт, в результате которого может появиться или не появиться событие A, вероятность которого равна p. Рассматривается случайная величина X — число появлений события A (характеристическая случайная величина события A). Определить ее характеристики: математическое ожидание, дисперсию, среднее квадратическое отклонение.

Решение. Ряд распределения величины имеет вид:

x_i	0	1
p_i	q	p

где $q = 1 - p$ — вероятность непоявления события A.

По формуле (5.6.1) находим математическое ожидание величины X:

$$m_x = M[X] = 0 \cdot q + 1 \cdot p = p.$$

Дисперсию величины X определим по формуле (5.7.15):

$$D_x = D[X] = (0-p)^2 \cdot q + (1-p)^2 p = pq,$$

откуда

$$\sigma_x = \sqrt{pq}.$$

(Предлагаем читателю получить тот же результат, выразив дисперсию через второй начальный момент.)

Пример 2. Производится три независимых выстрела по мишени; вероятность попадания при каждом выстреле равна 0,4. Случайная величина X — число попаданий. Определить характеристики величины X — математическое ожидание, дисперсию, с. к. о., асимметрию.

Решение. Ряд распределения величины X имеет вид:

x_i	0	1	2	3
p_i	0,216	0,432	0,288	0,064

Вычисляем числовые характеристики величины X:

$$m_x = 0 \cdot 0{,}216 + 1 \cdot 0{,}432 + 2 \cdot 0{,}288 + 3 \cdot 0{,}064 = 1{,}2;$$

$$D_x = (0-1{,}2)^2 \cdot 0{,}216 + (1-1{,}2)^2 \cdot 0{,}432 + (2-1{,}2)^2 \cdot 0{,}288 + (3-1{,}2)^2 \cdot 0{,}064 = 0{,}72;$$

$$\sigma_x = \sqrt{D_x} = \sqrt{0{,}72} = 0{,}848;$$

$$\mu_3 = (0-1{,}2)^3 \cdot 0{,}216 + (1-1{,}2)^3 \cdot 0{,}432 + (2-1{,}2)^3 \cdot 0{,}288 + (3-1{,}2)^3 \cdot 0{,}064 = 0{,}144;$$

$$Sk = \frac{\mu_3}{\sigma_x^3} = \frac{0{,}144}{0{,}72 \cdot 0{,}848} \approx 0{,}236.$$

Заметим, что те же характеристики могли бы быть вычислены значительно проще с помощью теорем о числовых характеристиках функций (см. главу 10).

Пример 3. Производится ряд независимых опытов до первого появления события A (см. пример 3 п° 5.1). Вероятность события A в каждом опыте равна p. Найти математическое ожидание, дисперсию и с. к. о. числа опытов, которое будет произведено.

Решение. Ряд распределения величины X имеет вид:

x_i	1	2	3	. . .	i	. . .
p_i	p	qp	q^2p	. . .	$q^{i-1}p$	. . .

Математическое ожидание величины X выражается суммой ряда

$$m_x = 1 \cdot p + 2 \cdot qp + 3 \cdot q^2 p + \ldots + i \cdot q^{i-1} p + \ldots = p(1 + 2q + 3q^2 + \ldots + iq^{i-1} + \ldots).$$

Нетрудно видеть, что ряд, стоящий в скобках, представляет собой результат дифференцирования геометрической прогрессии:

$$q+q^2+q^3+\ \dots\ +q^i+\ \dots\ =\frac{q}{1-q}.$$

Следовательно,

$$1+2q+3q^2+\ \dots\ +iq^{i-1}+\ \dots\ =\frac{d}{dq}\frac{1}{1-q}=\frac{1}{(1-q)^2}=\frac{1}{p^2},$$

откуда

$$m_x=\frac{p}{p^2}=\frac{1}{p}.$$

Для определения дисперсии величины X вычислим сначала ее второй начальный момент:

$$\alpha_2=\sum_{i=1}^{\infty}x_i^2p_i=1^2\cdot p+2^2\cdot qp+3^2\cdot q^2p+\ \dots\ +i^2\cdot q^{i-1}p+\ \dots\ =$$

$$=p\,(1^2+2^2q+3^2q^2+\ \dots\ +i^2q^{i-1}+\ \dots).$$

Для вычисления ряда, стоящего в скобках, умножим на q ряд:

$$1+2q+3q^2+\ \dots\ +iq^{i-1}+\ \dots\ =\frac{1}{(1-q)^2}.$$

Получим:

$$q+2q^2+3q^3+\ \dots\ +iq^i+\ \dots\ =\frac{q}{(1-q)^2}.$$

Дифференцируя этот ряд по q, имеем:

$$1^2+2^2q+3^2q^2+\ \dots\ +i^2q^{i-1}+\ \dots\ =\frac{d}{dq}\frac{q}{(1-q)^2}=\frac{1+q}{(1-q)^3}.$$

Умножая на $p=1-q$, получим:

$$\alpha_2=\frac{1+q}{(1-q)^2}.$$

По формуле (5.7.18) выразим дисперсию:

$$D_x=\alpha_2-m_x^2=\frac{1+q}{(1-q)^2}-\frac{1}{(1-q)^2}=\frac{q}{(1-q)^2}=\frac{q}{p^2},$$

откуда

$$\sigma_x=\sqrt{D_x}=\frac{\sqrt{q}}{p}.$$

Пример 4. Непрерывная случайная величина X подчинена закону распределения с плотностью:

$$f(x)=Ae^{-|x|}$$

(рис. 5.7.3).

Найти коэффициент A. Определить м. о., дисперсию, с. к. о., асимметрию, эксцесс величины X.

Решение. Для определения A воспользуемся свойством плотности распределения:

$$\int_{-\infty}^{\infty}f(x)\,dx=2A\int_0^{\infty}e^{-x}\,dx=2A=1;$$

отсюда

$$A = \frac{1}{2}.$$

Так как функция $xe^{-|x|}$ нечетная, то м. о. величины X равно нулю:

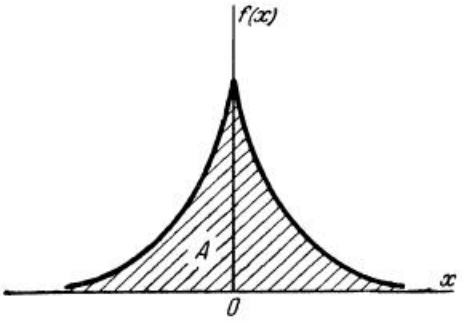


Рис. 5.7.3.

$$m_x = \int_{-\infty}^{\infty} \frac{1}{2} xe^{-|x|}\, dx = 0.$$

Дисперсия и с. к. о. равны, соответственно:

$$D_x = 2\int_0^{\infty} \frac{1}{2} x^2 e^{-x}\, dx = 2;$$

$$\sigma_x = \sqrt{D_x} = \sqrt{2}.$$

Так как распределение симметрично, то $Sk = 0$.

Для вычисления эксцесса находим μ_4:

$$\mu_4 = 2\int_0^{\infty} \frac{1}{2} x^4 e^{-x}\, dx = 24,$$

откуда

$$Ex = \frac{\mu_4}{\sigma_x^4} - 3 = 3.$$

Пример 5. Случайная величина X подчинена закону распределения, плотность которого задана графически на рис. 5.7.4.

Написать выражение плотности распределения. Найти м. о., дисперсию, с. к. о. и асимметрию распределения.

Решение. Выражение плотности распределения имеет вид:

$$f(x) = \begin{cases} ax & \text{при } 0 < x < 1, \\ 0 & \text{при } x < 0 \text{ или } x > 1. \end{cases}$$

Пользуясь свойством плотности распределения, находим $a = 2$.

Математическое ожидание величины X:

$$m_x = \int_0^1 2x^2\, dx = \frac{2}{3}.$$

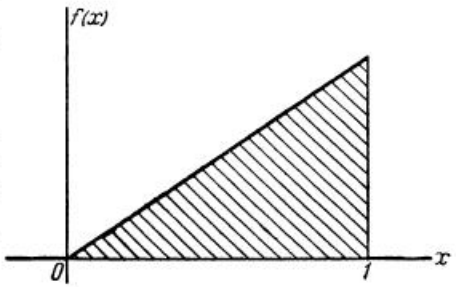


Рис. 5.7.4.

Дисперсию найдем через второй начальный момент:

$$\alpha_2 = 2\int_0^1 x^3\, dx = \frac{1}{2}; \quad D_x = \alpha_2 - m_x^2 = \frac{1}{18};$$

откуда

$$\sigma_x = \frac{1}{3\sqrt{2}}.$$

Третий начальный момент равен

$$\alpha_3 = 2\int_0^1 x^4\,dx = \frac{2}{5}.$$

Пользуясь третьей из формул (5.7.10), выражающей μ_3 через начальные моменты, имеем:

$$\mu_3 = \alpha_3 - 3m_x\alpha_2 + 2m_x^3 = -\frac{1}{135},$$

откуда

$$Sk = \frac{\mu_3}{\sigma_x^3} = -\frac{2}{5}\sqrt{2}.$$

5.8. Закон равномерной плотности

В некоторых задачах практики встречаются непрерывные случайные величины, о которых заранее известно, что их возможные значения лежат в пределах некоторого определенного интервала; кроме того, известно, что в пределах этого интервала все значения случайной величины одинаково вероятны (точнее, обладают одной и той же плотностью вероятности). О таких случайных величинах говорят, что они распределяются *по закону равномерной плотности.*

Приведем несколько примеров подобных случайных величин.

Пример 1. Произведено взвешивание тела на точных весах, но в распоряжении взвешивающего имеются только разновески весом не менее 1 *г*; результат взвешивания показывает, что вес тела заключен между k и $(k+1)$ граммами. Вес тела принят равным $\left(k+\frac{1}{2}\right)$ граммам. Допущенная при этом ошибка X, очевидно, есть случайная величина, распределенная с равномерной плотностью на участке $\left(-\frac{1}{2}, \frac{1}{2}\right)$ *г*.

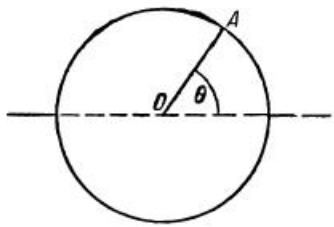


Рис. 5.8.1.

Пример 2. Вертикально поставленное симметричное колесо (рис. 5.8.1) приводится во вращение и затем останавливается вследствие трения. Рассматривается случайная величина θ — угол, который после остановки будет составлять с горизонтом фиксированный радиус колеса OA. Очевидно, величина θ распределена с равномерной плотностью на участке $(0, 2\pi)$.

Пример 3. Поезда метрополитена идут с интервалом 2 *мин*. Пассажир выходит на платформу в некоторый момент времени. Время T, в течение которого ему придется ждать поезда, представляет собой случайную величину, распределенную с равномерной плотностью на участке (0, 2) минут.

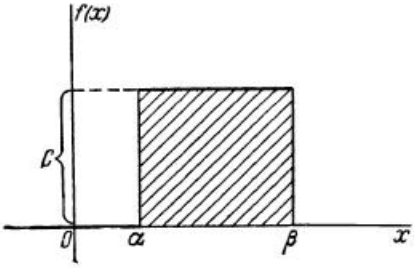


Рис. 5.8.2.

Рассмотрим случайную величину X, подчиненную закону равномерной плотности на участке от α до β (рис. 5.8.2), и напишем для нее выражение плотности распределения $f(x)$. Плотность $f(x)$ постоянна и равна c на отрезке (α, β); вне этого отрезка она равна нулю:

$$f(x) = \begin{cases} c \text{ при } \alpha < x < \beta, \\ 0 \text{ при } x < \alpha \text{ или } x > \beta. \end{cases}$$

Так как площадь, ограниченная кривой распределения, равна единице:

$$c\,(\beta - \alpha) = 1,$$

то

$$c = \frac{1}{\beta - \alpha},$$

и плотность распределения $f(x)$ имеет вид:

$$\left.\begin{aligned} f(x) &= \frac{1}{\beta - \alpha} \quad \text{при} \quad \alpha < x < \beta, \\ f(x) &= 0 \quad \text{при} \quad x < \alpha \quad \text{или} \quad x > \beta. \end{aligned}\right\} \tag{5.8.1}$$

Формула (5.8.1) и выражает закон равномерной плотности на участке (α, β).

Напишем выражение для функции распределения $F(x)$. Функция распределения выражается площадью кривой распределения, лежащей левее точки x. Следовательно,

$$F(x) = \begin{cases} 0 & \text{при} \quad x < \alpha, \\ \dfrac{x - \alpha}{\beta - \alpha} & \text{при} \quad \alpha < x < \beta, \\ 1 & \text{при} \quad x > \beta. \end{cases}$$

График функции $F(x)$ приведен на рис. 5.8.3.

Определим основные числовые характеристики случайной величины X, подчиненной закону равномерной плотности на участке от α до β.

Математическое ожидание величины X равно:

$$m_x = \int_\alpha^\beta \frac{x}{\beta-\alpha}\,dx = \frac{\alpha+\beta}{2}. \qquad (5.8.2)$$

В силу симметричности равномерного распределения медиана величины X также равна $\frac{\alpha+\beta}{2}$.

Моды закон равномерной плотности не имеет.

По формуле (5.7.16) находим дисперсию величины X:

$$D_x = \mu_2 = \frac{1}{\beta-\alpha}\int_\alpha^\beta \left(x-\frac{\alpha+\beta}{2}\right)^2 dx = \frac{(\beta-\alpha)^2}{12}, \qquad (5.8.3)$$

откуда среднее квадратическое отклонение

$$\sigma_x = \sqrt{D_x} = \frac{\beta-\alpha}{2\sqrt{3}}. \qquad (5.8.4)$$

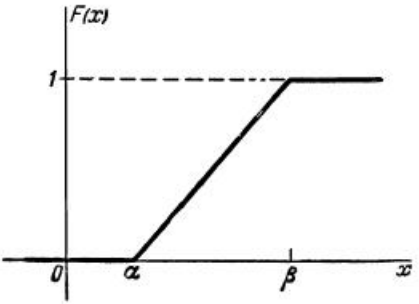


Рис. 5.8.3.

В силу симметричности распределения его асимметрия равна нулю:

$$Sk = \frac{\mu_3}{\sigma_x^3} = 0. \qquad (5.8.5)$$

Для определения эксцесса находим четвертый центральный момент:

$$\mu_4 = \frac{1}{\beta-\alpha}\int_\alpha^\beta \left(x-\frac{\alpha+\beta}{2}\right)^4 dx = \frac{(\beta-\alpha)^4}{80},$$

откуда

$$Ex = \frac{\mu_4}{\sigma_x^4} - 3 = -1{,}2. \qquad (5.8.6)$$

Определяем среднее арифметическое отклонение:

$$v_1 \quad \frac{1}{\beta-\alpha}\int_\alpha^\beta \left|x-\frac{\alpha+\beta}{2}\right| dx = \frac{2}{\beta-\alpha}\int_{\frac{\alpha+\beta}{2}}^\beta \left(x-\frac{\alpha+\beta}{2}\right) dx = \frac{\beta-\alpha}{4}. \qquad (5.8.7)$$

Наконец, найдем вероятность попадания случайной величины X, распределенной по закону равномерной плотности, на участок (a, b),

представляющий собой часть участка (α, β) (рис. 5.8.4). Геометрически эта вероятность представляет собой площадь, заштрихованную на рис. 5.8.4. Очевидно, она равна:

$$P(a < X < b) = \frac{b-a}{\beta-\alpha}, \qquad (5.8.8)$$

т. е. отношению длины отрезка (a, b) ко всей длине участка (α, β), на котором задано равномерное распределение.

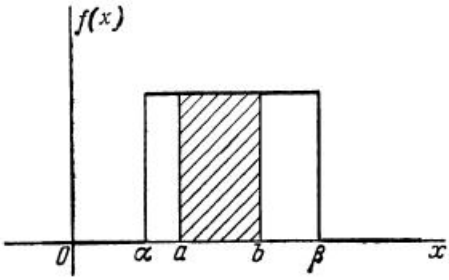


Рис. 5.8.4.

5.9. Закон Пуассона

Во многих задачах практики приходится иметь дело со случайными величинами, распределенными по своеобразному закону, который называется *законом Пуассона*.

Рассмотрим прерывную случайную величину X, которая может принимать только целые, неотрицательные значения:

$$0, 1, 2, \ldots, m, \ldots,$$

причем последовательность этих значений теоретически не ограничена.

Говорят, что случайная величина X распределена по закону Пуассона, если вероятность того, что она примет определенное значение m, выражается формулой

$$P_m = \frac{a^m}{m!} e^{-a} \qquad (m = 0, 1, \ldots), \qquad (5.9.1)$$

где a — некоторая положительная величина, называемая *параметром* закона Пуассона.

Ряд распределения случайной величины X, распределенной по закону Пуассона, имеет вид:

x_m	0	1	2	. . .	m	. . .
P_m	e^{-a}	$\frac{a}{1!} e^{-a}$	$\frac{a}{2!} e^{-a}$	. . .	$\frac{a^m}{m!} e^{-a}$	. . .

Убедимся прежде всего, что последовательность вероятностей, задаваемая формулой (5.9.1), может представлять собой ряд распре-

деления, т. е. что сумма всех вероятностей P_m равна единице. Имеем:

$$\sum_{m=0}^{\infty} P_m = \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!}.$$

Но

$$\sum_{m=0}^{\infty} \frac{a^m}{m!} = e^a,$$

откуда

$$\sum_{m=0}^{\infty} P_m = e^{-a} e^a = 1.$$

На рис. 5.9.1 показаны многоугольники распределения случайной величины X, распределенной по закону Пуассона, соответствующие различным значениям параметра a. В таблице 8 приложения приведены значения P_m для различных a.

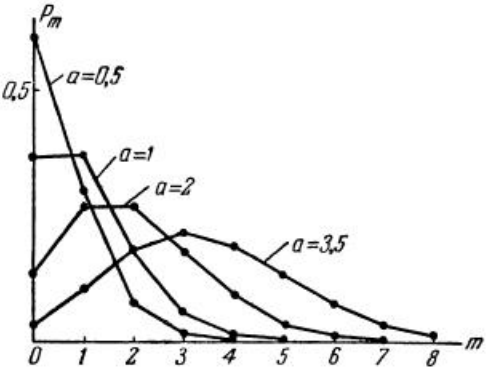


Рис. 5.9.1.

Определим основные характеристики — математическое ожидание и дисперсию — случайной величины X, распределенной по закону Пуассона. По определению математического ожидания

$$m_x = M[X] = \sum_{m=0}^{\infty} m P_m = \sum_{m=0}^{\infty} m \frac{a^m}{m!} e^{-a}.$$

Первый член суммы (соответствующий $m = 0$) равен нулю, следовательно, суммирование можно начинать с $m = 1$:

$$m_x = \sum_{m=1}^{\infty} m \frac{a^m}{m!} e^{-a} = a e^{-a} \sum_{m=1}^{\infty} \frac{m a^{m-1}}{m!} = a e^{-a} \sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!}.$$

Обозначим $m - 1 = k$; тогда

$$m_x = a e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} = a e^{-a} e^a = a. \qquad (5.9.2)$$

Таким образом, параметр a представляет собой не что иное, как математическое ожидание случайной величины X.

Для определения дисперсии найдем сначала второй начальный момент величины X:

$$\alpha_2 = \sum_{m=0}^{\infty} m^2 \frac{a^m}{m!} e^{-a} = a \sum_{m=1}^{\infty} m \frac{a^{m-1}}{(m-1)!} e^{-a} =$$

$$= a \sum_{m=1}^{\infty} [(m-1)+1] \frac{a^{m-1}}{(m-1)!} e^{-a} =$$

$$= a \left[\sum_{m=1}^{\infty} (m-1) \frac{a^{m-1}}{(m-1)!} e^{-a} + \sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} e^{-a} \right].$$

По ранее доказанному

$$\sum_{m=1}^{\infty} (m-1) \frac{a^{m-1}}{(m-1)!} e^{-a} = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} = a;$$

кроме того,

$$\sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} e^{-a} = e^{-a} e^{a} = 1,$$

следовательно,

$$\alpha_2 = a(a+1).$$

Далее находим дисперсию величины X:

$$D_x = \alpha_2 - m_x^2 = a^2 + a - a^2 = a. \qquad (5.9.3)$$

Таким образом, *дисперсия случайной величины, распределенной по закону Пуассона, равна ее математическому ожиданию a.*

Это свойство распределения Пуассона часто применяется на практике для решения вопроса, правдоподобна ли гипотеза о том, что случайная величина X распределена по закону Пуассона. Для этого определяют из опыта статистические характеристики — математическое ожидание и дисперсию — случайной величины [1]). Если их значения близки, то это может служить доводом в пользу гипотезы о пуассоновском распределении; резкое различие этих характеристик, напротив, свидетельствует против гипотезы.

Определим для случайной величины X, распределенной по закону Пуассона, вероятность того, что она примет значение не меньше

---

[1]) О способах экспериментального определения этих характеристик см. ниже, гл. 7 и 14.

заданного k. Обозначим эту вероятность R_k:

$$R_k = P(X \geqslant k).$$

Очевидно, вероятность R_k может быть вычислена как сумма

$$R_k = P_k + P_{k+1} + \ldots = \sum_{m=k}^{\infty} P_m.$$

Однако значительно проще определить ее из вероятности противоположного события:

$$R_k = 1 - (P_0 + P_1 + \ldots + P_{k-1}) = 1 - \sum_{m=0}^{k-1} P_m. \quad (5.9.4)$$

В частности, вероятность того, что величина X примет положительное значение, выражается формулой

$$R_1 = 1 - P_0 = 1 - e^{a}. \quad (5.9.5)$$

Мы уже упоминали о том, что многие задачи практики приводят к распределению Пуассона. Рассмотрим одну из типичных задач такого рода.

Пусть на оси абсцисс Ox случайным образом распределяются точки (рис. 5.9.2). Допустим, что случайное распределение точек удовлетворяет следующим условиям:

1. Вероятность попадания того или иного числа точек на отрезок l зависит только от длины этого отрезка, но не зависит от его положения на оси абсцисс. Иными словами, точки распределены на оси

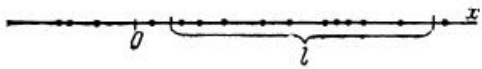


Рис. 5.9.2.

абсцисс с одинаковой средней плотностью. Обозначим эту плотность (т. е. математическое ожидание числа точек, приходящихся на единицу длины) через λ.

2. Точки распределяются на оси абсцисс независимо друг от друга, т. е. вероятность попадания того или другого числа точек на заданный отрезок не зависит от того, сколько их попало на любой другой отрезок, не перекрывающийся с ним.

3. Вероятность попадания на малый участок Δx двух или более точек пренебрежимо мала по сравнению с вероятностью попадания одной точки (это условие означает практическую невозможность совпадения двух или более точек).

Выделим на оси абсцисс определенный отрезок длины l и рассмотрим дискретную случайную величину X — число точек, попадающих на этот отрезок. Возможные значения величины будут

$$0,\ 1,\ 2,\ \ldots,\ m,\ \ldots \tag{5.9.6}$$

Так как точки попадают на отрезок независимо друг от друга, то теоретически не исключено, что их там окажется сколь угодно много, т. е. ряд (5.9.6) продолжается неограниченно.

Докажем, что случайная величина X имеет закон распределения Пуассона. Для этого вычислим вероятность P_m того, что на отрезок l попадет ровно m точек.

Сначала решим более простую задачу. Рассмотрим на оси Ox малый участок Δx и вычислим вероятность того, что на этот участок попадет хотя бы одна точка. Будем рассуждать следующим образом. Математическое ожидание числа точек, попадающих на этот участок, очевидно, равно $\lambda \Delta x$ (т. к. на единицу длины попадает в среднем λ точек). Согласно условию 3 для малого отрезка Δx можно пренебречь возможностью попадания на него двух или больше точек. Поэтому математическое ожидание $\lambda \Delta x$ числа точек, попадающих на участок Δx, будет приближенно равно вероятности попадания на него одной точки (или, что в наших условиях равнозначно, хотя бы одной).

Таким образом, с точностью до бесконечно малых высшего порядка, при $\Delta x \to 0$ можно считать вероятность того, что на участок Δx попадет одна (хотя бы одна) точка, равной $\lambda \Delta x$, а вероятность того, что не попадет ни одной, равной $1 - \lambda \Delta x$.

Воспользуемся этим для вычисления вероятности P_m попадания на отрезок l ровно m точек. Разделим отрезок l на n равных частей длиной $\Delta x = \frac{l}{n}$. Условимся называть элементарный отрезок Δx «пустым», если в него не попало ни одной точки, и «занятым», если в него попала хотя бы одна. Согласно вышедоказанному вероятность того, что отрезок Δx окажется «занятым», приближенно равна $\lambda \Delta x = \frac{\lambda l}{n}$; вероятность того, что он окажется «пустым», равна $1 - \frac{\lambda l}{n}$. Так как, согласно условию 2, попадания точек в неперекрывающиеся отрезки независимы, то наши n отрезков можно рассмотреть как n независимых «опытов», в каждом из которых отрезок может быть «занят» с вероятностью $p = \frac{\lambda l}{n}$. Найдем вероятность того, что среди n отрезков будет ровно m «занятых». По теореме о повторении опытов эта вероятность равна

$$C_n^m \left(\frac{\lambda l}{n}\right)^m \left(1 - \frac{\lambda l}{n}\right)^{n-m},$$

или, обозначая $\lambda l = a$,

$$C_n^m \left(\frac{a}{n}\right)^m \left(1 - \frac{a}{n}\right)^{n-m}. \tag{5.9.7}$$

При достаточно большом n эта вероятность приближенно равна вероятности попадания на отрезок l ровно m точек, так как попадание двух или больше точек на отрезок Δx имеет пренебрежимо малую вероятность. Для того чтобы найти точное значение P_m, нужно в выражении (5.9.7) перейти к пределу при $n \to \infty$:

$$P_m = \lim_{n \to \infty} C_n^m \left(\frac{a}{n}\right)^m \left(1 - \frac{a}{n}\right)^{n-m}. \tag{5.9.8}$$

Преобразуем выражение, стоящее под знаком предела:

$$C_n^m \left(\frac{a}{n}\right)^m \left(1 - \frac{a}{n}\right)^{n-m} = \frac{n(n-1)\ldots(n-m+1)}{m!} \frac{a^m}{n^m} \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^m} =$$

$$= \frac{n(n-1)\ldots(n-m+1)}{n^m} \frac{a^m}{m!} \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^m}. \tag{5.9.9}$$

Первая дробь и знаменатель последней дроби в выражении (5.9.9) при $n \to \infty$, очевидно, стремятся к единице. Выражение $\frac{a^m}{m!}$ от n не зависит. Числитель последней дроби можно преобразовать так:

$$\left(1 - \frac{a}{n}\right)^n = \left[\left(1 - \frac{a}{n}\right)^{\frac{n}{a}}\right]^a. \tag{5.9.10}$$

При $n \to \infty$ $\frac{n}{a} \to \infty$, и выражение (5.9.10) стремится к e^{-a}. Таким образом, доказано, что вероятность попадания ровно m точек в отрезок l выражается формулой

$$P_m = \frac{a^m}{m!} e^{-a},$$

где $a = \lambda l$, т. е. величина X распределена по закону Пуассона с параметром $a = \lambda l$.

Отметим, что величина a по смыслу представляет собой среднее число точек, приходящееся на отрезок l.

Величина R_1 (вероятность того, что величина X примет положительное значение) в данном случае выражает *вероятность того, что на отрезок l попадет хотя бы одна точка*:

$$R_1 = 1 - e^{-a}. \tag{5.9.11}$$

Таким образом, мы убедились, что распределение Пуассона возникает там, где какие-то точки (или другие элементы) занимают случайное положение независимо друг от друга, и подсчитывается количество этих точек, попавших в какую-то область. В нашем случае такой «областью» был отрезок l на оси абсцисс. Однако наш вывод легко распространить и на случай распределения точек на плоскости (случайное плоское поле точек) и в пространстве (случайное пространственное поле точек). Нетрудно доказать, что если соблюдены условия:

1) точки распределены в поле статистически равномерно со средней плотностью λ;

2) точки попадают в неперекрывающиеся области независимым образом;

3) точки появляются поодиночке, а не парами, тройками и т. д., то число точек X, попадающих в любую область D (плоскую или пространственную), распределяется по закону Пуассона:

$$P_m = \frac{a^m}{m!} e^{-a} \quad (m = 0,\ 1,\ 2,\ \ldots),$$

где a — среднее число точек, попадающих в область D.

Для плоского случая

$$a = S_D \lambda,$$

где S_D — площадь области D;

для пространственного

$$a = V_D \cdot \lambda,$$

где V_D — объем области D.

Заметим, что для наличия пуассоновского распределения числа точек, попадающих в отрезок или область, условие постоянной плотности ($\lambda = \text{const}$) несущественно. Если выполнены два других условия, то закон Пуассона все равно имеет место, только параметр a в нем приобретает другое выражение: он получается не простым умножением плотности λ на длину, площадь или объем области, а интегрированием переменной плотности по отрезку, площади или объему. (Подробнее об этом см. п° 19.4.)

Наличие случайных точек, разбросанных на линии, на плоскости или объеме — не единственное условие, при котором возникает распределение Пуассона. Можно, например, доказать, что закон Пуассона является предельным для биномиального распределения:

$$P_{m,\,n} = C_n^m p^m (1-p)^{n-m}, \tag{5.9.12}$$

если одновременно устремлять число опытов n к бесконечности, а вероятность p — к нулю, причем их произведение np сохраняет постоянное значение:

$$np = a. \tag{5.9.13}$$

Действительно, это предельное свойство биномиального распределения можно записать в виде:

$$\lim_{n\to\infty} C_n^m p^m (1-p)^{n-m} = \frac{a^m}{m!} e^{-a}. \tag{5.9.14}$$

Но из условия (5.9.13) следует, что

$$p = \frac{a}{n}. \tag{5.9.15}$$

Подставляя (5.9.15) в (5.9.14), получим равенство

$$\lim_{n\to\infty} C_n^m \left(\frac{a}{n}\right)^m \left(1-\frac{a}{n}\right)^{n-m} = \frac{a^m}{m!} e^{-a}, \tag{5.9.16}$$

которое только что было доказано нами по другому поводу.

Это предельное свойство биномиального закона часто находит применение на практике. Допустим, что производится большое количество независимых опытов n, в каждом из которых событие A имеет очень малую вероятность p. Тогда для вычисления вероятности $P_{m,n}$ того, что событие A появится ровно m раз, можно воспользоваться приближенной формулой

$$P_{m,n} \approx \frac{(np)^m}{m!} e^{-np}, \tag{5.9.17}$$

где $np = a$ — параметр того закона Пуассона, которым приближенно заменяется биномиальное распределение.

От этого свойства закона Пуассона — выражать биномиальное распределение при большом числе опытов и малой вероятности события — происходит его название, часто применяемое в учебниках статистики: *закон редких явлений*.

Рассмотрим несколько примеров, связанных с пуассоновским распределением, из различных областей практики.

Пример 1. На автоматическую телефонную станцию поступают вызовы со средней плотностью K вызовов в час. Считая, что число вызовов на любом участке времени распределено по закону Пуассона, найти вероятность того, что за две минуты на станцию поступит ровно три вызова.

Решение. Среднее число вызовов за две минуты равно:

$$a = \frac{2K}{60} = \frac{K}{30}.$$

По формуле (5.9.1) вероятность поступления ровно трех вызовов равна:

$$P_3 = \frac{\left(\frac{K}{30}\right)^3}{1\cdot 2\cdot 3} e^{-\frac{K}{30}}.$$

Пример 2. В условиях предыдущего примера найти вероятность того, что за две минуты придет хотя бы один вызов.

Решение. По формуле (5.9.4) имеем:

$$R_1 = 1 - e^{-a} = 1 - e^{-\frac{K}{30}}.$$

Пример 3. В тех же условиях найти вероятность того, что за две минуты придет не менее трех вызовов.

Решение. По формуле (5.9.4) имеем:

$$R_3 = 1 - (P_0 + P_1 + P_2) = 1 - e^{-\frac{K}{30}}\left[1 + \frac{K}{30} + \frac{1}{2}\left(\frac{K}{30}\right)^2\right].$$

Пример 4. На ткацком станке нить обрывается в среднем 0,375 раза в течение часа работы станка. Найти вероятность того, что за смену (8 часов) число обрывов нити будет заключено в границах 2 и 4 (не менее 2 и не более 4 обрывов).

Решение. Очевидно,

$$a = 0{,}375 \cdot 8 = 3;$$

имеем:

$$P(2 \leqslant X \leqslant 4) = P_2 + P_3 + P_4.$$

По таблице 8 приложения при $a = 3$

$$P_2 = 0{,}224; \quad P_3 = 0{,}224; \quad P_4 = 0{,}168,$$

$$P(2 \leqslant X \leqslant 4) = 0{,}616.$$

Пример 5. С накаленного катода за единицу времени вылетает в среднем $q(t)$ электронов, где t — время, протекшее с начала опыта. Найти вероятность того, что за промежуток времени длительности τ, начинающийся в момент t_0, с катода вылетит ровно m электронов.

Решение. Находим среднее число электронов a, вылетающих с катода за данный отрезок времени. Имеем:

$$a = \int_{t_0}^{t_0+\tau} q(t)\, dt.$$

По вычисленному a определяем искомую вероятность:

$$P_m = \frac{a^m}{m!} e^{-a}.$$

Пример 6. Число осколков, попадающих в малоразмерную цель при заданном положении точки разрыва, распределяется по закону Пуассона. Средняя плотность осколочного поля, в котором оказывается цель при данном положении точки разрыва, равна 3 оск./м². Площадь цели равна $S = 0{,}5$ м². Для поражения цели достаточно попадания в нее хотя бы одного осколка. Найти вероятность поражения цели при данном положении точки разрыва.

Решение. $a = \lambda S = 1{,}5$. По формуле (5.9.4) находим вероятность попадания хотя бы одного осколка:

$$R_1 = 1 - e^{-1{,}5} = 1 - 0{,}223 = 0{,}777.$$

(Для вычисления значения показательной функции e^{-a} пользуемся таблицей 2 приложения.)

Пример 7. Средняя плотность болезнетворных микробов в одном кубическом метре воздуха равна 100. Берется на пробу $2\,дм^3$ воздуха. Найти вероятность того, что в нем будет обнаружен хотя бы один микроб.

Решение. Принимая гипотезу о пуассоновском распределении числа микробов в объеме, находим:

$$a = 0{,}1; \quad R_1 = 1 - e^{-0{,}2} \approx 1 - 0{,}819 \approx 0{,}18.$$

Пример 8. По некоторой цели производится 50 независимых выстрелов Вероятность попадания в цель при одном выстреле равна 0,04. Пользуясь предельным свойством биномиального распределения (формула (5.9.17)), найти приближенно вероятность того, что в цель попадет: ни одного снаряда, один снаряд, два снаряда.

Решение. Имеем $a = np = 50 \cdot 0{,}04 = 2$. По таблице 8 приложения находим вероятности:

$$P_0 = 0{,}135; \quad P_1 = 0{,}271; \quad P_2 = 0{,}271.$$

ГЛАВА 6

НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

6.1. Нормальный закон и его параметры

Нормальный закон распределения (часто называемый законом Гаусса) играет исключительно важную роль в теории вероятностей и занимает среди других законов распределения особое положение. Это — наиболее часто встречающийся на практике закон распределения. Главная особенность, выделяющая нормальный закон среди других законов, состоит в том, что он является *предельным законом*, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.

Можно доказать, что сумма достаточно большого числа независимых (или слабо зависимых) случайных величин, подчиненных каким угодно законам распределения (при соблюдении некоторых весьма нежестких ограничений), приближенно подчиняется нормальному закону, и это выполняется тем точнее, чем большее количество случайных величин суммируется. Большинство встречающихся на практике случайных величин, таких, например, как ошибки измерений, ошибки стрельбы и т. д., могут быть представлены как суммы весьма большого числа сравнительно малых слагаемых — элементарных ошибок, каждая из которых вызвана действием отдельной причины, не зависящей от остальных. Каким бы законам распределения ни были подчинены отдельные элементарные ошибки, особенности этих распределений в сумме большого числа слагаемых нивелируются, и сумма оказывается подчиненной закону, близкому к нормальному. Основное ограничение, налагаемое на суммируемые ошибки, состоит в том, чтобы они все равномерно играли в общей сумме относительно малую роль. Если это условие не выполняется и, например, одна из случайных ошибок окажется по своему влиянию на сумму резко превалирующей над всеми другими, то закон распределения этой превалирующей ошибки наложит свое влияние на сумму и определит в основных чертах ее закон распределения.

Теоремы, устанавливающие нормальный закон как предельный для суммы независимых равномерно малых случайных слагаемых, будут подробнее рассмотрены в главе 13.

Нормальный закон распределения характеризуется плотностью вероятности вида:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}. \qquad (6.1.1)$$

Кривая распределения по нормальному закону имеет симметричный холмообразный вид (рис. 6.1.1). Максимальная ордината кривой, равная $\frac{1}{\sigma\sqrt{2\pi}}$, соответствует точке $x = m$; по мере удаления от точки m плотность распределения падает, и при $x \to \pm\infty$ кривая асимптотически приближается к оси абсцисс.

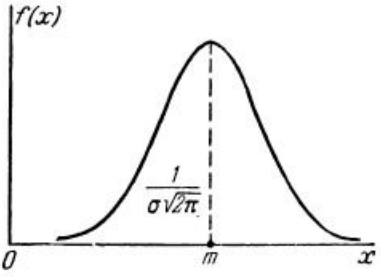


Рис. 6.1.1.

Выясним смысл численных параметров m и σ, входящих в выражение нормального закона (6.1.1); докажем, что величина m есть не что иное, как *математическое ожидание*, а величина σ — *среднее квадратическое отклонение* величины X. Для этого вычислим основные числовые характеристики величины X — математическое ожидание и дисперсию.

$$M[X] = \int_{-\infty}^{\infty} x f(x)\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Применяя замену переменной

$$\frac{x-m}{\sigma\sqrt{2}} = t,$$

имеем:

$$M[X] = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sigma\sqrt{2}t + m) e^{-t^2} dt =$$

$$= \frac{\sigma\sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t e^{-t^2} dt + \frac{m}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt. \qquad (6.1.2)$$

Нетрудно убедиться, что первый из двух интервалов в формуле (6.1.2) равен нулю; второй представляет собой известный интеграл Эйлера — Пуассона:

$$\int_{-\infty}^{\infty} e^{-t^2} dt = 2\int_{0}^{\infty} e^{-t^2} dt = \sqrt{\pi}. \qquad (6.1.3)$$

Следовательно,

$$M[X]=m,$$

т. е. параметр m представляет собой математическое ожидание величины X. Этот параметр, особенно в задачах стрельбы, часто называют *центром рассеивания* (сокращенно — ц. р.).

Вычислим дисперсию величины X:

$$D[X]=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}}dx.$$

Применив снова замену переменной

$$\frac{x-m}{\sigma\sqrt{2}}=t,$$

имеем:

$$D[X]=\frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty}t^2e^{-t^2}dt.$$

Интегрируя по частям, получим:

$$D[X]=\frac{\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty}t\cdot 2te^{-t^2}dt=$$

$$=\frac{\sigma^2}{\sqrt{\pi}}\left\{-te^{-t^2}\Big|_{-\infty}^{\infty}+\int_{-\infty}^{\infty}e^{-t^2}dt\right\}.$$

Первое слагаемое в фигурных скобках равно нулю (так как e^{-t^2} при $t\to\infty$ убывает быстрее, чем возрастает любая степень t), второе слагаемое по формуле (6.1.3) равно $\sqrt{\pi}$, откуда

$$D[X]=\sigma^2.$$

Следовательно, параметр σ в формуле (6.1.1) есть не что иное, как среднее квадратическое отклонение величины X.

Выясним смысл параметров m и σ нормального распределения. Непосредственно из формулы (6.1.1) видно, что центром симметрии распределения является центр рассеивания m. Это ясно из того, что при изменении знака разности $(x-m)$ на обратный выражение (6.1.1) не меняется. Если изменять центр рассеивания m, кривая распределения будет смещаться вдоль оси абсцисс, не изменяя своей формы (рис. 6.1.2). Центр рассеивания характеризует положение распределения на оси абсцисс.

Размерность центра рассеивания — та же, что размерность случайной величины X.

Параметр σ характеризует не положение, а самую форму кривой распределения. Это есть характеристика рассеивания. Наибольшая ордината кривой распределения обратно пропорциональна σ; при увеличении σ максимальная ордината уменьшается. Так как площадь

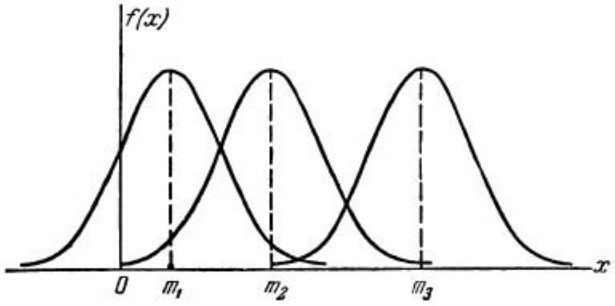


Рис. 6.1.2.

кривой распределения всегда должна оставаться равной единице, то при увеличении σ кривая распределения становится более плоской, растягиваясь вдоль оси абсцисс; напротив, при уменьшении σ кривая распределения вытягивается вверх, одновременно сжимаясь с боков, и становится более иглообразной. На рис. 6.1.3 показаны три нормальные кривые (*I*, *II*, *III*) при $m = 0$; из них кривая *I* соответствует

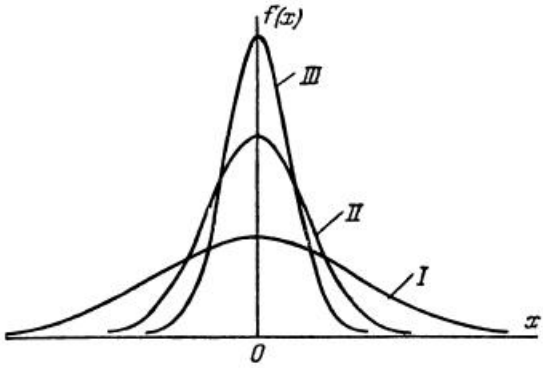


Рис. 6.1.3.

самому большому, а кривая *III* — самому малому значению σ. Изменение параметра σ равносильно изменению масштаба кривой распределения — увеличению масштаба по одной оси и такому же уменьшению по другой.

Размерность параметра σ, естественно, совпадает с размерностью случайной величины X.

В некоторых курсах теории вероятностей в качестве характеристики рассеивания для нормального закона вместо среднего квадратического отклонения применяется так называемая *мера точности*. Мерой точности называется величина, обратно пропорциональная среднему квадратическому отклонению σ:

$$h = \frac{1}{\sigma\sqrt{2}}.$$

Размерность меры точности обратна размерности случайной величины.

Термин «мера точности» заимствован из теории ошибок измерений: чем точнее измерение, тем больше мера точности. Пользуясь мерой точности h, можно записать нормальный закон в виде:

$$f(x) = \frac{h}{\sqrt{\pi}} e^{-h^2(x-m)^2}.$$

6.2. Моменты нормального распределения

Выше мы доказали, что математическое ожидание случайной величины, подчиненной нормальному закону (6.1.1), равно m, а среднее квадратическое отклонение равно σ.

Выведем общие формулы для центральных моментов любого порядка.

По определению:

$$\mu_s = \int_{-\infty}^{\infty} (x-m)^s f(x)\,dx =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-m)^s e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Делая замену переменной

$$\frac{x-m}{\sigma\sqrt{2}} = t,$$

получим:

$$\mu_s = \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^s e^{-t^2} dt. \qquad (6.2.1)$$

Применим к выражению (6.2.1) формулу интегрирования по частям:

$$\mu_s = \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-1} t e^{-t^2} dt =$$

$$= \frac{(\sigma\sqrt{2})^s}{\sqrt{\pi}} \left\{ -\frac{1}{2} e^{-t^2} t^{s-1} \Big|_{-\infty}^{\infty} + \frac{s-1}{2} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2} dt \right\}.$$

Имея в виду, что первый член внутри скобок равен нулю, получим:

$$\mu_s = \frac{(s-1)(\sigma\sqrt{2})^s}{2\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2}\, dt. \qquad (6.2.2)$$

Из формулы (6.2.1) имеем следующее выражение для μ_{s-2}:

$$\mu_{s-2} = \frac{(\sigma\sqrt{2})^{s-2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{s-2} e^{-t^2}\, dt. \qquad (6.2.3)$$

Сравнивая правые части формул (6.2.2) и (6.2.3), видим, что они отличаются между собой только множителем $(s-1)\sigma^2$; следовательно,

$$\mu_s = (s-1)\sigma^2 \mu_{s-2}. \qquad (6.2.4)$$

Формула (6.2.4) представляет собой простое рекуррентное соотношение, позволяющее выражать моменты высших порядков через моменты низших порядков. Пользуясь этой формулой и имея в виду, что $\mu_0 = 1$ [1]) и $\mu_1 = 0$, можно вычислить центральные моменты всех порядков. Так как $\mu_1 = 0$, то из формулы (6.2.4) следует, что все нечетные моменты нормального распределения равны нулю. Это, впрочем, непосредственно следует из симметричности нормального закона.

Для четных s из формулы (6.2.4) вытекают следующие выражения для последовательных моментов:

$$\mu_2 = \sigma^2;$$
$$\mu_4 = 3\sigma^4;$$
$$\mu_6 = 15\sigma^6$$

и т. д.

Общая формула для момента s-го порядка при любом четном s имеет вид:

$$\mu_s = (s-1)!!\,\sigma^s,$$

где под символом $(s-1)!!$ понимается произведение всех нечетных чисел от 1 до $s-1$.

Так как для нормального закона $\mu_3 = 0$, то асимметрия его также равна нулю:

$$Sk = \frac{\mu_3}{\sigma^3} = 0.$$

Из выражения четвертого момента

$$\mu_4 = 3\sigma^4$$

имеем:

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = 0,$$

[1]) Нулевой момент любой случайной величины равен единице как математическое ожидание нулевой степени этой величины.

т. е. эксцесс нормального распределения равен нулю. Это и естественно, так как назначение эксцесса — характеризовать сравнительную крутость данного закона по сравнению с нормальным.

6.3. Вероятность попадания случайной величины, подчиненной нормальному закону, на заданный участок. Нормальная функция распределения

Во многих задачах, связанных с нормально распределенными случайными величинами, приходится определять вероятность попадания случайной величины X, подчиненной нормальному закону с параметрами m, σ, на участок от α до β. Для вычисления этой вероятности воспользуемся общей формулой

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha), \tag{6.3.1}$$

где $F(x)$ — функция распределения величины X.

Найдем функцию распределения $F(x)$ случайной величины X, распределенной по нормальному закону с параметрами m, σ. Плотность распределения величины X равна:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}. \tag{6.3.2}$$

Отсюда находим функцию распределения

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-m)^2}{2\sigma^2}} dx. \tag{6.3.3}$$

Сделаем в интеграле (6.3.3) замену переменной

$$\frac{x-m}{\sigma} = t$$

и приведем его к виду:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{t^2}{2}} dt. \tag{6.3.4}$$

Интеграл (6.3.4) не выражается через элементарные функции, но его можно вычислить через специальную функцию, выражающую определенный интеграл от выражения e^{-t^2} или $e^{-\frac{t^2}{2}}$ (так называемый *интеграл вероятностей*), для которого составлены таблицы. Существует много разновидностей таких функций, например:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt; \qquad \Phi_1(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

и т. д. Какой из этих функций пользоваться — вопрос вкуса. Мы выберем в качестве такой функции

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \qquad (6.3.5)$$

Нетрудно видеть, что эта функция представляет собой не что иное, как функцию распределения для нормально распределенной случайной величины с параметрами $m = 0$, $\sigma = 1$.

Условимся называть функцию $\Phi^*(x)$ *нормальной функцией распределения*. В приложении (табл. 1) приведены таблицы значений функции $\Phi^*(x)$ [1]).

Выразим функцию распределения (6.3.3) величины X с параметрами m и σ через нормальную функцию распределения $\Phi^*(x)$. Очевидно,

$$F(x) = \Phi^*\left(\frac{x-m}{\sigma}\right). \qquad (6.3.6)$$

Теперь найдем вероятность попадания случайной величины X на участок от α до β. Согласно формуле (6.3.1)

$$P(\alpha < x < \beta) = \Phi^*\left(\frac{\beta-m}{\sigma}\right) - \Phi^*\left(\frac{\alpha-m}{\sigma}\right). \qquad (6.3.7)$$

Таким образом, мы выразили вероятность попадания на участок случайной величины X, распределенной по нормальному закону с любыми параметрами, через стандартную функцию распределения $\Phi^*(x)$, соответствующую простейшему нормальному закону с параметрами 0,1. Заметим, что аргументы функции Φ^* в формуле (6.3.7) имеют очень простой смысл: $\frac{\beta-m}{\sigma}$ есть расстояние от правого конца участка β до центра рассеивания, выраженное в средних квадратических отклонениях; $\frac{\alpha-m}{\sigma}$ — такое же расстояние для левого конца участка, причем это расстояние считается положительным, если конец расположен справа от центра рассеивания, и отрицательным, если слева.

Как и всякая функция распределения, функция $\Phi^*(x)$ обладает свойствами:

1. $\Phi^*(-\infty) = 0$. 2. $\Phi^*(+\infty) = 1$. 3. $\Phi^*(x)$ — неубывающая функция.

Кроме того, из симметричности нормального распределения с параметрами $m = 0$, $\sigma = 1$ относительно начала координат следует, что

$$\Phi^*(-x) = 1 - \Phi^*(x). \qquad (6.3.8)$$

---

[1]) Для облегчения интерполяции в таблицах рядом со значениями функции приведены ее приращения за один шаг таблиц Δ.

Пользуясь этим свойством, собственно говоря, можно было бы ограничить таблицы функции $\Phi^*(x)$ только положительными значениями аргумента, но, чтобы избежать лишней операции (вычитание из единицы), в таблице 1 приложения приводятся значения $\Phi^*(x)$ как для положительных, так и для отрицательных аргументов.

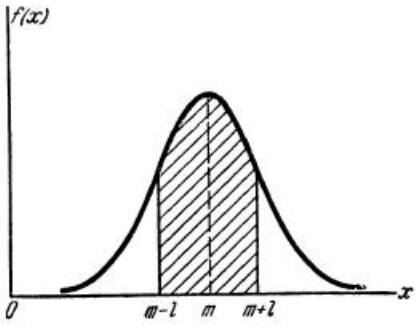


Рис. 6.3.1.

На практике часто встречается задача вычисления вероятности попадания нормально распределенной случайной величины на участок, симметричный относительно центра рассеивания m. Рассмотрим такой участок длины $2l$ (рис. 6.3.1). Вычислим вероятность попадания на этот участок по формуле (6.3.7):

$$P(m-l<X<m+l)=$$
$$=\Phi^*\left(\frac{l}{\sigma}\right)-\Phi^*\left(-\frac{l}{\sigma}\right). \qquad (6.3.9)$$

Учитывая свойство (6.3.8) функции $\Phi^*(x)$ и придавая левой части формулы (6.3.9) более компактный вид, получим формулу для вероятности попадания случайной величины, распределенной по нормальному закону, на участок, симметричный относительно центра рассеивания:

$$P(|X-m|<l)=2\Phi^*\left(\frac{l}{\sigma}\right)-1. \qquad (6.3.10)$$

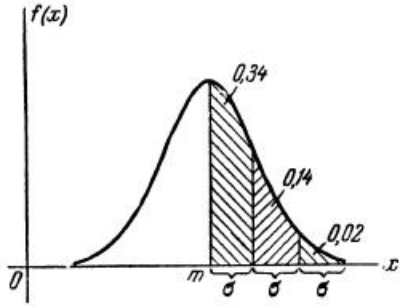


Рис. 6.3.2.

Решим следующую задачу. Отложим от центра рассеивания m последовательные отрезки длиной σ (рис. 6.3.2) и вычислим вероятность попадания случайной величины X в каждый из них. Так как кривая нормального закона симметрична, достаточно отложить такие отрезки только в одну сторону.

По формуле (6.3.7) находим:

$$\left.\begin{aligned}
&P(m<X<m+\sigma)=\Phi^*(1)-\Phi^*(0)=\\
&\qquad=0{,}8413-0{,}5000\approx 0{,}341;\\
&P(m+\sigma<X<m+2\sigma)=\Phi^*(2)-\Phi^*(1)\approx 0{,}136;\\
&P(m+2\sigma<X<m+3\sigma)=\Phi^*(3)-\Phi^*(2)\approx 0{,}012;\\
&P(m+3\sigma<X<m+4\sigma)=\Phi^*(4)-\Phi^*(3)\approx 0{,}001.
\end{aligned}\right\} \qquad (6.3.11)$$

Как видно из этих данных, вероятности попадания на каждый из следующих отрезков (пятый, шестой и т. д.) с точностью до 0,001 равны нулю.

Округляя вероятности попадания в отрезки до 0,01 (до 1%), получим три числа, которые легко запомнить:

$$0{,}34; \quad 0{,}14; \quad 0{,}02.$$

Сумма этих трех значений равна 0,5. Это значит, что для нормально распределенной случайной величины все рассеивание (с точностью до долей процента) укладывается на участке $m \pm 3\sigma$.

Это позволяет, зная среднее квадратическое отклонение и математическое ожидание случайной величины, ориентировочно указать интервал ее практически возможных значений. Такой способ оценки диапазона возможных значений случайной величины известен в математической статистике под названием «*правило трех сигма*». Из правила трех сигма вытекает также ориентировочный способ определения среднего квадратического отклонения случайной величины: берут максимальное практически возможное отклонение от среднего и делят его на три. Разумеется, этот грубый прием может быть рекомендован, только если нет других, более точных способов определения σ.

Пример 1. Случайная величина X, распределенная по нормальному закону, представляет собой ошибку измерения некоторого расстояния. При измерении допускается систематическая ошибка в сторону завышения на 1,2 (*м*); среднее квадратическое отклонение ошибки измерения равно 0,8 (*м*). Найти вероятность того, что отклонение измеренного значения от истинного не превзойдет по абсолютной величине 1,6 (*м*).

Решение. Ошибка измерения есть случайная величина X, подчиненная нормальному закону с параметрами $m = 1{,}2$ и $\sigma = 0{,}8$. Нужно найти вероятность попадания этой величины на участок от $\alpha = -1{,}6$ до $\beta = +1{,}6$. По формуле (6.3.7) имеем:

$$P(-1{,}6 < X < 1{,}6) = \Phi^*\left(\frac{1{,}6 - 1{,}2}{0{,}8}\right) - \Phi^*\left(\frac{-1{,}6 - 1{,}2}{0{,}8}\right) =$$

$$= \Phi^*(0{,}5) - \Phi^*(-3{,}5).$$

Пользуясь таблицами функции $\Phi^*(x)$ (приложение, табл. 1), найдем:

$$\Phi^*(0{,}5) = 0{,}6915; \quad \Phi^*(-3{,}5) = 0{,}0002,$$

откуда

$$P(-1{,}6 < X < 1{,}6) = 0{,}6915 - 0{,}0002 = 0{,}6913 \approx 0{,}691.$$

Пример 2. Найти ту же вероятность, что в предыдущем примере, но при условии, что систематической ошибки нет.

Решение. По формуле (6.3.10), полагая $l = 1{,}6$, найдем:

$$P(|X| < 1{,}6) = 2\Phi^*\left(\frac{1{,}6}{0{,}8}\right) - 1 \approx 0{,}955.$$

Пример 3. По цели, имеющей вид полосы (автострада), ширина которой равна 20 *м*, ведется стрельба в направлении, перпендикулярном автостраде. Прицеливание ведется по средней линии автострады. Среднее

квадратическое отклонение в направлении стрельбы равно $\sigma = 8$ *м*. Имеется систематическая ошибка в направлении стрельбы: недолет 3 *м*. Найти вероятность попадания в автостраду при одном выстреле.

Р е ш е н и е. Выберем начало координат в любой точке на средней линии автострады (рис. 6.3.3) и направим ось абсцисс перпендикулярно автостраде. Попадание или непопадание снаряда в автостраду определяется значением только одной координаты точки падения X (другая координата Y нам безразлична). Случайная величина X распределена по нормальному закону

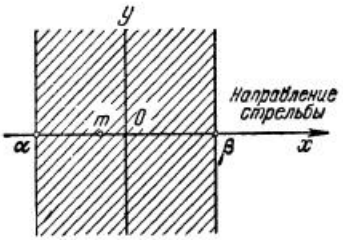


Рис. 6.3.3.

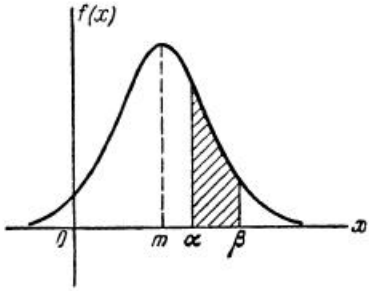


Рис. 6.3.4.

с параметрами $m = -3$, $\sigma = 8$. Попадание снаряда в автостраду соответствует попаданию величины X на участок от $\alpha = -10$ до $\beta = +10$. Применяя формулу (6.3.7), имеем:

$$P(-10 < X < 10) = \Phi^*\left(\frac{13}{8}\right) - \Phi^*\left(-\frac{7}{8}\right) = \Phi^*(1{,}625) - \Phi^*(-0{,}875) \approx 0{,}757.$$

П р и м е р 4. Имеется случайная величина X, нормально распределенная, с центром рассеивания m (рис. 6.3.4) и некоторый участок (α, β) оси абсцисс. Каково должно быть среднее квадратическое отклонение σ случайной величины X для того, чтобы вероятность попадания p на участок (α, β) достигала максимума?

Р е ш е н и е. Имеем:

$$p = P(\alpha < X < \beta) = \Phi^*\left(\frac{\beta - m}{\sigma}\right) - \Phi^*\left(\frac{\alpha - m}{\sigma}\right) = \varphi(\sigma).$$

Продифференцируем эту функцию величины σ:

$$\varphi'(\sigma) = \left[\Phi^*\left(\frac{\beta - m}{\sigma}\right)\right]'_\sigma - \left[\Phi^*\left(\frac{\alpha - m}{\sigma}\right)\right]'_\sigma,$$

но

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

Применяя правило дифференцирования интеграла по переменной, входящей в его предел, получим:

$$\left[\Phi^*\left(\frac{\beta-m}{\sigma}\right)\right]'_\sigma = \left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{\beta-m}{\sigma}} e^{-\frac{t^2}{2}}\,dt\right\}'_\sigma =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(\beta-m)^2}{2\sigma^2}}\left(-\frac{\beta-m}{\sigma^2}\right) =$$

$$= -\frac{\beta-m}{\sigma^2\sqrt{2\pi}} e^{-\frac{(\beta-m)^2}{2\sigma^2}}.$$

Аналогично

$$\left[\Phi^*\left(\frac{\alpha-m}{\sigma}\right)\right]'_\sigma = -\frac{\alpha-m}{\sigma^2\sqrt{2\pi}} e^{-\frac{(\alpha-m)^2}{2\sigma^2}}.$$

Для нахождения экстремума положим:

$$\varphi'(\sigma) = \frac{1}{\sigma^2\sqrt{2\pi}}\left\{(\alpha-m)\,e^{-\frac{(\alpha-m)^2}{2\sigma^2}} - (\beta-m)\,e^{-\frac{(\beta-m)^2}{2\sigma^2}}\right\}. \qquad (6.3.12)$$

При $\sigma=\infty$ это выражение обращается в нуль и вероятность p достигает минимума. Максимум p получим из условия

$$(\alpha-m)\,e^{-\frac{(\alpha-m)^2}{2\sigma^2}} - (\beta-m)\,e^{-\frac{(\beta-m)^2}{2\sigma^2}} = 0. \qquad (6.3.13)$$

Уравнение (6.3.13) можно решить численно или графически.

6.4. Вероятное (срединное) отклонение

В ряде областей практических применений теории вероятностей (в частности, в теории стрельбы) часто, наряду со средним квадратическим отклонением, пользуются еще одной характеристикой рассеивания, так называемым *вероятным*, или *срединным*, отклонением. Вероятное отклонение обычно обозначается буквой E (иногда B).

Вероятным (*срединным*) отклонением случайной величины X, распределенной по нормальному закону, называется половина длины участка, симметричного относительно центра рассеивания, вероятность попадания в который равна половине.

Геометрическая интерпретация вероятного отклонения показана на рис. 6.4.1. Вероятное отклонение E — это половина длины участка оси абсцисс, симметричного относительно точки m, на который опирается половина площади кривой распределения.

Поясним смысл термина «срединное отклонение» или «срединная ошибка», которым часто пользуются в артиллерийской практике вместо «вероятного отклонения».

Рассмотрим случайную величину X, распределенную по нормальному закону. Вероятность того, что она отклонится от центра рассеивания m меньше чем на E, по определению вероятного отклонения E, равна $\frac{1}{2}$:

$$P(|X-m|<E)=\frac{1}{2}. \tag{6.4.1}$$

Вероятность того, что она отклонится от m больше чем на E, тоже равна $\frac{1}{2}$:

$$P(|X-m|>E)=\frac{1}{2}.$$

Таким образом, при большом числе опытов в среднем половина значений случайной величины X отклонится от m больше чем на E, а половина — меньше. Отсюда и термины «срединная ошибка», «срединное отклонение».

Очевидно, вероятное отклонение, как характеристика рассеивания, должно находиться в прямой зависимости от среднего квадратического отклонения σ. Установим эту зависимость. Вычислим вероятность события $|X-m|<E$ в уравнении (6.4.1) по формуле (6.3.10). Имеем:

$$P(|X-m|<E)=2\Phi^*\left(\frac{E}{\sigma}\right)-1=\frac{1}{2}.$$

Отсюда

$$\Phi^*\left(\frac{E}{\sigma}\right)=\frac{3}{4}=0{,}75. \tag{6.4.2}$$

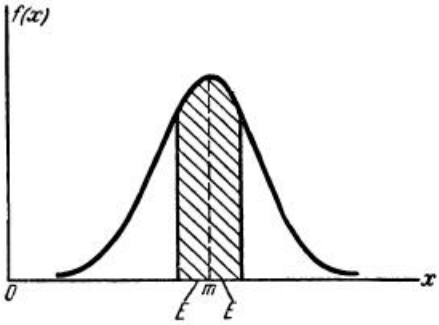


Рис. 6.4.1.

По таблицам функции $\Phi^*(x)$ можно найти такое значение аргумента x, при котором она равна 0,75. Это значение аргумента приближенно равно 0,674; отсюда

$$\frac{E}{\sigma}=0{,}674; \qquad E=0{,}674\sigma. \tag{6.4.3}$$

Таким образом, зная значение σ, можно сразу найти пропорциональное ему значение E. Часто пользуются еще такой формой записи этой зависимости:

$$E=\rho\sqrt{2}\,\sigma, \tag{6.4.4}$$

где ρ — такое значение аргумента, при котором одна из форм интеграла вероятностей — так называемая *функция Лапласа*

$$\Phi(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$$

равна половине. Численное значение величины ρ приближенно равно 0,477.

В настоящее время вероятное отклонение, как характеристика рассеивания, все больше вытесняется более универсальной характеристикой σ. В ряде областей приложений теории вероятностей она сохраняется лишь по традиции.

Если в качестве характеристики рассеивания принято вероятное отклонение E, то плотность нормального распределения записывается в виде:

$$f(x) = \frac{\rho}{E\sqrt{\pi}} e^{-\frac{\rho^2}{E^2}(x-m)^2}, \qquad (6.4.5)$$

а вероятность попадания на участок от α до β чаще всего записывается в виде:

$$P(\alpha < x < \beta) = \frac{1}{2}\left[\hat{\Phi}\left(\frac{\beta - m}{E}\right) - \hat{\Phi}\left(\frac{\alpha - m}{E}\right)\right], \qquad (6.4.6)$$

где

$$\hat{\Phi}(x) = \frac{2}{\sqrt{\pi}} \int_0^{\rho x} e^{-t^2} dt \qquad (6.4.7)$$

— так называемая *приведенная функция Лапласа.*

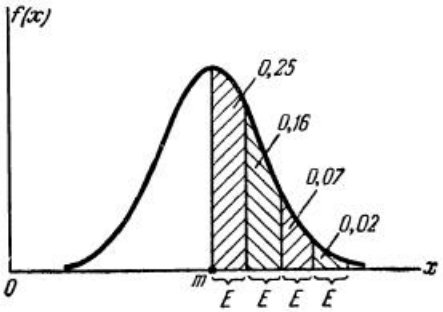


Рис. 6.4.2.

Сделаем подсчет, аналогичный выполненному в предыдущем п° для среднего квадратического отклонения σ: отложим от центра рассеивания m последовательные отрезки длиной в одно вероятное отклонение E (рис. 6.4.2) и подсчитаем вероятности попадания в эти отрезки с точностью до 0,01. Получим:

$$P(m < X < m + E) \approx 0{,}25;$$
$$P(m + E < X < m + 2E) \approx 0{,}16;$$
$$P(m + 2E < X < m + 3E) \approx 0{,}07;$$
$$P(m + 3E < X < m + 4E) \approx 0{,}02.$$

Отсюда видно, что с точностью до 0,01 все значения нормально распределенной случайной величины укладываются на участке $m \pm 4E$.

Пример. Самолет-штурмовик производит обстрел колонны войск противника, ширина которой равна 8 *м*. Полет — вдоль колонны, прицеливание — по средней линии колонны; вследствие скольжения имеется систематическая ошибка: 2 *м* вправо по направлению полета. Главные вероятные отклонения: по направлению полета $B_д = 15$ *м*, в боковом направлении

$B_б = 5$ *м*. Не имея в своем распоряжении никаких таблиц интеграла вероятностей, а зная только числа:

$$25\%,\ 16\%,\ 7\%,\ 2\%,$$

оценить грубо-приближенно вероятность попадания в колонну при одном выстреле и вероятность хотя бы одного попадания при трех независимых выстрелах.

Решение. Для решения задачи достаточно рассмотреть одну координату точки попадания — абсциссу X в направлении, перпендикулярном колонне. Эта абсцисса распределена по нормальному закону с центром рассеивания $m = 2$ и вероятным отклонением $B_б = E = 5$ (*м*). Отложим мысленно от центра рассеивания в ту и другую сторону отрезки длиной в 5 *м*. Вправо от центра рассеивания цель занимает участок 2 *м*, который составляет 0,4 вероятного отклонения. Вероятность попадания на этот участок приближенно равна:

$$0{,}4 \cdot 25\% = 0{,}1.$$

Влево от центра рассеивания цель занимает участок 6 *м*. Это — целое вероятное отклонение (5 *м*), вероятность попадания в которое равна 25% плюс часть длиной 1 *м* следующего (второго от центра) вероятного отклонения, вероятность попадания в которое равна 16%. Вероятность попадания в часть длиной 1 *м* приближенно равна:

$$\frac{1}{5} \cdot 16\% = 0{,}03.$$

Таким образом, вероятность попадания в колонну приближенно равна:

$$0{,}1 + 0{,}25 + 0{,}03 = 0{,}38.$$

Вероятность хотя бы одного попадания при трех выстрелах равна:

$$R_1 = 1 - (1 - 0{,}38)^3 \approx 0{,}76.$$

ГЛАВА 7

ОПРЕДЕЛЕНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН НА ОСНОВЕ ОПЫТНЫХ ДАННЫХ

7.1. Основные задачи математической статистики

Математические законы теории вероятностей не являются беспредметными абстракциями, лишенными физического содержания; они представляют собой математическое выражение реальных закономерностей, фактически существующих в массовых случайных явлениях природы.

До сих пор, говоря о законах распределения случайных величин, мы не затрагивали вопроса о том, откуда берутся, на каком основании устанавливаются эти законы распределения. Ответ на вопрос вполне определенен — в основе всех этих характеристик лежит опыт; каждое исследование случайных явлений, выполняемое методами теории вероятностей, прямо или косвенно опирается на экспериментальные данные. Оперируя такими понятиями, как события и их вероятности, случайные величины, их законы распределения и числовые характеристики, теория вероятностей дает возможность теоретическим путем определять вероятности одних событий через вероятности других, законы распределения и числовые характеристики одних случайных величин через законы распределения и числовые характеристики других. Такие косвенные методы позволяют значительно экономить время и средства, затрачиваемые на эксперимент, но отнюдь не исключают самого эксперимента. Каждое исследование в области случайных явлений, как бы отвлеченно оно ни было, корнями своими всегда уходит в эксперимент, в опытные данные, в систему наблюдений.

Разработка методов регистрации, описания и анализа статистических экспериментальных данных, получаемых в результате наблюдения массовых случайных явлений, составляет предмет специальной науки — *математической статистики.*

Все задачи математической статистики касаются вопросов обработки наблюдений над массовыми случайными явлениями, но в зависимости от характера решаемого практического вопроса и от объема

имеющегося экспериментального материала эти задачи могут принимать ту или иную форму.

Охарактеризуем вкратце некоторые типичные задачи математической статистики, часто встречаемые на практике.

1. Задача определения закона распределения случайной величины (или системы случайных величин) по статистическим данным

Мы уже указывали, что закономерности, наблюдаемые в массовых случайных явлениях, проявляются тем точнее и отчетливее, чем больше объем статистического материала. При обработке обширных по своему объему статистических данных часто возникает вопрос об определении законов распределения тех или иных случайных величин. Теоретически при достаточном количестве опытов свойственные этим случайным величинам закономерности будут осуществляться сколь угодно точно. На практике нам всегда приходится иметь дело с ограниченным количеством экспериментальных данных; в связи с этим результаты наших наблюдений и их обработки всегда содержат больший или меньший элемент случайности. Возникает вопрос о том, какие черты наблюдаемого явления относятся к постоянным, устойчивым и действительно присущи ему, а какие являются случайными и проявляются в данной серии наблюдений только за счет ограниченного объема экспериментальных данных. Естественно, к методике обработки экспериментальных данных следует предъявить такие требования, чтобы она, по возможности, сохраняла типичные, характерные черты наблюдаемого явления и отбрасывала все несущественное, второстепенное, связанное с недостаточным объемом опытного материала. В связи с этим возникает характерная для математической статистики задача *сглаживания* или *выравнивания* статистических данных, представления их в наиболее компактном виде с помощью простых аналитических зависимостей.

2. Задача проверки правдоподобия гипотез

Эта задача тесно связана с предыдущей; при решении такого рода задач мы обычно не располагаем настолько обширным статистическим материалом, чтобы выявляющиеся в нем статистические закономерности были в достаточной мере свободны от элементов случайности. Статистический материал может с бóльшим или меньшим правдоподобием подтверждать или не подтверждать справедливость той или иной гипотезы. Например, может возникнуть такой вопрос: согласуются ли результаты эксперимента с гипотезой о том, что данная случайная величина подчинена закону распределения $F(x)$? Другой подобный вопрос: указывает ли наблюденная в опыте тен-

денция к зависимости между двумя случайными величинами на наличие действительной объективной зависимости между ними или же она объясняется случайными причинами, связанными с недостаточным объемом наблюдений? Для решения подобных вопросов математическая статистика выработала ряд специальных приемов.

3. Задача нахождения неизвестных параметров распределения

Часто при обработке статистического материала вовсе не возникает вопрос об определении законов распределения исследуемых случайных величин. Обыкновенно это бывает связано с крайне недостаточным объемом экспериментального материала. Иногда же характер закона распределения качественно известен до опыта, из теоретических соображений; например, часто можно утверждать заранее, что случайная величина подчинена нормальному закону. Тогда возникает более узкая задача обработки наблюдений — определить только некоторые параметры (числовые характеристики) случайной величины или системы случайных величин. При небольшом числе опытов задача более или менее точного определения этих параметров не может быть решена; в этих случаях экспериментальный материал содержит в себе неизбежно значительный элемент случайности; поэтому случайными оказываются и все параметры, вычисленные на основе этих данных. В таких условиях может быть поставлена только задача об определении так называемых «оценок» или «подходящих значений» для искомых параметров, т. е. таких приближенных значений, которые при массовом применении приводили бы в среднем к меньшим ошибкам, чем всякие другие. С задачей отыскания «подходящих значений» числовых характеристик тесно связана задача оценки их точности и надежности. С подобными задачами мы встретимся в главе 14.

Таков далеко не полный перечень основных задач математической статистики. Мы перечислили только те из них, которые наиболее важны для нас по своим практическим применениям. В настоящей главе мы вкратце познакомимся с некоторыми, наиболее элементарными задачами математической статистики и с методами их решения.

7.2. Простая статистическая совокупность. Статистическая функция распределения

Предположим, что изучается некоторая случайная величина X, закон распределения которой в точности неизвестен, и требуется определить этот закон из опыта или проверить экспериментально гипотезу о том, что величина X подчинена тому или иному закону. С этой целью над случайной величиной X производится ряд независимых опытов (наблюдений). В каждом из этих опытов случайная

величина X принимает определенное значение. Совокупность наблюденных значений величины и представляет собой первичный статистический материал, подлежащий обработке, осмыслению и научному анализу. Такая совокупность называется «простой статистической совокупностью» или «простым статистическим рядом». Обычно простая статистическая совокупность оформляется в виде таблицы с одним входом, в первом столбце которой стоит номер опыта l, а во втором — наблюденное значение случайной величины.

Пример 1. Случайная величина β — угол скольжения самолета в момент сбрасывания бомбы [1]). Произведено 20 бомбометаний, в каждом из которых зарегистрирован угол скольжения β в тысячных долях радиана. Результаты наблюдений сведены в простой статистический ряд:

l	β_l	l	β_l	l	β_l
1	—20	8	—30	15	—10
2	—60	9	120	16	20
3	—10	10	—100	17	30
4	30	11	—80	18	—80
5	60	12	20	19	60
6	70	13	40	20	70
7	—10	14	—60		

Простой статистический ряд представляет собой первичную форму записи статистического материала и может быть обработан различными способами. Одним из способов такой обработки является построение *статистической функции распределения* случайной величины.

Статистической функцией распределения случайной величины X называется частота события $X < x$ в данном статистическом материале:

$$F^*(x) = P^*(X < x). \tag{7.2.1}$$

Для того чтобы найти значение статистической функции распределения при данном x, достаточно подсчитать число опытов, в которых величина X приняла значение, меньшее чем x, и разделить на общее число n произведенных опытов.

Пример 2. Построить статистическую функцию распределения для случайной величины β, рассмотренной в предыдущем примере [2]).

---

[1]) Под углом скольжения подразумевается угол, составленный вектором скорости и плоскостью симметрии самолета.

[2]) Здесь и во многих случаях далее, при рассмотрении конкретных практических примеров, мы не будем строго придерживаться правила — обозначать случайные величины большими буквами, а их возможные значения — соответствующими малыми буквами. Если это не может привести к недоразумениям, мы в ряде случаев будем обозначать случайную величину и ее возможное значение одной и той же буквой.

Решение. Так как наименьшее наблюденное значение величины равно -100, то $F(-100) = 0$. Значение -100 наблюдено один раз, его частота равна $\frac{1}{20}$; следовательно, в точке -100 $F^*(\beta)$ имеет скачок, равный $\frac{1}{20}$. В промежутке от -100 до -80 функция $F^*(\beta)$ имеет значение $\frac{1}{20}$; в точке -80 происходит скачок функции $F^*(\beta)$ на $\frac{2}{20}$, так как значение -80 наблюдено дважды, и т. д.

График статистической функции распределения величины представлен на рис. 7.2.1.

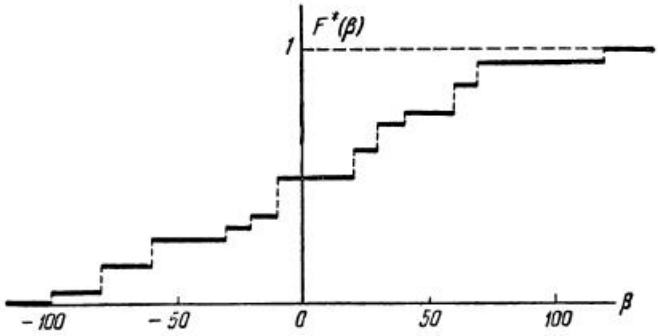


Рис. 7.2.1.

Статистическая функция распределения любой случайной величины — прерывной или непрерывной — представляет собой прерывную ступенчатую функцию, скачки которой соответствуют наблюденным значениям случайной величины и по величине равны частотам этих значений. Если каждое отдельное значение случайной величины X было наблюдено только один раз, скачок статистической функции распределения в каждом наблюденном значении равен $\frac{1}{n}$, где n — число наблюдений.

При увеличении числа опытов n, согласно теореме Бернулли, при любом x частота события $X < x$ приближается (сходится по вероятности) к вероятности этого события. Следовательно, при увеличении n статистическая функция распределения $F^*(x)$ приближается (сходится по вероятности) к подлинной функции распределения $F(x)$ случайной величины X.

Если X — непрерывная случайная величина, то при увеличении числа наблюдений n число скачков функции $F^*(x)$ увеличивается, самые скачки уменьшаются и график функции $F^*(x)$ неограниченно приближается к плавной кривой $F(x)$ — функции распределения величины X.

В принципе построение статистической функции распределения уже решает задачу описания экспериментального материала. Однако при большом числе опытов n построение $F^*(x)$ описанным выше способом весьма трудоемко. Кроме того, часто бывает удобно — в смысле наглядности — пользоваться другими характеристиками статистических распределений, аналогичными не функции распределения $F(x)$, а плотности $f(x)$. С такими способами описания статистических данных мы познакомимся в следующем параграфе.

7.3. Статистический ряд. Гистограмма

При большом числе наблюдений (порядка сотен) простая статистическая совокупность перестает быть удобной формой записи статистического материала — она становится слишком громоздкой и мало наглядной. Для придания ему большей компактности и наглядности статистический материал должен быть подвергнут дополнительной обработке — строится так называемый «статистический ряд».

Предположим, что в нашем распоряжении результаты наблюдений над непрерывной случайной величиной X, оформленные в виде простой статистической совокупности. Разделим весь диапазон наблюденных значений X на интервалы или «разряды» и подсчитаем количество значений m_l, приходящееся на каждый l-й разряд. Это число разделим на общее число наблюдений n и найдем частоту, соответствующую данному разряду:

$$p_l^* = \frac{m_l}{n}. \tag{7.3.1}$$

Сумма частот всех разрядов, очевидно, должна быть равна единице.

Построим таблицу, в которой приведены разряды в порядке их расположения вдоль оси абсцисс и соответствующие частоты. Эта таблица называется *статистическим рядом*:

l_l	$x_1; x_2$	$x_2; x_3$	. . .	$x_l; x_{l+1}$	. . .	$x_k; x_{k+1}$
p_l^*	p_1^*	p_2^*	. . .	p_l^*	. . .	p_k^*

Здесь l_l — обозначение l-го разряда; x_l, x_{l+1} — его границы; p_l^* — соответствующая частота; k — число разрядов.

Пример 1. Произведено 500 измерений боковой ошибки наводки при стрельбе с самолета по наземной цели. Результаты измерений (в тысячных долях радиана) сведены в статистический ряд:

I_l	$-4; -3$	$-3; -2$	$-2; -1$	$-1; 0$	$0; 1$	$1; 2$	$2; 3$	$3; 4$
m_l	6	25	72	133	120	88	46	10
p_l^*	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

Здесь I_l обозначены интервалы значений ошибки наводки; m_l — число наблюдений в данном интервале, $p_l^* = \frac{m_l}{n}$ — соответствующие частоты.

При группировке наблюденных значений случайной величины по разрядам возникает вопрос о том, к какому разряду отнести значение, находящееся в точности на границе двух разрядов. В этих случаях можно рекомендовать (чисто условно) считать данное значение принадлежащим в равной мере к обоим разрядам и прибавлять к числам m_l того и другого разряда по $\frac{1}{2}$.

Число разрядов, на которые следует группировать статистический материал, не должно быть слишком большим (тогда ряд распределения становится невыразительным, и частоты в нем обнаруживают незакономерные колебания); с другой стороны, оно не должно быть слишком малым (при малом числе разрядов свойства распределения описываются статистическим рядом слишком грубо). Практика показывает, что в большинстве случаев рационально выбирать число разрядов порядка 10 — 20. Чем богаче и однороднее статистический материал, тем большее число разрядов можно выбирать при составлении статистического ряда. Длины разрядов могут быть как одинаковыми, так и различными. Проще, разумеется, брать их одинаковыми. Однако при оформлении данных о случайных величинах, распределенных крайне неравномерно, иногда бывает удобно выбирать в области наибольшей плотности распределения разряды более узкие, чем в области малой плотности.

Статистический ряд часто оформляется графически в виде так называемой *гистограммы*. Гистограмма строится следующим образом. По оси абсцисс откладываются разряды, и на каждом из разрядов как их основании строится прямоугольник, площадь которого равна частоте данного разряда. Для построения гистограммы нужно частоту каждого разряда разделить на его длину и полученное число взять в качестве высоты прямоугольника. В случае равных по длине

разрядов высоты прямоугольников пропорциональны соответствующим частотам. Из способа построения гистограммы следует, что полная площадь ее равна единице.

В качестве примера можно привести гистограмму для ошибки наводки, построенную по данным статистического ряда, рассмотренного в примере 1 (рис. 7.3.1).

Очевидно, при увеличении числа опытов можно выбирать все более и более мелкие разряды; при этом гистограмма будет все более приближаться к некоторой кривой, ограничивающей площадь,

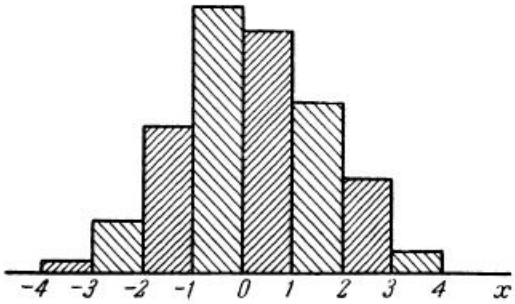


Рис. 7.3.1.

равную единице. Нетрудно убедиться, что эта кривая представляет собой график плотности распределения величины X.

Пользуясь данными статистического ряда, можно приближенно построить и статистическую функцию распределения величины X. Построение точной статистической функции распределения с несколькими сотнями скачков во всех наблюденных значениях X слишком трудоемко и себя не оправдывает. Для практики обычно достаточно построить статистическую функцию распределения по нескольким точкам. В качестве этих точек удобно взять границы $x_1, x_2, \ldots$ разрядов, которые фигурируют в статистическом ряде. Тогда, очевидно,

$$\left.\begin{aligned} &F^*(x_1) = 0;\\ &F^*(x_2) = p_1^*;\\ &F^*(x_3) = p_1^* + p_2^*;\\ &\cdots\cdots\cdots\\ &F^*(x_k) = \sum_{i=1}^{k-1} p_i^*;\\ &F^*(x_{k+1}) = \sum_{i=1}^{k} p_i = 1. \end{aligned}\right\} \qquad (7.3.2)$$

Соединяя полученные точки ломаной линией или плавной кривой, получим приближенный график статистической функции распределения.

Пример 2. Построить приближенно статистическую функцию распределения ошибки наводки по данным статистического ряда примера 1.

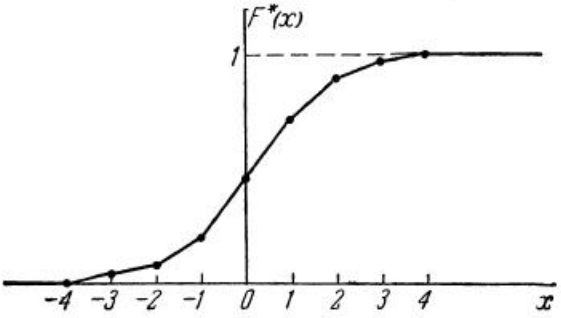


Рис. 7.3.2.

Решение. Применяя формулы (7.3.2), имеем:

$$F^*(-4)=0;\quad F^*(-3)=0{,}012;\quad F^*(-2)=0{,}012+0{,}050=0{,}062;$$
$$F^*(-1)=0{,}206;\quad F^*(0)=0{,}472;\quad F^*(1)=0{,}712;\quad F^*(2)=0{,}888;$$
$$F^*(3)=0{,}980;\quad F^*(4)=1{,}000.$$

Приближенный график статистической функции распределения дан на рис. 7.3.2.

7.4. Числовые характеристики статистического распределения

В главе 5 мы ввели в рассмотрение различные числовые характеристики случайных величин: математическое ожидание, дисперсию, начальные и центральные моменты различных порядков. Эти числовые характеристики играют большую роль в теории вероятностей. Аналогичные числовые характеристики существуют и для статистических распределений. Каждой числовой характеристике случайной величины X соответствует ее статистическая аналогия. Для основной характеристики положения — математического ожидания случайной величины — такой аналогией является среднее арифметическое наблюденных значений случайной величины:

$$M^*[X]=\frac{\sum_{i=1}^{n} x_i}{n}, \tag{7.4.1}$$

где x_i — значение случайной величины, наблюденное в i-м опыте, n — число опытов.

Эту характеристику мы будем в дальнейшем называть *статистическим средним* случайной величины.

Согласно закону больших чисел, при неограниченном увеличении числа опытов статистическое среднее приближается (сходится по вероятности) к математическому ожиданию. При достаточно большом n статистическое среднее может быть принято приближенно равным математическому ожиданию. При ограниченном числе опытов статистическое среднее является случайной величиной, которая, тем не менее, связана с математическим ожиданием и может дать о нем известное представление.

Подобные статистические аналогии существуют для всех числовых характеристик. Условимся в дальнейшем эти статистические аналогии обозначать теми же буквами, что и соответствующие числовые характеристики, но снабжать их значком *.

Рассмотрим, например, дисперсию случайной величины. Она представляет собой математическое ожидание случайной величины $\mathring{X}^2 = (X - m_x)^2$:

$$D[X] = M[\mathring{X}^2] = M[(X - m_x)^2]. \tag{7.4.2}$$

Если в этом выражении заменить математическое ожидание его статистической аналогией — средним арифметическим, мы получим статистическую дисперсию случайной величины X:

$$D^*[X] = \frac{\sum_{i=1}^{n} (x_i - m_x^*)^2}{n}, \tag{7.4.3}$$

где $m_x^* = M^*[X]$ — статистическое среднее.

Аналогично определяются статистические начальные и центральные моменты любых порядков:

$$\alpha_s^*[X] = \frac{\sum_{i=1}^{n} x_i^s}{n}, \tag{7.4.4}$$

$$\mu_s^*[X] = \frac{\sum_{i=1}^{n} (x_i - m_x^*)^s}{n}. \tag{7.4.5}$$

Все эти определения полностью аналогичны данным в главе 5 определениям числовых характеристик случайной величины, с той разницей, что в них везде вместо математического ожидания фигурирует среднее арифметическое. При увеличении числа наблюдений, очевидно, все статистические характеристики будут сходиться по вероятности к соответствующим математическим характеристикам и при достаточном n могут быть приняты приближенно равными им.

Нетрудно доказать, что для статистических начальных и центральных моментов справедливы те же свойства, которые были выведены в главе 5 для математических моментов. В частности, статистический первый центральный момент всегда равен нулю:

$$\mu_1^* = \frac{\sum_{i=1}^{n}(x_i - m_x^*)}{n} = \frac{\sum_{i=1}^{n} x_i}{n} - m_x^* = m_x^* - m_x^* = 0.$$

Соотношения между центральными и начальными моментами также сохраняются:

$$\mu_2^* = D_x^* = \frac{\sum_{i=1}^{n}(x_i - m_x^*)^2}{n} =$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{n} - 2m_x^* \frac{\sum_{i=1}^{n} x_i}{n} + (m_x^*)^2 = \alpha_2^* - (m_x^*)^2 \qquad (7.4.6)$$

и т. д.

При очень большом количестве опытов вычисление характеристик по формулам (7.4.1)—(7.4.5) становится чрезмерно громоздким, и можно применить следующий прием: воспользоваться теми же разрядами, на которые был расклассифицирован статистический материал для построения статистического ряда или гистограммы, и считать приближенно значение случайной величины в каждом разряде постоянным и равным среднему значению, которое выступает в роли «представителя» разряда. Тогда статистические числовые характеристики будут выражаться приближенными формулами:

$$m_x^* = M^*[X] = \sum_{i=1}^{k} \tilde{x}_i p_i^*, \qquad (7.4.7)$$

$$D_x^* = D^*[X] = \sum_{i=1}^{k} (\tilde{x}_i - m_x^*)^2 p_i^*, \qquad (7.4.8)$$

$$\alpha_s^*[X] = \sum_{i=1}^{k} \tilde{x}_i^s p_i^*, \qquad (7.4.9)$$

$$\mu_s^*[X] = \sum_{i=1}^{k} (\tilde{x}_i - m_x^*)^s p_i^*, \qquad (7.4.10)$$

где $\tilde{x}_i$ — «представитель» i-го разряда, p_i^* — частота i-го разряда, k — число разрядов.

Как видно, формулы (7.4.7)—(7.4.10) полностью аналогичны формулам п°п° 5.6 и 5.7, определяющим математическое ожидание, дисперсию, начальные и центральные моменты прерывной случайной

величины X, с той только разницей, что вместо вероятностей p_i в них стоят частоты p_i^*, вместо математического ожидания m_x — статистическое среднее m_x^*, вместо числа возможных значений случайной величины — число разрядов.

В большинстве руководств по теории вероятностей и математической статистике при рассмотрении вопроса о статистических аналогиях для характеристик случайных величин применяется терминология, несколько отличная от принятой в настоящей книге, а именно, статистическое среднее именуется «выборочным средним», статистическая дисперсия—«выборочной дисперсией» и т. д. Происхождение этих терминов следующее. В статистике, особенно сельскохозяйственной и биологической, часто приходится исследовать распределение того или иного признака для весьма большой совокупности индивидуумов, образующих статистический коллектив (таким признаком может быть, например, содержание белка в зерне пшеницы, вес того же зерна, длина или вес тела какого-либо из группы животных и т. д.). Данный признак является случайной величиной, значение которой от индивидуума к индивидууму меняется. Однако, для того, чтобы составить представление о распределении этой случайной величины или о ее важнейших характеристиках, нет необходимости обследовать каждый индивидуум данной обширной совокупности; можно обследовать некоторую в ы б о р к у достаточно большого объема для того, чтобы в ней были выявлены существенные черты изучаемого распределения. Та обширная совокупность, из которой производится выборка, носит в статистике название *генеральной совокупности*. При этом предполагается, что число членов (индивидуумов) N в генеральной совокупности весьма велико, а число членов n в выборке ограничено. При достаточно большом N оказывается, что свойства выборочных (статистических) распределений и характеристик практически не зависят от N; отсюда естественно вытекает математическая идеализация, состоящая в том, что генеральная совокупность, из которой осуществляется выбор, имеет бесконечный объем. При этом отличают точные характеристики (закон распределения, математическое ожидание, дисперсию и т. д.), относящиеся к генеральной совокупности, от аналогичных им «выборочных» характеристик. Выборочные характеристики отличаются от соответствующих характеристик генеральной совокупности за счет ограниченности объема выборки n; при неограниченном увеличении n, естественно, все выборочные характеристики приближаются (сходятся по вероятности) к соответствующим характеристикам генеральной совокупности. Часто возникает вопрос о том, каков должен быть объем выборки n для того, чтобы по выборочным характеристикам можно было с достаточной точностью судить о неизвестных характеристиках генеральной совокупности или о том, с какой степенью точности при заданном объеме выборки можно судить о характеристиках генеральной совокупности. Такой методический прием, состоящий в параллельном рассмотрении бесконечной генеральной совокупности, из которой осуществляется выбор, и ограниченной по объему выборки, является совершенно естественным в тех областях статистики, где фактически приходится осуществлять выбор из весьма многочисленных совокупностей индивидуумов. Для практических задач, связанных с вопросами стрельбы и вооружения, гораздо более характерно другое положение, когда над исследуемой случайной величиной (или системой случайных величин) производится ограниченное число опытов с целью определить те или иные характеристики этой величины, например, когда с целью исследования закона рассеивания при стрельбе производится некоторое количество выстрелов, или с целью исследования ошибки наводки производится серия опытов, в каждом из которых ошибка наводки регистрируется с помощью фотопулемета, и т. д. При этом ограни-

ченное число опытов связано не с трудностью регистрации и обработки, а со сложностью и дороговизной каждого отдельного опыта. В этом случае с известной натяжкой можно также произведенные n опытов мысленно рассматривать как «выборку» из некоторой чисто условной «генеральной совокупности», состоящей из бесконечного числа возможных или мыслимых опытов, которые можно было бы произвести в данных условиях. Однако искусственное введение такой гипотетической «генеральной совокупности» при данной постановке вопроса не вызвано необходимостью и вносит в рассмотрение вопроса, по существу, излишний элемент идеализации, не вытекающий из непосредственной реальности задачи.

Поэтому мы в данном курсе не пользуемся терминами «выборочное среднее», «выборочная дисперсия», «выборочные характеристики» и т. д., заменяя их терминами «статистическое среднее», «статистическая дисперсия», «статистические характеристики».

7.5. Выравнивание статистических рядов

Во всяком статистическом распределении неизбежно присутствуют элементы случайности, связанные с тем, что число наблюдений ограничено, что произведены именно те, а не другие опыты, давшие именно те, а не другие результаты. Только при очень большом числе наблюдений эти элементы случайности сглаживаются, и случайное явление обнаруживает в полной мере присущую ему закономерность. На практике мы почти никогда не имеем дела с таким большим числом наблюдений и вынуждены считаться с тем, что любому статистическому распределению свойственны в большей или меньшей мере черты случайности. Поэтому при обработке статистического материала часто приходится решать вопрос о том, как подобрать для данного статистического ряда теоретическую кривую распределения, выражающую лишь существенные черты статистического материала, но не случайности, связанные с недостаточным объемом экспериментальных данных. Такая задача называется задачей *выравнивания* (сглаживания) *статистических рядов*.

Задача выравнивания заключается в том, чтобы подобрать теоретическую плавную кривую распределения, с той или иной точки зрения наилучшим образом описывающую данное статистическое распределение (рис. 7.5.1).

Задача о наилучшем выравнивании статистических рядов, как и вообще задача о наилучшем аналитическом представлении эмпирических функций, есть задача в значительной мере неопределенная, и решение ее зависит от того, что условиться считать «наилучшим». Например, при сглаживании эмпирических зависимостей очень часто исходят из так называемого принципа или метода наименьших квадратов (см. n° 14.5), считая, что наилучшим приближением к эмпирической зависимости в данном классе функций является такое, при котором сумма квадратов отклонений обращается в минимум. При этом вопрос о том, в каком именно классе функций следует искать наилучшее приближение, решается уже не из математических сообра-

жений, а из соображений, связанных с физикой решаемой задачи, с учетом характера полученной эмпирической кривой и степени точности произведенных наблюдений. Часто принципиальный характер функции, выражающей исследуемую зависимость, известен заранее из теоретических соображений, из опыта же требуется получить лишь некоторые численные параметры, входящие в выражение функции; именно эти параметры подбираются с помощью метода наименьших квадратов.

Аналогично обстоит дело и с задачей выравнивания статистических рядов. Как правило, принципиальный вид теоретической кривой выбирается заранее из соображений, связанных с существом задачи,

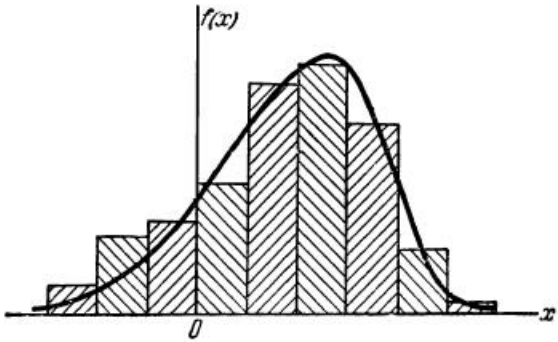


Рис. 7.5.1.

а в некоторых случаях просто с внешним видом статистического распределения. Аналитическое выражение выбранной кривой распределения зависит от некоторых параметров; задача выравнивания статистического ряда переходит в задачу рационального выбора тех значений параметров, при которых соответствие между статистическим и теоретическим распределениями оказывается наилучшим.

Предположим, например, что исследуемая величина X есть ошибка измерения, возникающая в результате суммирования воздействий множества независимых элементарных ошибок; тогда из теоретических соображений можно считать, что величина X подчиняется нормальному закону:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad (7.5.1)$$

и задача выравнивания переходит в задачу о рациональном выборе параметров m и σ в выражении (7.5.1).

Бывают случаи, когда заранее известно, что величина X распределяется статистически приблизительно равномерно на некотором

интервале; тогда можно поставить задачу о рациональном выборе параметров того закона равномерной плотности

$$f(x)=\begin{cases} \frac{1}{\beta-\alpha} & \text{при} \quad \alpha<x<\beta, \\ 0 & \text{при} \quad x<\alpha \quad \text{или} \quad x>\beta, \end{cases}$$

которым можно наилучшим образом заменить (выровнять) заданное статистическое распределение.

Следует при этом иметь в виду, что любая аналитическая функция $f(x)$, с помощью которой выравнивается статистическое распределение, должна обладать основными свойствами плотности распределения:

$$\left.\begin{aligned} f(x) &\geqslant 0; \\ \int_{-\infty}^{\infty} f(x)\,dx &= 1. \end{aligned}\right\} \tag{7.5.2}$$

Предположим, что, исходя из тех или иных соображений, нами выбрана функция $f(x)$, удовлетворяющая условиям (7.5.2), с помощью которой мы хотим выровнять данное статистическое распределение; в выражение этой функции входит несколько параметров a, b, ...; требуется подобрать эти параметры так, чтобы функция $f(x)$ наилучшим образом описывала данный статистический материал. Один из методов, применяемых для решения этой задачи, — это так называемый *метод моментов*.

Согласно методу моментов, параметры a, b, ... выбираются с таким расчетом, чтобы несколько важнейших числовых характеристик (моментов) теоретического распределения были равны соответствующим статистическим характеристикам. Например, если теоретическая кривая $f(x)$ зависит только от двух параметров a и b, эти параметры выбираются так, чтобы математическое ожидание m_x и дисперсия D_x теоретического распределения совпадали с соответствующими статистическими характеристиками $\overset{*}{m}_x$ и $\overset{*}{D}_x$. Если кривая $f(x)$ зависит от трех параметров, можно подобрать их так, чтобы совпали первые три момента, и т. д. При выравнивании статистических рядов может оказаться полезной специально разработанная система *кривых Пирсона*, каждая из которых зависит в общем случае от четырех параметров. При выравнивании эти параметры выбираются с тем расчетом, чтобы сохранить первые четыре момента статистического распределения (математическое ожидание, дисперсию, третий и четвертый моменты) [1]. Оригинальный набор кривых распределения, построенных

[1]) См., например, В. И. Романовский, Математическая статистика, ОНТИ, 1939.

по иному принципу, дал Н. А. Бородачев [1]). Принцип, на котором строится система кривых Н. А. Бородачева, заключается в том, что выбор типа теоретической кривой основывается не на внешних формальных признаках, а на анализе физической сущности случайного явления или процесса, приводящего к тому или иному закону распределения.

Следует заметить, что при выравнивании статистических рядов нерационально пользоваться моментами порядка выше четвертого, так как точность вычисления моментов резко падает с увеличением их порядка.

Пример. 1. В n° 7.3 (стр. 137) приведено статистическое распределение боковой ошибки наводки X при стрельбе с самолета по наземной цели. Требуется выровнять это распределение с помощью нормального закона:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Решение. Нормальный закон зависит от двух параметров: m и σ. Подберем эти параметры так, чтобы сохранить первые два момента — математическое ожидание и дисперсию — статистического распределения.

Вычислим приближенно статистическое среднее ошибки наводки по формуле (7.4.7), причем за представителя каждого разряда примем его середину:

$$m_x^* = -3{,}5 \cdot 0{,}012 - 2{,}5 \cdot 0{,}050 - 1{,}5 \cdot 0{,}144 - 0{,}5 \cdot 0{,}266 + 0{,}5 \cdot 0{,}240 + \\ + 1{,}5 \cdot 0{,}176 + 2{,}5 \cdot 0{,}092 + 3{,}5 \cdot 0{,}020 = 0{,}168.$$

Для определения дисперсии вычислим сначала второй начальный момент по формуле (7.4.9), полагая $s = 2$, $k = 8$

$$\alpha_2^* = \sum_{l=1}^{\infty} \tilde{x}_l^2 p_l^* = 2{,}126.$$

Пользуясь выражением дисперсии через второй начальный момент (формула (7.4.6)), получим:

$$D_x^* = \alpha_2^* - (m_x^*)^2 = 2{,}126 - 0{,}028 = 2{,}098.$$

Выберем параметры m и σ нормального закона так, чтобы выполнялись условия:

$$m = m_x^*, \quad \sigma^2 = D_x^*,$$

то есть примем:

$$m = 0{,}168; \quad \sigma = 1{,}448.$$

Напишем выражение нормального закона:

$$f(x) = \frac{1}{1{,}448\sqrt{2\pi}} e^{-\frac{(x-0{,}168)^2}{2 \cdot 1{,}448^2}}.$$

---

[1]) Н. А. Бородачев, Основные вопросы теории точности производства, АН СССР, М. — Л., 1950.

Пользуясь в табл. 3 приложения, вычислим значения $f(x)$ на границах разрядов

x	—4	—3	—2	—1	0	1	2	3	4
$f(x)$	0,004	0,025	0,090	0,199	0,274	0,234	0,124	0,041	0,008

Построим на одном графике (рис. 7.5.2) гистограмму и выравнивающую ее кривую распределения.

Из графика видно, что теоретическая кривая распределения $f(x)$, сохраняя, в основном существенные особенности статистического распределения, свободна от случайных неправильностей хода гистограммы, которые, по-видимому, могут быть отнесены за счет случайных причин; более серьезное обоснование последнему суждению будет дано в следующем параграфе.

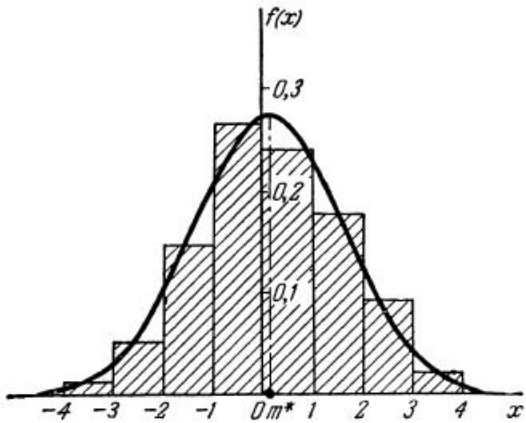


Рис. 7.5.2.

Примечание. В данном примере при определении D_x^* мы воспользовались выражением (7.4.6) статистической дисперсии через второй начальный момент. Этот прием можно рекомендовать только в случае, когда математическое ожидание m_x^* исследуемой случайной величины X сравнительно невелико; в противном случае формула (7.4.6) выражает дисперсию D_x^* как разность близких чисел и дает весьма малую точность. В случае, когда это имеет место, рекомендуется либо вычислять D_x^* непосредственно по формуле (7.4.3), либо перенести начало координат в какую-либо точку, близкую к m_x^*, и затем применить формулу (7.4.6). Пользование формулой (7.4.3) равносильно перенесению начала координат в точку m_x^*; это может

оказаться неудобным, так как выражение m_x^* может быть дробным, и вычитание m_x^* из каждого x_l при этом излишне осложняет вычисления; поэтому рекомендуется переносить начало координат в какое-либо круглое значение x, близкое к m_x^*.

Пример 2. С целью исследования закона распределения ошибки измерения дальности с помощью радиодальномера произведено 400 измерений дальности. Результаты опытов представлены в виде статистического ряда:

l_i (м)	20; 30	30; 40	40; 50	50; 60	60; 70	70; 80	80; 90	90; 100
m_l	21	72	66	38	51	56	64	32
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

Выровнять статистический ряд с помощью закона равномерной плотности.

Решение. Закон равномерной плотности выражается формулой

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{при } \alpha < x < \beta; \\ 0 & \text{при } x < \alpha \text{ или } x > \beta \end{cases}$$

и зависит от двух параметров α и β. Эти параметры следует выбрать так, чтобы сохранить первые два момента статистического распределения — математическое ожидание m_x^* и дисперсию $D_{x'}^*$ Из примера п° 5.8 имеем выражения математического ожидания и дисперсии для закона равномерной плотности:

$$m_x = \frac{\alpha + \beta}{2};$$

$$D_x = \frac{(\beta - \alpha)^2}{12}.$$

Для того чтобы упростить вычисления, связанные с определением статистических моментов, перенесем начало отсчета в точку $x_0 = 60$ и примем за представителя каждого разряда его середину. Ряд распределения примет вид:

$\tilde{x}_l'$	—35	—25	—15	—5	5	15	25	35
p_l^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

где $\tilde{x}_l'$ — среднее для разряда значение ошибки радиодальномера X' при новом начале отсчета.

Приближенное значение статистического среднего ошибки X' равно:

$$m^*_{x'} = \sum_{i=1}^{k} \tilde{x}'_i p^*_i = 0{,}26.$$

Второй статистический момент величины X' равен:

$$\alpha^*_2 = \sum_{i=1}^{k} (\tilde{x}'_i)^2 p^*_i = 447{,}8,$$

откуда статистическая дисперсия:

$$D^*_{x'} = \alpha^*_2 - (m^*_{x'})^2 = 447{,}7.$$

Переходя к прежнему началу отсчета, получим новое статистическое среднее:

$$m^*_x = m^*_{x'} + 60 = 60{,}26$$

и ту же статистическую дисперсию:

$$D^*_x = D^*_{x'} = 447{,}7.$$

Параметры закона равномерной плотности определяются уравнениями:

$$\frac{\alpha+\beta}{2} = 60{,}26; \quad \frac{(\beta-\alpha)^2}{12} = 447{,}7.$$

Решая эти уравнения относительно α и β, имеем:

$$\alpha \approx 23{,}6; \quad \beta \approx 96{,}9,$$

откуда

$$f(x) = \frac{1}{\beta-\alpha} = \frac{1}{73{,}3} \approx 0{,}0136.$$

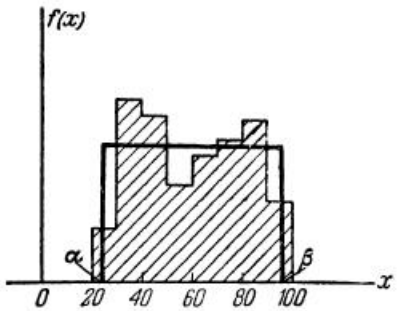


Рис. 7.5.3.

На рис. 7.5.3. показаны гистограмма и выравнивающий ее закон равномерной плотности $f(x)$.

7.6. Критерии согласия

В настоящем п° мы рассмотрим один из вопросов, связанных с проверкой правдоподобия гипотез, а именно—вопрос о согласованности теоретического и статистического распределения.

Допустим, что данное статистическое распределение выравнено с помощью некоторой теоретической кривой $f(x)$ (рис. 7.6.1). Как бы хорошо ни была подобрана теоретическая кривая, между нею и статистическим распределением неизбежны некоторые расхождения. Естественно возникает вопрос: объясняются ли эти расхождения только случайными обстоятельствами, связанными с ограниченным числом наблюдений, или они являются существенными и связаны с тем, что подобранная нами кривая плохо выравнивает данное статистическое распределение. Для ответа на такой вопрос служат так называемые «критерии согласия».

Идея применения критериев согласия заключается в следующем. На основании данного статистического материала нам предстоит проверить гипотезу H, состоящую в том, что случайная величина X подчиняется некоторому определенному закону распределения. Этот закон может быть задан в той или иной форме: например, в виде функции распределения $F(x)$ или в виде плотности распределения $f(x)$, или же в виде совокупности вероятностей p_l, где p_l — вероятность того, что величина X попадет в пределы l-го разряда.

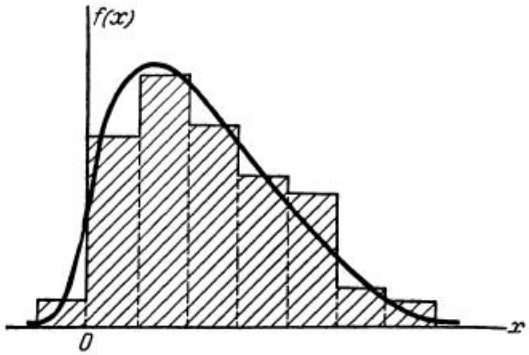


Рис. 7.6.1.

Так как из этих форм функция распределения $F(x)$ является наиболее общей и определяет собой любую другую, будем формулировать гипотезу H, как состоящую в том, что величина X имеет функцию распределения $F(x)$.

Для того чтобы принять или опровергнуть гипотезу H, рассмотрим некоторую величину U, характеризующую степень расхождения теоретического и статистического распределений. Величина U может быть выбрана различными способами; например, в качестве U можно взять сумму квадратов отклонений теоретических вероятностей p_l от соответствующих частот p_l^* или же сумму тех же квадратов с некоторыми коэффициентами («весами»), или же максимальное отклонение статистической функции распределения $F^*(x)$ от теоретической $F(x)$ и т. д. Допустим, что величина U выбрана тем или иным способом. Очевидно, это есть некоторая *случайная величина*. Закон распределения этой случайной величины зависит от закона распределения случайной величины X, над которой производились опыты, и от числа опытов n. Если гипотеза H верна, то закон распределения величины U определяется законом распределения величины X (функцией $F(x)$) и числом n.

Допустим, что этот закон распределения нам известен. В результате данной серии опытов обнаружено, что выбранная нами мера

расхождения U приняла некоторое значение u. Спрашивается, можно ли объяснить это случайными причинами или же это расхождение слишком велико и указывает на наличие существенной разницы между теоретическим и статистическим распределениями и, следовательно, на непригодность гипотезы H? Для ответа на этот вопрос предположим, что гипотеза H верна, и вычислим в этом предположении вероятность того, что за счет случайных причин, связанных с недостаточным объемом опытного материала, мера расхождения U окажется не меньше, чем наблюденное нами в опыте значение u, т. е. вычислим вероятность события:

$$U \geqslant u.$$

Если эта вероятность весьма мала, то гипотезу H следует отвергнуть как мало правдоподобную; если же эта вероятность значительна, следует признать, что экспериментальные данные не противоречат гипотезе H.

Возникает вопрос о том, каким же способом следует выбирать меру расхождения U? Оказывается, что при некоторых способах ее выбора закон распределения величины U обладает весьма простыми свойствами и при достаточно большом n практически не зависит от функции $F(x)$. Именно такими мерами расхождения и пользуются в математической статистике в качестве критериев согласия.

Рассмотрим один из наиболее часто применяемых критериев согласия — так называемый «критерий χ^2» Пирсона.

Предположим, что произведено n независимых опытов, в каждом из которых случайная величина X приняла определенное значение. Результаты опытов сведены в k разрядов и оформлены в виде статистического ряда:

I_l	$x_1; x_2$	$x_2; x_3$	. . .	$x_k; x_{k+1}$
p_l^*	p_1^*	p_2^*	. . .	p_k^*

Требуется проверить, согласуются ли экспериментальные данные с гипотезой о том, что случайная величина X имеет данный закон распределения (заданный функцией распределения $F(x)$ или плотностью $f(x)$). Назовем этот закон распределения «теоретическим».

Зная теоретический закон распределения, можно найти теоретические вероятности попадания случайной величины в каждый из разрядов:

$$p_1, p_2, \ldots, p_k.$$

Проверяя согласованность теоретического и статистического распределений, мы будем исходить из расхождений между теоретическими вероятностями p_l и наблюденными частотами p_l^*. Естественно

выбрать в качестве меры расхождения между теоретическим и статистическим распределениями сумму квадратов отклонений $(p_i^* - p_i)$, взятых с некоторыми «весами» c_i:

$$U = \sum_{i=1}^{k} c_i (p_i^* - p_i)^2. \tag{7.6.1}$$

Коэффициенты c_i («веса» разрядов) вводятся потому, что в общем случае отклонения, относящиеся к различным разрядам, нельзя считать равноправными по значимости. Действительно, одно и то же по абсолютной величине отклонение $p_i^* - p_i$ может быть мало значительным, если сама вероятность p_i велика, и очень заметным, если она мала. Поэтому естественно «веса» c_i взять обратно пропорциональными вероятностям разрядов p_i.

Далее возникает вопрос о том, как выбрать коэффициент пропорциональности.

К. Пирсон показал, что если положить

$$c_i = \frac{n}{p_i}, \tag{7.6.2}$$

то при больших n закон распределения величины U обладает весьма простыми свойствами: он практически не зависит от функции распределения $F(x)$ и от числа опытов n, а зависит только от числа разрядов k, а именно, этот закон при увеличении n приближается к так называемому «распределению χ^2» [1]).

При таком выборе коэффициентов c_i мера расхождения обычно обозначается χ^2:

$$\chi^2 = n \sum_{i=1}^{k} \frac{(p_i^* - p_i)^2}{p_i}. \tag{7.6.3}$$

Для удобства вычислений (чтобы не иметь дела с дробными величинами с большим числом нулей) можно ввести n под знак суммы

---

[1]) Распределением χ^2 с r степенями свободы называется распределение суммы квадратов r независимых случайных величин, каждая из которых подчинена нормальному закону с математическим ожиданием, равным нулю, и дисперсией, равной единице. Это распределение характеризуется плотностью

$$k_r(u) = \begin{cases} \dfrac{1}{2^{\frac{r}{2}}\Gamma\left(\frac{r}{2}\right)} u^{\frac{r}{2}-1} e^{-\frac{u}{2}} & \text{при } u > 0, \\ 0 & \text{при } u < 0, \end{cases}$$

где $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt$ — известная гамма-функция.

и, учитывая, что $p_l^* = \frac{m_l}{n}$, где m_l — число значений в l-м разряде, привести формулу (7.6.3) к виду:

$$U = \chi^2 = \sum_{l=1}^{k} \frac{(m_l - np_l)^2}{np_l}. \qquad (7.6.4)$$

Распределение χ^2 зависит от параметра r, называемого числом «степеней свободы» распределения. Число «степеней свободы» r равно числу разрядов k минус число независимых условий («связей»), наложенных на частоты p_l^*. Примерами таких условий могут быть

$$\sum_{l=1}^{k} p_l^* = 1,$$

если мы требуем только того, чтобы сумма частот была равна единице (это требование накладывается во всех случаях);

$$\sum_{l=1}^{k} \tilde{x}_l p_l^* = m_x^*,$$

если мы подбираем теоретическое распределение с тем условием, чтобы совпадали теоретическое и статистическое средние значения;

$$\sum_{l=1}^{k} (\tilde{x}_l - m_x^*)^2 p_l^* = D_x^*,$$

если мы требуем, кроме того, совпадения теоретической и статистической дисперсий и т. д.

Для распределения χ^2 составлены специальные таблицы (см. табл. 4 приложения). Пользуясь этими таблицами, можно для каждого значения χ^2 и числа степеней свободы r найти вероятность p того, что величина, распределенная по закону χ^2, превзойдет это значение. В табл. 4 входами являются: значение вероятности p и число степеней свободы r. Числа, стоящие в таблице, представляют собой соответствующие значения χ^2.

Распределение χ^2 дает возможность оценить степень согласованности теоретического и статистического распределений. Будем исходить из того, что величина X *действительно* распределена по закону $F(x)$. Тогда вероятность p, определенная по таблице, есть вероятность того, что за счет чисто случайных причин мера расхождения теоретического и статистического распределений (7.6.4) будет не меньше, чем фактически наблюденное в данной серии опытов значение χ^2. Если эта вероятность p весьма мала (настолько мала, что событие с такой вероятностью можно считать практически невозможным), то результат опыта следует считать *противоречащим* гипотезе H о том, что закон распределения величины X есть $F(x)$. Эту

гипотезу следует отбросить как неправдоподобную. Напротив, если вероятность p сравнительно велика, можно признать расхождения между теоретическим и статистическим распределениями несущественными и отнести их за счет случайных причин. Гипотезу H о том, что величина X распределена по закону $F(x)$, можно считать правдоподобной или, по крайней мере, не противоречащей опытным данным.

Таким образом, схема применения критерия χ^2 к оценке согласованности теоретического и статистического распределений сводится к следующему:

1) Определяется мера расхождения χ^2 по формуле (7.6.4).

2) Определяется число степеней свободы r как число разрядов k минус число наложенных связей s:

$$r = k - s.$$

3) По r и χ^2 с помощью табл. 4 определяется вероятность того, что величина, имеющая распределение χ^2 с r степенями свободы, превзойдет данное значение χ^2. Если эта вероятность весьма мала, гипотеза отбрасывается как неправдоподобная. Если эта вероятность относительно велика, гипотезу можно признать не противоречащей опытным данным.

Насколько мала должна быть вероятность p для того, чтобы отбросить или пересмотреть гипотезу, — вопрос неопределенный; он не может быть решен из математических соображений, так же как и вопрос о том, насколько мала должна быть вероятность события для того, чтобы считать его практически невозможным. На практике, если p оказывается меньшим чем 0,1, рекомендуется проверить эксперимент, если возможно — повторить его и в случае, если заметные расхождения снова появятся, пытаться искать более подходящий для описания статистических данных закон распределения.

Следует особо отметить, что с помощью критерия χ^2 (или любого другого критерия согласия) можно только в некоторых случаях опровергнуть выбранную гипотезу H и отбросить ее как явно несогласную с опытными данными; если же вероятность p велика, то этот факт сам по себе ни в коем случае не может считаться доказательством справедливости гипотезы H, а указывает только на то, что гипотеза не противоречит опытным данным.

С первого взгляда может показаться, что чем больше вероятность p, тем лучше согласованность теоретического и статистического распределений и тем более обоснованным следует считать выбор функции $F(x)$ в качестве закона распределения случайной величины. В действительности это не так. Допустим, например, что, оценивая согласие теоретического и статистического распределений по критерию χ^2, мы получили $p = 0{,}99$. Это значит, что с вероятностью 0,99 за счет чисто случайных причин при данном числе опытов

должны были получиться расхождения бóльшие, чем наблюденные. Мы же получили относительно весьма малые расхождения, которые слишком малы для того, чтобы признать их правдоподобными. Разумнее признать, что столь близкое совпадение теоретического и статистического распределений не является случайным и может быть объяснено определенными причинами, связанными с регистрацией и обработкой опытных данных (в частности, с весьма распространенной на практике «подчисткой» опытных данных, когда некоторые результаты произвольно отбрасываются или несколько изменяются).

Разумеется, все эти соображения применимы только в тех случаях, когда количество опытов n достаточно велико (порядка нескольких сотен) и когда имеет смысл применять сам критерий, основанный на предельном распределении меры расхождения при $n \to \infty$. Заметим, что при пользовании критерием χ^2 достаточно большим должно быть не только общее число опытов n, но и числа наблюдений m_l в отдельных разрядах. На практике рекомендуется иметь в каждом разряде не менее 5 — 10 наблюдений. Если числа наблюдений в отдельных разрядах очень малы (порядка 1 — 2), имеет смысл объединить некоторые разряды.

Пример 1. Проверить согласованность теоретического и статистического распределений для примера 1 n° 7.5 (стр. 137, 146).

Решение. Пользуясь теоретическим нормальным законом распределения с параметрами

$$m = 0{,}168, \qquad \sigma = 1{,}448,$$

находим вероятности попадания в разряды по формуле

$$p_l = \Phi^*\left(\frac{x_{l+1} - m}{\sigma}\right) - \Phi^*\left(\frac{x_l - m}{\sigma}\right),$$

где x_l, x_{l+1} — границы l-го разряда.

Затем составляем сравнительную таблицу чисел попаданий в разряды m_l и соответствующих значений np_l ($n = 500$).

l_l	—4; —3	—3; —2	—2; —1	—1; 0	0; 1	1; 2	2; 3	3; 4
m_l	6	25	72	133	120	88	46	10
np_l	6,2	26,2	71,2	122,2	131,8	90,5	38,2	10,5

По формуле (7.6.4) определяем значение меры расхождения

$$\chi^2 = \sum_{l=1}^{8} \frac{(m_l - np_l)^2}{np_l} = 3{,}94.$$

Определяем число степеней свободы как число разрядов минус число наложенных связей s (в данном случае $s = 3$):

$$r = 8 - 3 = 5.$$

По табл. 4 приложения находим для $r = 5$:

при $\chi^2 = 3{,}00$ $p = 0{,}70$;
при $\chi^2 = 4{,}35$ $p = 0{,}50$.

Следовательно, искомая вероятность p при $\chi^2 = 3{,}94$ приближенно равна 0,56. Эта вероятность малой не является; поэтому гипотезу о том, что величина X распределена по нормальному закону, можно считать правдоподобной.

Пример 2. Проверить согласованность теоретического и статистического распределений для условий примера 2 п° 7.5 (стр. 149).

Решение. Значения p_i вычисляем как вероятности попадания на участки (20; 30), (30; 40) и т. д. для случайной величины, распределенной по закону равномерной плотности на отрезке (23,6; 96,9). Составляем сравнительную таблицу значений m_i и np_i ($n = 400$):

I_i	20; 30	30; 40	40; 50	50; 60	60; 70	70; 80	80; 90	90; 100
m_i	21	72	66	38	51	56	64	32
np_i	34,9	54,6	54,6	54,6	54,6	54,6	54,6	38,0

По формуле (7.6.4) находим χ^2:

$$\chi^2 = \sum_{i=1}^{8} \frac{(m_i - np_i)^2}{np_i} \approx 20{,}9.$$

Число степеней свободы:

$$r = 8 - 3 = 5.$$

По табл. 4 приложения имеем:

при $\chi^2 = 20{,}5$ и $r = 5$ $p = 0{,}001$,

Следовательно, наблюденное нами расхождение между теоретическим и статистическим распределениями могло бы за счет чисто случайных причин появиться лишь с вероятностью $p \approx 0{,}001$. Так как эта вероятность очень мала, следует признать экспериментальные данные противоречащими гипотезе о том, что величина X распределена по закону равномерной плотности.

Кроме критерия χ^2, для оценки степени согласованности теоретического и статистического распределений на практике применяется еще ряд других критериев. Из них мы вкратце остановимся на критерии А. Н. Колмогорова.

В качестве меры расхождения между теоретическим и статистическим распределениями А. Н. Колмогоров рассматривает максимальное значение модуля разности между статистической функцией распределения $F^*(x)$ и соответствующей теоретической функцией распределения:

$$D = \max |F^*(x) - F(x)|.$$

Основанием для выбора в качестве меры расхождения величины D является простота ее вычисления. Вместе с тем она имеет достаточно

простой закон распределения. А. Н. Колмогоров доказал, что, какова бы ни была функция распределения $F(x)$ непрерывной случайной величины X, при неограниченном возрастании числа независимых наблюдений n вероятность неравенства

$$D\sqrt{n} \geqslant \lambda$$

стремится к пределу

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2}. \qquad (7.6.5)$$

Значения вероятности $P(\lambda)$, подсчитанные по формуле (7.6.5), приведены в таблице 7.6.1.

Таблица 7.6.1

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,0	1,000	0,7	0,711	1,4	0,040
0,1	1,000	0,8	0,544	1,5	0,022
0,2	1,000	0,9	0,393	1,6	0,012
0,3	1,000	1,0	0,270	1,7	0,006
0,4	0,997	1,1	0,178	1,8	0,003
0,5	0,964	1,2	0,112	1.9	0,002
0,6	0,864	1,3	0,068	2,0	0,001

Схема применения критерия А. Н. Колмогорова следующая: строятся статистическая функция распределения $F^*(x)$ и предполагаемая

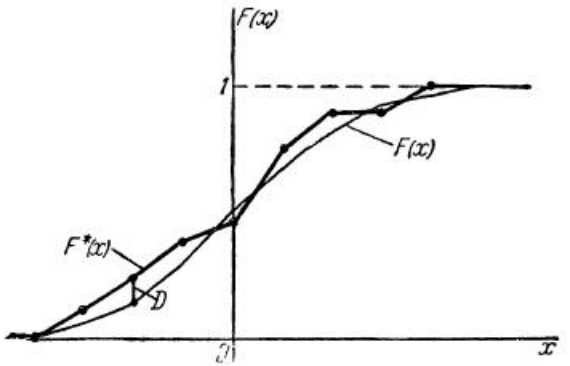


Рис. 7.6.2.

теоретическая функция распределения $F(x)$, и определяется максимум D модуля разности между ними (рис. 7.6.2).

Далее, определяется величина

$$\lambda = D\sqrt{n}$$

и по таблице 7.6.1 находится вероятность $P(\lambda)$. Это есть вероятность того, что (если величина X действительно распределена по закону $F(x)$) за счет чисто случайных причин максимальное расхождение между $F^*(x)$ и $F(x)$ будет не меньше, чем фактически наблюденное. Если вероятность $P(\lambda)$ весьма мала, гипотезу следует отвергнуть как неправдоподобную; при сравнительно больших $P(\lambda)$ ее можно считать совместимой с опытными данными.

Критерий А. Н. Колмогорова своей простотой выгодно отличается от описанного ранее критерия χ^2; поэтому его весьма охотно применяют на практике. Следует, однако, оговорить, что этот критерий можно применять только в случае, когда гипотетическое распределение $F(x)$ полностью известно заранее из каких-либо теоретических соображений, т. е. когда известен не только вид функции распределения $F(x)$, но и все входящие в нее параметры. Такой случай сравнительно редко встречается на практике. Обычно из теоретических соображений известен только общий вид функции $F(x)$, а входящие в нее числовые параметры определяются по данному статистическому материалу. При применении критерия χ^2 это обстоятельство учитывается соответствующим уменьшением числа степеней свободы распределения χ^2. Критерий А. Н. Колмогорова такого согласования не предусматривает. Если все же применять этот критерий в тех случаях, когда параметры теоретического распределения выбираются по статистическим данным, критерий дает заведомо завышенные значения вероятности $P(\lambda)$; поэтому мы в ряде случаев рискуем принять как правдоподобную гипотезу, в действительности плохо согласующуюся с опытными данными.

ГЛАВА 8

СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

8.1. Понятие о системе случайных величин

В практических применениях теории вероятностей очень часто приходится сталкиваться с задачами, в которых результат опыта описывается не одной случайной величиной, а двумя или более случайными величинами, образующими *комплекс* или *систему*. Например, точка попадания снаряда определяется не одной случайной величиной, а двумя: абсциссой и ординатой — и может быть рассмотрена как комплекс двух случайных величин. Аналогично точка разрыва дистанционного снаряда определяется комплексом трех случайных величин. При стрельбе группой из n выстрелов совокупность точек попадания на плоскости может рассматриваться как комплекс или система $2n$ случайных величин: n абсцисс и n ординат точек попадания. Осколок, образовавшийся при разрыве снаряда, характеризуется рядом случайных величин: весом, размерами, начальной скоростью, направлением полета и т. д. Условимся систему нескольких случайных величин $X, Y, \ldots, W$ обозначать $(X, Y, \ldots, W)$.

Свойства системы нескольких случайных величин не исчерпываются свойствами отдельных величин, ее составляющих: помимо этого, они включают также взаимные связи (зависимости) между случайными величинами.

При рассмотрении вопросов, связанных с системами случайных величин, удобно пользоваться геометрической интерпретацией системы. Например, систему двух случайных величин (X, Y) можно изображать случайной точкой на плоскости с координатами X и Y (рис. 8.1.1). Аналогично система трех случайных величин может быть изображена случайной точкой в трехмерном пространстве. Часто бывает удобно говорить о системе n случайных величин как о «случайной точке в пространстве n измерений». Несмотря на то, что последняя интерпретация не обладает непосредственной наглядностью, пользование ею дает некоторый выигрыш в смысле общности терминологии и упрощения записей.

Часто вместо образа *случайной точки* для геометрической интерпретации системы случайных величин пользуются образом *случайного вектора*. Систему двух случайных величин при этом рассматривают как случайный вектор на плоскости xOy, составляющие которого по осям представляют собой случайные величины X, Y (рис. 8.1.2). Система трех случайных величин изображается случайным вектором в трехмерном пространстве, система n случайных величин —

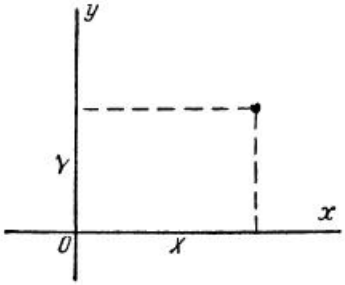


Рис. 8.1.1.

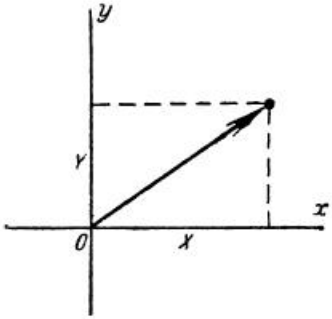


Рис. 8.1.2.

случайным вектором в пространстве n измерений. При этом теория систем случайных величин рассматривается как *теория случайных векторов*.

В данном курсе мы будем в зависимости от удобства изложения пользоваться как одной, так и другой интерпретацией.

Занимаясь системами случайных величин, мы будем рассматривать как полные, исчерпывающие вероятностные характеристики — законы распределения, так и неполные — числовые характеристики.

Изложение начнем с наиболее простого случая системы двух случайных величин.

8.2. Функция распределения системы двух случайных величин

Функцией распределения системы двух случайных величин (X, Y) называется вероятность совместного выполнения двух неравенств $X < x$ и $Y < y$:

$$F(x, y) = P\big((X < x)(Y < y)\big). \tag{8.2.1}$$

Если пользоваться для геометрической интерпретации системы образом случайной точки, то функция распределения $F(x, y)$ есть не что иное, как вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной в точке (x, y), лежащий левее

и ниже ее (рис. 8.2.1). В аналогичной интерпретации функция распределения одной случайной величины X — обозначим ее $F_1(x)$ — представляет собой вероятность попадания случайной точки в полуплоскость, ограниченную справа абсциссой x (рис. 8.2.2); функция распределения одной величины Y — $F_2(y)$ — вероятность попадания в полуплоскость, ограниченную сверху ординатой y (рис. 8.2.3).

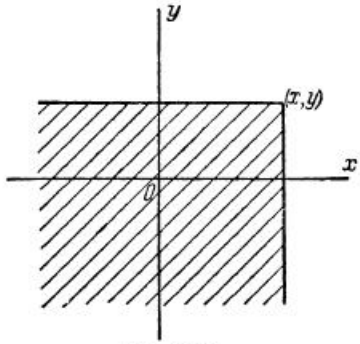


Рис. 8.2.1.

В п° 5.2 мы привели основные свойства функции распределения $F(x)$ для одной случайной величины. Сформулируем аналогичные свойства для функции распределения системы случайных величин и снова воспользуемся геометрической интерпретацией для наглядной иллюстрации этих свойств.

1. Функция распределения $F(x, y)$ есть неубывающая функция обоих своих аргументов, т. е.

$$\text{при } x_2 > x_1 \quad F(x_2, y) \geqslant F(x_1, y);$$
$$\text{при } y_2 > y_1 \quad F(x, y_2) \geqslant F(x, y_1).$$

В этом свойстве функции $F(x)$ можно наглядно убедиться, пользуясь геометрической интерпретацией функции распределения как вероятности попадания в квадрант с вершиной (x, y) (рис. 8.2.1). Действительно, увеличивая x (смещая

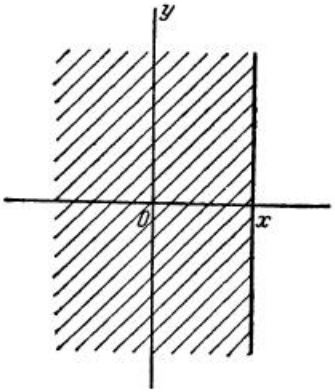


Рис. 8.2.2.

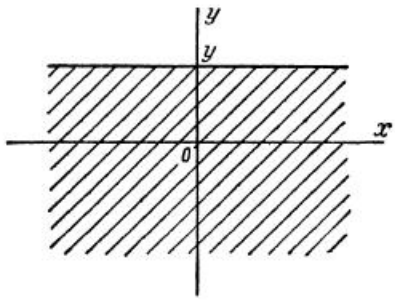


Рис. 8.2.3.

правую границу квадранта вправо) или увеличивая y (смещая верхнюю границу вверх), мы, очевидно, не можем уменьшить вероятность попадания в этот квадрант.

2. Повсюду на $-\infty$ функция распределения равна нулю:

$$F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0.$$

В этом свойстве мы наглядно убеждаемся, неограниченно отодвигая влево правую границу квадранта ($x \to -\infty$) или вниз его верхнюю границу ($y \to -\infty$) или делая это одновременно с обеими границами; при этом вероятность попадания в квадрант стремится к нулю.

3. При одном из аргументов, равном $+\infty$, функция распределения системы превращается в функцию распределения случайной величины, соответствующей другому аргументу:

$$F(x, +\infty) = F_1(x), \quad F(+\infty, y) = F_2(y),$$

где $F_1(x)$, $F_2(y)$ — соответственно функции распределения случайных величин X и Y.

В этом свойстве функции распределения можно наглядно убедиться, смещая ту или иную из границ квадранта на $+\infty$; при этом в пределе квадрант превращается в полуплоскость, вероятность попадания в которую есть функция распределения одной из величин, входящих в систему.

4. Если оба аргумента равны $+\infty$, функция распределения системы равна единице:

$$F(+\infty, +\infty) = 1.$$

Действительно, при $x \to +\infty$, $y \to +\infty$ квадрант с вершиной (x, y) в пределе обращается во всю плоскость xOy, попадание в которую есть достоверное событие.

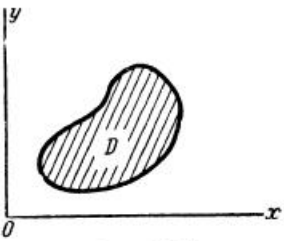


Рис. 8.2.4.

При рассмотрении законов распределения отдельных случайных величин (глава 5) мы вывели выражение для вероятности попадания случайной величины в пределы заданного участка. Эту вероятность мы выразили как через функцию распределения, так и через плотность распределения.

Аналогичным вопросом для системы двух случайных величин является вопрос о вероятности попадания случайной точки (X, Y) в пределы заданной области D на плоскости xOy (рис. 8.2.4).

Условимся событие, состоящее в попадании случайной точки (X, Y) в область D, обозначать символом $(X, Y) \subset D$.

Вероятность попадания случайной точки в заданную область выражается наиболее просто в том случае, когда эта область представляет собой прямоугольник со сторонами, параллельными координатным осям.

Выразим через функцию распределения системы вероятность попадания случайной точки (X, Y) в прямоугольник R, ограниченный абсциссами α и β и ординатами γ и δ (рис. 8.2.5).

При этом следует условиться, куда мы будем относить границы прямоугольника. Аналогично тому, как мы делали для одной случайной величины, условимся включать в прямоугольник R его нижнюю и левую границы и не включать верхнюю и правую [1]). Тогда событие $(X, Y) \subset R$ будет равносильно произведению двух событий: $\alpha \leqslant X < \beta$ и $\gamma \leqslant Y < \delta$. Выразим вероятность этого события через

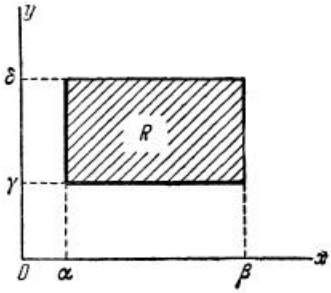


Рис. 8.2.5.

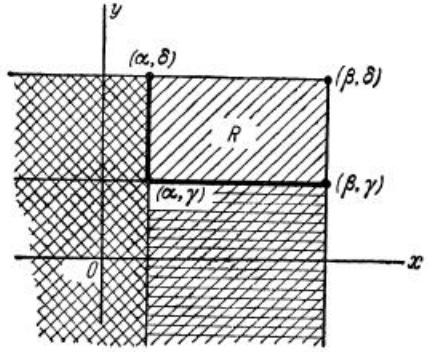


Рис. 8.2.6.

функцию распределения системы. Для этого рассмотрим на плоскости xOy четыре бесконечных квадранта с вершинами в точках (β, δ); (α, δ); (β, γ) и (α, γ) (рис. 8.2.6).

Очевидно, вероятность попадания в прямоугольник R равна вероятности попадания в квадрант (β, δ) минус вероятность попадания в квадрант (α, δ) минус вероятность попадания в квадрант (β, γ) плюс вероятность попадания в квадрант (α, γ) (так как мы дважды вычли вероятность попадания в этот квадрант). Отсюда получаем формулу, выражающую вероятность попадания в прямоугольник через функцию распределения системы:

$$P((X, Y) \subset R) = F(\beta, \delta) - F(\alpha, \delta) - F(\beta, \gamma) + F(\alpha, \gamma). \quad (8.2.2)$$

В дальнейшем, когда будет введено понятие плотности распределения системы, мы выведем формулу для вероятности попадания случайной точки в область произвольной формы.

8.3. Плотность распределения системы двух случайных величин

Введенная в предыдущем n° характеристика системы — функция распределения — существует для систем любых случайных величин, как прерывных, так и непрерывных. Основное практическое значение

[1]) На рис. 8.2.5 границы, включенные в прямоугольник, даны жирными линиями.

имеют системы непрерывных случайных величин. Распределение системы непрерывных величин обычно характеризуют не функцией распределения, а **плотностью распределения**.

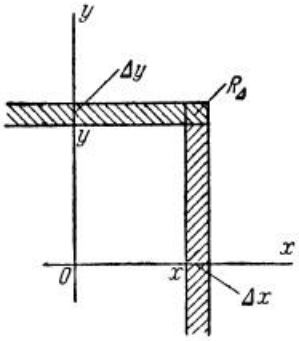


Рис. 8.3.1.

Вводя в рассмотрение плотность распределения для одной случайной величины, мы определяли ее как предел отношения вероятности попадания на малый участок к длине этого участка при ее неограниченном уменьшении. Аналогично определим плотность распределения системы двух величин.

Пусть имеется система двух непрерывных случайных величин (X, Y), которая интерпретируется случайной точкой на плоскости xOy. Рассмотрим на этой плоскости малый прямоугольник R_Δ со сторонами Δx и Δy, примыкающий к точке с координатами (x, y) (рис. 8.3.1). Вероятность попадания в этот прямоугольник по формуле (8.2.2) равна

$$P((X, Y) \subset R_\Delta) = F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y) - F(x, y+\Delta y) + F(x, y).$$

Разделим вероятность попадания в прямоугольник R_Δ на площадь этого прямоугольника и перейдем к пределу при $\Delta x \to 0$ и $\Delta y \to 0$:

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P((X, Y) \subset R_\Delta)}{\Delta x \, \Delta y} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y) - F(x, y+\Delta y) + F(x, y)}{\Delta x \, \Delta y}. \tag{8.3.1}$$

Предположим, что функция $F(x, y)$ не только непрерывна, но и дифференцируема; тогда правая часть формулы (8.3.1) представляет собой *вторую смешанную частную производную* функции $F(x, y)$ по x и y. Обозначим эту производную $f(x, y)$:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \, \partial y} = F''_{xy}(x, y). \tag{8.3.2}$$

Функция $f(x, y)$ называется *плотностью распределения* системы.

Таким образом, плотность распределения системы представляет собой предел отношения вероятности попадания в малый прямоугольник к площади этого прямоугольника, когда оба его размера стремятся к нулю; она может быть выражена как вторая смешанная частная производная функции распределения системы по обоим аргументам.

Если воспользоваться «механической» интерпретацией распределения системы как распределения единичной массы по плоскости xOy, функция $f(x, y)$ представляет собой плотность распределения массы в точке (x, y).

Геометрически функцию $f(x, y)$ можно изобразить некоторой поверхностью (рис. 8.3.2). Эта поверхность аналогична кривой распределения для одной случайной величины и называется *поверхностью распределения*.

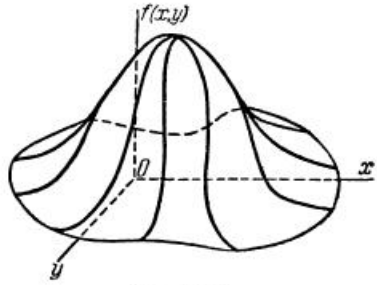


Рис. 8.3.2.

Если пересечь поверхность распределения $f(x, y)$ плоскостью, параллельной плоскости xOy, и спроектировать полученное сечение на плоскость xOy, получится кривая, в каждой точке которой плотность распределения постоянна. Такие кривые называются *кривыми равной плотности*. Кривые равной плотности, очевидно, представляют собой *горизонтали* поверхности распределения. Часто бывает удобно задавать распределение семейством кривых равной плотности.

Рассматривая плотность распределения $f(x)$ для одной случайной величины, мы ввели понятие «элемента вероятности» $f(x)\,dx$. Это есть вероятность попадания случайной величины X на элементарный участок dx, прилежащий к точке x. Аналогичное понятие «элемента вероятности» вводится и для системы двух случайных величин. Элементом вероятности в данном случае называется выражение

$$f(x, y)\,dx\,dy.$$

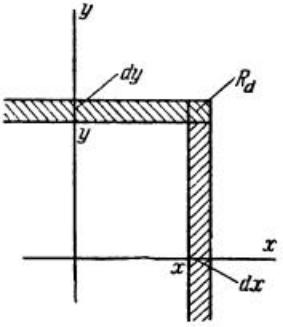


Рис. 8.3.3.

Очевидно, элемент вероятности есть не что иное, как вероятность попадания в элементарный прямоугольник со сторонами dx, dy, примыкающий к точке (x, y) (рис. 8.3.3). Эта вероятность равна объему элементарного параллелепипеда, ограниченного сверху поверхностью $f(x, y)$ и опирающегося на элементарный прямоугольник $dx\,dy$ (рис. 8.3.4).

Пользуясь понятием элемента вероятности, выведем выражение для вероятности попадания случайной точки в произвольную область D. Эта вероятность, очевидно, может быть получена суммированием (интегрированием) элементов вероятности по всей области D:

$$P((X, Y) \subset D) = \iint\limits_{(D)} f(x, y)\,dx\,dy. \qquad (8.3.3)$$

Геометрически вероятность попадания в область D изображается объемом цилиндрического тела C, ограниченного сверху поверхностью распределения и опирающегося на область D (рис. 8.3.5).

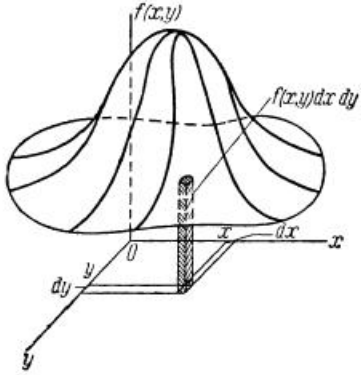


Рис. 8.3.4.

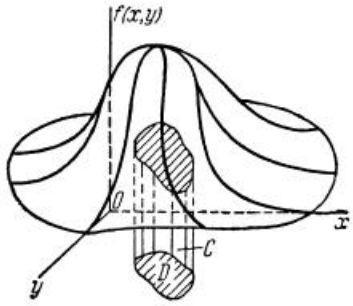


Рис. 8.3.5.

Из общей формулы (8.3.3) вытекает формула для вероятности попадания в прямоугольник R, ограниченный абсциссами α и β и ординатами γ и δ (рис. 8.2.5):

$$P((X, Y) \subset R) = \int_{\alpha}^{\beta} \int_{\gamma}^{\delta} f(x, y)\, dx\, dy. \qquad (8.3.4)$$

Воспользуемся формулой (8.3.4) для того, чтобы выразить функцию распределения системы $F(x, y)$ через плотность распределения $f(x, y)$. Функция распределения $F(x, y)$ есть вероятность попадания в бесконечный квадрант; последний можно рассматривать как прямоугольник, ограниченный абсциссами $-\infty$ и x и ординатами $-\infty$ и y. По формуле (8.3.4) имеем:

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dx\, dy. \qquad (8.3.5)$$

Легко убедиться в следующих свойствах плотности распределения системы:

1. Плотность распределения системы есть функция неотрицательная:

$$f(x, y) \geqslant 0.$$

Это ясно из того, что плотность распределения есть предел отношения двух неотрицательных величин: вероятности попадания в прямоугольник и площади прямоугольника — и, следовательно, отрицательной быть не может.

2. Двойной интеграл в бесконечных пределах от плотности распределения системы равен единице:

$$\int_{-\infty}^{\infty}\int f(x,\ y)\,dx\,dy = 1. \qquad (8.3.6)$$

Это видно из того, что интеграл (8.3.6) есть не что иное, как вероятность попадания во всю плоскость xOy, т. е. вероятность достоверного события.

Геометрически это свойство означает, что полный объем тела, ограниченного поверхностью распределения и плоскостью xOy, равен единице.

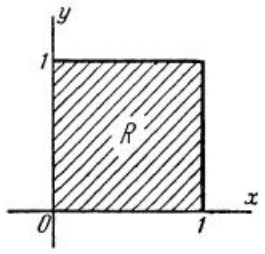


Рис. 8.3.6.

Пример 1. Система двух случайных величин (X, Y) подчинена закону распределения с плотностью

$$f(x, y) = \frac{1}{\pi^2 (1+x^2)(1+y^2)}.$$

Найти функцию распределения $F(x, y)$. Определить вероятность попадания случайной точки (X, Y) в квадрат R (рис. 8.3.6).

Решение. Функцию распределения $F(x, y)$ находим по формуле (8.3.5)

$$F(x, y) = \frac{1}{\pi^2}\int_{-\infty}^{x}\int_{-\infty}^{y}\frac{dx\,dy}{(1+x^2)(1+y^2)} = \left(\frac{1}{\pi}\operatorname{arctg} x + \frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg} y + \frac{1}{2}\right).$$

Вероятность попадания в прямоугольник R находим по формуле (8.3.4):

$$P((X, Y) \subset R) = \frac{1}{\pi^2}\int_0^1\int_0^1\frac{dx\,dy}{(1+x^2)(1+y^2)} = \frac{1}{\pi^2}\int_0^1\frac{dx}{1+x^2}\int_0^1\frac{dy}{1+y^2} = \frac{1}{16}.$$

Пример 2. Поверхность распределения системы (X, Y) представляет собой прямой круговой конус, основанием которого служит круг радиуса R с центром в начале координат. Написать выражение плотности распределения. Определить вероятность того, что случайная точка (x, y) попадет в круг K радиуса a (рис. 8.3.7), причем $a < R$.

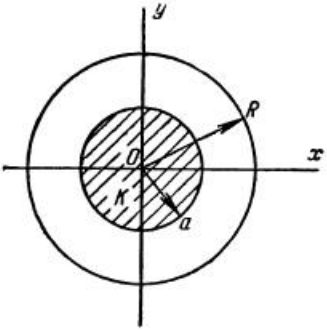


Рис. 8.3.7.

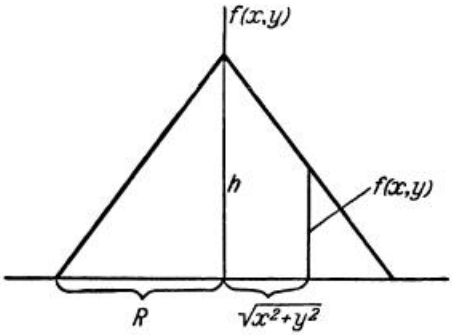


Рис. 8.3.8.

Решение. Выражение плотности распределения внутри круга K находим из рис. 8.3.8:

$$f(x, y) = \frac{h}{R}(R - \sqrt{x^2 + y^2}),$$

где h — высота конуса. Величину h определяем так, чтобы объем конуса был равен единице: $\frac{1}{3}\pi R^2 h = 1$, откуда

$$h = \frac{3}{\pi R^2},$$

и

$$f(x, y) = \frac{3}{\pi R^3}(R - \sqrt{x^2 + y^2}).$$

Вероятность попадания в круг K определяем по формуле (8.3.4):

$$P((X, Y) \subset K) = \int\limits_{(K)} \int f(x, y)\, dx\, dy. \tag{8.3.7}$$

Для вычисления интеграла (8.3.7) удобно перейти к полярной системе координат r, φ:

$$P((X, Y) \subset K) =$$

$$= \int\limits_0^a \int\limits_0^{2\pi} \frac{3}{\pi R^3}(R - r)\, r\, dr\, d\varphi = \frac{6}{R^3} \int\limits_0^a (R - r)\, r\, dr = 3\left(\frac{a}{R}\right)^2 - 2\left(\frac{a}{R}\right)^3.$$

8.4. Законы распределения отдельных величин, входящих в систему. Условные законы распределения

Зная закон распределения системы двух случайных величин, можно всегда определить законы распределения отдельных величин, входящих в систему. В п° 8.2 мы уже вывели выражения для функций распределения отдельных величин, входящих в систему, через функцию распределения системы, а именно, мы показали, что

$$F_1(x) = F(x, \infty); \quad F_2(y) = F(\infty, y). \tag{8.4.1}$$

Выразим теперь плотность распределения каждой из величин, входящих в систему, через плотность распределения системы. Пользуясь формулой (8.3.5), выражающей функцию распределения через плотность распределения, напишем:

$$F_1(x) = F(x, \infty) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{\infty} f(x, y)\, dx\, dy,$$

откуда, дифференцируя по x, получим выражение для плотности распределения величины X:

$$f_1(x) = F_1'(x) = \int\limits_{-\infty}^{\infty} f(x, y)\, dy. \tag{8.4.2}$$

Аналогично

$$f_2(y) = F_2'(y) = \int_{-\infty}^{\infty} f(x, y)\, dx. \qquad (8.4.3)$$

Таким образом, для того чтобы получить плотность распределения одной из величин, входящих в систему, нужно плотность распределения системы проинтегрировать в бесконечных пределах по аргументу, соответствующему другой случайной величине.

Формулы (8.4.1), (8.4.2) и (8.4.3) дают возможность, зная закон распределения системы (заданный в виде функции распределения или плотности распределения), найти законы распределения отдельных величин, входящих в систему. Естественно, возникает вопрос об обратной задаче: нельзя ли по законам распределения отдельных величин, входящих в систему, восстановить закон распределения системы? Оказывается, что в общем случае этого сделать нельзя: зная только законы распределения отдельных величин, входящих в систему, не всегда можно найти закон распределения системы. Для того чтобы исчерпывающим образом охарактеризовать систему, недостаточно знать распределение каждой из величин, входящих в систему; нужно еще знать зависимость между величинами, входящими в систему. Эта зависимость может быть охарактеризована с помощью так называемых *условных законов распределения*.

Условным законом распределения величины X, входящей в систему (X, Y), называется ее закон распределения, вычисленный при условии, что другая случайная величина Y приняла определенное значение y.

Условный закон распределения можно задавать как функцией распределения, так и плотностью. Условная функция распределения обозначается $F(x \mid y)$, условная плотность распределения $f(x \mid y)$. Так как системы непрерывных величин имеют основное практическое значение, мы в данном курсе ограничимся рассмотрением условных законов, заданных плотностью распределения.

Чтобы нагляднее пояснить понятие условного закона распределения, рассмотрим пример. Система случайных величин L и Q представляет собой длину и вес осколка снаряда. Пусть нас интересует длина осколка L безотносительно к его весу; это есть случайная величина, подчиненная закону распределения с плотностью $f_1(l)$. Этот закон распределения мы можем исследовать, рассматривая все без исключения осколки и оценивая их только по длине; $f_1(l)$ есть безусловный закон распределения длины осколка. Однако нас может интересовать и закон распределения длины осколка вполне определенного веса, например 10 *г*. Для того чтобы его определить, мы будем исследовать не все осколки, а только определенную весовую группу, в которой вес приблизительно равен 10 *г*, и получим условный закон распределения длины осколка при весе 10 *г*

с плотностью $f(l \mid q)$ при $q = 10$. Этот условный закон распределения вообще отличается от безусловного $f_1(l)$; очевидно, более тяжелые осколки должны в среднем обладать и большей длиной; следовательно, условный закон распределения длины осколка существенно зависит от веса q.

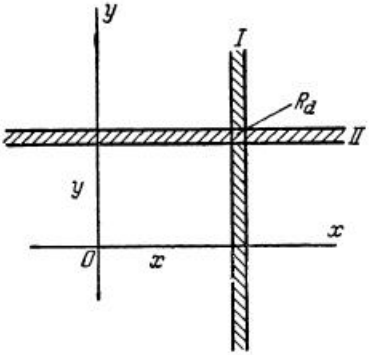


Рис. 8.4.1.

Зная закон распределения одной из величин, входящих в систему, и условный закон распределения второй, можно составить закон распределения системы. Выведем формулу, выражающую это соотношение, для непрерывных случайных величин. Для этого воспользуемся понятием об элементе вероятности. Рассмотрим прилежащий к точке (x, y) элементарный прямоугольник R_d со сторонами dx, dy (рис. 8.4.1). Вероятность попадания в этот прямоугольник — элемент вероятности $f(x, y)\,dx\,dy$ — равна вероятности одновременного попадания случайной точки (X, Y) в элементарную полосу I, опирающуюся на отрезок dx, и в полосу II, опирающуюся на отрезок dy:

$$f(x, y)\,dx\,dy = P\big((X, Y) \subset R_d\big) = \\ = P\big((x < X < x + dx)(y < Y < y + dy)\big).$$

Вероятность произведения этих двух событий, по теореме умножения вероятностей, равна вероятности попадания в элементарную полосу I, умноженной на условную вероятность попадания в элементарную полосу II, вычисленную при условии, что первое событие имело место. Это условие в пределе равносильно условию $X = x$; следовательно,

$$f(x, y)\,dx\,dy = f_1(x)\,dx\,f(y \mid x)\,dy,$$

откуда

$$f(x, y) = f_1(x)\,f(y \mid x), \tag{8.4.4}$$

т. е. плотность распределения системы двух величин равна плотности распределения одной из величин, входящих в систему, умноженной на условную плотность распределения другой величины, вычисленную при условии, что первая величина приняла заданное значение.

Формулу (8.4.4) часто называют *теоремой умножения законов распределения*. Эта теорема в схеме случайных величин аналогична теореме умножения вероятностей в схеме событий.

Очевидно, формуле (8.4.4) можно придать другой вид, если задать значение не величины X, а величины Y:

$$f(x,\ y) = f_2(y)\, f(x \mid y). \tag{8.4.5}$$

Разрешая формулы (8.4.4) и (8.4.5) относительно $f(y \mid x)$ и $f(x \mid y)$, получим выражения условных законов распределения через безусловные:

$$\left.\begin{aligned} f(y \mid x) &= \frac{f(x, y)}{f_1(x)}, \\ f(x \mid y) &= \frac{f(x, y)}{f_2(y)}, \end{aligned}\right\} \tag{8.4.6}$$

или, применяя формулы (8.4.2) и (8.4.3),

$$\left.\begin{aligned} f(y \mid x) &= \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)\, dy}; \\ f(x \mid y) &= \frac{f(x, y)}{\int\limits_{-\infty}^{\infty} f(x, y)\, dx}. \end{aligned}\right\} \tag{8.4.7}$$

8.5. Зависимые и независимые случайные величины

При изучении систем случайных величин всегда следует обращать внимание на степень и характер их зависимости. Эта зависимость может быть более или менее ярко выраженной, более или менее тесной. В некоторых случаях зависимость между случайными величинами может быть настолько тесной, что, зная значение одной случайной величины, можно в точности указать значение другой. В другом крайнем случае зависимость между случайными величинами является настолько слабой и отдаленной, что их можно практически считать независимыми.

Понятие о независимых случайных величинах — одно из важных понятий теории вероятностей.

Случайная величина Y называется независимой от случайной величины X, если закон распределения величины Y не зависит от того, какое значение приняла величина X.

Для непрерывных случайных величин условие независимости Y от X может быть записано в виде:

$$f(y \mid x) = f_2(y)$$

при любом y.

Напротив, в случае, если Y зависит от X, то

$$f(y \mid x) \neq f_2(y).$$

Докажем, что *зависимость или независимость случайных величин всегда взаимны*: если величина Y не зависит от X, то и величина X не зависит от Y.

Действительно, пусть Y не зависит от X:

$$f(y \mid x) = f_2(y). \tag{8.5.1}$$

Из формул (8.4.4) и (8.4.5) имеем:

$$f_1(x) f(y \mid x) = f_2(y) f(x \mid y),$$

откуда, принимая во внимание (8.5.1), получим.

$$f(x \mid y) = f_1(x),$$

что и требовалось доказать.

Так как зависимость и независимость случайных величин всегда взаимны, можно дать новое определение независимых случайных величин.

Случайные величины X и Y называются независимыми, если закон распределения каждой из них не зависит от того, какое значение приняла другая. В противном случае величины X и Y называются *зависимыми*.

Для независимых непрерывных случайных величин теорема умножения законов распределения принимает вид:

$$f(x, y) = f_1(x) f_2(y). \tag{8.5.2}$$

т. е. *плотность распределения системы независимых случайных величин равна произведению плотностей распределения отдельных величин, входящих в систему.*

Условие (8.5.2) может рассматриваться как необходимое и достаточное условие независимости случайных величин.

Часто по самому виду функции $f(x, y)$ можно заключить, что случайные величины X, Y являются независимыми, а именно, если плотность распределения $f(x, y)$ распадается на произведение двух функций, из которых одна зависит только от x, другая — только от y, то случайные величины независимы.

Пример. Плотность распределения системы (X, Y) имеет вид:

$$f(x, y) = \frac{1}{\pi^2 (x^2 + y^2 + x^2 y^2 + 1)}.$$

Определить, зависимы или независимы случайные величины X и Y.

Решение. Разлагая знаменатель на множители, имеем:

$$f(x, y) = \frac{1}{\pi (1 + x^2)} \cdot \frac{1}{\pi (1 + y^2)}.$$

Из того, что функция $f(x, y)$ распалась на произведение двух функций, из которых одна зависит только от x, а другая — только от y, заключаем,

что величины X и Y должны быть независимы. Действительно, применяя формулы (8.4.2) и (8.4.3), имеем:

$$f_1(x) = \frac{1}{\pi(1+x^2)} \int_{-\infty}^{\infty} \frac{dy}{\pi(1+y^2)} = \frac{1}{\pi(1+x^2)};$$

аналогично

$$f_2(y) = \frac{1}{\pi(1+y^2)},$$

откуда убеждаемся, что

$$f(x, y) = f_1(x) f_2(y)$$

и, следовательно, величины X и Y независимы.

Вышеизложенный критерий суждения о зависимости или независимости случайных величин исходит из предположения, что закон распределения системы нам известен. На практике чаще бывает наоборот: закон распределения системы (X, Y) не известен; известны только законы распределения отдельных величин, входящих в систему, и имеются основания считать, что величины X и Y независимы. Тогда можно написать плотность распределения системы как произведение плотностей распределения отдельных величин, входящих в систему.

Остановимся несколько подробнее на важных понятиях о «зависимости» и «независимости» случайных величин.

Понятие «зависимости» случайных величин, которым мы пользуемся в теории вероятностей, несколько отличается от обычного понятия «зависимости» величин, которым мы оперируем в математике. Действительно, обычно под «зависимостью» величин подразумевают только один тип зависимости — полную, жесткую, так называемую *функциональную* зависимость. Две величины X и Y называются функционально зависимыми, если, зная значение одной из них, можно точно указать значение другой.

В теории вероятностей мы встречаемся с другим, более общим, типом зависимости — с *вероятностной* или «стохастической» зависимостью. Если величина Y связана с величиной X вероятностной зависимостью, то, зная значение X, нельзя указать точно значение Y, а можно указать только ее закон распределения, зависящий от того, какое значение приняла величина X.

Вероятностная зависимость может быть более или менее тесной; по мере увеличения тесноты вероятностной зависимости она все более приближается к функциональной. Таким образом, функциональную зависимость можно рассматривать как крайний, предельный случай наиболее тесной вероятностной зависимости. Другой крайний случай — полная независимость случайных величин. Между этими двумя крайними случаями лежат все градации вероятностной зависимости — от самой сильной до самой слабой. Те физические величины, которые на практике мы считаем функционально зависимыми, в действительности связаны весьма тесной вероятностной зависимостью: при

заданном значении одной из этих величин другая колеблется в столь узких пределах, что ее практически можно считать вполне определенной. С другой стороны, те величины, которые мы на практике считаем независимыми, в действительности часто находятся в некоторой взаимной зависимости, но эта зависимость настолько слаба, что ею для практических целей можно пренебречь.

Вероятностная зависимость между случайными величинами очень часто встречается на практике. Если случайные величины X и Y находятся в вероятностной зависимости, это не означает, что с изменением величины X величина Y изменяется вполне определенным образом; это лишь означает, что с изменением величины X величина Y имеет тенденцию также изменяться (например, возрастать или убывать при возрастании X). Эта тенденция соблюдается лишь «в среднем», в общих чертах, и в каждом отдельном случае от нее возможны отступления.

Рассмотрим, например, две такие случайные величины: X — рост наугад взятого человека, Y — его вес. Очевидно, величины X и Y находятся в определенной вероятностной зависимости; она выражается в том, что в общем люди с бóльшим ростом имеют бóльший вес. Можно даже составить эмпирическую формулу, приближенно заменяющую эту вероятностную зависимость функциональной. Такова, например, общеизвестная формула, приближенно выражающая зависимость между ростом и весом:

$$Y\,(кг) = X\,(см) - 100.$$

Формулы подобного типа, очевидно, не являются точными и выражают лишь некоторую среднюю, массовую закономерность, тенденцию, от которой в каждом отдельном случае возможны отступления.

В вышеприведенном примере мы имели дело со случаем явно выраженной зависимости. Рассмотрим теперь такие две случайные величины: X — рост наугад взятого человека; Z — его возраст. Очевидно, для взрослого человека величины X и Z можно считать практически независимыми; напротив, для ребенка величины X и Z являются зависимыми.

Приведем еще несколько примеров случайных величин, находящихся в различных степенях зависимости.

1. Из камней, составляющих кучу щебня, выбирается наугад один камень. Случайная величина Q — вес камня; случайная величина L — наибольшая длина камня. Величины Q и L находятся в явно выраженной вероятностной зависимости.

2. Производится стрельба ракетой в заданный район океана. Величина ΔX — продольная ошибка точки попадания (недолет, перелет); случайная величина ΔV — ошибка в скорости ракеты в конце активного участка движения. Величины ΔX и ΔV явно зависимы, так как ошибка ΔV является одной из главных причин, порождающих продольную ошибку ΔX.

3. Летательный аппарат, находясь в полете, измеряет высоту над поверхностью Земли с помощью барометрического прибора. Рассматриваются две случайные величины: ΔH — ошибка измерения высоты и G — вес топлива, сохранившегося в топливных баках к моменту измерения. Величины ΔH и G практически можно считать независимыми.

В следующем п° мы познакомимся с некоторыми числовыми характеристиками системы случайных величин, которые дадут нам возможность оценивать степень зависимости этих величин.

8.6. Числовые характеристики системы двух случайных величин. Корреляционный момент. Коэффициент корреляции

В главе 5 мы ввели в рассмотрение числовые характеристики одной случайной величины X — начальные и центральные моменты различных порядков. Из этих характеристик важнейшими являются две: математическое ожидание m_x и дисперсия D_x.

Аналогичные числовые характеристики — начальные и центральные моменты различных порядков — можно ввести и для системы двух случайных величин.

Начальным моментом порядка k, s системы (X, Y) называется математическое ожидание произведения X^k на Y^s:

$$\alpha_{k,s} = M[X^k Y^s]. \tag{8.6.1}$$

Центральным моментом порядка k, s системы (X, Y) называется математическое ожидание произведения k-й и s-й степени соответствующих центрированных величин:

$$\mu_{k,s} = M[\mathring{X}^k \mathring{Y}^s], \tag{8.6.2}$$

где $\mathring{X} = X - m_x$, $\mathring{Y} = Y - m_y$.

Выпишем формулы, служащие для непосредственного подсчета моментов. Для прерывных случайных величин

$$\alpha_{k,s} = \sum_i \sum_j x_i^k y_j^s p_{ij}, \tag{8.6.3}$$

$$\mu_{k,s} = \sum_i \sum_j (x_i - m_x)^k (y_j - m_y)^s p_{ij}, \tag{8.6.4}$$

где $p_{ij} = P\big((X = x_i)(Y = y_j)\big)$ — вероятность того, что система (X, Y) примет значения (x_i, y_j), а суммирование распространяется по всем возможным значениям случайных величин X, Y.

Для непрерывных случайных величин:

$$\alpha_{k,s} = \int_{-\infty}^{\infty}\int x^k y^s f(x, y)\,dx\,dy, \qquad (8.6.5)$$

$$\mu_{k,s} = \int_{-\infty}^{\infty}\int (x - m_x)^k (y - m_y)^s f(x, y)\,dx\,dy, \qquad (8.6.6)$$

где $f(x, y)$ — плотность распределения системы.

Помимо чисел k и s, характеризующих порядок момента по отношению к отдельным величинам, рассматривается еще суммарный порядок момента $k+s$, равный сумме показателей степеней при X и Y. Соответственно суммарному порядку моменты классифицируются на первые, вторые и т. д. На практике обычно применяются только первые и вторые моменты.

Первые начальные моменты представляют собой уже известные нам математические ожидания величин X и Y, входящих в систему:

$$m_x = \alpha_{1,0} = M[X^1 Y^0] = M[X],$$
$$m_y = \alpha_{0,1} = M[X^0 Y^1] = M[Y].$$

Совокупность математических ожиданий m_x, m_y представляет собой *характеристику положения* системы. Геометрически это координаты средней точки на плоскости, вокруг которой происходит рассеивание точки (X, Y).

Кроме первых начальных моментов, на практике широко применяются еще *вторые центральные моменты* системы. Два из них представляют собой уже известные нам *дисперсии* величин X и Y:

$$D_x = \mu_{2,0} = M[\mathring{X}^2 \mathring{Y}^0] = M[\mathring{X}^2] = D[X],$$

$$D_y = \mu_{0,2} = M[\mathring{X}^0 \mathring{Y}^2] = M[\mathring{Y}^2] = D[Y],$$

характеризующие *рассеивание* случайной точки в направлении осей Ox и Oy.

Особую роль как характеристика системы играет *второй смешанный центральный момент*:

$$\mu_{1,1} = M[\mathring{X}\,\mathring{Y}],$$

т. е. математическое ожидание произведения центрированных величин.

Ввиду того, что этот момент играет важную роль в теории систем случайных величин, введем для него особое обозначение:

$$K_{xy} = M[\mathring{X}\,\mathring{Y}] = M[(X - m_x)(Y - m_y)]. \qquad (8.6.7)$$

Характеристика K_{xy} называется *корреляционным моментом* (иначе — «моментом связи») случайных величин X, Y.

Для прерывных случайных величин корреляционный момент выражается формулой

$$K_{xy} = \sum_i \sum_j (x_i - m_x)(y_j - m_y) p_{ij}, \qquad (8.6.8)$$

а для непрерывных — формулой

$$K_{xy} = \int_{-\infty}^{\infty}\int (x - m_x)(y - m_y) f(x, y)\, dx\, dy. \qquad (8.6.9)$$

Выясним смысл и назначение этой характеристики.

Корреляционный момент есть характеристика системы случайных величин, описывающая, помимо рассеивания величин X и Y, еще и связь между ними. Для того чтобы убедиться в этом, докажем, что *для независимых случайных величин корреляционный момент равен нулю.*

Доказательство проведем для непрерывных случайных величин [1]). Пусть X, Y — независимые непрерывные величины с плотностью распределения $f(x, y)$. В п° 8.5 мы доказали, что для независимых величин

$$f(x, y) = f_1(x) f_2(y), \qquad (8.6.10)$$

где $f_1(x)$, $f_2(y)$ — плотности распределения соответственно величин X и Y.

Подставляя выражение (8.6.10) в формулу (8.6.9), видим, что интеграл (8.6.9) превращается в произведение двух интегралов:

$$K_{xy} = \int_{-\infty}^{\infty} (x - m_x) f_1(x)\, dx \int_{-\infty}^{\infty} (y - m_y) f_2(y)\, dy.$$

Интеграл

$$\int_{-\infty}^{\infty} (x - m_x) f_1(x)\, dx$$

представляет собой не что иное, как первый центральный момент величины X, и, следовательно, равен нулю; по той же причине равен нулю и второй сомножитель; следовательно, для независимых случайных величин $K_{xy} = 0$.

Таким образом, если корреляционный момент двух случайных величин отличен от нуля, это есть признак наличия зависимости между ними.

Из формулы (8.6.7) видно, что корреляционный момент характеризует не только зависимость величин, но и их рассеивание. Действительно, если, например, одна из величин (X, Y) весьма мало отклоняется от своего математического ожидания (почти не случайна),

---

[1]) Для прерывных оно может быть выполнено аналогичным способом.

то корреляционный момент будет мал, какой бы тесной зависимостью ни были связаны величины (X, Y). Поэтому для характеристики связи между величинами (X, Y) в чистом виде переходят от момента K_{xy} к безразмерной характеристике

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}, \tag{8.6.11}$$

где σ_x, σ_y — средние квадратические отклонения величин X, Y. Эта характеристика называется *коэффициентом корреляции* величин X и Y. Очевидно, коэффициент корреляции обращается в нуль одновременно с корреляционным моментом; следовательно, *для независимых случайных величин коэффициент корреляции равен нулю.*

Случайные величины, для которых корреляционный момент (а значит, и коэффициент корреляции) равен нулю, называются *некоррелированными* (иногда — «несвязанными»).

Выясним, эквивалентно ли понятие некоррелированности случайных величин понятию независимости. Выше мы доказали, что две независимые случайные величины всегда являются некоррелированными. Остается выяснить: справедливо ли обратное положение, вытекает ли из некоррелированности величин их независимость? Оказывается — нет. Можно построить примеры таких случайных величин, которые являются некоррелированными, но зависимыми. Равенство нулю коэффициента корреляции — необходимое, но не достаточное условие независимости случайных величин. Из независимости случайных величин вытекает их некоррелированность; напротив, из некоррелированности величин еще не следует их независимость. Условие независимости случайных величин — более жесткое, чем условие некоррелированности.

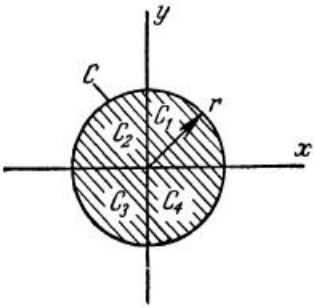


Рис. 8.6.1.

Убедимся в этом на примере. Рассмотрим систему случайных величин (X, Y), распределенную с равномерной плотностью внутри круга C радиуса r с центром в начале координат (рис. 8.6.1).

Плотность распределения величин (X, Y) выражается формулой

$$f(x, y) = \begin{cases} c & \text{при} \quad x^2 + y^2 < r^2, \\ 0 & \text{при} \quad x^2 + y^2 > r^2. \end{cases}$$

Из условия $\int\limits_{-\infty}^{\infty}\int f(x, y)\,dx\,dy = \int\limits_{(C)}\int c\,dx\,dy = 1$ находим $c = \frac{1}{\pi r^2}$.

Нетрудно убедиться, что в данном примере величины являются зависимыми. Действительно, непосредственно ясно, что если величина X приняла, например, значение 0, то величина Y может с равной вероятностью принимать все значения от $-r$ до $+r$; если же величина X приняла значение r, то величина Y может принять только одно-единственное значение, в точности равное нулю; вообще, диапазон возможных значений Y зависит от того, какое значение приняла X.

Посмотрим, являются ли эти величины коррелированными. Вычислим корреляционный момент. Имея в виду, что по соображениям симметрии $m_x = m_y = 0$, получим:

$$K_{xy} = \int\limits_{(C)}\int xyf(x, y)\,dx\,dy = \frac{1}{\pi r^2}\int\limits_{(C)}\int xy\,dx\,dy. \quad (8.6.12)$$

Для вычисления интеграла разобьем область интегрирования (круг C) на четыре сектора C_1, C_2, C_3, C_4, соответствующие четырем координатным углам. В секторах C_1 и C_3 подынтегральная функция положительна, в секторах C_2 и C_4 — отрицательна; по абсолютной же величине интегралы по этим секторам равны; следовательно, интеграл (8.6.12) равен нулю, и величины (X, Y) не коррелированы.

Таким образом, мы видим, что из некоррелированности случайных величин не всегда следует их независимость.

Коэффициент корреляции характеризует не всякую зависимость, а только так называемую *линейную* зависимость. Линейная вероятностная зависимость случайных величин заключается в том, что при возрастании одной случайной величины другая имеет тенденцию возрастать (или же убывать) по линейному закону. Эта тенденция к линейной зависимости может быть более или менее ярко выраженной, более или менее приближаться к функциональной, т. е. самой тесной линейной зависимости. Коэффициент корреляции характеризует степень тесноты линейной зависимости между случайными величинами. Если случайные величины X и Y связаны точной линейной функциональной зависимостью:

$$Y = aX + b,$$

то $r_{xy} = \pm 1$, причем знак «плюс» или «минус» берется в зависимости от того, положителен или отрицателен коэффициент a. В общем случае, когда величины X и Y связаны произвольной вероятностной зависимостью, коэффициент корреляции может иметь значение в пределах:

$$-1 < r_{xy} < 1\ ^{1}).$$

---

[1]) Доказательство этих положений будет дано ниже, в п° 10.3, после того как мы познакомимся с некоторыми теоремами теории вероятностей, которые позволят провести его очень просто.

В случае $r_{xy} > 0$ говорят о *положительной корреляции* величин X и Y, в случае $r_{xy} < 0$ — об *отрицательной корреляции*. Положительная корреляция между случайными величинами означает, что при возрастании одной из них другая имеет тенденцию в среднем возрастать; отрицательная корреляция означает, что при возрастании одной из случайных величин другая имеет тенденцию в среднем убывать.

В рассмотренном примере двух случайных величин (X, Y), распределенных внутри круга с равномерной плотностью, несмотря на наличие зависимости между X и Y, линейная зависимость отсутствует; при возрастании X меняется только диапазон изменения Y,

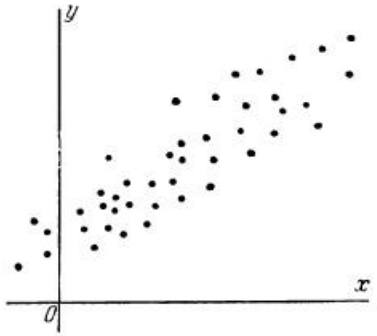


Рис. 8.6.2.

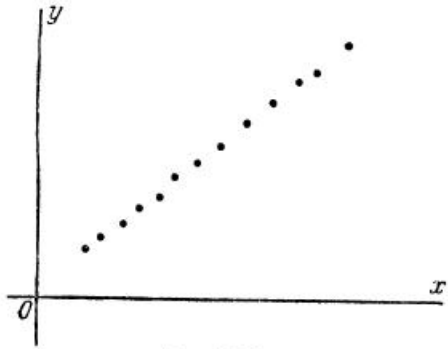


Рис. 8.6.3.

а его среднее значение не меняется; естественно, величины (X, Y) оказываются некоррелированными.

Приведем несколько примеров случайных величин с положительной и отрицательной корреляцией.

1. Вес и рост человека связаны положительной корреляцией.

2. Время, потраченное на регулировку прибора при подготовке его к работе, и время его безотказной работы связаны положительной корреляцией (если, разумеется, время потрачено разумно). Наоборот, время, потраченное на подготовку, и количество неисправностей, обнаруженное при работе прибора, связаны отрицательной корреляцией.

3. При стрельбе залпом координаты точек попадания отдельных снарядов связаны положительной корреляцией (так как имеются общие для всех выстрелов ошибки прицеливания, одинаково отклоняющие от цели каждый из них).

4. Производится два выстрела по цели; точка попадания первого выстрела регистрируется, и в прицел вводится поправка, пропорциональная ошибке первого выстрела с обратным знаком. Координаты

точек попадания первого и второго выстрелов будут связаны отрицательной корреляцией.

Если в нашем распоряжении имеются результаты ряда опытов над системой двух случайных величин (X, Y), то о наличии или отсутствии существенной корреляции между ними легко судить в первом приближении по графику, на котором изображены в виде точек все полученные из опыта пары значений случайных величин. Например, если наблюденные пары значений величин расположились так, как показано на рис. 8.6.2, то это указывает на наличие явно выраженной положительной корреляции между величинами. Еще более ярко выраженную положительную корреляцию, близкую к линейной функциональной зависимости, наблюдаем на рис. 8.6.3. На рис. 8.6.4 показан случай сравнительно слабой отрицательной корреляции. Наконец, на рис. 8.6.5 иллюстрируется случай практически

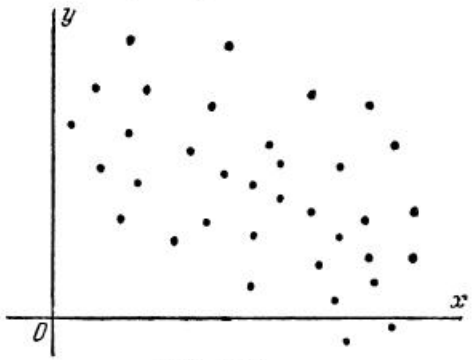


Рис. 8.6.4.

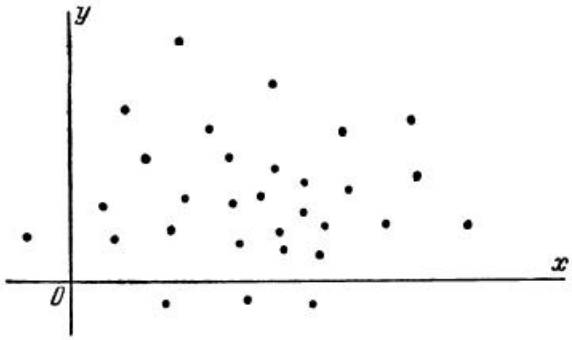


Рис. 8.6.5.

некоррелированных случайных величин. На практике, перед тем как исследовать корреляцию случайных величин, всегда полезно предварительно построить наблюденные пары значений на графике для первого качественного суждения о типе корреляции.

Способы определения характеристик системы случайных величин из опыта будут освещены в гл. 14.

8.7. Система произвольного числа случайных величин

На практике часто приходится рассматривать системы более чем двух случайных величин. Эти системы интерпретируются как случайные точки или случайные векторы в пространстве того или иного числа измерений.

Приведем примеры.

1. Точка разрыва дистанционного снаряда в пространстве характеризуется тремя декартовыми координатами (X, Y, Z) или тремя сферическими координатами (R, Φ, Θ).

2. Совокупность n последовательных измерений изменяющейся величины X — система n случайных величин $(X_1, X_2, \ldots, X_n)$.

3. Производится стрельба очередью из n снарядов. Совокупность координат n точек попадания на плоскости — система $2n$ случайных величин (абсцисс и ординат точек попадания):

$$(X_1, X_2, \ldots, X_n, \quad Y_1, Y_2, \ldots, Y_n).$$

4. Начальная скорость осколка — случайный вектор, характеризуемый тремя случайными величинами: величиной скорости V_0 и двумя углами Φ и Θ, определяющими направление полета осколка в сферической системе координат.

Полной характеристикой системы произвольного числа случайных величин служит *закон распределения* системы, который может быть задан функцией распределения или плотностью распределения.

Функцией распределения системы n случайных величин $(X_1, X_2, \ldots, X_n)$ называется вероятность совместного выполнения n неравенств вида $X_i < x_i$:

$$F(x_1, x_2, \ldots, x_n) = P\big((X_1 < x_1)(X_2 < x_2) \ldots (X_n < x_n)\big). \quad (8.7.1)$$

Плотностью распределения системы n непрерывных случайных величин называется n-я смешанная частная производная функции $F(x_1, x_2, \ldots, x_n)$, взятая один раз по каждому аргументу:

$$f(x_1, x_2, \ldots, x_n) = \frac{\partial^n F(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}. \quad (8.7.2)$$

Зная закон распределения системы, можно определить законы распределения отдельных величин, входящих в систему. Функция распределения каждой из величин, входящих в систему, получится, если в функции распределения системы положить все остальные аргументы равными ∞:

$$F_1(x_1) = F(x_1, \infty, \ldots, \infty). \quad (8.7.3)$$

Если выделить из системы величин $(X_1, X_2, \ldots, X_n)$ частную систему $(X_1, X_2, \ldots, X_k)$, то функция распределения этой системы определяется по формуле

$$F_{1,2,\ldots,k}(x_1, x_2, \ldots, x_k) = F(x_1, x_2, \ldots, x_k, \infty, \ldots, \infty). \quad (8.7.4)$$

Плотность распределения каждой из величин, входящих в систему, получится, если плотность распределения системы проинтегрировать в бесконечных пределах по всем остальным аргументам:

$$f_1(x_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n)\, dx_2 \dots dx_n. \qquad (8.7.5)$$

Плотность распределения частной системы $(X_1, X_2, \dots, X_k)$, выделенной из системы $(X_1, X_2, \dots, X_n)$, равна:

$$f_{1,2,\dots,k}(x_1, x_2, \dots, x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n)\, dx_{k+1} \dots dx_n. \qquad (8.7.6)$$

Условным законом распределения частной системы $(X_1, X_2, \dots, X_k)$ называется ее закон распределения, вычисленный при условии, что остальные величины $X_{k+1}, \dots, X_n$ приняли значения $x_{k+1}, \dots, x_n$.

Условная плотность распределения может быть вычислена по формуле

$$f(x_1, \dots, x_k \mid x_{k+1}, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n)}{f_{k+1,\dots,n}(x_{k+1}, \dots, x_n)}. \qquad (8.7.7)$$

Случайные величины $(X_1, X_2, \dots, X_n)$ называются *независимыми*, если закон распределения каждой частной системы, выделенной из системы $(X_1, X_2, \dots, X_n)$, не зависит от того, какие значения приняли остальные случайные величины.

Плотность распределения системы независимых случайных величин равна произведению плотностей распределения отдельных величин, входящих в систему:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n). \qquad (8.7.8)$$

Вероятность попадания случайной точки $(X_1, X_2, \dots, X_n)$ в пределы n-мерной области D выражается n-кратным интегралом:

$$P((X_1, X_2, \dots, X_n) \subset D) = \int \dots \int_{(D)} f(x_1, x_2, \dots, x_n)\, dx_1 dx_2 \dots dx_n. \qquad (8.7.9)$$

Формула (8.7.9) по существу является основной формулой для вычисления вероятностей событий, не сводящихся к схеме случаев. Действительно, если интересующее нас событие A не сводится к схеме случаев, то его вероятность не может быть вычислена непосредственно. Если при этом нет возможности поставить достаточное число однородных опытов и приближенно определить вероятность события A по его частоте, то типичная схема вычисления вероятности

события сводится к следующему. Переходят от схемы событий к схеме случайных величин (чаще всего — непрерывных) и сводят событие A к событию, состоящему в том, что система случайных величин $(X_1, X_2, \ldots, X_n)$ окажется в пределах некоторой области D. Тогда вероятность события A может быть вычислена по формуле (8.7.9).

Пример 1. Самолет поражается дистанционным снарядом при условии, если разрыв снаряда произошел не далее чем на расстоянии R от самолета (точнее, от условной точки на оси самолета, принимаемой за его центр). Закон распределения точек разрыва дистанционного снаряда в системе координат, связанной с целью, имеет плотность $f(x, y, z)$. Определить вероятность поражения самолета.

Решение. Обозначая поражение самолета буквой A, имеем:

$$P(A) = \iiint\limits_{(C)} f(x, y, z)\, dx\, dy\, dz,$$

где интегрирование распространяется по шару C радиуса R с центром в начале координат.

Пример 2. Метеорит, встретившийся на пути искусственного спутника Земли, пробивает его оболочку, если: 1) угол θ, под которым метеорит встречается с поверхностью спутника, заключен в определенных пределах (θ_1, θ_2); 2) метеорит имеет вес не менее q_0 $(г)$ и 3) относительная скорость встречи метеорита со спутником не меньше v_0 $(м/сек)$. Скорость встречи v, вес метеорита q и угол встречи θ представляют собой систему случайных величин с плотностью распределения $f(v, q, \theta)$. Найти вероятность p того, что отдельный метеорит, попавший в спутник, пробьет его оболочку.

Решение. Интегрируя плотность распределения $f(v, q, \theta)$ по трехмерной области, соответствующей пробиванию оболочки, получим:

$$p = \int\limits_{v_0}^{v_{\max}} \int\limits_{q_0}^{q_{\max}} \int\limits_{\theta_1}^{\theta_2} f(v, q, \theta)\, dv\, dq\, d\theta,$$

где $q_{\max}$ — максимальный вес метеорита, $v_{\max}$ — максимальная скорость встречи.

8.8. Числовые характеристики системы нескольких случайных величин

Закон распределения системы (заданный функцией распределения или плотностью распределения) является полной, исчерпывающей характеристикой системы нескольких случайных величин. Однако очень часто такая исчерпывающая характеристика не может быть применена. Иногда ограниченность экспериментального материала не дает возможности построить закон распределения системы. В других случаях исследование вопроса с помощью сравнительно громоздкого аппарата законов распределения не оправдывает себя в связи с невысокими требованиями к точности результата. Наконец, в ряде

задач примерный тип закона распределения (нормальный закон) известен заранее и требуется только найти его характеристики.

Во всех таких случаях вместо законов распределения применяют неполное, приближенное описание системы случайных величин с помощью минимального количества числовых характеристик.

Минимальное число характеристик, с помощью которых может быть охарактеризована система n случайных величин $X_1, X_2, \ldots, X_n$, сводится к следующему:

1) n математических ожиданий

$$m_1, m_2, \ldots, m_n,$$

характеризующих средние значения величин;

2) n дисперсий

$$D_1, D_2, \ldots, D_n,$$

характеризующих их рассеивание;

3) $n(n-1)$ корреляционных моментов

$$K_{ij} = M[\mathring{X}_i \mathring{X}_j] \quad (i \neq j),$$

где

$$\mathring{X}_i = X_i - m_i; \quad \mathring{X}_j = X_j - m_j,$$

характеризующих попарную корреляцию всех величин, входящих в систему.

Заметим, что дисперсия каждой из случайных величин X_i есть, по существу, не что иное, как частный случай корреляционного момента, а именно корреляционный момент величины X_i и той же величины X_i:

$$D_i = K_{ii} = M[\mathring{X}_i^2] = M[\mathring{X}_i \mathring{X}_i].$$

Все корреляционные моменты и дисперсии удобно расположить в виде прямоугольной таблицы (так называемой *матрицы*):

$$\begin{Vmatrix} K_{11} & K_{12} & \ldots & K_{1n} \\ K_{21} & K_{22} & \ldots & K_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ K_{n1} & K_{n2} & \ldots & K_{nn} \end{Vmatrix}.$$

Эта таблица называется *корреляционной матрицей* системы случайных величин $(X_1, X_2, \ldots, X_n)$.

Очевидно, что не все члены корреляционной матрицы различны. Из определения корреляционного момента ясно, что $K_{ij} = K_{ji}$, т. е. *элементы корреляционной матрицы, расположенные симметрично по отношению к главной диагонали, равны.* В связи

с этим часто заполняется не вся корреляционная матрица, а лишь ее половина, считая от главной диагонали:

$$\begin{Vmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ & K_{22} & \dots & K_{2n} \\ & & \dots & \\ & & & K_{nn} \end{Vmatrix}.$$

Корреляционную матрицу, составленную из элементов K_{ij}, часто сокращенно обозначают символом $\|K_{ij}\|$.

По главной диагонали корреляционной матрицы стоят дисперсии случайных величин $X_1, X_2, \dots, X_n$.

В случае, когда случайные величины $X_1, X_2, \dots, X_n$ не коррелированы, все элементы корреляционной матрицы, кроме диагональных, равны нулю:

$$\|K_{ij}\| = \begin{Vmatrix} D_1 & 0 & 0 \dots 0 \\ & D_2 & 0 \dots 0 \\ & & D_3 \dots 0 \\ & & \dots \\ & & D_n \end{Vmatrix}.$$

Такая матрица называется *диагональной*.

В целях наглядности суждения именно о коррелированности случайных величин безотносительно к их рассеиванию часто вместо корреляционной матрицы $\|K_{ij}\|$ пользуются *нормированной корреляционной матрицей* $\|r_{ij}\|$, составленной не из корреляционных моментов, а из коэффициентов корреляции:

$$r_{ij} = \frac{K_{ij}}{\sigma_i \sigma_j},$$

где

$$\sigma_i = \sqrt{D_i}, \quad \sigma_j = \sqrt{D_j}.$$

Все диагональные элементы этой матрицы, естественно, равны единице. Нормированная корреляционная матрица имеет вид:

$$\|r_{ij}\| = \begin{Vmatrix} 1 & r_{12} & r_{13} \dots & r_{1n} \\ & 1 & r_{23} \dots & r_{2n} \\ & & 1 \dots & r_{3n} \\ & & \dots & \\ & & & 1 \end{Vmatrix}.$$

Введем понятие о некоррелированных системах случайных величин (иначе — *о некоррелированных случайных векторах*). Рассмотрим две системы случайных величин:

$$(X_1, X_2, \ldots, X_n); \quad (Y_1, Y_2, \ldots, Y_n)$$

или два случайных вектора в n-мерном пространстве: $\vec{X}$ с составляющими $(X_1, X_2, \ldots, X_n)$ и $\vec{Y}$ с составляющими $(Y_1, Y_2, \ldots, Y_n)$. Случайные векторы $\vec{X}$ и $\vec{Y}$ называются *некоррелированными*, если каждая из составляющих вектора $\vec{X}$ не коррелирована с каждой из составляющих вектора $\vec{Y}$:

$$K_{x_i y_j} = M[\mathring{X}_i \mathring{Y}_j] = 0 \quad \text{при} \quad i = 1, \ldots, n; \quad j = 1, \ldots, n.$$

ГЛАВА 9

НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ ДЛЯ СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

9.1. Нормальный закон на плоскости

Из законов распределения системы двух случайных величин имеет смысл специально рассмотреть нормальный закон, как имеющий наибольшее распространение на практике. Так как система двух случайных величин изображается случайной точкой на плоскости, нормальный закон для системы двух величин часто называют «нормальным законом на плоскости».

В общем случае плотность нормального распределения двух случайных величин выражается формулой

$$f(x,\ y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\, e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r\,(x-m_x)\,(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]}. \tag{9.1.1}$$

Этот закон зависит от пяти параметров: m_x, m_y, σ_x, σ_y и r. Смысл этих параметров нетрудно установить. Докажем, что параметры m_x, m_y представляют собой математические ожидания (центры рассеивания) величин X и Y; σ_x, σ_y — их средние квадратические отклонения; r — коэффициент корреляции величин X и Y.

Для того чтобы убедиться в этом, найдем прежде всего плотность распределения для каждой из величин, входящих в систему. Согласно формуле (8.4.2)

$$f_1(x) = \int_{-\infty}^{\infty} f(x,\ y)\,dy =$$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r\,(x-m_x)\,(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]} dy.$$

Вычислим интеграл

$$I = \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]} dy.$$

Положим:

$$\frac{x-m_x}{\sigma_x\sqrt{2}} = u, \quad \frac{y-m_y}{\sigma_y\sqrt{2}} = v; \tag{9.1.2}$$

тогда

$$I = \sigma_y\sqrt{2}\int_{-\infty}^{\infty} e^{-\frac{1}{1-r^2}[u^2-2ruv+v^2]}\,dv.$$

Из интегрального исчисления известно, что

$$\int_{-\infty}^{\infty} e^{-Ax^2\pm 2Bx+C}\,dx = \sqrt{\frac{\pi}{A}}\,e^{-\frac{AC-B^2}{A}}\,^{1}). \tag{9.1.3}$$

В нашем случае

$$A = \frac{1}{1-r^2}; \quad B = \frac{ru}{1-r^2}; \quad C = \frac{u^2}{1-r^2}.$$

Подставляя эти значения в формулу (9.1.3), имеем:

$$I = \sigma_y\sqrt{2}\sqrt{\pi(1-r^2)}\,e^{-u^2},$$

откуда

$$f_1(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\,e^{-u^2},$$

или, учитывая (9.1.2),

$$f_1(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\,e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} \tag{9.1.4}$$

Таким образом, величина X подчинена нормальному закону с центром рассеивания m_x и средним квадратическим отклонением σ_x.

Аналогично покажем, что

$$f_2(y) = \frac{1}{\sigma_y\sqrt{2\pi}}\,e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}, \tag{9.1.5}$$

т. е. величина Y подчинена нормальному закону с центром рассеивания m_y и средним квадратическим отклонением σ_y.

---

[1]) Для вычисления интеграла (9.1.3) достаточно дополнить показатель степени до полного квадрата и после замены переменной воспользоваться интегралом Эйлера — Пуассона (6.1.3).

Остается доказать, что параметр r в формуле (9.1.1) представляет собой коэффициент корреляции величин X и Y. Для этого вычислим корреляционный момент:

$$K_{xy} = \int_{-\infty}^{\infty}\int (x - m_x)(y - m_y) f(x, y)\, dx\, dy,$$

где m_x, m_y — математические ожидания величин X и Y.

Подставляя в эту формулу выражение $f(x, y)$, получим:

$$K_{xy} = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \int_{-\infty}^{\infty}\int (x - m_x)(y - m_y)\, e^{-A(x, y)}\, dx\, dy, \qquad (9.1.6)$$

где

$$A(x, y) = -\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right].$$

Произведем в двойном интеграле (9.1.6) замену переменных, положив:

$$\frac{x - m_x}{\sigma_x\sqrt{2}} = u; \quad \frac{1}{\sqrt{2(1-r^2)}}\left(\frac{y - m_y}{\sigma_y} - r\frac{x - m_x}{\sigma_x}\right) = w. \qquad (9.1.7)$$

Якобиан преобразования равен

$$2\sigma_x\sigma_y\sqrt{1-r^2},$$

следовательно,

$$K_{xy} = \frac{1}{\pi}\int_{-\infty}^{\infty}\int (u\sigma_x\sqrt{2})\,\sigma_y\sqrt{2(1-r^2)}\left(w + \frac{ru}{\sqrt{1-r^2}}\right) e^{-u^2-w^2}\, du\, dw =$$

$$= \frac{2\sigma_x\sigma_y\sqrt{1-r^2}}{\pi}\int_{-\infty}^{\infty} ue^{-u^2}\, du \int_{-\infty}^{\infty} we^{-w^2}\, dw +$$

$$+ \frac{2\sigma_x\sigma_y r}{\pi}\int_{-\infty}^{\infty} u^2 e^{-u^2}\, du \int_{-\infty}^{\infty} e^{-w^2}\, dw.$$

Учитывая, что

$$\int_{-\infty}^{\infty} ue^{-u^2}\, du = \int_{-\infty}^{\infty} we^{-w^2}\, dw = 0; \quad \int_{-\infty}^{\infty} u^2 e^{-u^2}\, du = \frac{\sqrt{\pi}}{2};$$

$$\int_{-\infty}^{\infty} e^{-w^2}\, dw = \sqrt{\pi},$$

имеем:

$$K_{xy} = r\sigma_x\sigma_y; \quad r = \frac{K_{xy}}{\sigma_x\sigma_y}. \tag{9.1.8}$$

Таким образом, доказано, что параметр r в формуле (9.1.1) представляет собой коэффициент корреляции величин X и Y.

Предположим теперь, что случайные величины X и Y, подчиненные нормальному закону на плоскости, не коррелированы; положим в формуле (9.1.1) $r = 0$. Получим:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-m_x)^2}{2\sigma_x^2} - \frac{(y-m_y)^2}{2\sigma_y^2}}. \tag{9.1.9}$$

Легко убедиться, что случайные величины (X, Y), подчиненные закону распределения с плотностью (9.1.9), не только не коррелированы, но и независимы. Действительно,

$$f(x, y) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}} = f_1(x) f_2(y),$$

т. е. плотность распределения системы равна произведению плотностей распределения отдельных величин, входящих в систему, а это значит, что случайные величины (X, Y) независимы.

Таким образом, для системы случайных величин, подчиненных нормальному закону, из некоррелированности величин вытекает также их независимость. Термины «некоррелированные» и «независимые» величины для случая нормального распределения эквивалентны.

При $r \neq 0$ случайные величины (X, Y) зависимы. Нетрудно убедиться, вычисляя условные законы распределения по формулам (8.4.6), что

$$f(y \mid x) = \frac{1}{\sigma_y\sqrt{1-r^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r^2)}\left(\frac{y-m_y}{\sigma_y} - r\frac{x-m_x}{\sigma_x}\right)^2},$$

$$f(x \mid y) = \frac{1}{\sigma_x\sqrt{1-r^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r^2)}\left(\frac{x-m_x}{\sigma_x} - r\frac{y-m_y}{\sigma_y}\right)^2}.$$

Проанализируем один из этих условных законов распределения, например $f(y \mid x)$. Для этого преобразуем выражение плотности $f(y \mid x)$ к виду:

$$f(y \mid x) = \frac{1}{\sigma_y\sqrt{1-r^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r^2)\sigma_y^2}\left[y - m_y - r\frac{\sigma_y}{\sigma_x}(x-m_x)\right]^2}.$$

Очевидно, это есть плотность нормального закона с центром рассеивания

$$m_{y|x} = m_y + r\frac{\sigma_y}{\sigma_x}(x - m_x) \tag{9.1.10}$$

и средним квадратическим отклонением

$$\sigma_{y|x} = \sigma_y\sqrt{1 - r^2}. \tag{9.1.11}$$

Формулы (9.1.10) и (9.1.11) показывают, что в условном законе распределения величины Y при фиксированном значении $X = x$ от этого значения зависит только математическое ожидание, но не дисперсия.

Величина $m_{y|x}$ называется *условным математическим ожиданием* величины Y при данном x. Зависимость (9.1.10) можно изобразить на плоскости xOy, откладывая условное математическое ожидание $m_{y|x}$ по оси ординат. Получится прямая, которая называется *линией регрессии* Y на X. Аналогично прямая

$$x = m_x + r\frac{\sigma_x}{\sigma_y}(y - m_y) \tag{9.1.12}$$

есть линия регрессии X на Y.

Линии регрессии совпадают только при наличии линейной функциональной зависимости Y от X. При независимых X и Y линии регрессии параллельны координатным осям.

Рассматривая выражение (9.1.1) для плотности нормального распределения на плоскости, мы видим, что нормальный закон на плоскости полностью определяется заданием пяти параметров: двух координат центра рассеивания m_x, m_y, двух средних квадратических отклонений σ_x, σ_y и одного коэффициента корреляции r. В свою очередь последние три параметра σ_x, σ_y и r полностью определяются элементами корреляционной матрицы: дисперсиями D_x, D_y и корреляционным моментом K_{xy}. Таким образом, минимальное количество числовых характеристик системы — математические ожидания, дисперсии и корреляционный момент — в случае, когда система подчинена нормальному закону, определяет собой полностью закон распределения, т. е. образует *исчерпывающую* систему характеристик.

Так как на практике нормальный закон весьма распространен, то очень часто для полной характеристики закона распределения системы оказывается достаточно задать минимальное число — всего пять — числовых характеристик.

9.2. Эллипсы рассеивания. Приведение нормального закона к каноническому виду

Рассмотрим поверхность распределения, изображающую функцию (9.1.1). Она имеет вид холма, вершина которого находится над точкой (m_x, m_y) (рис. 9.2.1).

В сечении поверхности распределения плоскостями, параллельными оси $f(x, y)$, получаются кривые, подобные нормальным кривым распределения. В сечении поверхности распределения плоскостями, параллельными плоскости xOy, получаются эллипсы. Напишем уравнение проекции такого эллипса на плоскость xOy:

$$\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} = \text{const},$$

или, обозначая константу λ^2,

$$\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} = \lambda^2. \quad (9.2.1)$$

Уравнение эллипса (9.2.1) можно проанализировать обычными методами аналитической геометрии. Применяя их, убеждаемся, что центр эллипса (9.2.1) находится в точке с координатами (m_x, m_y); что касается направления осей симметрии эллипса, то они составляют с осью Ox углы, определяемые уравнением

$$\operatorname{tg} 2\alpha = \frac{2r\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}\ ^{1}). \quad (9.2.2)$$

Это уравнение дает два значения углов: α и α_1, различающиеся на $\frac{\pi}{2}$.

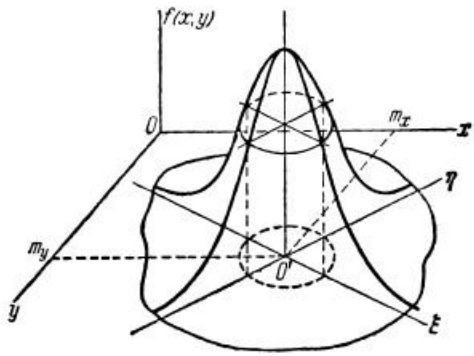


Рис. 9.2.1.

Таким образом, ориентация эллипса (9.2.1) относительно координатных осей находится в прямой зависимости от коэффициента корреляции r системы (X, Y); если величины не коррелированы (т. е. в данном случае и независимы), то оси симметрии эллипса параллельны координатным осям; в противном случае они составляют с координатными осями некоторый угол.

---

[1]) Обоснование формулы (9.2.2) другим способом см. в п° 14.7.

Пересекая поверхность распределения плоскостями, параллельными плоскости xOy, и проектируя сечения на плоскость xOy, мы получим целое семейство подобных и одинаково расположенных эллипсов с общим центром (m_x, m_y). Во всех точках каждого из таких эллипсов плотность распределения $f(x, y)$ постоянна. Поэтому такие эллипсы называются *эллипсами равной плотности* или, короче, *эллипсами рассеивания*. Общие оси всех эллипсов рассеивания называются *главными осями рассеивания*.

Известно, что уравнение эллипса принимает наиболее простой, так называемый «канонический» вид, если координатные оси совпадают с осями симметрии эллипса. Для того чтобы привести уравнение эллипса рассеивания к каноническому виду, достаточно перенести начало координат в точку (m_x, m_y) и повернуть координатные оси на угол α, определяемый уравнением (9.2.2). При этом координатные оси совпадут с главными осями рассеивания, и нормальный закон на плоскости преобразуется к так называемому «каноническому» виду.

Каноническая форма нормального закона на плоскости имеет вид:

$$f(\xi, \eta) = \frac{1}{2\pi\sigma_\xi\sigma_\eta} e^{-\frac{\xi^2}{2\sigma_\xi^2}-\frac{\eta^2}{2\sigma_\eta^2}}, \qquad (9.2.3)$$

где σ_ξ, σ_η — так называемые *главные средние квадратические отклонения*, т. е. средние квадратические отклонения случайных величин (Ξ, H), представляющих собой координаты случайной точки в системе координат, определяемой главными осями рассеивания $O\xi$, $O\eta$. Главные средние квадратические отклонения σ_ξ и σ_η выражаются через средние квадратические отклонения в прежней системе координат формулами:

$$\left.\begin{aligned} \sigma_\xi^2 &= \sigma_x^2\cos^2\alpha + r\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2\sin^2\alpha, \\ \sigma_\eta^2 &= \sigma_x^2\sin^2\alpha - r\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2\cos^2\alpha. \end{aligned}\right\}^{1)} \qquad (9.2.4)$$

Обычно, рассматривая нормальный закон на плоскости, стараются заранее выбрать координатные оси Ox, Oy так, чтобы они совпали с главными осями рассеивания. При этом средние квадратические отклонения по осям σ_x, σ_y и будут главными средними квадратическими отклонениями, и нормальный закон будет иметь вид:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}}. \qquad (9.2.5)$$

В некоторых случаях координатные оси выбирают параллельно главным осям рассеивания, но начало координат с центром рассеива-

---

[1]) Вывод этих формул см. в главе 14, п° 14.7.

ния не совмещают. При этом случайные величины (X, Y) также оказываются независимыми, но выражение нормального закона имеет вид:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-m_x)^2}{2\sigma_x^2} - \frac{(y-m_y)^2}{2\sigma_y^2}}, \qquad (9.2.6)$$

где m_x, m_y — координаты центра рассеивания.

Перейдем в канонической форме нормального закона (9.2.5) от средних квадратических отклонений к вероятным отклонениям:

$$E_x = \rho\sqrt{2}\,\sigma_x; \quad E_y = \rho\sqrt{2}\,\sigma_y.$$

Величины E_x, E_y называются *главными вероятными отклонениями*. Подставляя выражения σ_x, σ_y через E_x, E_y в уравнение (9.2.5), получим другую каноническую форму нормального закона:

$$f(x, y) = \frac{\rho^2}{\pi E_x E_y} e^{-\rho^2\left(\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2}\right)}. \qquad (9.2.7)$$

В такой форме нормальный закон часто применяется в теории стрельбы.

Напишем уравнение эллипса рассеивания в каноническом виде:

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} = k^2 \quad \text{или} \quad \frac{x^2}{(k\sigma_x)^2} + \frac{y^2}{(k\sigma_y)^2} = 1, \qquad (9.2.8)$$

где k — постоянное число.

Из уравнения видно, что полуоси эллипса рассеивания пропорциональны главным средним квадратическим отклонениям (а значит, и главным вероятным отклонениям).

Назовем «единичным» эллипсом рассеивания тот из эллипсов равной плотности вероятности, полуоси которого равны главным средним квадратическим отклонениям σ_x, σ_y. (Если пользоваться в качестве характеристик рассеивания не главными средними квадратическими, а главными вероятными отклонениями, то естественно будет назвать «единичным» тот эллипс, полуоси которого равны E_x, E_y.)

Кроме единичного эллипса рассеивания иногда рассматривают еще «полный» эллипс рассеивания, под которым понимают тот из эллипсов равной плотности вероятности, в который с практической достоверностью укладывается все рассеивание. Размеры этого эллипса, разумеется, зависят от того, что понимать под «практической достоверностью». В частности, если принять за «практическую достоверность» вероятность порядка 0,99, то «полным эллипсом рассеивания» можно считать эллипс с полуосями $3\sigma_x$, $3\sigma_y$.

Рассмотрим специально один частный случай, когда главные средние квадратические отклонения равны друг другу:

$$\sigma_x = \sigma_y = \sigma.$$

Тогда все эллипсы рассеивания обращаются в круги, и рассеивание называется *круговым*. При круговом рассеивании каждая из осей, проходящих через центр рассеивания, может быть принята за главную ось рассеивания, или, другими словами, направление главных осей рассеивания неопределенно. При некруговом рассеивании случайные величины X, Y, подчиненные нормальному закону на плоскости, независимы тогда и только тогда, когда координатные оси параллельны главным осям рассеивания; при круговом рассеивании случайные величины (X, Y) независимы при любом выборе прямоугольной системы координат. Эта особенность кругового рассеивания приводит к тому, что оперировать с круговым рассеиванием гораздо удобнее, чем с эллиптическим. Поэтому на практике, где только возможно, стремятся приближенно заменять некруговое рассеивание круговым.

9.3. Вероятность попадания в прямоугольник со сторонами, параллельными главным осям рассеивания

Пусть случайная точка (X, Y) на плоскости подчинена нормальному закону

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-m_x)^2}{2\sigma_x^2} - \frac{(y-m_y)^2}{2\sigma_y^2}} \tag{9.3.1}$$

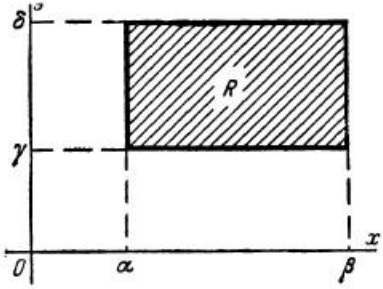


Рис. 9.3.1.

При этом главные оси рассеивания параллельны координатным осям и величины X и Y независимы.

Требуется вычислить вероятность попадания случайной точки (X, Y) в прямоугольник R, стороны которого параллельны координатным осям xOy, а следовательно и главным осям рассеивания (рис. 9.3.1). Согласно общей формуле (8.3.4) имеем:

$$P((X, Y) \subset R) = \int_\alpha^\beta \int_\gamma^\delta f(x, y)\, dx\, dy =$$

$$= \int_\alpha^\beta \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} dx \int_\gamma^\delta \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}} dy,$$

откуда, применяя формулу (6.3.3) для вероятности попадания на участок, находим:

$$P((X, Y) \subset R) = \left[\Phi^*\left(\frac{\beta - m_x}{\sigma_x}\right) - \Phi^*\left(\frac{\alpha - m_x}{\sigma_x}\right)\right]\left[\Phi^*\left(\frac{\delta - m_y}{\sigma_y}\right) - \Phi^*\left(\frac{\gamma - m_y}{\sigma_y}\right)\right], \quad (9.3.2)$$

где $\Phi^*(x)$ — нормальная функция распределения.

Если нормальный закон на плоскости дан в канонической форме, то $m_x = m_y = 0$, и формула (9.3.2) принимает вид

$$P((X, Y) \subset R) = \left[\Phi^*\left(\frac{\beta}{\sigma_x}\right) - \Phi^*\left(\frac{\alpha}{\sigma_x}\right)\right]\left[\Phi^*\left(\frac{\delta}{\sigma_y}\right) - \Phi^*\left(\frac{\gamma}{\sigma_y}\right)\right]. \quad (9.3.3)$$

Если стороны прямоугольника не параллельны координатным осям, то формулы (9.3.2) и (9.3.3) уже неприменимы. Только при круговом рассеивании вероятность попадания в прямоугольник любой ориентации вычисляется по формулам (9.3.2) или (9.3.3).

Формулы (9.3.2) и (9.3.3) широко применяются при вычислении вероятностей попадания в цели: прямоугольные, близкие к прямоугольным, составленные из прямоугольников или приближенно заменяемые таковыми.

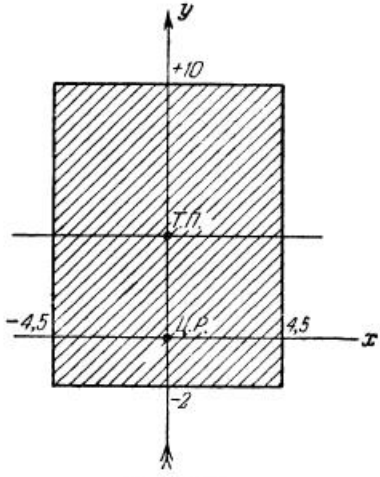


Рис. 9.3.2.

Пример. Производится стрельба с самолета по прямоугольному щиту размером 9 $м \times 12$ $м$, лежащему на земле горизонтально. Главные вероятные отклонения: в продольном направлении $B_д = 10$ $м$, в боковом направлении $B_б = 5$ $м$. Прицеливание — по центру мишени, заход — вдоль мишени. Вследствие несовпадения дальности пристрелки и дальности фактической стрельбы средняя точка попадания смещается в сторону недолета на 4 $м$. Найти вероятность попадания в мишень при одном выстреле.

Решение. На чертеже (рис. 9.3.2) наносим мишень, точку прицеливания (т. п.) и центр рассеивания (ц. р.). Через ц. р. проводим главные оси рассеивания: по направлению полета и перпендикулярно к нему.

Перейдем от главных вероятных отклонений $B_д$ и $B_б$ к главным средним квадратическим:

$$\sigma_x = \frac{B_б}{0{,}674} \approx 7{,}4 \text{ м}; \qquad \sigma_y = \frac{B_д}{0{,}674} \approx 14{,}8 \text{ м}.$$

По формуле (9.3.3) имеем:

$$P((X, Y) \subset R) = \left[\Phi^*\left(\frac{4{,}5}{7{,}4}\right) - \Phi^*\left(\frac{-4{,}5}{7{,}4}\right)\right]\left[\Phi^*\left(\frac{10}{14{,}8}\right) - \Phi^*\left(\frac{-2}{14{,}8}\right)\right] \approx 0{,}138.$$

9.4. Вероятность попадания в эллипс рассеивания

К числу немногих плоских фигур, вероятность попадания в которые может быть вычислена в конечном виде, принадлежит эллипс рассеивания (эллипс равной плотности).

Пусть нормальный закон на плоскости задан в канонической форме:

$$f(x,\ y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}. \qquad (9.4.1)$$

Рассмотрим эллипс рассеивания B_k, уравнение которого

$$\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}=k^2,$$

где параметр k представляет собой отношение полуосей эллипса рассеивания к главным средним квадратическим отклонениям. По общей формуле (8.3.3) имеем:

$$P\big((X,\ Y)\subset B_k\big)=\int\limits_{(B_k)}\int f(x,\ y)\,dx\,dy=$$

$$=\int\limits_{(B_k)}\int\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}dx\,dy. \qquad (9.4.2)$$

Сделаем в интеграле (9.4.2) замену переменных

$$\frac{x}{\sigma_x}=u;\qquad \frac{y}{\sigma_y}=v.$$

Этой подстановкой эллипс B_k преобразуется в круг C_k радиуса k. Следовательно,

$$P\big((X,\ Y)\subset B_k\big)=\frac{1}{2\pi}\int\limits_{(C_k)}\int e^{-\frac{u^2}{2}-\frac{v^2}{2}}du\,dv. \qquad (9.4.3)$$

Перейдем в интеграле (9.4.3) от декартовой системы координат к полярной, положив

$$u=r\cos\theta;\qquad v=r\sin\theta. \qquad (9.4.4)$$

Якобиан преобразования (9.4.4) равен r. Производя замену переменных, получим:

$$P\big((X,\ Y)\subset B_k\big)=\frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\int\limits_{0}^{k}re^{-\frac{r^2}{2}}dr\,d\theta=\int\limits_{0}^{k}re^{-\frac{r^2}{2}}=1-e^{-\frac{k^2}{2}}.$$

Таким образом, вероятность попадания случайной точки в эллипс рассеивания, полуоси которого равны k средним квадратическим отклонениям, равна:

$$P\big((X, Y) \subset B_k\big) = 1 - e^{-\frac{k^2}{2}}. \qquad (9.4.5)$$

В качестве примера найдем вероятность попадания случайной точки, распределенной по нормальному закону на плоскости xOy, в единичный эллипс рассеивания, полуоси которого равны средним квадратическим отклонениям:

$$a = \sigma_x; \quad b = \sigma_y.$$

Для такого эллипса $k = 1$. Имеем:

$$P\big((X, Y) \subset B_1\big) = 1 - e^{-\frac{1}{2}}.$$

Пользуясь таблицей 2 приложения, находим:

$$P\big((X, Y) \subset B_1\big) \approx 0{,}393.$$

Формула (9.4.5) чаще всего применяется для вычисления вероятности попадания в круг при круговом рассеивании.

Пример. На пути быстро движущейся малоразмерной цели площади 1,2 $м$ ставится осколочное поле в форме плоского диска радиуса $R = 30$ $м$. Внутри диска плотность осколков постоянна и равна 2 оск./$м^2$. Если цель накрыта диском, то число осколков, попадающих в нее, можно считать распределенным по закону Пуассона. В силу малости цели можно рассматривать ее как точечную и считать, что она или полностью накрывается осколочным полем (если ее центр попадает в круг), или совсем не накрывается (если ее центр не попадает в круг). Попадание осколка гарантирует поражение цели. При прицеливании центр круга O_1 стремятся совместить в плоскости xOy с началом координат O (центром цели), но вследствие ошибок точка O_1 рассеивается около O (рис. 9.4.1). Закон рассеивания нормальный, рассеивание круговое, $\sigma = 20$ $м$. Определить вероятность поражения цели $P(A)$.

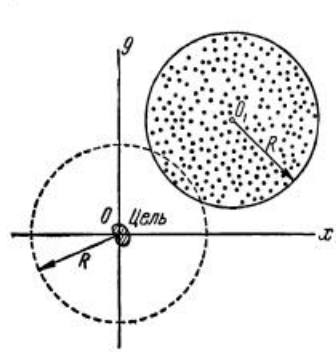


Рис. 9.4.1.

Решение. Чтобы цель была поражена осколками, необходимо совмещение двух событий: 1) попадание цели (точки O) в осколочное поле (круг радиуса R) и 2) поражение цели при условии, что попадание произошло.

Вероятность попадания цели в круг, очевидно, равна вероятности того, что центр круга (случайная точка O_1) попадет в круг радиуса R, описанный вокруг начала координат. Применим формулу (9.4.5). Имеем:

$$k = \frac{30}{20} = 1{,}5.$$

Вероятность попадания цели в осколочное поле равна:

$$p = 1 - e^{-\frac{1{,}5^2}{2}} \approx 0{,}675.$$

Далее найдем вероятность поражения цели p^* при условии, что она накрыта осколочным диском. Среднее число осколков a, попадающих в накрытую полем цель, равно произведению площади цели на плотность поля осколков:

$$1{,}2 \cdot 2 = 2{,}4.$$

Условная вероятность поражения цели p^* есть не что иное, как вероятность попадания в нее хотя бы одного осколка. Пользуясь формулой (5.9.5) главы 5, имеем:

$$p^* = R_1 = 1 - e^{-a} \approx 0{,}909.$$

Вероятность поражения цели равна:

$$P(A) = 0{,}675 \cdot 0{,}909 \approx 0{,}613.$$

Воспользуемся формулой (9.4.5) для вероятности попадания в круг, чтобы вывести одно важное для практики распределение: так называемое ***распределение Релея***.

Рассмотрим на плоскости xOy (рис. 9.4.2) случайную точку (X, Y), рассеивающуюся вокруг начала координат O по круговому нормальному закону со средним квадратическим отклонением σ. Найдем закон распределения случайной величины R — расстояния от точки (X, Y) до начала координат, т. е. длины случайного вектора с составляющими X, Y.

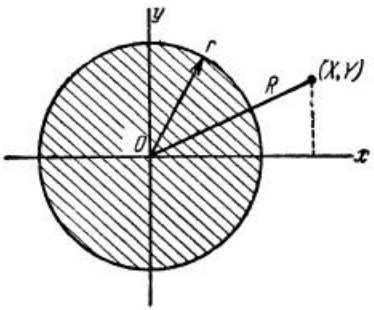


Рис. 9.4.2.

Найдем сначала функцию распределения $F(r)$ величины R. По определению

$$F(r) = P(R < r).$$

Это есть не что иное, как вероятность попадания случайной точки (X, Y) внутрь круга радиуса r (рис. 9.4.2). По формуле (9.4.5) эта вероятность равна:

$$F(r) = 1 - e^{-\frac{k^2}{2}},$$

где $k = \frac{r}{\sigma}$, т. е.

$$F(r) = 1 - e^{-\frac{r^2}{2\sigma^2}}. \qquad (9.4.6)$$

Данное выражение функции распределения имеет смысл только при положительных значениях r; при отрицательных r нужно положить $F(r) = 0$.

Дифференцируя функцию распределения $F(r)$ по r, найдем плотность распределения

$$f(r)=\begin{cases}\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}} & \text{при} \quad r>0,\\ 0 & \text{при} \quad r<0.\end{cases} \tag{9.4.7}$$

Закон Релея (9.4.7) встречается в разных областях практики: в стрельбе, радиотехнике, электротехнике и др.

График функции $f(r)$ (плотности закона Релея) приведен на рис. 9.4.3.

Найдем числовые характеристики величины R, распределенной по закону Релея, а именно: ее моду $\mathcal{M}$ и математическое ожидание m_r. Для того чтобы найти моду — абсциссу точки, в которой плотность вероятности максимальна, продифференцируем $f(r)$ и приравняем производную нулю:

$$1-\frac{r^2}{\sigma^2}=0; \qquad \sigma^2=r^2.$$

Корень этого уравнения и есть искомая мода

$$\mathcal{M}=\sigma. \tag{9.4.8}$$

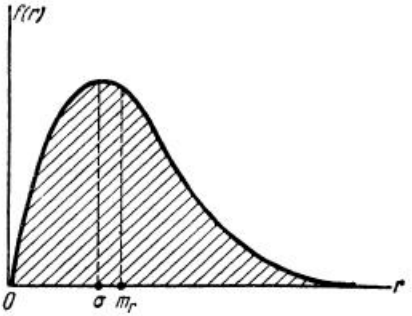


Рис. 9.4.3.

Таким образом, наивероятнейшее значение расстояния R случайной точки (X, Y) от начала координат равно среднему квадратическому отклонению рассеивания.

Математическое ожидание m_r найдем по формуле

$$m_r=\int_0^\infty rf(r)\,dr=\int_0^\infty \frac{r^2}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}\,dr.$$

Производя замену переменной

$$\frac{r}{\sigma\sqrt{2}}=t,$$

получим:

$$m_r=\sigma\sqrt{2}\int_0^\infty 2t^2e^{-t^2}\,dt=\sigma\sqrt{2}\int_0^\infty t\cdot 2te^{-t^2}\,dt.$$

Интегрируя по частям, найдем математическое ожидание расстояния R:

$$m_r=\sigma\sqrt{2}\,\frac{\sqrt{\pi}}{2}=\sigma\sqrt{\frac{\pi}{2}}\approx 1{,}25\sigma. \tag{9.4.9}$$

9.5. Вероятность попадания в область произвольной формы

При стрельбе ударными снарядами вычисление вероятности попадания в цель сводится к вычислению вероятности попадания случайной точки (X, Y) в некоторую область D. Пусть случайная точка (X, Y) подчинена нормальному закону в каноническом виде. Вероятность попадания точки (X, Y) в область D выражается интегралом

$$P((X, Y) \subset D) = \frac{1}{2\pi\sigma_x\sigma_y} \iint\limits_{(D)} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}} dx\, dy. \qquad (9.5.1)$$

В отдельных частных случаях (например, когда область D есть прямоугольник со сторонами, параллельными главным осям рассеивания, или эллипс рассеивания, а также в некоторых других, имеющих меньшее практическое значение) интеграл (9.5.1) может быть выражен через известные функции; в общем же случае этот интеграл через известные функции не выражается. На практике для вычисления вероятности попадания в область произвольной формы применяются следующие приближенные способы.

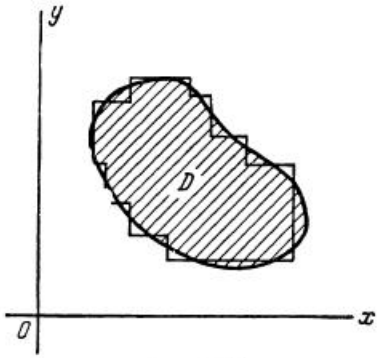


Рис. 9.5.1.

1. Область D приближенно заменяется областью, составленной из прямоугольников, стороны которых параллельны главным осям рассеивания (рис. 9.5.1). Вероятность попадания в каждый из таких прямоугольников вычисляется по формуле (9.3.3). Этот способ можно рекомендовать тогда, когда число прямоугольников, на которые приближенно разбивается цель D, не слишком велико.

2. Вся плоскость xOy с помощью некоторой системы линий (прямых или кривых) заранее разбивается на ряд ячеек, вероятности попадания в которые могут быть выражены точно через известные функции, и вычисляется вероятность попадания в каждую ячейку. Такая система линий с соответствующими ей вероятностями попадания в ячейки называется *сеткой рассеивания*. Работа с сеткой заключается в том, что изображение сетки накладывается на изображение цели, после чего производится суммирование вероятностей попадания в ячейки, накрытые целью; если цель накрывает часть ячейки, то берется часть вероятности попадания в ячейку, пропорциональная накрытой площади.

Сетку рассеивания можно применять двояким образом: а) строить цель в масштабе сетки, б) строить сетку в масштабе цели.

Если цель имеет сложные очертания и, особенно, если она сравнительно невелика, бывает обычно удобнее построить на изображении цели в том же масштабе ту часть сетки, которая занята целью. Если же цель имеет сравнительно простые очертания и довольно велика (занимает значительную часть полного эллипса рассеивания),

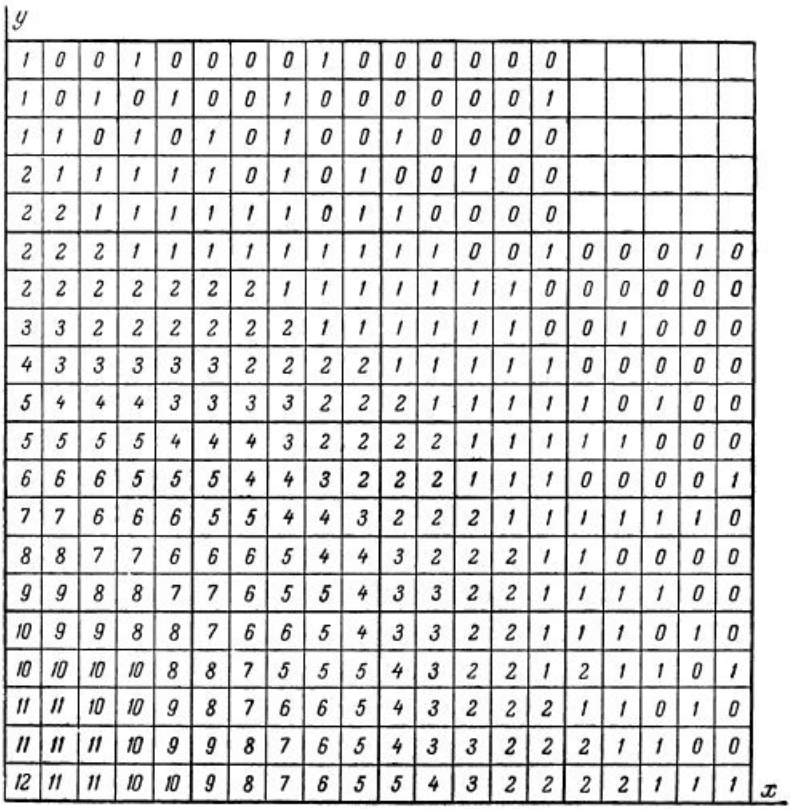


Рис. 9.5.2.

обычно удобнее построить цель в масштабе сетки. Так как стандартная сетка строится для кругового рассеивания, а на практике рассеивание в общем случае круговым не является, при построении цели в масштабе сетки приходится в общем случае пользоваться двумя разными масштабами по осям Ox и Oy. При этом способе удобно иметь в распоряжении сетку рассеивания, выполненную на прозрачной бумаге, и накладывать ее на перестроенное изображение цели. Прямолинейная сетка рассеивания для одного координатного угла дана на рис. 9.5.2. Сторона ячейки равна $0{,}2E \approx 0{,}133\sigma$.

В ячейках проставлены вероятности попадания в них, выраженные в сороковых долях процента.

3. В случае, когда размеры области D невелики по сравнению со средними квадратическими отклонениями (не превышают 0,5—1 с. к. о. в направлении соответствующих осей), вероятность попадания в эту область может быть приближенно вычислена по формуле, не содержащей операции интегрирования.

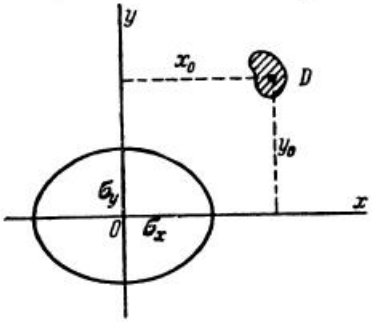


Рис. 9.5.3.

Рассмотрим на плоскости xOy малую цель D произвольной формы (рис. 9.5.3). Допустим, что размеры этой цели невелики по сравнению с вероятными отклонениями E_x, E_y. По общей формуле (8.3.3) имеем:

$$P\big((X, Y) \subset D\big) = \iint\limits_{(D)} f(x, y)\, dx\, dy, \qquad (9.5.2)$$

где $f(x, y)$ — плотность распределения системы (X, Y). Применим к интегралу (9.5.2) теорему о среднем значении:

$$P\big((X, Y) \subset D\big) = f(x_0, y_0) \iint\limits_{(D)} dx\, dy = f(x_0, y_0)\, S_D,$$

где (x_0, y_0) — некоторая точка внутри области D; S_D — площадь области D.

В случае, когда система (X, Y) подчинена нормальному закону в каноническом виде, имеем:

$$P\big((X, Y) \subset D\big) = \frac{S_D}{2\pi\sigma_x\sigma_y} e^{-\frac{x_0^2}{2\sigma_x^2} - \frac{y_0^2}{2\sigma_y^2}}. \qquad (9.5.3)$$

При сравнительно малых размерах области D плотность распределения $f(x, y)$ в пределах этой области изменяется мало и практически может быть принята постоянной. Тогда в качестве точки (x_0, y_0) можно выбрать любую точку в пределах области D (например, приблизительный центр цели).

Формулы типа (9.5.3) широко применяются на практике. Для областей, наибольшие размеры которых не превышают 0,5 среднего квадратического отклонения в соответствующем направлении, они дают вполне приемлемые по точности результаты. В отдельных случаях их применяют и для более крупных областей (порядка одного с. к. о.). При условии внесения некоторых поправок (а именно, замены величин σ_x, σ_y несколько увеличенными значениями) область применимости этой формулы может быть расширена на области размером порядка двух с. к. о.

9.6. Нормальный закон в пространстве трех измерений. Общая запись нормального закона для системы произвольного числа случайных величин

При исследовании вопросов, связанных со стрельбой дистанционными снарядами, приходится иметь дело с законом распределения точек разрыва дистанционного снаряда в пространстве. При условии применения обычных дистанционных взрывателей этот закон распределения может считаться нормальным.

В данном п° мы рассмотрим лишь каноническую форму нормального закона в пространстве:

$$f(x,\ y,\ z)=\frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}+\frac{z^2}{\sigma_z^2}\right)}, \qquad (9.6.1)$$

где σ_x, σ_y, σ_z — главные средние квадратические отклонения.

Переходя от средних квадратических отклонений к вероятным, имеем:

$$f(x,\ y,\ z)=\frac{\rho^3}{\pi^{3/2}E_xE_yE_z}e^{-\rho^2\left(\frac{x^2}{E_x^2}+\frac{y^2}{E_y^2}+\frac{z^2}{E_z^2}\right)}. \qquad (9.6.2)$$

При решении задач, связанных со стрельбой дистанционными снарядами, иногда приходится вычислять вероятность разрыва дистанционного снаряда в пределах заданной области D. В общем случае эта вероятность выражается тройным интегралом:

$$P\big((X,\ Y,\ Z)\subset D\big)=\int\limits_{(D)}\int\int f(x,\ y,\ z)\,dx\,dy\,dz. \qquad (9.6.3)$$

Интеграл (9.6.3) обычно не выражается через элементарные функции. Однако существует ряд областей, вероятность попадания в которые вычисляется сравнительно просто.

1. Вероятность попадания в прямоугольный параллелепипед со сторонами, параллельными главным осям рассеивания

Пусть область R представляет собой прямоугольный параллелепипед, ограниченный абсциссами α, β, ординатами γ, δ и аппликатами ε, η (рис. 9.6.1). Вероятность попадания в область R, очевидно, равна:

$$P\big((X,\ Y,\ Z)\subset R\big)=$$

$$=\left[\Phi^*\left(\frac{\beta}{\sigma_x}\right)-\Phi^*\left(\frac{\alpha}{\sigma_x}\right)\right]\left[\Phi^*\left(\frac{\delta}{\sigma_y}\right)-\Phi^*\left(\frac{\gamma}{\sigma_y}\right)\right]\left[\Phi^*\left(\frac{\eta}{\sigma_z}\right)-\Phi^*\left(\frac{\varepsilon}{\sigma_z}\right)\right]. \qquad (9.6.4)$$

2. Вероятность попадания в эллипсоид равной плотности

Рассмотрим эллипсоид равной плотности B_k, уравнение которого

$$\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}+\frac{z^2}{\sigma_z^2}=k^2.$$

Полуоси этого эллипсоида пропорциональны главным средним квадратическим отклонениям:

$$a=k\sigma_x; \quad b=k\sigma_y; \quad c=k\sigma_z.$$

Пользуясь формулой (9.6.1) для $f(x, y, z)$, выразим вероятность

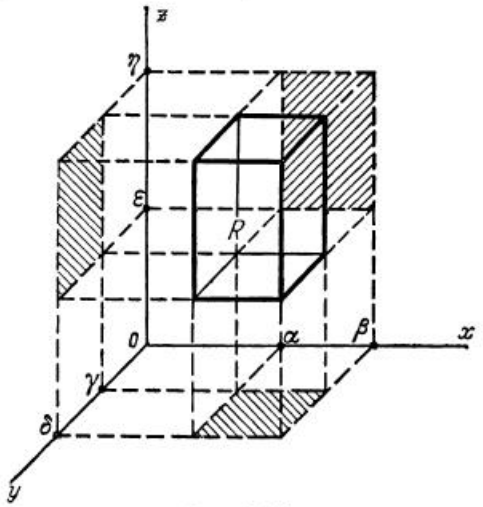


Рис. 9.6.1.

попадания в эллипсоид B_k:

$$P((X, Y, Z)\subset B_k)=\frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z}\iiint\limits_{(B_k)} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}+\frac{z^2}{\sigma_z^2}\right)}dx\,dy\,dz.$$

Перейдем от декартовых координат к полярным (сферическим) заменой переменных

$$\left.\begin{aligned}\frac{x}{\sigma_x}&=r\cos\theta\cos\varphi,\\ \frac{y}{\sigma_y}&=r\cos\theta\sin\varphi,\\ \frac{z}{\sigma_z}&=r\sin\theta.\end{aligned}\right\} \qquad (9.6.5)$$

Якобиан преобразования (9.6.5) равен:

$$J = r^2 \cos\theta \sigma_x \sigma_y \sigma_z.$$

Переходя к новым переменным, имеем:

$$P\big((X, Y, Z) \subset B_k\big) =$$

$$= \frac{1}{(2\pi)^{\frac{3}{2}}} \int_0^k \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} r^2 \cos\theta e^{-\frac{r^2}{2}} dr\, d\theta\, d\varphi = \sqrt{\frac{2}{\pi}} \int_0^k r^2 e^{-\frac{r^2}{2}} dr.$$

Интегрируя по частям, получим:

$$P\big((X, Y, Z) \subset B_k\big) = \sqrt{\frac{2}{\pi}} \left\{ -ke^{-\frac{k^2}{2}} + \int_0^k e^{-\frac{r^2}{2}} dr \right\} =$$

$$= 2\Phi^*(k) - 1 - \sqrt{\frac{2}{\pi}} ke^{-\frac{k^2}{2}}. \quad (9.6.6)$$

3. Вероятность попадания в цилиндрическую область с образующей, параллельной одной из главных осей рассеивания

Рассмотрим цилиндрическую область C, образующая которой параллельна одной из главных осей рассеивания (например, оси Oz) а направляющая есть контур произвольной области D на плоскости xOy (рис. 9.6.2). Пусть область C ограничена двумя плоскостями $z = \varepsilon$ и $z = \eta$. Вычислим вероятность попадания в область C; это есть вероятность произведения двух событий, первое из которых состоит в попадании точки (X, Y) в область D, а второе — в попадании величины Z на участок (ε, η). Так как величины (X, Y, Z), подчиненные нормальному закону в канонической форме, независимы, то независимы и эти два события. Поэтому

$$P\big((X, Y, Z) \subset C\big) =$$

$$= P\big((X, Y) \subset D\big) P(\varepsilon < Z < \eta) =$$

$$= P\big((X, Y) \subset D\big) \left[\Phi^*\left(\frac{\eta}{\sigma_z}\right) - \Phi^*\left(\frac{\varepsilon}{\sigma_z}\right) \right]. \quad (9.6.7)$$

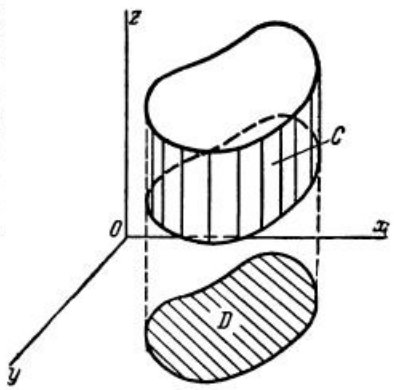


Рис. 9.6.2.

Вероятность $P\big((X, Y) \subset D\big)$ в формуле (9.6.7) может быть вычислена любым из способов вычисления вероятности попадания в плоскую область.

На формуле (9.6.7) основан следующий способ вычисления вероятности попадания в пространственную область G произвольной формы: область G приближенно разбивается на ряд цилиндрических областей G_1, G_2, ... (рис. 9.6.3), и вероятность попадания в каждую

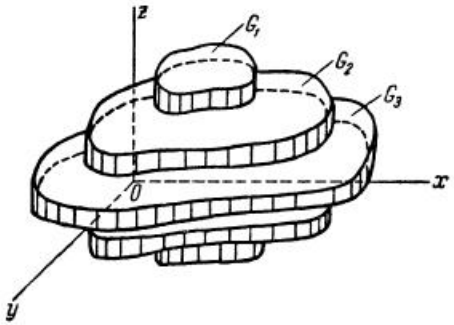


Рис. 9.6.3.

из них вычисляется по формуле (9.6.7). Для применения этого способа достаточно начертить ряд фигур, представляющих собой сечения области G плоскостями, параллельными одной из координатных плоскостей. Вероятность попадания в каждую из них вычисляется по сетке рассеивания.

В заключение данной главы напишем общее выражение для нормального закона в пространстве любого числа измерений n.

Плотность распределения такого закона имеет вид:

$$f(x_1, x_2, \ldots, x_n) = \frac{\sqrt{|C|}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(x_i - m_{x_i})(x_j - m_{x_j})}, \quad (9.6.8)$$

где $|C|$ — определитель матрицы C, $C = \|C_{ij}\|$ — матрица, обратная корреляционной матрице K, т. е. если корреляционная матрица

$$K = \|K_{ij}\|,$$

то

$$C_{ij} = (-1)^{i+j}\frac{M_{ij}}{|K|},$$

где $|K|$ — определитель корреляционной матрицы, а M_{ij} — минор этого определителя, получаемый из него вычеркиванием i-й строки

и j-го столбца. Заметим, что

$$|C| = \frac{1}{|K|}.$$

Из общего выражения (9.6.8) вытекают все формы нормального закона для любого числа измерений и для любых видов зависимости между случайными величинами. В частности, при $n = 2$ (рассеивание на плоскости) корреляционная матрица есть

$$K = \begin{Vmatrix} \sigma_x^2 & \sigma_x\sigma_y r \\ \sigma_x\sigma_y r & \sigma_y^2 \end{Vmatrix},$$

где r — коэффициент корреляции. Отсюда

$$|K| = \sigma_x^2\sigma_y^2(1-r^2); \quad |C| = \frac{1}{\sigma_x^2\sigma_y^2(1-r^2)};$$

$$C = \begin{Vmatrix} \frac{1}{\sigma_x^2(1-r^2)} & \frac{-r}{\sigma_x\sigma_y(1-r^2)} \\ \frac{-r}{\sigma_x\sigma_y(1-r^2)} & \frac{1}{\sigma_y^2(1-r^2)} \end{Vmatrix}.$$

Подставляя определитель матрицы $|C|$ и ее члены в (9.6.8), получим формулу (9.1.1) для нормального закона на плоскости, с которой мы начали п° 9.1.

ГЛАВА 10

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ФУНКЦИЙ СЛУЧАЙНЫХ ВЕЛИЧИН

10.1. Математическое ожидание функции. Дисперсия функции

При решении различных задач, связанных со случайными явлениями, современная теория вероятностей широко пользуется аппаратом случайных величин. Для того чтобы пользоваться этим аппаратом, необходимо знать законы распределения фигурирующих в задаче случайных величин. Вообще говоря, эти законы могут быть определены из опыта, но обычно опыт, целью которого является определение закона распределения случайной величины или системы случайных величин (особенно в области военной техники), оказывается и сложным и дорогостоящим. Естественно возникает задача — свести объем эксперимента к минимуму и составлять суждение о законах распределения случайных величин косвенным образом, на основании уже известных законов распределения других случайных величин. Такие косвенные методы исследования случайных величин играют весьма большую роль в теории вероятностей. При этом обычно интересующая нас случайная величина представляется как функция других случайных величин; зная законы распределения аргументов, часто удается установить закон распределения функции. С рядом задач такого типа мы встретимся в дальнейшем (см. главу 12).

Однако на практике часто встречаются случаи, когда нет особой надобности полностью определять закон распределения функции случайных величин, а достаточно только указать его числовые характеристики: математическое ожидание, дисперсию, иногда — некоторые из высших моментов. К тому же очень часто самые законы распределения аргументов бывают известны недостаточно хорошо. В связи с этим часто возникает задача об определении только числовых характеристик функций случайных величин.

Рассмотрим такую задачу: случайная величина Y есть функция нескольких случайных величин $X_1, X_2, \ldots, X_n$:

$$Y = \varphi(X_1, X_2, \ldots, X_n).$$

Пусть нам известен закон распределения системы аргументов $(X_1, X_2, \ldots, X_n)$; требуется найти числовые характеристики величины Y, в первую очередь — математическое ожидание и дисперсию.

Представим себе, что нам удалось тем или иным способом найти закон распределения $g(y)$ величины Y. Тогда задача об определении числовых характеристик становится тривиальной; они находятся по формулам:

$$m_y = \int_{-\infty}^{\infty} y g(y)\, dy;$$

$$D_y = \int_{-\infty}^{\infty} (y - m_y)^2 g(y)\, dy$$

и т. д.

Однако самая задача нахождения закона распределения $g(y)$ величины Y часто оказывается довольно сложной. К тому же для решения поставленной нами задачи нахождение закона распределения величины Y как такового вовсе и не нужно: чтобы найти только числовые характеристики величины Y, нет надобности знать ее закон распределения; достаточно знать закон распределения аргументов $(X_1, X_2, \ldots, X_n)$. Более того, в некоторых случаях, для того чтобы найти числовые характеристики функции, не требуется даже знать закона распределения ее аргументов; достаточно бывает знать лишь некоторые числовые характеристики аргументов.

Таким образом, возникает задача определения числовых характеристик функций случайных величин помимо законов распределения этих функций.

Рассмотрим задачу об определении числовых характеристик функции при заданном законе распределения аргументов. Начнем с самого простого случая — функции одного аргумента — и поставим следующую задачу.

Имеется случайная величина X с заданным законом распределения; другая случайная величина Y связана с X функциональной зависимостью:

$$Y = \varphi(X).$$

Требуется, не находя закона распределения величины Y, определить ее математическое ожидание:

$$m_y = M[\varphi(X)]. \qquad (10.1.1)$$

Рассмотрим сначала случай, когда X есть прерывная случайная величина с рядом распределения:

x_i	x_1	x_2	$\ldots$	x_n
p_i	p_1	p_2	$\ldots$	p_n

Выпишем возможные значения величины Y и вероятности этих значений:

$$\begin{array}{c|c|c|c|c|c} \varphi(x_i) & \varphi(x_1) & \varphi(x_2) & \varphi(x_3) & \dots & \varphi(x_n) \\ \hline p_i & p_1 & p_2 & p_3 & \dots & p_n \end{array} \qquad (10.1.2)$$

Таблица (10.1.2) не является в строгом смысле слова рядом распределения величины Y, так как в общем случае некоторые из значений

$$\varphi(x_1),\ \varphi(x_2),\ \dots,\ \varphi(x_n) \qquad (10.1.3)$$

могут совпадать между собой; к тому же эти значения в верхнем столбце таблицы (10.1.2) не обязательно идут в возрастающем порядке. Для того чтобы от таблицы (10.1.2) перейти к подлинному ряду распределения величины Y, нужно было бы расположить значения (10.1.3) в порядке возрастания, объединить столбцы, соответствующие равным между собой значениям Y, и сложить соответствующие вероятности. Но в данном случае нас не интересует закон распределения величины Y как таковой; для наших целей — определения математического ожидания — достаточно такой «неупорядоченной» формы ряда распределения, как (10.1.2). Математическое ожидание величины Y можно определить по формуле

$$M[\varphi(X)] = \sum_i \varphi(x_i)\, p_i. \qquad (10.1.4)$$

Очевидно, величина $m_y = M[\varphi(X)]$, определяемая по формуле (10.1.4), не может измениться от того, что под знаком суммы некоторые члены будут объединены заранее, а порядок членов изменен.

В формуле (10.1.4) для математического ожидания функции не содержится в явном виде закона распределения самой функции, а содержится только закон распределения аргумента. Таким образом, *для определения математического ожидания функции вовсе не требуется знать закон распределения этой функции, а достаточно знать закон распределения аргумента.*

Заменяя в формуле (10.1.4) сумму интегралом, а вероятность p_i — элементом вероятности, получим аналогичную формулу для непрерывной случайной величины:

$$M[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x) f(x)\, dx, \qquad (10.1.5)$$

где $f(x)$ — плотность распределения величины X.

Аналогично может быть определено математическое ожидание функции $\varphi(X, Y)$ от двух случайных аргументов X и Y. Для прерывных величин

$$M[\varphi(X, Y)] = \sum_i \sum_j \varphi(x_i, y_j)\, p_{ij}, \qquad (10.1.6)$$

где $p_{ij} = P\big((X = x_i)(Y = y_j)\big)$ — вероятность того, что система (X, Y) примет значения (x_i, y_j).

Для непрерывных величин

$$M[\varphi(X, Y)] = \int_{-\infty}^{\infty}\int \varphi(x, y) f(x, y)\, dx\, dy, \qquad (10.1.7)$$

где $f(x, y)$ — плотность распределения системы (X, Y).

Совершенно аналогично определяется математическое ожидание функции от произвольного числа случайных аргументов. Приведем соответствующую формулу только для непрерывных величин:

$$M[\varphi(X_1, X_2, \ldots, X_n)] =$$

$$= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \varphi(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n, \qquad (10.1.8)$$

где $f(x_1, x_2, \ldots, x_n)$ — плотность распределения системы $(X_1, X_2, \ldots, X_n)$.

Формулы типа (10.1.8) весьма часто встречаются в практическом применении теории вероятностей, когда речь идет об осреднении каких-либо величин, зависящих от ряда случайных аргументов.

Таким образом, математическое ожидание функции любого числа случайных аргументов может быть найдено помимо закона распределения функции. Аналогично могут быть найдены и другие числовые характеристики функции — моменты различных порядков. Так как каждый момент представляет собой математическое ожидание некоторой функции исследуемой случайной величины, то вычисление любого момента может быть осуществлено приемами, совершенно аналогичными вышеизложенным. Здесь мы приведем расчетные формулы только для дисперсии, причем лишь для случая непрерывных случайных аргументов.

Дисперсия функции одного случайного аргумента выражается формулой

$$D[\varphi(X)] = \int_{-\infty}^{\infty} [\varphi(x) - m_\varphi]^2 f(x)\, dx, \qquad (10.1.9)$$

где $m_\varphi = M[\varphi(x)]$ — математическое ожидание функции $\varphi(X)$; $f(x)$ — плотность распределения величины X.

Аналогично выражается дисперсия функции двух аргументов:

$$D[\varphi(X, Y)] = \int_{-\infty}^{\infty}\int [\varphi(x, y) - m_\varphi]^2 f(x, y)\, dx\, dy, \qquad (10.1.10)$$

где m_φ — математическое ожидание функции $\varphi(X, Y)$;
$f(x, y)$ — плотность распределения системы (X, Y).

Наконец, в случае произвольного числа аргументов, в аналогичных обозначениях:

$$D[\varphi(X_1, X_2, \ldots, X_n)] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} [\varphi(x_1, x_2, \ldots, x_n) - m_\varphi]^2 f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n. \quad (10.1.11)$$

Заметим, что часто при вычислении дисперсии бывает удобно пользоваться соотношением между начальным и центральным моментами второго порядка (см. главу 5) и писать:

$$D[\varphi(X)] = \int_{-\infty}^{\infty} [\varphi(x)]^2 f(x)\, dx - m_\varphi^2; \quad (10.1.12)$$

$$D[\varphi(X, Y)] = \int_{-\infty}^{\infty}\int [\varphi(x, y)]^2 f(x, y)\, dx\, dy - m_\varphi^2; \quad (10.1.13)$$

$$D[\varphi(X_1, X_2, \ldots, X_n)] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} [\varphi(x_1, x_2, \ldots, x_n)]^2 f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n - m_\varphi^2. \quad (10.1.14)$$

Формулы (10.1.12) — (10.1.14) можно рекомендовать тогда, когда они не приводят к разностям близких чисел, т. е. когда m_φ сравнительно невелико.

Рассмотрим несколько примеров, иллюстрирующих применение изложенных выше методов для решения практических задач.

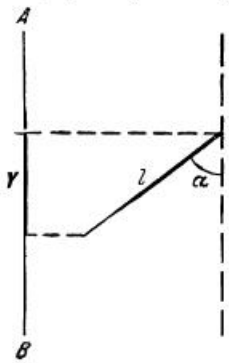


Рис. 10.1.1.

Пример 1. На плоскости задан отрезок длины l (рис. 10.1.1), вращающийся случайным образом так, что все направления его одинаково вероятны. Отрезок проектируется на неподвижную ось AB. Определить среднее значение длины проекции отрезка.

Решение. Длина проекции равна:

$$Y = l\,|\cos\alpha|,$$

где угол α — случайная величина, распределенная с равномерной плотностью на участке $0, 2\pi$.

По формуле (10.1.5) имеем:

$$m = M[l\,|\cos\alpha|] = \int_0^{2\pi} l\,|\cos\alpha|\,\frac{d\alpha}{2\pi} = \frac{2l}{\pi}\int_0^{\frac{\pi}{2}} \cos\alpha\, d\alpha = \frac{2l}{\pi} \approx 0{,}637l.$$

Пример 2. Удлиненный осколок снаряда, который можно схематически изобразить отрезком длины l, летит, вращаясь вокруг центра массы таким образом, что все его ориентации в пространстве одинаково вероятны. На своем пути осколок встречает плоский экран, перпендикулярный к направлению его движения, и оставляет в нем пробоину. Найти математическое ожидание длины этой пробоины.

Решение. Прежде всего дадим математическую формулировку утверждения, заключающегося в том, что «все ориентации осколка в пространстве одинаково вероятны». Направление отрезка l будем характеризовать единичным вектором $\vec{r}$ (рис. 10.1.2). Направление вектора $\vec{r}$ в сферической системе координат, связанной с плоскостью P, на которую производится проектирование, определяется двумя углами: углом θ, лежащим в плоскости P, и углом φ, лежащим в плоскости, перпендикулярной к P. При равной вероятности всех направлений вектора $\vec{r}$ все положения его конца на поверхности сферы единичного радиуса C должны обладать одинаковой плотностью вероятности; следовательно, элемент вероятности

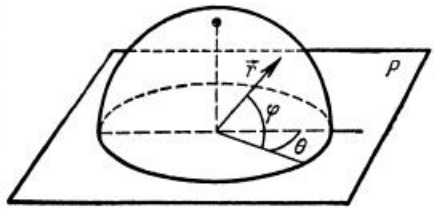


Рис. 10.1.2.

$$f(\theta, \varphi)\, d\theta\, d\varphi,$$

где $f(\theta, \varphi)$ — плотность распределения углов θ, φ, должен быть пропорционален элементарной площадке ds на сфере C; эта элементарная площадка равна

$$ds = d\theta\, d\varphi \cos\varphi,$$

откуда

$$f(\theta, \varphi)\, d\theta\, d\varphi = A \cos\varphi\, d\theta\, d\varphi; \quad f(\theta, \varphi) = A \cos\varphi,$$

где A — коэффициент пропорциональности.

Значение коэффициента A найдем из соотношения

$$\int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta, \varphi)\, d\theta\, d\varphi = 1,$$

откуда

$$A = \frac{1}{4\pi}.$$

Таким образом, плотность распределения углов θ, φ выражается формулой

$$f(\theta, \varphi) = \frac{1}{4\pi} \cos\varphi \quad \text{при} \quad \begin{cases} 0 < \theta < 2\pi, \\ -\frac{\pi}{2} < \varphi < \frac{\pi}{2}. \end{cases} \tag{10.1.15}$$

Спроектируем отрезок на плоскость P; длина проекции равна:

$$Y = l \cos\varphi.$$

Рассматривая Y как функцию двух аргументов θ и φ и применяя формулу (10.1.7), получим:

$$m_y = M[l\cos\varphi] = \frac{l}{4\pi}\int_0^{2\pi} d\theta \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\varphi\, d\varphi = \frac{l}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\varphi\, d\varphi = \frac{l\pi}{4} \approx 0{,}785l.$$

Таким образом, средняя длина пробоины, оставляемой осколком в экране, равна 0,785 длины осколка.

Пример 3. Плоская фигура площади S беспорядочно вращается в пространстве так, что все ориентации этой фигуры одинаково вероятны. Найти среднюю площадь проекции фигуры S на неподвижную плоскость P (рис. 10.1.3).

Решение. Направление плоскости фигуры S в пространстве будем характеризовать направлением нормали N к этой плоскости. С плоскостью P свяжем ту же сферическую систему координат, что в предыдущем примере. Направление нормали N к площадке S характеризуется случайными углами θ и φ, распределенными с плотностью (10.1.5). Площадь Z проекции фигуры S на плоскость P равна

$$Z = S\left|\cos\left(\frac{\pi}{2}-\varphi\right)\right| = S|\sin\varphi|,$$

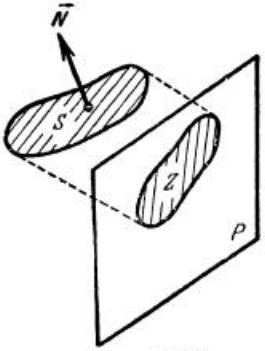


Рис. 10.1.3.

а средняя площадь проекции

$$m_z = M[S|\sin\varphi|] = \frac{S}{4\pi}\int_0^{2\pi} d\theta \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\varphi\,|\sin\varphi|\, d\varphi = \frac{S}{2}.$$

Таким образом, средняя площадь проекции произвольно ориентированной плоской фигуры на неподвижную плоскость равна половине площади этой фигуры.

Пример 4. В процессе слежения радиолокатором за определенным объектом пятно, изображающее объект, все время удерживается в пределах экрана. Экран представляет собой круг K радиуса R. Пятно занимает на экране случайное положение с постоянной плотностью вероятности. Найти среднее расстояние от пятна до центра экрана.

Решение. Обозначая расстояние D, имеем $D = \sqrt{X^2+Y^2}$, где X, Y — координаты пятна; $f(x, y) = \frac{1}{\pi R^2}$ в пределах круга K и равна нулю за его пределами. Применяя формулу (10.1.7) и переходя в интеграле к полярным координатам, имеем:

$$m_D = \frac{1}{\pi R^2}\iint_{(K)} \sqrt{x^2+y^2}\, dx\, dy = \frac{1}{\pi R^2}\int_0^{2\pi} d\varphi \int_0^R r^2\, dr = \frac{2}{3}R.$$

Пример 5. Надежность (вероятность безотказной работы) технического устройства есть определенная функция $p(X, Y, Z)$ трех параметров, характеризующих работу регулятора. Параметры X, Y, Z представляют собой случайные величины с известной плотностью распределения $f(x, y, z)$.

Найти среднее значение (математическое ожидание) надежности устройства и среднее квадратическое отклонение, характеризующее ее устойчивость.

Р е ш е н и е. Надежность устройства $p(X, Y, Z)$ есть функция трех случайных величин (параметров) X, Y, Z. Ее среднее значение (математическое ожидание) найдется по формуле (10.1.8):

$$m_p = M[p(X, Y, Z)] = \int_{-\infty}^{\infty}\int\int p(x, y, z) f(x, y, z)\, dx\, dy\, dz. \quad (10.1.16)$$

По формуле (10.1.14) имеем:

$$D_p = D[p(X, Y, Z)] = \int_{-\infty}^{\infty}\int\int [p(x, y, z)]^2 f(x, y, z)\, dx\, dy\, dz - m_p^2,$$

$$\sigma = \sqrt{D_p}.$$

Формула (10.1.16), выражающая среднюю (полную) вероятность безотказной работы устройства с учетом случайных величин, от которых зависит эта вероятность в каждом конкретном случае, представляет собой частный случай так называемой *интегральной формулы полной вероятности*, обобщающей обычную формулу полной вероятности на случай бесконечного (несчетного) числа гипотез.

Выведем здесь эту формулу в общем виде.

Предположим, что опыт, в котором может появиться или не появиться интересующее нас событие A, протекает в случайных, заранее неизвестных условиях. Пусть эти условия характеризуются непрерывными случайными величинами

$$X_1, X_2, \ldots, X_n, \quad (10.1.17)$$

плотность распределения которых

$$f(x_1, x_2, \ldots, x_n).$$

Вероятность P_A появления события A есть некоторая функция случайных величин (10.1.17):

$$P_A(X_1, X_2, \ldots, X_n). \quad (10.1.18)$$

Нам нужно найти среднее значение этой вероятности или, другими словами, полную вероятность события A:

$$P_A = M[P_A(X_1, X_2, \ldots, X_n)].$$

Применяя формулу (10.1.8) для математического ожидания функции, найдем:

$$P_A = \int_{-\infty}^{\infty}\int\ldots\int P_A(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n, \quad (10.1.19)$$

Формула (10.1.19) называется *интегральной формулой полной вероятности*. Нетрудно заметить, что по своей структуре она сходна с формулой полной вероятности, если заменить дискретный ряд гипотез непрерывной гаммой, сумму — интегралом, вероятность гипотезы — элементом вероятности:

$$f(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2 \ldots dx_n,$$

а условную вероятность события при данной гипотезе — условной вероятностью события при фиксированных значениях случайных величин:

$$P_A(x_1, x_2, \ldots, x_n).$$

Не менее часто, чем интегральной формулой полной вероятности, пользуются *интегральной формулой полного математического ожидания*. Эта формула выражает среднее (полное) математическое ожидание случайной величины Z, значение которой принимается в опыте, условия которого заранее неизвестны (случайны). Если эти условия характеризуются непрерывными случайными величинами

$$X_1, X_2, \ldots, X_n$$

с плотностью распределения

$$f(x_1, x_2, \ldots, x_n),$$

а математическое ожидание величины Z есть функция от величин $X_1, X_2, \ldots, X_n$:

$$m_z(X_1, X_2, \ldots, X_n),$$

то полное математическое ожидание величины Z вычисляется по формуле

$$m_z = \int_{-\infty}^{\infty} \int \ldots \int m_z(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2 \ldots dx_n, \tag{10.1.20}$$

которая называется *интегральной формулой полного математического ожидания*.

Пример 6. Математическое ожидание расстояния D, на котором будет обнаружен объект с помощью четырех радиолокационных станций, зависит от некоторых технических параметров этих станций:

$$X_1, X_2, X_3, X_4,$$

которые представляют собой независимые случайные величины с плотностью распределения

$$f(x_1, x_2, x_3, x_4) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4).$$

При фиксированных значениях параметров $X_1 = x_1$, $X_2 = x_2$, $X_3 = x_3$, $X_4 = x_4$ математическое ожидание дальности обнаружения равно

$$m_D(x_1, x_2, x_3, x_4).$$

Найти среднее (полное) математическое ожидание дальности обнаружения.

Решение. По формуле (10.1.20) имеем:

$$m_D = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} m_D(x_1, x_2, x_3, x_4) f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)\, dx_1\, dx_2\, dx_3\, dx_4.$$

10.2. Теоремы о числовых характеристиках

В предыдущем п° мы привели ряд формул, позволяющих находить числовые характеристики функций, когда известны законы распределения аргументов. Однако во многих случаях для нахождения числовых характеристик функций не требуется знать даже законов распределения аргументов, а достаточно знать только некоторые их числовые характеристики; при этом мы вообще обходимся без каких бы то ни было законов распределения. Определение числовых характеристик функций по заданным числовым характеристикам аргументов широко применяется в теории вероятностей и позволяет значительно упрощать решение ряда задач. По преимуществу такие упрощенные методы относятся к линейным функциям; однако некоторые элементарные нелинейные функции также допускают подобный подход.

В настоящем п° мы изложим ряд теорем о числовых характеристиках функций, представляющих в своей совокупности весьма простой аппарат вычисления этих характеристик, применимый в широком круге условий.

1. Математическое ожидание неслучайной величины

Если c — неслучайная величина, то

$$M[c] = c.$$

Сформулированное свойство является достаточно очевидным; доказать его можно, рассматривая неслучайную величину c как частный вид случайной, при одном возможном значении с вероятностью единица; тогда по общей формуле для математического ожидания:

$$M[c] = c \cdot 1 = c.$$

2. Дисперсия неслучайной величины

Если c — неслучайная величина, то

$$D[c] = 0.$$

Доказательство. По определению дисперсии

$$D[c] = M[\mathring{c}^2] = M[(c - m_c)^2] = M[(c - c)^2] = M[0] = 0.$$

3. Вынесение неслучайной величины за знак математического ожидания

Если c — неслучайная величина, а X — случайная, то

$$M[cX] = cM[X], \tag{10.2.1}$$

т. е. *неслучайную величину можно выносить за знак математического ожидания.*

Доказательство.

а) Для прерывных величин

$$M[cX] = \sum_i cx_i p_i = c\sum_i x_i p_i = cM[X].$$

б) Для непрерывных величин

$$M[cX] = \int_{-\infty}^{\infty} cxf(x)\,dx = c\int_{-\infty}^{\infty} xf(x)\,dx = cM[X].$$

4. Вынесение неслучайной величины за знак дисперсии и среднего квадратического отклонения

Если c — неслучайная величина, а X — случайная, то

$$D[cX] = c^2 D[X], \tag{10.2.2}$$

т. е. *неслучайную величину можно выносить за знак дисперсии, возводя ее в квадрат.*

Доказательство. По определению дисперсии

$$D[cX] = M[(cX - M[cX])^2] =$$

$$= M[(cX - cm_x)^2] = c^2 M[(X - m_x)^2] = c^2 D[X].$$

Следствие

$$\sigma[cX] = |c|\sigma[X],$$

т. е. неслучайную величину можно выносить за знак среднего квадратического отклонения ее абсолютным значением. Доказательство получим, извлекая корень квадратный из формулы (10.2.2) и учитывая, что с.к.о. — существенно положительная величина.

5. Математическое ожидание суммы случайных величин

Докажем, что для любых двух случайных величин X и Y

$$M[X+Y]=M[X]+M[Y], \tag{10.2.3}$$

т. е. *математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий.*

Это свойство известно под названием *теоремы сложения математических ожиданий.*

Доказательство.

а) Пусть (X, Y) — система прерывных случайных величин. Применим к сумме случайных величин общую формулу (10.1.6) для математического ожидания функции двух аргументов:

$$M[X+Y]=\sum_i\sum_j (x_i+y_j)\,p_{ij}=$$

$$=\sum_i\sum_j x_i p_{ij}+\sum_i\sum_j y_j p_{ij}=$$

$$=\sum_i x_i\sum_j p_{ij}+\sum_j y_j\sum_i p_{ij}.$$

Но $\sum_j p_{ij}$ представляет собой не что иное, как полную вероятность того, что величина X примет значение x_i:

$$\sum_j p_{ij}=P(X=x_i)=p_i;$$

следовательно,

$$\sum_i x_i\sum_j p_{ij}=\sum_i x_i p_i=M[X].$$

Аналогично докажем, что

$$\sum_j y_j\sum_i p_{ij}=M[Y],$$

и теорема доказана.

б) Пусть (X, Y) — система непрерывных случайных величин. По формуле (10.1.7)

$$M[X+Y]=\int_{-\infty}^{\infty}\int (x+y)\,f(x,\,y)\,dx\,dy=$$

$$=\int_{-\infty}^{\infty}\int xf(x,\,y)\,dx\,dy+\int_{-\infty}^{\infty}\int yf(x,\,y)\,dx\,dy. \tag{10.2.4}$$

Преобразуем первый из интегралов (10.2.4):

$$\int_{-\infty}^{\infty}\int xf(x, y)\,dx\,dy = \int_{-\infty}^{\infty} x\left\{\int_{-\infty}^{\infty} f(x, y)\,dy\right\}dx =$$

$$= \int_{-\infty}^{\infty} xf_1(x)\,dx = M[X];$$

аналогично

$$\int_{-\infty}^{\infty}\int yf(x, y)\,dx\,dy = M[Y],$$

и теорема доказана.

Следует специально отметить, что теорема сложения математических ожиданий справедлива для любых случайных величин — как зависимых, так и независимых.

Теорема сложения математических ожиданий обобщается на произвольное число слагаемых:

$$M\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} M[X_i], \qquad (10.2.5)$$

т. е. *математическое ожидание суммы нескольких случайных величин равно сумме их математических ожиданий.*

Для доказательства достаточно применить метод полной индукции.

6. Математическое ожидание линейной функции

Рассмотрим линейную функцию нескольких случайных аргументов $X_1, X_2, \ldots, X_n$:

$$\sum_{i=1}^{n} a_i X_i + b,$$

где a_i, b — неслучайные коэффициенты. Докажем, что

$$M\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i M[X_i] + b, \qquad (10.2.6)$$

т. е. математическое ожидание линейной функции равно той же линейной функции от математических ожиданий аргументов.

Доказательство. Пользуясь теоремой сложения м. о. и правилом вынесения неслучайной величины за знак м. о., получим:

$$M\left[\sum_{i=1}^{n} a_i X_i + b\right] = M\left[\sum_{i=1}^{n} a_i X_i\right] + M[b] =$$

$$= \sum_{i=1}^{n} M[a_i X_i] + b = \sum_{i=1}^{n} a_i M[X_i] + b.$$

7. Дисперсия суммы случайных величин

Дисперсия суммы двух случайных величин равна сумме их дисперсий плюс удвоенный корреляционный момент:

$$D[X+Y] = D[X] + D[Y] + 2K_{xy}. \qquad (10.2.7)$$

Доказательство. Обозначим

$$Z = X + Y. \qquad (10.2.8)$$

По теореме сложения математических ожиданий

$$m_z = m_x + m_y. \qquad (10.2.9)$$

Перейдем от случайных величин X, Y, Z к соответствующим центрированным величинам $\mathring{X}$, $\mathring{Y}$, $\mathring{Z}$. Вычитая почленно из равенства (10.2.8) равенство (10.2.9), имеем:

$$\mathring{Z} = \mathring{X} + \mathring{Y}.$$

По определению дисперсии

$$D[X+Y] = D[Z] = M[\mathring{Z}^2] =$$

$$= M[\mathring{X}^2] + 2M[\mathring{X}\mathring{Y}] + M[\mathring{Y}^2] = D[X] + 2K_{xy} + D[Y],$$

что и требовалось доказать.

Формула (10.2.7) для дисперсии суммы может быть обобщена на любое число слагаемых:

$$D\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} D[X_i] + 2\sum_{i<j} K_{ij}, \qquad (10.2.10)$$

где K_{ij} — корреляционный момент величин X_i, X_j; знак $i<j$ под суммой обозначает, что суммирование распространяется на все возможные попарные сочетания случайных величин (X_1, X_2, ..., X_n).

Доказательство аналогично предыдущему и вытекает из формулы для квадрата многочлена.

Формула (10.2.10) может быть записана еще в другом виде:

$$D\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} K_{ij}, \qquad (10.2.11)$$

где двойная сумма распространяется на все элементы корреляционной матрицы системы величин $(X_1, X_2, \ldots, X_n)$, содержащей как корреляционные моменты, так и дисперсии.

Если все случайные величины $(X_1, X_2, \ldots, X_n)$, входящие в систему, некоррелированы (т. е. $K_{ij}=0$ при $i\neq j$), формула (10.2.10) принимает вид:

$$D\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} D[X_i], \qquad (10.2.12)$$

т. е. *дисперсия суммы некоррелированных случайных величин равна сумме дисперсий слагаемых.*

Это положение известно под названием *теоремы сложения дисперсий.*

8. Дисперсия линейной функции

Рассмотрим линейную функцию нескольких случайных величин.

$$\sum_{i=1}^{n} a_i X_i + b,$$

где a_i, b — неслучайные величины.

Докажем, что дисперсия этой линейной функции выражается формулой

$$D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i^2 D[X_i] + 2\sum_{i<j} a_i a_j K_{ij}, \qquad (10.2.13)$$

где K_{ij} — корреляционный момент величин X_i, X_j.

Доказательство. Введем обозначение:

$$Y_i = a_i X_i.$$

Тогда

$$\sum_{i=1}^{n} a_i X_i + b = \sum_{i=1}^{n} Y_i + b. \qquad (10.2.14)$$

Применяя к правой части выражения (10.2.14) формулу (10.2.10) для дисперсии суммы и учитывая, что $D[b]=0$, получим:

$$D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} D[Y_i] + 2\sum_{i<j} K_{ij}^{(Y)} = \sum_{i=1}^{n} a_i^2 D[X_i] + 2\sum_{i<j} K_{ij}^{(Y)}, \qquad (10.2.15)$$

где $K_{ij}^{(y)}$ — корреляционный момент величин Y_i, Y_j:

$$K_{ij}^{(y)} = M[\mathring{Y}_i \mathring{Y}_j].$$

Вычислим этот момент. Имеем:

$$\mathring{Y}_i = Y_i - m_{y_i} = a_i X_i - a_i m_{x_i} = a_i \mathring{X}_i;$$

аналогично

$$\mathring{Y}_j = a_j \mathring{X}_j.$$

Отсюда

$$K_{ij}^{(y)} = M[a_i a_j \mathring{X}_i \mathring{X}_j] = a_i a_j M[\mathring{X}_i \mathring{X}_j] = a_i a_j K_{ij}.$$

Подставляя это выражение в (10.2.15), приходим к формуле (10.2.13).

В частном случае, когда все величины $(X_1, X_2, \ldots, X_n)$ некоррелированны, формула (10.2.13) принимает вид:

$$D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i^2 D[X_i], \qquad (10.2.16)$$

т. е. *дисперсия линейной функции некоррелированных случайных величин равна сумме произведений квадратов коэффициентов на дисперсии соответствующих аргументов*[1]).

9. Математическое ожидание произведения случайных величин

Математическое ожидание произведения двух случайных величин равно произведению их математических ожиданий плюс корреляционный момент:

$$M[XY] = M[X]\, M[Y] + K_{xy}. \qquad (10.2.17)$$

Доказательство. Будем исходить из определения корреляционного момента:

$$K_{xy} = M[\mathring{X}\mathring{Y}] = M[(X - m_x)(Y - m_y)],$$

где

$$m_x = M[X]; \quad m_y = M[Y].$$

Преобразуем это выражение, пользуясь свойствами математического ожидания:

$$K_{xy} = M[(X - m_x)(Y - m_y)] =$$

$$= M[XY] - m_x M[Y] - m_y M[X] + m_x m_y = M[XY] - M[X]\, M[Y],$$

что, очевидно, равносильно формуле (10.2.17).

---

[1]) Так как независимые величины всегда являются некоррелированными, то все свойства, доказываемые в данном п° для некоррелированных величин, справедливы для независимых величин.

Если случайные величины (X, Y) некоррелированны ($K_{xy}=0$), то формула (10.2.17) принимает вид:

$$M[XY]=M[X]M[Y], \tag{10.2.18}$$

т. е. *математическое ожидание произведения двух некоррелированных случайных величин равно произведению их математических ожиданий.*

Это положение известно под названием *теоремы умножения математических ожиданий.*

Формула (10.2.17) представляет собой не что иное, как выражение второго смешанного центрального момента системы через второй смешанный начальный момент и математические ожидания:

$$K_{xy}=\mu_{11}=\alpha_{11}-m_x m_y. \tag{10.2.19}$$

Это выражение часто применяется на практике при вычислении корреляционного момента аналогично тому, как для одной случайной величины дисперсия часто вычисляется через второй начальный момент и математическое ожидание.

Теорема умножения математических ожиданий обобщается и на произвольное число сомножителей, только в этом случае для ее применения недостаточно того, чтобы величины были некоррелированны, а требуется, чтобы обращались в нуль и некоторые высшие смешанные моменты, число которых зависит от числа членов в произведении. Эти условия заведомо выполнены при независимости случайных величин, входящих в произведение. В этом случае

$$M\left[\prod_{i=1}^{n} X_i\right]=\prod_{i=1}^{n} M[X_i], \tag{10.2.20}$$

т. е. *математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий.*

Это положение легко доказывается методом полной индукции.

10. Дисперсия произведения независимых случайных величин

Докажем, что для независимых величин X, Y

$$D[XY]=D[X]D[Y]+m_x^2 D[Y]+m_y^2 D[X]. \tag{10.2.21}$$

Доказательство. Обозначим $XY=Z$. По определению дисперсии

$$D[XY]=D[Z]=M[\mathring{Z}^2]=M[(Z-m_z)^2].$$

Так как величины X, Y независимы, $m_z = m_x m_y$ и

$$D[XY] = M[(XY - m_x m_y)^2] = M[X^2Y^2] - 2m_x m_y M[XY] + m_x^2 m_y^2.$$

При независимых X, Y величины X^2, Y^2 тоже независимы [1]); следовательно,

$$M[X^2Y^2] = M[X^2]\,M[Y^2], \quad M[XY] = m_x m_y$$

и

$$D[XY] = M[X^2]\,M[Y^2] - m_x^2 m_y^2. \qquad (10.2.22)$$

Но $M[X^2]$ есть не что иное, как второй начальный момент величины X, и, следовательно, выражается через дисперсию:

$$M[X^2] = D[X] + m_x^2;$$

аналогично

$$M[Y^2] = D[Y] + m_y^2.$$

Подставляя эти выражения в формулу (10.2.22) и приводя подобные члены, приходим к формуле (10.2.21).

В случае, когда перемножаются центрированные случайные величины (величины с математическими ожиданиями, равными нулю), формула (10.2.21) принимает вид:

$$D[\mathring{X}\mathring{Y}] = D[\mathring{X}]\,D[\mathring{Y}], \qquad (10.2.23)$$

т. е. *дисперсия произведения независимых центрированных случайных величин равна произведению их дисперсий.*

11. Высшие моменты суммы случайных величин

В некоторых случаях приходится вычислять высшие моменты суммы независимых случайных величин. Докажем некоторые относящиеся сюда соотношения.

1) Если величины X, Y независимы, то

$$\mu_3[X+Y] = \mu_3[X] + \mu_3[Y]. \qquad (10.2.24)$$

Доказательство.

$$\begin{aligned}\mu_3[X+Y] &= M[(X+Y-m_x-m_y)^3] = \\ &= M[\{(X-m_x)+(Y-m_y)\}^3] = \\ &= M[(X-m_x)^3] + 3M[(X-m_x)^2(Y-m_y)] + \\ &\quad + 3M[(X-m_x)(Y-m_y)^2] + M[(Y-m_y)^3],\end{aligned}$$

[1]) Можно доказать, что любые функции от независимых случайных величин также независимы.

откуда по теореме умножения математических ожиданий

$$\mu_3[X+Y]=\mu_3[X]+3\mu_2[X]\mu_1[Y]+3\mu_1[Y]\mu_2[X]+\mu_3[Y].$$

Но первый центральный момент μ_1 для любой величины равен нулю; два средних члена обращаются в нуль, и формула (10.2.24) доказана.

Соотношение (10.2.24) методом индукции легко обобщается на произвольное число независимых слагаемых:

$$\mu_3\left[\sum_{i=1}^{n} X_i\right]=\sum_{i=1}^{n}\mu_3[X_i]. \qquad (10.2.25)$$

2) Четвертый центральный момент суммы двух независимых случайных величин выражается формулой

$$\mu_4[X+Y]=\mu_4[X]+\mu_4[Y]+6D_xD_y, \qquad (10.2.26)$$

где D_x, D_y — дисперсии величин X и Y.

Доказательство совершенно аналогично предыдущему.

Методом полной индукции легко доказать обобщение формулы (10.2.26) на произвольное число независимых слагаемых:

$$\mu_4\left[\sum_{i=1}^{n} X_i\right]=\sum_{i=1}^{n}\mu_4[X_i]+6\sum_{i<j} D_{x_i}D_{x_j}. \qquad (10.2.27)$$

Аналогичные соотношения в случае необходимости легко вывести и для моментов более высоких порядков.

12. Сложение некоррелированных случайных векторов

Рассмотрим на плоскости xOy два некоррелированных случайных вектора: вектор $\vec{V}_1$ с составляющими (X_1, Y_1) и вектор $\vec{V}_2$ с составляющими (X_2, Y_2) (рис. 10.2.1).

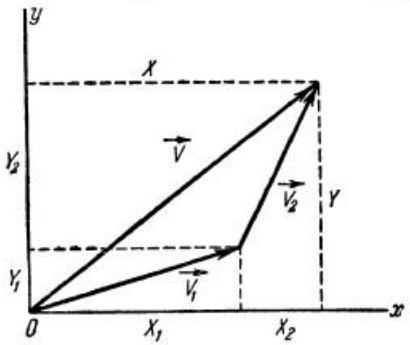


Рис. 10.2.1.

Рассмотрим их векторную сумму:

$$\vec{V}=\vec{V}_1+\vec{V}_2,$$

т. е. вектор с составляющими:

$$X=X_1+X_2,$$
$$Y=Y_1+Y_2.$$

Требуется определить числовые характеристики случайного вектора $\vec{V}$ — математические ожидания m_x, m_y, дисперсии и корреляционный момент составляющих: D_x, D_y, K_{xy}.

По теореме сложения математических ожиданий:

$$m_x = m_{x_1} + m_{x_2};$$
$$m_y = m_{y_1} + m_{y_2}.$$

По теореме сложения дисперсий

$$D_x = D_{x_1} + D_{x_2};$$
$$D_y = D_{y_1} + D_{y_2}.$$

Докажем, что корреляционные моменты также складываются:

$$K_{xy} = K_{x_1y_1} + K_{x_2y_2}, \qquad (10.2.28)$$

где $K_{x_1y_1}$, $K_{x_2y_2}$ — корреляционные моменты составляющих каждого из векторов $\vec{V}_1$ и $\vec{V}_2$.

Доказательство. По определению корреляционного момента:

$$K_{xy} = M[\mathring{X}\mathring{Y}] = M[(\mathring{X}_1 + \mathring{X}_2)(\mathring{Y}_1 + \mathring{Y}_2)] =$$
$$= M[\mathring{X}_1\mathring{Y}_1] + M[\mathring{X}_2\mathring{Y}_1] + M[\mathring{X}_1\mathring{Y}_2] + M[\mathring{X}_2\mathring{Y}_2]. \qquad (10.2.29)$$

Так как векторы $\vec{V}_1$, $\vec{V}_2$ некоррелированны, то два средних члена в формуле (10.2.29) равны нулю; два оставшихся члена представляют собой $K_{x_1y_1}$ и $K_{x_2y_2}$; формула (10.2.28) доказана.

Формулу (10.2.28) иногда называют «теоремой сложения корреляционных моментов».

Теорема легко обобщается на произвольное число слагаемых. Если имеется две некоррелированные системы случайных величин, т. е. два n-мерных случайных вектора:

$\vec{X}$ с составляющими $X_1, X_2, \ldots, X_n$,

$\vec{Y}$ с составляющими $Y_1, Y_2, \ldots, Y_n$,

то их векторная сумма

$$\vec{Z} = \vec{X} + \vec{Y}$$

имеет корреляционную матрицу, элементы которой получаются суммированием элементов корреляционных матриц слагаемых:

$$K_{ij}^{(z)} = K_{ij}^{(x)} + K_{ij}^{(y)}, \qquad (10.2.30)$$

где $K_{ij}^{(z)}$, $K_{ij}^{(x)}$, $K_{ij}^{(y)}$ обозначают соответственно корреляционные моменты величин (Z_i, Z_j); (X_i, X_j), (Y_i, Y_j).

Формула (10.2.30) справедлива как при $i = j$, так и при $i \neq j$.

Действительно, составляющие вектора $\vec{Z}$ равны:

$$Z_1 = X_1 + Y_1;$$
$$Z_2 = X_2 + Y_2;$$
$$\ldots\ldots$$
$$Z_n = X_n + Y_n.$$

По теореме сложения дисперсий

$$D_{z_i} = D_{x_i} + D_{y_i},$$

или в других обозначениях

$$K_{ii}^{(z)} = K_{ii}^{(x)} + K_{ii}^{(y)}.$$

По теореме сложения корреляционных моментов при $i \neq j$

$$K_{ij}^{(z)} = K_{ij}^{(x)} + K_{ij}^{(y)}.$$

В математике *суммой двух матриц* называется матрица, элементы которой получены сложением соответствующих элементов этих матриц. Пользуясь этой терминологией, можно сказать, что *корреляционная матрица суммы двух некоррелированных случайных векторов равна сумме корреляционных матриц слагаемых*:

$$\|K_{ij}^{(z)}\| = \|K_{ij}^{(x)}\| + \|K_{ij}^{(y)}\|. \qquad (10.2.31)$$

Это правило по аналогии с предыдущими можно назвать «теоремой сложения корреляционных матриц».

10.3. Применения теорем о числовых характеристиках

В данном п° мы продемонстрируем применение аппарата числовых характеристик к решению ряда задач. Некоторые из этих задач имеют самостоятельное теоретическое значение и найдут применение в дальнейшем. Другие задачи носят характер примеров и приводятся для иллюстрации выведенных общих формул на конкретном цифровом материале.

Задача 1. Коэффициент корреляции линейно зависимых случайных величин.

Доказать, что если случайные величины X и Y связаны линейной функциональной зависимостью

$$Y = aX + b,$$

то их коэффициент корреляции равен $+1$ или -1, смотря по знаку коэффициента a.

Решение. Имеем:

$$K_{xy} = M[\mathring{X}\mathring{Y}] = M[(X - m_x)(Y - m_y)] =$$
$$= M[(X - m_x)(aX + b - am_x - b)] = aM[(X - m_x)^2] = aD_x,$$

где D_x — дисперсия величины X.

Для коэффициента корреляции имеем выражение:

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}. \qquad (10.3.1)$$

Для определения σ_y найдем дисперсию величины Y:

$$D_y = D[aX + b] = a^2 D_x,$$

$$\sigma_y = |a|\,\sigma_x$$

Подставляя в формулу (10.3.1), имеем:

$$r_{xy} = \frac{aD_x}{|a|\sigma_x^2} = \frac{a}{|a|}.$$

Величина $\frac{a}{|a|}$ равна $+1$, когда a положительно, и -1, когда a отрицательно, что и требовалось доказать.

Задача 2. Границы коэффициента корреляции.

Доказать, что для любых случайных величин

$$|r_{xy}| \leqslant 1.$$

Решение. Рассмотрим случайную величину:

$$Z = \sigma_y X \pm \sigma_x Y,$$

где σ_x, σ_y — средние квадратические отклонения величин X, Y. Определим дисперсию величины Z. По формуле (10.2.13) имеем:

$$D_z = \sigma_y^2 D_x + \sigma_x^2 D_y \pm 2\sigma_x\sigma_y K_{xy},$$

или

$$D_z = 2\sigma_x^2\sigma_y^2 \pm 2\sigma_x\sigma_y K_{xy}.$$

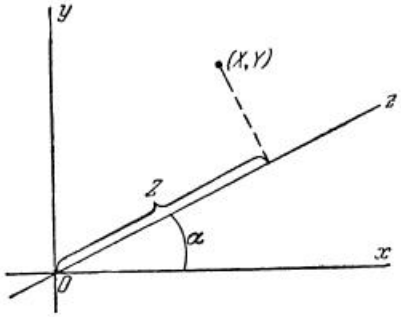


Рис. 10.3.1.

Так как дисперсия любой случайной величины не может быть отрицательна, то

$$2\sigma_x^2\sigma_y^2 \pm 2\sigma_x\sigma_y K_{xy} \geqslant 0,$$

или

$$\sigma_x\sigma_y \pm K_{xy} \geqslant 0,$$

откуда

$$|K_{xy}| \leqslant \sigma_x\sigma_y,$$

а следовательно,

$$|r_{xy}| \leqslant 1,$$

что и требовалось доказать.

Задача 3. Проектирование случайной точки на плоскости на произвольную прямую.

Дана случайная точка на плоскости с координатами (X, Y) (рис. 10.3.1). Спроектируем эту точку на ось Oz, проведенную через начало координат под углом α к оси Ox. Проекция точки (X, Y) на ось Oz также есть случайная точка; ее расстояние Z от начала координат есть случайная величина. Требуется найти математическое ожидание и дисперсию величины Z.

Решение. Имеем:

$$Z = X\cos\alpha + Y\sin\alpha.$$

Так как Z есть линейная функция аргументов X и Y, то

$$m_z = m_x\cos\alpha + m_y\sin\alpha;$$

$$D_z = D_x\cos^2\alpha + D_y\sin^2\alpha + 2K_{xy}\sin\alpha\cos\alpha =$$

$$= D_x\cos^2\alpha + D_y\sin^2\alpha + K_{xy}\sin 2\alpha,$$

где D_x, D_y, K_{xy} — дисперсии и корреляционный момент величин (X, Y).

Переходя к средним квадратическим отклонениям, получим:

$$\sigma_z^2 = \sigma_x^2\cos^2\alpha + \sigma_y^2\sin^2\alpha + r_{xy}\sigma_x\sigma_y\sin 2\alpha. \qquad (10.3.2)$$

В случае некоррелированных случайных величин (при $r_{xy} = 0$)

$$\sigma_z^2 = \sigma_x^2\cos^2\alpha + \sigma_y^2\sin^2\alpha. \qquad (10.3.3)$$

Задача 4. Математическое ожидание числа появлений события при нескольких опытах.

Производится n опытов, в каждом из которых может появиться или не появиться событие A. Вероятность появления события A в i-м опыте равна p_i. Найти математическое ожидание числа появлений события.

Решение. Рассмотрим прерывную случайную величину X — число появлений события во всей серии опытов. Очевидно,

$$X = X_1 + X_2 + \ldots + X_n,$$

где X_1 — число появлений события в первом опыте,

X_2 — число появлений события во втором опыте,

. . . . . .

X_n — число появлений события в n-м опыте,

или, короче,

$$X = \sum_{i=1}^{n} X_i,$$

где X_i — число появлений события в i-м опыте [1]).

Каждая из величин X_i есть прерывная случайная величина с двумя возможными значениями: 0 и 1. Ряд распределения величины X_i имеет вид:

$$\begin{array}{c|c} 0 & 1 \\ \hline q_i & p_i \end{array}, \qquad (10.3.4)$$

где $q_i = 1 - p_i$ — вероятность непоявления события A в i-м опыте.

[1]) Иначе — характеристическая случайная величина события A в i-м опыте.

По теореме сложения математических ожиданий

$$m_x = M[X] = \sum_{i=1}^{n} m_{x_i}, \tag{10.3.5}$$

где m_{x_i} — математическое ожидание величины X_i.

Вычислим математическое ожидание величины X_i. По определению математического ожидания

$$m_{x_i} = 0 \cdot q_i + 1 \cdot p_i = p_i.$$

Подставляя это выражение в формулу (10.3.5), имеем

$$m_x = \sum_{i=1}^{n} p_i, \tag{10.3.6}$$

т. е. *математическое ожидание числа появлений события при нескольких опытах равно сумме вероятностей события в отдельных опытах.*

В частности, когда условия опытов одинаковы и

$$p_1 = p_2 = \ldots = p_n = p,$$

формула (10.3.5) принимает вид

$$m_x = np. \tag{10.3.7}$$

Так как теорема сложения математических ожиданий применима к любым случайным величинам — как зависимым, так и независимым, формулы (10.3.6) и (10.3.7) применимы к любым опытам — зависимым и независимым.

Выведенная теорема часто применяется в теории стрельбы, когда требуется найти среднее число попаданий при нескольких выстрелах — зависимых или независимых. *Математическое ожидание числа попаданий при нескольких выстрелах равно сумме вероятностей попадания при отдельных выстрелах.*

Задача 5. Дисперсия числа появлений события при нескольких независимых опытах.

Производится n независимых опытов, в каждом из которых может появиться событие A, причем вероятность появления события A в i-м опыте равна p_i. Найти дисперсию и среднее квадратическое отклонение числа появлений события A.

Решение. Рассмотрим случайную величину X — число появлений события A. Так же как в предыдущей задаче, представим величину X в виде суммы:

$$X = \sum_{i=1}^{n} X_i,$$

где X_i — число появлений события в i-м опыте.

В силу независимости опытов случайные величины $X_1, X_2, \ldots, X_n$ независимы и к ним применима теорема сложения дисперсий:

$$D_x = \sum_{i=1}^{n} D_{x_i}.$$

Найдем дисперсию случайной величины X_i. Из ряда распределения (10.3.4) имеем:

$$D_{x_i} = (0 - p_i)^2 q_i + (1 - p_i)^2 p_i = p_i q_i,$$

откуда

$$D_x = \sum_{i=1}^{n} p_i q_i, \tag{10.3.8}$$

т. е. дисперсия числа появлений события при нескольких независимых опытах равна сумме вероятностей появления и непоявления события в каждом опыте.

Из формулы (10.3.8) находим среднее квадратическое отклонение числа появлений события A:

$$\sigma_x = \sqrt{\sum_{i=1}^{n} p_i q_i}. \tag{10.3.9}$$

При неизменных условиях опытов, когда $p_1 = p_2 = \ldots = p_n = p$, формулы (10.3.8) и (10.3.9) упрощаются и принимают вид:

$$\left.\begin{aligned} D_x &= npq, \\ \sigma_x &= \sqrt{npq}. \end{aligned}\right\} \tag{10.3.10}$$

Задача 6. Дисперсия числа появлений события при зависимых опытах.

Производится n зависимых опытов, в каждом из которых может появиться событие A, причем вероятность события A в i-м опыте равна p_i ($i = 1, 2, \ldots, n$). Определить дисперсию числа появлений события.

Решение. Для того чтобы решить задачу, снова представим число появлений события X в виде суммы:

$$X = \sum_{i=1}^{n} X_i, \tag{10.3.11}$$

где

$$X_i = \begin{cases} 1, & \text{если в } i\text{-м опыте событие } A \text{ появилось,} \\ 0, & \text{если в } i\text{-м опыте событие } A \text{ не появилось.} \end{cases}$$

Так как опыты зависимы, то нам недостаточно задать вероятности

$$p_1, p_2, \ldots, p_n$$

того, что событие A произойдет в первом, втором, третьем и т. д. опытах. Нужно еще задать характеристики зависимости опытов. Оказывается, для решения нашей задачи достаточно задать вероятности P_{ij} совместного появления события A как в i-м, так и в j-м опыте: $P((X_i=1)(X_j=1)=P_{ij}$. Предположим, что эти вероятности заданы. Применим к выражению (10.3.11) теорему о дисперсии суммы (формулу (10.2.10)):

$$D_x=\sum_{i=1}^{n} D_{x_i}+2\sum_{i<j} K_{ij}, \qquad (10.3.12)$$

где K_{ij} — корреляционный момент величин X_i, X_j:

$$K_{ij}=M[\mathring{X}_i\mathring{X}_j].$$

По формуле (10.2.19)

$$K_{ij}=M[X_iX_j]-m_{x_i}m_{x_j}=M[X_iX_j]-p_ip_j. \qquad (10.3.13)$$

Рассмотрим случайную величину X_iX_j. Очевидно она равна нулю, если хотя бы одна из величин X_i, X_j равна нулю, т. е. хотя бы в одном из опытов (i-м или j-м) событие A не появилось. Для того чтобы величина X_iX_j была равна единице, требуется, чтобы в обоих опытах (i-м и j-м) событие A появилось. Вероятность этого равна P_{ij}. Следовательно,

$$M[X_iX_j]=P_{ij},$$

и

$$K_{ij}=P_{ij}-p_ip_j.$$

Подставляя это выражение в формулу (10.3.12), получим:

$$D_x=\sum_{i=1}^{n} p_iq_i+2\sum_{i<j}(P_{ij}-p_ip_j). \qquad (10.3.14)$$

Формула (10.3.14) и выражает дисперсию числа появлений события при зависимых опытах. Проанализируем структуру этой формулы. Первый член в правой части формулы представляет собой дисперсию числа появлений события при независимых опытах, а второй дает «поправку на зависимость». Если вероятность P_{ij} равна p_ip_j, то эта поправка равна нулю. Если вероятность P_{ij} больше, чем p_ip_j, это значит, что условная вероятность появления события A в j-м опыте при условии, что в i-м опыте оно появилось, больше, чем простая (безусловная) вероятность появления события в j-м опыте p_j (между появлениями события в i-м и j-м опытах имеется положительная корреляция). Если это так для любой пары опытов, то поправочный член в формуле (10.3.14) положителен и дисперсия числа появлений события при зависимых опытах больше, чем при независимых.

Если вероятность P_{ij} меньше, чем $p_i p_j$ (между появлениями события в i-м и j-м опытах существует отрицательная корреляция), то соответствующее слагаемое отрицательно. Если это так для любой пары опытов, то дисперсия числа появлений события при зависимых опытах меньше, чем при независимых.

Рассмотрим частный случай, когда $p_1 = p_2 = \ldots = p_n = p$, $P_{11} = P_{12} = \ldots = P$, т. е. условия всех опытов одинаковы. Формула (10.3.14) принимает вид:

$$D_x = npq + 2\sum_{i<j}(P - p^2) = npq + n(n-1)(P - p^2), \quad (10.3.15)$$

где P — вероятность появления события A сразу в паре опытов (все равно каких).

В этом частном случае особый интерес представляют два подслучая:

1. Появление события A в любом из опытов влечет за собой с достоверностью его появление в каждом из остальных. Тогда $P = p$, и формула (10.3.15) принимает вид:

$$D_x = npq + n(n-1)(p - p^2) = npq + n(n-1)pq = n^2pq.$$

2. Появление события A в любом из опытов исключает его появление в каждом из остальных. Тогда $P = 0$, и формула (10.3.15) принимает вид:

$$D_x = npq - n(n-1)p^2 = np[q - (n-1)p] = np(1 - np).$$

Задача 7. Математическое ожидание числа объектов, приведенных в заданное состояние.

На практике часто встречается следующая задача. Имеется некоторая группа, состоящая из n объектов, по которым осуществляется какое-то воздействие. Каждый из объектов в результате воздействия может быть приведен в определенное состояние S (например, поражен, исправлен, обнаружен, обезврежен и т. п.). Вероятность того, что i-й объект будет приведен в состояние S, равна p_i. Найти математическое ожидание числа объектов, которые в результате воздействия по группе будут приведены в состояние S.

Решение. Свяжем с каждым из объектов случайную величину X_i, которая принимает значения 0 или 1:

$$X_i = \begin{cases} 1, & \text{если } i\text{-й объект приведен в состояние } S, \\ 0, & \text{если } i\text{-й объект не приведен в состояние } S. \end{cases}$$

Случайная величина X — число объектов, приведенных в состояние S, — может быть представлена в виде суммы:

$$X = \sum_{i=1}^{n} X_i.$$

Отсюда, пользуясь теоремой сложения математических ожиданий, получим:

$$m_x = \sum_{i=1}^{n} m_{x_i}.$$

Математическое ожидание каждой из случайных величин X_i известно:

$$m_{x_i} = p_i.$$

Следовательно,

$$m_x = \sum_{i=1}^{n} p_i, \qquad (10.3.16)$$

т. е. математическое ожидание числа объектов, приведенных в состояние S, равно сумме вероятностей перехода в это состояние для каждого из объектов.

Особо подчеркнем, что для справедливости доказанной формулы вовсе не нужно, чтобы объекты переходили в состояние S независимо друг от друга. Формула справедлива для любого вида воздействия.

Задача 8. Дисперсия числа объектов, приведенных в заданное состояние.

Если в условиях предыдущей задачи переход каждого из объектов в состояние S происходит независимо от всех других, то, применяя теорему сложения дисперсий к величине

$$X = \sum_{i=1}^{n} X_i,$$

получим дисперсию числа объектов, приведенных в состояние S:

$$D_x = \sum_{i=1}^{n} D_{x_i} = \sum_{i=1}^{n} p_i q_i, \qquad q_i = 1 - p_i. \qquad (10.3.17)$$

Если же воздействие по объектам производится так, что переходы в состояние S для отдельных объектов зависимы, то дисперсия числа объектов, переведенных в состояние S, выразится формулой (см. задачу 6)

$$D_x = \sum_{i=1}^{n} p_i q_i + 2 \sum_{i<j} (P_{ij} - p_i p_j), \qquad (10.3.18)$$

где P_{ij} — вероятность того, что в результате воздействия i-й и j-й объекты вместе перейдут в состояние S.

Задача 9. Математическое ожидание числа опытов до k-го появления события.

Производится ряд независимых опытов, в каждом из которых может с вероятностью p появиться событие A. Опыты проводятся до тех пор, пока событие A не появится k раз, после чего опыты

прекращаются. Определить математическое ожидание, дисперсию и с. к. о. числа опытов X, которое будет произведено.

Решение. В примере 3 п° 5.7 были определены математическое ожидание и дисперсия числа опытов до первого появления события A:

$$m = \frac{1}{p}, \quad D = \frac{q}{p^2},$$

где p — вероятность появления события в одном опыте,
$q = 1 - p$ — вероятность непоявления.

Рассмотрим случайную величину X — число опытов до k-го появления события A. Ее можно представить в виде суммы:

$$X = X_1 + X_2 + \ldots + X_k,$$

где X_1 — число опытов до первого появления события A,
X_2 — число опытов от первого до второго появления события A (считая второе),

. . . . . .

X_k — число опытов от $(k-1)$-го до k-го появления события A (считая k-е).

Очевидно, величины $X_1, X_2, \ldots, X_k$ независимы; каждая из них распределена по тому же закону, что и первая из них (число опытов до первого появления события) и имеет числовые характеристики

$$m_{x_i} = \frac{1}{p}, \quad D_{x_i} = \frac{q}{p^2}.$$

Применяя теоремы сложения математических ожиданий и дисперсий, получим:

$$\left.\begin{aligned} m_x &= \sum_{i=1}^{k} m_{x_i} = \frac{k}{p}; \\ D_x &= \sum_{i=1}^{k} D_{x_i} = \frac{kq}{p^2}; \\ \sigma_x &= \frac{\sqrt{kq}}{p}. \end{aligned}\right\} \qquad (10.3.19)$$

Задача 10. Средний расход средств до достижения заданного результата.

В предыдущей задаче был рассмотрен случай, когда предпринимается ряд опытов с целью получения вполне определенного результата — k появлений события A, которое в каждом опыте имеет одну и ту же вероятность. Эта задача является частным случаем другой, когда производится ряд опытов с целью достижения любого результата B, вероятность которого с увеличением числа опытов n возрастает по любому закону $P(n)$. Предположим, что на каждый опыт расходуется определенное количество средств a. Требуется

найти математическое ожидание количества средств, которое будет израсходовано.

Решение. Для того чтобы решить задачу, сначала предположим, что число производимых опытов ничем не ограничено, и что они продолжаются и после достижения результата B. Тогда некоторые из этих опытов будут излишними. Условимся называть опыт «необходимым», если он производится при еще не достигнутом результате B, и «излишним», если он производится при уже достигнутом результате B.

Свяжем с каждым (l-м) опытом случайную величину X_l, которая равна нулю или единице в зависимости от того, «необходимым» или «излишним» оказался этот опыт. Положим

$$X_l = \begin{cases} 1, & \text{если опыт оказался «необходимым»,} \\ 0, & \text{если он оказался «излишним».} \end{cases}$$

Рассмотрим случайную величину X — число опытов, которое придется произвести для получения результата B. Очевидно, ее можно представить в виде суммы:

$$X = X_1 + X_2 + \ldots + X_l + \ldots \qquad (10.3.20)$$

Из величин в правой части (10.3.20) первая (X_1) является неслучайной и всегда равна единице (первый опыт всегда «необходим»). Каждая из остальных — случайная величина с возможными значениями 0 и 1. Построим ряд распределения случайной величины X_l ($l > 1$). Он имеет вид:

0	1
$P(l-1)$	$1-P(l-1)$

(10.3.21)

где $P(l-1)$ — вероятность достижения результата B после $l-1$ опытов.

Действительно, если результат B уже был достигнут при предыдущих $l-1$ опытах, то $X_l = 0$ (опыт излишен), если не достигнут, то $X_l = 1$ (опыт необходим).

Найдем математическое ожидание величины X_l. Из ряда распределения (10.3.21) имеем:

$$m_{x_l} = 0 \cdot P(l-1) + 1 \cdot [1 - P(l-1)] = 1 - P(l-1).$$

Нетрудно убедиться, что та же формула будет справедлива и при $l = 1$, так как $P(0) = 0$.

Применим к выражению (10.3.20) теорему сложения математических ожиданий. Получим:

$$m_x = \sum_{l=1}^{\infty} m_{x_l} = \sum_{l=1}^{\infty} [1 - P(l-1)]$$

или, обозначая $l-1=k$,

$$m_x=\sum_{k=0}^{\infty}[1-P(k)]. \tag{10.3.22}$$

Каждый опыт требует затраты средств a. Умножая полученную величину m_x на a, определим среднюю затрату средств на достижение результата B:

$$S_B=a\sum_{k=0}^{\infty}[1-P(k)]. \tag{10.3.23}$$

Данная формула выведена в предположении, что стоимость каждого опыта одна и та же. Если это не так, то можно применить другой прием — представить суммарную затрату средств S_B как сумму затрат на выполнение отдельных опытов, которая принимает два значения: a_l, если l-й опыт «необходим», и нуль, если он «излишен». Средний расход средств S_B представится в виде:

$$S_B=\sum_{k=0}^{\infty}a_k[1-P(k)]. \tag{10.3.24}$$

Задача 11. Математическое ожидание суммы случайного числа случайных слагаемых.

В ряде практических приложений теории вероятностей приходится встречаться с суммами случайных величин, в которых число слагаемых заранее неизвестно, случайно.

Поставим следующую задачу. Случайная величина Z представляет собой сумму Y случайных величин:

$$Z=\sum_{l=1}^{Y}X_l, \tag{10.3.25}$$

причем Y — также случайная величина. Допустим, что нам известны математические ожидания m_{x_l} всех слагаемых:

$$m_{x_l}=M[X_l],$$

и что величина Y не зависит ни от одной из величин X_l.

Требуется найти математическое ожидание величины Z.

Решение. Число слагаемых в сумме есть дискретная случайная величина. Предположим, что нам известен ее ряд распределения:

y_k	1	2	. . .	k	. . .
p_k	p_1	p_2	. . .	p_k	. . .

где p_k — вероятность того, что величина Y приняла значение k. Зафиксируем значение $Y=k$ и найдем при этом условии математи-

ческое ожидание величины Z (условное математическое ожидание):

$$M[Z|k] = \sum_{i=1}^{k} m_{x_i}. \qquad (10.3.26)$$

Теперь применим формулу полного математического ожидания, для чего умножим каждое условное математическое ожидание на вероятность соответствующей гипотезы p_k и сложим:

$$M[Z] = \sum_{k} p_k \sum_{i=1}^{k} m_{x_i}. \qquad (10.3.27)$$

Особый интерес представляет случай, когда все случайные величины $X_1, X_2, \ldots$ имеют одно и то же математическое ожидание:

$$m_{x_1} = m_{x_2} = \ldots = m_x.$$

Тогда формула (10.3.26) принимает вид:

$$M[Z|k] = \sum_{i=1}^{k} m_x = km_x$$

и

$$M[Z] = m_x \sum_{k} kp_k. \qquad (10.3.28)$$

Сумма в выражении (10.3.28) представляет собой не что иное, как математическое ожидание величины Y:

$$m_y = \sum_{k} kp_k.$$

Отсюда

$$m_z = M[Z] = m_x \cdot m_y, \qquad (10.3.29)$$

т. е. математическое ожидание суммы случайного числа случайных слагаемых с одинаковыми средними значениями (если только число слагаемых не зависит от их значений) равно произведению среднего значения каждого из слагаемых на среднее число слагаемых.

Снова отметим, что полученный результат справедлив как для независимых, так и для зависимых слагаемых $X_1, X_2, \ldots$, лишь бы число слагаемых Y не зависело от самих слагаемых.

Ниже мы решим ряд конкретных примеров из разных областей практики, на которых продемонстрируем конкретное применение общих методов оперирования с числовыми характеристиками, вытекающих из доказанных теорем, и специфических приемов, связанных с решенными выше общими задачами.

Пример 1. Монета бросается 10 раз. Определить математическое ожидание и среднее квадратическое отклонение числа X выпавших гербов.

Решение. По формулам (10.3.7) и (10.3.10) найдем:

$$m_x = 10 \cdot 0{,}5 = 5; \quad D_x = 10 \cdot 0{,}5 \cdot 0{,}5 = 2{,}5; \quad \sigma_x = 1{,}58.$$

Пример 2. Производится 5 независимых выстрелов по круглой мишени диаметром 20 *см*. Прицеливание — по центру мишени, систематическая

ошибка отсутствует, рассеивание — круговое, среднее квадратическое отклонение $\sigma = 16$ *см*. Найти математическое ожидание и с. к. о. числа попаданий.

Решение. Вероятность попадания в мишень при одном выстреле вычислим по формуле (9.4.5):

$$p = 1 - e^{-\frac{k^2}{2}} = 1 - e^{-0,625} \approx 0,465.$$

Пользуясь формулами (10.3.7) и (10.3.10), получим:

$$m_x = 5p \approx 2,32; \quad D_x = 5p(1-p) \approx 1,25; \quad \sigma_x \approx 1,12.$$

Пример 3. Производится отражение воздушного налета, в котором участвует 20 летательных аппаратов типа 1 и 30 летательных аппаратов типа 2. Летательные аппараты типа 1 атакуются истребительной авиацией. Число атак, приходящееся на каждый аппарат, подчинено закону Пуассона с параметром $a_1 = 2,5$. Каждой атакой истребителя летательный аппарат типа 1 поражается с вероятностью $p_1 = 0,6$. Летательные аппараты типа 2 атакуются зенитными управляемыми ракетами. Число ракет, направляемых на каждый аппарат, подчинено закону Пуассона с параметром $a_2 = 1,8$, каждая ракета поражает летательный аппарат типа 2 с вероятностью $p_2 = 0,8$. Все аппараты, входящие в состав налета, атакуются и поражаются независимо друг от друга. Найти:

1) математическое ожидание, дисперсию и с. к. о. числа пораженных летательных аппаратов типа 1;

2) математическое ожидание, дисперсию и с. к. о. числа пораженных летательных аппаратов типа 2;

3) математическое ожидание, дисперсию и с. к. о. числа пораженных летательных аппаратов обоих типов.

Решение. Рассмотрим вместо «числа атак» на каждый аппарат типа 1 «число поражающих атак», тоже распределенное по закону Пуассона, но с другим параметром:

$$a_1^* = p_1 a_1 = 0,6 \cdot 2,5 = 1,5.$$

Вероятность поражения каждого из летательных аппаратов типа 1 будет равна вероятности того, что на него придется хотя бы одна поражающая атака:

$$R_1^{(1)} = 1 - e^{-1,5} \approx 0,777.$$

Вероятность поражения каждого из летательных аппаратов типа 2 найдем аналогично:

$$R_1^{(2)} = 1 - e^{-0,8 \cdot 1,8} = 1 - e^{-1,44} \approx 0,763.$$

Математическое ожидание числа пораженных аппаратов типа 1 будет:

$$m_1 = 20 \cdot 0,777 = 15,5.$$

Дисперсия и с. к. о. этого числа:

$$D_1 = 20 \cdot 0,777 \cdot 0,223 = 3,46, \quad \sigma_1 \approx 1,86.$$

Математическое ожидание, дисперсия числа и с. к. о. пораженных аппаратов типа 2:

$$m_2 = 30 \cdot 0,763 = 22,8, \quad D_2 = 5,41, \quad \sigma_2 \approx 2,33.$$

Математическое ожидание, дисперсия и с. к. о. общего числа пораженных аппаратов обоих типов:

$$m = m_1 + m_2 = 38,3, \quad D = D_1 + D_2 = 8,87, \quad \sigma_2 \approx 2,97.$$

Пример 4. Случайные величины X и Y представляют собой элементарные ошибки, возникающие на входе прибора. Они имеют математические ожидания $m_x = -2$ и $m_y = 4$, дисперсии $D_x = 4$ и $D_y = 9$; коэффициент корреляции этих ошибок равен $r_{xy} = -0{,}5$. Ошибка на выходе прибора связана с ошибками на входе функциональной зависимостью:

$$Z = 3X^2 - 2XY + Y^2 - 3.$$

Найти математическое ожидание ошибки на выходе прибора.

Решение.

$$m_z = M[Z] = 3M[X^2] - 2M[XY] + M[Y^2] - 3.$$

Пользуясь связью между начальными и центральными моментами и формулой (10.2.17), имеем:

$$M[X^2] = \alpha_2[X] = D_x + m_x^2 = 8;$$

$$M[Y^2] = \alpha_2[Y] = D_y + m_y^2 = 25;$$

$$M[XY] = m_x m_y + K_{xy} = m_x m_y + r_{xy}\sigma_x\sigma_y = -11,$$

откуда

$$m_z = 3 \cdot 8 - 2 \cdot (-11) + 25 - 3 = 68.$$

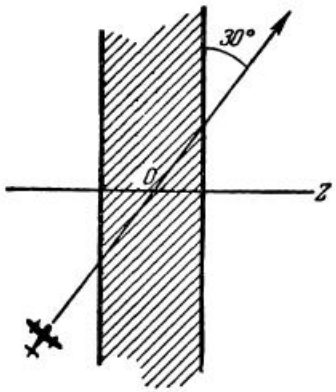


Рис. 10.3.2.

Пример 5. Самолет производит бомбометание по автостраде, ширина которой 30 $м$ (рис. 10.3.2). Направление полета составляет угол 30° с направлением автострады. Прицеливание — по средней линии автострады, систематические ошибки отсутствуют. Рассеивание задано главными вероятными отклонениями: по направлению полета $B_д = 50$ $м$ и в боковом направлении $B_б = 25$ $м$. Найти вероятность попадания в автостраду при сбрасывании одной бомбы.

Решение. Спроектируем случайную точку попадания на ось Oz, перпендикулярную к автостраде и применим формулу (10.3.3). Она, очевидно, остается справедливой, если в нее вместо средних квадратических подставить вероятные отклонения:

$$E_z^2 = E_x^2 \cos 60^\circ + E_y^2 \sin 60^\circ = 50^2 \cdot 0{,}25 + 25^2 \cdot 0{,}75 \approx 1093.$$

Отсюда

$$E_z \approx 33, \quad \sigma_z = \frac{E_z}{0{,}674} \approx 48{,}9.$$

Вероятность попадания в автостраду найдем по формуле (6.3.10):

$$p = 2\Phi^*\left(\frac{15}{48{,}9}\right) - 1 \approx 0{,}238.$$

Примечание. Примененный здесь прием пересчета рассеивания к другим осям пригоден только для вычисления вероятности попадания в область, имеющую вид полосы; для прямоугольника, стороны которого повернуты под углом к осям рассеивания, он уже не годится. Вероятность попадания в каждую из полос, пересечением которых образован прямоугольник, может быть вычислена с помощью этого приема, однако вероятность

попадания в прямоугольник уже не равна произведению вероятностей попадания в полосы, так как эти события зависимы.

Пример 6. Производится наблюдение с помощью системы радиолокационных станций за группой объектов в течение некоторого времени; группа состоит из четырех объектов; каждый из них за время t обнаруживается с вероятностью, равной соответственно:

$$p_1 = 0{,}2, \quad p_2 = 0{,}25, \quad p_3 = 0{,}35, \quad p_4 = 0{,}42.$$

Найти математическое ожидание числа объектов, которые будут обнаружены через время t.

Решение. По формуле (10.3.16) имеем:

$$m_x = \sum_{i=1}^{4} p_i = 0{,}2 + 0{,}25 + 0{,}35 + 0{,}42 = 1{,}22.$$

Пример 7. Предпринимается ряд мероприятий, каждое из которых, если оно состоится, приносит случайный чистый доход X, распределенный по нормальному закону со средним значением $m = 2$ (условных единиц). Число мероприятий за данный период времени случайно и распределено по закону

y_i	1	2	3	4
p_i	0,2	0,3	0,4	0,1

причем не зависит от доходов, приносимых мероприятиями. Определить средний ожидаемый доход за весь период.

Решение. На основе задачи 11 данного п° находим математическое ожидание полного дохода Z:

$$m_z = m_x \cdot m_y,$$

где m_x — средний доход от одного мероприятия, m_y — среднее ожидаемое число мероприятий.

Имеем:

$$m_y = 1 \cdot 0{,}2 + 2 \cdot 0{,}3 + 3 \cdot 0{,}4 + 4 \cdot 0{,}1 = 2{,}4,$$
$$m_x = 2,$$
$$m_z = 2 \cdot 2{,}4 = 4{,}8.$$

Пример 8. Ошибка прибора выражается функцией

$$U = 3Z + 2X - Y - 4, \tag{10.3.30}$$

где X, Y, Z — так называемые «первичные ошибки», представляющие собой систему случайных величин (случайный вектор).

Случайный вектор (X, Y, Z) характеризуется математическими ожиданиями

$$m_x = -4; \quad m_y = 1; \quad m_z = 1$$

и корреляционной матрицей:

$$\begin{Vmatrix} D_x & K_{xy} & K_{xz} \\ & D_y & K_{yz} \\ & & D_z \end{Vmatrix} = \begin{Vmatrix} 2 & 1 & -1 \\ & 3 & 1 \\ & & 4 \end{Vmatrix}.$$

Определить математическое ожидание, дисперсию и среднее квадратическое отклонение ошибки прибора.

Решение. Так как функция (10.3.30) линейна, применяя формулы (10.2.6) и (10.2.13), находим:

$$m_u = 3m_z + 2m_x - m_y - 4 = -10,$$

$$D_u = 3^2 \cdot D_z + 2^2 \cdot D_x + 1^2 \cdot D_y + 2[3 \cdot 2 \cdot K_{xz} + 3(-1) K_{yz} + 2(-1) K_{xy}] = 25,$$

$$\sigma_u = 5.$$

Пример 9. Для обнаружения источника неисправности в вычислительной машине проводятся пробы (тесты). В каждой пробе неисправность независимо от других проб локализуется с вероятностью $p = 0{,}2$. На каждую пробу в среднем уходит 3 минуты. Найти математическое ожидание времени, которое потребуется для локализации неисправности.

Решение. Пользуясь результатом задачи 9 данного п° (математическое ожидание числа опытов до k-го появления события A), полагая $k = 1$, найдем среднее число проб

$$m_x = \frac{1}{p} = \frac{1}{0{,}2} = 5.$$

На эти пять проб потребуется в среднем

$$5 \cdot 3 = 15 \text{ (минут)}.$$

Пример 10. Производится стрельба по резервуару с горючим. Вероятность попадания при каждом выстреле равна 0,3. Выстрелы независимы. При первом попадании в резервуар появляется только течь горючего, при втором попадании горючее воспламеняется. После воспламенения горючего стрельба прекращается. Найти математическое ожидание числа произведенных выстрелов.

Решение. Пользуясь той же формулой, что и в предыдущем примере, найдем математическое ожидание числа выстрелов до 2-го попадания:

$$m_x = \frac{2}{0{,}3} = 6{,}67.$$

Пример 11. Вероятность обнаружения объекта радиолокатором с ростом числа циклов обзора растет по закону:

$$P(n) = 1 - 0{,}8^n,$$

где n — число циклов, начиная с начала наблюдения.

Найти математическое ожидание числа циклов, после которого объект будет обнаружен.

Решение. Пользуясь результатами задачи 10 данного параграфа получим:

$$m_x = \sum_{k=0}^{\infty} [1 - P(k)] = \sum_{k=0}^{\infty} 0{,}8^n = \frac{1}{1 - 0{,}8} = 5.$$

Пример 12. Для того чтобы выполнить определенную задачу по сбору информации, в заданный район высылается несколько разведчиков. Каждый посланный разведчик достигает района назначения с вероятностью 0,7. Для выполнения задачи достаточно наличия в районе трех разведчиков. Один разведчик с задачей вообще справиться не может, а два разведчика выполняют ее с вероятностью 0,4. Обеспечена непрерывная связь с районом, и дополнительные разведчики посылаются, только если задача еще не выполнена.

Требуется найти математическое ожидание числа разведчиков, которые будут посланы.

Решение. Обозначим X — число *прибывших* в район разведчиков, которое оказалось достаточным для выполнения задачи. В задаче 10 данного п° было найдено математическое ожидание числа опытов, которое нужно для того, чтобы достигнуть определенного результата, вероятность которого с увеличением числа опытов возрастает по закону $P(n)$. Это математическое ожидание равно:

$$m_x = \sum_{k=0}^{\infty} [1 - P(k)].$$

В нашем случае:

$$P(0) = 0; \quad P(1) = 0; \quad P(2) = 0{,}4; \quad P(3) = 1;$$
$$P(4) = P(5) = \ldots = 1.$$

Математическое ожидание величины X равно:

$$m_x = M[X] = \sum_{k=0}^{\infty} [1 - P(k)] = 1 + 1 + 0{,}6 = 2{,}6.$$

Итак, для того чтобы задача была выполнена, необходимо, чтобы в район прибыло в среднем 2,6 разведчика.

Теперь решим следующую задачу. Сколько разведчиков придется в среднем выслать в район для того, чтобы их в среднем прибыло m_x?

Пусть послано Y разведчиков. Число прибывших разведчиков можно представить в виде

$$X = Z_1 + Z_2 + \ldots + Z_Y,$$

где случайная величина Z_i принимает значение 1, если i-й разведчик прибыл, и 0, если не прибыл. Величина X есть не что иное, как сумма случайного числа случайных слагаемых (см. задачу 11 данного п°). С учетом этого имеем:

$$M[X] = M[Z_i]\, M[Y],$$

откуда

$$M[Y] = \frac{M[X]}{M[Z_i]} = \frac{2{,}6}{M[Z_i]},$$

но $M[Z_i] = p$, где p — вероятность прибытия отправленного разведчика (в нашем случае $p = 0{,}7$). Величина m_x нами только что найдена и равна 2,6. Имеем:

$$m_y = M[Y] = \frac{2{,}6}{0{,}7} \approx 3{,}71.$$

Пример 13. Радиолокационная станция просматривает область пространства, в которой находится N объектов. За один цикл обзора она обнаруживает каждый из объектов (независимо от других циклов) с вероятностью p. На один цикл требуется время τ. Сколько времени потребуется на то, чтобы из N объектов обнаружить в среднем k?

Решение. Найдем прежде всего математическое ожидание числа обнаруженных объектов после n циклов обзора. За n циклов один (любой) из объектов обнаруживается с вероятностью

$$P_n = 1 - (1 - p)^n,$$

а среднее число объектов, обнаруженных за n циклов, по теореме сложения математических ожиданий (см. задачу 5 данного п°) равно:

$$M[X] = N[1-(1-p)^n].$$

Полагая

$$N[1-(1-p)^n] = k,$$

получим необходимое число циклов n из уравнения

$$1-(1-p)^n = \frac{k}{N},$$

решая которое, найдем:

$$n = \frac{\lg\left(1-\frac{k}{N}\right)}{\lg(1-p)},$$

откуда время, необходимое для обнаружения в среднем k объектов, будет равно:

$$t_k = \tau n = \tau \frac{\lg\left(1-\frac{k}{N}\right)}{\lg(1-p)}.$$

Пример 14. Изменим условия примера 13. Пусть радиолокационная станция ведет наблюдение за областью только до тех пор, пока не будет обнаружено k объектов, после чего наблюдение прекращается или продолжается в новом режиме. Найти математическое ожидание времени, которое для этого понадобится.

Для того чтобы решить эту задачу, недостаточно задаться вероятностью обнаружения одного объекта в одном цикле, а надо еще указать, как растет с увеличением числа циклов вероятность того, что из N объектов будет обнаружено не менее k. Проще всего вычислить эту вероятность, если предположить, что объекты обнаруживаются независимо друг от друга. Сделаем такое допущение и решим задачу.

Решение. При независимых обнаружениях можно наблюдение за N объектами представить как N независимых опытов. После n циклов каждый из объектов обнаруживается с вероятностью

$$P_n = 1-(1-p)^n.$$

Вероятность того, что после n циклов будет обнаружено не менее k объектов из N, найдем по теореме о повторении опытов:

$$R^{(n)}_{k,N} = \sum_{m=k}^{N} C_N^m P_n^m (1-P_n)^{N-m}.$$

Среднее число циклов, после которых будет обнаружено не менее k объектов, определится по формуле (10.3.22):

$$n^{(k)}_{\text{ср}} = \sum_{n=0}^{\infty}\left[1-R^{(n)}_{k,N}\right] = \sum_{n=0}^{N}\left[1-\sum_{m=k}^{N} C_N^m P_n^m (1-P_n)^{N-m}\right].$$

Пример 15. На плоскости xOy случайная точка M с координатами (X, Y) отклоняется от требуемого положения (начало координат) под влиянием трех независимых векторных ошибок $\vec{V}_1$, $\vec{V}_2$ и $\vec{V}_3$. Каждый из векторов характеризуется двумя составляющими:

$$\vec{V}_1(X_1, Y_1), \quad \vec{V}_2(X_2, Y_2), \quad \vec{V}_3(X_3, Y_3)$$

(рис. 10.3.3). Числовые характеристики этих трех векторов равны:

$$m_{x_1}=2, \quad m_{y_1}=-3, \quad \sigma_{x_1}=2, \quad \sigma_{y_1}=3, \quad r_{x_1y_1}=-0{,}3,$$
$$m_{x_2}=-1, \quad m_{y_2}=-2, \quad \sigma_{x_2}=4, \quad \sigma_{y_2}=1, \quad r_{x_2y_2}=0{,}5,$$
$$m_{x_3}=3, \quad m_{y_3}=1, \quad \sigma_{x_3}=2, \quad \sigma_{y_3}=2, \quad r_{x_3y_3}=0{,}2.$$

Найти характеристики суммарной ошибки (вектора, отклоняющего точку M от начала координат).

Решение. Применяя теоремы сложения математических ожиданий, дисперсий и корреляционных моментов, получим:

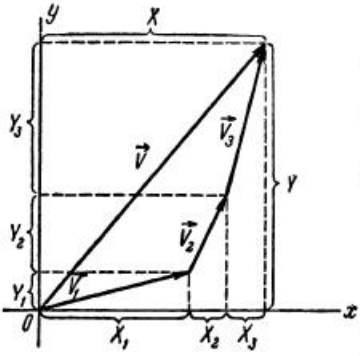


Рис. 10.3.3.

$$m_x = m_{x_1} + m_{x_2} + m_{x_3} = 4,$$
$$m_y = m_{y_1} + m_{y_2} + m_{y_3} = -4,$$
$$\sigma_x^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 = 24, \quad \sigma_x = \sqrt{24} \approx 4{,}90,$$
$$\sigma_y^2 = \sigma_{y_1}^2 + \sigma_{y_2}^2 + \sigma_{y_3}^2 = 14, \quad \sigma_y = \sqrt{14} \approx 3{,}75,$$
$$K_{xy} = K_{x_1y_1} + K_{x_2y_2} + K_{x_3y_3},$$

где

$$K_{x_1y_1} = r_{x_1y_1}\sigma_{x_1}\sigma_{y_1} = -0{,}3 \cdot 2 \cdot 3 = -1{,}8,$$
$$K_{x_2y_2} = r_{x_2y_2}\sigma_{x_2}\sigma_{y_2} = 0{,}5 \cdot 4 \cdot 1 = 2{,}0,$$
$$K_{x_3y_3} = r_{x_3y_3}\sigma_{x_3}\sigma_{y_3} = 0{,}2 \cdot 2 \cdot 2 = 0{,}8,$$

откуда

$$K_{xy} = -1{,}8 + 2{,}0 + 0{,}8 = 1{,}0$$

и

$$r_{xy} = \frac{1{,}0}{4{,}90 \cdot 3{,}75} \approx 0{,}054.$$

Пример 16. Тело, которое имеет форму прямоугольного параллелепипеда с размерами a, b, c, летит в пространстве, беспорядочно вращаясь вокруг центра массы так, что все его ориентации одинаково вероятны. Тело находится в потоке частиц, и среднее число частиц, встречающихся с телом, пропорционально средней площади, которую тело подставляет потоку. Найти математическое ожидание площади проекции тела на плоскость, перпендикулярную направлению его движения.

Решение. Так как все ориентации тела в пространстве одинаково вероятны, то направление плоскости проекций безразлично. Очевидно, площадь проекции тела равна половине суммы проекций всех граней параллелепипеда (так как каждая точка проекции представляет собой проекцию двух точек на поверхности тела). Применяя теорему сложения математических ожиданий и формулу для средней площади проекции плоской фигуры (см. пример 3 п° 10.1), получим:

$$m_S = \frac{ab}{2} + \frac{ac}{2} + \frac{bc}{2} = \frac{S_п}{4},$$

где $S_п$ — полная площадь поверхности параллелепипеда.

Заметим, что выведенная формула справедлива не только для параллелепипеда, но и для любого выпуклого тела: средняя площадь проекции такого тела при беспорядочном вращении равна одной четверти полной его поверхности. Рекомендуем читателю в качестве упражнения доказать это положение.

Пример 17. На оси абсцисс Ox движется случайным образом точка x по следующему закону. В начальный момент она находится в начале координат и начинает двигаться с вероятностью $\frac{1}{2}$ вправо и с вероятностью $\frac{1}{2}$ влево. Пройдя единичное расстояние, точка с вероятностью p продолжает двигаться в том же направлении, а с вероятностью $q=1-p$ меняет его на противоположное. Пройдя единичное расстояние, точка снова с вероятностью p продолжает движение в том направлении, в котором двигалась, а с вероятностью $1-p$ меняет его на противоположное и т. д.

В результате такого случайного блуждания по оси абсцисс точка x после n шагов займет случайное положение, которое мы обозначим X_n. Требуется найти характеристики случайной величины X_n: математическое ожидание и дисперсию.

Решение. Прежде всего, из соображений симметрии задачи ясно, что $M[X_n]=0$. Чтобы найти $D[X_n]$, представим X_n в виде суммы n слагаемых:

$$X_n=U_1+U_2+\ldots+U_n=\sum_{i=1}^{n}U_i, \tag{10.3.31}$$

где U_i — расстояние, пройденное точкой на i-м шаге, т. е. $+1$, если точка двигалась на этом шаге вправо, и -1, если она двигалась влево.

По теореме о дисперсии суммы (см. формулу (10.2.10)) имеем:

$$D[X_n]=\sum_{i=1}^{n}D[U_i]+2\sum_{i<j}K_{u_iu_j}.$$

Ясно, что $D[U_i]=1$, так как величина U_i принимает значения $+1$ и -1 с одинаковой вероятностью (из тех же соображений симметрии). Найдем корреляционные моменты

$$K_{u_iu_j}=M[\mathring{U}_i\mathring{U}_j]=M[U_iU_j].$$

Начнем со случая $j=i+1$, когда величины U_i и U_j стоят рядом в сумме (10.3.31). Ясно, что U_iU_{i+1} принимает значение $+1$ с вероятностью p и значение -1 с вероятностью q. Имеем:

$$K_{u_iu_{i+1}}=M[U_iU_{i+1}]=1\cdot p+(-1)\cdot q=p-q.$$

Рассмотрим, далее, случай $j=i+2$. В этом случае произведение U_iU_j равно $+1$, если оба перемещения — на i-м и $i+2$-м шаге — происходят в одном и том же направлении. Это может произойти двумя способами. Или точка x все три шага — i-й, $(i+1)$-й и $(i+2)$-й — двигалась в одном и том же направлении, или же она дважды изменила за эти три шага свое направление. Найдем вероятность того, что $U_iU_{i+2}=1$:

$$P(U_iU_{i+2}=1)=P\big((U_iU_{i+1}=1)(U_{i+1}U_{i+2}=1)\big)+$$
$$+P\big((U_iU_{i+1}=-1)(U_{i+1}U_{i+2}=-1)\big)=p^2+q^2.$$

Найдем теперь вероятность того, что $U_iU_{i+2}=-1$. Это тоже может произойти двумя способами: или точка изменила свое направление при переходе от i-го шага к $(i+1)$-му, а при переходе от $(i+1)$-го шага к $(i+2)$-му сохранила его, или наоборот. Имеем:

$$P(U_iU_{i+2}=1)=P\big((U_iU_{i+1}=-1)(U_{i+1}U_{i+2}=1)\big)+$$
$$+P\big((U_iU_{i+1}=1)(U_{i+1}U_{i+2}=-1)\big)=2pq.$$

Таким образом, величина U_iU_{i+2} имеет два возможных значения $+1$ и -1, которые она принимает с вероятностями соответственно p^2+q^2 и $2pq$. Ее математическое ожидание равно:

$$K_{u_iu_{i+2}} = M[U_iU_{i+2}] = p^2+q^2-2pq = (p-q)^2.$$

Легко доказать по индукции, что для любого расстояния k между шагами в ряду $U_1, U_2, \ldots, U_n$ справедливы формулы:

$$P(U_iU_{i+k}=1) = p^k + C_k^2 p^{k-2}q^2 + C_k^4 p^{k-4}q^4 + \ldots,$$

$$P(U_iU_{i+k}=-1) = C_k^1 p^{k-1}q + C_k^3 p^{k-3}q^3 + \ldots,$$

и, следовательно,

$$K_{u_iu_{i+k}} = (p-q)^k.$$

Таким образом, корреляционная матрица системы случайных величин $U_1, U_2, \ldots, U_n$ будет иметь вид:

$$\|K_{u_iu_j}\| = \begin{Vmatrix} 1 & p-q & (p-q)^2 & \ldots & (p-q)^{n-1} \\ & 1 & p-q & \ldots & (p-q)^{n-2} \\ & & 1 & \ldots & (p-q)^{n-3} \\ & & & \ldots & \ldots \\ & & & & 1 \end{Vmatrix}.$$

Дисперсия случайной величины X_n будет равна:

$$D[X_n] = \sum_{i=1}^{n} D[U_i] + 2\sum_{i<j} K_{u_iu_j} = n + 2\sum_{i<j}(p-q)^{j-i},$$

или же, производя суммирование элементов, стоящих на одном расстоянии от главной диагонали,

$$D[X_n] = n + 2\sum_{k=1}^{n-1}(n-k)(p-q)^k.$$

Пример 18. Найти асимметрию биномиального распределения

$$P(X=m) = C_n^m p^m q^{n-m} \qquad (q=1-p). \tag{10.3.32}$$

Решение. Известно, что биномиальное распределение (10.3.32) представляет собой распределение числа появлений в n независимых опытах некоторого события, которое в одном опыте имеет вероятность p. Представим случайную величину X — число появлений события в n опытах — как сумму n случайных величин:

$$X = \sum_{i=1}^{n} X_i,$$

где

$$X_i = \begin{cases} 1, & \text{если в } i\text{-м опыте событие появилось,} \\ 0, & \text{если в } i\text{-м опыте событие не появилось.} \end{cases}$$

По теореме сложения третьих центральных моментов

$$\mu_3[X] = \sum_{i=1}^{n} \mu_3[X_i]. \tag{10.3.33}$$

Найдем третий центральный момент случайной величины X_i. Она имеет ряд распределения

0	1
q	p

Третий центральный момент величины X_i равен:

$$(0-p)^3 q + (1-p)^3 p = -p^3 q + q^3 p = pq(q-p).$$

Подставляя в (10.3.33), получим:

$$\mu_3[X] = \sum_{i=1}^{n} pq(q-p) = npq(q-p).$$

Чтобы получить асимметрию, нужно разделить третий центральный момент величины X на куб среднего квадратического отклонения:

$$Sk = \frac{npq(q-p)}{(npq)^{\frac{3}{2}}} = \frac{q-p}{\sqrt{npq}}.$$

Пример 19. Имеется n положительных, одинаково распределенных независимых случайных величин:

$$X_1, X_2, \ldots, X_n.$$

Найти математическое ожидание случайной величины

$$Z_1 = \frac{X_1}{X_1 + X_2 + \ldots + X_n}.$$

Решение. Ясно, что математическое ожидание величины Z_1 существует, так как она заключена между нулем и единицей. Кроме того, легко видеть, что закон распределения системы величин $(X_1, X_2, \ldots, X_n)$, каков бы он ни был, симметричен относительно своих переменных, т. е. не меняется при любой их перестановке. Рассмотрим случайные величины:

$$Z_1 = \frac{X_1}{X_1 + X_2 + \ldots + X_n}, \quad Z_2 = \frac{X_2}{X_1 + X_2 + \ldots + X_n}, \ldots$$

$$\ldots, Z_n = \frac{X_n}{X_1 + X_2 + \ldots + X_n}.$$

Очевидно, их закон распределения тоже должен обладать свойством симметрии, т. е. не меняться при замене одного аргумента любым другим и наоборот. Отсюда, в частности, вытекает, что

$$M[Z_1] = M[Z_2] = \ldots = M[Z_n].$$

Вместе с тем нам известно, что в сумме случайные величины $Z_1, Z_2, \ldots, Z_n$ образуют единицу, следовательно, по теореме сложения математических ожиданий,

$$M[Z_1] + M[Z_2] + \ldots + M[Z_n] = M[1] = 1,$$

откуда

$$M[Z_1] = M[Z_2] = \ldots = M[Z_n] = \frac{1}{n}.$$

ГЛАВА 11

ЛИНЕАРИЗАЦИЯ ФУНКЦИЙ

11.1. Метод линеаризации функций случайных аргументов

В предыдущей главе мы познакомились с весьма удобным математическим аппаратом теории вероятностей — с аппаратом числовых характеристик. Этот аппарат во многих случаях позволяет находить числовые характеристики функций случайных величин (в первую очередь — математическое ожидание и дисперсию) по числовым характеристикам аргументов, оставляя совершенно в стороне законы распределения. Такие методы непосредственного определения числовых характеристик применимы главным образом к линейным функциям.

На практике очень часто встречаются случаи, когда исследуемая функция случайных величин хотя и не является строго линейной, но практически мало отличается от линейной и при решении задачи может быть приближенно заменена линейной. Это связано с тем, что во многих практических задачах случайные изменения фигурирующих в них величин выступают как незначительные «погрешности», накладывающиеся на основную закономерность. Вследствие сравнительной малости этих погрешностей обычно фигурирующие в задаче функции, не будучи линейными во всем диапазоне изменения своих аргументов, оказываются почти линейными в узком диапазоне их случайных изменений.

Действительно, из математики известно, что любая непрерывная дифференцируемая функция в достаточно узких пределах изменения аргументов может быть приближенно заменена линейной (линеаризована). Ошибка, возникающая при этом, тем меньше, чем уже границы изменения аргументов и чем ближе функция к линейной. Если область практически возможных значений случайных аргументов настолько мала, что в этой области функция может быть с достаточной для практики точностью линеаризована, то, заменив нелинейную функцию линейной, можно применить к последней тот аппарат числовых характеристик, который разработан для линейных функций. Зная числовые характеристики аргументов, можно будет найти числовые характеристики функции. Конечно, при этом мы получим лишь приближен-

ное решение задачи, но в большинстве случаев точного решения и не требуется.

При решении практических задач, в которых случайные факторы сказываются в виде незначительных возмущений, налагающихся на основные закономерности, линеаризация почти всегда оказывается возможной именно в силу малости случайных возмущений.

Рассмотрим, например, задачу внешней баллистики о движении центра массы снаряда. Дальность полета снаряда X определяется как некоторая функция условий стрельбы — угла бросания θ_0, начальной скорости v_0 и баллистического коэффициента c:

$$X = \varphi(\theta_0, v_0, c). \qquad (11.1.1)$$

Функция (11.1.1) нелинейна, если рассматривать ее на всем диапазоне изменения аргументов. Поэтому, когда речь идет о решении основной задачи внешней баллистики, функция (11.1.1) выступает как нелинейная и никакой линеаризации не подлежит. Однако есть задачи, в которых такие функции линеаризуются; это — задачи, связанные с исследованием ошибок или погрешностей. Пусть нас интересует случайная ошибка в дальности полета снаряда X, связанная с наличием ряда случайных факторов: неточностью установки угла θ_0, колебаниями ствола при выстреле, баллистической неоднородностью снарядов, различными весами зарядов и т. д. Тогда мы зафиксируем определенные номинальные условия стрельбы и будем рассматривать случайные отклонения от этих условий. Диапазон таких случайных изменений, как правило, невелик, и функция φ, не будучи линейной во всей области изменения своих аргументов, может быть линеаризована в малой области их случайных изменений.

Метод линеаризации функций, зависящих от случайных аргументов, находит самое широкое применение в различных областях техники. Очень часто, получив решение задачи обычными методами «точных наук», желательно оценить возможные погрешности в этом решении, связанные с влиянием не учтенных при решении задачи случайных факторов. В этом случае, как правило, задача оценки погрешности успешно решается методом линеаризации, так как случайные изменения фигурирующих в задаче величин обычно невелики. Если бы это было не так, и случайные изменения аргументов выходили за пределы области примерной линейности функций, следовало бы считать техническое решение неудовлетворительным, так как оно содержало бы слишком большой элемент неопределенности.

11.2. Линеаризация функции одного случайного аргумента

На практике необходимость в линеаризации функции одного случайного аргумента встречается сравнительно редко: обычно приходится учитывать совокупное влияние нескольких случайных факторов. Однако из методических соображений удобно начать с этого наиболее

простого случая. Пусть имеется случайная величина X и известны ее числовые характеристики: математическое ожидание m_x и дисперсия D_x.

Допустим, что практически возможные значения случайной величины X ограничены пределами α, β, т. е.

$$P(\alpha < X < \beta) \approx 1.$$

Имеется другая случайная величина Y, связанная с X функциональной зависимостью:

$$Y = \varphi(X)\ ^{1}), \tag{11.2.1}$$

причем функция φ хотя не является линейной, но мало отличается от линейной на участке (α, β).

Требуется найти числовые характеристики величины Y — математическое ожидание m_y и дисперсию D_y.

Рассмотрим кривую $y = \varphi(x)$ на участке (α, β) (рис. 11.2.1) и заменим ее приближенно касательной, проведенной в точке M с абсциссой m_x. Уравнение касательной имеет вид:

$$y = \varphi(m_x) + \varphi'(m_x)(x - m_x). \tag{11.2.2}$$

Предположим, что интервал практически возможных значений аргумента (α, β) настолько узок, что в пределах этого интервала кривая и касательная различаются мало, так что участок кривой практически можно заменить участком касательной; короче, на участке (α, β) функция $y = \varphi(x)$ почти линейна. Тогда случайные величины X и Y приближенно связаны линейной зависимостью:

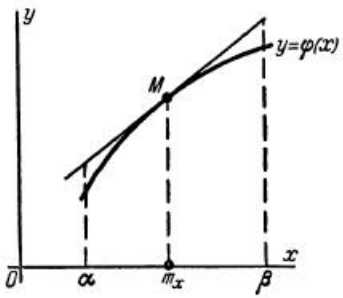


Рис. 11.2.1.

$$Y = \varphi(m_x) + \varphi'(m_x)(X - m_x),$$

или, обозначая $X - m_x = \mathring{X}$,

$$Y = \varphi(m_x) + \varphi'(m_x)\mathring{X}. \tag{11.2.3}$$

К линейной функции (11.2.3) можно применить известные приемы определения числовых характеристик линейных функций (см. п° 10.2). Математическое ожидание этой линейной функции найдем, подставляя в ее выражение (11.2.3) математическое ожидание аргумента $\mathring{X}$, равное нулю. Получим:

$$m_y = \varphi(m_x). \tag{11.2.4}$$

---

[1]) Функцию φ на участке (α, β) предполагаем непрерывной и дифференцируемой.

Дисперсия величины Y определится по формуле

$$D_y = [\varphi'(m_x)]^2 D_x. \tag{11.2.5}$$

Переходя к среднему квадратическому отклонению, имеем:

$$\sigma_y = |\varphi'(m_x)| \sigma_x. \tag{11.2.6}$$

Формулы (11.2.4), (11.2.5), (11.2.6), разумеется, являются приближенными, поскольку приближенной является и сама замена нелинейной функции линейной.

Таким образом, мы решили поставленную задачу и пришли к следующим выводам.

Чтобы найти математическое ожидание почти линейной функции, нужно в выражение функции вместо аргумента подставить его математическое ожидание. Чтобы найти дисперсию почти линейной функции, нужно дисперсию аргумента умножить на квадрат производной функции в точке, соответствующей математическому ожиданию арумента.

11.3. Линеаризация функции нескольких случайных аргументов

Имеется система n случайных величин:

$$(X_1, X_2, \ldots, X_n)$$

и заданы числовые характеристики системы: математические ожидания

$$m_{x_1}, m_{x_2}, \ldots, m_{x_n}$$

и корреляционная матрица

$$\|K_{ij}\| = \begin{Vmatrix} K_{11} & K_{12} \ldots K_{1n} \\ & K_{22} \ldots K_{2n} \\ & \ldots \\ & K_{nn} \end{Vmatrix}.$$

Случайная величина Y есть функция аргументов $X_1, X_2, \ldots, X_n$:

$$Y = \varphi(X_1, X_2, \ldots, X_n), \tag{11.3.1}$$

причем функция φ не линейна, но мало отличается от линейной в области практически возможных значений всех аргументов (короче, «почти линейная» функция). Требуется приближенно найти числовые характеристики величины Y — математическое ожидание m_y и дисперсию D_y.

Для решения задачи подвергнем линеаризации функцию

$$y = \varphi(x_1, x_2, \ldots, x_n). \tag{11.3.2}$$

В данном случае нет смысла пользоваться геометрической интерпретацией, так как за пределами трехмерного пространства она уже не обладает преимуществами наглядности. Однако качественная сторона вопроса остается совершенно той же, что и в предыдущем п°.

Рассмотрим функцию $y=\varphi(x_1, x_2, \ldots, x_n)$ в достаточно малой окрестности точки $m_{x_1}, m_{x_2}, \ldots, m_{x_n}$. Так как функция в этой окрестности почти линейна, ее можно приближенно заменить линейной. Это равносильно тому, чтобы в разложении функции в ряд Тейлора около точки $(m_{x_1}, m_{x_2}, \ldots, m_{x_n})$ сохранить только члены первого порядка, а все высшие отбросить:

$$y=\varphi(x_1, x_2, \ldots, x_n)\approx$$

$$\approx\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n})+\sum_{i=1}^{n}\varphi'_{x_i}(m_{x_1}, m_{x_2}, \ldots, m_{x_n})(x_i-m_{x_i}).$$

Значит, и зависимость (11.3.1) между случайными величинами можно приближенно заменить линейной зависимостью:

$$Y=\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n})+$$

$$+\sum_{i=1}^{n}\varphi'_{x_i}(m_{x_1}, m_{x_2}, \ldots, m_{x_n})(X_i-m_{x_i}). \qquad (11.3.3)$$

Введем для краткости обозначение:

$$\varphi'_{x_i}(m_{x_1}, m_{x_2}, \ldots, m_{x_n})=\left(\frac{\partial\varphi}{\partial x_i}\right)_m.$$

Учитывая, что $X_i-m_{x_i}=\mathring{X}_i$, перепишем формулу (11.3.3) в виде:

$$Y=\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n})+\sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m\mathring{X}_i. \qquad (11.3.4)$$

К линейной функции (11.3.4) применим способы определения числовых характеристик линейных функций, выведенные в п° 10.2. Имея в виду, что центрированные аргументы $(\mathring{X}_1, \mathring{X}_2, \ldots, \mathring{X}_n)$ имеют математические ожидания, равные нулю, и ту же корреляционную матрицу $\|K_{ij}\|$, получим:

$$m_y=\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n}), \qquad (11.3.5)$$

$$D_y=\sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m^2 D_{x_i}+2\sum_{i<j}\left(\frac{\partial\varphi}{\partial x_i}\right)_m\left(\frac{\partial\varphi}{\partial x_j}\right)_m K_{ij}. \qquad (11.3.6)$$

Переходя в последней формуле от дисперсий к средним квадратическим отклонениям, получим:

$$\sigma_y^2 = \sum_{l=1}^{n} \left(\frac{\partial \varphi}{\partial x_l}\right)_m^2 \sigma_{x_l}^2 + 2 \sum_{l<j} \left(\frac{\partial \varphi}{\partial x_l}\right)_m \left(\frac{\partial \varphi}{\partial x_j}\right)_m r_{lj} \sigma_{x_l} \sigma_{x_j}, \quad (11.3.7)$$

где r_{lj} — коэффициент корреляции величин X_l, X_j.

Особенно простой вид принимает формула (11.3.7), когда величины $X_1, X_2, \ldots, X_n$ не коррелированы, т. е. $r_{lj} = 0$ при $l \neq j$. В этом случае

$$\sigma_y^2 = \sum_{l=1}^{n} \left(\frac{\partial \varphi}{\partial x_l}\right)_m^2 \sigma_{x_l}^2. \quad (11.3.8)$$

Формулы типа (11.3.7) и (11.3.8) находят широкое применение в различных прикладных вопросах: при исследовании ошибок разного вида приборов и механизмов, а также при анализе точности стрельбы и бомбометания.

Пример 1. Относ бомбы X (рис. 11.3.1) выражается приближенной аналитической формулой:

$$X = v_0 \sqrt{\frac{2H}{g}} (1 - 1{,}8 \cdot 10^{-5} cH), \quad (11.3.9)$$

где v_0 — скорость самолета (*м/сек*), H — высота сбрасывания (*м*), c — баллистический коэффициент.

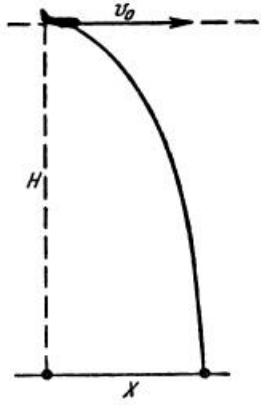


Рис. 11.3.1.

Высота H определяется по высотомеру, скорость самолета v_0 — по указателю скорости, баллистический коэффициент c принимается его номинальным значением $c = 1{,}00$. Высотомер показывает 4000 *м*, указатель скорости 150 *м/сек*. Показания высотомера характеризуются систематической ошибкой $+50$ *м* и средним квадратическим отклонением $\sigma_H = 40$ *м*; показания указателя скорости — систематической ошибкой -2 *м/сек* и средним квадратическим отклонением 1 *м/сек*; разброс возможных значений баллистического коэффициента c, обусловленный неточностью изготовления бомбы, характеризуется средним квадратическим отклонением $\sigma_c = 0{,}05$. Ошибки приборов независимы между собой.

Найти систематическую ошибку и среднее квадратическое отклонение точки падения бомбы вследствие неточности в определении параметров H, v_0 и c. Определить, какой из этих факторов оказывает наибольшее влияние на разброс точки падения бомбы.

Решение. Величины H, v_0 и c представляют собой некоррелированные случайные величины с числовыми характеристиками:

$$m_H = 3950 \text{ м}; \qquad \sigma_H = 40 \text{ м};$$
$$m_{v_0} = 152 \text{ м/сек}; \quad \sigma_{v_0} = 1 \text{ м/сек};$$
$$m_c = 1{,}00; \qquad \sigma_c = 0{,}05.$$

Так как диапазон возможных изменений случайных аргументов сравнительно невелик, для решения задачи можно применить метод линеаризации.

Подставляя в формулу (11.3.9) вместо величин H, v_0 и c их математические ожидания, найдем математическое ожидание величины X:

$$m_x = 4007 \ (м).$$

Для сравнения вычислим номинальное значение:

$$X_{ном} = 150 \sqrt{\frac{8000}{9{,}81}} (1 - 1{,}8 \cdot 10^{-5} \cdot 1 \cdot 4000) = 3975 \ (м).$$

Разность между математическим ожиданием и номинальным значением и представляет собой систематическую ошибку точки падения:

$$\Delta_x = m_x - X_{ном} = 4007 - 3975 = 32 \ (м).$$

Для определения дисперсии величины X вычислим частные производные:

$$\frac{\partial X}{\partial H} = \frac{v_0}{\sqrt{2Hg}} (1 - 1{,}8 \cdot 10^{-5} cH) - v_0 \sqrt{\frac{2H}{g}} \cdot 1{,}8 \cdot 10^{-5} c,$$

$$\frac{\partial X}{\partial v_0} = \sqrt{\frac{2H}{g}} (1 - 1{,}8 \cdot 10^{-5} cH),$$

$$\frac{\partial X}{\partial c} = - v_0 \sqrt{\frac{2H}{g}}$$

и подставим в эти выражения вместо каждого аргумента его математическое ожидание:

$$\left(\frac{\partial X}{\partial H}\right)_m = 0{,}429; \qquad \left(\frac{\partial X}{\partial v_0}\right)_m = 26{,}4; \qquad \left(\frac{\partial X}{\partial c}\right)_m = -307.$$

По формуле (11.3.8) вычислим среднее квадратическое отклонение величины X:

$$\sigma_x^2 = 0{,}429^2 \cdot 40^2 + 26{,}4^2 \cdot 1^2 + 307^2 \cdot 0{,}05^2 = 294{,}4 + 697{,}0 + 235{,}5 = 1126{,}9,$$

откуда

$$\sigma_x \approx 33{,}6 \ (м).$$

Сравнивая слагаемые, образующие σ_x^2, приходим к заключению, что наибольшее из них (697,0) обусловлено наличием ошибок в скорости v_0; следовательно, в данных условиях из рассмотренных случайных факторов, обусловливающих разброс точки падения бомбы, наиболее существенной является ошибка указателя скорости.

Пример 2. Абсцисса точки попадания (в метрах) при стрельбе по самолету выражается формулой

$$X = X_н + 1{,}8\omega D + X_б, \tag{11.3.10}$$

где $X_н$ — ошибка наводки ($м$), ω — угловая скорость цели ($рад/сек$), D — дальность стрельбы ($м$), $X_б$ — ошибка, связанная с баллистикой снаряда ($м$).

Величины $X_н$, ω, D, $X_б$ представляют собой случайные величины с математическими ожиданиями:

$$m_{x_н} = 0; \quad m_\omega = 0{,}1; \quad m_д = 1000; \quad m_{x_б} = 0$$

и средними квадратическими отклонениями:

$$\sigma_{x_н} = 4; \quad \sigma_\omega = 0{,}005; \quad \sigma_д = 50; \quad \sigma_{x_б} = 3.$$

Нормированная корреляционная матрица системы ($X_н$, ω, D, $X_б$) (т. е. матрица, составленная из коэффициентов корреляции) имеет вид:

$$\|r_{ij}\| = \begin{Vmatrix} 1 & 0{,}5 & 0{,}3 & 0 \\ & 1 & 0{,}4 & 0 \\ & & 1 & 0 \\ & & & 1 \end{Vmatrix}.$$

Требуется найти математическое ожидание и среднее квадратическое отклонение величины X.

Р е ш е н и е. Подставляя в формулу (11.3.10) математические ожидания аргументов, имеем:

$$m_x = 1{,}8 \cdot 0{,}1 \cdot 1000 = 180 \ (м).$$

Для определения среднего квадратического отклонения величины X найдем частные производные:

$$\left(\frac{\partial X}{\partial X_н}\right)_m = 1; \quad \left(\frac{\partial X}{\partial \omega}\right)_m = 1{,}8 m_д = 1800;$$

$$\left(\frac{\partial X}{\partial D}\right)_m = 1{,}8 m_\omega = 0{,}18; \quad \left(\frac{\partial X}{\partial X_б}\right)_m = 1.$$

Применяя формулу (11.3.7), имеем:

$$\sigma_x^2 = (1 \cdot 4)^2 + (1800 \cdot 0{,}005)^2 + (0{,}18 \cdot 50)^2 + (1 \cdot 3)^2 + 2 \cdot 1 \cdot 1800 \cdot 0{,}5 \cdot 4 \cdot 0{,}005 +$$
$$+ 2 \cdot 1 \cdot 0{,}18 \cdot 0{,}3 \cdot 4 \cdot 50 + 2 \cdot 1800 \cdot 0{,}18 \cdot 0{,}4 \cdot 0{,}005 \cdot 50 =$$
$$= 16 + 81 + 81 + 9 + 36 + 21{,}6 + 64{,}8 = 309{,}4,$$

откуда

$$\sigma_x \approx 17{,}6 \ (м).$$

11.4. Уточнение результатов, полученных методом линеаризации

В некоторых задачах практики возникает сомнение в применимости метода линеаризации в связи с тем, что диапазон изменений случайных аргументов не настолько мал, чтобы в его пределах функция могла быть с достаточной точностью линеаризована.

В этих случаях для проверки применимости метода линеаризации и для уточнения полученных результатов может быть применен метод, основанный на сохранении в разложении функции не только линейных членов, но и некоторых последующих членов более высоких порядков и оценке погрешностей, связанных с этими членами.

Для того чтобы пояснить этот метод, рассмотрим сначала наиболее простой случай функции одного случайного аргумента. Случайная величина Y есть функция случайного аргумента X:

$$Y = \varphi(X), \tag{11.4.1}$$

причем функция φ сравнительно мало отличается от линейной на участке практически возможных значений аргумента X, но все же

отличается настолько, что возникает сомнение в применимости метода линеаризации. Для проверки этого обстоятельства применим более точный метод, а именно: разложим функцию φ в ряд Тейлора в окрестности точки m_x и сохраним в разложении первые три члена:

$$y = \varphi(x) \approx \varphi(m_x) + \varphi'(m_x)(x - m_x) + \frac{1}{2}\varphi''(m_x)(x - m_x)^2. \quad (11.4.2)$$

Та же формула будет, очевидно, приближенно связывать случайные величины Y и X:

$$Y = \varphi(m_x) + \varphi'(m_x)(X - m_x) + \frac{1}{2}\varphi''(m_x)(X - m_x)^2 =$$

$$= \varphi(m_x) + \varphi'(m_x)\mathring{X} + \frac{1}{2}\varphi''(m_x)\mathring{X}^2. \quad (11.4.3)$$

Пользуясь выражением (11.4.3), найдем математическое ожидание и дисперсию величины Y. Применяя теоремы о числовых характеристиках, имеем:

$$m_y = \varphi(m_x) + \frac{1}{2}\varphi''(m_x) M[\mathring{X}^2] = \varphi(m_x) + \frac{1}{2}\varphi''(m_x) D_x. \quad (11.4.4)$$

По формуле (11.4.4) можно найти уточненное значение математического ожидания и сравнить его с тем значением $\varphi(m_x)$, которое получается методом линеаризации; поправкой, учитывающей нелинейность функции, является второй член формулы (11.4.4).

Определяя дисперсию правой и левой части формулы (11.4.3), имеем:

$$D_y = [\varphi'(m_x)]^2 D_x + \frac{1}{4}[\varphi''(m_x)]^2 D[\mathring{X}^2] +$$

$$+ \varphi'(m_x)\varphi''(m_x) K[\mathring{X}, \mathring{X}^2], \quad (11.4.5)$$

где $K[\mathring{X}, \mathring{X}^2]$ — корреляционный момент величин $\mathring{X}$, $\mathring{X}^2$.

Выразим входящие в формулу (11.4.5) величины через центральные моменты величины X:

$$D[\mathring{X}^2] = M[\mathring{X}^4] - [M[\mathring{X}^2]]^2 = \mu_4[X] - D_x^2,$$

$$K[\mathring{X}, \mathring{X}^2] = M[\mathring{X}\{\mathring{X}^2 - M[\mathring{X}^2]\}] = \mu_3[X].$$

Окончательно имеем:

$$D_y = [\varphi'(m_x)]^2 D_x + \frac{1}{4}[\varphi''(m_x)]^2 (\mu_4[X] - D_x^2) +$$

$$+ \varphi'(m_x)\varphi''(m_x)\mu_3[X]. \quad (11.4.6)$$

Формула (11.4.6) дает уточненное значение дисперсии по сравнению с методом линеаризации; ее второй и третий члены представ-

ляют собой поправку на нелинейность функции. В формулу, кроме дисперсии аргумента D_x, входят еще третий и четвертый центральные моменты $\mu_3[X]$, $\mu_4[X]$. Если эти моменты известны, то поправка к дисперсии может быть найдена непосредственно по формуле (11.4.6). Однако зачастую нет необходимости в ее точном определении; достаточно лишь знать ее порядок. На практике часто встречаются случайные величины, распределенные приблизительно по нормальному закону. Для случайной величины, подчиненной нормальному закону,

$$\mu_3[X]=0; \quad \mu_4[X]=3\sigma_x^4=3D_x^2, \tag{11.4.7}$$

и формула (11.4.6) принимает вид:

$$D_y=[\varphi'(m_x)]^2 D_x+\frac{1}{2}[\varphi''(m_x)]^2 D_x^2. \tag{11.4.8}$$

Формулой (11.4.8) можно пользоваться для приближенной оценки погрешности метода линеаризации в случае, когда аргумент распределен по закону, близкому к нормальному.

Совершенно аналогичный метод может быть применен по отношению к функции нескольких случайных аргументов:

$$Y=\varphi(X_1, X_2, \ldots, X_n). \tag{11.4.9}$$

Разлагая функцию

$$y=\varphi(x_1, x_2, \ldots, x_n)$$

в ряд Тейлора в окрестности точки $m_{x_1}, m_{x_2}, \ldots, m_{x_n}$ и сохраняя в разложении члены не выше второго порядка, имеем приближенно:

$$Y=\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n})+\sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m(X_i-m_{x_i})+$$

$$+\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m(X_i-m_{x_i})^2+\sum_{i<j}\left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}\right)_m(X_i-m_{x_i})(X_j-m_{x_j}),$$

или, вводя центрированные величины,

$$Y=\varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n})+\sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m\mathring{X}_i+$$

$$+\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m\mathring{X}_i^2+\sum_{i<j}\left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}\right)_m\mathring{X}_i\mathring{X}_j, \tag{11.4.10}$$

где индекс m по-прежнему обозначает, что в выражение частной производной вместо аргументов X_i подставлены их математические ожидания m_{x_i}.

Применяя к формуле (11.4.10) операцию математического ожидания, имеем:

$$m_y = \varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n}) + \\ + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m D_{x_i} + \sum_{i<j}\left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}\right)_m K_{ij}, \quad (11.4.11)$$

где K_{ij} — корреляционный момент величин X_i, X_j.

В наиболее важном для практики случае, когда аргументы X_1, X_2, ..., X_n некоррелированны, формула (11.4.11) принимает вид:

$$m_y = \varphi(m_{x_1}, m_{x_2}, \ldots, m_{x_n}) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m D_{x_i}. \quad (11.4.12)$$

Второй член формулы (11.4.12) представляет собой поправку на нелинейность функции.

Перейдем к определению дисперсии величины Y. Чтобы получить выражение дисперсии в наиболее простом виде, предположим, что величины X_1, X_2, ..., X_n не только некоррелированны, но и независимы. Определяя дисперсию правой и левой части (11.4.10) и пользуясь теоремой о дисперсии произведения (см. п° 10.2), получим:

$$D_y = \sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m^2 D_{x_i} + \frac{1}{4}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m^2 (\mu_4[X_i] - D_{x_i}^2) + \\ + \sum_{i<j}\left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}\right)_m^2 D_{x_i}D_{x_j} + \sum\left(\frac{\partial\varphi}{\partial x_i}\right)_m\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m \mu_3[X_i]. \quad (11.4.13)$$

Для величин, распределенных по закону, близкому к нормальному, можно воспользоваться формулой (11.4.7) и преобразовать выражение (11.4.13) к виду:

$$D_y = \sum_{i=1}^{n}\left(\frac{\partial\varphi}{\partial x_i}\right)_m^2 D_{x_i} + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2\varphi}{\partial x_i^2}\right)_m^2 D_{x_i}^2 + \\ + \sum_{i<j}\left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}\right)^2 D_{x_i}D_{x_j}. \quad (11.4.14)$$

Последние два члена в выражении (11.4.14) представляют собой «поправку на нелинейность функции» и могут служить для оценки точности метода линеаризации при вычислении дисперсии.

ГЛАВА 12

ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ФУНКЦИЙ СЛУЧАЙНЫХ АРГУМЕНТОВ

12.1. Закон распределения монотонной функции одного случайного аргумента

В предыдущих главах мы познакомились с методами определения числовых характеристик функций случайных величин; главное удобство этих методов в том, что они не требуют нахождения законов распределения функций. Однако иногда возникает необходимость в определении не только числовых характеристик, но и законов распределения функций.

Начнем с рассмотрения наиболее простой задачи, относящейся к этому классу: задачи о законе распределения функции одного случайного аргумента. Так как для практики наибольшее значение имеют непрерывные случайные величины, будем решать задачу именно для них.

Имеется непрерывная случайная величина X с плотностью распределения $f(x)$. Другая случайная величина Y связана с нею функциональной зависимостью:

$$Y = \varphi(X)^{1}).$$

Требуется найти плотность распределения величины Y.

Рассмотрим участок оси абсцисс (a, b), на котором лежат все возможные значения величины X, т. е.

$$P(a < X < b) = 1.$$

В частном случае, когда область возможных значений X ничем не ограничена, $a = -\infty$; $b = +\infty$.

Способ решения поставленной задачи зависит от поведения функции φ на участке (a, b): возрастает ли она на этом участке или убывает, или колеблется.

---

[1]) Функцию φ предполагаем непрерывной и дифференцируемой.

В данном п° мы рассмотрим случай, когда функция $y=\varphi(x)$ на участке (a, b) монотонна [1]). При этом отдельно проанализируем два случая: монотонного возрастания и монотонного убывания функции.

1. Функция $y=\varphi(x)$ на участке (a, b) монотонно возрастает (рис. 12.1.1). Когда величина X принимает различные значения на участке (a, b), случайная точка (X, Y) перемещается только по кривой $y=\varphi(x)$; ордината этой случайной точки полностью определяется ее абсциссой.

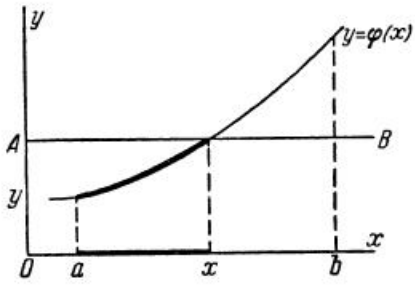


Рис. 12.1.1.

Обозначим $g(y)$ плотность распределения величины Y. Для того чтобы определить $g(y)$, найдем сначала функцию распределения величины Y:

$$G(y)=P(Y<y).$$

Проведем прямую AB, параллельную оси абсцисс на расстоянии y от нее (рис. 12.1.1). Чтобы выполнялось условие $Y<y$, случайная точка (X, Y) должна попасть на тот участок кривой, который лежит ниже прямой AB; для этого необходимо и достаточно, чтобы случайная величина X попала на участок оси абсцисс от a до x, где x — абсцисса точки пересечения кривой $y=\varphi(x)$ и прямой AB. Следовательно,

$$G(y)=P(Y<y)=P(a<X<x)=\int_a^x f(x)\,dx.$$

Верхний предел интеграла x можно выразить через y:

$$x=\psi(y),$$

где ψ — функция, обратная функции φ. Тогда

$$G(y)=\int_a^{\psi(y)} f(x)\,dx. \qquad (12.1.1)$$

Дифференцируя интеграл (12.1.1) по переменной y, входящей в верхний предел, получим:

$$g(y)=G'(y)=f\big(\psi(y)\big)\psi'(y). \qquad (12.1.2)$$

[1]) Случай немонотонной функции будет рассмотрен в п° 12.3.

2. Функция $y=\varphi(x)$ на участке (a, b) монотонно убывает (рис. 12.1.2). В этом случае

$$G(y)=P(Y<y)=P(x<X<b)=\int_{x}^{b} f(x)\,dx=\int_{\varphi(y)}^{b} f(x)\,dx,$$

откуда

$$g(y)=G'(y)=-f\big(\psi(y)\big)\psi'(y). \qquad (12.1.3)$$

Сравнивая формулы (12.1.2) и (12.1.3), замечаем, что они могут быть объединены в одну:

$$g(y)=f\big(\psi(y)\big)\,|\psi'(y)|. \qquad (12.1.4)$$

Действительно, когда φ возрастает, ее производная (а значит, и ψ') положительна. При убывающей функции φ производная ψ' отрицательна, но зато перед ней в формуле (12.1.3) стоит минус. Следовательно, формула (12.1.4), в которой производная берется по модулю, верна в обоих случаях. Таким образом, задача о законе распределения монотонной функции решена.

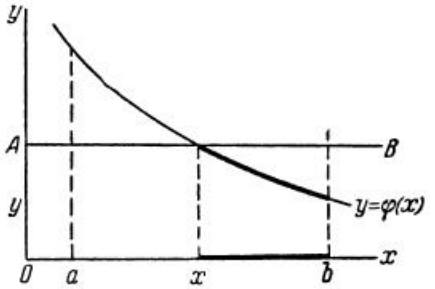


Рис. 12.1.2.

Пример. Случайная величина X подчинена закону Коши с плотностью распределения:

$$f(x)=\frac{1}{\pi(1+x^2)}.$$

Величина Y связана с X зависимостью

$$Y=1-X^3.$$

Найти плотность распределения величины Y.

Решение. Так как функция $y=1-x^3$ монотонна на участке $(-\infty, +\infty)$, можно применить формулу (12.1.4). Решение задачи оформим в виде двух столбцов: в левом будут помещены обозначения функций, принятые в общем решении задачи, в правом — конкретные функции, соответствующие данному примеру:

$f(x)$	$\dfrac{1}{\pi(1+x^2)}$
$y=\varphi(x)$	$y=1-x^3$
$x=\psi(y)$	$x=\sqrt[3]{1-y}$
$\psi'(y)$	$\dfrac{-1}{3\sqrt[3]{(1-y)^2}}$
$\lvert\psi'(y)\rvert$	$\dfrac{1}{3\sqrt[3]{(1-y)^2}}$
$g(y)=f\big(\psi(y)\big)\lvert\psi'(y)\rvert$	$g(y)=\dfrac{1}{3\pi\big(1+\sqrt[3]{(1-y)^2}\big)\sqrt[3]{(1-y)^2}}.$

12.2. Закон распределения линейной функции от аргумента, подчиненного нормальному закону

Пусть случайная величина X подчинена нормальному закону с плотностью:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}, \tag{12.2.1}$$

а случайная величина Y связана с нею линейной функциональной зависимостью:

$$Y = aX + b, \tag{12.2.2}$$

где a и b — неслучайные коэффициенты.

Требуется найти закон распределения величины Y.

Оформим решение в виде двух столбцов, аналогично примеру предыдущего п°:

$f(x)$	$\frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}$
$y = \varphi(x)$	$y = ax + b$
$x = \psi(y)$	$x = \frac{y-b}{a}$
$\psi'(y)$	$\frac{1}{a}$
$\lvert\psi'(y)\rvert$	$\frac{1}{\lvert a\rvert}$
$g(y) = f(\psi(y))\lvert\psi'(y)\rvert$	$g(y) = \frac{1}{\lvert a\rvert \sigma_x \sqrt{2\pi}} e^{-\frac{\left(\frac{y-b}{a}-m_x\right)^2}{2\sigma_x^2}}$.

Преобразуя выражение $g(y)$, имеем:

$$g(y) = \frac{1}{|a|\sigma_x \sqrt{2\pi}} e^{-\frac{[y-(am_x+b)]^2}{2|a|^2\sigma_x^2}},$$

а это есть не что иное, как нормальный закон с параметрами:

$$\left.\begin{aligned} a_y &= am_x + b, \\ \sigma_y &= |a|\sigma_x. \end{aligned}\right\} \tag{12.2.3}$$

Если перейти от средних квадратических отклонений к пропорциональным им вероятным отклонениям, получим:

$$E_y = |a| E_x. \tag{12.2.4}$$

Таким образом мы убедились, что *линейная функция от аргумента, подчиненного нормальному закону, также подчинена нормальному закону*. Чтобы найти центр рассеивания этого закона, нужно в выражение линейной функции вместо аргумента подставить его центр рассеивания. Чтобы найти среднее квадратическое отклонение этого закона, нужно среднее квадратическое отклонение аргумента умножить на модуль коэффициента при аргументе в выражении линейной функции. То же правило справедливо и для вероятных отклонений.

12.3. Закон распределения немонотонной функции одного случайного аргумента

Имеется непрерывная случайная величина X с плотностью распределения $f(x)$; другая величина Y связана с X функциональной зависимостью:

$$Y = \varphi(X),$$

причем функция $y = \varphi(x)$ на участке (a, b) возможных значений аргумента не монотонна (рис. 12.3.1).

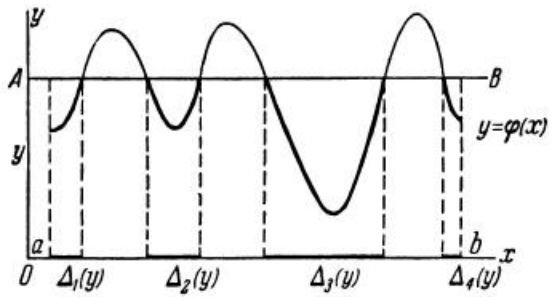


Рис. 12.3.1.

Найдем функцию распределения $G(y)$ величины Y. Для этого снова проведем прямую AB, параллельную оси абсцисс, на расстоянии y от нее и выделим те участки кривой $y = \varphi(x)$, на которых выполняется условие $Y < y$. Пусть этим участкам соответствуют участки оси абсцисс: $\Delta_1(y)$, $\Delta_2(y)$, ...

Событие $Y < y$ равносильно попаданию случайной величины X на один из участков $\Delta_1(y)$, $\Delta_2(y)$, ... — безразлично, на какой именно. Поэтому

$$G(y) = P(Y < y) = P\Big((X \subset \Delta_1(y)) + (X \subset \Delta_2(y)) + \ldots\Big) =$$

$$= \sum_i P(X \subset \Delta_i(y)) = \sum_i \int_{\Delta_i(y)} f(x)\,dx.$$

Таким образом, для функции распределения величины $Y=\varphi(X)$ имеем формулу:

$$G(y)=\sum_i \int_{\Delta_i(y)} f(x)\,dx. \qquad (12.3.1)$$

Границы интервалов $\Delta_i(y)$ зависят от y и при заданном конкретном виде функции $y=\varphi(x)$ могут быть выражены как явные функции y. Дифференцируя $G(y)$ по величине y, входящей в пределы интегралов, получим плотность распределения величины Y:

$$g(y)=G'(y). \qquad (12.3.2)$$

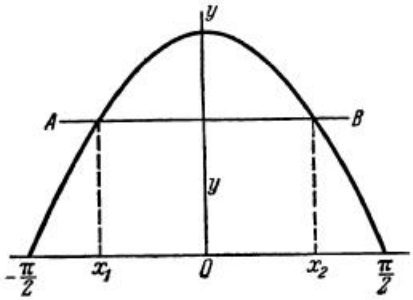


Рис. 12.3.2.

Пример. Величина X подчинена закону равномерной плотности на участке от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$:

$$f(x)=\begin{cases}\frac{1}{\pi} \text{ при } |x|<\frac{\pi}{2};\\ 0 \text{ при } |x|>\frac{\pi}{2}.\end{cases}$$

Найти закон распределения величины

$$Y=\cos X.$$

Решение. Строим график функции $y=\cos x$ (рис. 12.3.2). Очевидно, $a=-\frac{\pi}{2}$, $b=\frac{\pi}{2}$, и в интервале (a, b) функция $y=\cos x$ немонотонна.

Применяя формулу (12.3.1), имеем:

$$G(y)=\int_{-\frac{\pi}{2}}^{x_1} f(x)\,dx+\int_{x_2}^{\frac{\pi}{2}} f(x)\,dx.$$

Выразим пределы x_1 и x_2 через y:

$$x_1=-\arccos y;\quad x_2=\arccos y.$$

Отсюда

$$G(y)=\int_{-\frac{\pi}{2}}^{-\arccos y} f(x)\,dx+\int_{\arccos y}^{\frac{\pi}{2}} f(x)\,dx. \qquad (12.3.3)$$

Чтобы найти плотность $g(y)$, не будем вычислять интегралы в формуле (12.3.3), а непосредственно продифференцируем это выражение по переменной y, входящей в пределы интегралов:

$$g(y)=G'(y)=f(-\arccos y)\frac{1}{\sqrt{1-y^2}}+f(\arccos y)\frac{1}{\sqrt{1-y^2}}.$$

Имея в виду, что $f(x) = \frac{1}{\pi}$, получим:

$$g(y) = \frac{2}{\pi} \frac{1}{\sqrt{1-y^2}}. \qquad (12.3.4)$$

Указывая для Y закон распределения (12.3.4), следует оговорить, что он действителен лишь в пределах от 0 до 1, т. е. в тех пределах, в которых меняется $Y = \cos X$ при аргументе X, заключенном между $-\frac{\pi}{2}$ и $\frac{\pi}{2}$. Вне этих пределов плотность $g(y)$ равна нулю.

График функции $g(y)$ дан на рис. 12.3.3. При $y = 1$ кривая $g(y)$ имеет ветвь, уходящую на бесконечность.

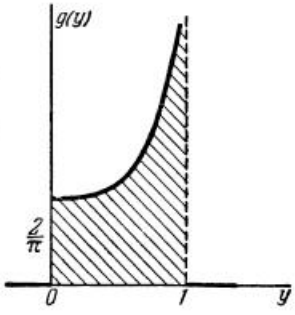


Рис. 12.3.3.

12.4. Закон распределения функции двух случайных величин

Задача определения закона распределения функции нескольких случайных аргументов значительно сложнее аналогичной задачи для функции одного аргумента. Здесь мы изложим общий метод решения этой задачи для наиболее простого случая функции двух аргументов.

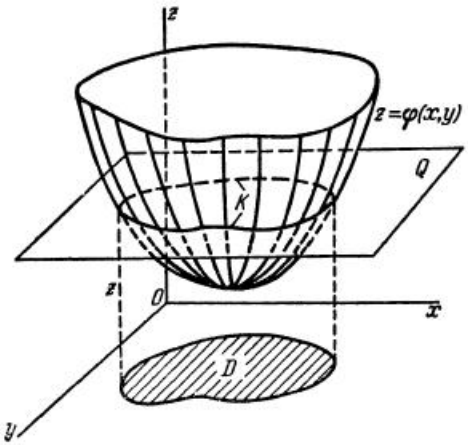


Рис. 12.4.1.

Имеется система двух непрерывных случайных величин (X, Y) с плотностью распределения $f(x, y)$. Случайная величина Z связана с X и Y функциональной зависимостью:

$$Z = \varphi(X, Y).$$

Требуется найти закон распределения величины Z.

Для решения задачи воспользуемся геометрической интерпретацией, аналогичной той, которую мы применяли в случае одного аргумента. Функция $z = \varphi(x, y)$ изобразится уже не кривой, а поверхностью (рис. 12.4.1).

Найдем функцию распределения величины Z:

$$G(z) = P(Z < z) = P\big(\varphi(X, Y) < z\big). \qquad (12.4.1)$$

Проведем плоскость Q, параллельную плоскости xOy, на расстоянии z от нее. Эта плоскость пересечет поверхность $z=\varphi(x, y)$ по некоторой кривой K [1]). Спроектируем кривую K на плоскость xOy. Эта проекция, уравнение которой $\varphi(x, y)=z$, разделит плоскость xOy на две области; для одной из них высота поверхности над плоскостью xOy будет меньше, а для другой — больше z. Обозначим D ту область, для которой эта высота меньше z. Чтобы выполнялось неравенство (12.4.1), случайная точка (X, Y), очевидно, должна попасть в область D; следовательно,

$$G(z)=P\big((X, Y)\subset D\big)=\iint\limits_{(D)} f(x, y)\,dx\,dy. \qquad (12.4.2)$$

В выражение (12.4.2) величина z входит неявно, через пределы интегрирования.

Дифференцируя $G(z)$ по z, получим плотность распределения величины Z:

$$g(z)=G'(z).$$

Зная конкретный вид функции $z=\varphi(x, y)$, можно выразить пределы интегрирования через z и написать выражение $g(z)$ в явном виде.

Для того чтобы найти закон распределения функции двух аргументов, нет необходимости каждый раз строить поверхность $z=\varphi(x, y)$, подобно тому, как это сделано на рис. 12.4.1, и пересекать ее плоскостью, параллельной xOy. На практике достаточно построить на плоскости xOy кривую, уравнение которой $z=\varphi(x, y)$, отдать себе отчет в том, по какую сторону этой кривой $Z<z$, а по какую $Z>z$, и интегрировать по области D, для которой $Z<z$.

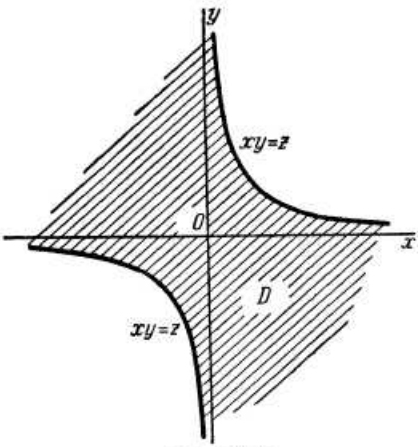


Рис. 12.4.2.

Пример. Система величин (X, Y) подчинена закону распределения с плотностью $f(x, y)$. Величина Z есть произведение величин X, Y:

$$Z=XY.$$

Найти плотность распределения величины Z.

Решение. Зададимся некоторым значением z и построим на плоскости xOy кривую, уравнение которой $xy=z$ (рис. 12.4.2). Это — гипербола, асимптоты которой совпадают с осями координат. Область D на рис. 12.4.2 заштрихована.

[1]) На нашем чертеже кривая K замкнута; вообще она может быть и незамкнутой, а может состоять и из нескольких ветвей.

Функция распределения величины Z имеет вид:

$$G(z) = \iint\limits_{(D)} f(x, y)\,dx\,dy = \int\limits_{-\infty}^{0}\int\limits_{\frac{z}{x}}^{\infty} f(x, y)\,dx\,dy + \int\limits_{0}^{\infty}\int\limits_{-\infty}^{\frac{z}{x}} f(x, y)\,dx\,dy.$$

Дифференцируя это выражение по z, имеем:

$$g(z) = -\int\limits_{-\infty}^{0} \frac{1}{x} f\left(x, \frac{z}{x}\right) dx + \int\limits_{0}^{\infty} \frac{1}{x} f\left(x, \frac{z}{x}\right) dx. \qquad (12.4.3)$$

12.5. Закон распределения суммы двух случайных величин. Композиция законов распределения

Воспользуемся изложенным выше общим методом для решения одной важной для практики частной задачи, а именно для нахождения закона распределения суммы двух случайных величин.

Имеется система двух случайных величин (X, Y) с плотностью распределения $f(x, y)$. Рассмотрим сумму случайных величин X и Y:

$$Z = X + Y,$$

и найдем закон распределения величины Z. Для этого построим на плоскости xOy линию, уравнение которой $x + y = z$ (рис. 12.5.1). Это — прямая, отсекающая на осях отрезки, равные z. Прямая $x + y = z$ делит плоскость xOy на две части; правее и выше ее $X + Y > z$; левее и ниже $X + Y < z$. Область D в данном случае — левая нижняя часть плоскости xOy, заштрихованная на рис. 12.5.1. Согласно формуле (12.4.2) имеем:

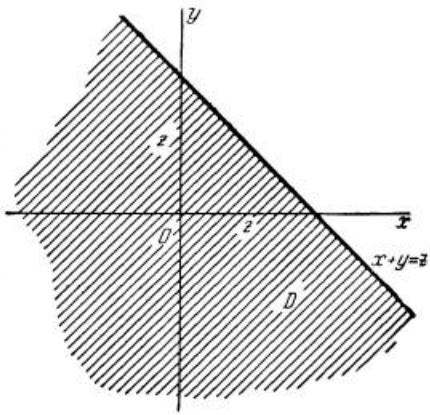


Рис. 12.5.1.

$$G(z) = \iint\limits_{(D)} f(x, y)\,dx\,dy =$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{z-x} f(x, y)\,dx\,dy = \int\limits_{-\infty}^{\infty}\left\{\int\limits_{-\infty}^{z-x} f(x, y)\,dy\right\} dx.$$

Дифференцируя это выражение по переменной z, входящей в верхний предел внутреннего интеграла, получим:

$$g(z) = \int_{-\infty}^{\infty} f(x,\ z-x)\,dx. \qquad (12.5.1)$$

Это — общая формула для плотности распределения суммы двух случайных величин.

Из соображений симметричности задачи относительно X и Y можно написать другой вариант той же формулы:

$$g(z) = \int_{-\infty}^{\infty} f(z-y,\ y)\,dy, \qquad (12.5.2)$$

который равносилен первому и может применяться вместо него.

Особое практическое значение имеет случай, когда складываемые случайные величины (X, Y) независимы. Тогда говорят о ***композиции законов распределения.***

Произвести композицию двух законов распределения это значит найти закон распределения суммы двух независимых случайных величин, подчиненных этим законам распределения.

Выведем формулу для композиции двух законов распределения. Имеются две независимые случайные величины X и Y, подчиненные соответственно законам распределения $f_1(x)$ и $f_2(y)$; требуется произвести композицию этих законов, т. е. найти плотность распределения величины

$$Z = X + Y.$$

Так как величины X и Y независимы, то

$$f(x,\ y) = f_1(x)\,f_2(y),$$

и формулы (12.5.1) и (12.5.2) принимают вид:

$$g(z) = \int_{-\infty}^{\infty} f_1(x)\,f_2(z-x)\,dx, \qquad (12.5.3)$$

$$g(z) = \int_{-\infty}^{\infty} f_1(z-y)\,f_2(y)\,dy. \qquad (12.5.4)$$

Для обозначения композиции законов распределения часто применяют символическую запись:

$$g = f_1 * f_2,$$

где $*$ — символ композиции.

Формулы (12.5.3) и (12.5.4) для композиции законов распределения удобны только тогда, когда законы распределения $f_1(x)$ и $f_2(y)$ (или, по крайней мере, один из них) заданы одной формулой на всем диапазоне значений аргумента (от $-\infty$ до ∞). Если же оба закона заданы на различных участках различными уравнениями (например, два закона равномерной плотности), то удобнее пользоваться непосредственно общим методом, изложенным в п° 12.4, т. е. вычислить функцию распределения $G(z)$ величины $Z = X + Y$ и продифференцировать эту функцию.

Пример 1. Составить композицию нормального закона:

$$f_1(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

и закона равномерной плотности:

$$f_2(y) = \frac{1}{\beta - \alpha} \quad \text{при} \quad \alpha < y < \beta.$$

Решение. Применим формулу композиции законов распределения в виде (12.5.4):

$$g(z) = \frac{1}{\beta-\alpha}\int_\alpha^\beta \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-y-m)^2}{2\sigma^2}} dy = \frac{1}{\beta-\alpha}\int_\alpha^\beta \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[y-(z-m)]^2}{2\sigma^2}} dy. \quad (12.5.5)$$

Подынтегральная функция в выражении (12.5.5) есть не что иное, как нормальный закон с центром рассеивания $z - m$ и средним квадратическим

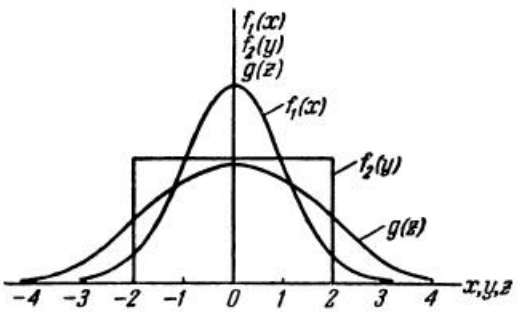


Рис. 12.5.2.

отклонением σ, а интеграл в выражении (12.5.5) есть вероятность попадания случайной величины, подчиненной этому закону, на участок от α до β; следовательно,

$$g(z) = \frac{1}{\beta-\alpha}\left[\Phi^*\left(\frac{\beta-(z-m)}{\sigma}\right) - \Phi^*\left(\frac{\alpha-(z-m)}{\sigma}\right)\right].$$

Графики законов $f_1(x)$, $f_2(y)$ и $g(z)$ при $\alpha = -2$, $\beta = 2$, $m = 0$, $\sigma = 1$ приведены на рис. 12.5.2.

Пример 2. Составить композицию двух законов равномерной плотности, заданных на одном и том же участке (0, 1):

$$f_1(x) = 1 \quad \text{при} \quad 0 < x < 1,$$
$$f_2(y) = 1 \quad \text{при} \quad 0 < y < 1.$$

Решение. Так как законы $f_1(x)$ и $f_2(y)$ заданы только на определенных участках осей Ox и Oy, для решения этой задачи удобнее воспользоваться не формулами (12.5.3) и (12.5.4), а общим методом, изложенным в п° 12.4, и найти функцию распределения $G(z)$ величины $Z = X + Y$.

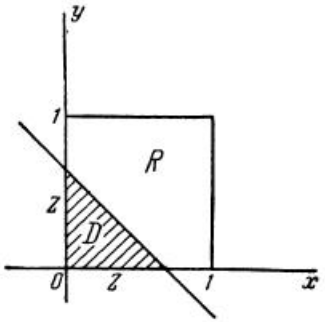


Рис. 12.5.3.

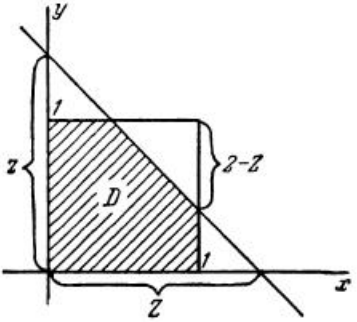


Рис. 12.5.4.

Рассмотрим случайную точку (X, Y) на плоскости xOy. Область ее возможных положений — квадрат R со стороной, равной 1 (рис. 12.5.3). Имеем:

$$G(z) = \iint\limits_{(D)} f(x, y)\, dx\, dy = \iint\limits_{(D)} dx\, dy,$$

где область D — часть квадрата R, лежащая левее и ниже прямой $x + y = z$. Очевидно,

$$G(z) = S_D,$$

где S_D — площадь области D.

Составим выражение для площади области D при различных значениях z, пользуясь рис. 12.5.3 и 12.5.4:

1) при $z < 0$ $\quad G(z) = 0$;
2) при $0 < z < 1$ $\quad G(z) = \frac{z^2}{2}$;
3) при $1 < z < 2$ $\quad G(z) = 1 - \frac{(2-z)^2}{2}$;
4) при $z > 2$ $\quad G(z) = 1$.

Дифференцируя эти выражения, получим:

1) при $z < 0$ $\quad g(z) = 0$;
2) при $0 < z < 1$ $\quad g(z) = z$;
3) при $1 < z < 2$ $\quad g(z) = 2 - z$;
4) при $z > 2$ $\quad g(z) = 0$.

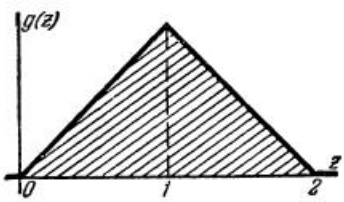


Рис. 12.5.5.

График закона распределения $g(z)$ дан на рис. 12.5.5. Такой закон носит название «закона Симпсона» или «закона треугольника».

12.6. Композиция нормальных законов

Рассмотрим две независимые случайные величины X и Y, подчиненные нормальным законам:

$$f_1(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}, \tag{12.6.1}$$

$$f_2(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}. \tag{12.6.2}$$

Требуется произвести композицию этих законов, т. е. найти закон распределения величины:

$$Z = X + Y.$$

Применим общую формулу (12.5.3) для композиции законов распределения:

$$g(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z-x)\, dx =$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} e^{-\frac{(x-m_x)^2}{2\sigma_x^2} - \frac{(z-x-m_y)^2}{2\sigma_y^2}} dx. \tag{12.6.3}$$

Если раскрыть скобки в показателе степени подынтегральной функции и привести подобные члены, получим:

$$g(z) = \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} e^{-Ax^2+2Bx-C}\, dx,$$

где

$$A = \frac{1}{2} \frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2 \sigma_y^2};$$

$$B = \frac{m_x}{2\sigma_x^2} + \frac{z-m_y}{2\sigma_y^2};$$

$$C = \frac{m_x^2}{2\sigma_x^2} + \frac{(z-m_y)^2}{2\sigma_y^2}.$$

Подставляя эти выражения в уже встречавшуюся нам формулу (9.1.3):

$$\int_{-\infty}^{\infty} e^{-Ax^2 \pm 2Bx - C}\, dx = \sqrt{\frac{\pi}{A}}\, e^{-\frac{AC-B^2}{A}}, \tag{12.6.4}$$

после преобразований получим:

$$g(z) = \frac{1}{\sqrt{\sigma_x^2 + \sigma_y^2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{[z-(m_x+m_y)]^2}{2(\sigma_x^2+\sigma_y^2)}}, \qquad (12.6.5)$$

а это есть не что иное, как нормальный закон с центром рассеивания

$$m_z = m_x + m_y \qquad (12.6.6)$$

и средним квадратическим отклонением

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}. \qquad (12.6.7)$$

К тому же выводу можно прийти значительно проще с помощью следующих качественных рассуждений.

Не раскрывая скобок и не производя преобразований в подынтегральной функции (12.6.3), сразу приходим к выводу, что показатель степени есть квадратный трехчлен относительно x вида

$$\varphi(x) = -Ax^2 + 2Bx - C,$$

где в коэффициент A величина z не входит совсем, в коэффициент B входит в первой степени, а в коэффициент C — в квадрате. Имея это в виду и применяя формулу (12.6.4), приходим к заключению, что $g(z)$ есть показательная функция, показатель степени которой — квадратный трехчлен относительно z, а плотность распределения такого вида соответствует нормальному закону. Таким образом, мы приходим к чисто качественному выводу: закон распределения величины z должен быть нормальным.

Чтобы найти параметры этого закона — m_z и σ_z — воспользуемся теоремой сложения математических ожиданий и теоремой сложения дисперсий. По теореме сложения математических ожиданий

$$m_z = m_x + m_y. \qquad (12.6.8)$$

По теореме сложения дисперсий

$$D_z = D_x + D_y,$$

или

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2, \qquad (12.6.9)$$

откуда следует формула (12.6.7).

Переходя от средних квадратических отклонений к пропорциональным им вероятным отклонениям, получим:

$$E_z^2 = E_x^2 + E_y^2. \qquad (12.6.10)$$

Таким образом, мы пришли к следующему правилу: *при композиции нормальных законов получается снова нормальный закон, причем математические ожидания и дисперсии (или квадраты вероятных отклонений) суммируются.*

Правило композиции нормальных законов может быть обобщено на случай произвольного числа независимых случайных величин.

Если имеется n независимых случайных величин:

$$X_1, X_2, \ldots, X_n,$$

подчиненных нормальным законам с центрами рассеивания

$$m_{x_1}, m_{x_2}, \ldots, m_{x_n}$$

и средними квадратическими отклонениями

$$\sigma_{x_1}, \sigma_{x_2}, \ldots, \sigma_{x_n},$$

то величина

$$Z = \sum_{i=1}^{n} X_i$$

также подчинена нормальному закону с параметрами

$$m_z = \sum_{i=1}^{n} m_{x_i}, \tag{12.6.11}$$

$$\sigma_z^2 = \sum_{i=1}^{n} \sigma_{x_i}^2. \tag{12.6.12}$$

Вместо формулы (12.6.12) можно применять равносильную ей формулу:

$$E_z^2 = \sum_{i=1}^{n} E_{x_i}^2. \tag{12.6.13}$$

Если система случайных величин (X, Y) распределена по нормальному закону, но величины X, Y зависимы, то нетрудно доказать, так же как раньше, исходя из общей формулы (12.5.1), что закон распределения величины

$$Z = X + Y$$

есть тоже нормальный закон. Центры рассеивания по-прежнему складываются алгебраически, но для средних квадратических отклонений правило становится более сложным:

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2r\sigma_x\sigma_y, \tag{12.6.14}$$

где r — коэффициент корреляции величин X и Y.

При сложении нескольких зависимых случайных величин, подчиненных в своей совокупности нормальному закону, закон распределения суммы также оказывается нормальным с параметрами

$$m_z = \sum_{i=1}^{n} m_{x_i}, \tag{12.6.15}$$

$$\sigma_z^2 = \sum_{i=1}^{n} \sigma_{x_i}^2 + 2 \sum_{i<j} r_{ij} \sigma_{x_i} \sigma_{x_j}, \tag{12.6.16}$$

или в вероятных отклонениях

$$E_z^2 = \sum_{i=1}^{n} E_{x_i}^2 + 2 \sum_{i<j} r_{ij} E_{x_i} E_{x_j}, \tag{12.6.17}$$

где r_{ij} — коэффициент корреляции величин X_i, X_j, а суммирование распространяется на все различные попарные комбинации величин $X_1, X_2, \ldots, X_n$.

Мы убедились в весьма важном свойстве нормального закона: при композиции нормальных законов получается снова нормальный закон. Это — так называемое «свойство устойчивости». Закон распределения называется *устойчивым*, если при композиции двух законов этого типа получается снова закон того же типа[1]). Выше мы показали, что нормальный закон является устойчивым. Свойством устойчивости обладают весьма немногие законы распределения. В предыдущем п° (пример 2) мы убедились, что, например, закон равномерной плотности неустойчив: при композиции двух законов равномерной плотности на участках от 0 до 1 мы получили закон Симпсона.

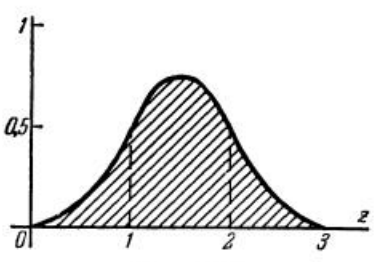


Рис. 12.6.1.

Устойчивость нормального закона — одно из существенных условий его широкого распространения на практике. Однако свойством устойчивости, кроме нормального, обладают и некоторые другие законы распределения. Особенностью нормального закона является то, что при композиции достаточно большого числа практически произвольных законов распределения суммарный закон оказывается сколь угодно близок к нормальному вне зависимости от того, каковы были законы распределения слагаемых. Это можно проиллюстрировать, например, составляя композицию трех законов равномерной плотности на участках от 0 до 1. Получающийся при этом закон распределения $g(z)$ изображен на рис. 12.6.1. Как видно из чертежа, график функции $g(z)$ весьма напоминает график нормального закона.

[1]) Под «законами одного и того же типа» мы подразумеваем законы, различающиеся только масштабами по осям и началом отсчета по оси абсцисс.

12.7. Линейные функции от нормально распределенных аргументов

Дана система случайных величин $(X_1, X_2, \ldots, X_n)$, подчиненная нормальному закону распределения (или, короче, «распределенная нормально»); случайная величина Y представляет собой линейную функцию этих величин:

$$Y = \sum_{i=1}^{n} a_i X_i + b. \tag{12.7.1}$$

Требуется найти закон распределения величины Y.

Нетрудно убедиться, что это нормальный закон. Действительно, величина Y представляет собой сумму линейных функций, каждая из которых зависит от одного нормально распределенного аргумента X, а выше было доказано, что такая линейная функция также распределена нормально. Складывая несколько нормально распределенных случайных величин, мы снова получим величину, распределенную нормально.

Остается найти параметры величины Y — центр рассеивания m_y и среднее квадратическое отклонение σ_y. Применяя теоремы о математическом ожидании и дисперсии линейной функции, получим:

$$m_y = \sum_{i=1}^{n} a_i m_{x_i} + b, \tag{12.7.2}$$

$$\sigma_y^2 = \sum_{i=1}^{n} a_i^2 \sigma_{x_i}^2 + 2 \sum_{i<j} a_i a_j r_{ij} \sigma_{x_i} \sigma_{x_j}, \tag{12.7.3}$$

где r_{ij} — коэффициент корреляции величин X_i, X_j.

В случае, когда величины $(X_1, X_2, \ldots, X_n)$ некоррелированы (а значит, при нормальном законе, и независимы), формула (12.7.3) принимает вид:

$$\sigma_y^2 = \sum_{i=1}^{n} a_i^2 \sigma_{x_i}^2. \tag{12.7.4}$$

Средние квадратические отклонения в формулах (12.7.3) и (12.7.4) могут быть заменены пропорциональными им вероятными отклонениями.

На практике часто встречается случай, когда законы распределения случайных величин $X_1, X_2, \ldots, X_n$, входящих в формулу (12.7.1), в точности не известны, а известны только их числовые характеристики: математические ожидания и дисперсии. Если при этом величины $X_1, X_2, \ldots, X_n$ независимы, а их число n достаточно велико, то, как правило, можно утверждать, что, безотносительно к виду законов распределения величин X_i, закон распределения величины Y

близок к нормальному. На практике для получения закона распределения, который приближенно может быть принят за нормальный, обычно оказывается достаточным наличие $5 \div 10$ слагаемых в выражении (12.7.1). Следует оговориться, что это не относится к случаю, когда дисперсия одного из слагаемых в формуле (12.7.1) подавляюще велика по сравнению со всеми другими; предполагается, что случайные слагаемые в сумме (12.7.1) по своему рассеиванию имеют примерно один и тот же порядок. Если эти условия соблюдены, то для величины Y может быть приближенно принят нормальный закон с параметрами, определяемыми формулами (12.7.2) и (12.7.4).

Очевидно, все вышеприведенные соображения о законе распределения линейной функции справедливы (разумеется, приближенно) и для того случая, когда функция не является в точности линейной, но может быть линеаризована.

12.8. Композиция нормальных законов на плоскости

Пусть в системе координат xOy заданы два независимых случайных вектора: $\vec{V}_1$ с составляющими (X_1, Y_1) и $\vec{V}_2$ с составляющими (X_2, Y_2). Допустим, что каждый из них распределен нормально, причем параметры первого вектора равны

$$m_{x_1},\ m_{y_1},\ \sigma_{x_1},\ \sigma_{y_1},\ r_{x_1y_1},$$

а параметры второго —

$$m_{x_2},\ m_{y_2},\ \sigma_{x_2},\ \sigma_{y_2},\ r_{x_2y_2}.$$

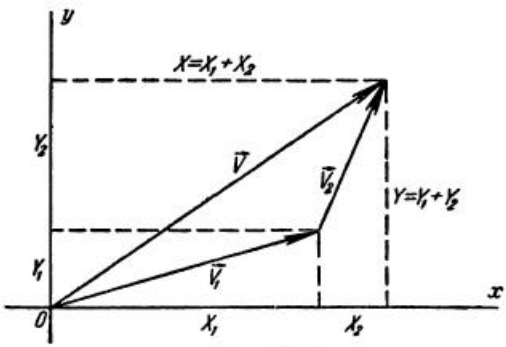


Рис. 12.8.1.

Требуется определить закон распределения случайного вектора $\vec{V} = \vec{V}_1 + \vec{V}_2$ (рис. 12.8.1), составляющие которого равны:

$$X = X_1 + X_2;$$
$$Y = Y_1 + Y_2.$$

Не представляет трудности качественно доказать (аналогично тому, как мы это сделали для случая композиции двух нормальных законов в п° 12.6), что вектор $\vec{V}$ также распределен нормально. Мы примем это положение без специального доказательства.

Определим параметры закона распределения вектора $\vec{V}$.

По теореме сложения математических ожиданий

$$\left.\begin{aligned} m_x &= m_{x_1} + m_{x_2}, \\ m_y &= m_{y_1} + m_{y_2}. \end{aligned}\right\} \qquad (12.8.1)$$

По теореме сложения дисперсий

$$\left.\begin{aligned} \sigma_x^2 &= \sigma_{x_1}^2 + \sigma_{x_2}^2, \\ \sigma_y^2 &= \sigma_{y_1}^2 + \sigma_{y_2}^2. \end{aligned}\right\} \tag{12.8.2}$$

По теореме сложения корреляционных моментов

$$K_{xy} = K_{x_1y_1} + K_{x_2y_2},$$

или, переходя к коэффициентам корреляции,

$$r_{xy}\sigma_x\sigma_y = r_{x_1y_1}\sigma_{x_1}\sigma_{y_1} + r_{x_2y_2}\sigma_{x_2}\sigma_{y_2},$$

откуда

$$r_{xy} = \frac{r_{x_1y_1}\sigma_{x_1}\sigma_{y_1} + r_{x_2y_2}\sigma_{x_2}\sigma_{y_2}}{\sqrt{(\sigma_{x_1}^2 + \sigma_{x_2}^2)(\sigma_{y_1}^2 + \sigma_{y_2}^2)}}. \tag{12.8.3}$$

Таким образом, задача композиции нормальных законов на плоскости решается формулами (12.8.1), (12.8.2) и (12.8.3).

Эти формулы выведены для того случая, когда оба исходных нормальных закона (для векторов $\vec{V}_1$ и $\vec{V}_2$) заданы в одной и той же координатной системе xOy. На практике иногда встречается случай, когда нужно произвести композицию двух нормальных законов на плоскости, каждый из которых задан в своей системе координат, а именно в своих главных осях рассеивания. Дадим способ композиции нормальных законов для этого случая.

Пусть на плоскости xOy (рис. 12.8.2) даны два нормально распределенных некоррелированных случайных вектора $\vec{V}_1$ и $\vec{V}_2$. Каждый

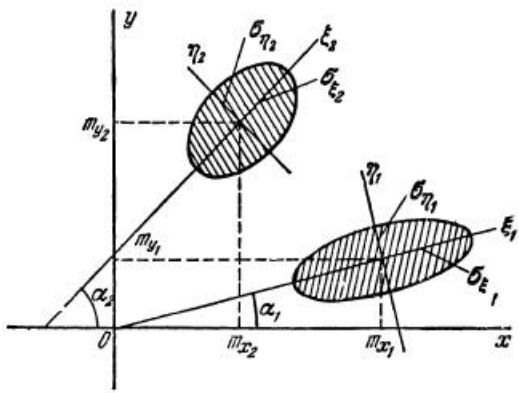


Рис. 12.8.2.

из векторов характеризуется своим единичным эллипсом рассеивания: вектор $\vec{V}_1$ — эллипсом с центром в точке m_{x_1}, m_{y_1} с полуосями σ_{ξ_1}, σ_{η_1}, из которых первая образует с осью Ox угол α_1; аналогичные

характеристики для вектора $\vec{V}_2$ будут: m_{x_2}, m_{y_2}, σ_{ξ_2}, σ_{η_2}, α_2. Требуется найти параметры единичного эллипса рассеивания, характеризующего вектор $\vec{V} = \vec{V}_1 + \vec{V}_2$. Обозначим их

$$m_x,\ m_y,\ \sigma_x,\ \sigma_y,\ \alpha.$$

Так как положение центра рассеивания не зависит от выбора системы координат, очевидно, по-прежнему будут справедливы соотношения:

$$m_x = m_{x_1} + m_{x_2},$$
$$m_y = m_{y_1} + m_{y_2}.$$

Для того чтобы найти элементы корреляционной матрицы вектора $\vec{V}$, спроектируем случайные точки, соответствующие векторам $\vec{V}_1$ и $\vec{V}_2$, на оси Ox и Oy. Пользуясь формулой (10.3.3), получим:

$$\left.\begin{aligned}
\sigma_{x_1}^2 &= \sigma_{\xi_1}^2 \cos^2\alpha_1 + \sigma_{\eta_1}^2 \sin^2\alpha_1,\\
\sigma_{y_1}^2 &= \sigma_{\xi_1}^2 \sin^2\alpha_1 + \sigma_{\eta_1}^2 \cos^2\alpha_1,\\
\sigma_{x_2}^2 &= \sigma_{\xi_2}^2 \cos^2\alpha_2 + \sigma_{\eta_2}^2 \sin^2\alpha_2,\\
\sigma_{y_2}^2 &= \sigma_{\xi_2}^2 \sin^2\alpha_2 + \sigma_{\eta_2}^2 \cos^2\alpha_2.
\end{aligned}\right\} \qquad (12.8.4)$$

Коэффициенты корреляции составляющих векторов $\vec{V}_1$ и $\vec{V}_2$ в системе координат xOy найдем из соотношения (9.2.2):

$$\left.\begin{aligned}
r_{x_1y_1} &= \frac{\operatorname{tg} 2\alpha_1 \left(\sigma_{x_1}^2 - \sigma_{y_1}^2\right)}{2\sigma_{x_1}\sigma_{y_1}},\\
r_{x_2y_2} &= \frac{\operatorname{tg} 2\alpha_2 \left(\sigma_{x_2}^2 - \sigma_{y_2}^2\right)}{2\sigma_{x_2}\sigma_{y_2}}.
\end{aligned}\right\} \qquad (12.8.5)$$

Далее задача композиции нормальных законов на плоскости сводится к предыдущей. Зная σ_x, σ_y, r_{xy}, можно найти углы, составленные осями суммарного эллипса с осью абсцисс, по формуле (9.2.2):

$$\operatorname{tg} 2\alpha = \frac{2r_{xy}\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}\ ^{1}), \qquad (12.8.6)$$

и главные средние квадратические отклонения — по формулам (9.2.4):

$$\left.\begin{aligned}
\sigma_\xi &= \sqrt{\sigma_x^2 \cos^2\alpha + r_{xy}\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \sin^2\alpha},\\
\sigma_\eta &= \sqrt{\sigma_x^2 \sin^2\alpha - r_{xy}\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \cos^2\alpha}.
\end{aligned}\right\} \qquad (12.8.7)$$

---

[1]) Так как тангенс имеет период π, то значения α, определяемые по формуле (12.8.6), могут различаться на $\frac{\pi}{2}$, что соответствует двум главным осям эллипса.

Последние соотношения справедливы не только для средних квадратических отклонений, но и для пропорциональных им вероятных отклонений:

$$\left.\begin{aligned} E_\xi &= \sqrt{E_x^2\cos^2\alpha + r_{xy}E_xE_y\sin 2\alpha + E_y^2\sin^2\alpha}, \\ E_\eta &= \sqrt{E_x^2\sin^2\alpha - r_{xy}E_xE_y\sin 2\alpha + E_y^2\cos^2\alpha}. \end{aligned}\right\} \quad (12.8.8)$$

Перейдем к композиции произвольного числа нормальных законов на плоскости.

С наиболее простым случаем композиции произвольного числа нормальных законов мы встречаемся тогда, когда главные оси рассеивания для всех законов, подлежащих композиции, параллельны друг другу. Тогда, выбирая координатные оси параллельно этим главным осям рассеивания, мы будем иметь дело с системами независимых случайных величин, и композиция нормальных законов выполняется по простым формулам:

$$\left.\begin{aligned} m_x &= \sum_{i=1}^{n} m_{x_i}; & \sigma_x^2 &= \sum_{i=1}^{n} \sigma_{x_i}^2; \\ m_y &= \sum_{i=1}^{n} m_{y_i}; & \sigma_y^2 &= \sum_{i=1}^{n} \sigma_{y_i}^2, \end{aligned}\right\} \quad (12.8.9)$$

где σ_x, σ_{x_i}, σ_y, σ_{y_i} — главные средние квадратические отклонения соответствующих законов.

В случае, когда направления главных осей не совпадают, можно составить композицию нескольких нормальных законов тем же методом, которым мы пользовались выше для двух законов, т. е. проектируя складываемые случайные векторы на оси одной и той же системы координат.

На практике часто встречаются случаи, когда в числе законов, подлежащих композиции, встречаются так называемые «вырожденные» законы, т. е. законы, характеризующиеся эллипсом рассеивания, имеющим только одну полуось (другая равна нулю). Такие «вырожденные» законы дают рассеивание только в одном направлении. При композиции таких законов нужно поступать так же, как при композиции обычных законов, полагая некоторые параметры (средние квадратические или вероятные отклонения) равными нулю.

Пример 1. Ошибка бомбометания вызвана совместным действием следующих факторов:

1) техническое рассеивание бомб;
2) неточность прицеливания по дальности;
3) неточная наводка в боковом направлении.

Все эти факторы независимы. Техническое рассеивание бомб дает единичный эллипс рассеивания в виде круга радиусом 20 *м*. Ошибка прицеливания

по дальности действует только в направлении полета и имеет среднее квадратическое отклонение 40 *м*; центр рассеивания сдвинут вперед по полету на 5 *м*. Ошибка боковой наводки действует только в направлении, перпендикулярном к полету, и имеет среднее квадратическое отклонение 30 *м*; центр рассеивания смещен вправо на 10 *м*. Найти параметры нормального закона, которому подчинена суммарная ошибка бомбометания, вызванная совместным действием всех перечисленных факторов.

Решение. Так как главные оси всех перечисленных в задаче эллипсов (из которых второй и третий вырождены) параллельны, то можно применить правило композиции нормальных законов с независимыми составляющими (формулы (12.8.9)). Выбирая ось Oy по направлению полета, ось Ox — перпендикулярно к нему, имеем:

$$m_x = 10, \qquad m_y = 5,$$

$$\sigma_x^2 = 20^2 + 30^2 = 1300, \qquad \sigma_x = 36{,}1\ (м),$$

$$\sigma_y^2 = 20^2 + 40^2 = 2000, \qquad \sigma_y = 44{,}7\ (м).$$

Пример 2. Производится воздушная стрельба с самолета по самолету; рассеивание точек попадания рассматривается на вертикальной плоскости, перпендикулярной к направлению стрельбы. Причины рассеивания точек попадания состоят в следующем:

1) ошибки, связанные с неоднородностью баллистики снарядов и колебаниями установки;
2) ошибки наводки;
3) ошибки, вызванные неточностью определения дальности;
4) инструментальные ошибки прицела.

Главные оси рассеивания, вызванного первой причиной, расположены горизонтально и вертикально, и главные средние квадратические отклонения равны соответственно 1 и 2 *м*; ошибка наводки дает круговое рассеивание со средним квадратическим отклонением 3 *м*; ошибка, вызванная неточностью определения дальности, дает рассеивание только вдоль оси, наклоненной к горизонту под углом 30°, со с. к. о. 4 *м*; инструментальные ошибки прицела дают круговое рассеивание со с. к. о. 2 *м*. Систематические ошибки равны нулю.

Требуется найти параметры закона распределения суммарной ошибки, вызванной всеми перечисленными факторами.

Решение. Выбираем систему координат с горизонтальной осью Ox и вертикальной Oy. Эти оси являются главными осями рассеивания для всех законов, кроме третьего (ошибки вследствие неточности определения дальности). Обозначим составляющие каждой ошибки в системе координат xOy соответственно:

$$X_1,\ Y_1;\ X_2,\ Y_2;\ X_3,\ Y_3;\ X_4,\ Y_4.$$

Параметры этих составляющих равны соответственно:

$$m_{x_1} = m_{x_2} = m_{x_3} = m_{x_4} = m_{y_1} = m_{y_2} = m_{y_3} = m_{y_4} = 0;$$

$$\sigma_{x_1} = 1; \qquad \sigma_{y_1} = 2; \qquad \sigma_{x_2} = \sigma_{y_2} = 3; \qquad \sigma_{x_4} = \sigma_{y_4} = 2.$$

Что касается величин σ_{x_3} и σ_{y_3}, то их мы определяем, проектируя случайную точку (X_3, Y_3) на оси Ox и Oy по формулам (12.8.4):

$$\sigma_{x_3}^2 = 4^2 \cos^2 30^\circ = 12;$$

$$\sigma_{y_3}^2 = 4^2 \sin^2 30^\circ = 4.$$

Коэффициент корреляции величин (X_3, Y_3) найдем по формуле (12.8.5)

$$r_{x_3y_3} = \frac{\operatorname{tg} 60^\circ (12-4)}{2\sqrt{48}} = 1,$$

что и естественно, так как рассеивание сосредоточено на одной прямой и, следовательно, величины X_3 и Y_3 зависимы функционально.

Применяя теорему сложения дисперсий, имеем:

$$\sigma_x^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 + \sigma_{x_4}^2 = 26; \qquad \sigma_x = 5{,}09 \ (м);$$

$$\sigma_y^2 = \sigma_{y_1}^2 + \sigma_{y_2}^2 + \sigma_{y_3}^2 + \sigma_{y_4}^2 = 21; \qquad \sigma_y = 4{,}58 \ (м).$$

Коэффициент корреляции r_{xy} найдем, применяя теорему сложения корреляционных моментов:

$$K_{xy} = K_{x_1y_1} + K_{x_2y_2} + K_{x_3y_3} + K_{x_4y_4} = 0 + 0 + \frac{4\sqrt{3} \cdot 1}{2\rho^2} + 0,$$

откуда

$$r_{xy} = \frac{4\sqrt{3} \cdot 2\rho^2}{2\rho^2 \cdot 5{,}09 \cdot 4{,}58} = 0{,}297.$$

Определим угол α, который составляет с осью первая главная ось рассеивания:

$$\operatorname{tg} 2\alpha = \frac{2 \cdot 0{,}297 \cdot 5{,}09 \cdot 4{,}58}{5{,}09^2 - 4{,}58^2} = 1{,}340,$$

$$2\alpha \approx 69^\circ 30'; \qquad \alpha \approx 34^\circ 45'.$$

По формулам (12.8.8) имеем:

$$\sigma_\xi = \sqrt{\sigma_x^2 \cos^2\alpha + r_{xy}\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \sin^2\alpha} = 5{,}55 \ (м);$$

$$\sigma_\eta = \sqrt{\sigma_x^2 \sin^2\alpha - r_{xy}\sigma_x\sigma_y \sin 2\alpha + \sigma_y^2 \cos^2\alpha} = 4{,}22 \ (м).$$

ГЛАВА 13

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

13.1. Закон больших чисел и центральная предельная теорема

В начале курса мы уже говорили о том, что математические законы теории вероятностей получены абстрагированием реальных статистических закономерностей, свойственных массовым случайным явлениям. Наличие этих закономерностей связано именно с массовостью явлений, то есть с большим числом выполняемых однородных опытов или с большим числом складывающихся случайных воздействий, порождающих в своей совокупности случайную величину, подчиненную вполне определенному закону. Свойство устойчивости массовых случайных явлений известно человечеству еще с глубокой древности. В какой бы области оно ни проявлялось, суть его сводится к следующему: конкретные особенности каждого отдельного случайного явления почти не сказываются на среднем результате массы таких явлений; случайные отклонения от среднего, неизбежные в каждом отдельном явлении, в массе взаимно погашаются, нивелируются, выравниваются. Именно эта устойчивость средних и представляет собой физическое содержание «закона больших чисел», понимаемого в широком смысле слова: при очень большом числе случайных явлений средний их результат практически перестает быть случайным и может быть предсказан с большой степенью определенности.

В узком смысле слова под «законом больших чисел» в теории вероятностей понимается ряд математических теорем, в каждой из которых для тех или иных условий устанавливается факт приближения средних характеристик большого числа опытов к некоторым определенным постоянным.

В п° 2.3 мы уже формулировали простейшую из этих теорем — теорему Я. Бернулли. Она утверждает, что при большом числе опытов частота события приближается (точнее — сходится по вероятности) к вероятности этого события. С другими, более общими формами закона больших чисел мы познакомимся в данной главе. Все они устанавливают факт и условия сходимости по вероятности тех или иных случайных величин к постоянным, не случайным величинам.

Закон больших чисел играет важную роль в практических применениях теории вероятностей. Свойство случайных величин при определенных условиях вести себя практически как не случайные позволяет уверенно оперировать с этими величинами, предсказывать результаты массовых случайных явлений почти с полной определенностью.

Возможности таких предсказаний в области массовых случайных явлений еще больше расширяются наличием другой группы предельных теорем, касающихся уже не предельных значений случайных величин, а *предельных законов распределения*. Речь идет о группе теорем, известных под названием «центральной предельной теоремы». Мы уже говорили о том, что при суммировании достаточно большого числа случайных величин закон распределения суммы неограниченно приближается к нормальному при соблюдении некоторых условий. Эти условия, которые математически можно формулировать различным образом — в более или менее общем виде, — по существу сводятся к требованию, чтобы влияние на сумму отдельных слагаемых было *равномерно малым*, т. е. чтобы в состав суммы не входили члены, явно преобладающие над совокупностью остальных по своему влиянию на рассеивание суммы. Различные формы центральной предельной теоремы различаются между собой теми условиями, для которых устанавливается это предельное свойство суммы случайных величин.

Различные формы закона больших чисел вместе с различными формами центральной предельной теоремы образуют совокупность так называемых *предельных теорем* теории вероятностей. Предельные теоремы дают возможность не только осуществлять научные прогнозы в области случайных явлений, но и оценивать точность этих прогнозов.

В данной главе мы рассмотрим только некоторые, наиболее простые формы предельных теорем. Сначала будут рассмотрены теоремы, относящиеся к группе «закона больших чисел», затем — теоремы, относящиеся к группе «центральной предельной теоремы».

13.2. Неравенство Чебышева

В качестве леммы, необходимой для доказательства теорем, относящихся к группе «закона больших чисел», мы докажем одно весьма общее неравенство, известное под названием *неравенства Чебышева*.

Пусть имеется случайная величина X с математическим ожиданием m_x и дисперсией D_x. Неравенство Чебышева утверждает, что, каково бы ни было положительное число α, вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на α, ограничена сверху величиной $\frac{D_x}{\alpha^2}$:

$$P(|X - m_x| \geqslant \alpha) \leqslant \frac{D_x}{\alpha^2}. \qquad (13.2.1)$$

Доказательство. 1. Пусть величина X прерывная, с рядом распределения

$$\frac{x_l \mid x_1 \mid x_2 \mid \ldots \mid x_n}{p_l \mid p_1 \mid p_2 \mid \ldots \mid p_n}.$$

Изобразим возможные значения величины X и ее математическое ожидание m_x в виде точек на числовой оси Ox (рис. 13.2.1).

Зададимся некоторым значением $\alpha > 0$ и вычислим вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на α:

$$P(|X - m_x| \geqslant \alpha). \tag{13.2.2}$$

Для этого отложим от точки m_x вправо и влево по отрезку длиной α; получим отрезок AB. Вероятность (13.2.2) есть не что иное, как вероятность того, что случайная точка X попадет не внутрь отрезка AB, а вовне его:

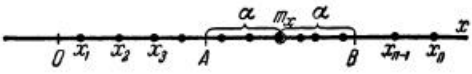


Рис. 13.2.1.

$$P(|X - m_x| \geqslant \alpha) = P(X \notin AB)^{1}).$$

Для того чтобы найти эту вероятность, нужно просуммировать вероятности всех тех значений x_l, которые лежат *вне* отрезка AB. Это мы запишем следующим образом:

$$P(|X - m_x| \geqslant \alpha) = \sum_{|x_l - m_x| \geqslant \alpha} p_l, \tag{13.2.3}$$

где запись $|x_l - m_x| \geqslant \alpha$ под знаком суммы означает, что суммирование распространяется на все те значения l, для которых точки x_l лежат вне отрезка AB.

С другой стороны, напишем выражение дисперсии величины X По определению:

$$D_x = M[(X - m_x)^2] = \sum_{l=1}^{n} (x_l - m_x)^2 p_l = \sum_{l=1}^{n} |x_l - m_x|^2 p_l. \tag{13.2.4}$$

Так как все члены суммы (13.2.4) неотрицательны, она может только уменьшиться, если мы распространим ее не на все значения x_l, а только на некоторые, в частности на те, которые лежат вне отрезка AB:

$$D_x \geqslant \sum_{|x_l - m_x| \geqslant \alpha} |x_l - m_x|^2 p_l. \tag{13.2.5}$$

---

[1]) Концы отрезка AB мы в него не включаем.

Заменим под знаком суммы выражение $|x_i - m_x|$ через α. Так как для всех членов суммы $|x_i - m_x| \geqslant \alpha$, то от такой замены сумма тоже может только уменьшиться; значит,

$$D_x \geqslant \sum_{|x_i - m_x| \geqslant \alpha} \alpha^2 p_i = \alpha^2 \sum_{|x_i - m_x| \geqslant \alpha} p_i. \qquad (13.2.6)$$

Но согласно формуле (13.2.3) сумма, стоящая в правой части (13.2.6), есть не что иное, как вероятность попадания случайной точки вовне отрезка AB; следовательно,

$$D_x \geqslant \alpha^2 P(|X - m_x| \geqslant \alpha),$$

откуда непосредственно вытекает доказываемое неравенство.

2. В случае, когда величина X непрерывна, доказательство проводится аналогичным образом с заменой вероятностей p_i элементом вероятности, а конечных сумм — интегралами. Действительно,

$$P(|X - m_x| > \alpha) = \int_{|x - m_x| > \alpha} f(x)\,dx\ ^1), \qquad (13.2.7)$$

где $f(x)$ — плотность распределения величины X. Далее, имеем:

$$D_x = \int_{-\infty}^{\infty} (x - m_x)^2 f(x)\,dx = \int_{-\infty}^{\infty} |x - m_x|^2 f(x)\,dx \geqslant$$

$$\geqslant \int_{|x - m_x| > \alpha} |x - m_x|^2 f(x)\,dx,$$

где знак $|x - m_x| > \alpha$ под интегралом означает, что интегрирование распространяется на внешнюю часть отрезка AB.

Заменяя $|x - m_x|$ под знаком интеграла через α, получим:

$$D_x \geqslant \alpha^2 \int_{|x - m_x| > \alpha} f(x)\,dx = \alpha^2 P(|X - m_x| > \alpha),$$

откуда и вытекает неравенство Чебышева для непрерывных величин.

Пример. Дана случайная величина X с математическим ожиданием m_x и дисперсией σ_x^2. Оценить сверху вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на $3\sigma_x$.

Решение. Полагая в неравенстве Чебышева $\alpha = 3\sigma_x$, имеем:

$$P(|X - m_x| \geqslant 3\sigma_x) \leqslant \frac{D_x}{9\sigma_x^2} = \frac{1}{9},$$

---

[1]) Знак $\geqslant$ заменен знаком $>$, так как для непрерывной величины вероятность точного равенства равна нулю.

т. е. вероятность того, что отклонение случайной величины от ее математического ожидания выйдет за пределы трех средних квадратических отклонений, не может быть больше $\frac{1}{9}$.

Примечание. Неравенство Чебышева дает только верхнюю границу вероятности данного отклонения. Выше этой границы вероятность не может быть ни при каком законе распределения. На практике в большинстве случаев вероятность того, что величина X выйдет за пределы участка $m_x \pm 3\sigma_x$, значительно меньше $\frac{1}{9}$. Например, для нормального закона эта вероятность приблизительно равна 0,003. На практике чаще всего мы имеем дело со случайными величинами, значения которых только крайне редко выходят за пределы $m_x \pm 3\sigma_x$. Если закон распределения случайной величины неизвестен, а известны только m_x и σ_x, на практике обычно считают отрезок $m_x \pm 3\sigma_x$ участком практически возможных значений случайной величины (так называемое «правило трех сигма»).

13.3. Закон больших чисел (теорема Чебышева)

В данном п° мы докажем одну из простейших, но вместе с тем наиболее важных форм закона больших чисел — теорему Чебышева. Эта теорема устанавливает связь между средним арифметическим наблюденных значений случайной величины и ее математическим ожиданием.

Предварительно решим следующую вспомогательную задачу.

Имеется случайная величина X с математическим ожиданием m_x и дисперсией D_x. Над этой величиной производится n независимых опытов и вычисляется среднее арифметическое всех наблюденных значений величины X. Требуется найти числовые характеристики этого среднего арифметического — математическое ожидание и дисперсию — и выяснить, как они изменяются с увеличением n.

Обозначим:

X_1 — значение величины X в первом опыте;

X_2 — значение величины X во втором опыте, и т. д.

Очевидно, совокупность величин $X_1, X_2, \ldots, X_n$ представляет собой n независимых случайных величин, каждая из которых распределена по тому же закону, что и сама величина X. Рассмотрим среднее арифметическое этих величин:

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Случайная величина Y есть линейная функция независимых случайных величин $X_1, X_2, \ldots, X_n$. Найдем математическое ожидание

и дисперсию этой величины. Согласно правилам n° 10 для определения числовых характеристик линейных функций получим:

$$m_y = M[Y] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = \frac{1}{n}\, n m_x = m_x;$$

$$D_y = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i] = \frac{D_x}{n}.$$

Итак, математическое ожидание величины Y не зависит от числа опытов n и равно математическому ожиданию наблюдаемой величины X; что касается дисперсии величины Y, то она неограниченно убывает с увеличением числа опытов и при достаточно большом n может быть сделана сколь угодно малой. Мы убеждаемся, что среднее арифметическое есть случайная величина со сколь угодно малой дисперсией и при большом числе опытов ведет себя почти как не случайная.

Теорема Чебышева и устанавливает в точной количественной форме это свойство устойчивости среднего арифметического. Она формулируется следующим образом:

При достаточно большом числе независимых опытов среднее арифметическое наблюденных значений случайной величины сходится по вероятности к ее математическому ожиданию.

Запишем теорему Чебышева в виде формулы. Для этого напомним смысл термина «сходится по вероятности». Говорят, что случайная величина X_n сходится по вероятности к величине a, если при увеличении n вероятность того, что X_n и a будут сколь угодно близки, неограниченно приближается к единице, а это значит, что при достаточно большом n

$$P(|X_n - a| < \varepsilon) > 1 - \delta,$$

где ε, δ — произвольно малые положительные числа.

Запишем в аналогичной форме теорему Чебышева. Она утверждает, что при увеличении n среднее арифметическое $\dfrac{\sum_{i=1}^{n} X_i}{n}$ сходится по вероятности к m_x, т. е.

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| < \varepsilon\right) > 1 - \delta. \qquad (13.3.1)$$

Докажем это неравенство.

Доказательство. Выше было показано, что величина

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}$$

имеет числовые характеристики

$$m_y = m_x; \quad D_y = \frac{D_x}{n}.$$

Применим к случайной величине Y неравенство Чебышева, полагая $\alpha = \varepsilon$:

$$P(|Y - m_y| \geqslant \varepsilon) \leqslant \frac{D_y}{\varepsilon^2} = \frac{D_x}{n\varepsilon^2}.$$

Как бы мало ни было число ε, можно взять n таким большим, чтобы выполнялось неравенство

$$\frac{D_x}{n\varepsilon^2} < \delta,$$

где δ — сколь угодно малое число.

Тогда

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| \geqslant \varepsilon\right) < \delta,$$

откуда, переходя к противоположному событию, имеем:

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| < \varepsilon\right) > 1 - \delta,$$

что и требовалось доказать.

13.4. Обобщенная теорема Чебышева. Теорема Маркова

Теорема Чебышева легко может быть обобщена на более сложный случай, а именно когда закон распределения случайной величины X от опыта к опыту не остается одним и тем же, а изменяется. Тогда вместо среднего арифметического наблюденных значений одной и той же величины X с постоянными математическим ожиданием и дисперсией мы имеем дело со средним арифметическим n различных случайных величин, с различными математическими ожиданиями и дисперсиями. Оказывается, что и в этом случае при соблюдении некоторых условий среднее арифметическое является устойчивым и сходится по вероятности к определенной неслучайной величине.

Обобщенная теорема Чебышева формулируется следующим образом. *Если*

$$X_1,\ X_2,\ \ldots,\ X_n —$$

независимые случайные величины с математическими ожиданиями

$$m_{x_1},\ m_{x_2},\ \ldots,\ m_{x_n}$$

и дисперсиями

$$D_{x_1},\ D_{x_2},\ \ldots,\ D_{x_n}$$

и если все дисперсии ограничены сверху одним и тем же числом L:

$$D_{x_i} < L \qquad (i = 1,\ 2,\ \ldots,\ n),$$

то при возрастании n *среднее арифметическое наблюденных значений величин* $X_1,\ X_2,\ \ldots,\ X_n$ *сходится по вероятности к среднему арифметическому их математических ожиданий* [1]).

Запишем эту теорему в виде формулы. Пусть ε, δ — сколь угодно малые положительные числа. Тогда при достаточно большом n

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| < \varepsilon\right) > 1 - \delta. \qquad (13.4.1)$$

Доказательство. Рассмотрим величину

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Ее математическое ожидание равно:

$$m_y = \frac{\sum_{i=1}^{n} m_{x_i}}{n},$$

а дисперсия

$$D_y = \frac{1}{n^2} \sum_{i=1}^{n} D_{x_i}.$$

---

[1]) То есть разность между тем и другим средним арифметическим сходится по вероятности к нулю.

Применим к величине Y неравенство Чебышева:

$$P(|Y - m_y| \geqslant \varepsilon) \leqslant \frac{D_y}{\varepsilon^2},$$

или

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| \geqslant \varepsilon\right) \leqslant \frac{\sum_{i=1}^{n} D_{x_i}}{n^2\varepsilon^2}. \qquad (13.4.2)$$

Заменим в правой части неравенства (13.4.2) каждую из величин D_{x_i} большей величиной L. Тогда неравенство только усилится:

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| \geqslant \varepsilon\right) < \frac{L}{n\varepsilon^2}.$$

Как бы мало ни было ε, можно выбрать n настолько большим, чтобы выполнялось неравенство

$$\frac{L}{n\varepsilon^2} < \delta;$$

тогда

$$P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| \geqslant \varepsilon\right) < \delta,$$

откуда, переходя к противоположному событию, получим доказываемое неравенство (13.4.1).

Закон больших чисел может быть распространен и на зависимые случайные величины. Обобщение закона больших чисел на случай зависимых случайных величин принадлежит А. А. Маркову.

Теорема Маркова. *Если имеются зависимые случайные величины* $X_1, X_2, \ldots, X_n$ *и если при* $n \to \infty$

$$\frac{D\left[\sum_{i=1}^{n} X_i\right]}{n^2} \to 0,$$

то среднее арифметическое наблюденных значений случайных величин $X_1, X_2, \ldots, X_n$ *сходится по вероятности к среднему арифметическому их математических ожиданий.*

Доказательство. Рассмотрим величину

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Очевидно,

$$D_y = \frac{D\left[\sum_{i=1}^{n} X_i\right]}{n^2}.$$

Применим к величине Y неравенство Чебышева:

$$P(|Y - m_y| \geqslant \varepsilon) \leqslant \frac{D_y}{\varepsilon^2}.$$

Так как по условию теоремы при $n \to \infty$ $D_y \to 0$, то при достаточно большом n

$$P(|Y - m_y| \geqslant \varepsilon) < \delta,$$

или, переходя к противоположному событию,

$$P(|Y - m_y| < \varepsilon) = P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| < \varepsilon\right) > 1 - \delta,$$

что и требовалось доказать.

13.5. Следствия закона больших чисел: теоремы Бернулли и Пуассона

Известная теорема Я. Бернулли, устанавливающая связь между частотой события и его вероятностью, может быть доказана как прямое следствие закона больших чисел.

Пусть производится n независимых опытов, в каждом из которых может появиться или не появиться некоторое событие A, вероятность которого в каждом опыте равна p. Теорема Я. Бернулли утверждает, что *при неограниченном увеличении числа опытов n частота события A сходится по вероятности к его вероятности p.*

Обозначим частоту события A в n опытах через P^* и запишем теорему Я. Бернулли в виде формулы

$$P(|P^* - p| < \varepsilon) > 1 - \delta, \qquad (13.5.1)$$

где ε, δ — сколь угодно малые положительные числа.

Требуется доказать справедливость этой формулы при достаточно большом n.

Доказательство. Рассмотрим независимые случайные величины:

X_1 — число появлений события A в первом опыте;

X_2 — число появлений события A во втором опыте, и т. д.

Все эти величины прерывны и имеют один и тот же закон распределения, выражаемый рядом вида:

$$\begin{array}{c|c} 0 & 1 \\ \hline q & p \end{array},$$

где $q = 1 - p$. Математическое ожидание каждой из величин X_i равно p, а ее дисперсия pq (см. п° 10.3).

Частота P^* представляет собой не что иное, как среднее арифметическое величин $X_1, X_2, \ldots, X_n$:

$$P^* = \frac{\sum_{i=1}^{n} X_i}{n},$$

и, согласно закону больших чисел, сходится по вероятности к общему математическому ожиданию этих случайных величин. Отсюда и следует справедливость неравенства (13.5.1).

Теорема Я. Бернулли утверждает устойчивость частоты при постоянных условиях опыта. Но при изменяющихся условиях опыта аналогичная устойчивость также существует. Теорема, устанавливающая свойство устойчивости частот при переменных условиях опыта, называется *теоремой Пуассона* и формулируется следующим образом:

Если производится n независимых опытов и вероятность появления события A в i-м опыте равна p_i, то при увеличении n частота события A сходится по вероятности к среднему арифметическому вероятностей p_i.

Теорема Пуассона выводится из обобщенной теоремы Чебышева точно так же, как теорема Бернулли была выведена из закона больших чисел.

Теорема Пуассона имеет большое принципиальное значение для практического применения теории вероятностей. Дело в том, что зачастую вероятностные методы применяются для исследования явлений, которые в одних и тех же условиях не имеют шансов повториться достаточно много раз, но повторяются многократно при весьма разнообразных условиях, причем вероятности интересующих нас событий сильно зависят от этих условий. Например, вероятность поражения цели в воздушном бою существенно зависит от дальности стрельбы, ракурса цели, высоты полета, скорости стреляющего самолета и цели и т. д. Комплекс этих условий слишком многочислен для того, чтобы можно было рассчитывать на многократное осуществление воздушного боя именно в данных фиксированных условиях. И все же, несмотря на это, в данном явлении налицо определенная устойчивость частот, а именно частота поражения цели в реальных воздушных боях, осуществляемых в самых разных условиях, будет

приближаться к средней вероятности поражения цели, характерной для данной группы условий. Поэтому те методы организации стрельбы, которые основаны на максимальной вероятности поражения цели, будут оправданы и в данном случае, несмотря на то, что нельзя ожидать подлинной массовости опытов в каждом определенном комплексе условий.

Аналогичным образом обстоит дело в области опытной проверки вероятностных расчетов. На практике очень часто встречается случай, когда требуется проверить на опыте соответствие вычисленной вероятности какого-либо события A его фактической частоте. Чаще всего это делается для того, чтобы проверить правильность той или иной теоретической схемы, положенной в основу метода вычисления вероятности события. Зачастую при такой экспериментальной проверке не удается воспроизвести достаточно много раз одни и те же условия опыта. И все же эта проверка может быть осуществлена, если сравнить наблюденную в опыте частоту события не с его вероятностью для фиксированных условий, а со средним арифметическим вероятностей, вычисленных для различных условий.

13.6. Массовые случайные явления и центральная предельная теорема

В предыдущих п°п° мы рассмотрели различные формы закона больших чисел. Все эти формы, как бы они ни были различны, утверждают одно: факт сходимости по вероятности тех или иных случайных величин к определенным постоянным. Ни в одной из форм закона больших чисел мы не имеем дела с законами распределения случайных величин. Предельные законы распределения составляют предмет другой группы теорем — центральной предельной теоремы, которую иногда называют «количественной формой закона больших чисел».

Все формы центральной предельной теоремы посвящены установлению условий, при которых возникает нормальный закон распределения. Так как эти условия на практике весьма часто выполняются, нормальный закон является самым распространенным из законов распределения, наиболее часто встречающимся в случайных явлениях природы. Он возникает во всех случаях, когда исследуемая случайная величина может быть представлена в виде суммы достаточно большого числа независимых (или слабо зависимых) элементарных слагаемых, каждое из которых в отдельности сравнительно мало влияет на сумму.

В теории стрельбы нормальный закон распределения играет особо важную роль, так как в большинстве случаев практики координаты точек попадания и точек разрыва снарядов распределяются по нормальному закону. Объяснить это можно на следующем примере.

Пусть производится стрельба по некоторой плоской мишени, с центром которой (точкой прицеливания) связано начало координат. Точка попадания характеризуется двумя случайными величинами: X и Y. Рассмотрим одну из них, например отклонение X точки попадания от цели в направлении оси Ox. Это отклонение вызвано совокупным действием очень большого количества сравнительно малых факторов, как-то: ошибка наводки, ошибка в определении дальности до цели, вибрации орудия и установки при стрельбе, ошибки изготовления снаряда, атмосферные условия и т. д. Каждая из этих причин создает элементарную ошибку — отклонение снаряда от цели, и координата снаряда X может быть представлена как сумма таких элементарных отклонений:

$$X = X_1 + X_2 + \ldots + X_n + \ldots, \qquad (13.6.1)$$

где X_1, X_2, ... — отклонения, вызванные отдельными факторами. Так как этих факторов очень много, между собой они являются в основном независимыми и по влиянию на сумму отдельные слагаемые можно считать приблизительно равномерно малыми, то налицо условия применимости центральной предельной теоремы, и величина (13.6.1) должна подчиняться закону распределения, близкому к нормальному.

Остановимся несколько подробнее на нашем утверждении о приблизительно равномерно малом влиянии каждого из слагаемых на сумму. Смысл его в том, что среди элементарных ошибок стрельбы нет ни одной резко превалирующей над суммой всех остальных. Действительно, если бы такая ошибка была, нужно думать, что, составляя правила стрельбы или конструируя прицельный прибор, мы постарались бы ликвидировать эту ошибку и учесть заранее самую значительную причину, отклоняющую снаряд от цели. Неучтенные случайные факторы, создающие рассеивание, обычно характерны своей равномерной малостью и отсутствием среди них резко преобладающих. Именно поэтому закон распределения точек попадания снарядов (или закон распределения точек разрыва снарядов при дистанционной стрельбе) обычно принимается нормальным [1]).

Нормальный закон распределения является доминирующим не только в теории стрельбы, но и во многих других областях, например в теории ошибок измерения. Именно исходя из теории ошибок измерения нормальный закон и был впервые обоснован Лапласом и Гауссом. Действительно, в большинстве случаев ошибки, возникаю-

[1]) В некоторых случаях стрельбы фактическое распределение точек попадания на плоскости может сильно отличаться от нормального, например, при стрельбе в резко переменных условиях, когда центр рассеивания и вероятное отклонение в процессе стрельбы заметно меняются. Однако в таких случаях мы фактически имеем дело не с законом распределения координат точки попадания при одном выстреле, а со средним из таких законов для различных выстрелов.

щие при измерении тех или иных физических величин, распределяются именно по нормальному закону; причина этого в том, что такие ошибки, как правило, складываются из многочисленных независимых элементарных ошибок, порождаемых различными причинами. Долгое время нормальный закон считался единственным и универсальным законом ошибок. В настоящее время взгляд на нормальный закон как на единственный и универсальный должен быть пересмотрен (опыт показывает, что в ряде процессов измерения и производства наблюдаются законы распределения, отличные от нормального), но все же нормальный закон остается самым распространенным и самым важным для практики законом ошибок.

13.7. Характеристические функции

Одна из наиболее общих форм центральной предельной теоремы была доказана А. М. Ляпуновым в 1900 г. Для доказательства этой теоремы А. М. Ляпунов создал специальный метод характеристических функций. В дальнейшем этот метод приобрел самостоятельное значение и оказался весьма мощным и гибким методом, пригодным для решения самых различных вероятностных задач.

Характеристической функцией случайной величины X называется функция

$$g(t) = M[e^{itX}], \tag{13.7.1}$$

где i — мнимая единица. Функция $g(t)$ представляет собой математическое ожидание некоторой комплексной случайной величины

$$U = e^{itX},$$

функционально связанной с величиной X. При решении многих задач теории вероятностей оказывается удобнее пользоваться характеристическими функциями, чем законами распределения.

Зная закон распределения случайной величины X, легко найти ее характеристическую функцию.

Если X — прерывная случайная величина с рядом распределения

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

то ее характеристическая функция

$$g(t) = \sum_{k=1}^{n} e^{itx_k} p_k. \tag{13.7.2}$$

Если X — непрерывная случайная величина с плотностью распределения $f(x)$, то ее характеристическая функция

$$g(t)=\int_{-\infty}^{\infty} e^{itx} f(x)\,dx. \tag{13.7.3}$$

Пример 1. Случайная величина X — число попаданий при одном выстреле. Вероятность попадания равна p. Найти характеристическую функцию случайной величины X.

Решение. По формуле (13.7.2) имеем:

$$g(t)=e^{it\cdot 0}(1-p)+e^{it\cdot 1}p=q+e^{it}p,$$

где $q=1-p$.

Пример 2. Случайная величина X имеет нормальное распределение:

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}. \tag{13.7.4}$$

Определить ее характеристическую функцию.

Решение. По формуле (13.7.3) имеем:

$$g(t)=\int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx-\frac{x^2}{2}}dx. \tag{13.7.5}$$

Пользуясь известной формулой

$$\int_{-\infty}^{\infty} e^{-Ax^2\pm 2Bx-C}dx=\sqrt{\frac{\pi}{A}}e^{-\frac{AC-B^2}{A}}$$

и имея в виду, что $i^2=-1$, получим:

$$g(t)=e^{-\frac{t^2}{2}}. \tag{13.7.6}$$

Формула (13.7.3) выражает характеристическую функцию $g(t)$ непрерывной случайной величины X через ее плотность распределения $f(x)$. Преобразование (13.7.3), которому нужно подвергнуть $f(x)$, чтобы получить $g(t)$, называется *преобразованием Фурье*. В курсах математического анализа доказывается, что если функция $g(t)$ выражается через $f(x)$ с помощью преобразования Фурье, то, в свою очередь, функция $f(x)$ выражается через $g(t)$ с помощью так называемого *обратного преобразования Фурье*:

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}g(t)\,dt\ ^{1}). \tag{13.7.7}$$

---

[1]) В п°п° 17.3, 17.4 будут выведены те же преобразования Фурье, связывающие корреляционную функцию и спектральную плотность.

Сформулируем и докажем основные свойства характеристических функций.

1. *Если случайные величины X и Y связаны соотношением*

$$Y = aX,$$

где a — *неслучайный множитель, то их характеристические функции связаны соотношением*:

$$g_y(t) = g_x(at). \tag{13.7.8}$$

Доказательство:

$$g_y(t) = M[e^{itY}] = M[e^{itaX}] = M[e^{i(at)X}] = g_x(at).$$

2. *Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых.*

Доказательство. Даны $X_1, X_2, \ldots, X_n$ — независимые случайные величины с характеристическими функциями

$$g_{x_1}(t),\ g_{x_2}(t),\ \ldots,\ g_{x_n}(t)$$

и их сумма

$$Y = \sum_{k=1}^{n} X_k.$$

Требуется доказать, что

$$g_y(t) = \prod_{k=1}^{n} g_{x_k}(t). \tag{13.7.9}$$

Имеем

$$g_y(t) = M[e^{itY}] = M\left[e^{it\sum_{k=1}^{n} X_k}\right] = M\left[\prod_{k=1}^{n} e^{itX_k}\right].$$

Так как величины X_k независимы, то независимы и их функции e^{itX_k}. По теореме умножения математических ожиданий получим:

$$g_y(t) = \prod_{k=1}^{n} M[e^{itX_k}] = \prod_{k=1}^{n} g_{x_k}(t),$$

что и требовалось доказать.

Аппарат характеристических функций часто применяется для композиции законов распределения. Пусть, например, имеются две независимые случайные величины X и Y с плотностями распределения $f_1(x)$ и $f_2(y)$. Требуется найти плотность распределения величины

$$Z = X + Y.$$

Это можно выполнить следующим образом: найти характеристические функции $g_x(t)$ и $g_y(t)$ случайных величин X и Y и, перемножив

их, получить характеристическую функцию величины Z:

$$g_z(t) = g_x(t)\, g_y(t),$$

после чего, подвергнув $g_z(t)$ обратному преобразованию Фурье, найти плотность распределения величины Z:

$$f_3(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} g_z(t)\, dt.$$

Пример 3. Найти с помощью характеристических функций композицию двух нормальных законов:

$f_1(x)$ с характеристиками $m_x = 0$; σ_x;

$f_2(y)$ с характеристиками $m_y = 0$; σ_y.

Решение. Находим характеристическую функцию величины X. Для этого представим ее в виде

$$X = \sigma_x U,$$

где $m_u = 0$; $\sigma_u = 1$.

Пользуясь результатом примера 2, найдем

$$g_u(t) = e^{-\frac{t^2}{2}}.$$

Согласно свойству 1 характеристических функций,

$$g_x(t) = g_u(\sigma_x t) = e^{-\frac{(\sigma_x t)^2}{2}}.$$

Аналогично

$$g_y(t) = e^{-\frac{(\sigma_y t)^2}{2}}.$$

Перемножая $g_x(t)$ и $g_y(t)$, имеем:

$$g_z(t) = e^{-\frac{\left(\sigma_x^2 + \sigma_y^2\right)}{2} t^2},$$

а это есть характеристическая функция нормального закона с параметрами $m_z = 0$; $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$. Таким образом, получена композиция нормальных законов гораздо более простыми средствами, чем в п° 12.6.

13.8. Центральная предельная теорема для одинаково распределенных слагаемых

Различные формы центральной предельной теоремы отличаются между собой условиями, накладываемыми на распределения образующих сумму случайных слагаемых. Здесь мы сформулируем и докажем одну из самых простых форм центральной предельной теоремы, относящуюся к случаю одинаково распределенных слагаемых.

Теорема. *Если $X_1, X_2, \ldots, X_n, \ldots$ — независимые случайные величины, имеющие один и тот же закон распределения с математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон распределения суммы*

$$Y_n = \sum_{k=1}^{n} X_k \tag{13.8.1}$$

неограниченно приближается к нормальному.

Доказательство.

Проведем доказательство для случая непрерывных случайных величин $X_1, \ldots, X_n$ (для прерывных оно будет аналогичным).

Согласно второму свойству характеристических функций, доказанному в предыдущем n°, характеристическая функция величины Y_n представляет собой произведение характеристических функций слагаемых. Случайные величины $X_1, \ldots, X_n$ имеют один и тот же закон распределения с плотностью $f(x)$ и, следовательно, одну и ту же характеристическую функцию

$$g_x(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx. \tag{13.8.2}$$

Следовательно, характеристическая функция случайной величины Y_n будет

$$g_{y_n}(t) = [g_x(t)]^n. \tag{13.8.3}$$

Исследуем более подробно функцию $g_x(t)$. Представим ее в окрестности точки $t = 0$ по формуле Маклорена с тремя членами:

$$g_x(t) = g_x(0) + g_x'(0)\, t + \left[\frac{g_x''(0)}{2} + \alpha(t)\right] t^2, \tag{13.8.4}$$

где $\alpha(t) \to 0$ при $t \to 0$.

Найдем величины $g_x(0)$, $g_x'(0)$, $g_x''(0)$. Полагая в формуле (13.8.2) $t = 0$, имеем:

$$g_x(0) = \int_{-\infty}^{\infty} f(x)\, dx = 1. \tag{13.8.5}$$

Продифференцируем (13.8.2) по t:

$$g_x'(t) = \int_{-\infty}^{\infty} ixe^{itx} f(x)\, dx = i \int_{-\infty}^{\infty} xe^{itx} f(x)\, dx. \tag{13.8.6}$$

Полагая в (13.8.6) $t=0$, получим:

$$g_x'(0)=i\int_{-\infty}^{\infty}xf(x)\,dx=iM[X]=im. \tag{13.8.7}$$

Очевидно, не ограничивая общности, можно положить $m=0$ (для этого достаточно перенести начало отсчета в точку m). Тогда

$$g_x'(0)=0.$$

Продифференцируем (13.8.6) еще раз:

$$g_x''(t)=-\int_{-\infty}^{\infty}x^2e^{itx}f(x)\,dx,$$

отсюда

$$g_x''(0)=-\int_{-\infty}^{\infty}x^2f(x)\,dx. \tag{13.8.8}$$

При $m=0$ интеграл в выражении (13.8.8) есть не что иное, как дисперсия величины X с плотностью $f(x)$, следовательно

$$g_x''(0)=-\sigma^2. \tag{13.8.9}$$

Подставляя в (13.8.4) $g_x(0)=1$, $g_x'(0)=0$ и $g_x''(0)=-\sigma^2$, получим:

$$g_x(t)=1-\left[\frac{\sigma^2}{2}-\alpha(t)\right]t^2. \tag{13.8.10}$$

Обратимся к случайной величине Y_n. Мы хотим доказать, что ее закон распределения при увеличении n приближается к нормальному. Для этого перейдем от величины Y_n к другой («нормированной») случайной величине

$$Z_n=\frac{Y_n}{\sigma\sqrt{n}}. \tag{13.8.11}$$

Эта величина удобна тем, что ее дисперсия не зависит от n и равна единице при любом n. В этом нетрудно убедиться, рассматривая величину Z_n как линейную функцию независимых случайных величин $X_1, X_2, \ldots, X_n$, каждая из которых имеет дисперсию σ^2. Если мы докажем, что закон распределения величины Z_n приближается к нормальному, то, очевидно, это будет справедливо и для величины Y_n, связанной с Z_n линейной зависимостью (13.8.11).

Вместо того чтобы доказывать, что закон распределения величины Z_n при увеличении n приближается к нормальному, покажем, что ее характеристическая функция приближается к характеристической функции нормального закона[1]).

---

[1]) Здесь мы принимаем без доказательства, что из сходимости характеристических функций следует сходимость законов распределения. Доказательство см., например, Б. В. Гнеденко, Курс теории вероятностей, 1961.

Найдем характеристическую функцию величины Z_n. Из соотношения (13.8.11), согласно первому свойству характеристических функций (13.7.8), получим

$$g_{z_n}(t) = g_{y_n}\left(\frac{t}{\sigma\sqrt{n}}\right), \tag{13.8.12}$$

где $g_{y_n}(t)$ — характеристическая функция случайной величины Y_n.

Из формул (13.8.12) и (13.8.3) получим

$$g_{z_n}(t) = \left[g_x\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \tag{13.8.13}$$

или, пользуясь формулой (13.8.10),

$$g_{z_n}(t) = \left\{1 - \left[\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\right]\frac{t^2}{n\sigma^2}\right\}^n. \tag{13.8.14}$$

Прологарифмируем выражение (13.8.14):

$$\ln g_{z_n}(t) = n \ln\left\{1 - \left[\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\right]\frac{t^2}{n\sigma^2}\right\}.$$

Введем обозначение

$$\left[\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\right]\frac{t^2}{n\sigma^2} = \varkappa. \tag{13.8.15}$$

Тогда

$$\ln g_{z_n}(t) = n \ln\{1 - \varkappa\}. \tag{13.8.16}$$

Будем неограниченно увеличивать n. При этом величина $\varkappa$, согласно формуле (13.8.15), стремится к нулю. При значительном n ее можно считать весьма малой. Разложим $\ln\{1 - \varkappa\}$ в ряд и ограничимся одним членом разложения (остальные при $n \to \infty$ станут пренебрежимо малыми):

$$\ln\{1 - \varkappa\} \approx -\varkappa.$$

Тогда получим

$$\lim_{n\to\infty} \ln g_{z_n}(t) = \lim_{n\to\infty} n \cdot (-\varkappa) =$$

$$= \lim_{n\to\infty}\left\{-\frac{t^2}{2} + \alpha\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{\sigma^2}\right\} = -\frac{t^2}{2} + \lim_{n\to\infty}\frac{t^2}{\sigma^2}\alpha\left(\frac{t}{\sigma\sqrt{n}}\right).$$

По определению функция $\alpha(t)$ стремится к нулю при $t \to 0$; следовательно,

$$\lim_{n\to\infty} \alpha\left(\frac{t}{\sigma\sqrt{n}}\right) = 0$$

и

$$\lim_{n\to\infty} \ln g_{z_n}(t) = -\frac{t^2}{2},$$

откуда

$$\lim_{n\to\infty} g_{z_n}(t) = e^{-\frac{t^2}{2}}. \qquad (13.8.17)$$

Это есть не что иное, как характеристическая функция нормального закона с параметрами $m=0$, $\sigma=1$ (см. пример 2, п° 13.7).

Таким образом, доказано, что при увеличении n характеристическая функция случайной величины Z_n неограниченно приближается к характеристической функции нормального закона; отсюда заключаем, что и закон распределения величины Z_n (а значит и величины Y_n) неограниченно приближается к нормальному закону. Теорема доказана.

Мы доказали центральную предельную теорему для частного, но важного случая одинаково распределенных слагаемых. Однако в достаточно широком классе условий она справедлива и для неодинаково распределенных слагаемых. Например, А. М. Ляпунов доказал центральную предельную теорему для следующих условий:

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} b_k}{\left\{\sum_{k=1}^{n} D_k\right\}^{3/2}} = 0, \qquad (13.8.18)$$

где b_k — третий абсолютный центральный момент величины X_k:

$$b_k = \nu_3[X_k] = M[|\mathring{X}_k|^3] \quad (k=1, \ldots, n),$$

D_k — дисперсия величины X_k.

Наиболее общим (необходимым и достаточным) условием справедливости центральной предельной теоремы является условие Линдеберга: *при любом* $\tau > 0$

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int_{|x-m_k|>\tau B_n} (x-m_k)^2 f_k(x)\,dx = 0,$$

где m_k — *математическое ожидание,* $f_k(x)$ — *плотность распределения случайной величины* X_k, $B_n = \sqrt{\sum_{k=1}^{n} D_k}$.

13.9. Формулы, выражающие центральную предельную теорему и встречающиеся при ее практическом применении

Согласно центральной предельной теореме, закон распределения суммы достаточно большого числа независимых случайных величин (при соблюдении некоторых нежестких ограничений) сколь угодно близок к нормальному.

Практически центральной предельной теоремой можно пользоваться и тогда, когда речь идет о сумме сравнительно небольшого числа случайных величин. При суммировании независимых случайных величин, сравнимых по своему рассеиванию, с увеличением числа слагаемых закон распределения суммы очень скоро становится приблизительно нормальным. На практике вообще широко применяется приближенная замена одних законов распределения другими; при той сравнительно малой точности, которая требуется от вероятностных расчетов, такая замена тоже может быть сделана крайне приближенно. Опыт показывает, что когда число слагаемых порядка десяти (а часто и меньше), закон распределения суммы обычно может быть заменен нормальным.

В практических задачах часто применяют центральную предельную теорему для вычисления вероятности того, что сумма нескольких случайных величин окажется в заданных пределах.

Пусть $X_1, X_2, \ldots, X_n$ — независимые случайные величины с математическими ожиданиями

$$m_1, m_2, \ldots, m_n$$

и дисперсиями

$$D_1, D_2, \ldots, D_n.$$

Предположим, что условия центральной предельной теоремы выполнены (величины $X_1, X_2, \ldots, X_n$ сравнимы по порядку своего влияния на рассеивание суммы) и число слагаемых n достаточно для того, чтобы закон распределения величины

$$Y = \sum_{i=1}^{n} X_i \tag{13.9.1}$$

можно было считать приближенно нормальным.

Тогда вероятность того, что случайная величина Y попадает в пределы участка (α, β), выражается формулой

$$P(\alpha < Y < \beta) = \Phi^*\left(\frac{\beta - m_y}{\sigma_y}\right) - \Phi^*\left(\frac{\alpha - m_y}{\sigma_y}\right), \tag{13.9.2}$$

где m_y, σ_y — математическое ожидание и среднее квадратическое отклонение величины Y, Φ^* — нормальная функция распределения.

Согласно теоремам сложения математических ожиданий и дисперсий

$$\left.\begin{aligned} m_y &= \sum_{i=1}^{n} m_i, \\ \sigma_y &= \sqrt{D_y} = \sqrt{\sum_{i=1}^{n} D_i}. \end{aligned}\right\} \tag{13.9.3}$$

Таким образом, для того чтобы приближенно найти вероятность попадания суммы большого числа случайных величин на заданный участок, не требуется знать законы распределения этих величин; достаточно знать лишь их характеристики. Разумеется, это относится только к случаю, когда выполнено основное условие центральной предельной теоремы — равномерно малое влияние слагаемых на рассеивание суммы.

Кроме формул типа (13.9.2), на практике часто применяются формулы, в которых вместо суммы случайных величин X_i фигурирует их нормированная сумма

$$Z = \frac{\mathring{Y}}{\sigma_y} = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} m_i}{\sqrt{\sum_{i=1}^{n} D_i}}. \qquad (13.9.4)$$

Очевидно,

$$M[Z] = 0; \quad D[Z] = \sigma_z = 1.$$

Если закон распределения величины Y близок к нормальному с параметрами (13.9.3), то закон распределения величины Z близок к нормальному с параметрами $m_z = 0$, $\sigma_z = 1$. Отсюда

$$P(\alpha < Z < \beta) = \Phi^*(\beta) - \Phi^*(\alpha). \qquad (13.9.5)$$

Заметим, что центральная предельная теорема может применяться не только к непрерывным, но и к дискретным случайным величинам при условии, что мы будем оперировать не плотностями, а функциями распределения. Действительно, если величины $X_1, X_2, \ldots, X_n$ дискретны, то их сумма X — также дискретная случайная величина и поэтому, строго говоря, не может подчиняться нормальному закону. Однако все формулы типа (13.9.2), (13.9.5) остаются в силе, так как в них фигурируют не плотности, а функции распределения. Можно доказать, что если дискретные случайные величины удовлетворяют условиям центральной предельной теоремы, то функция распределения их нормированной суммы Z (см. формулу (13.9.4)) при увеличении n неограниченно приближается к нормальной функции распределения с параметрами $m_z = 0$, $\sigma_z = 1$.

Частным случаем центральной предельной теоремы для дискретных случайных величин является теорема Лапласа.

Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p, то справедливо соотношение

$$P\left(\alpha < \frac{Y - np}{\sqrt{npq}} < \beta\right) = \Phi^*(\beta) - \Phi^*(\alpha), \qquad (13.9.6)$$

где Y — число появлений события A в n опытах, $q = 1 - p$.

Доказательство. Пусть производится n независимых опытов, в каждом из которых с вероятностью p может появиться событие A. Представим случайную величину Y — общее число появлений события в n опытах — в виде суммы

$$Y = \sum_{i=1}^{n} X_i, \tag{13.9.7}$$

где X_i — число появлений события A в i-м опыте.

Согласно доказанной в п° 13.8 теореме, закон распределения суммы одинаково распределенных слагаемых при увеличении их числа приближается к нормальному закону. Следовательно, при достаточно большом n справедлива формула (13.9.5), где

$$Z = \frac{Y - m_y}{\sigma_y}. \tag{13.9.8}$$

В п° 10.3 мы доказали, что математическое ожидание и дисперсия числа появлений события в n независимых опытах равны:

$$m_y = np; \quad D_y = npq \quad (q = 1 - p).$$

Подставляя эти выражения в (13.9.8), получим

$$Z = \frac{Y - np}{\sqrt{npq}},$$

и формула (13.9.5) примет вид:

$$P\left(\alpha < \frac{Y - np}{\sqrt{npq}} < \beta\right) = \Phi^*(\beta) - \Phi^*(\alpha).$$

Теорема доказана.

Пример 1. По полосе укреплений противника сбрасывается 100 серий бомб. При сбрасывании одной такой серии математическое ожидание числа попаданий равно 2, а среднее квадратическое отклонение числа попаданий равно 1,5. Найти приближенно вероятность того, что при сбрасывании 100 серий в полосу попадает от 180 до 220 бомб.

Решение. Представим общее число попаданий как сумму чисел попаданий бомб в отдельных сериях:

$$X = X_1 + X_2 + \ldots + X_{100} = \sum_{i=1}^{100} X_i,$$

где X_i — число попаданий i-й серии.

Условия центральной предельной теоремы соблюдены, так как величины $X_1, X_2, \ldots, X_{100}$ распределены одинаково. Будем считать число $n = 100$ достаточным для того, чтобы можно было применить предельную теорему (на практике она обычно применима и при гораздо меньших n). Имеем:

$$m_x = \sum_{i=1}^{100} m_i = 200, \qquad \sum_{i=1}^{n} D_i = \sum_{i=1}^{100} 1{,}5^2 = 225.$$

Применяя формулу (13.9.6), получим:

$$P(180 < X < 220) = \Phi^*\left(\frac{220-200}{\sqrt{225}}\right) - \Phi^*\left(\frac{180-200}{\sqrt{225}}\right) \approx 0{,}82,$$

т. е. с вероятностью 0,82 можно утверждать, что общее число попаданий в полосу не выйдет за пределы $180 \div 220$.

Пример 2. Происходит групповой воздушный бой, в котором участвуют 50 бомбардировщиков и 100 истребителей. Каждый бомбардировщик атакуется двумя истребителями; таким образом, воздушный бой распадается на 50 элементарных воздушных боев, в каждом из которых участвует один бомбардировщик и два истребителя. В каждом элементарном бою вероятность сбития бомбардировщика равна 0,4; вероятность того, что в элементарном бою будут сбиты оба истребителя, равна 0,2; вероятность того, что будет сбит ровно один истребитель, равна 0,5. Требуется: 1) найти вероятность того, что в воздушном бою будет сбито не менее 35% бомбардировщиков; 2) оценить границы, в которых с вероятностью 0,9 будет заключено число сбитых истребителей.

Решение. 1) Обозначим X — число сбитых бомбардировщиков:

$$X = \sum_{i=1}^{50} X_i,$$

где X_i — число бомбардировщиков, сбитых в i-м элементарном бою.

Ряд распределения величины X_i имеет вид:

0	1
0,6	0,4

Отсюда

$$m_{x_i} = 0{,}4; \quad D_{x_i} = 0{,}4 \cdot 0{,}6 = 0{,}24; \quad m_x = 50 \cdot 0{,}4 = 20;$$

$$\sigma_x = \sqrt{50 \cdot 0{,}24} \approx 3{,}464.$$

Применяя формулу (13.9.6) и полагая $\beta = 50$ (или, что в данном случае равносильно, $\beta = \infty$), $\alpha = 17$, находим:

$$P(17 < X) = \frac{1}{2} - \Phi^*\left(\frac{17-20}{3{,}464}\right) \approx 0{,}807.$$

2) Обозначим Y число сбитых истребителей:

$$Y = \sum_{i=1}^{50} Y_i,$$

где Y_i — число истребителей, сбитых в i-м элементарном бою.

Ряд распределения величины Y_i имеет вид:

0	1	2
0,3	0,5	0,2

Отсюда находим математическое ожидание и дисперсию величины Y_i:

$$m_{y_i} = 0{,}9; \quad D_{y_i} = 0{,}49.$$

Для величины Y:

$$m_y = 50 \cdot 0{,}9 = 45; \quad D_y = 24{,}5; \quad \sigma_y = 4{,}96.$$

Определим границы участка, симметричного относительно $m_y = 45$, в который с вероятностью 0,9 попадет величина Y. Обозначим половину длины этого участка l. Тогда

$$P(|Y - m_y| < l) = 2\Phi^*\left(\frac{l}{\sigma_y}\right) - 1 = 0{,}9,$$

$$\Phi^*\left(\frac{l}{\sigma_y}\right) = 0{,}95.$$

По таблицам функции $\Phi^*\left(\frac{l}{\sigma_y}\right)$ находим то значение аргумента, для которого $\Phi^*(x) = 0{,}95$; это значение приближенно равно

$$x = 1{,}645,$$

т. е.

$$\frac{l}{\sigma_y} = 1{,}645,$$

откуда

$$l = 8{,}14 \approx 8.$$

Следовательно, с вероятностью около 0,9 можно утверждать, что число сбитых истребителей будет заключено в пределах $m_y \pm l$, т. е. в пределах от 37 до 53.

ГЛАВА 14

ОБРАБОТКА ОПЫТОВ

14.1. Особенности обработки ограниченного числа опытов. Оценки для неизвестных параметров закона распределения

В главе 7 мы уже рассмотрели некоторые задачи математической статистики, относящиеся к обработке опытных данных. Это были главным образом задачи о нахождении законов распределения случайных величин по результатам опытов. Чтобы найти закон распределения, нужно располагать достаточно обширным статистическим материалом, порядка нескольких сотен опытов (наблюдений). Однако на практике часто приходится иметь дело со статистическим материалом весьма ограниченного объема — с двумя-тремя десятками наблюдений, часто даже меньше. Это обычно связано с дороговизной и сложностью постановки каждого опыта. Такого ограниченного материала явно недостаточно для того, чтобы найти заранее неизвестный закон распределения случайной величины; но все же этот материал может быть обработан и использован для получения некоторых сведений о случайной величине. Например, на основе ограниченного статистического материала можно определить — хотя бы ориентировочно — важнейшие числовые характеристики случайной величины: математическое ожидание, дисперсию, иногда — высшие моменты. На практике часто бывает, что вид закона распределения известен заранее, а требуется найти только некоторые параметры, от которых он зависит. Например, если заранее известно, что закон распределения случайной величины нормальный, то задача обработки сводится к определению двух его параметров m и σ. Если заранее известно, что величина распределена по закону Пуассона, то подлежит определению только один его параметр: математическое ожидание a. Наконец, в некоторых задачах вид закона распределения вообще несуществен, а требуется знать только его числовые характеристики.

В данной главе мы рассмотрим ряд задач об определении неизвестных параметров, от которых зависит закон распределения случайной величины, по ограниченному числу опытов.

Прежде всего нужно отметить, что любое значение искомого параметра, вычисленное на основе ограниченного числа опытов, всегда будет содержать элемент случайности. Такое приближенное, случайное значение мы будем называть *оценкой* параметра. Например, оценкой для математического ожидания может служить среднее арифметическое наблюденных значений случайной величины в n независимых опытах. При очень большом числе опытов среднее арифметическое будет с большой вероятностью весьма близко к математическому ожиданию. Если же число опытов n невелико, то замена математического ожидания средним арифметическим приводит к какой-то ошибке. Эта ошибка в среднем тем больше, чем меньше число опытов. Так же будет обстоять дело и с оценками других неизвестных параметров. Любая из таких оценок случайна; при пользовании ею неизбежны ошибки. Желательно выбрать такую оценку, чтобы эти ошибки были по возможности минимальными.

Рассмотрим следующую общую задачу. Имеется случайная величина X, закон распределения которой содержит неизвестный параметр a. Требуется найти подходящую оценку для параметра a по результатам n независимых опытов, в каждом из которых величина X приняла определенное значение.

Обозначим наблюденные значения случайной величины

$$X_1,\ X_2,\ \ldots,\ X_n. \tag{14.1.1}$$

Их можно рассматривать как n «экземпляров» случайной величины X, то есть n независимых случайных величин, каждая из которых распределена по тому же закону, что и случайная величина X.

Обозначим $\tilde{a}$ оценку для параметра a. Любая оценка, вычисляемая на основе материала (14.1.1), должна представлять собой функцию величин $X_1,\ X_2,\ \ldots,\ X_n$:

$$\tilde{a} = \tilde{a}(X_1,\ X_2,\ \ldots,\ X_n) \tag{14.1.2}$$

и, следовательно, сама является величиной случайной. Закон распределения $\tilde{a}$ зависит, во-первых, от закона распределения величины X (и, в частности, от самого неизвестного параметра a); во-вторых, от числа опытов n. В принципе этот закон распределения может быть найден известными методами теории вероятностей.

Предъявим к оценке $\tilde{a}$ ряд требований, которым она должна удовлетворять, чтобы быть в каком-то смысле «доброкачественной» оценкой.

Естественно потребовать от оценки $\tilde{a}$, чтобы при увеличении числа опытов n она приближалась (сходилась по вероятности) к параметру a. Оценка, обладающая таким свойством, называется *состоятельной*.

Кроме того, желательно, чтобы, пользуясь величиной $\tilde{a}$ вместо a, мы по крайней мере не делали систематической ошибки в сторону

завышения или занижения, т. е. чтобы выполнялось условие

$$M[\tilde{a}] = a. \tag{14.1.3}$$

Оценка, удовлетворяющая такому условию, называется *несмещенной*.

Наконец, желательно, чтобы выбранная несмещенная оценка обладала по сравнению с другими наименьшей дисперсией, т. е.

$$D[\tilde{a}] = \min. \tag{14.1.4}$$

Оценка, обладающая таким свойством, называется *эффективной*.

На практике не всегда удается удовлетворить всем этим требованиям. Например, может оказаться, что, даже если эффективная оценка существует, формулы для ее вычисления оказываются слишком сложными, и приходится удовлетворяться другой оценкой, дисперсия которой несколько больше. Иногда применяются — в интересах простоты расчетов — незначительно смещенные оценки. Однако выбору оценки всегда должно предшествовать ее критическое рассмотрение со всех перечисленных выше точек зрения.

14.2. Оценки для математического ожидания и дисперсии

Пусть имеется случайная величина X с математическим ожиданием m и дисперсией D; оба параметра неизвестны. Над величиной X произведено n независимых опытов, давших результаты $X_1, X_2, \ldots, X_n$. Требуется найти состоятельные и несмещенные оценки для параметров m и D.

В качестве оценки для математического ожидания естественно предложить среднее арифметическое наблюденных значений (ранее мы его обозначали m^*):

$$\tilde{m} = m^* = \frac{\sum_{i=1}^{n} X_i}{n}. \tag{14.2.1}$$

Нетрудно убедиться, что эта оценка является состоятельной: согласно закону больших чисел, при увеличении n величина $\tilde{m}$ сходится по вероятности к m. Оценка $\tilde{m}$ является также и несмещенной, так как

$$M[\tilde{m}] = \frac{\sum_{i=1}^{n} m}{n} = m. \tag{14.2.2}$$

Дисперсия этой оценки равна:

$$D[\tilde{m}] = \frac{1}{n} D. \tag{14.2.3}$$

Эффективность или неэффективность оценки зависит от вида закона распределения величины X. Можно доказать, что если величина X распределена по нормальному закону, дисперсия (14.2.3) будет минимально возможной, т. е. оценка $\tilde{m}$ является эффективной. Для других законов распределения это может быть и не так.

Перейдем к оценке для дисперсии D. На первый взгляд наиболее естественной оценкой представляется статистическая дисперсия:

$$D^* = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^2}{n}, \tag{14.2.4}$$

где

$$\tilde{m} = \frac{\sum_{i=1}^{n} X_i}{n}. \tag{14.2.5}$$

Проверим, является ли эта оценка состоятельной. Выразим ее через второй начальный момент (по формуле (7.4.6) гл. 7):

$$D^* = \frac{\sum_{i=1}^{n} X_i^2}{n} - \tilde{m}^2. \tag{14.2.6}$$

Первый член в правой части есть среднее арифметическое n наблюденных значений случайной величины X^2; он сходится по вероятности к $M[X^2] = \alpha_2[X]$. Второй член сходится по вероятности к m^2; вся величина (14.2.6) сходится по вероятности к величине

$$\alpha_2[X] - m^2 = D.$$

Это означает, что оценка (14.2.4) состоятельна.

Проверим, является ли оценка D^* также и несмещенной. Подставим в формулу (14.2.6) вместо $\tilde{m}$ его выражение (14.2.5) и произведем указанные действия:

$$\begin{aligned} D^* &= \frac{\sum_{i=1}^{n} X_i^2}{n} - \left(\frac{\sum_{i=1}^{n} X_i}{n}\right)^2 = \\ &= \frac{\sum_{i=1}^{n} X_i^2}{n} - \frac{\sum_{i=1}^{n} X_i^2}{n^2} - 2\frac{\sum_{i<j} X_i X_j}{n^2} = \\ &= \frac{n-1}{n^2}\sum_{i=1}^{n} X_i^2 - 2\frac{\sum_{i<j} X_i X_j}{n^2}. \end{aligned} \tag{14.2.7}$$

Найдем математическое ожидание величины (14.2.7):

$$M[D^*] = \frac{n-1}{n^2}\sum_{i=1}^{n} M[X_i^2] - \frac{2}{n^2}\sum_{i<j} M[X_iX_j]. \qquad (14.2.8)$$

Так как дисперсия D^* не зависит от того, в какой точке выбрать начало координат, выберем его в точке m. Тогда

$$M[X_i^2] = M[\mathring{X}_i^2] = D; \quad \sum_{i=1}^{n} M[X_i^2] = nD, \qquad (14.2.9)$$

$$M[X_iX_j] = M[\mathring{X}_i\mathring{X}_j] = K_{ij} = 0. \qquad (14.2.10)$$

Последнее равенство следует из того, что опыты независимы.

Подставляя (14.2.9) и (14.2.10) в (14.2.8), получим:

$$M[D^*] = \frac{n-1}{n}D. \qquad (14.2.11)$$

Отсюда видно, что величина D^* не является несмещенной оценкой для D: ее математическое ожидание не равно D, а несколько меньше. Пользуясь оценкой D^* вместо дисперсии D, мы будем совершать некоторую систематическую ошибку в меньшую сторону. Чтобы ликвидировать это смещение, достаточно ввести поправку, умножив величину D^* на $\frac{n}{n-1}$. Получим:

$$\frac{n}{n-1}D^* = \frac{\sum_{i=1}^{n}(X_i-\tilde{m})^2}{n}\cdot\frac{n}{n-1} = \frac{\sum_{i=1}^{n}(X_i-\tilde{m})^2}{n-1}.$$

Такую «исправленную» статистическую дисперсию мы и выберем в качестве оценки для D:

$$\tilde{D} = \frac{\sum_{i=1}^{n}(X_i-\tilde{m})^2}{n-1}. \qquad (14.2.12)$$

Так как множитель $\frac{n}{n-1}$ стремится к единице при $n\to\infty$, а оценка D^* состоятельна, то оценка $\tilde{D}$ также будет состоятельной [1]).

На практике часто вместо формулы (14.2.12) бывает удобнее применять другую, равносильную ей, в которой статистическая

[1]) Оценка $\tilde{D}$ для дисперсии не является эффективной. Однако в случае нормального распределения она является «асимптотически эффективной», то есть при увеличении n отношение ее дисперсии к минимально возможной неограниченно приближается к единице.

дисперсия выражена через второй начальный момент:

$$\tilde{D} = \left[\frac{\sum_{i=1}^{n} X_i^2}{n} - \tilde{m}^2 \right] \frac{n}{n-1}. \qquad (14.2.13)$$

При больших значениях n, естественно, обе оценки — смещенная D^* и несмещенная $\tilde{D}$ — будут различаться очень мало и введение поправочного множителя теряет смысл.

Таким образом, мы пришли к следующим правилам обработки ограниченного по объему статистического материала.

Если даны значения $x_1, x_2, \ldots, x_n$, принятые в n независимых опытах случайной величиной X с неизвестными математическим ожиданием m и дисперсией D, то для определения этих параметров следует пользоваться приближенными значениями (оценками):

$$\left.\begin{aligned} \tilde{m} &= \frac{\sum_{i=1}^{n} x_i}{n}; \\ \tilde{D} &= \frac{\sum_{i=1}^{n} (x_i - \tilde{m})^2}{n-1} \end{aligned}\right\} \qquad (14.2.14)$$

или

$$\tilde{D} = \left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \tilde{m}^2 \right) \frac{n}{n-1}.$$

14.3. Доверительный интервал. Доверительная вероятность

В предыдущих п°п° мы рассмотрели вопрос об оценке неизвестного параметра a одним числом. Такая оценка называется «точечной». В ряде задач требуется не только найти для параметра a подходящее численное значение, но и оценить его точность и надежность. Требуется знать — к каким ошибкам может привести замена параметра a его точечной оценкой $\tilde{a}$ и с какой степенью уверенности можно ожидать, что эти ошибки не выйдут за известные пределы?

Такого рода задачи особенно актуальны при малом числе наблюдений, когда точечная оценка $\tilde{a}$ в значительной мере случайна и приближенная замена a на $\tilde{a}$ может привести к серьезным ошибкам.

Чтобы дать представление о точности и надежности оценки $\tilde{a}$, в математической статистике пользуются так называемыми доверительными интервалами и доверительными вероятностями.

Пусть для параметра a получена из опыта несмещенная оценка $\tilde{a}$. Мы хотим оценить возможную при этом ошибку. Назначим некоторую достаточно большую вероятность β (например, $\beta = 0{,}9$, $0{,}95$ или $0{,}99$) такую, что событие с вероятностью β можно считать практически достоверным, и найдем такое значение ε, для которого

$$P(|\tilde{a} - a| < \varepsilon) = \beta. \qquad (14.3.1)$$

Тогда диапазон практически возможных значений ошибки, возникающей при замене a на $\tilde{a}$, будет $\pm\varepsilon$; бóльшие по абсолютной величине ошибки будут появляться только с малой вероятностью $\alpha = 1 - \beta$.

Перепишем (14.3.1) в виде:

$$P(\tilde{a} - \varepsilon < a < \tilde{a} + \varepsilon) = \beta. \qquad (14.3.2)$$

Равенство (14.3.2) означает, что с вероятностью β неизвестное значение параметра a попадает в интервал

$$I_\beta = (\tilde{a} - \varepsilon;\ \ \tilde{a} + \varepsilon). \qquad (14.3.3)$$

При этом необходимо отметить одно обстоятельство. Ранее мы неоднократно рассматривали вероятность попадания случайной величины в заданный неслучайный интервал. Здесь дело обстоит иначе: величина a не случайна, зато случаен интервал I_β. Случайно его положение на оси абсцисс, определяемое его центром $\tilde{a}$; случайна вообще и длина интервала 2ε, так как величина ε вычисляется, как правило, по опытным данным. Поэтому в данном случае лучше будет толковать величину β не как вероятность «попадания» точки a в интервал I_β, а как вероятность того, что случайный интервал I_β накроет точку a (рис. 14.3.1).

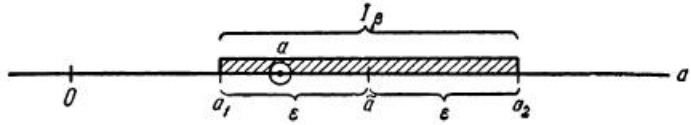


Рис. 14.3.1.

Вероятность β принято называть *доверительной вероятностью*, а интервал I_β — *доверительным интервалом* [1]). Границы интервала I_β: $a_1 = \tilde{a} - \varepsilon$ и $a_2 = \tilde{a} + \varepsilon$ называются *доверительными границами*.

Дадим еще одно истолкование понятию доверительного интервала: его можно рассматривать как интервал значений параметра a, совместимых с опытными данными и не противоречащих им. Действительно, если условиться считать событие с вероятностью

[1]) На рис. 14.3.1 рассматривается доверительный интервал, симметричный относительно $\tilde{a}$. Вообще, как мы увидим дальше, это не обязательно.

$\alpha = 1 - \beta$ практически невозможным, то те значения параметра a, для которых $|\tilde{a} - a| > \varepsilon$, нужно признать противоречащими опытным данным, а те, для которых $|\tilde{a} - a| < \varepsilon$, — совместимыми с ними.

Перейдем к вопросу о нахождении доверительных границ a_1 и a_2.

Пусть для параметра a имеется несмещенная оценка $\tilde{a}$. Если бы нам был известен закон распределения величины $\tilde{a}$, задача нахождения доверительного интервала была бы весьма проста: достаточно было бы найти такое значение ε, для которого

$$P(|\tilde{a} - a| < \varepsilon) = \beta.$$

Затруднение состоит в том, что закон распределения оценки $\tilde{a}$ зависит от закона распределения величины X и, следовательно, от его неизвестных параметров (в частности, и от самого параметра a).

Чтобы обойти это затруднение, можно применить следующий грубо приближенный прием: заменить в выражении для ε неизвестные параметры их точечными оценками. При сравнительно большом числе опытов n (порядка $20 \div 30$) этот прием обычно дает удовлетворительные по точности результаты.

В качестве примера рассмотрим задачу о доверительном интервале для математического ожидания.

Пусть произведено n независимых опытов над случайной величиной X, характеристики которой — математическое ожидание m и дисперсия D — неизвестны. Для этих параметров получены оценки:

$$\tilde{m} = \frac{\sum_{i=1}^{n} X_i}{n}; \qquad \tilde{D} = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^2}{n-1}. \tag{14.3.4}$$

Требуется построить доверительный интервал I_β, соответствующий доверительной вероятности β, для математического ожидания m величины X.

При решении этой задачи воспользуемся тем, что величина $\tilde{m}$ представляет собой сумму n независимых одинаково распределенных случайных величин X_i, и, согласно центральной предельной теореме, при достаточно большом n ее закон распределения близок к нормальному. На практике даже при относительно небольшом числе слагаемых (порядка $10 \div 20$) закон распределения суммы можно приближенно считать нормальным. Будем исходить из того, что величина $\tilde{m}$ распределена по нормальному закону. Характеристики этого закона — математическое ожидание и дисперсия — равны соответственно m и $\frac{D}{n}$ (см. гл. 13 п°13.3). Предположим, что величина D нам известна,

и найдем такую величину ε_β, для которой

$$P(|\tilde{m} - m| < \varepsilon_\beta) = \beta. \qquad (14.3.5)$$

Применяя формулу (6.3.5) главы 6, выразим вероятность в левой части (14.3.5) через нормальную функцию распределения

$$P(|\tilde{m} - m| < \varepsilon_\beta) = 2\Phi^*\left(\frac{\varepsilon_\beta}{\sigma_{\tilde{m}}}\right) - 1, \qquad (14.3.6)$$

где $\sigma_{\tilde{m}} = \sqrt{\frac{D}{n}}$ — среднее квадратическое отклонение оценки $\tilde{m}$.

Из уравнения

$$2\Phi^*\left(\frac{\varepsilon_\beta}{\sigma_{\tilde{m}}}\right) - 1 = \beta$$

находим значение ε_β:

$$\varepsilon_\beta = \sigma_{\tilde{m}} \arg \Phi^*\left(\frac{1+\beta}{2}\right), \qquad (14.3.7)$$

где $\arg \Phi^*(x)$ — функция, обратная $\Phi^*(x)$, т. е. такое значение аргумента, при котором нормальная функция распределения равна x.

Дисперсия D, через которую выражена величина $\sigma_{\tilde{m}}$, нам в точности не известна; в качестве ее ориентировочного значения можно воспользоваться оценкой $\tilde{D}$ (14.3.4) и положить приближенно:

$$\sigma_{\tilde{m}} = \sqrt{\frac{\tilde{D}}{n}}. \qquad (14.3.8)$$

Таким образом, приближенно решена задача построения доверительного интервала, который равен:

$$I_\beta = (\tilde{m} - \varepsilon_\beta;\ \tilde{m} + \varepsilon_\beta), \qquad (14.3.9)$$

где ε_β определяется формулой (14.3.7).

Чтобы избежать при вычислении ε_β обратного интерполирования в таблицах функции $\Phi^*(x)$, удобно составить специальную таблицу (см. табл. 14.3.1), где приводятся значения величины

$$t_\beta = \arg \Phi^*\left(\frac{1+\beta}{2}\right)$$

в зависимости от β. Величина t_β определяет для нормального закона число средних квадратических отклонений, которое нужно отложить вправо и влево от центра рассеивания для того, чтобы вероятность попадания в полученный участок была равна β.

Через величину t_β доверительный интервал выражается в виде:

$$I_\beta = (\tilde{m} - t_\beta \sigma_{\tilde{m}};\ \tilde{m} + t_\beta \sigma_{\tilde{m}}).$$

Таблица 14.3.1

β	t_β	β	t_β	β	t_β	β	t_β
0,80	1,282	0,86	1,475	0,91	1,694	0,97	2,169
0,81	1,310	0,87	1,513	0,92	1,750	0,98	2,325
0,82	1,340	0,88	1,554	0,93	1,810	0,99	2,576
0,83	1,371	0,89	1,597	0,94	1,880	0,9973	3,000
0,84	1,404	0,90	1,643	0,95	1,960	0,999	3,290
0,85	1,439			0,96	2,053		

Пример 1. Произведено 20 опытов над величиной X; результаты приведены в таблице 14.3.2.

Таблица 14.3.2

i	x_i	i	x_i	i	x_i	i	x_i
1	10,5	6	10,6	11	10,6	16	10,9
2	10,8	7	10,9	12	11,3	17	10,8
3	11,2	8	11,0	13	10,5	18	10,7
4	10,9	9	10,3	14	10,7	19	10,9
5	10,4	10	10,8	15	10,8	20	11,0

Требуется найти оценку $\tilde{m}$ для математического ожидания m величины X и построить доверительный интервал, соответствующий доверительной вероятности $\beta = 0{,}8$.

Решение. Имеем:

$$\tilde{m} = \frac{1}{20}\sum_{i=1}^{20} x_i = 10{,}78.$$

Выбрав за начало отсчета $x = 10$, по третьей формуле (14.2.14) находим несмещенную оценку $\tilde{D}$:

$$\tilde{D} = \left(\frac{13{,}38}{20} - 0{,}78^2\right)\frac{20}{19} = 0{,}064;$$

$$\sigma_{\tilde{m}} = \sqrt{\frac{\tilde{D}}{n}} = 0{,}0564.$$

По таблице 14.3.1 находим $t_\beta = 1{,}282$;

$$\varepsilon_\beta = t_\beta \sigma_{\tilde{m}} = 0{,}072.$$

Доверительные границы:

$$m_1 = \tilde{m} - 0{,}072 = 10{,}71;$$

$$m_2 = \tilde{m} + 0{,}072 = 10{,}85.$$

Доверительный интервал:

$$I_\beta = (10{,}71;\ 10{,}85).$$

Значения параметра m, лежащие в этом интервале, являются совместимыми с опытными данными, приведенными в таблице 14.3.2.

Аналогичным способом может быть построен доверительный интервал и для дисперсии.

Пусть произведено n независимых опытов над случайной величиной X с неизвестными параметрами m и D, и для дисперсии D получена несмещенная оценка:

$$\tilde{D} = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^2}{n-1}, \tag{14.3.11}$$

где

$$\tilde{m} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Требуется приближенно построить доверительный интервал для дисперсии.

Из формулы (14.3.11) видно, что величина $\tilde{D}$ представляет собой сумму n случайных величин вида $\frac{(X_i - \tilde{m})^2}{n-1}$. Эти величины не являются независимыми, так как в любую из них входит величина $\tilde{m}$, зависящая от всех остальных. Однако можно показать, что при увеличении n закон распределения их суммы тоже приближается к нормальному. Практически при $n = 20 \div 30$ он уже может считаться нормальным.

Предположим, что это так, и найдем характеристики этого закона: математическое ожидание и дисперсию. Так как оценка $\tilde{D}$ — несмещенная, то

$$M[\tilde{D}] = D.$$

Вычисление дисперсии $D[\tilde{D}]$ связано со сравнительно сложными выкладками, поэтому приведем ее выражение без вывода:

$$D[\tilde{D}] = \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)} D^2, \tag{14.3.12}$$

где μ_4 — четвертый центральный момент величины X.

Чтобы воспользоваться этим выражением, нужно подставить в него значения μ_4 и D (хотя бы приближенные). Вместо D можно воспользоваться его оценкой $\tilde{D}$. В принципе четвертый центральный момент μ_4 тоже можно заменить его оценкой, например величиной вида:

$$\mu_4^* = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^4}{n}, \tag{14.3.13}$$

но такая замена даст крайне невысокую точность, так как вообще при ограниченном числе опытов моменты высокого порядка определяются с большими ошибками. Однако на практике часто бывает, что вид закона распределения величины X известен заранее: неизвестны лишь его параметры. Тогда можно попытаться выразить μ_4 через D.

Возьмем наиболее часто встречающийся случай, когда величина X распределена по нормальному закону. Тогда ее четвертый центральный момент выражается через дисперсию (см. гл. 6 п° 6.2):

$$\mu_4 = 3D^2,$$

и формула (14.3.12) дает

$$D[\tilde{D}] = \frac{3}{n} D^2 - \frac{n-3}{n(n-1)} D^2$$

или

$$D[\tilde{D}] = \frac{2}{n-1} D^2. \qquad (14.3.14)$$

Заменяя в (14.3.14) неизвестное D его оценкой $\tilde{D}$, получим:

$$D[\tilde{D}] = \frac{2}{n-1} \tilde{D}^2, \qquad (14.3.15)$$

откуда

$$\sigma_{\tilde{D}} = \sqrt{\frac{2}{n-1}}\, \tilde{D}. \qquad (14.3.16)$$

Момент μ_4 можно выразить через D также и в некоторых других случаях, когда распределение величины X не является нормальным, но вид его известен. Например, для закона равномерной плотности (см. главу 5) имеем:

$$\mu_4 = \frac{(\beta-\alpha)^4}{80}; \quad D = \frac{(\beta-\alpha)^2}{12},$$

где (α, β) — интервал, на котором задан закон. Следовательно,

$$\mu_4 = 1{,}8D^2.$$

По формуле (14.3.12) получим:

$$D[\tilde{D}] = \frac{0{,}8n+1{,}2}{n(n-1)} D^2,$$

откуда находим приближенно

$$\sigma_{\tilde{D}} \approx \sqrt{\frac{0{,}8n+1{,}2}{n(n-1)}}\, \tilde{D}. \qquad (14.3.17)$$

В случаях, когда вид закона распределения величины X неизвестен, при ориентировочной оценке величины $\sigma_{\tilde{D}}$ рекомендуется все же пользоваться формулой (14.3.16), если нет специальных оснований считать, что этот закон сильно отличается от нормального (обладает заметным положительным или отрицательным эксцессом).

Если ориентировочное значение $\sigma_{\tilde{D}}$ тем или иным способом получено, то можно построить доверительный интервал для дисперсии, аналогично тому, как мы строили его для математического ожидания:

$$I_\beta = (\tilde{D} - t_\beta \sigma_{\tilde{D}};\ \tilde{D} + t_\beta \sigma_{\tilde{D}}), \qquad (14.3.18)$$

где величина t_β в зависимости от заданной вероятности β находится по таблице 14.3.1.

Пример 2. Найти приближенно 80%-й доверительный интервал для дисперсии случайной величины X в условиях примера 1, если известно, что величина X распределена по закону, близкому к нормальному.

Решение. Величина t_β остается той же, что в примере 1:

$$t_\beta = 1{,}282.$$

По формуле (14.3.16)

$$\sigma_{\tilde{D}} = \sqrt{\frac{2}{19}} \cdot 0{,}064 = 0{,}0207.$$

По формуле (14.3.18) находим доверительный интервал:

$$I_\beta = (0{,}043;\ 0{,}085).$$

Соответствующий интервал значений среднего квадратического отклонения: (0,21; 0,29).

14.4. Точные методы построения доверительных интервалов для параметров случайной величины, распределенной по нормальному закону

В предыдущем п° мы рассмотрели грубо приближенные методы построения доверительных интервалов для математического ожидания и дисперсии. В данном п° мы дадим представление о точных методах решения той же задачи. Подчеркнем, что для точного нахождения доверительных интервалов совершенно необходимо знать заранее вид закона распределения величины X, тогда как для применения приближенных методов это не обязательно.

Идея точных методов построения доверительных интервалов сводится к следующему. Любой доверительный интервал находится из условия, выражающего вероятность выполнения некоторых неравенств, в которые входит интересующая нас оценка $\tilde{a}$. Закон распределения оценки $\tilde{a}$ в общем случае зависит от самих неизвестных параметров величины X. Однако иногда удается перейти в неравенствах от слу-

чайной величины $\tilde{a}$ к какой-либо другой функции наблюденных значений $X_1, X_2, \ldots, X_n$, закон распределения которой не зависит от неизвестных параметров, а зависит только от числа опытов n и от вида закона распределения величины X. Такого рода случайные величины играют большую роль в математической статистике; они наиболее подробно изучены для случая нормального распределения величины X.

Например, доказано, что при нормальном распределении величины X случайная величина

$$T = \sqrt{n}\,\frac{\tilde{m} - m}{\sqrt{\tilde{D}}}, \tag{14.4.1}$$

где

$$\tilde{m} = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad \tilde{D} = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^2}{n-1},$$

подчиняется так называемому *закону распределения Стьюдента* с $n-1$ степенями свободы; плотность этого закона имеет вид

$$S_{n-1}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{(n-1)\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}, \tag{14.4.2}$$

где $\Gamma(x)$ — известная гамма-функция:

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u}\,du.$$

Доказано также, что случайная величина

$$V = \frac{(n-1)\,\tilde{D}}{D} \tag{14.4.3}$$

имеет «распределение χ^2» с $n-1$ степенями свободы (см. гл. 7, стр. 145), плотность которого выражается формулой

$$k_{n-1}(v) = \begin{cases} \dfrac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}\, v^{\frac{n-1}{2}-1} e^{-\frac{v}{2}} & \text{при } v > 0, \\ 0 & \text{при } v < 0. \end{cases} \tag{14.4.4}$$

Не останавливаясь на выводах распределений (14.4.2) и (14.4.4), покажем, как их можно применить при построении доверительных интервалов для параметров m и $\tilde{D}$.

Пусть произведено n независимых опытов над случайной величиной X, распределенной по нормальному закону с неизвестными параметрами m и D. Для этих параметров получены оценки

$$\tilde{m} = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad \tilde{D} = \frac{\sum_{i=1}^{n} (X_i - \tilde{m})^2}{n-1}.$$

Требуется построить доверительные интервалы для обоих параметров, соответствующие доверительной вероятности β.

Построим сначала доверительный интервал для математического ожидания. Естественно этот интервал взять симметричным относительно $\tilde{m}$; обозначим ε_β половину длины интервала. Величину ε_β нужно выбрать так, чтобы выполнялось условие

$$P(|\tilde{m} - m| < \varepsilon_\beta) = \beta. \tag{14.4.5}$$

Попытаемся перейти в левой части равенства (14.4.5) от случайной величины $\tilde{m}$ к случайной величине T, распределенной по закону Стьюдента. Для этого умножим обе части неравенства $|\tilde{m} - m| < \varepsilon_\beta$ на положительную величину $\frac{\sqrt{n}}{\sqrt{\tilde{D}}}$:

$$P\left(\frac{\sqrt{n}\,|\tilde{m} - m|}{\sqrt{\tilde{D}}} < \frac{\varepsilon_\beta}{\sqrt{\frac{\tilde{D}}{n}}}\right) = \beta$$

или, пользуясь обозначением (14.4.1),

$$P\left(|T| < \frac{\varepsilon_\beta}{\sqrt{\frac{\tilde{D}}{n}}}\right) = \beta. \tag{14.4.6}$$

Найдем такое число t_β, что

$$P(|T| < t_\beta) = \beta. \tag{14.4.7}$$

Величина t_β найдется из условия

$$P(|T| < t_\beta) = \int_{-t_\beta}^{t_\beta} S_{n-1}(t)\,dt = \beta. \tag{14.4.8}$$

Из формулы (14.4.2) видно, что $S_{n-1}(t)$ — четная функция; поэтому (14.4.8) дает

$$2\int_0^{t_\beta} S_{n-1}(t)\,dt = \beta. \tag{14.4.9}$$

Равенство (14.4.9) определяет величину t_β в зависимости от β. Если иметь в своем распоряжении таблицу значений интеграла

$$\Psi(x) = 2\int_0^x S_{n-1}(t)\,dt,$$

то величину t_β можно найти обратным интерполированием в этой таблице. Однако удобнее составить заранее таблицу значений t_β. Такая таблица дается в приложении (см. табл. 5). В этой таблице приведены значения t_β в зависимости от доверительной вероятности β и числа степеней свободы $n-1$. Определив t_β по таблице 5 и полагая

$$\varepsilon_\beta = t_\beta \sqrt{\frac{\tilde{D}}{n}}, \qquad (14.4.10)$$

мы найдем половину ширины доверительного интервала I_β и сам интервал

$$I_\beta = \left(\tilde{m} - t_\beta \sqrt{\frac{\tilde{D}}{n}};\ \ \tilde{m} + t_\beta \sqrt{\frac{\tilde{D}}{n}}\right). \qquad (14.4.11)$$

Пример 1. Произведено 5 независимых опытов над случайной величиной X, распределенной нормально с неизвестными параметрами m и σ. Результаты опытов приведены в таблице 14.4.1.

Таблица 14.4.1

l	1	2	3	4	5
x_l	—2,5	3,4	—2,0	1,0	2,1

Найти оценку $\tilde{m}$ для математического ожидания и построить для него 90%-й доверительный интервал I_β (т. е. интервал, соответствующий доверительной вероятности $\beta = 0{,}9$).

Решение. Имеем

$$\tilde{m} = 0{,}4; \qquad \tilde{D} = 6{,}6.$$

По таблице 5 приложения для $n-1=4$ и $\beta = 0{,}9$ находим

$$t_\beta = 2{,}13,$$

откуда

$$\varepsilon_\beta = t_\beta \sqrt{\frac{\tilde{D}}{n}} \approx 2{,}45.$$

Доверительный интервал будет

$$I_\beta = (\tilde{m} - \varepsilon_\beta;\ \ \tilde{m} + \varepsilon_\beta) = (-2{,}05;\ 2{,}85).$$

Пример 2. Для условий примера 1 п° 14.3, предполагая величину X распределенной нормально, найти точный доверительный интервал.

Решение. По таблице 5 приложения находим при $n-1=19$ и $\beta=0{,}8$ $t_\beta=1{,}328$; отсюда $\varepsilon_\beta=t_\beta\sqrt{\frac{\tilde{D}}{n}}\approx 0{,}0704$.

Сравнивая с решением примера 1 п° 14.3 ($\varepsilon_\beta=0{,}072$), убеждаемся, что расхождение весьма незначительно. Если сохранить точность до второго знака после запятой, то доверительные интервалы, найденные точным и приближенным методами, совпадают:

$$I_\beta=(10{,}71;\ \ 10{,}85).$$

Перейдем к построению доверительного интервала для дисперсии. Рассмотрим несмещенную оценку дисперсии

$$\tilde{D}=\frac{\sum_{i=1}^{n}(X_i-\tilde{m})^2}{n-1}$$

и выразим случайную величину $\tilde{D}$ через величину V (14.4.3), имеющую распределение χ^2 (14.4.4):

$$\tilde{D}=V\frac{D}{n-1}. \qquad (14.4.12)$$

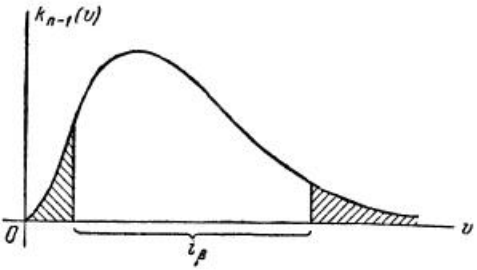


Рис. 14.4.1.

Зная закон распределения величины V, можно найти интервал i_β, в который она попадает с заданной вероятностью β.

Закон распределения $k_{n-1}(v)$ величины V имеет вид, изображенный на рис. 14.4.1.

Возникает вопрос: как выбрать интервал i_β? Если бы закон распределения величины V был симметричным (как нормальный закон или распределение Стьюдента), естественно было бы взять интервал i_β симметричным относительно математического ожидания. В данном случае закон $k_{n-1}(v)$ несимметричен. Условимся выбирать интервал i_β так, чтобы вероятности выхода величины V за пределы интервала вправо и влево (заштрихованные площади на рис. 14.4.1) были одинаковы и равны

$$\frac{\alpha}{2}=\frac{1-\beta}{2}.$$

Чтобы построить интервал l_β с таким свойством, воспользуемся таблицей 4 приложения: в ней приведены числа χ^2 такие, что

$$P(V > \chi^2) = p$$

для величины V, имеющей χ^2-распределение с r степенями свободы. В нашем случае $r = n - 1$. Зафиксируем $r = n - 1$ и найдем в соответствующей строке табл. 4 два значения χ^2: одно, отвечающее вероятности $p_1 = \frac{\alpha}{2}$; другое — вероятности $p_2 = 1 - \frac{\alpha}{2}$. Обозначим эти значения χ_1^2 и χ_2^2. Интервал l_β имеет χ_2^2 своим левым, а χ_1^2 — правым концом.

Теперь найдем по интервалу l_β искомый доверительный интервал I_β для дисперсии с границами D_1 и D_2, который накрывает точку D с вероятностью β:

$$P(D_1 < D < D_2) = \beta.$$

Построим такой интервал $I_\beta = (D_1, D_2)$, который накрывает точку D тогда и только тогда, когда величина V попадает в интервал l_β. Покажем, что интервал

$$I_\beta = \left(\frac{\tilde{D}(n-1)}{\chi_1^2}; \quad \frac{\tilde{D}(n-1)}{\chi_2^2} \right) \tag{14.4.13}$$

удовлетворяет этому условию. Действительно, неравенства

$$\frac{\tilde{D}(n-1)}{\chi_1^2} < D; \quad \frac{\tilde{D}(n-1)}{\chi_2^2} > D$$

равносильны неравенствам

$$V < \chi_1^2; \quad V > \chi_2^2,$$

а эти неравенства выполняются с вероятностью β. Таким образом, доверительный интервал для дисперсии найден и выражается формулой (14.4.13).

Пример 3. Найти доверительный интервал для дисперсии в условиях примера 2 п°14.3, если известно, что величина X распределена нормально.

Решение. Имеем $\beta = 0{,}8$; $\alpha = 0{,}2$; $\frac{\alpha}{2} = 0{,}1$. По таблице 4 приложения находим при $r = n - 1 = 19$

для $p_1 = \frac{\alpha}{2} = 0{,}1$ $\quad \chi_1^2 = 27{,}2$;

для $p_2 = 1 - \frac{\alpha}{2} = 0{,}9$ $\quad \chi_2^2 = 11{,}65$.

По формуле (14.4.13) находим доверительный интервал для дисперсии

$$I_\beta = (0{,}045;\ 0{,}104).$$

Соответствующий интервал для среднего квадратического отклонения: (0,21; 0,32). Этот интервал лишь незначительно превосходит полученный в примере 2 п° 14.3 приближенным методом интервал (0,21; 0,29).

14.5. Оценка вероятности по частоте

На практике часто приходится оценивать неизвестную вероятность p события A по его частоте p^* в n независимых опытах.

Эта задача близко примыкает к рассмотренным в предыдущих п°п°. Действительно, частота события A в n независимых опытах есть не что иное, как среднее арифметическое наблюденных значений величины X, которая в каждом отдельном опыте принимает значение 1, если событие A появилось, и 0, если не появилось:

$$p^* = \frac{\sum_{i=1}^{n} X_i}{n}. \tag{14.5.1}$$

Напомним, что математическое ожидание величины X равно p; ее дисперсия pq, где $q = 1 - p$. Математическое ожидание среднего арифметического также равно p

$$M[p^*] = p, \tag{14.5.2}$$

т. е. оценка p^* для p является несмещенной.

Дисперсия величины p^* равна

$$D[p^*] = \frac{pq}{n}. \tag{14.5.3}$$

Можно доказать, что эта дисперсия является минимально возможной, т. е. оценка p^* для p является эффективной.

Таким образом, в качестве точечной оценки для неизвестной вероятности p разумно во всех случаях принимать частоту p^*. Возникает вопрос о точности и надежности такой оценки, т. е. о построении доверительного интервала для вероятности p.

Хотя эта задача и представляет собой частный случай ранее рассмотренной задачи о доверительном интервале для математического ожидания, все же целесообразно решать ее отдельно. Специфика здесь в том, что величина X — прерывная случайная величина только с двумя возможными значениями: 0 и 1. Кроме того, ее математическое ожидание p и дисперсия $pq = p(1 - p)$ связаны функциональной зависимостью. Это упрощает задачу построения доверительного интервала.

Рассмотрим сначала наиболее простой случай, когда число опытов n сравнительно велико, а вероятность p не слишком велика и не слишком мала. Тогда можно считать, что частота события p^* есть случайная величина, распределение которой близко к нормальному [1]). Расчеты показывают, что этим допущением можно пользоваться даже при не очень больших значениях n: достаточно, чтобы обе величины np и nq были больше четырех. Будем исходить из того, что эти условия выполнены и частоту p^* можно считать распределенной по нормальному закону. Параметрами этого закона будут:

$$m_{p^*} = p; \quad \sigma_{p^*} = \sqrt{\frac{pq}{n}}. \tag{14.5.4}$$

Предположим сначала, что величина p нам известна. Назначим доверительную вероятность β и найдем такой интервал $(p - \varepsilon_\beta,\ p + \varepsilon_\beta)$, чтобы величина p^* попадала в этот интервал с вероятностью β:

$$P(|p^* - p| < \varepsilon_\beta) = \beta. \tag{14.5.5}$$

Так как величина p^* распределена нормально, то

$$P(|p^* - p| < \varepsilon_\beta) = 2\Phi^*\left(\frac{\varepsilon_\beta}{\sigma_{p^*}}\right) - 1 = \beta,$$

откуда, как и в п° 14.3,

$$\varepsilon_\beta = \sigma_{p^*} \arg \Phi^*\left(\frac{1+\beta}{2}\right),$$

где $\arg \Phi^*$ — функция, обратная нормальной функции распределения Φ^*.

Для определения β, как и в п° 14,3, можно обозначить

$$t_\beta = \arg \Phi^*\left(\frac{1+\beta}{2}\right).$$

Тогда

$$\varepsilon_\beta = t_\beta \sigma_{p^*}, \tag{14.5.6}$$

где t_β определяется из таблицы 14.3.1.

Таким образом, с вероятностью β можно утверждать, что

$$|p^* - p| < t_\beta \sqrt{\frac{pq}{n}}. \tag{14.5.7}$$

Фактически величина p нам неизвестна; однако неравенство (14.5.7) будет иметь вероятность β независимо от того, известна нам или неизвестна вероятность p. Получив из опыта конкретное значение частоты p^*, можно, пользуясь неравенством (14.5.7), найти интервал I_β, который с вероятностью β накрывает точку p. Действительно,

[1]) Частота события при n опытах представляет собой прерывную случайную величину; говоря о близости ее закона распределения к нормальному, мы имеем в виду функцию распределения, а не плотность.

преобразуем это неравенство к виду

$$(p^* - p)^2 < \frac{t_\beta^2}{n} p(1-p) \qquad (14.5.8)$$

и дадим ему геометрическую интерпретацию. Будем откладывать по оси абсцисс частоту p^*, а по оси ординат — вероятность p (рис. 14.5.1). Геометрическим местом точек, координаты которых p^* и p удовлетворяют неравенству (14.5.8), будет внутренняя часть эллипса, проходящего через точки (0, 0) и (1, 1) и имеющего в этих точках касательные, параллельные оси Op^*. Так как величина p^* не может быть ни отрицательной, ни большей единицы, то область D, соответствующую неравенству (14.5.8), нужно еще ограничить слева и справа прямыми $p^* = 0$ и $p^* = 1$. Теперь можно для любого значения p^*, полученного из опыта, построить доверительный интервал I_β, который с вероятностью β накроет неизвестное значение p. Для этого проведем через точку p^* прямую, параллельную оси ординат; на этой прямой границы области D отсекут доверительный интервал $I_\beta = (p_1, p_2)$. Действительно, точка M со случайной абсциссой p^* и неслучайной (но неизвестной) ординатой p с вероятностью β попадет внутрь эллипса, т. е. интервал I_β с вероятностью β накроет точку p.

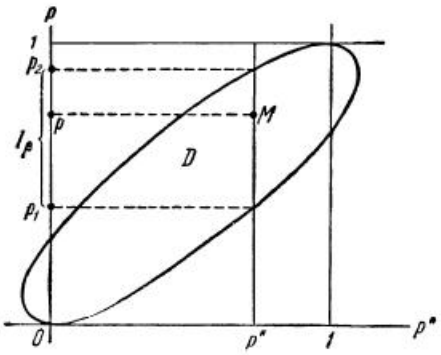


Рис. 14.5.1.

Размеры и конфигурация «доверительного эллипса» зависят от числа опытов n. Чем больше n, тем больше вытянут эллипс и тем уже доверительный интервал.

Доверительные границы p_1 и p_2 можно найти из соотношения (14.5.8), заменив в нем знак неравенства равенством. Решая полученное квадратное уравнение относительно p, получим два корня:

$$\left.\begin{aligned} p_1 &= \frac{p^* + \frac{1}{2}\frac{t_\beta^2}{n} - t_\beta \sqrt{\frac{p^*(1-p^*)}{n} + \frac{1}{4}\frac{t_\beta^2}{n^2}}}{1 + \frac{t_\beta^2}{n}}; \\ p_2 &= \frac{p^* + \frac{1}{2}\frac{t_\beta^2}{n} + t_\beta \sqrt{\frac{p^*(1-p^*)}{n} + \frac{1}{4}\frac{t_\beta^2}{n^2}}}{1 + \frac{t_\beta^2}{n}}. \end{aligned}\right\} \qquad (14.5.9)$$

Доверительный интервал для вероятности p будет

$$I_\beta = (p_1,\ p_2).$$

Пример 1. Частота события A в серии из 100 опытов оказалась $p^* = 0{,}78$. Определить 90%-й доверительный интервал для вероятности p события A.

Решение. Прежде всего проверяем применимость нормального закона; для этого оценим величины np и nq. Полагая ориентировочно $p \approx p^*$, получим

$$np \approx np^* = 78; \quad nq \approx n(1 - p^*) = 22.$$

Обе величины значительно больше четырех; нормальный закон применим. Из таблицы 14.3.1 для $\beta = 0{,}9$ находим $t_\beta = 1{,}643$. По формулам (14.5.9) имеем

$$p_1 = 0{,}705; \quad p_2 = 0{,}840; \quad I_\beta = (0{,}705;\ 0{,}840).$$

Заметим, что при увеличении n величины $\frac{t_\beta^2}{n}$ и $\frac{1}{4}\frac{t_\beta^2}{n}$ в формулах (14.5.9) стремятся к нулю; в пределе формулы принимают вид

$$\left.\begin{aligned} p_1 &= p^* - t_\beta \sqrt{\frac{p^*(1-p^*)}{n}}, \\ p_2 &= p^* + t_\beta \sqrt{\frac{p^*(1-p^*)}{n}}. \end{aligned}\right\} \qquad (14.5.10)$$

Эти формулы могут быть получены и непосредственно, если воспользоваться приближенным способом построения доверительного интервала для математического ожидания, данным в п° 14.3. Формулами (14.5.10) можно пользоваться при больших n (порядка сотен) если только вероятность p не слишком велика и не слишком мала (например, когда обе величины np и nq порядка 10 или более).

Пример 2. Произведено 200 опытов; частота события A оказалась $p^* = 0{,}34$. Построить 85%-й доверительный интервал для вероятности события приближенно (по формулам (14.5.10)). Сравнить результат с точным, соответствующим формулам (14.5.9).

Решение. $\beta = 0{,}85$; по таблице 14.3.1 находим $t_\beta = 1{,}439$. Умножая его на

$$\sqrt{\frac{p^*(1-p^*)}{n}} \approx 0{,}0335,$$

получим

$$t_\beta \sqrt{\frac{p^*(1-p^*)}{n}} \approx 0{,}048,$$

откуда находим приближенно доверительный интервал

$$I_\beta \approx (0{,}292;\ 0{,}388).$$

По формулам (14.5.9) найдем более точные значения $p_1 = 0{,}294$; $p_2 = 0{,}389$, которые почти не отличаются от приближенных.

Выше мы рассмотрели вопрос о построении доверительного интервала для случая достаточно большого числа опытов, когда частоту можно считать распределенной нормально. При малом числе опытов (а также если вероятность p очень велика или очень мала) таким допущением пользоваться нельзя. В этом случае доверительный интервал строят, исходя не из приближенного, а из точного закона распределения частоты. Нетрудно убедиться, что это есть биномиальное распределение, рассмотренное в главах 3 и 4. Действительно, число появлений события A в n опытах распределено по биномиальному закону: вероятность того, что событие A появится ровно m раз, равна

$$P_{m,n} = C_n^m p^m q^{n-m}, \qquad (14.5.11)$$

а частота p^* есть не что иное, как число появлений события, деленное на число опытов.

Исходя из этого распределения, можно построить доверительный интервал I_β аналогично тому, как мы строили его, исходя из нормального закона для больших n (стр. 331).

Предположим сначала, что вероятность p нам известна, и найдем интервал частот p_1^*, p_2^*, в который с вероятностью $\beta = 1 - \alpha$ попадет частота события p^*.

Для случая большого n мы пользовались нормальным законом распределения и брали интервал симметричным относительно математического ожидания. Биномиальное распределение (14.5.11) не обладает симметрией. К тому же (в связи с тем, что частота — прерывная случайная величина) интервала, вероятность попадания в который в точности равна β, может и не существовать. Поэтому выберем в качестве интервала p_1^*, p_2^* самый малый интервал, вероятность попадания левее которого и правее которого будет больше $\frac{\alpha}{2}$.

Аналогично тому, как мы строили область D для нормального закона (рис. 14.5.1), можно будет для каждого n и заданного β построить область, внутри которой значение вероятности p совместимо с наблюденным в опыте значением частоты p^*.

На рис. 14.5.2 изображены кривые, ограничивающие такие области для различных n при доверительной вероятности $\beta = 0{,}9$. По оси абсцисс откладывается частота p^*, по оси ординат — вероятность p. Каждая пара кривых, соответствующая данному n, определяет доверительный интервал вероятностей, отвечающий данному значению частоты. Строго говоря, границы областей должны быть ступенчатыми (ввиду прерывности частоты), но для удобства они изображены в виде плавных кривых.

Для того чтобы, пользуясь такими кривыми, найти доверительный интервал I_β, нужно произвести следующее построение (см. рис. 14.5.2): по оси абсцисс отложить наблюденное в опыте значение частоты p^*,

провести через эту точку прямую, параллельную оси ординат, и отметить точки пересечения прямой с парой кривых, соответствующих данному числу опытов n; проекции этих точек на ось ординат и дадут границы p_1, p_2 доверительного интервала I_β.

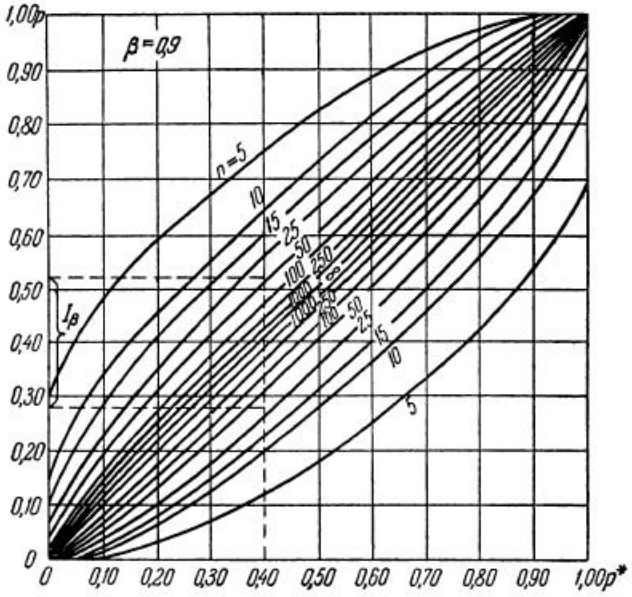


Рис. 14.5.2.

При заданном n кривые, ограничивающие «доверительную область», определяются уравнениями:

$$\sum_{m=k}^{n} C_n^m p^m (1-p)^{n-m} = \frac{\alpha}{2}; \qquad (14.5.12)$$

$$\sum_{m=0}^{k} C_n^m p^m (1-p)^{n-m} = \frac{\alpha}{2}, \qquad (14.5.13)$$

где k — число появлений события:

$$k = np^*.$$

Разрешая уравнение (14.5.12) относительно p, можно найти нижнюю границу p_1 «доверительной области»; аналогично из (14.5.13) можно найти p_2.

Чтобы не решать эти уравнения каждый раз заново, удобно заранее затабулировать (или представить графически) решения для нескольких

типичных значений доверительной вероятности β. Например, в книге И. В. Дунина-Барковского и Н. В. Смирнова «Теория вероятностей и математическая статистика в технике» имеются таблицы p_1 и p_2 для $\beta = 0{,}95$ и $\beta = 0{,}99$. Из той же книги заимствован график рис. 14.5.2.

Пример 3. Найти доверительные границы p_1 и p_2 для вероятности события, если в 50 опытах частота его оказалась $p^* = 0{,}4$. Доверительная вероятность $\beta = 0{,}9$.

Решение. Построением (см. пунктир на рис. 14.5.2) для $p^* = 0{,}4$ и $n = 50$ находим: $p_1 \approx 0{,}28$; $p_2 \approx 0{,}52$.

Пользуясь методом доверительных интервалов, можно приближенно решить и другой важный для практики вопрос: каково должно быть число опытов n для того, чтобы с доверительной вероятностью β ожидать, что ошибка от замены вероятности частотой не превзойдет заданного значения?

При решении подобных задач удобнее не пользоваться непосредственно графиками типа рис. 14.5.2, а перестроить их, представив доверительные границы как функции от числа опытов n.

Пример 4. Проведено 25 опытов, в которых событие A произошло 12 раз. Найти ориентировочно число опытов n, которое понадобится для

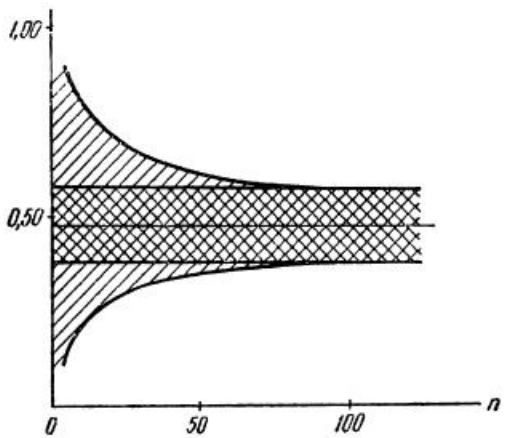


Рис. 14.5.3.

того, чтобы с вероятностью $\beta = 0{,}9$ ошибка от замены вероятности частотой не превзошла 20%.

Решение. Определяем предельно допустимую ошибку:

$$\Delta = 0{,}2 \cdot 0{,}48 = 0{,}096 \approx 0{,}1.$$

Пользуясь кривыми рис. 14.5.2, построим новый график: по оси абсцисс отложим число опытов n, по оси ординат — доверительные границы для вероятности (рис. 14.5.3). Средняя прямая, параллельная оси абсцисс, соответствует

наблюденной частоте события $p^* = \frac{12}{25} = 0{,}48$. Выше и ниже прямой $p = p^* = 0{,}48$ проведены кривые $p_1(n)$ и $p_2(n)$, изображающие нижнюю и верхнюю доверительные границы в зависимости от n. Область между кривыми, определяющая доверительный интервал, заштрихована. В непосредственной близости от прямой $p = 0{,}48$ двойной штриховкой показана более узкая область 20%-й допустимой ошибки. Из рис. 14.5.3 видно, что ошибка падает до допустимой величины при числе опытов n порядка 100.

Заметим, что после выполнения потребного числа опытов может понадобиться новая проверка точности определения вероятности по частоте, так как будет получено в общем случае уже другое значение частоты p^*, отличное от наблюденного в ранее проведенных опытах. При этом может оказаться, что число опытов все еще недостаточно для обеспечения необходимой точности, и его придется несколько увеличить. Однако первое приближение, полученное описанным выше методом, может служить для ориентировочного предварительного планирования серии опытов с точки зрения требуемого на них времени, денежных затрат и т. д.

На практике иногда приходится встречаться со своеобразной задачей определения доверительного интервала для вероятности события, когда полученная из опыта частота равна нулю. Такая задача обычно связана с опытами, в которых вероятность интересующего нас события очень мала (или, наоборот, очень велика — тогда мала вероятность противоположного события).

Пусть, например, проводятся испытания какого-то изделия на безотказность работы. В результате испытаний изделие не отказало ни разу. Требуется найти максимальную практически возможную вероятность отказа.

Поставим эту задачу в общем виде. Произведено n независимых опытов, ни в одном из которых событие A не произошло. Задана доверительная вероятность β; требуется построить доверительный интервал для вероятности p события A, точнее — найти его верхнюю границу p_2, так как нижняя p_1, естественно, равна нулю.

Поставленная задача является частным случаем общей задачи о доверительном интервале для вероятности, но ввиду своих особенностей заслуживает отдельного рассмотрения. Прежде всего, приближенный метод построения доверительного интервала (на основе замены закона распределения частоты нормальным), изложенный в начале данного п°, здесь неприменим, так как вероятность p очень мала. Точный метод построения доверительного интервала на основе биномиального распределения в данном случае применим, но может быть существенно упрощен.

Будем рассуждать следующим образом. В результате n опытов наблюдено событие B, состоящее в том, что A не появилось ни разу. Требуется найти максимальное значение $p = p_2$, которое «совместимо» с наблюденным в опыте событием B, если считать

«несовместимыми» с B те значения p, для которых вероятность события B меньше, чем $\alpha = 1 - \beta$.

Очевидно, для любой вероятности p события A вероятность наблюденного события B равна

$$P(B) = (1-p)^n.$$

Полагая $P(B) = \alpha$, получим уравнение для p_2:

$$(1-p_2)^n = 1-\beta, \tag{14.5.14}$$

откуда

$$p_2 = 1 - \sqrt[n]{1-\beta}. \tag{14.5.15}$$

Пример 5. Вероятность p самопроизвольного срабатывания взрывателя при падении снаряда с высоты h неизвестна, но предположительно весьма мала. Произведено 100 опытов, в каждом из которых снаряд роняли с высоты h, но ни в одном опыте взрыватель не сработал. Определить верхнюю границу p_2 90%-го доверительного интервала для вероятности p.

Решение. По формуле (14.5.15)

$$p_2 = 1 - \sqrt[100]{1-0{,}9} = 1 - \sqrt[100]{0{,}1},$$

$$\lg \sqrt[100]{0{,}1} = 0{,}01 \lg 0{,}1 = 1{,}9900;$$

$$\sqrt[100]{0{,}1} \approx 0{,}977; \quad p_2 = 1 - 0{,}977 = 0{,}023.$$

Рассмотрим еще одну задачу, связанную с предыдущей. Событие A с малой вероятностью p не наблюдалось в серии из n опытов ни разу. Задана доверительная вероятность β. Каково должно быть число опытов n для того, чтобы верхняя доверительная граница для вероятности события была равна заданному значению p_2?

Решение сразу получается из формулы (14.5.14):

$$n = \frac{\lg(1-\beta)}{\lg(1-p_2)}. \tag{14.5.16}$$

Пример 6. Сколько раз нужно убедиться в безотказной работе изделия для того, чтобы с гарантией 95% утверждать, что в практическом применении оно будет отказывать не более чем в 5% всех случаев?

Решение. По формуле (14.5.16) при $\beta = 0{,}95$, $p_2 = 0{,}05$ имеем:

$$n = \frac{\lg 0{,}05}{\lg 0{,}95} \approx 58{,}4.$$

Округляя в бо́льшую сторону, получим:

$$n = 59.$$

Имея в виду ориентировочный характер всех расчетов подобного рода, можно предложить вместо формул (14.5.15) и (14.5.16) более простые приближенные формулы. Их можно получить, предполагая,

что число появлений события A при n опытах распределено по закону Пуассона с математическим ожиданием $a = np$. Это предположение приближенно справедливо в случае, когда вероятность p очень мала (см. гл. 5, п° 5.9). Тогда

$$P(B) \approx e^{-np},$$

и вместо формулы (14.5.15) получим:

$$p_2 \approx \frac{-\ln(1-\beta)}{n}, \qquad (14.5.17)$$

а вместо формулы (14.5.16)

$$n \approx -\frac{\ln(1-\beta)}{p_2}. \qquad (14.5.18)$$

Пример 7. Найти приближенно значение p_2 для условий примера 5. Решение. По формуле (14.5.14) имеем:

$$p_2 = \frac{-\ln 0{,}1}{100} = \frac{2{,}303}{100} = 0{,}023,$$

т. е. тот же результат, который получен по точной формуле в примере 5.

Пример 8. Найти приближенно значение n для условий примера 6. Решение. По формуле (14.5.18) имеем:

$$n \approx \frac{-\ln 0{,}05}{0{,}05} = \frac{2{,}996}{0{,}05} = 59{,}9.$$

Округляя в бо́льшую сторону, находим $n = 60$, что мало отличается от результата $n = 59$, полученного в примере 6.

14.6. Оценки для числовых характеристик системы случайных величин

В п° п° 14.1—14.4 мы рассмотрели задачи, связанные с оценками для числовых характеристик одной случайной величины при ограниченном числе опытов и построением для этих характеристик доверительных интервалов.

Аналогичные вопросы возникают и при обработке ограниченного числа наблюдений над двумя и более случайными величинами.

Здесь мы ограничимся рассмотрением только точечных оценок для характеристик системы.

Рассмотрим сначала случай двух случайных величин.

Имеются результаты n независимых опытов над системой случайных величин (X, Y), давшие результаты:

$$(x_1, y_1); (x_2, y_2); \ldots; (x_n, y_n).$$

Требуется найти оценки для числовых характеристик системы: математических ожиданий m_x, m_y, дисперсий D_x, D_y и корреляционного момента K_{xy}.

Этот вопрос решается аналогично тому, как мы решили его для одной случайной величины. Несмещенными оценками для математических ожиданий будут средние арифметические:

$$\tilde{m}_x = \frac{\sum_{i=1}^{n} x_i}{n}; \qquad \tilde{m}_y = \frac{\sum_{i=1}^{n} y_i}{n}, \tag{14.6.1}$$

а для элементов корреляционной матрицы —

$$\left.\begin{aligned} \tilde{D}_x &= \frac{\sum_{i=1}^{n} (x_i - \tilde{m}_x)^2}{n-1}; \\ \tilde{D}_y &= \frac{\sum_{i=1}^{n} (y_i - \tilde{m}_y)^2}{n-1}; \\ \tilde{K}_{xy} &= \frac{\sum_{i=1}^{n} (x_i - \tilde{m}_x)(y_i - \tilde{m}_y)}{n-1}. \end{aligned}\right\} \tag{14.6.2}$$

Доказательство может быть проведено аналогично п° 14.2.

При непосредственном вычислении оценок для дисперсий и корреляционного момента часто бывает удобно воспользоваться связью между центральными и начальными статистическими моментами:

$$\left.\begin{aligned} D_x^* &= \alpha_2^*[X] - (m_x^*)^2; \\ D_y^* &= \alpha_2^*[Y] - (m_y^*)^2; \\ K_{xy}^* &= \alpha_{1,1}^*[X, Y] - m_x^* m_y^*, \end{aligned}\right\} \tag{14.6.3}$$

где

$$\left.\begin{aligned} m_x^* &= \frac{\sum_{i=1}^{n} x_i}{n}; & \alpha_2^*[X] &= \frac{\sum_{i=1}^{n} x_i^2}{n}; \\ m_y^* &= \frac{\sum_{i=1}^{n} y_i}{n}; & \alpha_2^*[Y] &= \frac{\sum_{i=1}^{n} y_i^2}{n}; \\ & & \alpha_{1,1}^*[X, Y] &= \frac{\sum_{i=1}^{n} x_i y_i}{n}. \end{aligned}\right\} \tag{14.6.4}$$

Вычислив статистические моменты по формулам (14.6.3), можно затем найти несмещенные оценки для элементов корреляционной

магрицы по формулам:

$$\left.\begin{aligned} \tilde{D}_x &= D_x^* \frac{n}{n-1}; \\ \tilde{D}_y &= D_y^* \frac{n}{n-1}; \\ \tilde{K}_{xy} &= K_{xy}^* \frac{n}{n-1}. \end{aligned}\right\} \qquad (14.6.5)$$

Пример. Произведены стрельбы с самолета по земле одиночными выстрелами. Зарегистрированы координаты точек попадания и одновременно записаны соответствующие значения угла скольжения самолета. Наблюденные значения угла скольжения β (в тысячных радиана) и абсциссы точки попадания X (в метрах) приведены в таблице 14.6.1.

Таблица 14.6.1

i	β_i	x_i	i	β_i	x_i
1	−8	−10	11	+3	−1
2	+10	−2	12	−2	+4
3	+22	+4	13	+28	+12
4	+55	+10	14	+62	+20
5	+2	−1	15	−10	−11
6	−30	−16	16	−8	+2
7	−15	−8	17	+22	+14
8	+5	−2	18	+3	+6
9	+10	+6	19	−32	−12
10	+18	+8	20	+8	+1

Найти оценки для числовых характеристик системы (β, X).

Решение. Для наглядности наносим все пары значений (β, X) на график (рис. 14.6.1). Расположение точек на графике уже свидетельствует о наличии определенной зависимости (положительной корреляции) между β и X.

По формулам (14.6.1) вычисляем средние значения величин β и X — оценки для математических ожиданий:

$$m_\beta = 7{,}15; \quad m_x = 0{,}5.$$

Далее находим статистические вторые начальные моменты:

$$\alpha_2^*[\beta] = \frac{\sum_{i=1}^{20} \beta_i^2}{20} = 581{,}8;$$

$$\alpha_2^*[X] = \frac{\sum_{i=1}^{20} x_i^2}{20} = 85{,}4,$$

По формулам (14.6.3) находим статистические дисперсии:

$$D_\beta^* = 581{,}8 - 51{,}0 = 530{,}8;$$

$$D_x^* = 85{,}4 - 0{,}2 = 85{,}2.$$

Для нахождения несмещенных оценок умножим статистические дисперсии на $\frac{n}{n-1} = \frac{20}{19}$; получим:

$$\tilde{D}_\beta = 559,$$

$$\tilde{D}_x = 89{,}7.$$

Соответственно средние квадратические отклонения равны:

$$\tilde{\sigma}_\beta = \sqrt{\tilde{D}_\beta} = 23{,}6; \quad \tilde{\sigma}_x = 9{,}46.$$

По последней формуле (14.6.4) находим статистический начальный момент:

$$\alpha_{1,1}^*[\beta, X] = \frac{\sum_{i=1}^{n} \beta_i x_i}{20} = 190{,}8$$

и статистический корреляционный момент:

$$K_{\beta x}^* = \alpha_{1,1}^*[\beta, X] - \tilde{m}_\beta \tilde{m}_x = 190{,}8 - 3{,}6 = 187{,}2.$$

Для определения несмещенной оценки умножаем его на $\frac{n}{n-1} = \frac{20}{19}$; получаем:

$$\tilde{K}_{\beta x} = 197{,}0,$$

откуда оценка для коэффициента корреляции равна:

$$\tilde{r}_{\beta x} = \frac{197{,}0}{23{,}6 \cdot 9{,}46} \approx 0{,}88.$$

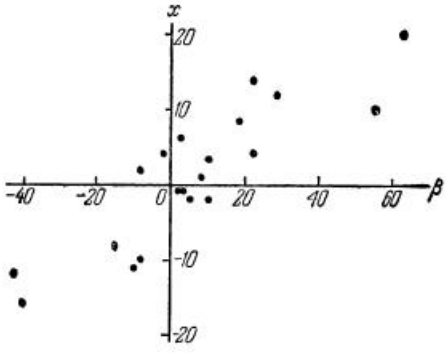


Рис. 14.6.1.

Полученное сравнительно большое значение $\tilde{r}_{\beta x}$ указывает на наличие существенной связи между β и X; на этом основании можно предполагать, что скольжение является основной причиной боковых отклонений снарядов.

Перейдем к случаю обработки наблюдений над системой произвольного числа случайных величин.

Имеется система m случайных величин

$$(X_1, X_2, \ldots, X_m).$$

Над системой произведено n независимых наблюдений; результаты этих наблюдений оформлены в виде таблицы, каждая строка которой содержит m значений, принятых случайными величинами $X_1, X_2, \ldots, X_m$ в одном наблюдении (табл. 14.6.2).

Таблица 14.6.2

i	x_1	x_2	...	x_k	...	x_m
1	x_{11}	x_{21}	...	x_{k1}	...	x_{m1}
2	x_{12}	x_{22}	...	x_{k2}	...	x_{m2}
⋮	⋮	⋮	...	⋮	...	⋮
i	x_{1i}	x_{2i}	...	x_{ki}	...	x_{mi}
⋮	⋮	⋮	...	⋮	...	⋮
n	x_{1n}	x_{2n}	...	x_{kn}	...	x_{mn}

Числа, стоящие в таблице и занумерованные двумя индексами, представляют собой зарегистрированные результаты наблюдений; первый индекс обозначает номер случайной величины, второй—номер наблюдения, так что x_{ki} — это значение, принятое величиной X_k в i-м наблюдении.

Требуется найти оценки для числовых характеристик системы: математических ожиданий $m_{x_1}, m_{x_2}, \ldots, m_{x_m}$ и элементов корреляционной матрицы:

$$\|K_{ij}\| = \begin{Vmatrix} K_{11} K_{12} \ldots K_{1m} \\ K_{22} \ldots K_{2m} \\ \ldots \\ K_{mm} \end{Vmatrix}.$$

По главной диагонали корреляционной матрицы, очевидно, стоят дисперсии случайных величин $X_1, X_2, \ldots, X_m$:

$$K_{11} = D_1; \quad K_{22} = D_2; \quad \ldots; \quad K_{mm} = D_m.$$

Оценки для математических ожиданий найдутся как средние арифметические:

$$\tilde{m}_{x_k} = \frac{\sum_{i=1}^{n} x_{ki}}{n} \qquad (k = 1, 2, \ldots, m). \qquad (14.6.6)$$

Несмещенные оценки для дисперсий определятся по формулам

$$\tilde{D}_k = \frac{\sum_{i=1}^{n} (x_{ki} - \tilde{m}_{x_k})^2}{n-1}, \qquad (14.6.7)$$

а для корреляционных моментов — по формулам

$$\tilde{K}_{kl} = \frac{\sum_{i=1}^{n} (x_{ki} - \tilde{m}_{x_k})(x_{li} - \tilde{m}_{x_l})}{n-1}. \qquad (14.6.8)$$

По этим данным определяются оценки для элементов нормированной корреляционной матрицы:

$$\tilde{r}_{kl} = \frac{\tilde{K}_{kl}}{\tilde{\sigma}_k \tilde{\sigma}_l}, \qquad (14.6.9)$$

где

$$\tilde{\sigma}_k = \sqrt{\tilde{D}_k}; \quad \tilde{\sigma}_l = \sqrt{\tilde{D}_l}. \qquad (14.6.10)$$

Пример. Сброшено 10 серий бомб, по 5 бомб в каждой, и зарегистрированы точки попадания. Результаты опытов сведены в таблицу 14.6.3. В таблице буквой l обозначен номер серии; k — номер бомбы в серии.

Требуется определить подходящие значения числовых характеристик — математических ожиданий и элементов корреляционных матриц — для системы пяти случайных величин

$$(X_1, X_2, X_3, X_4, X_5)$$

и системы пяти случайных величин

$$(Y_1, Y_2, Y_3, Y_4, Y_5).$$

Решение. Оценки для математических ожиданий найдутся как средние арифметические по столбцам:

$$\tilde{m}_{x_1} = -74{,}3; \quad \tilde{m}_{x_2} = -19{,}9; \quad \tilde{m}_{x_3} = 27{,}7; \quad \tilde{m}_{x_4} = 85{,}8; \quad \tilde{m}_{x_5} = 147{,}0;$$

$$\tilde{m}_{y_1} = -3{,}9; \quad \tilde{m}_{y_2} = -1{,}6; \quad \tilde{m}_{y_3} = 12{,}2; \quad \tilde{m}_{y_4} = 13{,}3; \quad \tilde{m}_{y_5} = 9{,}9.$$

При вычислении элементов корреляционной матрицы мы не будем, как в прежних примерах, пользоваться соотношениями между начальными и центральными моментами; в данном случае ввиду сильно изменяющихся математических ожиданий пользование этим приемом не даст преимуществ. Будем вычислять оценки для моментов непосредственно по формулам (14.6.2). Для этого вычтем из каждого элемента таблицы 14.6.3 среднее значение соответствующего столбца. Результаты сведем в таблицу 14.6.4.

Таблица 14.6.3

	Абсцисса X					Ордината Y				
l \ k	1	2	3	4	5	1	2	3	4	5
1	—120	—20	2	60	180	—20	—15	— 8	—6	—2
2	—108	—75	—20	20	80	40	60	120	125	130
3	—200	—120	—80	—20	10	—25	—30	—20	—10	2
4	—55	—2	40	120	200	—100	—75	—35	2	2
5	5	60	100	165	220	—40	—30	—25	—30	—45
6	—240	—202	—140	—88	—30	80	30	25	10	2
7	10	65	120	160	205	14	25	25	30	10
8	—40	0	65	103	170	80	75	60	10	—4
9	—100	—40	—10	55	105	—70	—60	—30	—10	0
10	105	135	190	280	330	2	4	10	12	4

Таблица 14.6.4

	$x_{kl} - \tilde{m}_{x_k}$					$y_{kl} - \tilde{m}_{y_k}$				
l \ k	1	2	3	4	5	1	2	3	4	5
1	—45,7	—0,1	—25,7	—25,8	33,0	—16,1	—13,4	—20,2	—19,3	—11,9
2	—33,7	—55,1	—37,7	—65,8	—67,0	43,9	61,6	107,8	111,7	120,1
3	—125,7	—100,1	—107,7	—105,8	—137,0	—21,1	—28,4	—32,2	—23,3	—7,9
4	19,3	17,9	12,3	34,2	53,0	—96,1	—73,4	—47,2	—11,3	—7,9
5	79,3	79,9	72,3	79,2	73,0	—36,1	—28,4	—37,2	—43,3	—54,9
6	—165,7	—182,1	—167,7	—173,8	—177,0	83,9	31,6	12,8	—3,3	—7,9
7	84,3	84,9	92,3	74,2	58,0	17,9	26,6	12,8	16,7	0,1
8	34,3	19,9	37,3	17,2	23,0	83,9	76,6	47,8	—3,3	—13,9
9	—25,7	—20,1	—37,7	—30,8	—42,0	—66,1	—58,4	—42,2	—23,3	—9,9
10	179,3	154,9	162,3	194,2	183,0	5,9	5,6	—2,2	—1,3	—5,9

Возводя эти числа в квадрат, суммируя по столбцам и деля на $n - 1 = 9$, получим оценки для дисперсий и средних квадратических отклонений:

$$\tilde{D}_{x_1} = 104{,}2 \cdot 10^2; \quad \tilde{D}_{x_2} = 94{,}4 \cdot 10^2; \quad \tilde{D}_{x_3} = 94{,}3 \cdot 10^2;$$

$$\tilde{D}_{x_4} = 110{,}2 \cdot 10^2; \quad \tilde{D}_{x_5} = 114{,}4 \cdot 10^2;$$

$$\tilde{\sigma}_{x_1} = 102; \quad \tilde{\sigma}_{x_2} = 97; \quad \tilde{\sigma}_{x_3} = 97; \quad \tilde{\sigma}_{x_4} = 105; \quad \tilde{\sigma}_{x_5} = 107.$$

$$\tilde{D}_{y_1} = 35{,}5 \cdot 10^2; \quad \tilde{D}_{y_2} = 24{,}4 \cdot 10^2; \quad \tilde{D}_{y_3} = 23{,}4 \cdot 10^2;$$

$$\tilde{D}_{y_4} = 19{,}6 \cdot 10^2; \quad \tilde{D}_{y_5} = 20{,}1 \cdot 10^2;$$

$$\tilde{\sigma}_{y_1} = 60; \quad \tilde{\sigma}_{y_2} = 49; \quad \tilde{\sigma}_{y_3} = 48; \quad \tilde{\sigma}_{y_4} = 44; \quad \tilde{\sigma}_{y_5} = 45.$$

Чтобы найти оценку для корреляционного момента, например, между величинами X_1 и X_2, составим столбец попарных произведений чисел,

стоящих в первом и втором столбцах таблицы 14.6.4 [1]). Сложив все эти произведения и разделив сумму на $n-1$, получим:

$$\tilde{K}_{x_1x_2} = 0{,}959 \cdot 10^4.$$

Деля $\tilde{K}_{x_1x_2}$ на $\tilde{\sigma}_{x_1}\tilde{\sigma}_{x_2}$, получим:

$$\tilde{r}_{x_1x_2} \approx 0{,}97.$$

Аналогично находим все остальные элементы корреляционных матриц. Для удобства умножим все элементы обеих матриц моментов на 10^{-2}. Получим:

$$\|\tilde{K}_{x_ix_j}\| = \begin{Vmatrix} 104{,}2 & 96{,}0 & 98{,}4 & 105{,}1 & 106{,}9 \\ & 94{,}4 & 93{,}2 & 107{,}0 & 101{,}1 \\ & & 94{,}3 & 102{,}0 & 99{,}2 \\ & & & 110{,}2 & 108{,}2 \\ & & & & 111{,}4 \end{Vmatrix},$$

$$\|\tilde{K}_{y_iy_j}\| = \begin{Vmatrix} 35{,}5 & 27{,}8 & 22{,}9 & 11{,}0 & 8{,}3 \\ & 24{,}4 & 21{,}4 & 13{,}0 & 10{,}2 \\ & & 23{,}4 & 17{,}4 & 17{,}2 \\ & & & 19{,}6 & 19{,}3 \\ & & & & 20{,}1 \end{Vmatrix}$$

(Ввиду симметричности матриц они заполнены только наполовину.)

Нормированные корреляционные матрицы имеют вид:

$$\|\tilde{r}_{x_ix_j}\| = \begin{Vmatrix} 1{,}00 & 0{,}97 & 0{,}99 & 0{,}98 & 0{,}98 \\ & 1{,}00 & 0{,}99 & 0{,}99 & 0{,}97 \\ & & 1{,}00 & 0{,}98 & 0{,}96 \\ & & & 1{,}00 & 0{,}97 \\ & & & & 1{,}00 \end{Vmatrix},$$

$$\|\tilde{r}_{y_iy_j}\| = \begin{Vmatrix} 1{,}00 & 0{,}94 & 0{,}76 & 0{,}42 & 0{,}31 \\ & 1{,}00 & 0{,}89 & 0{,}59 & 0{,}46 \\ & & 1{,}00 & 0{,}81 & 0{,}80 \\ & & & 1{,}00 & 0{,}97 \\ & & & & 1{,}00 \end{Vmatrix}.$$

Рассматривая эти матрицы, убеждаемся, что величины $(X_1, X_2, X_3, X_4, X_5)$ находятся в весьма тесной зависимости, приближающейся к функциональной; величины $(Y_1, Y_2, Y_3, Y_4, Y_5)$ связаны менее тесно, и коэффициенты корреляции между ними убывают по мере удаления от главной диагонали корреляционной матрицы.

---

[1]) Если вычисление производится на арифмометре или клавишной счетной машине, то нет смысла выписывать отдельные произведения, а следует непосредственно вычислять суммы произведений, не останавливаясь на промежуточных результатах.

14.7. Обработка стрельб

Одной из важных практических задач, возникающих при изучении вопросов стрельбы и бомбометания, является задача обработки результатов экспериментальных стрельб (бомбометаний).

Здесь мы будем различать два случая: стрельбу ударными снарядами и стрельбу дистанционными снарядами.

При стрельбе ударными снарядами рассеивание характеризуется законом распределения системы двух случайных величин: абсциссы и ординаты точки попадания на некоторой плоскости (реальной или воображаемой). При стрельбе дистанционными снарядами рассеивание носит пространственный характер и описывается законом распределения системы трех координат точки разрыва снаряда.

Рассмотрим сначала задачу обработки стрельб ударными снарядами. Пусть произведено n независимых выстрелов по некоторой плоской мишени и зарегистрированы координаты n точек попадания (рис. 14.7.1):

(x_1, y_1); (x_2, y_2); ...; (x_n, y_n).

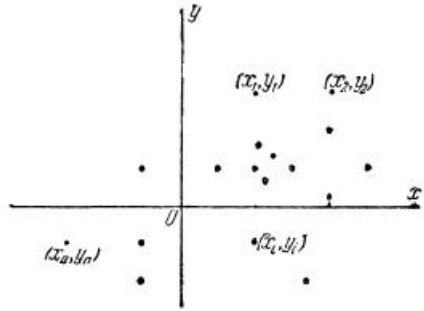


Рис. 14.7.1.

Предполагая, что закон распределения системы (X, Y) нормальный, требуется найти оценки для его параметров: координат центра рассеивания m_x, m_y, угла α, определяющего направление главных осей рассеивания ξ, η, и главных с.к.о. σ_ξ, σ_η.

Начнем с рассмотрения самого простого случая, когда направление главных осей рассеивания известно заранее. Этот случай часто встречается на практике, так как обычно направление главных осей рассеивания определяется самими условиями стрельбы (например, при бомбометании — направление полета и перпендикулярное к нему; при воздушной стрельбе — направление поперечной скорости цели и перпендикулярное к нему и т. д.). В этом случае задача обработки стрельб сильно упрощается. Зная заранее хотя бы ориентировочно направление главных осей, можно выбрать координатные оси параллельно им; в такой системе координат абсцисса и ордината точки попадания представляют собой независимые случайные величины, и их закон распределения определяется всего четырьмя параметрами: координатами центра рассеивания и главными средними квадратическими отклонениями σ_x, σ_y. Оценки для этих параметров

определяются формулами

$$\left.\begin{aligned} \tilde{m}_x &= \frac{\sum_{i=1}^{n} x_i}{n}; \quad \tilde{m}_y = \frac{\sum_{i=1}^{n} y_i}{n}; \\ \tilde{\sigma}_x &= \sqrt{\frac{\sum_{i=1}^{n} (x_i - m_x)^2}{n-1}}; \\ \tilde{\sigma}_y &= \sqrt{\frac{\sum_{i=1}^{n} (y_i - m_y)^2}{n-1}}. \end{aligned}\right\} \qquad (14.7.1)$$

Рассмотрим более сложный случай, когда направление главных осей рассеивания заранее неизвестно и тоже должно быть определено из опыта. В этом случае определению подлежат оценки всех пяти параметров: координат центра рассеивания m_x, m_y, угла α и главных средних квадратических отклонений σ_ξ, σ_η (рис. 14.7.2).

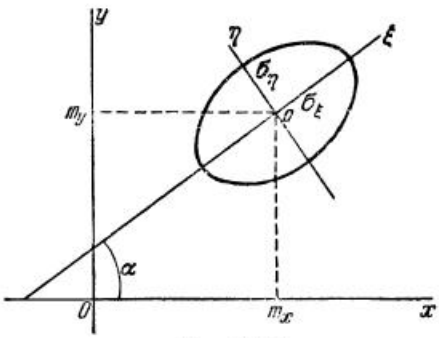


Рис. 14.7.2.

Оценки для координат центра рассеивания в этом случае определяются так же, как в предыдущем случае, по формулам

$$\tilde{m}_x = \frac{\sum_{i=1}^{n} x_i}{n}; \quad \tilde{m}_y = \frac{\sum_{i=1}^{n} y_i}{n}. \qquad (14.7.2)$$

Перейдем к оценке угла α. Предположим, что направления главных осей рассеивания известны, и проведем через точку (m_x, m_y) главные оси $O\xi$, $O\eta$ (рис. 14.7.2). В системе $\xi O\eta$ координаты случайной точки (X, Y) будут:

$$\left.\begin{aligned} \Xi &= (X - m_x)\cos\alpha + (Y - m_y)\sin\alpha, \\ \mathrm{H} &= -(X - m_x)\sin\alpha + (Y - m_y)\cos\alpha. \end{aligned}\right\}$$

или

$$\left.\begin{aligned} \Xi &= \mathring{X}\cos\alpha + \mathring{Y}\sin\alpha, \\ \mathrm{H} &= -\mathring{X}\sin\alpha + \mathring{Y}\cos\alpha. \end{aligned}\right\} \qquad (14.7.3)$$

Очевидно, величины Ξ, H будут иметь математические ожидания, равные нулю:

$$M[\Xi] = M[\mathrm{H}] = 0.$$

Так как оси $O\xi$, $O\eta$ — главные оси рассеивания, величины Ξ, H независимы. Но для величин, подчиненных нормальному закону, независимость эквивалентна некоррелированности; следовательно, нам достаточно найти такое значение угла α, при котором величины Ξ, H не коррелированы. Это значение и определит направление главных осей рассеивания.

Вычислим корреляционный момент величин (Ξ, H). Перемножая равенства (14.7.3) и применяя к их произведению операцию математического ожидания, получим:

$$K_{\xi\eta} = M[\Xi H] = -M[\mathring{X}^2]\sin\alpha\cos\alpha + M[\mathring{X}\mathring{Y}](\cos^2\alpha - \sin^2\alpha) + \\ + M[\mathring{Y}^2]\sin\alpha\cos\alpha = -\frac{1}{2}\sin 2\alpha\,(D_x - D_y) + K_{xy}\cos 2\alpha.$$

Приравнивая это выражение нулю и деля обе части на $\cos 2\alpha$, имеем:

$$\operatorname{tg} 2\alpha = \frac{2K_{xy}}{D_x - D_y}. \qquad (14.7.4)$$

Уравнение (14.7.4) даст два значения угла α: α_1 и α_2, различающиеся на $\frac{\pi}{2}$. Эти два угла и определяют направления главных осей рассеивания [1]).

Заменяя в равенстве (14.7.4) K_{xy}, D_x, D_y их оценками, получим оценку для угла α:

$$\tilde{\alpha} = \frac{1}{2}\operatorname{arctg}\frac{2\tilde{K}_{xy}}{\tilde{D}_x - \tilde{D}_y}.$$

Найдем оценки для главных средних квадратических отклонений σ_ξ, σ_η. Для этого найдем дисперсии величин Ξ, H, заданных формулами (14.7.3), по теореме о дисперсии линейной функции:

$$D_\xi = D_x\cos^2\alpha + D_y\sin^2\alpha + 2K_{xy}\sin\alpha\cos\alpha;$$
$$D_\eta = D_x\sin^2\alpha + D_y\cos^2\alpha - 2K_{xy}\sin\alpha\cos\alpha,$$

откуда находим оценки для главных дисперсий:

$$\left.\begin{aligned}\tilde{D}_\xi &= \tilde{D}_x\cos^2\tilde{\alpha} + \tilde{K}_{xy}\sin 2\tilde{\alpha} + \tilde{D}_y\sin^2\tilde{\alpha};\\ \tilde{D}_\eta &= \tilde{D}_x\sin^2\tilde{\alpha} - \tilde{K}_{xy}\sin 2\tilde{\alpha} + \tilde{D}_y\cos^2\tilde{\alpha}.\end{aligned}\right\} \qquad (14.7.5)$$

---

[1]) Заметим, что выражение (14.7.4) совпадает с приведенным в главе 9 выражением (9.2.2) для угла α, определяющего направление осей симметрии эллипса рассеивания.

Оценки для главных средних квадратических отклонений выразятся формулами:

$$\left.\begin{aligned}\tilde{\sigma}_\xi &= \sqrt{\tilde{D}_\xi};\\ \tilde{\sigma}_\eta &= \sqrt{\tilde{D}_\eta}.\end{aligned}\right\} \tag{14.7.6}$$

Выпишем отдельно полную сводку формул для обработки стрельб по плоской мишени в случае, когда направление главных осей рассеивания заранее неизвестно. Оценки искомых параметров определяются формулами:

$$\left.\begin{aligned}&\tilde{m}_x = \frac{\sum_{l=1}^{n} x_l}{n}; \quad \tilde{m}_y = \frac{\sum_{l=1}^{n} y_l}{n};\\ &\tilde{\alpha} = \frac{1}{2}\operatorname{arctg}\frac{2\tilde{K}_{xy}}{\tilde{D}_x - \tilde{D}_y};\\ &\tilde{\sigma}_\xi = \sqrt{\tilde{D}_x\cos^2\tilde{\alpha} + \tilde{K}_{xy}\sin 2\tilde{\alpha} + \tilde{D}_y\sin^2\tilde{\alpha}};\\ &\tilde{\sigma}_\eta = \sqrt{\tilde{D}_x\sin^2\tilde{\alpha} - \tilde{K}_{xy}\sin 2\tilde{\alpha} + \tilde{D}_y\cos^2\tilde{\alpha}},\end{aligned}\right\} \tag{14.7.7}$$

где

$$\left.\begin{aligned}&\tilde{D}_x = \frac{\sum_{l=1}^{n}(x_l - \tilde{m}_x)^2}{n-1}; \quad \tilde{D}_y = \frac{\sum_{l=1}^{n}(y_l - \tilde{m}_y)^2}{n-1};\\ &\tilde{K}_{xy} = \frac{\sum_{l=1}^{n}(x_l - \tilde{m}_x)(y_l - \tilde{m}_y)}{n-1}.\end{aligned}\right\} \tag{14.7.8}$$

В заключение следует заметить, что обработку стрельб по полным формулам (14.7.7) имеет смысл предпринимать только тогда, когда число опытов достаточно велико (порядка многих десятков); только в этом случае угол α может быть оценен с достаточной точностью. При малом числе наблюдений значение $\tilde{\alpha}$, полученное обработкой, является в значительной мере случайным.

Задачу обработки стрельб дистанционными снарядами мы здесь рассмотрим только в простейшем случае, когда направление главных осей рассеивания (хотя бы ориентировочно) известно заранее. Как правило, встречающиеся в практике стрельбы дистанционными снарядами задачи обработки опытов относятся к этому типу. Тогда можно выбрать координатные оси параллельно главным осям рассеивания и рассматривать три координаты точки разрыва снаряда как независимые случайные величины.

Пусть в результате n независимых выстрелов зарегистрированы координаты n точек разрыва дистанционных снарядов

$$(x_1, y_1, z_1);\ (x_2, y_2, z_2);\ \ldots,\ (x_n, y_n, z_n)$$

в системе координат с осями, параллельными главным осям рассеивания. Оценки для параметров нормального закона определятся формулами:

$$\left.\begin{aligned} \tilde{m}_x &= \frac{\sum_{i=1}^{n} x_i}{n}; \quad \tilde{m}_y = \frac{\sum_{i=1}^{n} y_i}{n}; \quad \tilde{m}_z = \frac{\sum_{i=1}^{n} z_i}{n}; \\ \tilde{\sigma}_x &= \sqrt{\frac{\sum_{i=1}^{n} (x_i - \tilde{m}_x)^2}{n-1}}; \\ \tilde{\sigma}_y &= \sqrt{\frac{\sum_{i=1}^{n} (y_i - \tilde{m}_y)^2}{n-1}}; \\ \tilde{\sigma}_z &= \sqrt{\frac{\sum_{i=1}^{n} (z_i - \tilde{m}_z)^2}{n-1}}. \end{aligned}\right\} \qquad (14.7.9)$$

На решении задачи обработки стрельб дистанционными снарядами в случае, когда направления главных осей рассеивания заранее неизвестны, мы останавливаться не будем, так как на практике эта задача встречается сравнительно редко.

14.8. Сглаживание экспериментальных зависимостей по методу наименьших квадратов

К вопросам, связанным с обработкой опытов, разобранным в данной главе, близко примыкает вопрос о сглаживании экспериментальных зависимостей.

Пусть производится опыт, целью которого является исследование зависимости некоторой физической величины y от физической величины x (например, зависимости пути, пройденного телом, от времени; начальной скорости снаряда от температуры заряда; подъемной силы от угла атаки и т. д.). Предполагается, что величины x и y связаны функциональной зависимостью:

$$y = \varphi(x). \qquad (14.8.1)$$

Вид этой зависимости и требуется определить из опыта.

Предположим, что в результате опыта мы получили ряд экспериментальных точек и построили график зависимости y от x (рис. 14.8.1). Обычно экспериментальные точки на таком графике располагаются не совсем правильным образом — дают некоторый «разброс», т. е. обнаруживают случайные отклонения от видимой общей закономер-

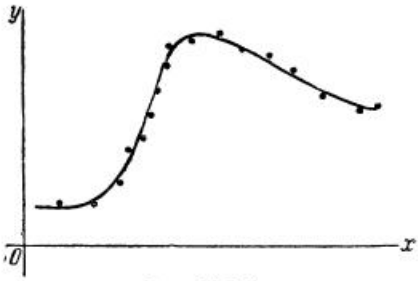


Рис. 14.8.1.

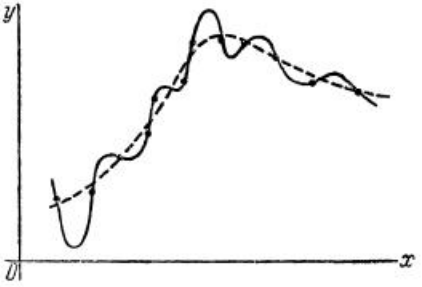


Рис. 14.8.2.

ности. Эти уклонения связаны с неизбежными при всяком опыте ошибками измерения.

Возникает вопрос, как по этим экспериментальным данным наилучшим образом воспроизвести зависимость y от x?

Известно, что через любые n точек с координатами (x_i, y_i) всегда можно провести кривую, выражаемую аналитически полиномом степени $(n-1)$, так, чтобы она в точности прошла через каждую из точек (рис. 14.8.2). Однако такое решение вопроса обычно не является удовлетворительным: как правило, нерегулярное поведение экспериментальных точек, подобное изображенному на рис. 14.8.1 и 14.8.2, связано не с объективным характером зависимости y от x, а исключительно с ошибками измерения. Это легко обнаружить, сравнивая наблюденные уклонения (разброс точек) с примерно известными ошибками измерительной аппаратуры.

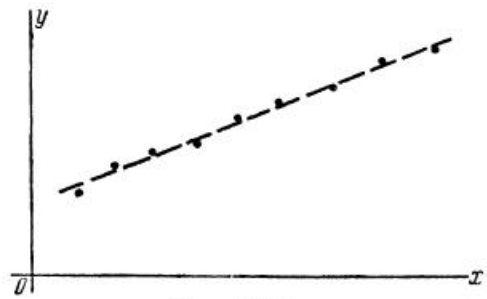


Рис. 14.8.3.

Тогда возникает весьма типичная для практики задача сглаживания экспериментальной зависимости. Желательно обработать экспериментальные данные так, чтобы по возможности точно отразить общую тенденцию зависимости y от x, но вместе с тем сгладить незакономерные, случайные уклонения, связанные с неизбежными погрешностями самого наблюдения.

Для решения подобных задач обычно применяется расчетный метод, известный под названием «метода наименьших квадратов». Этот метод дает возможность при заданном типе зависимости $y=\varphi(x)$ так выбрать ее числовые параметры, чтобы кривая $y=\varphi(x)$ в известном смысле наилучшим образом отображала экспериментальные данные.

Скажем несколько слов о том, из каких соображений может быть выбран тип кривой $y=\varphi(x)$. Часто этот вопрос решается непосредственно по внешнему виду экспериментальной зависимости. Например, экспериментальные точки, изображенные на рис. 14.8.3, явно наводят на мысль о прямолинейной зависимости вида $y=ax+b$. Зависимость, изображенная на рис. 14.8.4, хорошо может быть представлена полиномом второй степени $y=ax^2+bx+c$. Если речь идет о периодической функции, часто можно выбрать для ее изображения несколько гармоник тригонометрического ряда и т. д.

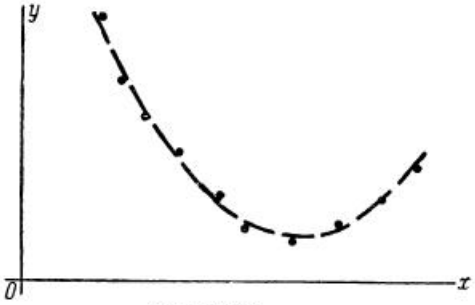


Рис. 14.8.4.

Очень часто бывает так, что вид зависимости (линейная, квадратичная, показательная и т. д.) бывает известен из физических соображений, связанных с существом решаемой задачи, а из опыта требуется установить только некоторые *параметры* этой зависимости.

Таблица 14.8.1

i	x_i	y_i
1	x_1	y_1
2	x_2	y_2
...	...	...
i	x_i	y_i
...	...	...
n	x_n	y_n

Задачу о рациональном выборе таких числовых параметров при данном виде зависимости мы и будем решать в настоящем п°.

Пусть имеются результаты n независимых опытов, оформленные в виде простой статистической таблицы (табл. 14.8.1), где i — номер опыта; x_i — значение аргумента; y_i — соответствующее значение функции.

Точки (x_i, y_i) нанесены на график (рис. 14.8.5).

Из теоретических или иных соображений выбран принципиальный вид зависимости $y=\varphi(x)$. Функция $y=\varphi(x)$ содержит ряд числовых параметров $a, b, c, \ldots$ Требуется так выбрать эти параметры, чтобы кривая $y=\varphi(x)$ в каком-то смысле наилучшим образом изображала зависимость, полученную в опыте.

Решение этой задачи, как и любой задачи выравнивания или сглаживания, зависит от того, что именно мы условимся считать «наилучшим». Можно, например, считать «наилучшим» такое взаимное расположение кривой и экспериментальных точек, при котором максимальное расстояние между ними обращается в минимум; можно потребовать, чтобы в минимум обращалась сумма абсолютных величин отклонений точек от кривой и т. д. При каждом из этих требований мы получим свое решение задачи, свои значения параметров $a, b, c, \ldots$

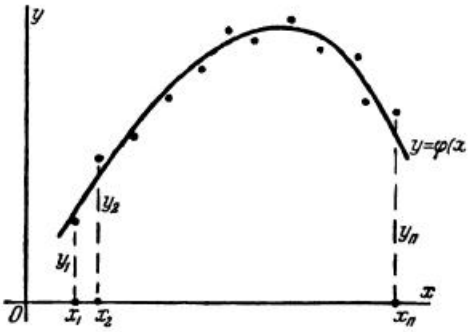


Рис. 14.8.5.

Однако общепринятым при решении подобных задач является так называемый *метод наименьших квадратов*, при котором требование наилучшего согласования кривой $y=\varphi(x)$ и экспериментальных точек сводится к тому, чтобы *сумма квадратов отклонений экспериментальных точек от сглаживающей кривой обращалась в минимум*. Метод наименьших квадратов имеет перед другими методами сглаживания существенные преимущества: во-первых, он приводит к сравнительно простому математическому способу определения параметров $a, b, c, \ldots$; во-вторых, он допускает довольно веское теоретическое обоснование с вероятностной точки зрения.

Изложим это обоснование. Предположим, что истинная зависимость y от x в точности выражается формулой $y=\varphi(x)$; экспериментальные точки уклоняются от этой зависимости вследствие неизбежных ошибок измерения. Мы уже упоминали о том, что ошибки измерения, как правило, подчиняются нормальному закону. Допустим, что это так. Рассмотрим какое-нибудь значение аргумента x_i. Результат опыта есть случайная величина Y_i, распределенная по нормальному закону с математическим ожиданием $\varphi(x_i)$ и со средним квадратическим отклонением σ_i, характеризующим ошибку измерения. Предположим, что точность измерения во всех точках одинакова:

$$\sigma_1=\sigma_2=\ldots=\sigma_n=\sigma.$$

Тогда нормальный закон, по которому распределяется величина Y_i, можно записать в виде:

$$f_i(y_i)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{[y_i-\varphi(x_i)]^2}{2\sigma^2}}, \qquad (14.8.2)$$

В результате нашего опыта — ряда измерений — произошло следующее событие: случайные величины $(Y_1, Y_2, \ldots, Y_n)$ приняли совокупность значений $(y_1, y_2, \ldots, y_n)$. Поставим задачу: так подобрать математические ожидания $\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n)$, чтобы вероятность этого события была максимальна[1]).

Строго говоря, вероятность любого из событий $Y_i = y_i$, так же как и их совмещения, равна нулю, так как величины Y_i непрерывны; поэтому мы будем пользоваться не вероятностями событий $Y_i = y_i$, а соответствующими элементами вероятностей:

$$f_i(y_i)\,dy_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[y_i-\varphi(x_i)]^2}{2\sigma^2}} dy_i. \qquad (14.8.3)$$

Найдем вероятность того, что система случайных величин $Y_1, Y_2, \ldots, Y_n$ примет совокупность значений, лежащих в пределах

$$(y_i,\ y_i+dy_i) \qquad (i=1,\ 2,\ \ldots,\ n).$$

Так как опыты независимы, эта вероятность равна произведению элементов вероятностей (14.8.3) для всех значений i:

$$\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[y_i-\varphi(x_i)]^2}{2\sigma^2}} dy_i = Ke^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i-\varphi(x_i)]^2}, \qquad (14.8.4)$$

где K — некоторый коэффициент, не зависящий от $\varphi(x_i)$.

Требуется так выбрать математические ожидания $\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n)$, чтобы выражение (14.8.4) обращалось в максимум.

Величина

$$e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i-\varphi(x_i)]^2}$$

всегда меньше единицы; очевидно, она имеет наибольшее значение, когда показатель степени по абсолютной величине минимален:

$$\frac{1}{2\sigma^2}\sum_{i=1}^{n}[y_i-\varphi(x_i)]^2 = \min.$$

Отсюда, отбрасывая постоянный множитель $\frac{1}{2\sigma^2}$, получаем требование метода наименьших квадратов: для того чтобы данная совокупность наблюденных значений

$$y_1,\ y_2,\ \ldots,\ y_n$$

была *наивероятнейшей*, нужно выбрать функцию $\varphi(x)$ так, чтобы *сумма квадратов отклонений наблюденных значений* y_i *от* $\varphi(x_i)$ *была минимальной*:

$$\sum_{i=1}^{n}[y_i-\varphi(x_i)]^2 = \min.$$

[1]) Так называемый «принцип максимального правдоподобия».

Таким образом обосновывается метод наименьших квадратов, исходя из нормального закона ошибок измерения и требования максимальной вероятности данной совокупности ошибок.

Перейдем к задаче определения параметров $a, b, c, \ldots$, исходя из принципа наименьших квадратов. Пусть имеется таблица экспериментальных данных (табл. 14.8.1) и пусть из каких-то соображений (связанных с существом явления или просто с внешним видом наблюденной зависимости) выбран общий вид функции $y=\varphi(x)$, зависящей от нескольких числовых параметров $a, b, c, \ldots$; именно эти параметры и требуется выбрать согласно методу наименьших квадратов так, чтобы сумма квадратов отклонений y_i от $\varphi(x_i)$ была минимальна. Запишем y как функцию не только аргумента x, но и параметров $a, b, c, \ldots$:

$$y=\varphi(x;\ a,\ b,\ c,\ \ldots). \tag{14.8.5}$$

Требуется выбрать $a, b, c, \ldots$ так, чтобы выполнялось условие:

$$\sum_{i=1}^{n}[y_i-\varphi(x_i;\ a,\ b,\ c,\ \ldots)]^2=\min. \tag{14.8.6}$$

Найдем значения $a, b, c, \ldots$, обращающие левую часть выражения (14.8.6) в минимум. Для этого продифференцируем ее по $a, b, c, \ldots$ и приравняем производные нулю:

$$\left.\begin{aligned}
&\sum_{i=1}^{n}[y_i-\varphi(x_i;\ a,\ b,\ c,\ \ldots)]\left(\frac{\partial\varphi}{\partial a}\right)_i=0,\\
&\sum_{i=1}^{n}[y_i-\varphi(x_i;\ a,\ b,\ c,\ \ldots)]\left(\frac{\partial\varphi}{\partial b}\right)_i=0,\\
&\sum_{i=1}^{n}[y_i-\varphi(x_i;\ a,\ b,\ c,\ \ldots)]\left(\frac{\partial\varphi}{\partial c}\right)_i=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\} \tag{14.8.7}$$

где $\left(\frac{\partial\varphi}{\partial a}\right)_i=\varphi'_a(x_i;\ a,\ b,\ c,\ \ldots)$ — значение частной производной функции φ по параметру a в точке x_i; $\left(\frac{\partial\varphi}{\partial b}\right)_i$, $\left(\frac{\partial\varphi}{\partial c}\right)_i$, $\ldots$ — аналогично.

Система уравнений (14.8.7) содержит столько же уравнений, сколько неизвестных $a, b, c, \ldots$

Решить систему (14.8.7) в общем виде нельзя; для этого необходимо задаться конкретным видом функции φ.

Рассмотрим два часто встречающихся на практике случая: когда функция φ линейна и когда она выражается полиномом второй степени (параболой).

1. Подбор параметров линейной функции методом наименьших квадратов

В опыте зарегистрирована совокупность значений (x_l, y_l) $(l = 1, 2, \ldots, n$; см. рис. 14.8.6).

Требуется подобрать по методу наименьших квадратов параметры a, b линейной функции

$$y = ax + b,$$

изображающей данную экспериментальную зависимость.

Решение. Имеем:

$$y = \varphi(x;\ a,\ b) = ax + b. \tag{14.8.8}$$

Дифференцируя выражение (14.8.8) по a и b, имеем:

$$\frac{\partial \varphi}{\partial a} = x; \quad \left(\frac{\partial \varphi}{\partial a}\right)_l = x_l;$$

$$\frac{\partial \varphi}{\partial b} = 1; \quad \left(\frac{\partial \varphi}{\partial b}\right)_l = 1.$$

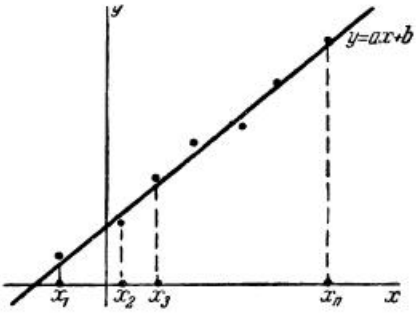


Рис. 14.8.6.

Подставляя в формулы (14.8.7), получим два уравнения для определения a и b:

$$\sum_{l=1}^{n} [y_l - (ax_l + b)]\, x_l = 0,$$

$$\sum_{l=1}^{n} [y_l - (ax_l + b)] = 0,$$

или, раскрывая скобки и производя суммирование,

$$\left.\begin{aligned} \sum_{l=1}^{n} x_l y_l - a \sum_{l=1}^{n} x_l^2 - b \sum_{l=1}^{n} x_l = 0, \\ \sum_{l=1}^{n} y_l - a \sum_{l=1}^{n} x_l - bn = 0. \end{aligned}\right\} \tag{14.8.9}$$

Разделим оба уравнения (14.8.9) на n:

$$\left.\begin{aligned} \frac{\sum_{l=1}^{n} x_l y_l}{n} - a \frac{\sum_{l=1}^{n} x_l^2}{n} - b \frac{\sum_{l=1}^{n} x_l}{n} = 0, \\ \frac{\sum_{l=1}^{n} y_l}{n} - a \frac{\sum_{l=1}^{n} x_l}{n} - b = 0. \end{aligned}\right\} \tag{14.8.10}$$

Суммы, входящие в уравнения (14.8.10), представляют собой не что иное, как уже знакомые нам статистические моменты:

$$\frac{\sum_{i=1}^{n} x_i}{n} = m_x^*; \quad \frac{\sum_{i=1}^{n} x_i^2}{n} = \alpha_2^*[X];$$

$$\frac{\sum_{i=1}^{n} y_i}{n} = m_y^*; \quad \frac{\sum_{i=1}^{n} x_i y_i}{n} = \alpha_{1,1}^*[X, Y].$$

Подставляя эти выражения в систему (14.8.10), получим:

$$\left.\begin{aligned} \alpha_{1,1}^*[X, Y] - a\alpha_2^*[X] - bm_x^* = 0, \\ m_y^* - am_x^* - b = 0. \end{aligned}\right\} \tag{14.8.11}$$

Выразим b из второго уравнения (14.8.11) и подставим в первое:

$$b = m_y^* - am_x^*;$$

$$\alpha_{1,1}^*[X, Y] - a\alpha_2^*[X] - (m_y^* - am_x^*)\, m_x^* = 0.$$

Решая последнее уравнение относительно a, имеем:

$$a = \frac{\alpha_{1,1}^*[X, Y] - m_x^* m_y^*}{\alpha_2^*[X] - (m_x^*)^2}. \tag{14.8.12}$$

Выражение (14.8.12) можно упростить, если ввести в него не начальные, а центральные моменты. Действительно,

$$\alpha_{1,1}^*[X, Y] - m_x^* m_y^* = K_{xy}^*, \qquad \alpha_2^*[X] - (m_x^*)^2 = D_x^*,$$

откуда

$$a = \frac{K_{xy}^*}{D_x^*}; \quad b = m_y^* - am_x^*, \tag{14.8.13}$$

где

$$\left.\begin{aligned} m_x^* &= \frac{\sum_{i=1}^{n} x_i}{n}; \quad m_y^* = \frac{\sum_{i=1}^{n} y_i}{n}; \\ K_{xy}^* &= \frac{\sum_{i=1}^{n} (x_i - m_x^*)(y_i - m_y^*)}{n}; \\ D_x^* &= \frac{\sum_{i=1}^{n} (x_i - m_x^*)^2}{n}. \end{aligned}\right\} \tag{14.8.14}$$

Таким образом, поставленная задача решена, и линейная зависимость, связывающая y и x, имеет вид:

$$y = \frac{K_{xy}^*}{D_x^*} x + m_y^* - \frac{K_{xy}^*}{D_x^*} m_x^*,$$

или, перенося m_y^* в левую часть,

$$y - m_y^* = \frac{K_{xy}^*}{D_x^*}(x - m_x^*). \tag{14.8.15}$$

Мы выразили коэффициенты линейной зависимости через центральные, а не через начальные вторые моменты только потому, что в таком виде формулы имеют более компактный вид. При практическом применении выведенных формул может оказаться удобнее вычислять моменты K_{xy}^* и D_x^* не по формулам (14.8.14), а через вторые начальные моменты:

$$\left.\begin{aligned} K_{xy}^* &= \frac{\sum_{i=1}^{n} x_i y_i}{n} - m_x^* m_y^*, \\ D_x^* &= \frac{\sum_{i=1}^{n} x_i^2}{n} - (m_x^*)^2. \end{aligned}\right\} \tag{14.8.16}$$

Для того чтобы формулы (14.8.16) не приводили к разностям близких чисел, рекомендуется перенести начало отсчета в точку, не слишком далекую от математических ожиданий m_x^*, m_y^*.

2. Подбор параметров параболы второго порядка методом наименьших квадратов

В опыте зарегистрированы значения (x_i, y_i) $(i = 1, 2, \ldots, n$; см. рис. 14.8.7). Требуется методом наименьших квадратов подобрать параметры квадратичной функции — параболы второго порядка:

$$y = ax^2 + bx + c,$$

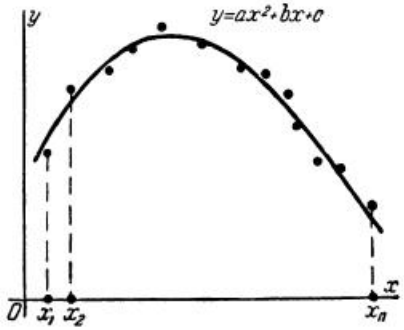


Рис. 14.8.7.

соответствующей наблюденной экспериментальной зависимости. Имеем:

$$y = \varphi(x;\ a,\ b,\ c) = ax^2 + bx + c,$$

$$\frac{\partial \varphi}{\partial a} = x^2; \quad \left(\frac{\partial \varphi}{\partial a}\right)_i = x_i^2;$$

$$\frac{\partial \varphi}{\partial b} = x; \quad \left(\frac{\partial \varphi}{\partial b}\right)_i = x_i;$$

$$\frac{\partial \varphi}{\partial c} = 1; \quad \left(\frac{\partial \varphi}{\partial c}\right)_i = 1.$$

Подставляя в уравнения (14.8.7), имеем:

$$\sum_{i=1}^{n}[y_i-(ax_i^2+bx_i+c)]x_i^2=0,$$

$$\sum_{i=1}^{n}[y_i-(ax_i^2+bx_i+c)]x_i=0,$$

$$\sum_{i=1}^{n}[y_i-(ax_i^2+bx_i+c)]=0,$$

или, раскрывая скобки, производя суммирование и деля на n,

$$\left.\begin{aligned}
&\frac{\sum_{i=1}^{n}x_i^2y_i}{n}-a\frac{\sum_{i=1}^{n}x_i^4}{n}-b\frac{\sum_{i=1}^{n}x_i^3}{n}-c\frac{\sum_{i=1}^{n}x_i^2}{n}=0,\\
&\frac{\sum_{i=1}^{n}x_iy_i}{n}-a\frac{\sum_{i=1}^{n}x_i^3}{n}-b\frac{\sum_{i=1}^{n}x_i^2}{n}-c\frac{\sum_{i=1}^{n}x_i}{n}=0,\\
&\frac{\sum_{i=1}^{n}y_i}{n}-a\frac{\sum_{i=1}^{n}x_i^2}{n}-b\frac{\sum_{i=1}^{n}x_i}{n}-c=0.
\end{aligned}\right\}\qquad(14.8.17)$$

Коэффициенты этой системы также представляют собой статистические моменты системы двух величин X, Y, а именно:

$$\frac{\sum_{i=1}^{n}x_i}{n}=m_x^*=\alpha_1^*[X];\quad\frac{\sum_{i=1}^{n}y_i}{n}=m_y^*=\alpha_1^*[Y];$$

$$\frac{\sum_{i=1}^{n}x_i^2}{n}=\alpha_2^*[X];\quad\frac{\sum_{i=1}^{n}x_i^3}{n}=\alpha_3^*[X];\quad\frac{\sum_{i=1}^{n}x_i^4}{n}=\alpha_4^*[X];$$

$$\frac{\sum_{i=1}^{n}x_iy_i}{n}=\alpha_{1,1}^*[X,Y];\quad\frac{\sum_{i=1}^{n}x_i^2y_i}{n}=\alpha_{2,1}^*[X,Y].$$

Пользуясь этими выражениями для коэффициентов через начальные моменты одной случайной величины и системы двух величин,

можно придать системе уравнений (14,8.7) достаточно компактный вид. Действительно, учитывая, что $\alpha_0^*[X]=1$; $\alpha_{0,1}^*[X, Y]=\alpha_1^*[Y]$ и перенося члены, не содержащие неизвестных, в правые части, приведем систему (14.8.17) к виду:

$$\left.\begin{aligned}
&\alpha_4^*[X]\,a+\alpha_3^*[X]\,b+\alpha_2^*[X]\,c=\alpha_{2,1}^*[X, Y],\\
&\alpha_3^*[X]\,a+\alpha_2^*[X]\,b+\alpha_1^*[X]\,c=\alpha_{1,1}^*[X, Y],\\
&\alpha_2^*[X]\,a+\alpha_1^*[X]\,b+\alpha_0^*[X]\,c=\alpha_{0,1}^*[X, Y].
\end{aligned}\right\} \qquad (14.8.18)$$

Закон образования коэффициентов в уравнениях (14.8.18) нетрудно подметить: в левой части фигурируют только моменты величины X в убывающем порядке; в правой части стоят моменты системы (X, Y), причем порядок момента по X убывает от уравнения к уравнению, а порядок по Y всегда остается первым [1]).

Аналогичными по структуре уравнениями будут определяться коэффициенты параболы любого порядка.

Мы видим, что в случае, когда экспериментальная зависимость выравнивается по методу наименьших квадратов полиномом некоторой степени, то коэффициенты этого полинома находятся решением системы линейных уравнений. Коэффициенты этой системы линейных уравнений представляют собой статистические моменты различных порядков, характеризующие систему величин (X, Y), если ее рассмотреть как систему случайных величин.

Почти так же просто решается задача сглаживания экспериментальной зависимости методом наименьших квадратов в случае, когда сглаживающая функция представляет собой не полином, а сумму произвольных заданных функций $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_k(x)$ с коэффициентами a_1, a_2, ..., a_k:

$$y=\varphi(x;\ a_1,\ a_2,\ \ldots,\ a_k)=$$

$$=a_1\varphi_1(x)+a_2\varphi_2(x)+\ \ldots\ +a_k\varphi_k(x)=\sum_{l=1}^{k}a_l\varphi_l(x), \qquad (14.8.19)$$

и когда требуется определить коэффициенты a_l.

---

[1]) Решения системы (14.8.18) относительно неизвестных a, b, c в общем виде мы не приводим, так как это решение слишком громоздко, а на практике обычно удобнее решать систему (14.8.18) не с помощью определителей, а последовательным исключением неизвестных.

Например, экспериментальную зависимость можно сглаживать тригонометрическим полиномом

$$\varphi(x;\ a_1,\ a_2,\ a_3,\ a_4) = \\ = a_1 \cos \omega x + a_2 \sin \omega x + a_3 \cos 2\omega x + a_4 \sin 2\omega x$$

или линейной комбинацией показательных функций

$$\varphi(x;\ a_1,\ a_2,\ a_3) = a_1 e^{\alpha t} + a_2 e^{\beta t} + a_3 e^{\gamma t},$$

и т. д.

В случае, если функция задается выражением типа (14.8.19), коэффициенты $a_1, a_2, \ldots, a_k$ находятся решением системы k линейных уравнений вида:

$$\sum_{i=1}^{n} \{y_i - [a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \ldots + a_k\varphi_k(x_i)]\}\,\varphi_1(x_i) = 0,$$

$$\sum_{i=1}^{n} \{y_i - [a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \ldots + a_k\varphi_k(x_i)]\}\,\varphi_2(x_i) = 0,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\sum_{i=1}^{n} \{y_i - [a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \ldots + a_k\varphi_k(x_i)]\}\,\varphi_k(x_i) = 0.$$

Выполняя почленное суммирование, имеем:

$$a_1 \sum_{i=1}^{n} [\varphi_1(x_i)]^2 + a_2 \sum_{i=1}^{n} \varphi_2(x_i)\varphi_1(x_i) + \ldots + a_k \sum_{i=1}^{n} \varphi_k(x_i)\varphi_1(x_i) = \\ = \sum_{i=1}^{n} y_i\varphi_1(x_i),$$

$$a_1 \sum_{i=1}^{n} \varphi_1(x_i)\varphi_2(x_i) + a_2 \sum_{i=1}^{n} [\varphi_2(x_i)]^2 + \ldots + a_k \sum_{i=1}^{n} \varphi_k(x_i)\varphi_2(x_i) = \\ = \sum_{i=1}^{n} y_i\varphi_2(x_i),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_1 \sum_{i=1}^{n} \varphi_1(x_i)\varphi_k(x_i) + a_2 \sum_{i=1}^{n} \varphi_2(x_i)\varphi_k(x_i) + \ldots + a_k \sum_{i=1}^{n} [\varphi_k(x_i)]^2 = \\ = \sum_{i=1}^{n} y_i\varphi_k(x_i),$$

или, короче,

$$\left.\begin{aligned}
&\sum_{j=1}^{k} a_j \sum_{i=1}^{n} \varphi_1(x_i)\varphi_j(x_i) = \sum_{i=1}^{n} y_i\varphi_1(x_i),\\
&\sum_{j=1}^{k} a_j \sum_{i=1}^{n} \varphi_2(x_i)\varphi_j(x_i) = \sum_{i=1}^{n} y_i\varphi_2(x_i),\\
&\qquad\qquad\cdots\cdots\cdots\cdots\cdots\\
&\sum_{j=1}^{k} a_j \sum_{i=1}^{n} \varphi_k(x_i)\varphi_j(x_i) = \sum_{i=1}^{n} y_i\varphi_k(x_i).
\end{aligned}\right\} \qquad (14.8.20)$$

Систему линейных уравнений (14.8.20) всегда можно решить и определить таким образом коэффициенты $a_1, a_2, \ldots, a_k$.

Сложнее решается задача о сглаживании методом наименьших квадратов, если в выражение функции $y=\varphi(x;\ a,\ b,\ c,\ \ldots)$ числовые параметры $a,\ b,\ c,\ \ldots$ входят нелинейно. Тогда решение системы (14.8.7) может оказаться сложным и трудоемким. Однако и в этом случае часто удается получить решение задачи с помощью сравнительно простых приемов.

Проиллюстрируем идею этих приемов на самом простом примере функции, нелинейно зависящей только от одного параметра a (например, $y=e^{-ax^2}$ или $y=\sin ax$, или $y=\frac{1}{ax}$). Имеем:

$$y=\varphi(x,\ a), \qquad (14.8.21)$$

где a — параметр, подлежащий подбору методом наименьших квадратов для наилучшего сглаживания заданной экспериментальной зависимости (рис. 14.8.8).

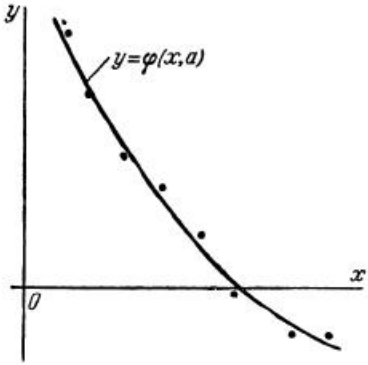


Рис. 14.8.8.

Будем решать задачу следующим образом. Зададимся рядом значений параметра a и для каждого из них найдем сумму квадратов отклонений y_i от $\varphi(x_i,\ a)$. Эта сумма квадратов есть некоторая функция a; обозначим ее $\Sigma(a)$:

$$\Sigma(a)=\sum_{i=1}^{n}[y_i-\varphi(x_i,\ a)]^2.$$

Нанесем значение $\Sigma(a)$ на график (рис. 14.8.9).

То значение a, для которого кривая $\Sigma(a)$ имеет минимум, и выбирается как подходящее значение параметра a в выражении (14.8.21).

Совершенно так же, в принципе, можно, не решая уравнений (14.8.7), подобрать совокупность двух параметров (a, b), удовлетворяющую

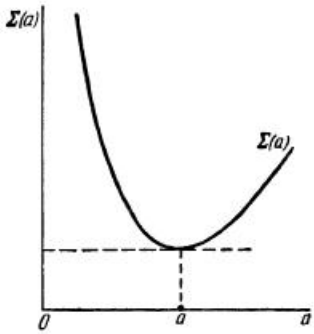


Рис. 14.8.9.

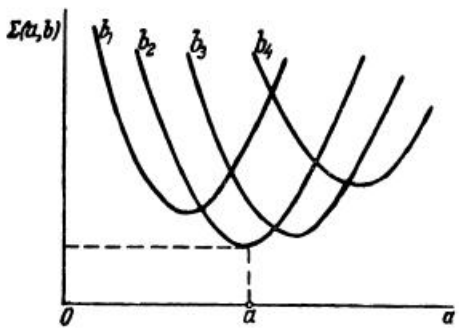


Рис. 14.8.10.

принципу наименьших квадратов; работа при этом лишь незначительно усложнится и сведется к построению не одного, а нескольких графиков (рис. 14.8.10); при этом придется искать совокупность значений a, b, обеспечивающую минимум минимального значения суммы квадратов отклонений $\Sigma(a, b)$.

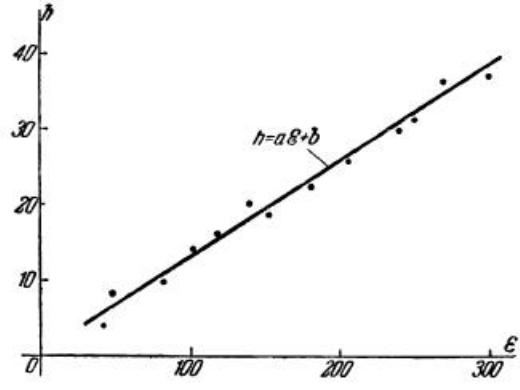


Рис. 14.8.11.

Пример 1. В опыте исследована зависимость глубины проникания h тела в преграду от удельной энергии $\mathcal{E}$ (энергии, приходящейся на квадратный сантиметр площади соударения). Экспериментальные данные приведены в таблице 14.8.2 и на графике (рис. 14.8.11).

Требуется по методу наименьших квадратов подобрать и построить прямую, изображающую зависимость h от $\mathcal{E}$.

Решение. Имеем:

$$m_{\mathcal{E}}^{*} = \frac{\sum_i \mathcal{E}_i}{13} = \frac{2137}{13} \approx 164{,}4, \quad m_h^{*} = \frac{\sum_i h_i}{13} = \frac{274}{13} \approx 21{,}1.$$

Для обработки по начальным моментам переносим начало координат в близкую к средней точку:

$$\mathcal{E}_0 = 150; \quad h_0 = 20.$$

Получаем новую таблицу значений величин:

$$\mathcal{E}' = \mathcal{E} - \mathcal{E}_0; \quad h' = h - h_0$$

(табл. 14.8.3).

Определяем моменты:

$$\alpha_2^*[\mathcal{E}'] = \frac{\sum_l \mathcal{E}_l'^2}{13} \approx 6869;$$

$$D_{\mathcal{E}}^* = 6869 - (m_{\mathcal{E}'}^*)^2 = 6869 - (164{,}4 - 150)^2 = 6662;$$

$$\alpha_{1,1}^*[\mathcal{E}', h'] = \frac{\sum_l \mathcal{E}_l' h_l'}{13} \approx 842;$$

Таблица 14.8.2

l	$\mathcal{E}_l$ (кгм/см)²	h_l (мм)
1	41	4
2	50	8
3	81	10
4	104	14
5	120	16
6	139	20
7	154	19
8	180	23
9	208	26
10	241	30
11	250	31
12	269	36
13	301	37

Таблица 14.8.3

l	$\mathcal{E}_l' = \mathcal{E}_l - 150$	$h_l' = h_l - 20$
1	—109	—16
2	—100	—12
3	—69	—10
4	—46	—6
5	—30	—4
6	—11	0
7	4	—1
8	30	3
9	58	6
10	91	10
11	100	11
12	119	16
13	151	17

$$K_{\mathcal{E}h}^* = \alpha_{1,1}^*[\mathcal{E}', h'] - m_{\mathcal{E}'}^* m_{h'}^* = 842 - (164{,}4 - 150)(21{,}1 - 20) \approx$$

$$\approx 842 - 16 = 826.$$

Уравнение прямой имеет вид:

$$h - m_h^* = \frac{K_{\mathcal{E}h}^*}{D_{\mathcal{E}}^*}(\mathcal{E} - m_{\mathcal{E}}^*),$$

или

$$h - 21{,}1 = 0{,}124\,(\mathcal{E} - 164{,}4). \qquad (14.8.22)$$

Прямая (14.8.22) показана на рис. 14.8.11.

Пример 2. Произведен ряд опытов по измерению перегрузки авиационной бомбы, проникающей в грунт, при различных скоростях встречи. Полученные значения перегрузки N в зависимости от скорости v приведены в таблице 14.8.4.

Построить по методу наименьших квадратов квадратичную зависимость вида:

$$N = av^2 + bv + c,$$

наилучшим образом согласующуюся с экспериментальными данными.

Р е ш е н и е. В целях удобства обработки удобно изменить единицы измерения так, чтобы не иметь дела с многозначными числами; для этого можно значения v выразить в сотнях *м/сек* (умножить на 10^{-2}), а N — в тысячах единиц (умножить на 10^{-3}) и всю обработку провести в этих условных единицах.

Находим коэффициенты уравнений (14.8.18).

В принятых условных единицах:

Таблица 14.8.4

i	v (*м/сек*)	N
1	120	540
2	131	590
3	140	670
4	161	760
5	174	850
6	180	970
7	200	1070
8	214	1180
9	219	1270
10	241	1390
11	250	1530
12	268	1600
13	281	1780
14	300	2030

$$\alpha_4^*[v] = \frac{\sum_i v_i^4}{n} = \frac{362{,}95}{14} = 25{,}92;$$

$$\alpha_3^*[v] = \frac{\sum_i v_i^3}{n} = \frac{148{,}36}{14} = 10{,}60;$$

$$\alpha_2^*[v] = \frac{\sum_i v_i^2}{n} = \frac{63{,}49}{14} = 4{,}535;$$

$$\alpha_1^*[v] = m_v^* = \frac{\sum_i v_i}{n} = \frac{28{,}79}{14} = 2{,}056;$$

$$\alpha_{0,\,1}^*[v,\ N] = \alpha_1^*[N] = m_N^* = \frac{\sum_i N_i}{14} = \frac{16{,}23}{14} = 1{,}159;$$

$$\alpha_{1,\,1}^*[v,\ N] = \frac{\sum_i v_i N_i}{n} = \frac{36{,}81}{14} = 2{,}629;$$

$$\alpha_{2,\,1}^*[v,\ N] = \frac{\sum_i v_i^2 N_i}{n} = \frac{88{,}02}{14} = 6{,}287.$$

Система уравнений (14.8.18) имеет вид:

$$\begin{aligned} 25{,}92a + 10{,}60b + 4{,}535c &= 6{,}287, \\ 10{,}60a + 4{,}535b + 2{,}056c &= 2{,}629, \\ 4{,}535a + 2{,}056b + \qquad c &= 1{,}159. \end{aligned}$$

Решая эту систему, находим:

$$a \approx 0{,}168; \qquad b \approx 0{,}102; \qquad c \approx 0{,}187.$$

На рис. 14.8.12 нанесены экспериментальные точки и зависимость

$$N = av^2 + bv + c,$$

построенная по методу наименьших квадратов.

П р и м е ч а н и е. В некоторых случаях может потребоваться провести кривую $y = \varphi(x)$ так, чтобы она точно проходила через некоторые заданные заранее точки. Тогда некоторые из числовых параметров a, b, c, ..., входящих в функцию $y = \varphi(x)$, могут быть определены из этих условий.

Например, в условиях примера 2 может понадобиться проэкстраполировать зависимость $N(v)$ на малые значения v; при этом естественно провести параболу второго порядка так, чтобы она проходила через начало координат (т. е. нулевой скорости встречи соответствовала нулевая перегрузка). Тогда, естественно, $c = 0$ и зависимость $N(v)$ приобретает вид:

$$N = av^2 + bv,$$

а система уравнений для определения a и b будет иметь вид:

$$25{,}92a + 10{,}60b = 6{,}287,$$

$$10{,}60a + 4{,}535b = 2{,}629.$$

Пример 3. Конденсатор, заряженный до напряжения $U_0 = 100$ *вольт*, разряжается через некоторое сопротивление. Зависимость напряжения U

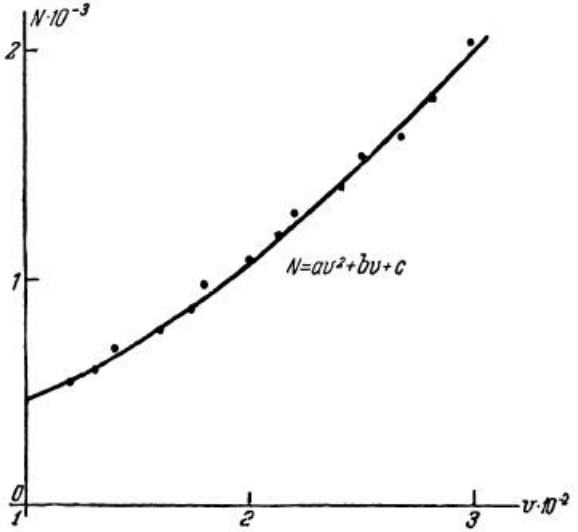


Рис. 14.8.12.

между обкладками конденсатора от времени t зарегистрирована на участке времени 10 *сек* с интервалом 1 *сек*. Напряжение измеряется с точностью до 5 *вольт*. Результаты измерений приведены в таблице 14.8.5.

Согласно теоретическим данным, зависимость напряжения от времени должна иметь вид:

$$U = U_0 e^{-\alpha t}.$$

Основываясь на опытных данных, подобрать методом наименьших квадратов значение параметра α.

Решение. По таблицам функции e^{-x} убеждаемся, что e^{-x} доходит приблизительно до 0,05 при $x \approx 3$; следовательно, коэффициент α должен иметь

Таблица 14.8.5

i	t_i (сек)	U_i (в)	i	t_i (сек)	U_i (в)
1	0	100	7	6	15
2	1	75	8	7	10
3	2	55	9	8	10
4	3	40	10	9	5
5	4	30	11	10	5
6	5	20			

порядок 0,3. Задаемся в районе $\alpha = 0{,}3$ несколькими значениями α:

$$\alpha = 0{,}28;\quad 0{,}29;\quad 0{,}30;\quad 0{,}31;\quad 0{,}32;\quad 0{,}33$$

и вычисляем для них значения функции

$$U = U_0 e^{-\alpha t}$$

в точках t_i (табл. 14.8.6). В нижней строке таблицы 14.8.6 помещены значения суммы квадратов отклонений $\Sigma(\alpha)$ в зависимости от α.

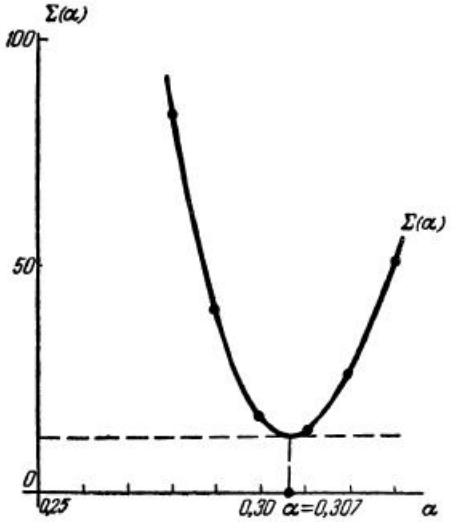


Рис. 14.8.13.

График функции $\Sigma(\alpha)$ приведен на рис. 14.8.13. Из графика видно, что значение α, отвечающее минимуму, приближенно равно 0,307. Таким образом,

по методу наименьших квадратов наилучшим приближением к опытным данным будет функция $U = U_0 e^{-0{,}307t}$. График этой функции вместе с экспериментальными точками дан на рис. 14.8.14.

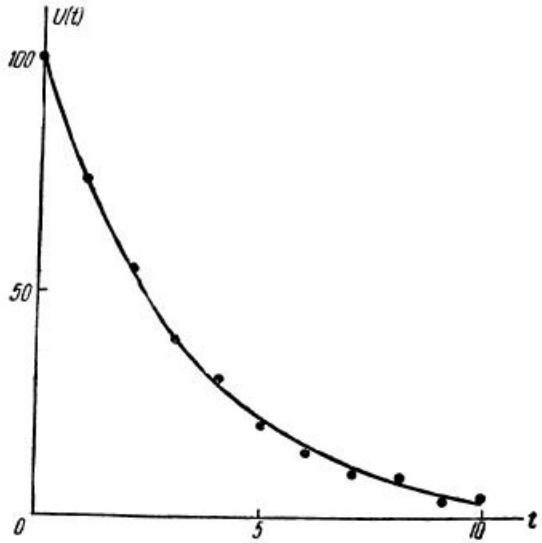


Рис. 14.8.14.

Таблица 14.8.6

t	t_t	$\alpha=0{,}28$	$\alpha=0{,}29$	$\alpha=0{,}30$	$\alpha=0{,}31$	$\alpha=0{,}32$	$\alpha=0{,}33$
1	0	100,0	100,0	100,0	100,0	100,0	100,0
2	1	75,5	74,8	74,1	73,3	72,6	71,9
3	2	57,1	56,0	54,9	53,8	52,7	51,7
4	3	43,2	41,9	40,7	39,5	38,3	37,2
5	4	32,6	31,3	30,1	28,9	27,8	26,7
6	5	24,6	23,5	22,3	21,2	20,2	19,2
7	6	18,6	17,6	16,5	15,6	14,7	13,8
8	7	14,1	13,1	12,2	11,4	10,6	9,9
9	8	10,7	9,8	9,1	8,4	7,7	7,1
10	9	8,0	7,4	6,7	6,1	5,6	5,1
11	10	6,1	5,5	5,0	4,5	4,1	3,7
	$\Sigma(\alpha)$	83,3	40,3	17,4	13,6	25,7	51,4

ГЛАВА 15

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ СЛУЧАЙНЫХ ФУНКЦИЙ

15.1. Понятие о случайной функции

До сих пор в нашем курсе теории вероятностей основным предметом исследования были случайные величины. Случайная величина характерна тем, что она в результате опыта принимает некоторое одно, заранее неизвестное, но единственное значение. Примерами таких случайных величин могут служить: абсцисса точки попадания при выстреле; ошибка радиодальномера при одном, единичном измерении дальности; горизонтальная ошибка наводки при одном выстреле и т. д.

Ограничиваясь рассмотрением подобных отдельных случайных величин, мы изучали случайные явления как бы «в статике», в каких-то фиксированных постоянных условиях отдельного опыта.

Однако такой элементарный подход к изучению случайных явлений в ряде практических задач является явно недостаточным. На практике часто приходится иметь дело со случайными величинами, непрерывно изменяющимися в процессе опыта. Примерами таких случайных величин могут служить: ошибка радиодальномера при непрерывном измерении меняющейся дальности; угол упреждения при непрерывном прицеливании по движущейся цели; отклонение траектории управляемого снаряда от теоретической в процессе управления или самонаведения.

Такие случайные величины, изменяющиеся в процессе опыта, мы будем в отличие от обычных случайных величин называть *случайными функциями*.

Изучением подобных случайных явлений, в которых случайность проявляется в форме процесса, занимается специальная отрасль теории вероятностей — теория случайных функций (иначе — теория случайных или стохастических процессов). Эту науку можно образно назвать «динамикой случайных явлений».

Теория случайных функций — новейший раздел теории вероятностей, развившийся, в основном, за последние два-три десятилетия.

В настоящее время эта теория продолжает развиваться и совершенствоваться весьма быстрыми темпами. Это связано с непосредственными требованиями практики, в частности с необходимостью решения ряда технических задач. Известно, что за последнее время в технике все большее распространение получают системы с автоматизированным управлением. Соответственно все бо́льшие требования предъявляются к теоретической базе этого вида техники — к теории автоматического управления. Развитие этой теории невозможно без анализа ошибок, неизбежно сопровождающих процессы управления, которые всегда протекают в условиях непрерывно воздействующих случайных возмущений (так называемых «помех»). Эти возмущения по своей природе являются случайными функциями. Для того чтобы рационально выбрать конструктивные параметры системы управления, необходимо изучить ее реакцию на непрерывно воздействующие случайные возмущения, а единственным аппаратом, пригодным для такого исследования, является аппарат теории случайных функций.

В данной главе мы познакомимся с основными понятиями этой теории и с общей постановкой ряда практических задач, требующих применения теории случайных функций. Кроме того, здесь будут изложены общие правила оперирования с характеристиками случайных функций, аналогичные правилам оперирования с числовыми характеристиками обычных случайных величин.

Первым из основных понятий, с которыми нам придется иметь дело, является само понятие *случайной функции*. Это понятие настолько же шире и богаче понятия случайной величины, насколько математические понятия переменной величины и функции шире и богаче понятия постоянной величины.

Вспомним определение случайной величины. Случайной величиной называется величина, которая в результате опыта может принять то или иное значение, неизвестно заранее — какое именно. Дадим аналогичное определение случайной функции.

Случайной функцией называется функция, которая в результате опыта может принять тот или иной конкретный вид, неизвестно заранее — какой именно.

Конкретный вид, принимаемый случайной функцией в результате опыта, называется *реализацией* случайной функции. Если над случайной функцией произвести группу опытов, то мы получим группу или «семейство» реализаций этой функции.

Приведем несколько примеров случайных функций.

Пример 1. Самолет-бомбардировщик на боевом курсе имеет теоретически постоянную воздушную скорость V. Фактически его скорость колеблется около этого среднего номинального значения и представляет собой случайную функцию времени. Полет на боевом курсе можно рассматривать как опыт, в котором случайная функция $V(t)$ принимает определенную реализацию (рис. 15.1.1). От опыта

к опыту вид реализации меняется. Если на самолете установлен самопишущий прибор, то он в каждом полете запишет новую, отличную от других реализацию случайной функции. В результате нескольких полетов можно получить семейство реализаций случайной функции $V(t)$ (рис. 15.1.2).

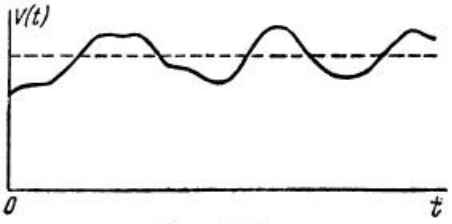


Рис. 15.1.1.

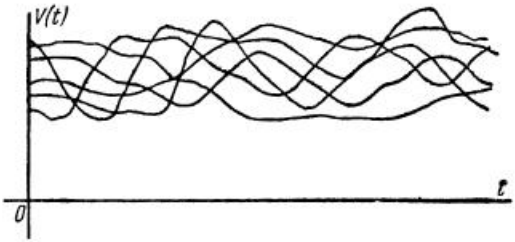


Рис. 15.1.2.

Пример 2. При наведении управляемого снаряда на цель ошибка наведения $R(t)$ представляет собой отклонение центра массы снаряда от теоретической траектории, т. е. случайную функцию времени (рис. 15.1.3).

В том же опыте случайными функциями времени являются, например, перегрузка снаряда $N(t)$, угол атаки $\alpha(t)$ и т. д.

Пример 3. При стрельбе с самолета по самолету перекрестие прицела в течение некоторого времени должно непрерывно совмещаться с целью — следить за ней. Операция слежения

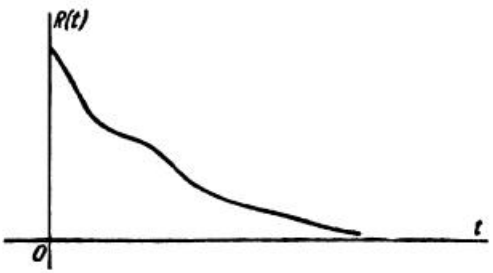


Рис. 15.1.3.

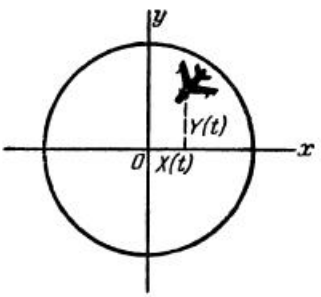


Рис. 15.1.4.

за целью сопровождается ошибками — так называемыми ошибками наводки (рис. 15.1.4). Горизонтальная и вертикальная ошибки наводки в процессе прицеливания непрерывно меняются и представляют собой две случайные функции $X(t)$ и $Y(t)$. Реализации этих случайных функ-

ций можно получить в результате дешифровки снимков фотопулемета, фотографирующего цель в течение всего процесса слежения.

Число примеров случайных функций, встречающихся в технике, можно было бы неограниченно увеличивать. Действительно, в любом случае, когда мы имеем дело с непрерывно работающей системой (системой измерения, управления, наведения, регулирования), при анализе точности работы этой системы нам приходится учитывать наличие случайных воздействий (помех). Как сами помехи, так и вызванная ими реакция системы представляют собой случайные функции времени.

До сих пор мы говорили только о случайных функциях, аргументом которых является время t. В ряде задач практики встречаются случайные функции, зависящие не от времени, а от других аргументов. Например, характеристики прочности неоднородного стержня могут рассматриваться как случайные функции абсциссы сечения x. Температура воздуха в различных слоях атмосферы может рассматриваться как случайная функция высоты H.

На практике встречаются также случайные функции, зависящие не от одного аргумента, а от нескольких. Например, аэрологические данные, характеризующие состояние атмосферы (температура, давление, ветер), представляют собой в общем случае случайные функции четырех аргументов: трех координат x, y, z и времени t.

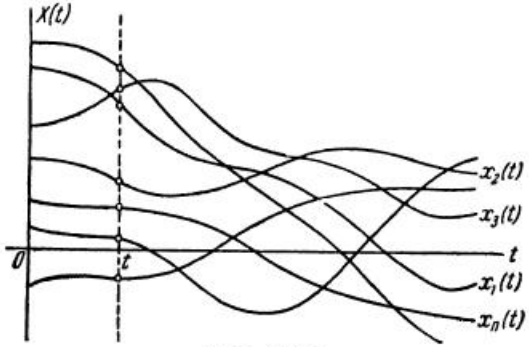


Рис. 15.1.5.

В данном курсе мы будем рассматривать только случайные функции одного аргумента. Так как этим аргументом чаще всего является время, будем обозначать его буквой t. Кроме того, условимся, как правило, обозначать случайные функции большими буквами ($X(t)$, $Y(t)$, ...) в отличие от неслучайных функций ($x(t)$, $y(t)$, ...).

Рассмотрим некоторую случайную функцию $X(t)$. Предположим, что над ней произведено n независимых опытов, в результате которых получено n реализаций (рис. 15.1.5). Обозначим их соответственно номеру опыта $x_1(t)$, $x_2(t)$, ..., $x_n(t)$.

Каждая реализация, очевидно, есть обычная (неслучайная) функция. Таким образом, в результате каждого опыта случайная функция $X(t)$ превращается в обычную, *неслучайную* функцию.

Зафиксируем теперь некоторое значение аргумента t и посмотрим, во что превратится при этом случайная функция $X(t)$. Очевидно, она превратится в *случайную величину* в обычном смысле слова. Условимся называть эту случайную величину *сечением случайной функции*, соответствующим данному t. Если провести «сечение» семейства реализаций при данном t (рис. 15.1.5), мы получим n значений, принятых случайной величиной $X(t)$ в n опытах.

Мы видим, что случайная функция совмещает в себе черты случайной величины и функции. Если зафиксировать значение аргумента, она превращается в обычную случайную величину; в результате каждого опыта она превращается в обычную (неслучайную) функцию.

В ходе дальнейшего изложения мы часто будем попеременно рассматривать одну и ту же функцию $X(t)$ то как случайную функцию, то как случайную величину, в зависимости от того, рассматривается ли она на всем диапазоне изменения t или при его фиксированном значении.

15.2. Понятие о случайной функции как расширение понятия о системе случайных величин. Закон распределения случайной функции

Рассмотрим некоторую случайную функцию $X(t)$ на определенном отрезке времени (рис. 15.2.1).

Строго говоря, случайную функцию мы не можем изображать с помощью кривой на графике: начертить мы можем лишь ее

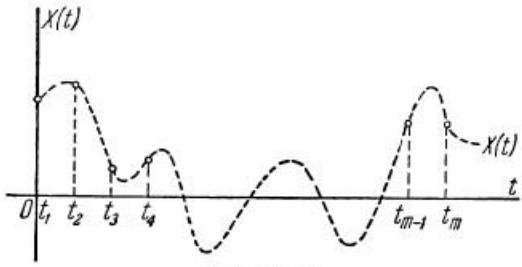


Рис. 15.2.1.

конкретные реализации. Однако в целях наглядности можно позволить себе условно изобразить на чертеже случайную функцию $X(t)$ в виде кривой, понимая под этой кривой не конкретную реализацию, а всю

совокупность возможных реализаций $X(t)$. Эту условность мы будем отмечать тем, что кривую, символически изображающую случайную функцию, будем проводить пунктиром.

Предположим, что ход изменения случайной функции регистрируется с помощью некоторого прибора, который не записывает случайную функцию непрерывно, а отмечает ее значения через определенные интервалы — в моменты времени $t_1, t_2, \ldots, t_m$.

Как было указано выше, при фиксированном значении t случайная функция превращается в обычную случайную величину. Следовательно, результаты записи в данном случае представляют собой систему m случайных величин:

$$X(t_1),\ X(t_2),\ \ldots,\ X(t_m). \qquad (15.2.1)$$

Очевидно, при достаточно высоком темпе работы регистрирующей аппаратуры запись случайной функции через такие интервалы даст достаточно точное представление о ходе ее изменения. Таким образом, рассмотрение случайной функции можно с некоторым приближением заменить рассмотрением системы случайных величин (15.2.1). По мере увеличения m такая замена становится все более и более точной. В пределе число значений аргумента — и соответственно число случайных величин (15.2.1) — становится бесконечным. Таким образом, понятие случайной функции можно рассматривать как естественное обобщение понятия системы случайных величин на случай бесконечного (несчетного) множества величин, входящих в систему.

Исходя из такого толкования случайной функции попытаемся ответить на вопрос: что же должен представлять собой *закон распределения* случайной функции?

Мы знаем, что закон распределения одной случайной величины есть функция одного аргумента, закон распределения системы двух величин — функция двух аргументов и т. д. Однако практическое пользование в качестве вероятностных характеристик функциями многих аргументов настолько неудобно, что даже для систем трех-четырех величин мы обычно отказываемся от пользования законами распределения и рассматриваем только числовые характеристики. Что касается закона распределения случайной функции, который представляет собой функцию бесчисленного множества аргументов, то такой закон в лучшем случае можно чисто формально записать в какой-либо символической форме; практическое же пользование подобной характеристикой, очевидно, совершенно исключено.

Можно, однако, для случайной функции построить некоторые вероятностные характеристики, аналогичные законам распределения. Идея построения этих характеристик заключается в следующем.

Рассмотрим случайную величину $X(t)$ — сечение случайной функции в момент t (рис. 15.2.2). Эта случайная величина, очевидно,

обладает законом распределения, который в общем случае зависит от t. Обозначим его $f(x, t)$. Функция $f(x, t)$ называется *одномерным* законом распределения случайной функции $X(t)$.

Очевидно, функция $f(x, t)$ не является полной, исчерпывающей характеристикой случайной функции $X(t)$. Действительно, эта функция характеризует только закон распределения $X(t)$ для данного,

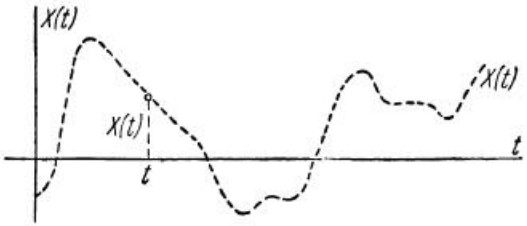


Рис. 15.2.2.

хотя и произвольного t; она не отвечает на вопрос о зависимости случайных величин $X(t)$ при различных t. С этой точки зрения более полной характеристикой случайной функции $X(t)$ является так называемый *двумерный* закон распределения:

$$f(x_1, x_2; t_1, t_2). \qquad (15.2.2)$$

Это — закон распределения системы двух случайных величин $X(t_1)$, $X(t_2)$, т. е. двух произвольных сечений случайной функции $X(t)$. Однако и эта характеристика в общем случае не является исчерпывающей; еще более полной характеристикой был бы трехмерный закон:

$$f(x_1, x_2, x_3; t_1, t_2, t_3). \qquad (15.2.3)$$

Очевидно, теоретически можно неограниченно увеличивать число аргументов и получать при этом все более подробную, все более исчерпывающую характеристику случайной функции, но оперировать со столь громоздкими характеристиками, зависящими от многих аргументов, крайне неудобно. Поэтому при исследовании законов распределения случайных функций обычно ограничиваются рассмотрением частных случаев, где для полной характеристики случайной функции достаточно, например, знания функции (15.2.2) (так называемые «процессы без последействия»).

В пределах настоящего элементарного изложения теории случайных функций мы вовсе не будем пользоваться законами распределения, а ограничимся рассмотрением простейших характеристик случайных функций, аналогичных числовым характеристикам случайных величин.

15.3. Характеристики случайных функций

Мы имели много случаев убедиться в том, какое большое значение в теории вероятностей имеют основные числовые характеристики случайных величин: математическое ожидание и дисперсия — для одной случайной величины, математические ожидания и корреляционная матрица — для системы случайных величин. Искусство пользоваться числовыми характеристиками, оставляя по возможности в стороне законы распределения, — основа прикладной теории вероятностей. Аппарат числовых характеристик представляет собой весьма гибкий и мощный аппарат, позволяющий сравнительно просто решать многие практические задачи.

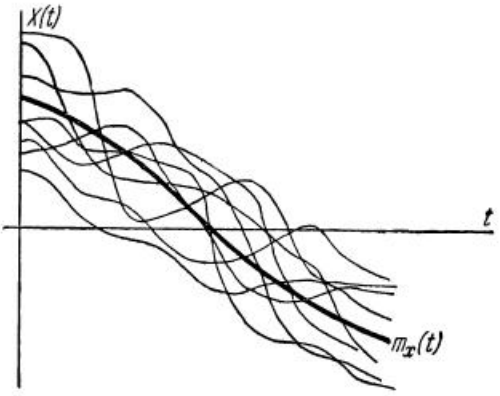


Рис. 15.3.1.

Совершенно аналогичным аппаратом пользуются и в теории случайных функций. Для случайных функций также вводятся простейшие основные характеристики, аналогичные числовым характеристикам случайных величин, и устанавливаются правила действий с этими характеристиками. Такой аппарат оказывается достаточным для решения многих практических задач.

В отличие от числовых характеристик случайных величин, представляющих собой определенные *числа*, характеристики случайных функций представляют собой в общем случае не числа, а *функции*.

Математическое ожидание случайной функции $X(t)$ определяется следующим образом. Рассмотрим сечение случайной функции $X(t)$ при фиксированном t. В этом сечении мы имеем обычную случайную величину; определим ее математическое ожидание. Очевидно, в общем случае оно зависит от t, т. е. представляет собой некоторую функцию t:

$$m_x(t) = M[X(t)]. \qquad (15.3.1)$$

Таким образом, *математическим ожиданием случайной функции* $X(t)$ *называется неслучайная функция* $m_x(t)$, *которая при каждом значении аргумента* t *равна математическому ожиданию соответствующего сечения случайной функции.*

По смыслу математическое ожидание случайной функции есть некоторая *средняя функция*, около которой различным образом варьируются конкретные реализации случайной функции.

На рис. 15.3.1 тонкими линиями показаны реализации случайной функции, жирной линией — ее математическое ожидание.

Аналогичным образом определяется дисперсия случайной функции.

Дисперсией случайной функции $X(t)$ называется неслучайная функция $D_x(t)$, значение которой для каждого t равно дисперсии соответствующего сечения случайной функции:

$$D_x(t) = D[X(t)]. \qquad (15.3.2)$$

Дисперсия случайной функции при каждом t характеризует разброс возможных реализаций случайной функции относительно среднего, иными словами, «степень случайности» случайной функции.

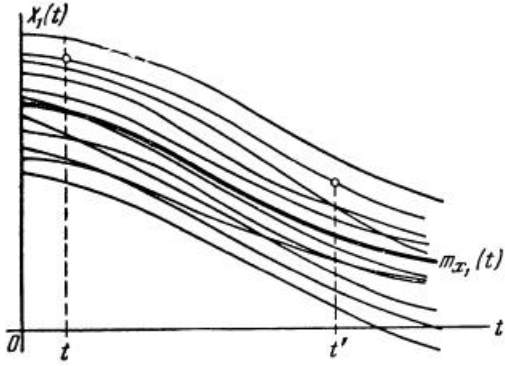


Рис. 15.3.2.

Очевидно, $D_x(t)$ есть неотрицательная функция. Извлекая из нее квадратный корень, получим функцию $\sigma_x(t)$ — среднее квадратическое отклонение случайной функции:

$$\sigma_x(t) = \sqrt{D_x(t)}. \qquad (15.3.3)$$

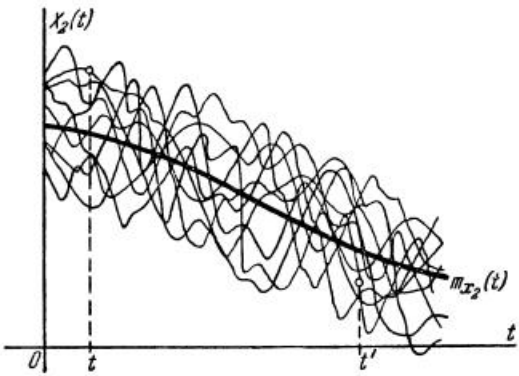


Рис. 15.3.3.

Математическое ожидание и дисперсия представляют собой весьма важные характеристики случайной функции; однако для описания основных особенностей случайной функции этих характеристик недостаточно. Чтобы убедиться в этом, рассмотрим две случайные функции $X_1(t)$ и $X_2(t)$, наглядно изображенные семействами реализаций на рис. 15.3.2 и 15.3.3.

У случайных функций $X_1(t)$ и $X_2(t)$ примерно одинаковые математические ожидания и дисперсии; однако характер этих случайных

функций резко различен. Для случайной функции $X_1(t)$ (рис. 15.3.2) характерно плавное, постепенное изменение. Если, например, в точке t случайная функция $X_1(t)$ приняла значение, заметно превышающее среднее, то весьма вероятно, что и в точке t' она также примет значение больше среднего. Для случайной функции $X_1(t)$ характерна ярко выраженная зависимость между ее значениями при различных t. Напротив, случайная функция $X_2(t)$ (рис. 15.3.3) имеет резко колебательный характер с неправильными, беспорядочными колебаниями. Для такой случайной функции характерно быстрое затухание зависимости между ее значениями по мере увеличения расстояния по t между ними.

Очевидно, внутренняя структура обоих случайных процессов совершенно различна, но это различие не улавливается ни математическим ожиданием, ни дисперсией; для его описания необходимо ввести специальную характеристику. Эта характеристика называется *корреляционной функцией* (иначе — *автокорреляционной* функцией). Корреляционная функция характеризует степень зависимости между сечениями случайной функции, относящимися к различным t.

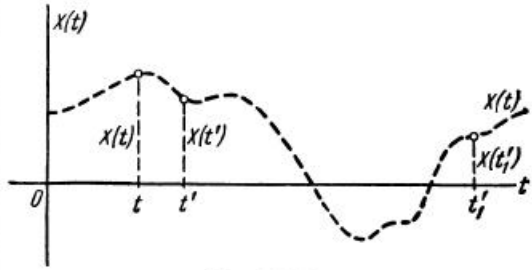


Рис. 15.3.4.

Пусть имеется случайная функция $X(t)$ (рис. 15.3.4); рассмотрим два ее сечения, относящихся к различным моментам: t и t', т. е. две случайные величины $X(t)$ и $X(t')$. Очевидно, что при близких значениях t и t' величины $X(t)$ и $X(t')$ связаны тесной зависимостью: если величина $X(t)$ приняла какое-то значение, то и величина $X(t')$ с большой вероятностью примет значение, близкое к нему. Очевидно также, что при увеличении интервала между сечениями t, t' зависимость величин $X(t)$ и $X(t')$ вообще должна убывать.

Степень зависимости величин $X(t)$ и $X(t')$ может быть в значительной мере охарактеризована их *корреляционным моментом*; очевидно, он является функцией двух аргументов t и t'. Эта функция и называется *корреляционной функцией*.

Таким образом, *корреляционной функцией случайной функции* $X(t)$ *называется неслучайная функция двух аргументов* $K_x(t, t')$, *которая при каждой паре значений* t, t' *равна корреляционному моменту соответствующих сечений случайной функции*:

$$K_x(t, t') = M[\mathring{X}(t)\,\mathring{X}(t')], \qquad (15.3.4)$$

где

$$\mathring{X}(t) = X(t) - m_x(t), \qquad \mathring{X}(t') = X(t') - m_x(t').$$

Вернемся к примерам случайных функций $X_1(t)$ и $X_2(t)$ (рис. 15.3.2 и 15.3.3). Мы видим теперь, что при одинаковых математических ожиданиях и дисперсиях случайные функции $X_1(t)$ и $X_2(t)$ имеют совершенно различные корреляционные функции. Корреляционная функция случайной функции $X_1(t)$ медленно убывает по мере увеличения промежутка (t, t'); напротив, корреляционная функция случайной функции $X_2(t)$ быстро убывает с увеличением этого промежутка.

Выясним, во что обращается корреляционная функция $K_x(t, t')$, когда ее аргументы совпадают. Полагая $t' = t$, имеем:

$$K_x(t, t) = M[(\mathring{X}(t))^2] = D_x(t), \tag{15.3.5}$$

т. е. *при $t' = t$ корреляционная функция обращается в дисперсию случайной функции.*

Таким образом, необходимость в дисперсии как отдельной характеристике случайной функции отпадает: в качестве основных характеристик случайной функции достаточно рассматривать ее математическое ожидание и корреляционную функцию.

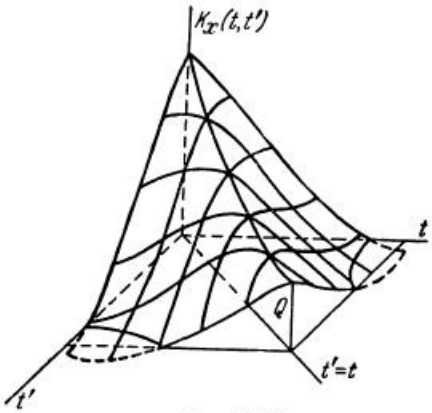


Рис. 15.3.5.

Так как корреляционный момент двух случайных величин $X(t)$ и $X(t')$ не зависит от последовательности, в которой эти величины рассматриваются, то корреляционная функция симметрична относительно своих аргументов, т. е. не меняется при перемене аргументов местами:

$$K_x(t, t') = K_x(t', t). \tag{15.3.6}$$

Если изобразить корреляционную функцию $K_x(t, t')$ в виде поверхности, то эта поверхность будет симметрична относительно вертикальной плоскости Q, проходящей через биссектрису угла tOt' (рис. 15.3.5).

Заметим, что свойства корреляционной функции естественно вытекают из свойств корреляционной матрицы системы случайных величин. Действительно, заменим приближенно случайную функцию $X(t)$ системой m случайных величин $X(t_1)$, $X(t_2)$, ..., $X(t_m)$. При увеличении m и соответственном уменьшении промежутков между

аргументами корреляционная матрица системы, представляющая собой таблицу о двух входах, в пределе переходит в функцию двух непрерывно изменяющихся аргументов, обладающую аналогичными свойствами. Свойство симметричности корреляционной матрицы относительно главной диагонали переходит в свойство симметричности корреляционной функции (15.3.6). По главной диагонали корреляционной матрицы стоят дисперсии случайных величин; аналогично при $t' = t$ корреляционная функция $K_x(t, t')$ обращается в дисперсию $D_x(t)$.

На практике, если требуется построить корреляционную функцию случайной функции $X(t)$, обычно поступают следующим образом: задаются рядом равноотстоящих значений аргумента и строят корреляционную матрицу полученной системы случайных величин. Эта матрица есть не что иное, как таблица значений корреляционной функции для прямоугольной сетки значений аргументов на плоскости (t, t'). Далее, путем интерполирования или аппроксимации можно построить функцию двух аргументов $K_x(t, t')$.

Вместо корреляционной функции $K_x(t, t')$ можно пользоваться *нормированной корреляционной функцией*:

$$r_x(t, t') = \frac{K_x(t, t')}{\sigma_x(t)\,\sigma_x(t')}, \tag{15.3.7}$$

которая представляет собой коэффициент корреляции величин $X(t)$, $X(t')$. Нормированная корреляционная функция аналогична нормированной корреляционной матрице системы случайных величин. При $t' = t$ нормированная корреляционная функция равна единице:

$$r_x(t, t) = \frac{K_x(t, t)}{[\sigma_x(t)]^2} = \frac{D_x(t)}{[\sigma_x(t)]^2} = 1. \tag{15.3.8}$$

Выясним, как меняются основные характеристики случайной функции при элементарных операциях над нею: при прибавлении неслучайного слагаемого и при умножении на неслучайный множитель. Эти неслучайные слагаемые и множители могут быть как постоянными величинами, так в общем случае и функциями t.

Прибавим к случайной функции $X(t)$ неслучайное слагаемое $\varphi(t)$. Получим новую случайную функцию:

$$Y(t) = X(t) + \varphi(t). \tag{15.3.9}$$

По теореме сложения математических ожиданий:

$$m_y(t) = m_x(t) + \varphi(t), \tag{15.3.10}$$

т. е. *при прибавлении к случайной функции неслучайного слагаемого к ее математическому ожиданию прибавляется то же неслучайное слагаемое.*

Определим корреляционную функцию случайной функции $Y(t)$:

$$K_y(t, t') = M[\mathring{Y}(t)\mathring{Y}(t')] =$$

$$= M[(Y(t) - m_y(t))(Y(t') - m_y(t'))] =$$

$$= M[(X(t) + \varphi(t) - m_x(t) - \varphi(t))(X(t') + \varphi(t') - m_x(t') - \varphi(t'))] =$$

$$= M[(X(t) - m_x(t))(X(t') - m_x(t'))] = K_x(t, t'), \quad (15.3.11)$$

т. е. *от прибавления неслучайного слагаемого корреляционная функция случайной функции не меняется.*

Умножим случайную функцию $X(t)$ на неслучайный множитель $\varphi(t)$:

$$Y(t) = \varphi(t) X(t). \quad (15.3.12)$$

Вынося неслучайную величину $\varphi(t)$ за знак математического ожидания, имеем:

$$m_y(t) = M[\varphi(t) X(t)] = \varphi(t) m_x(t), \quad (15.3.13)$$

т. е. *при умножении случайной функции на неслучайный множитель ее математическое ожидание умножается на тот же множитель.*

Определяем корреляционную функцию:

$$K_y(t, t') = M[\mathring{Y}(t)\mathring{Y}(t')] = M[(Y(t) - m_y(t))(Y(t') - m_y(t'))] =$$

$$= M[\varphi(t)\varphi(t')(X(t) - m_x(t))(X(t') - m_x(t'))] =$$

$$= \varphi(t)\varphi(t') K_x(t, t'), \quad (15.3.14)$$

т. е. *при умножении случайной функции на неслучайную функцию $\varphi(t)$ ее корреляционная функция умножается на $\varphi(t)\varphi(t')$.*

В частности, когда $\varphi(t) = c$ (не зависит от t), корреляционная функция умножается на c^2.

Пользуясь выведенными свойствами характеристик случайных функций, можно в ряде случаев значительно упростить операции с ними. В частности, когда требуется исследовать корреляционную функцию или дисперсию случайной функции, можно заранее перейти от нее к так называемой *центрированной функции*:

$$\mathring{X}(t) = X(t) - m_x(t). \quad (15.3.15)$$

Математическое ожидание центрированной функции тождественно равно нулю, а ее корреляционная функция совпадает с корреляцион-

ной функцией случайной функции $X(t)$:

$$K_{\mathring{x}}(t, t') = M[\mathring{X}(t)\,\mathring{X}(t')] = K_x(t, t'). \qquad (15.3.16)$$

При исследовании вопросов, связанных с корреляционными свойствами случайных функций, мы в дальнейшем всегда будем переходить от случайных функций к соответствующим центрированным функциям, отмечая это значком $^\circ$ вверху знака функции.

Иногда, кроме центрирования, применяется еще нормирование случайных функций. *Нормированной* называется случайная функция вида:

$$X_N(t) = \frac{\mathring{X}(t)}{\sigma_x(t)}. \qquad (15.3.17)$$

Корреляционная функция нормированной случайной функции $X_N(t)$ равна

$$K_{x_N}(t, t') = \frac{K_x(t, t')}{\sigma_x(t)\,\sigma_x(t')} = r_x(t, t'), \qquad (15.3.18)$$

а ее дисперсия равна единице.

15.4. Определение характеристик случайной функции из опыта

Пусть над случайной функцией $X(t)$ произведено n независимых опытов (наблюдений) и в результате получено n реализаций случайной функции (рис. 15.4.1).

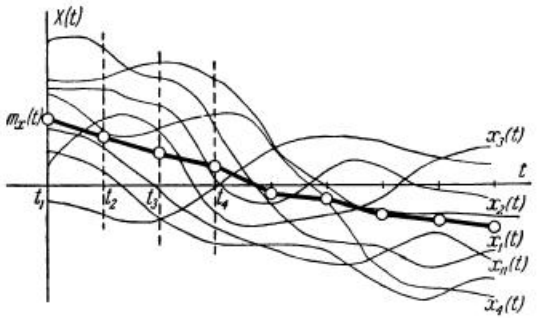


Рис. 15.4.1.

Требуется найти оценки для характеристик случайной функции: ее математического ожидания $m_x(t)$, дисперсии $D_x(t)$ и корреляционной функции $K_x(t, t')$.

Для этого рассмотрим ряд сечений случайной функции для моментов времени

$$t_1, t_2, \ldots, t_m$$

и зарегистрируем значения, принятые функцией $X(t)$ в эти моменты времени. Каждому из моментов $t_1, t_2, \ldots, t_m$ будет соответствовать n значений случайной функции.

Значения $t_1, t_2, \ldots, t_m$ обычно задаются равноотстоящими; величина интервала между соседними значениями выбирается в зависимости от вида экспериментальных кривых так, чтобы по выбранным точкам можно было восстановить основной ход кривых. Часто бывает так, что интервал между соседними значениями t задается независимо от задач обработки частотой работы регистрирующего прибора (например, темпом киноаппарата).

Зарегистрированные значения $X(t)$ заносятся в таблицу, каждая строка которой соответствует определенной реализации, а число столбцов равно числу опорных значений аргумента (табл. 15.4.1).

Таблица 15.4.1

t / $X(t)$	t_1	t_2	. . .	t_k	. . .	t_l	. . .	t_m
$x_1(t)$	$x_1(t_1)$	$x_1(t_2)$	. . .	$x_1(t_k)$	. . .	$x_1(t_l)$	. . .	$x_1(t_m)$
$x_2(t)$	$x_2(t_1)$	$x_2(t_2)$	. . .	$x_2(t_k)$	. . .	$x_2(t_l)$	. . .	$x_2(t_m)$
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
$x_i(t)$	$x_i(t_1)$	$x_i(t_2)$	. . .	$x_i(t_k)$	. . .	$x_i(t_l)$	. . .	$x_i(t_m)$
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
$x_n(t)$	$x_n(t_1)$	$x_n(t_2)$	. . .	$x_n(t_k)$	. . .	$x_n(t_l)$	. . .	$x_n(t_m)$

В таблице 15.4.1 в i-й строке помещены значения случайной функции, наблюденной в i-й реализации (i-м опыте) при значениях аргумента $t_1, t_2, \ldots, t_m$. Символом $x_i(t_k)$ обозначено значение, соответствующее i-й реализации в момент t_k.

Полученный материал представляет собой не что иное, как результаты n опытов над системой m случайных величин

$$X(t_1),\ X(t_2),\ \ldots,\ X(t_m),$$

и обрабатывается совершенно аналогично (см. п° 14.3). Прежде всего находятся оценки для математических ожиданий по формуле

$$\tilde{m}_x(t_k) = \frac{\sum_{i=1}^{n} x_i(t_k)}{n}, \qquad (15.4.1)$$

затем — для дисперсий

$$\tilde{D}_x(t_k) = \frac{\sum_{i=1}^{n} [x_i(t_k) - \tilde{m}_x(t_k)]^2}{n-1} \qquad (15.4.2)$$

и, наконец, для корреляционных моментов

$$\tilde{K}_x(t_k, t_l) = \frac{\sum_{i=1}^{n} [x_i(t_k) - \tilde{m}_x(t_k)]\,[x_i(t_l) - \tilde{m}_x(t_l)]}{n-1}. \qquad (15.4.3)$$

В ряде случаев бывает удобно при вычислении оценок для дисперсий и корреляционных моментов воспользоваться связью между начальными и центральными моментами и вычислять их по формулам:

$$\tilde{D}_x(t_k) = \left[\frac{\sum_{i=1}^{n} [x_i(t_k)]^2}{n} - [\tilde{m}_x(t_k)]^2\right] \frac{n}{n-1}; \qquad (15.4.4)$$

$$\tilde{K}_x(t_k, t_l) = \left[\frac{\sum_{i=1}^{n} x_i(t_k)\, x_i(t_l)}{n} - \tilde{m}_x(t_k)\,\tilde{m}_x(t_l)\right] \frac{n}{n-1}. \qquad (15.4.5)$$

При пользовании последними вариантами формул, чтобы избежать разности близких чисел, рекомендуется заранее перенести начало отсчета по оси ординат поближе к математическому ожиданию.

После того, как эти характеристики вычислены, можно, пользуясь рядом значений $\tilde{m}_x(t_1), \tilde{m}_x(t_2), \ldots, \tilde{m}_x(t_m)$, построить зависимость $\tilde{m}_x(t)$ (рис. 15.4.1). Аналогично строится зависимость $\tilde{D}_x(t)$. Функция двух аргументов $\tilde{K}_x(t, t')$ воспроизводится по ее значениям в прямоугольной сетке точек. В случае надобности все эти функции аппроксимируются какими-либо аналитическими выражениями.

15.5. Методы определения характеристик преобразованных случайных функций по характеристикам исходных случайных функций

В предыдущем п° мы познакомились с методом непосредственного определения характеристик случайной функции из опыта. Такой метод применяется далеко не всегда. Во-первых, постановка специальных опытов, предназначенных для исследования интересующих нас случайных функций, может оказаться весьма сложной и дорогостоящей. Во-вторых, часто нам требуется исследовать случайные функции, характеризующие ошибки приборов, прицельных приспособлений, систем управления и т. д., еще не существующих, а только проектируемых или разрабатываемых. При этом обычно исследование этих ошибок и предпринимается именно для того, чтобы рационально выбрать конструктивные параметры системы так, чтобы они приводили

к минимальным ошибкам. Ясно, что при этом непосредственное исследование случайных функций, характеризующих работу системы, нецелесообразно, а в ряде случаев вообще невозможно. В таких случаях в качестве основных рабочих методов применяются не прямые, а *косвенные* методы исследования случайных функций. Подобными косвенными методами мы уже пользовались при исследовании случайных величин: ряд глав нашего курса — гл. 10, 11, 12 — был посвящен нахождению законов распределения и числовых характеристик случайных величин косвенно, по законам распределения и числовым характеристикам других случайных величин, с ними связанных. Пользуясь совершенно аналогичными методами, можно определять характеристики случайных функций косвенно, по характеристикам других случайных функций, с ними связанных. Развитие таких косвенных методов и составляет главное содержание прикладной теории случайных функций.

Задача косвенного исследования случайных функций на практике обычно возникает в следующей форме.

Имеется некоторая динамическая система A; под «динамической системой» мы понимаем любой прибор, прицел, счетно-решающий механизм, систему автоматического управления и т. п. Эта система может быть механической, электрической или содержать любые другие элементы. Работу системы будем представлять себе следующим образом: на вход системы непрерывно поступают какие-то входные данные; система перерабатывает их и непрерывно выдает некоторый результат. Условимся называть поступающие на вход системы данные «воздействием», а выдаваемый результат «реакцией» системы на это воздействие. В качестве воздействий могут фигурировать изменяющиеся напряжения, угловые и линейные координаты каких-либо объектов, сигналы или команды, подаваемые на систему управления, и т. п. Равным образом и реакция системы может вырабатываться в той или иной форме: в виде напряжений, угловых перемещений и т. д. Например, для прицела воздушной стрельбы воздействием является угловая координата движущейся цели, непрерывно измеряемая в процессе слежения, реакцией — угол упреждения.

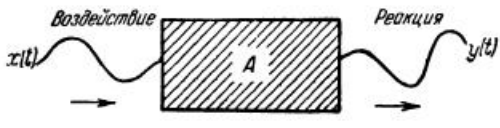


Рис. 15.5.1.

Рассмотрим самый простой случай: когда на вход системы A подается только одно воздействие, представляющее собой функцию времени $x(t)$; реакция системы на это воздействие есть другая функция времени $y(t)$. Схема работы системы A условно изображена на рис. 15.5.1.

Будем говорить, что система A осуществляет над входным воздействием некоторое *преобразование*, в результате которого

функция $x(t)$ преобразуется в другую функцию $y(t)$. Запишем это преобразование символически в виде:

$$y(t) = A[x(t)]. \qquad (15.5.1)$$

Преобразование A может быть любого вида и любой сложности. В наиболее простых случаях это, например, умножение на заданный множитель (усилители, множительные механизмы), дифференцирование или интегрирование (дифференцирующие или интегрирующие устройства). Однако на практике системы, осуществляющие в чистом виде такие простейшие преобразования, почти не встречаются; как правило, работа системы описывается дифференциальными уравнениями, и преобразование A сводится к решению дифференциального уравнения, связывающего воздействие $x(t)$ с реакцией $y(t)$.

При исследовании динамической системы в первую очередь решается основная задача: по заданному воздействию $x(t)$ определить реакцию системы $y(t)$. Однако для полного исследования системы и оценки ее технических качеств такой элементарный подход является недостаточным. В действительности воздействие $x(t)$ никогда не поступает на вход системы в чистом виде: оно всегда искажено некоторыми случайными ошибками (возмущениями), в результате которых на систему фактически воздействует не заданная функция $x(t)$, а случайная функция $X(t)$; соответственно этому система вырабатывает в качестве реакции случайную функцию $Y(t)$, также отличающуюся от теоретической реакции $y(t)$ (рис. 15.5.2).

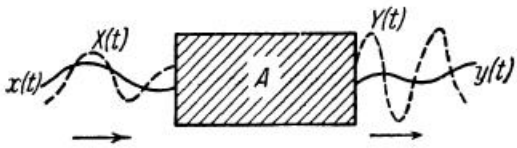


Рис. 15.5.2.

Естественно возникает вопрос: насколько велики будут случайные искажения реакции системы при наличии случайных возмущений на ее входе? И далее: как следует выбрать параметры системы для того, чтобы эти искажения были минимальными?

Решение подобных задач не может быть получено методами классической теории вероятностей; единственным подходящим математическим аппаратом для этой цели является аппарат теории случайных функций.

Из двух поставленных выше задач, естественно, более простой является первая — прямая — задача. Сформулируем ее следующим образом.

На вход динамической системы A поступает случайная функция $X(t)$; система подвергает ее известному преобразованию, в результате чего на выходе системы появляется случайная функция:

$$Y(t) = A[X(t)]. \qquad (15.5.2)$$

Известны характеристики случайной функции $X(t)$: математическое ожидание и корреляционная функция. Требуется найти аналогичные характеристики случайной функции $Y(t)$. Короче: по заданным характеристикам случайной функции на входе динамической системы найти характеристики случайной функции на выходе.

Поставленная задача может быть решена совершенно точно в одном частном, но весьма важном для практики случае: когда преобразование A принадлежит к классу так называемых *линейных преобразований* и соответственно система A принадлежит к классу *линейных систем*.

Содержание этих понятий будет пояснено в следующем п°.

15.6. Линейные и нелинейные операторы. Оператор динамической системы

При изложении теории преобразования случайных функций мы будем пользоваться широко применяемым в математике и технике понятием *оператора*.

Понятие оператора является обобщением понятия функции. Когда мы устанавливаем функциональную связь между двумя переменными y и x и пишем:

$$y = f(x), \tag{15.6.1}$$

то под символом f мы понимаем правило, по которому заданному значению x приводится в соответствие вполне определенное значение y. Знак f есть символ некоторого *преобразования*, которому нужно подвергнуть величину x, чтобы получить y. Соответственно виду этого преобразования функции могут быть линейными и нелинейными, алгебраическими, трансцендентными и т. д.

Аналогичные понятия и соответствующая символика применяются в математике и в тех случаях, когда преобразованию подвергаются не величины, а функции.

Рассмотрим некоторую функцию $x(t)$ и установим определенное правило A, согласно которому функция $x(t)$ преобразуется в другую функцию $y(t)$. Запишем это преобразование в следующем виде:

$$y(t) = A[x(t)]. \tag{15.6.2}$$

Примерами подобных преобразований могут быть, например, дифференцирование:

$$y(t) = \frac{dx(t)}{dt}, \tag{15.6.3}$$

интегрирование:

$$y(t) = \int_0^t x(\tau)\, d\tau, \tag{15.6.4}$$

и т. д.

Правило A, согласно которому функция $x(t)$ преобразуется в функцию $y(t)$, мы будем называть *оператором*; например, мы

будем говорить: оператор дифференцирования, оператор интегрирования, оператор решения дифференциального уравнения и т. д.

Определяя оператор, мы рассматривали только преобразование функции $x(t)$ в другую функцию y того же аргумента t. Следует заметить, что такое сохранение аргумента при определении оператора вовсе не является обязательным: оператор может преобразовывать функцию $x(t)$ в функцию другого аргумента $y(s)$, например:

$$y(s) = \int_a^b \varphi(t, s)\, x(t)\, dt, \tag{15.6.5}$$

где $\varphi(t, s)$ — некоторая функция, зависящая, помимо аргумента t, еще и от параметра s.

Но так как при анализе ошибок динамических систем наиболее естественным аргументом является время t, мы здесь ограничимся рассмотрением операторов, преобразующих одну функцию аргумента t в другую функцию того же аргумента.

Если динамическая система преобразует поступающую на ее вход функцию $x(t)$ в функцию $y(t)$:

$$y(t) = A\{x(t)\},$$

то оператор A называется *оператором динамической системы*.

В более общем случае на вход системы поступает не одна, а несколько функций; равным образом на выходе системы могут появляться несколько функций; в этом случае оператор системы преобразует одну совокупность функций в другую. Однако в целях простоты изложения мы рассмотрим здесь лишь наиболее элементарный случай преобразования одной функции в другую.

Преобразования или операторы, применяемые к функциям, могут быть различных типов. Наиболее важным для практики является класс так называемых *линейных операторов*.

Оператор L называется *линейным однородным*, если он обладает следующими свойствами:

1) к сумме функций оператор может применяться почленно:

$$L\{x_1(t) + x_2(t)\} = L\{x_1(t)\} + L\{x_2(t)\}; \tag{15.6.6}$$

2) постоянную величину c можно выносить за знак оператора:

$$L\{cx(t)\} = cL\{x(t)\}. \tag{15.6.7}$$

Из второго свойства, между прочим, следует, что для линейного однородного оператора справедливо свойство

$$L\{0\} = 0, \tag{15.6.8}$$

т. е. при нулевом входном воздействии реакция системы равна нулю.

Примеры линейных однородных операторов:

1) оператор дифференцирования:

$$y(t) = \frac{dx(t)}{dt};$$

2) оператор интегрирования:

$$y(t) = \int_0^t x(\tau)\, d\tau;$$

3) оператор умножения на определенную функцию $\varphi(t)$:

$$y(t) = \varphi(t)\, x(t);$$

4) оператор интегрирования с заданным «весом» $\varphi(t)$:

$$y(t) = \int_0^t x(\tau)\, \varphi(\tau)\, d\tau$$

и т. д.

Кроме линейных однородных операторов, существуют еще линейные неоднородные операторы.

Оператор L называется *линейным неоднородным*, если он состоит из линейного однородного оператора с прибавлением некоторой вполне определенной функции $\varphi(t)$:

$$L\{x(t)\} = L_0\{x(t)\} + \varphi(t), \tag{15.6.9}$$

где L_0 — линейный однородный оператор.

Примеры линейных неоднородных операторов:

1) $y(t) = \frac{dx(t)}{dt} + \varphi(t)$,

2) $y(t) = \int_0^t x(\tau)\, \varphi(\tau)\, d\tau + \varphi_1(t)$,

3) $y(t) = \varphi_1(t)\, x(t) + \varphi_2(t)$,

где $\varphi(t)$, $\varphi_1(t)$, $\varphi_2(t)$ — вполне определенные функции, а $x(t)$ — преобразуемая оператором функция.

В математике и технике широко применяется условная форма записи операторов, аналогичная алгебраической символике. Такая символика в ряде случаев позволяет избегать сложных преобразований и записывать формулы в простой и удобной форме.

Например, оператор дифференцирования часто обозначают буквой p:

$$p = \frac{d}{dt},$$

помещаемой в виде множителя перед выражением, подлежащим дифференцированию. При этом запись

$$y(t) = px(t)$$

равносильна записи

$$y(t) = \frac{dx(t)}{dt}.$$

Двойное дифференцирование обозначается множителем p^2:

$$p^2 x(t) = \frac{d^2 x(t)}{dt^2}$$

и т. д.

Пользуясь подобной символикой, в частности, очень удобно записывать дифференциальные уравнения.

Пусть, например, работа динамической системы A описывается линейным дифференциальным уравнением с постоянными коэффициентами, связывающими реакцию системы $y(t)$ с воздействием $x(t)$. В обычной форме записи это дифференциальное уравнение имеет вид:

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \ldots + a_1 \frac{dy(t)}{dt} + a_0 y(t) =$$

$$= b_m \frac{d^m x(t)}{dt^m} + b_{m-1} \frac{d^{m-1} x(t)}{dt^{m-1}} + \ldots + b_1 \frac{dx(t)}{dt} + b_0 x(t). \quad (15.6.10)$$

В символической форме это уравнение может быть записано в виде:

$$(a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_0)\, y(t) =$$
$$= (b_m p^m + b_{m-1} p^{m-1} + \ldots + b_1 p + b_0)\, x(t),$$

где $p = \frac{d}{dt}$ — оператор дифференцирования.

Обозначая для краткости полиномы относительно p, входящие в правую и левую части,

$$A_n(p) = a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_0,$$
$$B_m(p) = b_m p^m + b_{m-1} p^{m-1} + \ldots + b_1 p + b_0,$$

запишем уравнение в еще более компактной форме:

$$A_n(p)\, y(t) = B_m(p)\, x(t). \quad (15.6.11)$$

Наконец, формально разрешая уравнение (15.6.11) относительно $y(t)$, можно символически записать оператор решения линейного дифференциального уравнения в «явном» виде:

$$y(t) = \frac{B_m(p)}{A_n(p)}\, x(t). \quad (15.6.12)$$

Пользуясь аналогичной символикой, можно записать в операторной форме и линейное дифференциальное уравнение с переменными

коэффициентами. В обычной форме это уравнение имеет вид:

$$a_n(t)\frac{d^n y(t)}{dt^n}+a_{n-1}(t)\frac{d^{n-1}y(t)}{dt^{n-1}}+\ldots+a_1(t)\frac{dy(t)}{dt}+a_0(t)\,y(t)=$$

$$=b_m(t)\frac{d^m x(t)}{dt^m}+b_{m-1}(t)\frac{d^{m-1}x(t)}{dt^{m-1}}+\ldots+b_1(t)\frac{dx(t)}{dt}+b_0(t)\,x(t). \tag{15.6.13}$$

Обозначая многочлены относительно p, коэффициенты которых зависят от t,

$$A_n(p,\,t)=a_n(t)\,p^n+a_{n-1}(t)\,p^{n-1}+\ldots+a_1(t)\,p+a_0(t),$$
$$B_m(p,\,t)=b_m(t)\,p^m+b_{m-1}(t)\,p^{m-1}+\ldots+b_1(t)\,p+b_0(t),$$

можно записать оператор дифференциального уравнения в виде:

$$y(t)=\frac{B_m(p,\,t)}{A_n(p,\,t)}\,x(t). \tag{15.6.14}$$

В дальнейшем мы по мере надобности будем пользоваться такой символической формой записи операторов.

Встречающиеся в технике динамические системы часто описываются линейными дифференциальными уравнениями. В этом случае, как нетрудно убедиться, оператор системы является линейным.

Динамическая система, оператор которой является линейным, называется *линейной динамической системой*.

В противоположность линейным операторам и системам рассматриваются системы и операторы *нелинейные*. Примерами нелинейных операторов могут служить

$$y(t)=x^2(t),\quad y(t)=\int_0^t x^3(\tau)\,d\tau,\quad y(t)=\sin x(t),$$

а также решение нелинейного дифференциального уравнения, хотя бы

$$y'(t)+\alpha\cos y(t)=x(t).$$

Динамическая система, оператор которой не является линейным, называется *нелинейной системой*.

На практике линейные системы встречаются очень часто. В связи с линейностью этих систем к анализу их ошибок может быть с большой эффективностью применен аппарат теории случайных функций. Подобно тому как числовые характеристики линейных функций обычных случайных величин могут быть получены по числовым характеристикам аргументов, характеристики случайной функции на выходе линейной динамической системы могут быть определены, если известны оператор системы и характеристики случайной функции на ее входе.

Еще чаще, чем линейные системы, на практике встречаются системы не строго линейные, но в известных пределах допускающие линеаризацию. Если случайные возмущения на входе системы достаточно малы, то практически любая система может рассматриваться — в пределах этих малых возмущений — как приближенно линейная, подобно тому как при достаточно малых случайных изменениях аргументов практически любая функция может быть линеаризована.

Прием приближенной линеаризации дифференциальных уравнений широко применяется в теории ошибок динамических систем.

В дальнейшем мы будем рассматривать только линейные (или линеаризуемые) динамические системы и соответствующие им линейные операторы.

15.7. Линейные преобразования случайных функций

Пусть на вход линейной системы с оператором L воздействует случайная функция $X(t)$, причем известны ее характеристики: математическое ожидание $m_x(t)$ и корреляционная функция $K_x(t, t')$. Реакция системы представляет собой случайную функцию

$$Y(t) = L[X(t)]. \tag{15.7.1}$$

Требуется найти характеристики случайной функции $Y(t)$ на выходе системы: $m_y(t)$ и $K_y(t, t')$. Короче: по характеристикам случайной функции на входе линейной системы найти характеристики случайной функции на выходе.

Покажем сначала, что можно ограничиться решением этой задачи только для однородного оператора L. Действительно, пусть оператор L неоднороден и выражается формулой

$$L[X(t)] = L_0[X(t)] + \varphi(t), \tag{15.7.2}$$

где L_0 — линейный однородный оператор, $\varphi(t)$ — определенная неслучайная функция. Тогда

$$m_y(t) = M[L_0[X(t)]] + \varphi(t), \tag{15.7.3}$$

т. е. функция $\varphi(t)$ просто прибавляется к математическому ожиданию случайной функции на выходе линейной системы. Что же касается корреляционной функции, то, как известно, она не меняется от прибавления к случайной функции неслучайного слагаемого.

Поэтому в дальнейшем изложении под «линейными операторами» будем разуметь только *линейные однородные операторы*.

Решим задачу об определении характеристик на выходе линейной системы сначала для некоторых частных видов линейных операторов.

1. Интеграл от случайной функции

Дана случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$. Случайная функция $Y(t)$ связана с $X(t)$ линейным однородным оператором интегрирования:

$$Y(t) = \int_0^t X(\tau)\, d\tau. \tag{15.7.4}$$

Требуется найти характеристики случайной функции $Y(t)$: $m_y(t)$ и $K_y(t, t')$.

Представим интеграл (15.7.4) как предел суммы:

$$Y(t) = \int_0^t X(\tau)\, d\tau = \lim_{\Delta\tau \to 0} \sum_i X(\tau_i)\, \Delta\tau \tag{15.7.5}$$

и применим к равенству (15.7.5) операцию математического ожидания. По теореме сложения математических ожиданий имеем:

$$m_y(t) = M[Y(t)] = \lim_{\Delta\tau \to 0} \sum_i M[X(\tau_i)]\, \Delta\tau =$$

$$= \lim_{\Delta\tau \to 0} \sum_i m_x(\tau_i)\, \Delta\tau = \int_0^t m_x(\tau)\, d\tau\ ^{1}). \tag{15.7.6}$$

Итак,

$$m_y(t) = \int_0^t m_x(\tau)\, d\tau, \tag{15.7.7}$$

т. е. *математическое ожидание интеграла от случайной функции равно интегралу от ее математического ожидания.* Иными словами: операцию интегрирования и операцию математического ожидания можно менять местами. Это и естественно, так как операция интегрирования по своей природе не отличается от операции суммирования, которую, как мы раньше убедились, можно менять местами с операцией математического ожидания.

Найдем корреляционную функцию $K_y(t, t')$. Для этого перейдем к центрированным случайным функциям:

$$\mathring{X}(t) = X(t) - m_x(t); \quad \mathring{Y}(t) = Y(t) - m_y(t).$$

Нетрудно убедиться, что

$$\mathring{Y}(t) = \int_0^t \mathring{X}(\tau)\, d\tau. \tag{15.7.8}$$

---

[1]) При этом мы предполагаем, что математическое ожидание предела равно пределу от математического ожидания. На практике мы, как правило, имеем дело с функциями, для которых такая перестановка возможна.

По определению корреляционной функции,

$$K_y(t, t') = M[\mathring{Y}(t)\,\mathring{Y}(t')],$$

где

$$\mathring{Y}(t) = \int_0^t \mathring{X}(\tau)\,d\tau; \quad \mathring{Y}(t') = \int_0^{t'} \mathring{X}(\tau')\,d\tau' \text{ [1]}. \tag{15.7.9}$$

Перемножим выражения (15.7.9):

$$\mathring{Y}(t)\,\mathring{Y}(t) = \int_0^t \mathring{X}(\tau)\,d\tau \int_0^{t'} \mathring{X}(\tau')\,d\tau'. \tag{15.7.10}$$

Нетрудно убедиться, что произведение двух интегралов в правой части формулы (15.7.10) равно двойному интегралу

$$\int_0^t \int_0^{t'} \mathring{X}(\tau)\,\mathring{X}(\tau')\,d\tau\,d\tau'. \tag{15.7.11}$$

Действительно, в связи с тем, что подынтегральная функция в интеграле (15.7.11) распадается на два множителя, из которых первый зависит только от τ, второй — только от τ', двойной интеграл (15.7.11) распадается на произведение двух однократных интегралов (15.7.10). Следовательно,

$$\mathring{Y}(t)\,\mathring{Y}(t') = \int_0^t \int_0^{t'} \mathring{X}(\tau)\,\mathring{X}(\tau')\,d\tau\,d\tau'. \tag{15.7.12}$$

Применяя к равенству (15.7.12) операцию математического ожидания и меняя ее в правой части местами с операцией интегрирования, получим:

$$K_y(t, t') = M[\mathring{Y}(t)\,\mathring{Y}(t')] = \int_0^t \int_0^{t'} M[\mathring{X}(\tau)\,\mathring{X}(\tau')]\,d\tau\,d\tau',$$

или окончательно:

$$K_y(t, t') = \int_0^t \int_0^{t'} K_x(\tau, \tau')\,d\tau\,d\tau'. \tag{15.7.13}$$

Таким образом, для того чтобы найти корреляционную функцию интеграла от случайной функции, нужно дважды проинтегрировать корреляционную функцию исходной случайной функции: сначала по одному аргументу, затем — по другому.

---

[1]) Как известно, в определенном интеграле переменная интегрирования может быть обозначена любой буквой; в данном случае удобно обозначить ее τ и τ' соответственно.

2. Производная от случайной функции

Дана случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$. Случайная функция $Y(t)$ связана со случайной функцией $X(t)$ линейным однородным оператором дифференцирования:

$$Y(t) = \frac{dX(t)}{dt}. \tag{15.7.14}$$

Требуется найти $m_y(t)$ и $K_y(t, t')$.

Представим производную в виде предела:

$$Y(t) = \lim_{\Delta t \to 0} \frac{X(t+\Delta t) - X(t)}{\Delta t}. \tag{15.7.15}$$

Применяя к равенству (15.7.15) операцию математического ожидания, получим:

$$m_y(t) = M[Y(t)] = \lim_{\Delta t \to 0} \frac{m_x(t+\Delta t) - m_x(t)}{\Delta t} = \frac{dm_x(t)}{dt}\ ^{1}).$$

Итак,

$$m_y(t) = \frac{dm_x(t)}{dt}, \tag{15.7.16}$$

т. е. *математическое ожидание производной от случайной функции равно производной от ее математического ожидания.* Следовательно, операцию дифференцирования, как и операцию интегрирования, тоже можно менять местами с операцией математического ожидания.

Для определения $K_y(t, t')$ перейдем к центрированным случайным функциям $\mathring{Y}(t)$ и $\mathring{X}(t)$; очевидно,

$$\mathring{Y}(t) = \frac{d\mathring{X}(t)}{dt}. \tag{15.7.17}$$

По определению

$$K_y(t, t') = M[\mathring{Y}(t)\,\mathring{Y}(t')].$$

Подставим вместо $\mathring{Y}(t)$ и $\mathring{Y}(t')$ их выражения:

$$K_y(t, t') = M\left[\frac{d\mathring{X}(t)}{dt}\,\frac{d\mathring{X}(t')}{dt'}\right].$$

Представим выражение под знаком математического ожидания в виде второй смешанной частной производной:

$$\frac{d\mathring{X}(t)}{dt}\,\frac{d\mathring{X}(t')}{dt'} = \frac{\partial^2 \mathring{X}(t)\,\mathring{X}(t')}{\partial t\,\partial t'}. \tag{15.7.18}$$

Мы доказали, что математическое ожидание производной случайной функции равно производной от математического ожидания, т. е.

---

[1]) См. сноску [1]) на стр. 394.

знаки дифференцирования и математического ожидания можно менять местами. Следовательно,

$$K_y(t, t') = M\left[\frac{\partial^2 \mathring{X}(t)\,\mathring{X}(t')}{\partial t\,\partial t'}\right] = \frac{\partial^2}{\partial t\,\partial t'} M[\mathring{X}(t)\,\mathring{X}(t')] = \frac{\partial^2}{\partial t\,\partial t'} K_x(t, t'). \quad (15.7.19)$$

Таким образом,

$$K_y(t, t') = \frac{\partial^2 K_x(t, t')}{\partial t\,\partial t'}. \quad (15.7.20)$$

Итак, чтобы найти корреляционную функцию производной, нужно дважды продифференцировать корреляционную функцию исходной случайной функции: сначала по одному аргументу, затем — по другому.

Сравнивая правила нахождения математического ожидания и корреляционной функции для двух рассмотренных нами линейных однородных операторов, мы видим, что они совершенно аналогичны, а именно: для нахождения математического ожидания преобразованной случайной функции тот же линейный оператор применяется к математическому ожиданию исходной случайной функции; для нахождения корреляционной функции тот же линейный оператор применяется дважды к корреляционной функции исходной случайной функции. В первом частном случае это было двойное интегрирование, во втором — двойное дифференцирование.

Можно доказать, что такое правило является общим для всех линейных однородных операторов[1]). Мы здесь сформулируем это общее правило без доказательства.

Если случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$ преобразуется линейным однородным оператором L в случайную функцию

$$Y(t) = L\{X(t)\},$$

то для нахождения математического ожидания случайной функции $Y(t)$ нужно применить тот же оператор к математическому ожиданию случайной функции $X(t)$:

$$m_y(t) = L\{m_x(t)\}, \quad (15.7.21)$$

а для нахождения корреляционной функции нужно дважды применить тот же оператор к корреляционной функции случайной функции $X(t)$, сначала по одному аргументу, затем — по другому:

$$K_y(t, t') = L^{(t)}L^{(t')}\{K_x(t, t')\}. \quad (15.7.22)$$

В формуле (15.7.22) значки (t), (t') у знака оператора L указывают, по какому аргументу он применяется.

---

[1]) См., например, В. С. Пугачев, Теория случайных функций и ее применение к задачам автоматического управления, Физматгиз, 1960.

Во многих задачах практики нас, в конечном счете, интересует не корреляционная функция $K_y(t, t')$ на выходе линейной системы, а дисперсия $D_y(t)$, характеризующая точность работы системы в условиях наличия случайных возмущений. Дисперсию $D_y(t)$ можно найти, зная корреляционную функцию:

$$D_y(t) = K_y(t, t). \tag{15.7.23}$$

При этом нужно подчеркнуть, что, как правило, для определения дисперсии на выходе линейной системы недостаточно знать дисперсию на ее входе, а существенно важно знать корреляционную функцию. Действительно, линейная система может совершенно по-разному реагировать на случайные возмущения, поступающие на ее вход, в зависимости от того, какова внутренняя структура этих случайных возмущений; состоят ли они, например, по преимуществу из высокочастотных или низкочастотных колебаний. Внутренняя же структура случайного процесса описывается не его дисперсией, а корреляционной фукцией.

Пример. На вход дифференцирующего механизма поступает случайная функция $X(t)$ с математическим ожиданием $m_x(t) = \sin t$ и корреляционной функцией

$$K_x(t, t') = D_x e^{-\alpha(t'-t)^2},$$

где D_x — постоянная дисперсия случайной функции $X(t)$.

Определить математическое ожидание и дисперсию на выходе системы.

Решение. Случайная функция $Y(t)$ на выходе системы (реакция) связана с воздействием $X(t)$ оператором дифференцирования:

$$Y(t) = \frac{d}{dt} X(t).$$

Применяя общие правила, имеем:

$$m_y(t) = \frac{d}{dt} m_x(t) = \cos t;$$

$$K_y(t, t') = \frac{\partial^2}{\partial t\, \partial t'} K_x(t, t') = 2D_x \alpha e^{-\alpha(t'-t)^2} [1 - 2\alpha (t' - t)^2].$$

Полагая $t' = t$, имеем:

$$D_y(t) = 2D_x\alpha,$$

или, замечая, что $D_y(t)$ не зависит от t,

$$D_y = 2D_x\alpha.$$

Итак, дисперсия на выходе дифференцирующего механизма зависит не только от дисперсии D_x на входе, но также и от коэффициента α, характеризующего быстроту затухания корреляционной связи между сечениями случайной функции $X(t)$ при возрастании промежутка между ними. Если коэффициент α мал, корреляционная связь затухает медленно, случайная функция изменяется со временем сравнительно плавно, и, естественно, дифференцирование такой функции приводит к сравнительно малым ошибкам. Напротив, если коэффициент α велик, корреляционная функция убывает быстро; в составе случайной функции преобладают резкие, беспорядочные

высокочастотные колебания; естественно, дифференцирование такой функции приводит к большим случайным ошибкам. В таких случаях обычно прибегают к сглаживанию дифференцируемой функции, т. е. так меняют оператор системы, чтобы он давал меньшие случайные ошибки на выходе.

15.8. Сложение случайных функций

Во многих задачах практики мы встречаемся с тем, что на вход динамической системы поступает не одна случайная функция $X(t)$, а две или более случайные функции, каждая из которых связана с действием отдельного возмущающего фактора. Возникает задача сложения случайных функций, точнее — задача определения характеристик суммы по характеристикам слагаемых.

Эта задача решается элементарно просто, если две складываемые случайные функции независимы (точнее, некоррелированны) между собой. В общем же случае для ее решения необходимо знание еще одной характеристики — так называемой *взаимной корреляционной функции* (иначе — *корреляционной функции связи*).

Взаимной корреляционной функцией двух случайных функций $X(t)$ и $Y(t)$ называется неслучайная функция двух аргументов t и t', которая при каждой паре значений t, t' равна корреляционному моменту соответствующих сечений случайной функции $X(t)$ и случайной функции $Y(t)$:

$$R_{xy}(t, t') = M[\mathring{X}(t)\mathring{Y}(t')]. \qquad (15.8.1)$$

Взаимная корреляционная функция, так же как и обычная корреляционная функция, не изменяется при прибавлении к случайным функциям любых неслучайных слагаемых, а следовательно, и при центрировании случайных функций.

Из определения взаимной корреляционной функции вытекает следующее ее свойство:

$$R_{xy}(t, t') = R_{yx}(t', t). \qquad (15.8.2)$$

Вместо функции $R_{xy}(t, t')$ часто пользуются нормированной взаимной корреляционной функцией:

$$r_{xy}(t, t') = \frac{R_{xy}(t, t')}{\sigma_x(t)\,\sigma_y(t')}, \qquad (15.8.3)$$

Если взаимная корреляционная функция равна нулю при всех значениях t, t', то случайные функции $X(t)$ и $Y(t)$ называются некоррелированными (несвязанными).

На практике обычно суждение о некоррелированности случайных функций составляют не на основании равенства нулю взаимной корреляционной функции, а, наоборот, взаимную корреляционную функцию полагают равной нулю на основании физических соображений, свидетельствующих о независимости случайных функций.

Пусть, например, два самолета обстреливают наземную цель; обозначим $X(t)$, $Y(t)$ углы пикирования первого и второго самолетов в процессе выполнения боевой операции. Если самолеты заходят на цель поодиночке, естественно считать случайные функции $X(t)$ и $Y(t)$ некоррелированными; если же маневр выполняется самолетами совместно, причем один из них является ведущим, а другой — ведомым, очевидно наличие корреляционной связи между функциями $X(t)$ и $Y(t)$.

Вообще, если из физических соображений, связанных с существом решаемой задачи, вытекает наличие зависимости между фигурирующими в задаче случайными функциями, то их взаимные корреляционные функции должны быть обследованы.

Зная математические ожидания и корреляционные функции двух случайных функций $X(t)$ и $Y(t)$, а также их взаимную корреляционную функцию, можно найти характеристики суммы этих двух случайных функций:

$$Z(t) = X(t) + Y(t). \tag{15.8.4}$$

По теореме сложения математических ожиданий:

$$m_z(t) = m_x(t) + m_y(t), \tag{15.8.5}$$

т. е. *при сложении двух случайных функций их математические ожидания складываются.*

Для определения корреляционной функции $K_z(t, t')$ перейдем к центрированным случайным функциям $\mathring{Z}(t)$, $\mathring{X}(t)$, $\mathring{Y}(t)$. Очевидно,

$$\mathring{Z}(t) = \mathring{X}(t) + \mathring{Y}(t). \tag{15.8.6}$$

По определению корреляционной функции

$$K_z(t, t') = M[\mathring{Z}(t)\,\mathring{Z}(t')] =$$

$$= M[(\mathring{X}(t) + \mathring{Y}(t))(\mathring{X}(t') + \mathring{Y}(t'))] =$$

$$= M[\mathring{X}(t)\,\mathring{X}(t')] + M[\mathring{Y}(t)\,\mathring{Y}(t')] + M[\mathring{X}(t)\,\mathring{Y}(t')] + M[\mathring{X}(t')\,\mathring{Y}(t)],$$

или

$$K_z(t, t') = K_x(t, t') + K_y(t, t') + R_{xy}(t, t') + R_{xy}(t', t). \tag{15.8.7}$$

Формула (15.8.7) аналогична формуле (10.2.7) для дисперсии суммы случайных величин.

В случае, когда случайные функции $X(t)$ и $Y(t)$ некоррелированны, $R_{xy}(t, t') \equiv 0$, и формула (15.8.7) принимает вид:

$$K_z(t, t') = K_x(t, t') + K_y(t, t'), \tag{15.8.8}$$

т. е. *при сложении некоррелированных случайных функций их корреляционные функции складываются.*

Выведенные формулы могут быть обобщены на случай произвольного числа слагаемых. Если случайная функция $X(t)$ есть сумма n

случайных функций:

$$X(t) = \sum_{i=1}^{n} X_i(t), \tag{15.8.9}$$

то ее математическое ожидание выражается формулой

$$m_x(t) = \sum_{i=1}^{n} m_{x_i}(t), \tag{15.8.10}$$

а ее корреляционная функция — формулой

$$K_x(t, t') = \sum_{i=1}^{n} K_{x_i}(t, t') + \sum_{i \neq j} R_{x_i x_j}(t, t'), \tag{15.8.11}$$

где суммирование распространяется на все возможные размещения индексов i и j попарно.

В случае, когда все случайные функции $X_i(t)$ некоррелированны, формула (15.8.11) превращается в *теорему сложения корреляционных функций*:

$$K_x(t, t') = \sum_{i=1}^{n} K_{x_i}(t, t'), \tag{15.8.12}$$

т. е. *корреляционная функция суммы взаимно некоррелированных случайных функций равна сумме корреляционных функций слагаемых.*

Формула (15.8.12) аналогична теореме сложения дисперсий для обычных случайных величин.

Частным случаем сложения случайных функций является сложение случайной функции со случайной величиной.

Рассмотрим случайную функцию $X(t)$ с характеристиками $m_x(t)$ и $K_x(t, t')$ и случайную величину Y с математическим ожиданием m_y и дисперсией D_y. Предположим, что случайная функция $X(t)$ и случайная величина Y некоррелированны, т. е. при любом t

$$M[\mathring{X}(t)\mathring{Y}] = 0.$$

Сложим случайную функцию $X(t)$ со случайной величиной Y; получим случайную функцию

$$Z(t) = X(t) + Y. \tag{15.8.13}$$

Определим ее характеристики. Очевидно,

$$m_z(t) = m_x(t) + m_y. \tag{15.8.14}$$

Чтобы найти $K_z(t, t')$, пользуясь теоремой сложения корреляционных функций (15.8.8), рассмотрим случайную величину Y как частный случай случайной функции, не меняющейся во времени, и найдем ее корреляционную функцию:

$$K_y(t, t') = M[\mathring{Y}(t)\mathring{Y}(t')] = M[\mathring{Y}^2] = D_y. \tag{15.8.15}$$

Применяя формулу (15.8.8), получим:

$$K_z(t, t') = K_x(t, t') + D_y,$$

т. е. *при прибавлении к случайной функции некоррелированной с нею случайной величины к корреляционной функции прибавляется постоянное слагаемое, равное дисперсии этой случайной величины.*

15.9. Комплексные случайные функции

При практическом применении математического аппарата теории случайных функций часто оказывается удобным записывать как сами случайные функции, так и их характеристики не в действительной, а в комплексной форме. В связи с этим необходимо дать определение *комплексной случайной величины* и *комплексной случайной функции.*

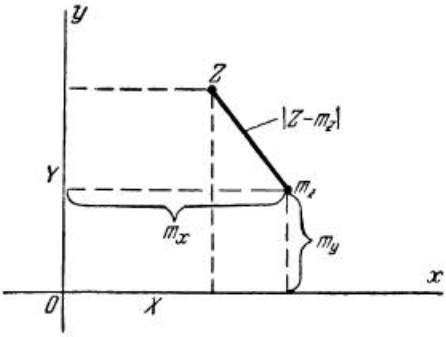


Рис. 15.9.1.

Комплексной случайной величиной называется случайная величина вида:

$$Z = X + iY, \quad (15.9.1)$$

где X, Y — действительные случайные величины; $i = \sqrt{-1}$ — мнимая единица.

Комплексную случайную величину можно геометрически интерпретировать как случайную точку Z на плоскости xOy (рис. 15.9.1).

Для того чтобы аппарат числовых характеристик был применим и к комплексным случайным величинам, необходимо обобщить основные понятия математического ожидания, дисперсии и корреляционного момента на случай комплексных случайных величин. Очевидно, эти обобщения должны быть сделаны так, чтобы в частном случае, когда $Y = 0$ и величина Z действительна, они сводились к обычным определениям характеристик действительных случайных величин.

Математическим ожиданием комплексной случайной величины $Z = X + iY$ называется комплексное число

$$m_z = m_x + im_y. \quad (15.9.2)$$

Это есть некоторое среднее значение величины Z или, геометрически, средняя точка m_z, вокруг которой происходит рассеивание случайной точки Z (рис. 15.9.1).

Дисперсией комплексной случайной величины называется математическое ожидание квадрата модуля соответствующей центрированной величины:

$$D_z = M[|\mathring{Z}|^2], \tag{15.9.3}$$

где

$$\mathring{Z} = Z - m_z.$$

Геометрически дисперсия комплексной случайной величины есть не что иное, как среднее значение квадрата расстояния от случайной точки Z до ее математического ожидания m_z (рис. 15.9.1). Эта величина характеризует разброс случайной точки Z около ее среднего положения.

Выразим дисперсию комплексной случайной величины через дисперсии ее действительной и мнимой частей. Очевидно,

$$\mathring{Z} = Z - m_z = X + iY - m_x - im_y = \mathring{X} + i\mathring{Y};$$

отсюда

$$D_z = M[|\mathring{Z}|^2] = M[\mathring{X}^2 + \mathring{Y}^2] = M[\mathring{X}^2] + M[\mathring{Y}^2]$$

или

$$D_z = D_x + D_y, \tag{15.9.4}$$

т. е. *дисперсия комплексной случайной величины равна сумме дисперсий ее действительной и мнимой частей.*

Из самого определения дисперсии следует, что дисперсия комплексной случайной величины всегда действительна и существенно положительна. Обращаться в нуль она может только в случае, если величина Z не случайна.

Данные выше определения математического ожидания и дисперсии, очевидно, удовлетворяют поставленному требованию: при $Y = 0$ и $Z = X$ они превращаются в обычные определения математического ожидания и дисперсии действительной случайной величины.

Попытаемся построить аналогичное определение корреляционного момента двух комплексных случайных величин Z_1 и Z_2:

$$Z_1 = X_1 + iY_1; \quad Z_2 = X_2 + iY_2. \tag{15.9.5}$$

Это определение, очевидно, должно быть построено так, чтобы при $Z_1 = Z_2 = Z$ корреляционный момент обращался в дисперсию величины Z. Оказывается, этому требованию нельзя было бы удовлетворить, если бы мы, как в случае действительных величин, назвали корреляционным моментом математическое ожидание произведения $\mathring{Z}_1\mathring{Z}_2$. Нетрудно убедиться, что при $Z_1 = Z_2 = Z$ математическое ожидание такого произведения будет не действительным, а комплексным, т. е. уже не дает дисперсии, которая, согласно определению, действительна и существенно положительна. Этого не будет, если назвать корреляционным моментом математическое

ожидание произведения $\mathring{Z}_1$ не на самую величину $\mathring{Z}_2$, а на соответствующую ей *комплексную сопряженную* величину:

$$\overline{\mathring{Z}_2} = \mathring{X}_2 - i\mathring{Y}_2. \tag{15.9.6}$$

Тогда при $Z_1 = Z_2 = Z$ корреляционный момент, очевидно, обратится в дисперсию величины Z:

$$K_{zz} = M[(\mathring{X} + i\mathring{Y})(\mathring{X} - i\mathring{Y})] = M[\mathring{X}^2] + M[\mathring{Y}^2] = D_z. \tag{15.9.7}$$

Таким образом, целесообразно дать следующее определение корреляционного момента двух комплексных случайных величин Z_1 и Z_2:

$$K_{z_1 z_2} = M[\mathring{Z}_1 \overline{\mathring{Z}_2}], \tag{15.9.8}$$

где чертой наверху обозначается комплексная сопряженная величина.

Выразим корреляционный момент двух комплексных случайных величин через корреляционные моменты их действительных и мнимых частей. Имеем:

$$K_{z_1 z_2} = M[\mathring{Z}_1 \overline{\mathring{Z}_2}] = M[(\mathring{X}_1 + i\mathring{Y}_1)(\mathring{X}_2 - i\mathring{Y}_2)] = \\ = K_{x_1 x_2} + K_{y_1 y_2} + i(K_{y_1 x_2} - K_{x_1 y_2}), \tag{15.9.9}$$

где $K_{x_1 x_2}$, $K_{y_1 y_2}$, $K_{y_1 x_2}$, $K_{x_1 y_2}$ — соответственно корреляционные моменты величин (X_1, X_2), (Y_1, Y_2); (Y_1, X_2), (X_1, Y_2).

Очевидно, в случае, когда все эти величины между собой не коррелированы, корреляционный момент величин Z_1, Z_2 также равен нулю.

Таким образом, определения основных характеристик комплексных случайных величин отличаются от обычных определений аналогичных характеристик для действительных величин только тем, что:

1) в качестве дисперсии рассматривается не математическое ожидание квадрата центрированной случайной величины, а математическое ожидание *квадрата ее модуля*;

2) в качестве корреляционного момента рассматривается не математическое ожидание произведения центрированных величин, а математическое ожидание произведения одной центрированной величины на *комплексную сопряженную* другой.

Перейдем к определению комплексной случайной функции и ее характеристик.

Комплексной случайной функцией называется функция вида:

$$Z(t) = X(t) + iY(t), \tag{15.9.10}$$

где $X(t)$, $Y(t)$ — действительные случайные функции.

Математическое ожидание комплексной случайной функции (15.9.10) равно:

$$m_z(t) = m_x(t) + im_y(t). \tag{15.9.11}$$

Дисперсия комплексной случайной функции $Z(t)$ определяется как математическое ожидание квадрата модуля соответствующей центрированной функции:

$$D_z(t) = M[|\mathring{Z}(t)|^2], \tag{15.9.12}$$

где

$$\mathring{Z}(t) = Z(t) - m_z(t) = \mathring{X}(t) + i\mathring{Y}(t). \tag{15.9.13}$$

Из определения (15.9.12) видно, что дисперсия комплексной случайной функции действительна и неотрицательна.

Из формулы (15.9.4) следует, что дисперсия комплексной случайной функции равна сумме дисперсий ее действительной и мнимой частей:

$$D_z(t) = D_x(t) + D_y(t). \tag{15.9.14}$$

Корреляционная функция комплексной случайной функции определяется как корреляционный момент ее сечений t и t':

$$K_z(t, t') = M[\mathring{Z}(t)\,\overline{\mathring{Z}}(t')], \tag{15.9.15}$$

где

$$\overline{\mathring{Z}}(t') = \mathring{X}(t') - i\mathring{Y}(t')$$

— комплексная величина, сопряженная величине $\mathring{Z}(t')$.

При $t' = t$ корреляционная функция, очевидно, обращается в дисперсию:

$$K_z(t, t) = D_z(t). \tag{15.9.16}$$

Пользуясь формулой (15.9.9), можно выразить корреляционную функцию комплексной случайной функции через характеристики ее действительной и мнимой частей. Рассматривая в качестве случайных величин Z_1 и Z_2, фигурирующих в этой формуле, сечения случайной функции $Z(t)$ и $Z(t')$, получим:

$$K_z(t, t') = K_x(t, t') + K_y(t, t') + i[R_{xy}(t', t) - R_{xy}(t, t')], \tag{15.9.17}$$

где $R_{xy}(t, t')$ — взаимная корреляционная функция случайных функций $X(t)$ и $Y(t)$ (действительной и мнимой частей случайной функции $Z(t)$).

В случае, когда действительная и мнимая части случайной функции не коррелированы ($R_{xy}(t, t') \equiv 0$), формула (15.9.17) имеет вид:

$$K_z(t, t') = K_x(t, t') + K_y(t, t'). \tag{15.9.18}$$

В дальнейшем изложении мы будем пользоваться как действительной, так и комплексной формой записи случайных функций. В последнем случае мы будем всегда оговаривать это.

ГЛАВА 16

КАНОНИЧЕСКИЕ РАЗЛОЖЕНИЯ СЛУЧАЙНЫХ ФУНКЦИЙ

16.1. Идея метода канонических разложений. Представление случайной функции в виде суммы элементарных случайных функций

В n° 15.7 мы познакомились с общими правилами линейных преобразований случайных функций. Эти правила сводятся к тому, что при линейном преобразовании случайной функции ее математическое ожидание подвергается тому же линейному преобразованию, а корреляционная функция подвергается этому преобразованию дважды: по одному и другому аргументу.

Правило преобразования математического ожидания очень просто и при практическом применении затруднений не вызывает. Что касается двойного преобразования корреляционной функции, то оно в ряде случаев приводит к чрезвычайно сложным и громоздким операциям, что затрудняет практическое применение изложенных общих методов.

Действительно, рассмотрим, например, простейший интегральный оператор:

$$Y(t) = \int_0^t X(\tau)\, d\tau. \tag{16.1.1}$$

Согласно общему правилу корреляционная функция преобразуется тем же оператором дважды:

$$K_y(t,\ t') = \int_0^t \int_0^{t'} K_x(\tau,\ \tau')\, d\tau\, d\tau'. \tag{16.1.2}$$

Очень часто бывает, что полученная из опыта корреляционная функция $K_x(t,\ t')$ не имеет аналитического выражения и задана таблично; тогда интеграл (16.1.2) приходится вычислять численно, определяя его как функцию обоих пределов. Это — задача очень

громоздкая и трудоемкая. Если даже аппроксимировать подынтегральную функцию каким-либо аналитическим выражением, то и в этом случае чаще всего интеграл (16.1.2) через известные функции не выражается. Так обстоит дело даже при простейшей форме оператора преобразования. Если же, как часто бывает, работа динамической системы описывается дифференциальными уравнениями, решение которых не выражается в явной форме, задача об определении корреляционной функции на выходе еще более осложняется: она требует интегрирования дифференциальных уравнений с частными производными.

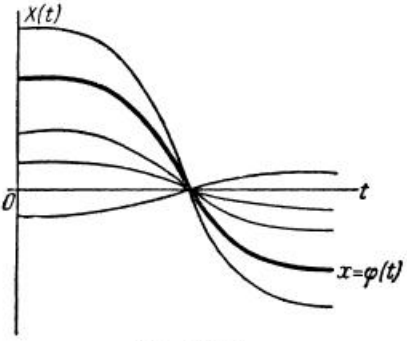


Рис. 16.1.1.

В связи с этим на практике применение изложенных общих методов линейных преобразований случайных функций, как правило, оказывается слишком сложным и себя не оправдывает. При решении практических задач значительно чаще применяются другие методы, приводящие к более простым преобразованиям. Один из них — так называемый *метод канонических разложений*, разработанный В. С. Пугачевым, и составляет содержание данной главы.

Идея метода канонических разложений состоит в том, что случайная функция, над которой нужно произвести те или иные преобразования, предварительно представляется в виде суммы так называемых *элементарных случайных функций*.

Элементарной случайной функцией называется функция вида:

$$X(t) = V\varphi(t), \tag{16.1.3}$$

где V — обычная случайная величина, $\varphi(t)$ — обычная (неслучайная) функция.

Элементарная случайная функция является наиболее простым типом случайной функции. Действительно, в выражении (16.1.3) случайным является только множитель V, стоящий перед функцией $\varphi(t)$; сама же зависимость от времени случайной не является.

Все возможные реализации элементарной случайной функции $X(t)$ могут быть получены из графика функции $x = \varphi(t)$ простым изменением масштаба по оси ординат (рис. 16.1.1). При этом ось абсцисс ($x = 0$) также представляет собой одну из возможных реализаций случайной функции $X(t)$, осуществляющуюся, когда случайная величина V принимает значение 0 (если это значение принадлежит к числу возможных значений величины V).

В качестве примеров элементарных случайных функций приведем функции $X(t) = V \sin t$ (рис. 16.1.2) и $X(t) = Vt^2$ (рис. 16.1.3).

Элементарная случайная функция характерна тем, что в ней разделены две особенности случайной функции: случайность вся сосредоточена в коэффициенте V, а зависимость от времени — в обычной функции $\varphi(t)$.

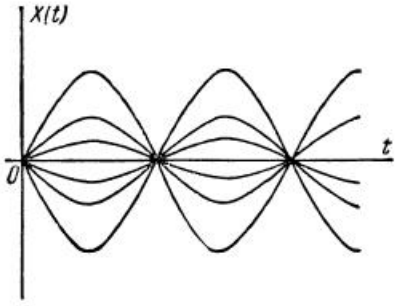


Рис. 16.1.2.

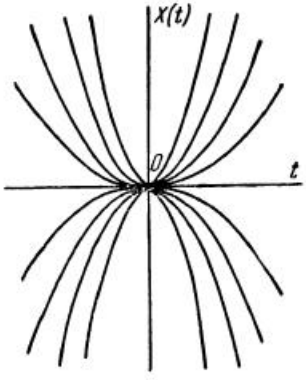


Рис. 16.1.3.

Определим характеристики элементарной случайной функции (16.1.3). Имеем:

$$m_x(t) = M[V\varphi(t)] = m_v \varphi(t),$$

где m_v — математическое ожидание случайной величины V.

Если $m_v = 0$, математическое ожидание случайной функции $X(t)$ также равно нулю, причем тождественно:

$$m_x(t) \equiv 0.$$

Мы знаем, что любую случайную функцию можно центрировать, т. е. привести к такому виду, когда ее математическое ожидание равно нулю. Поэтому в дальнейшем мы будем рассматривать только центрированные элементарные случайные функции, для которых $m_v = 0$; $V = \mathring{V}$; $m_x(t) \equiv 0$.

Определим корреляционную функцию элементарной случайной функции $X(t)$. Имеем:

$$K_x(t, t') = M[X(t)\,X(t')] = \varphi(t)\,\varphi(t')\,M[V^2] = \varphi(t)\,\varphi(t')\,D,$$

где D — дисперсия величины V.

Над элементарными случайными функциями весьма просто выполняются всевозможные линейные преобразования.

Например, продифференцируем случайную функцию (16.1.3). Случайная величина V, не зависящая от t, выйдет за знак производной, и мы получим:

$$X'(t) = V\varphi'(t).$$

Аналогично

$$\int_0^t X(\tau)\,d\tau = V\int_0^t \varphi(\tau)\,d\tau.$$

Вообще, если элементарная случайная функция (16.1.3) преобразуется линейным оператором L, то при этом случайный множитель V, как не зависящий от t, выходит за знак оператора, а неслучайная функция $\varphi(t)$ преобразуется тем же оператором L:

$$L\{X(t)\} = VL\{\varphi(t)\}. \qquad (16.1.4)$$

Значит, если элементарная случайная функция поступает на вход линейной системы, то задача ее преобразования сводится к простой задаче преобразования одной неслучайной функции $\varphi(t)$. Отсюда возникает идея: если на вход динамической системы поступает некоторая случайная функция общего вида, то можно ее представить — точно или приближенно — в виде суммы элементарных случайных функций и только затем подвергать преобразованию. Такая идея разложения случайной функции на сумму элементарных случайных функций и лежит в основе метода канонических разложений.

Пусть имеется случайная функция:

$$X(t) = m_x(t) + \mathring{X}(t). \qquad (16.1.5)$$

Допустим, что нам удалось — точно или приближенно — представить ее в виде суммы

$$X(t) = m_x(t) + \sum_{l=1}^{m} V_l \varphi_l(t), \qquad (16.1.6)$$

где V_l — случайные величины с математическими ожиданиями, равными нулю, $\varphi_l(t)$ — неслучайные функции, $m_x(t)$ — математическое ожидание функции $X(t)$.

Условимся называть представление случайной функции в форме (16.1.6) ***разложением случайной функции***. Случайные величины $V_1, V_2, \ldots, V_m$ будем называть ***коэффициентами*** разложения, а неслучайные функции $\varphi_1(t), \varphi_2(t), \ldots, \varphi_m(t)$ — ***координатными функциями***.

Определим реакцию линейной системы с оператором L на случайную функцию $X(t)$, заданную в виде разложения (16.1.6). Известно, что линейная система обладает так называемым ***свойством суперпозиции***, состоящим в том, что реакция системы на сумму нескольких воздействий равна сумме реакций системы на каждое отдельное воздействие. Действительно, оператор системы L, будучи линейным, может, по определению, применяться к сумме почленно.

Обозначая $Y(t)$ реакцию системы на случайное воздействие $X(t)$, имеем:

$$Y(t) = L\{X(t)\} = L\{m_x(t)\} + \sum_{l=1}^{m} V_l L\{\varphi_l(t)\}. \qquad (16.1.7)$$

Придадим выражению (16.1.7) несколько иную форму. Учитывая общее правило линейного преобразования математического ожидания, убеждаемся, что

$$L[m_x(t)] = m_y(t).$$

Обозначая

$$L[\varphi_i(t)] = \psi_i(t),$$

имеем:

$$Y(t) = m_y(t) + \sum_{i=1}^{m} V_i \psi_i(t). \qquad (16.1.8)$$

Выражение (16.1.8) представляет собой не что иное, как *разложение* случайной функции $Y(t)$ по элементарным функциям. Коэффициентами этого разложения являются те же случайные величины $V_1, V_2, \ldots, V_m$, а математическое ожидание и координатные функции получены из математического ожидания и координатных функций исходной случайной функции тем же линейным преобразованием L, какому подвергается случайная функция $X(t)$.

Резюмируя, получаем следующее правило преобразования случайной функции, заданной разложением.

Если случайная функция $X(t)$, заданная разложением по элементарным функциям, подвергается линейному преобразованию L, то коэффициенты разложения остаются неизменными, а математическое ожидание и координатные функции подвергаются тому же линейному преобразованию L.

Таким образом, смысл представления случайной функции в виде разложения сводится к тому, чтобы свести линейное преобразование случайной функции к таким же линейным преобразованиям нескольких неслучайных функций — математического ожидания и координатных функций. Это позволяет значительно упростить решение задачи нахождения характеристик случайной функции $Y(t)$ по сравнению с общим решением, данным в п° 15.7. Действительно, каждая из неслучайных функций $m_x(t)$, $\varphi_1(t)$, $\varphi_2(t)$, ..., $\varphi_m(t)$ в данном случае преобразуется только один раз в отличие от корреляционной функции $K_x(t, t')$, которая, согласно общим правилам, преобразуется дважды.

16.2. Каноническое разложение случайной функции

Рассмотрим случайную функцию $X(t)$, заданную разложением

$$X(t) = m_x(t) + \sum_{i=1}^{m} V_i \varphi_i(t), \qquad (16.2.1)$$

где коэффициенты $V_1, V_2, \ldots, V_m$ представляют собой систему случайных величин с математическими ожиданиями, равными нулю и с корреляционной матрицей $\|K_{ij}\|$.

Найдем корреляционную функцию и дисперсию случайной функции $X(t)$.

По определению

$$K_x(t,\ t') = M[\mathring{X}(t)\,\mathring{X}(t')], \tag{16.2.2}$$

где

$$\mathring{X}(t) = \sum_{i=1}^{m} V_i \varphi_i(t), \tag{16.2.3}$$

$$\mathring{X}(t') = \sum_{j=1}^{m} V_j \varphi_j(t'). \tag{16.2.4}$$

В формуле (16.2.4) индекс суммирования обозначен буквой j, чтобы подчеркнуть его независимость от индекса суммирования i в формуле (16.2.3).

Перемножая выражения (16.2.3) и (16.2.4) и применяя к произведению операцию математического ожидания, получим:

$$K_x(t,\ t') = M\Big[\sum_{i,\ j} V_i V_j \varphi_i(t)\,\varphi_j(t')\Big] = \sum_{i,\ j} M[V_i V_j]\,\varphi_i(t)\,\varphi_j(t'), \tag{16.2.5}$$

где суммирование распространяется на все пары значений i, j — как равные, так и неравные. В случае, когда $i = j$,

$$M[V_i V_j] = M[V_i^2] = K_{ii} = D_i,$$

где D_i — дисперсия случайной величины V_i. В случае, когда $i \neq j$,

$$M[V_i V_j] = K_{ij},$$

где K_{ij} — корреляционный момент случайных величин V_i, V_j.

Подставляя эти значения в формулу (16.2.5), получим выражение для корреляционной функции случайной функции $X(t)$, заданной разложением (16.2.1):

$$K_x(t,\ t') = \sum_{i=1}^{m} \varphi_i(t)\,\varphi_i(t')\,D_i + \sum_{i \neq j} \varphi_i(t)\,\varphi_j(t')\,K_{ij}. \tag{16.2.6}$$

Полагая в выражении (16.2.6) $t' = t$, получим дисперсию случайной функции $X(t)$:

$$D_x(t) = \sum_{i=1}^{m} [\varphi_i(t)]^2 D_i + \sum_{i \neq j} \varphi_i(t)\,\varphi_j(t)\,K_{ij}. \tag{16.2.7}$$

Очевидно, выражения (16.2.6) и (16.2.7) приобретают особенно простой вид, когда все коэффициенты V_i разложения (16.2.1) некоррелированны, т. е. $K_{ij}=0$ при $i \neq j$. В этом случае разложение случайной функции называется «каноническим».

Таким образом, *каноническим разложением* случайной функции $X(t)$ называется ее представление в виде:

$$X(t)=m_x(t)+\sum_{i=1}^{m} V_i \varphi_i(t), \tag{16.2.8}$$

где $m_x(t)$ — математическое ожидание случайной функции; $\varphi_1(t)$, $\varphi_2(t)$, ..., $\varphi_m(t)$ — координатные функции, а V_1, V_2, ..., V_m — некоррелированные случайные величины с математическими ожиданиями, равными нулю.

Если задано каноническое разложение случайной функции, то ее корреляционная функция $K_x(t, t')$ выражается весьма просто. Полагая в формуле (16.2.6) $K_{ij}=0$ при $i \neq j$, получим:

$$K_x(t, t')=\sum_{i=1}^{m} \varphi_i(t)\, \varphi_i(t')\, D_i. \tag{16.2.9}$$

Выражение (16.2.9) называется *каноническим разложением корреляционной функции*.

Полагая в формуле (16.2.9) $t'=t$, получим дисперсию случайной функции

$$D_x(t)=\sum_{i=1}^{m} [\varphi_i(t)]^2 D_i. \tag{16.2.10}$$

Таким образом, зная каноническое разложение случайной функции $X(t)$, можно сразу найти каноническое разложение ее корреляционной функции. Можно доказать, что обратное положение тоже справедливо, а именно: если задано каноническое разложение корреляционной функции (16.2.9), то для случайной функции $X(t)$ справедливо каноническое разложение вида (16.2.8) с координатными функциями $\varphi_i(t)$ и коэффициентами V_i с дисперсиями D_i. Мы примем это положение без специального доказательства [1]).

Число членов канонического разложения случайной функции может быть не только конечным, но и бесконечным. Примеры канонических разложений с бесконечным числом членов встретятся нам в главе 17. Кроме того, в ряде случаев применяются так называемые *интегральные канонические представления* случайных функций, в которых сумма заменяется интегралом.

---

[1]) Доказательство см.: В. С. Пугачев, Теория случайных функций и ее применение в задачах автоматического управления, Физматгиз, 1962.

Канонические разложения применяются не только для действительных, но и для комплексных случайных функций. Рассмотрим обобщение понятия канонического разложения на случай комплексной случайной функции.

Элементарной комплексной случайной функцией называется функция вида:

$$X(t) = V\varphi(t), \tag{16.2.11}$$

где как случайная величина V, так и функция $\varphi(t)$ комплексны.

Определим корреляционную функцию элементарной случайной функции (16.2.11). Пользуясь общим определением корреляционной функции комплексной случайной функции, имеем:

$$K_x(t, t') = M[V\varphi(t)\overline{V\varphi(t')}], \tag{16.2.12}$$

где чертой вверху, как и ранее, обозначена комплексная сопряженная величина. Имея в виду, что

$$\overline{V\varphi(t)} = \overline{V}\,\overline{\varphi(t)},$$

и вынося неслучайные величины $\varphi(t)$ и $\overline{\varphi(t')}$ за знак математического ожидания, получим:

$$K_x(t, t') = \varphi(t)\overline{\varphi(t')}\,M[|V|^2].$$

Но, согласно п° 15.9, $M[|V|^2]$ есть не что иное, как дисперсия комплексной случайной величины V:

$$M[|V|^2] = D,$$

следовательно,

$$K_x(t, t') = \varphi(t)\overline{\varphi(t')}\,D. \tag{16.2.13}$$

Каноническим разложением комплексной случайной функции называется ее представление в виде:

$$X(t) = m_x(t) + \sum_{l=1}^{m} V_l\varphi_l(t), \tag{16.2.14}$$

где $V_1, V_2, \ldots, V_m$ — некоррелированные комплексные случайные величины с математическими ожиданиями, равными нулю, а $m_x(t)$, $\varphi_1(t), \varphi_2(t), \ldots, \varphi_m(t)$ — комплексные неслучайные функции.

Если комплексная случайная функция представлена каноническим разложением (16.2.14), то ее корреляционная функция выражается формулой

$$K_x(t, t') = \sum_{l=1}^{m} \varphi_l(t)\overline{\varphi_l(t')}\,D_l, \tag{16.2.15}$$

где D_l — дисперсия величины V_l:

$$D_l = M[|V_l|^2]. \tag{16.2.16}$$

Формула (16.2.15) непосредственно следует из выражения (16.2.13) для корреляционной функции элементарной комплексной случайной функции.

Выражение (16.2.15) называется каноническим разложением корреляционной функции комплексной случайной функции.

Полагая в (16.2.15) $t' = t$, получим выражение для дисперсии комплексной случайной функции, заданной разложением (16.2.14):

$$D_x(t) = \sum_{i=1}^{m} |\varphi_i(t)|^2 D_i. \tag{16.2.17}$$

16.3. Линейные преобразования случайных функций, заданных каноническими разложениями

Пусть на вход некоторой линейной системы L поступает случайная функция $X(t)$ (рис. 16.3.1).

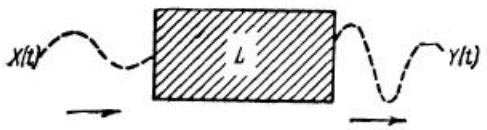


Рис. 16.3.1.

Система преобразует функцию $X(t)$ посредством линейного оператора L, и на выходе мы получаем случайную функцию

$$Y(t) = L\{X(t)\}. \tag{16.3.1}$$

Предположим, что случайная функция $X(t)$ задана ее каноническим разложением:

$$X(t) = m_x(t) + \sum_{i=1}^{k} V_i \varphi_i(t). \tag{16.3.2}$$

Определим реакцию системы на это воздействие. Так как оператор системы является линейным, то

$$Y(t) = L\{X(t)\} = L\{m_x(t)\} + \sum_{i=1}^{k} V_i L\{\varphi_i(t)\}. \tag{16.3.3}$$

Рассматривая выражение (16.3.3), легко убедиться, что оно представляет собой не что иное, как *каноническое разложение случайной функции* $Y(t)$ с математическим ожиданием

$$m_y(t) = L\{m_x(t)\} \tag{16.3.4}$$

и координатными функциями

$$\psi_i(t) = L[\varphi_i(t)]. \tag{16.3.5}$$

Таким образом, *при линейном преобразовании канонического разложения случайной функции $X(t)$ получается каноническое разложение случайной функции $Y(t)$, причем математическое ожидание и координатные функции подвергаются тому же линейному преобразованию.*

Если случайная функция $Y(t)$ на выходе линейной системы получена в виде канонического разложения

$$Y(t) = m_y(t) + \sum_{i=1}^{k} V_i \psi_i(t), \tag{16.3.6}$$

то ее корреляционная функция и дисперсия находятся весьма просто:

$$K_y(t, t') = \sum_{i=1}^{k} \psi_i(t)\,\psi_i(t')\,D_i, \tag{16.3.7}$$

$$D_y(t) = \sum_{i=1}^{k} [\psi_i(t)]^2 D_i. \tag{16.3.8}$$

Это делает особенно удобными именно каноническ ие разложения по сравнению с любыми другими разложениями по элементарным функциям.

Рассмотрим несколько подробнее применение метода канонических разложений к определению реакции динамической системы на случайное входное воздействие, когда работа системы описывается линейным дифференциальным уравнением, в общем случае — с переменными коэффициентами. Запишем это уравнение в операторной форме:

$$A_n(p, t)\,Y(t) = B_m(p, t)\,X(t). \tag{16.3.9}$$

Согласно вышеизложенным правилам линейных преобразований случайных функций математические ожидания воздействия и реакции должны удовлетворять тому же уравнению:

$$A_n(p, t)\,m_y(t) = B_m(p, t)\,m_x(t). \tag{16.3.10}$$

Аналогично каждая из координатных функций должна удовлетворять тому же дифференциальному уравнению:

$$A_n(p, t)\,\psi_i(t) = B_m(p, t)\,\varphi_i(t), \tag{16.3.11}$$
$$(i = 1, 2, \ldots, k).$$

Таким образом, задача определения реакции линейной динамической системы на случайное воздействие свелась к обычной математической задаче интегрирования $k+1$ обыкновенных дифференциальных

уравнений, содержащих обычные, не случайные функции. Так как при решении основной задачи анализа динамической системы — определения ее реакции на заданное воздействие — задача интегрирования дифференциального уравнения, описывающего работу системы, тем или другим способом решается, то при решении уравнений (16.3.10) и (16.3.11) новых математических трудностей не возникает. В частности, для решения этих уравнений могут быть с успехом применены те же интегрирующие системы или моделирующие устройства, которые применяются для анализа работы системы без случайных возмущений.

Остается осветить вопрос о начальных условиях, при которых следует интегрировать уравнения (16.3.10) и (16.3.11).

Сначала рассмотрим наиболее простой случай, когда начальные условия для данной динамической системы являются неслучайными. В этом случае при $t=0$ должны выполняться условия:

$$\left.\begin{aligned} Y(0) &= y_0, \\ Y'(0) &= y_1, \\ &\cdots\cdots\cdots \\ Y^{(r)}(0) &= y_r, \\ &\cdots\cdots\cdots \\ Y^{(n-1)}(0) &= y_{n-1}. \end{aligned}\right\} \tag{16.3.12}$$

где $y_0, y_1, \ldots, y_{n-1}$ — неслучайные числа.

Условия (16.3.12) можно записать более компактно:

$$Y^{(r)}(0) = y_r \quad (r = 0, 1, \ldots, n-1), \tag{16.3.13}$$

понимая при этом под «производной нулевого порядка» $Y^{(0)}(t)$ саму функцию $Y(t)$.

Выясним, при каких начальных условиях должны интегрироваться уравнения (16.3.10) и (16.3.11). Для этого найдем r-ю производную функции $Y(t)$ и положим в ней $t=0$:

$$Y^{(r)}(0) = m_y^{(r)}(0) + \sum_{i=1}^{k} V_i \psi_i^{(r)}(0).$$

Учитывая (16.3.12), имеем:

$$m_y^{(r)}(0) + \sum_{i=1}^{k} V_i \psi_i^{(r)}(0) = y_r. \tag{16.3.14}$$

Так как величина y_r не случайна, то дисперсия левой части равенства (16.3.14) должна быть равна нулю:

$$\sum_{i=1}^{k} D_i \left[\psi_i^{(r)}(0)\right]^2 = 0. \tag{16.3.15}$$

Так как все дисперсии D_l величин V_l положительны, то равенство (16.3.15) может осуществиться только, когда

$$\psi_l^{(r)}(0) = 0 \tag{16.3.16}$$

для всех l.

Подставляя $\psi_l^{(r)}(0) = 0$ в формулу (16.3.14), получим:

$$m_y^{(r)}(0) = y_r. \tag{16.3.17}$$

Из равенства (16.3.17) следует, что уравнение (16.3.10) для математического ожидания должно интегрироваться при заданных начальных условиях (16.3.12):

$$\left.\begin{aligned} m_y(0) &= y_0, \\ m_y'(0) &= y_1, \\ &\dots \\ m_y^{(r)}(0) &= y_r, \\ &\dots \\ m_y^{(n)}(0) &= y_n. \end{aligned}\right\} \tag{16.3.18}$$

Что касается уравнений (16.3.11), то они должны интегрироваться при нулевых начальных условиях:

$$\psi_l(0) = \psi_l'(0) = \dots = \psi_l^{(r)}(0) = \dots = \psi_l^{(n)}(0) = 0$$

$$(l = 1, 2, \dots, k). \tag{16.3.19}$$

Рассмотрим более сложный случай, когда начальные условия случайны:

$$\left.\begin{aligned} Y(0) &= Y_0, \\ Y'(0) &= Y_1. \\ &\dots \\ Y^{(r)}(0) &= Y_r, \\ &\dots \\ Y^{(n-1)}(0) &= Y_{n-1}, \end{aligned}\right\} \tag{16.3.20}$$

где Y_0, Y_1, ..., Y_{n-1} — случайные величины.

В этом случае реакция на выходе системы может быть найдена в виде суммы:

$$Y(t) = Y_{\text{I}}(t) + Y_{\text{II}}(t), \tag{16.3.21}$$

где $Y_{\text{I}}(t)$ — решение дифференциального уравнения (16.3.9) при нулевых начальных условиях; $Y_{\text{II}}(t)$ — решение того же дифференциального уравнения, но с нулевой правой частью при заданных начальных условиях (16.3.20). Как известно из теории дифференциальных

уравнений, это решение представляет собой линейную комбинацию начальных условий:

$$Y_{\text{II}}(t) = \sum_{j=0}^{n-1} Y_j f_j(t), \tag{16.3.22}$$

где $f_j(t)$ — неслучайные функции.

Решение $Y_{\text{I}}(t)$ может быть получено изложенным выше методом в форме канонического разложения. Корреляционная функция случайной функции $Y(t) = Y_{\text{I}}(t) + Y_{\text{II}}(t)$ может быть найдена обычными приемами сложения случайных функций (см. п° 15.8).

Следует заметить, что на практике весьма часто встречаются случаи, когда для моментов времени, достаточно удаленных от начала случайного процесса, начальные условия уже не оказывают влияния на его течение: вызванные ими переходные процессы успевают затухнуть. Системы, обладающие таким свойством, называются асимптотически устойчивыми. Если нас интересует реакция асимптотически устойчивой динамической системы на участках времени, достаточно удаленных от начала, то можно ограничиться исследованием решения $Y_{\text{I}}(t)$, полученного при нулевых начальных условиях. Для достаточно удаленных от начального моментов времени это решение будет справедливым и при любых начальных условиях.

ГЛАВА 17

СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ФУНКЦИИ

17.1. Понятие о стационарном случайном процессе

На практике очень часто встречаются случайные процессы, протекающие во времени приблизительно однородно и имеющие вид непрерывных случайных колебаний вокруг некоторого среднего значения, причем ни средняя амплитуда, ни характер этих колебаний не обнаруживают существенных изменений с течением времени. Такие случайные процессы называются *стационарными*.

В качестве примеров стационарных случайных процессов можно привести: 1) колебания самолета на установившемся режиме горизонтального полета; 2) колебания напряжения в электрической осветительной сети; 3) случайные шумы в радиоприемнике; 4) процесс качки корабля и т. п.

Каждый стационарный процесс можно рассматривать как продолжающийся во времени неопределенно долго; при исследовании стационарного процесса в качестве начала отсчета можно выбрать любой момент времени. Исследуя стационарный процесс на любом участке времени, мы должны получить одни и те же его характеристики. Образно выражаясь, стационарный процесс «не имеет ни начала, ни конца».

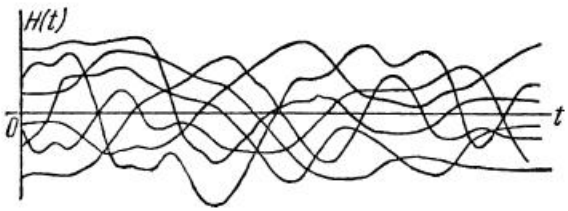


Рис. 17.1.1.

Примером стационарного случайного процесса может служить изменение высоты центра тяжести самолета на установившемся режиме горизонтального полета (рис. 17.1.1).

В противоположность стационарным случайным процессам можно указать другие, явно нестационарные, случайные процессы, например:

колебания самолета в режиме пикирования; процесс затухающих колебаний в электрической цепи; процесс горения порохового заряда в реактивной камере и т. д. Нестационарный процесс характерен тем, что он имеет определенную тенденцию развития во времени; характеристики такого процесса зависят от начала отсчета, зависят от времени.

На рис. 17.1.2 изображено семейство реализаций явно нестационарного случайного процесса — процесса изменения тяги двигателя реактивного снаряда во времени.

Заметим, что далеко не все нестационарные случайные процессы являются существенно нестационарными на всем протяжении своего развития. Существуют нестационарные процессы, которые (на известных отрезках времени и с известным приближением) могут быть приняты за стационарные.

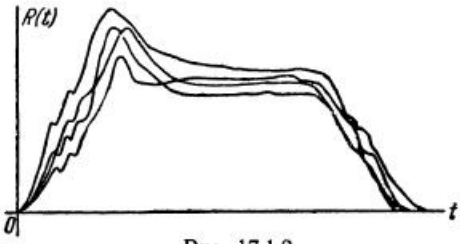


Рис. 17.1.2.

Например, процесс наводки перекрестия авиационного прицела на цель есть явно нестационарный процесс, если цель за короткое время с большой и резко меняющейся угловой скоростью проходит поле зрения прицела. В этом случае колебания оси прицела относительно цели не успевают установиться в некотором стабильном режиме; процесс начинается и заканчивается, не успев приобрести стационарный характер. Напротив, процесс наводки перекрестия прицела на неподвижную или движущуюся с постоянной угловой скоростью цель через некоторое время после начала слежения приобретает стационарный характер.

Вообще, как правило, случайный процесс в любой динамической системе начинается с нестационарной стадии — с так называемого «переходного процесса». После затухания переходного процесса система обычно переходит на установившийся режим, и тогда случайные процессы, протекающие в ней, могут считаться стационарными.

Стационарные случайные процессы очень часто встречаются в физических и технических задачах. По своей природе эти процессы проще, чем нестационарные, и описываются более простыми характеристиками. Линейные преобразования стационарных случайных процессов также обычно осуществляются проще, чем нестационарных. В связи с этим на практике получила широкое применение специальная *теория стационарных случайных процессов*, или, точнее, *теория стационарных случайных функций* (так как аргументом стационарной случайной функции в общем случае может быть и не время). Элементы этой теории и будут изложены в данной главе.

Случайная функция $X(t)$ называется *стационарной*, если все ее вероятностные характеристики не зависят от t (точнее, не меняются при любом сдвиге аргументов, от которых они зависят, по оси t).

В данном элементарном изложении теории случайных функций мы совсем не пользуемся такими вероятностными характеристиками, как законы распределения: единственными характеристиками, которыми мы пользуемся, являются математическое ожидание, дисперсия и корреляционная функция. Сформулируем определение стационарной случайной функции в терминах этих характеристик.

Так как изменение стационарной случайной функции должно протекать однородно по времени, то естественно потребовать, чтобы для стационарной случайной функции математическое ожидание было постоянным:

$$m_x(t) = m_x = \text{const}. \qquad (17.1.1)$$

Заметим, однако, что это требование не является существенным: мы знаем, что от случайной функции $X(t)$ всегда можно перейти к центрированной случайной функции $\mathring{X}(t)$, для которой математическое ожидание тождественно равно нулю и, следовательно, удовлетворяет условию (17.1.1). Таким образом, если случайный процесс нестационарен только за счет переменного математического ожидания, это не помешает нам изучать его как стационарный процесс.

Второе условие, которому, очевидно, должна удовлетворять стационарная случайная функция, — это условие постоянства дисперсии:

$$D_x(t) = D_x = \text{const}. \qquad (17.1.2)$$

Установим, какому условию должна удовлетворять корреляционная функция стационарной случайной функции. Рассмотрим случайную функцию $X(t)$ (рис. 17.1.3). Положим в выражении $K_x(t, t')$ $t' = t + \tau$ и рассмотрим $K_x(t, t+\tau)$ — корреляционный момент двух сечений случайной функции, разделенных интервалом времени τ. Очевидно, если случайный процесс $X(t)$ действительно стационарен, то этот корреляционный момент не должен зависеть от того, г д е именно на оси Ot мы взяли участок τ, а должен зависеть только от д л и н ы этого участка. Например, для участков I и II на рис. 17.1.3, имеющих одну и ту же длину τ, значения корреляционной функции $K_x(t, t+\tau)$ и $K_x(t_1, t_1+\tau)$ должны быть одинаковыми. Вообще, корреляционная функция стационарного случайного процесса должна зависеть не от п о л о ж е н и я t первого аргумента на оси абсцисс, а т о л ь к о о т п р о м е ж у т к а τ между первым и вторым аргументами:

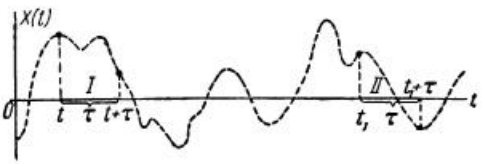


Рис. 17.1.3.

$$K_x(t, t+\tau) = k_x(\tau). \qquad (17.1.3)$$

Следовательно, корреляционная функция стационарного случайного процесса есть функция не двух, а всего одного аргумента. Это обстоятельство в ряде случаев сильно упрощает операции над стационарными случайными функциями.

Заметим, что условие (17.1.2), требующее от стационарной случайной функции постоянства дисперсии, является частным случаем условия (17.1.3). Действительно, полагая в формуле (17.1.3) $t+\tau=t$ ($\tau=0$), имеем

$$D_x(t)=K_x(t,\ t)=k_x(0)=\text{const}. \qquad (17.1.4)$$

Таким образом, условие (17.1.3) есть единственное существенное условие, которому должна удовлетворять стационарная случайная функция.

Поэтому в дальнейшем мы под стационарной случайной функцией будем понимать такую случайную функцию, корреляционная функция которой зависит не от обоих своих аргументов t и t', а только от разности τ между ними. Чтобы не накладывать специальных условий на математическое ожидание, мы будем рассматривать только центрированные случайные функции.

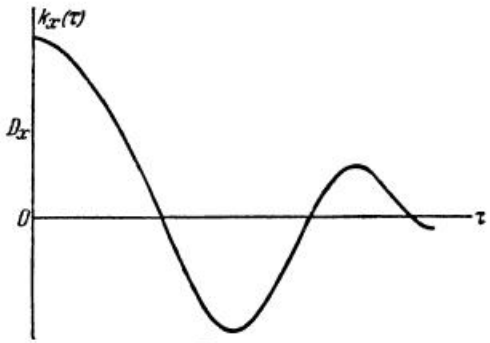


Рис. 17.1.4.

Мы знаем, что корреляционная функция любой случайной функции обладает свойством симметрии:

$$K_x(t,\ t')=K_x(t',\ t).$$

Отсюда для стационарного процесса, полагая $t'-t=\tau$, имеем:

$$k_x(\tau)=k_x(-\tau), \qquad (17.1.5)$$

т. е. корреляционная функция $k_x(\tau)$ есть *четная функция* своего аргумента. Поэтому обычно корреляционную функцию определяют только для положительных значений аргумента (рис. 17.1.4).

На практике, вместо корреляционной функции $k_x(\tau)$, часто пользуются *нормированной корреляционной функцией*

$$\rho_x(\tau)=\frac{k_x(\tau)}{D_x}, \qquad (17.1.6)$$

где $D_x=k_x(0)$ — постоянная дисперсия стационарного процесса. Функция $\rho_x(\tau)$ есть не что иное, как коэффициент корреляции между сечениями случайной функции, разделенными интервалом τ по времени. Очевидно, что $\rho_x(0)=1$.

В качестве примеров рассмотрим два образца приблизительно стационарных случайных процессов и построим их характеристики.

Пример 1. Случайная функция $X(t)$ задана совокупностью 12 реализаций (рис. 17.1.5). а) Найти ее характеристики $m_x(t)$, $K_x(t,\ t')$, $D_x(t)$ и

нормированную корреляционную функцию $r_X(t, t')$. б) Приближенно рассматривая случайную функцию $X(t)$ как стационарную, найти ее характеристики.

Решение. Так как случайная функция $X(t)$ меняется сравнительно плавно, можно брать сечения не очень часто, например через 0,4 *сек*. Тогда

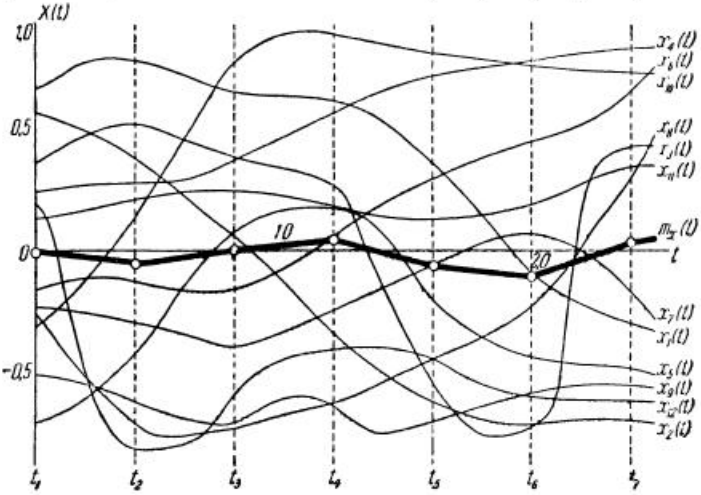


Рис. 17.1.5.

случайная функция будет сведена к системе семи случайных величин, отвечающих сечениям $t = 0$; 0,4; 0,8; 1,2; 1,6; 2,0; 2,4. Намечая эти сечения на графике и снимая с графика значения случайной функции в этих сечениях, получим таблицу (табл. 17.1.1).

Таблица 17.1.1

№ реализации \ t	0	0,4	0,8	1,2	1,6	2,0	2,4
1	0,64	0,74	0,62	0,59	0,35	—0,09	—0,39
2	0,54	0,37	0,06	—0,32	—0,60	—0,69	—0,67
3	0,34	0,50	0,37	0,26	—0,52	—0,72	0,42
4	0,23	0,26	0,35	0,55	0,69	0,75	0,80
5	0,12	0,20	0,24	0,18	—0,20	—0,42	—0,46
6	—0,16	—0,12	—0,15	0,05	0,29	0,43	0,63
7	—0,22	—0,29	—0,38	—0,24	—0,06	0,07	—0,16
8	—0,26	—0,69	—0,70	—0,61	—0,43	—0,22	0,29
9	—0,50	—0,60	—0,68	—0,62	—0,68	—0,56	—0,54
10	—0,30	0,13	0,75	0,84	0,78	0,73	0,71
11	—0,69	—0,40	0,08	0,16	0,12	0,18	0,33
12	0,18	—0,79	—0,56	—0,39	—0,42	—0,58	—0,53

Таблицу рекомендуется заполнять по строчкам, передвигаясь все время вдоль одной реализации.

Далее находим оценки для характеристик случайных величин $X(0)$, $X(0{,}4)$, ..., $X(2{,}4)$. Суммируя значения по столбцам и деля сумму на число

реализаций $n = 12$, найдем приближенно зависимость математического ожидания от времени:

t	0	0,4	0,8	1,2	1,6	2,0	2,4
$\tilde{m}_x(t)$	—0,007	—0,057	0,000	0,037	—0,057	—0,093	0,036

На графике рис. 17.1.5 математическое ожидание показано жирной линией.

Далее находим оценки для элементов корреляционной матрицы: дисперсий и корреляционных моментов. Вычисления удобнее всего производить по следующей схеме. Для вычисления статистической дисперсии суммируются квадраты чисел, стоящих в соответствующем столбце; сумма делится на $n = 12$; из результата вычитается квадрат соответствующего математического ожидания. Для получения несмещенной оценки результат множится на поправку $\frac{n}{n-1} = \frac{12}{11}$. Аналогично оцениваются корреляционные моменты. Для вычисления статистического момента, отвечающего двум заданным сечениям, перемножаются числа, стоящие в соответствующих столбцах; произведения складываются алгебраически; полученная сумма делится на $n = 12$; из результата вычитается произведение соответствующих математических ожиданий; для получения несмещенной оценки корреляционного момента результат множится на $\frac{n}{n-1}$. При выполнении расчетов на счетной машине или арифмометре промежуточные результаты умножений не записываются, а непосредственно суммируются [1]). Полученная таким способом корреляционная матрица системы случайных величин $X(0)$, $X(0{,}4)$, ..., $X(2{,}4)$ — она же таблица значений корреляционной функции $\tilde{K}_x(t, t')$ — приведена в таблице 17.1.2.

Таблица 17.1.2

t / t'	0	0,4	0,8	1,2	1,6	2,0	2,4
0	0,1632	0,1379	0,0795	0,0457	—0,0106	—0,0642	—0,0648
0,4		0,2385	0,2029	0,1621	0,0827	0,0229	0,0251
0,8			0,2356	0,2152	0,1527	0,0982	0,0896
1,2				0,2207	0,1910	0,1491	0,1322
1,6					0,2407	0,2348	0,1711
2,0						0,2691	0,2114
2,4							0,2878

По главной диагонали таблицы стоят оценки дисперсий:

t	0	0,4	0,8	1,2	1,6	2,0	2,4
$\tilde{D}_x(t)$	0,1632	0,2385	0,2356	0,2207	0,2407	0,2691	0,2878

[1]) В данном случае оказалось возможным непосредственно вести обработку через начальные моменты, так как математическое ожидание функции $X(t)$ близко к нулю. Если это не так, то перед обработкой необходимо перенести начало отсчета поближе к математическому ожиданию.

Извлекая из этих величин квадратные корни, найдем зависимость среднего квадратического отклонения $\tilde{\sigma}_x$ от времени:

t	0	0,4	0,8	1,2	1,6	2,0	2,4
$\tilde{\sigma}_x(t)$	0,404	0,488	0,485	0,470	0,491	0,519	0,536

Деля значения, стоящие в табл. 17.1.2, на произведения соответствующих средних квадратических отклонений, получим таблицу значений нормированной корреляционной функции $\tilde{r}_x(t, t')$ (табл. 17.1.3).

Таблица 17.1.3

t' \ t	0	0,4	0,8	1,2	1,6	2,0	2,4
0	1	0,700	0,405	0,241	—0,053	—0,306	—0,299
0,4		1	0,856	0,707	0,345	0,090	0,095
0,8			1	0,943	0,643	0,390	0,344
1,2				1	0,829	0,612	0,524
1,6					1	0,923	0,650
2,0						1	0,760
2,4							1

Проанализируем полученные данные под углом зрения предполагаемой стационарности случайной функции $X(t)$. Если судить непосредственно по данным, полученным в результате обработки, то можно прийти к выводу, что случайная функция $X(t)$ стационарной не является: ее математическое ожидание не вполне постоянно; дисперсия также несколько меняется со временем; значения нормированной корреляционной функции вдоль параллелей главной диагонали также не вполне постоянны. Однако, принимая во внимание весьма ограниченное число обработанных реализаций ($n = 12$) и в связи с этим наличие большого элемента случайности в полученных оценках, эти видимые отступления от стационарности вряд ли можно считать значимыми, тем более, что они не носят сколько-нибудь закономерного характера. Поэтому вполне целесообразной будет приближенная замена функции $X(t)$ стационарной. Для приведения функции к стационарной прежде всего осредним по времени оценки для математического ожидания:

$$\tilde{m}_x = \frac{\tilde{m}_x(0) + \tilde{m}_x(0{,}4) + \ \dots \ + \tilde{m}_x(2{,}4)}{7} \approx -0{,}02.$$

Аналогичным образом осредним оценки для дисперсии:

$$\tilde{D}_x = \frac{\tilde{D}_x(0) + \tilde{D}_x(0{,}4) + \ \dots \ + \tilde{D}_x(2{,}4)}{7} \approx 0{,}236.$$

Извлекая корень, найдем осредненную оценку с. к. о.:

$$\tilde{\sigma}_x \approx 0{,}486.$$

Перейдем к построению нормированной корреляционной функции того стационарного процесса, которым можно заменить случайную функцию $X(t)$. Для стационарного процесса корреляционная функция (а значит, и нормированная корреляционная функция) зависит только от $\tau = t' - t$; следовательно,

при постоянном τ корреляционная функция должна быть постоянной. В таблице 17.1.3 постоянному τ соответствуют: главная диагональ ($\tau = 0$) и параллели этой диагонали ($\tau = 0{,}4$; $\tau = 0{,}8$; $\tau = 1{,}2$ и т. д.). Осредняя оценки нормированной корреляционной функции вдоль этих параллелей главной диагонали, получим значения функции $\tilde{\rho}_x(\tau)$:

t	0	0,4	0,8	1,2	1,6	2,0	2,4
$\tilde{\rho}_x(\tau)$	1,00	0,84	0,60	0,38	0,13	—0,10	—0,30

График функции $\tilde{\rho}_x(\tau)$ представлен на рис. 17.1.6.

При рассмотрении рис. 17.1.6 обращает на себя внимание наличие для некоторых τ отрицательных значений корреляционной функции. Это указывает на то, что в структуре случайной функции имеется некоторый элемент периодичности, в связи с чем на расстоянии по времени, равном примерно половине периода основных колебаний, наблюдается отрицательная корреляция между значениями случайной функции: положительным отклонениям от среднего в одном сечении соответствуют отрицательные отклонения через определенный промежуток времени, и наоборот.

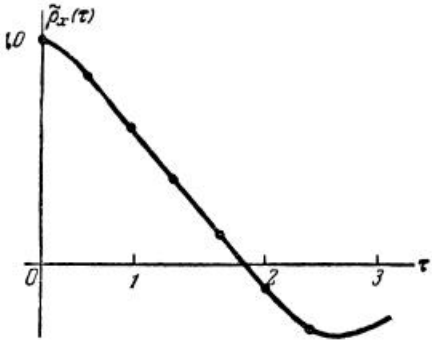


Рис. 17.1.6.

Такой характер корреляционной функции, с переходом на отрицательные значения, очень часто встречается на практике. Обычно в таких случаях по мере увеличения τ амплитуда колебаний корреляционной функции уменьшается и при дальнейшем увеличении τ корреляционная функция стремится к нулю.

Пример 2. Случайная функция $X(t)$ задана совокупностью 12 своих реализаций (рис. 17.1.7). Приближенно заменив функцию $X(t)$ стационарной, сравнить ее нормированную корреляционную функцию с функцией $\tilde{\rho}_x(\tau)$ предыдущего примера.

Решение. Так как случайная функция $X(t)$ отличается значительно менее плавным ходом по сравнению с функцией $X(t)$ предыдущего примера, промежуток между сечениями уже нельзя брать равным 0,4 *сек*, как в предыдущем примере, а следует взять по крайней мере вдвое меньше (например, 0,2 *сек*, как на рис. 17.1.7). В результате обработки получаем оценку для нормированной корреляционной функции $\tilde{\rho}_x(\tau)$:

τ	0	0,2	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4
$\tilde{\rho}_x(\tau)$	1,00	0,73	0,41	0,22	—0,01	—0,20	—0,19	—0,10	—0,06	—0,15	0,08	0,19	0,05

График функции $\tilde{\rho}_x(t)$ представлен на рис. 17.1.8. Из сравнения графиков рис. 17.1.8 и 17.1.6 видно, что корреляционная функция, изображенная

на рис. 17.1.8, убывает значительно быстрее. Это и естественно, так как характер изменения функции $X(t)$ в примере 1 гораздо более

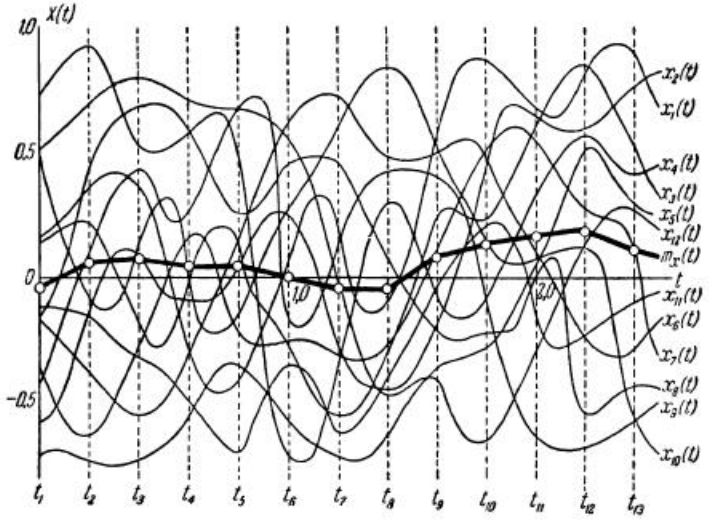


Рис. 17.1.7.

плавный и постепенный, чем в примере 2; в связи с этим корреляция между значениями случайной функции в примере 1 убывает медленнее.

При рассмотрении рис. 17.1.8 бросаются в глаза незакономерные колебания функции $\tilde{\rho}_x(\tau)$ для больших значений τ. Так как при больших значениях τ точки графика получены осреднением сравнительно очень небольшого числа данных, их нельзя считать надежными. В подобных случаях имеет смысл сгладить корреляционную функцию, как, например, показано пунктиром на рис. 17.1.8.

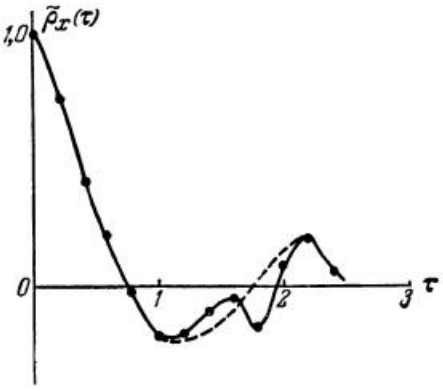


Рис. 17.1.8.

17.2. Спектральное разложение стационарной случайной функции на конечном участке времени. Спектр дисперсий

На двух примерах, приведенных в предыдущем п°, мы наглядно убедились в том, что существует связь между характером корреляционной функции и внутренней структурой соответствующего ей случайного процесса. В зависимости от того, какие частоты и

в каких соотношениях преобладают в составе случайной функции, ее корреляционная функция имеет тот или другой вид. Из таких соображений мы непосредственно приходим к понятию о *спектральном составе* случайной функции.

Понятие «спектра» встречается не только в теории случайных функций; оно широко применяется в математике, физике и технике.

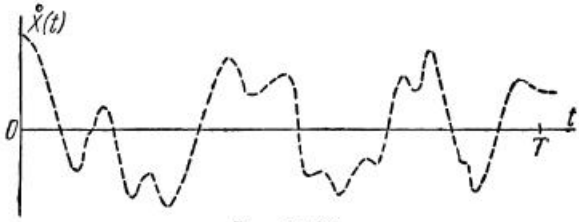


Рис. 17.2.1.

Если какой-либо колебательный процесс представляется в виде суммы гармонических колебаний различных частот (так называемых «гармоник»), то *спектром* колебательного процесса называется функция, описывающая распределение амплитуд по различным частотам. Спектр показывает, какого рода колебания преобладают в данном процессе, какова его внутренняя структура.

Совершенно аналогичное спектральное описание можно дать и стационарному случайному процессу; вся разница в том, что для случайного процесса амплитуды колебаний будут случайными величинами. Спектр стационарной случайной функции будет описывать *распределение дисперсий* по различным частотам.

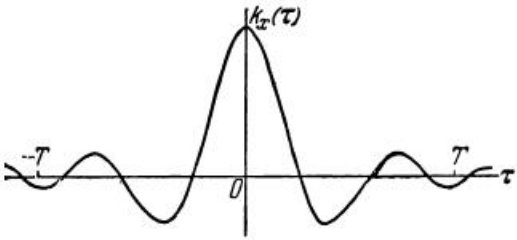


Рис. 17.2.2.

Подойдем к понятию о спектре стационарной случайной функции из следующих соображений.

Рассмотрим стационарную случайную функцию $\mathring{X}(t)$, которую мы наблюдаем на интервале $(0, T)$ (рис. 17.2.1).

Задана корреляционная функция случайной функции $X(t)$

$$K_x(t,\ t+\tau) = k_x(\tau).$$

Функция $k_x(\tau)$ есть четная функция:

$$k_x(\tau) = k_x(-\tau)$$

и, следовательно, на графике изобразится симметричной кривой (рис. 17.2.2).

При изменении t и t' от 0 до T аргумент $\tau = t' - t$ изменяется от $-T$ до $+T$.

Мы знаем, что четную функцию на интервале $(-T, T)$ можно разложить в ряд Фурье, пользуясь только четными (косинусными) гармониками:

$$k_x(\tau) = \sum_{k=0}^{\infty} D_k \cos \omega_k \tau, \tag{17.2.1}$$

где

$$\omega_k = k\omega_1; \qquad \omega_1 = \frac{2\pi}{2T} = \frac{\pi}{T}, \tag{17.2.2}$$

а коэффициенты D_k определяются формулами:

$$\left.\begin{aligned} D_0 &= \frac{1}{2T} \int_{-T}^{T} k_x(\tau)\, d\tau, \\ D_k &= \frac{1}{T} \int_{-T}^{T} k_x(\tau) \cos \omega_k \tau\, d\tau \quad \text{при} \quad k \neq 0. \end{aligned}\right\} \tag{17.2.3}$$

Имея в виду, что функции $k_x(\tau)$ и $\cos \omega_k \tau$ четные, можно преобразовать формулы (17.2.3) к виду:

$$\left.\begin{aligned} D_0 &= \frac{1}{T} \int_{0}^{T} k_x(\tau)\, d\tau, \\ D_k &= \frac{2}{T} \int_{0}^{T} k_x(\tau) \cos \omega_k \tau\, d\tau \quad \text{при} \quad k \neq 0. \end{aligned}\right\} \tag{17.2.4}$$

Перейдем в выражении (17.2.1) корреляционной функции $k_x(\tau)$ от аргумента τ снова к двум аргументам t и t'. Для этого положим

$$\cos \omega_k \tau = \cos \omega_k (t' - t) = \cos \omega_k t' \cos \omega_k t + \sin \omega_k t' \sin \omega_k t \tag{17.2.5}$$

и подставим выражение (17.2.5) в формулу (17.2.1):

$$K_x(t, t') = \sum_{k=0}^{\infty} (D_k \cos \omega_k t' \cos \omega_k t + D_k \sin \omega_k t' \sin \omega_k t). \tag{17.2.6}$$

Мы видим, что выражение (17.2.6) есть не что иное, как *каноническое разложение* корреляционной функции $K_x(t, t')$. Координатными функциями этого канонического разложения являются попеременно косинусы и синусы частот, кратных ω_1:

$$\cos \omega_k t, \ \sin \omega_k t \qquad (k = 0, 1, \ldots).$$

Мы знаем, что по каноническому разложению корреляционной функции можно построить каноническое разложение самой случайной функции с теми же координатными функциями и с дисперсиями, равными коэффициентам D_k в каноническом разложении корреляционной функции [1]).

Следовательно, случайная функция $\mathring{X}(t)$ может быть представлена в виде канонического разложения:

$$\mathring{X}(t) = \sum_{k=0}^{\infty} (U_k \cos \omega_k t + V_k \sin \omega_k t), \qquad (17.2.7)$$

где U_k, V_k — некоррелированные случайные величины с математическими ожиданиями, равными нулю, и дисперсиями, одинаковыми для каждой пары случайных величин с одним и тем же индексом k:

$$D[U_k] = D[V_k] = D_k. \qquad (17.2.8)$$

Дисперсии D_k при различных k определяются формулами (17.2.4).

Таким образом, мы получили на интервале $(0, T)$ каноническое разложение случайной функции $\mathring{X}(t)$, координатными функциями которого являются функции $\cos \omega_k t$, $\sin \omega_k t$ при различных ω_k. Разложение такого рода называется *спектральным разложением* стационарной случайной функции. На представлении случайных функций в виде спектральных разложений основана так называемая *спектральная теория* стационарных случайных процессов.

Спектральное разложение изображает стационарную случайную функцию разложенной на гармонические колебания различных частот:

$$\omega_1, \ \omega_2, \ \ldots, \ \omega_k, \ \ldots,$$

причем амплитуды этих колебаний являются случайными величинами.

Определим дисперсию случайной функции $\mathring{X}(t)$, заданной спектральным разложением (17.2.7). По теореме о дисперсии линейной функции некоррелированных случайных величин

$$D_x = D[\mathring{X}(t)] = \sum_{k=0}^{\infty} (\cos^2 \omega_k t + \sin^2 \omega_k t) D_k = \sum_{k=0}^{\infty} D_k. \qquad (17.2.9)$$

Таким образом, *дисперсия стационарной случайной функции равна сумме дисперсий всех гармоник ее спектрального разложения*. Формула (17.2.9) показывает, что дисперсия функции $\mathring{X}(t)$

[1]) Можно доказать, что для любой корреляционной функции $k_x(\tau)$ коэффициенты D_k, выражаемые формулами (17.2.4), неотрицательны.

известным образом распределена по различным частотам: одним частотам соответствуют бо́льшие дисперсии, другим — меньшие. Распределение дисперсий по частотам можно проиллюстрировать графически в виде так называемого *спектра* стационарной случайной функции (точнее — спектра дисперсий). Для этого по оси абсцисс откладываются частоты $\omega_0 = 0$, ω_1, ω_2, ..., ω_k, ..., а по оси ординат—соответствующие дисперсии (рис. 17.2.3).

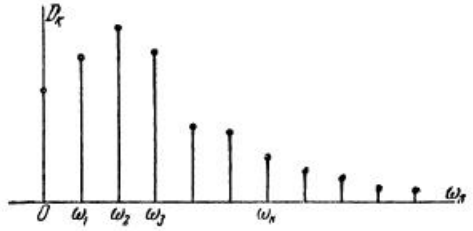


Рис. 17.2.3.

Очевидно, сумма всех ординат построенного таким образом спектра равна дисперсии случайной функции.

17.3. Спектральное разложение стационарной случайной функции на бесконечном участке времени. Спектральная плотность стационарной случайной функции

Строя спектральное разложение стационарной случайной функции $\mathring{X}(t)$ на конечном участке времени $(0, T)$, мы получили спектр дисперсий случайной функции в виде ряда отдельных дискретных линий, разделенных равными промежутками (так называемый «прерывистый» или «линейчатый» спектр).

Очевидно, чем бо́льший участок времени мы будем рассматривать, тем полнее будут наши сведения о случайной функции. Естественно поэтому в спектральном разложении попытаться перейти к пределу при $T \to \infty$ и посмотреть, во что при этом обратится спектр случайной функции. При $T \to \infty$ $\omega_1 = \frac{2\pi}{2T} \to 0$; поэтому расстояния между частотами ω_k, на которых строится спектр, будут при $T \to \infty$ неограниченно уменьшаться. При этом дискретный спектр будет приближаться к непрерывному, в котором каждому сколь угодно малому интервалу частот $\Delta\omega$ будет соответствовать элементарная дисперсия $\Delta D(\omega)$.

Попробуем изобразить непрерывный спектр графически. Для этого мы должны несколько перестроить график дискретного спектра при конечном T. А именно, будем откладывать по оси ординат уже не самую дисперсию D_k (которая безгранично уменьшается при $T \to \infty$), а *среднюю плотность дисперсии*, т. е. дисперсию, приходящуюся на единицу длины данного интервала частот. Обозначим расстояние между соседними частотами $\Delta\omega$:

$$\omega_1 = \frac{2\pi}{2T} = \Delta\omega$$

и на каждом отрезке $\Delta\omega$, как на основании, построим прямоугольник с площадью D_k (рис. 17.3.1). Получим ступенчатую диаграмму, напоминающую по принципу построения гистограмму статистического распределения.

Высота диаграммы на участке $\Delta\omega$, прилежащем к точке ω_k, равна

$$S_x(\omega_k) = \frac{D_k}{\Delta\omega} \qquad (17.3.1)$$

и представляет собой среднюю плотность дисперсии на этом участке. Суммарная площадь всей диаграммы, очевидно, равна дисперсии случайной функции.

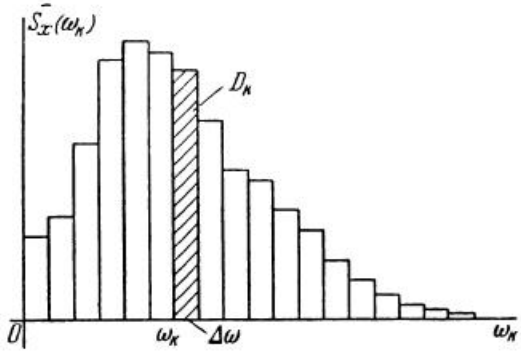


Рис. 17.3.1.

Будем неограниченно увеличивать интервал T. При этом $\Delta\omega \to 0$, и ступенчатая кривая будет неограниченно приближаться к плавной кривой $S_x(\omega)$ (рис. 17.3.2). Эта кривая изображает плотность распределения дисперсий по частотам непрерывного спектра, а сама функция $S_x(\omega)$ называется *спектральной плотностью дисперсии*, или, короче, *спектральной плотностью* стационарной случайной функции $\mathring{X}(t)$.

Очевидно, площадь, ограниченная кривой $S_x(\omega)$, по-прежнему должна равняться дисперсии D_x случайной функции $\mathring{X}(t)$:

$$D_x = \int_0^\infty S_x(\omega)\, d\omega. \qquad (17.3.2)$$

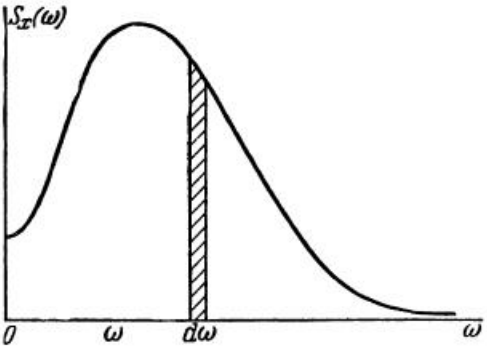


Рис. 17.3.2.

Формула (17.3.2) есть не что иное, как разложение дисперсии D_x на сумму элементарных слагаемых $S_x(\omega)\, d\omega$, каждое из которых представляет собой дисперсию, приходящуюся на элементарный участок частот $d\omega$, прилежащий к точке ω (рис. 17.3.2).

Таким образом, мы ввели в рассмотрение новую дополнительную характеристику стационарного случайного процесса — спектральную

плотность, описывающую частотный состав стационарного процесса. Однако эта характеристика не является самостоятельной; она полностью определяется корреляционной функцией данного процесса. Подобно тому, как ординаты дискретного спектра D_k выражаются формулами (17.2.4) через корреляционную функцию $k_x(\tau)$, спектральная плотность $S_x(\omega)$ также может быть выражена через корреляционную функцию.

Выведем это выражение. Для этого перейдем в каноническом разложении корреляционной функции к пределу при $T \to \infty$ и посмотрим, во что оно обратится. Будем исходить из разложения (17.2.1) корреляционной функции в ряд Фурье на конечном интервале $(-T, T)$:

$$k_x(\tau) = \sum_{k=0}^{\infty} D_k \cos \omega_k \tau, \tag{17.3.3}$$

где дисперсия, соответствующая частоте ω_k, выражается формулой

$$D_k = \frac{2}{T} \int_0^T k_x(\tau) \cos \omega_k \tau \, d\tau. \tag{17.3.4}$$

Перед тем как переходить к пределу при $T \to \infty$, перейдем в формуле (17.3.3) от дисперсии D_k к средней плотности дисперсии $\frac{D_k}{\Delta\omega}$. Так как эта плотность вычисляется еще при конечном значении T и зависит от T, обозначим ее:

$$S_x^{(T)}(\omega_k) = \frac{D_k}{\Delta\omega}. \tag{17.3.5}$$

Разделим выражение (17.3.4) на $\Delta\omega = \frac{\pi}{T}$; получим:

$$S_x^{(T)}(\omega_k) = \frac{2}{\pi} \int_0^T k_x(\tau) \cos \omega_k \tau \, d\tau. \tag{17.3.6}$$

Из (17.3.5) следует, что

$$D_k = S_x^{(T)}(\omega_k) \, \Delta\omega. \tag{17.3.7}$$

Подставим выражение (17.3.7) в формулу (17.3.3); получим:

$$k_x(\tau) = \sum_{k=0}^{\infty} S_x^{(T)}(\omega_k) \cos \omega_k \tau \, \Delta\omega. \tag{17.3.8}$$

Посмотрим, во что превратится выражение (17.3.8) при $T \to \infty$. Очевидно, при этом $\Delta\omega \to 0$; дискретный аргумент ω_k переходит в непрерывно меняющийся аргумент ω; сумма переходит в интеграл по переменной ω; средняя плотность дисперсии $S_x^{(T)}(\omega_k)$ стремится

к плотности дисперсии $S_x(\omega)$, и выражение (17.3.8) в пределе принимает вид:

$$k_x(\tau) = \int_0^\infty S_x(\omega) \cos \omega\tau \, d\omega, \tag{17.3.9}$$

где $S_x(\omega)$ — спектральная плотность стационарной случайной функции.

Переходя к пределу при $T \to \infty$ в формуле (17.3.6), получим выражение спектральной плотности через корреляционную функцию:

$$S_x(\omega) = \frac{2}{\pi} \int_0^\infty k_x(\tau) \cos \omega\tau \, d\tau. \tag{17.3.10}$$

Выражение типа (17.3.9) известно в математике под названием *интеграла Фурье*. Интеграл Фурье есть обобщение разложения в ряд Фурье для случая непериодической функции, рассматриваемой на бесконечном интервале, и представляет собой разложение функции на сумму элементарных гармонических колебаний с непрерывным спектром [1]).

Подобно тому как ряд Фурье выражает разлагаемую функцию через коэффициенты ряда, которые в свою очередь выражаются через разлагаемую функцию, формулы (17.3.9) и (17.3.10) выражают функции $k_x(\tau)$ и $S_x(\omega)$ взаимно: одна через другую. Формула (17.3.9) выражает корреляционную функцию через спектральную плотность; формула (17.3.10), наоборот, выражает спектральную плотность через корреляционную функцию. Формулы типа (17.3.9) и (17.3.10), связывающие взаимно две функции, называются *преобразованиями Фурье* [2]).

Таким образом, корреляционная функция и спектральная плотность выражаются одна через другую с помощью преобразований Фурье.

Заметим, что из общей формулы (17.3.9) при $\tau = 0$ получается ранее полученное разложение дисперсии по частотам (17.3.2).

На практике вместо спектральной плотности $S_x(\omega)$ часто пользуются *нормированной* спектральной плотностью:

$$s_x(\omega) = \frac{S_x(\omega)}{D_x}, \tag{17.3.11}$$

где D_x — дисперсия случайной функции.

---

[1]) Формула (17.3.9) является частным видом интеграла Фурье, обобщающим разложение в ряд Фурье четной функции по косинусным гармоникам. Аналогичное выражение может быть написано и для более общего случая.

[2]) В данном случае мы имеем дело с частным случаем преобразований Фурье — с так называемыми «косинус-преобразованиями Фурье».

Нетрудно убедиться, что нормированная корреляционная функция $\rho_X(\tau)$ и нормированная спектральная плотность $s_X(\omega)$ связаны теми же преобразованиями Фурье:

$$\left.\begin{aligned} \rho_X(\tau) &= \int_0^\infty s_X(\omega)\cos\omega\tau\, d\omega, \\ s_X(\omega) &= \frac{2}{\pi}\int_0^\infty \rho_X(\tau)\cos\omega\tau\, d\tau. \end{aligned}\right\} \qquad (17.3.12)$$

Полагая в первом из равенств (17.3.12) $\tau = 0$ и учитывая, что $\rho_X(0) = 1$, имеем:

$$\int_0^\infty s_X(\omega)\, d\omega = 1, \qquad (17.3.13)$$

т. е. полная площадь, ограниченная графиком нормированной спектральной плотности, равна единице.

Пример 1. Нормированная корреляционная функция $\rho_X(\tau)$ случайной функции $X(t)$ убывает по линейному закону от единицы до нуля при $0 < \tau < \tau_0$;

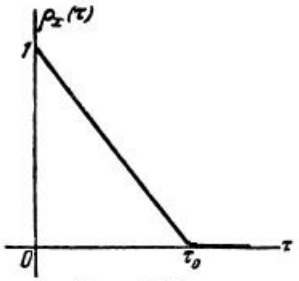


Рис. 17.3.3.

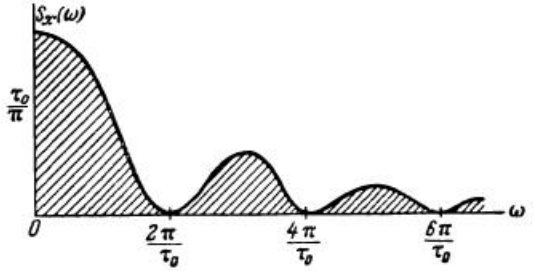


Рис. 17.3.4.

при $\tau > \tau_0$ $\rho_X(\tau) = 0$ (рис. 17.3.3). Определить нормированную спектральную плотность случайной функции $X(t)$.

Решение. Нормированная корреляционная функция выражается формулами:

$$\rho_X(\tau)\begin{cases} = 1 - \dfrac{\tau}{\tau_0} & \text{при } 0 < \tau < \tau_0, \\ = 0 & \text{при } \tau > \tau_0. \end{cases}$$

Из формул (17.3.12) имеем:

$$s_X(\omega) = \frac{2}{\pi}\int_0^\infty \rho_X(\tau)\cos\omega\tau\, d\tau =$$

$$= \frac{2}{\pi}\int_0^{\tau_0}\left(1 - \frac{\tau}{\tau_0}\right)\cos\omega\tau\, d\tau = \frac{2}{\pi\tau_0\omega^2}(1 - \cos\omega\tau_0).$$

График нормированной спектральной плотности представлен на рис. 17.3.4. Первый — абсолютный — максимум спектральной плотности достигается при $\omega = 0$; раскрытием неопределенности в этой точке убеждаемся, что он равен $\frac{\tau_0}{\pi}$. Далее при возрастании ω спектральная плотность достигает ряда относительных максимумов, высота которых убывает с возрастанием ω; при $\omega \to \infty$ $s_X(\omega) \to 0$.

Характер изменения спектральной плотности $s_X(\omega)$ (быстрое или медленное убывание) зависит от параметра τ_0. Полная площадь, ограниченная кривой $s_X(\omega)$, постоянна и равна единице. Изменение τ_0 равносильно изменению масштаба кривой $s_X(\omega)$ по обеим осям при сохранении ее площади. При увеличении τ_0 масштаб по оси ординат увеличивается, по оси абсцисс — уменьшается; преобладание в спектре случайной функции нулевой частоты становится более ярко выраженным. В пределе при $\tau_0 \to \infty$ случайная функция вырождается в обычную случайную величину; при этом $\rho_X(\tau) \equiv 1$, а спектр становится дискретным с одной-единственной частотой $\omega_0 = 0$.

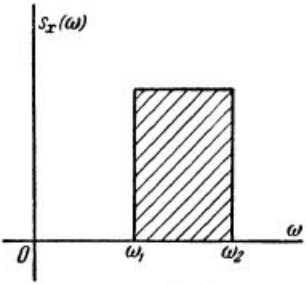


Рис. 17.3.5.

Пример 2. Нормированная спектральная плотность $s_X(\omega)$ случайной функции $\mathring{X}(t)$ постоянна на некотором интервале частот ω_1, ω_2 и равна нулю вне этого интервала (рис. 17.3.5). Определить нормированную корреляционную функцию случайной функции $\mathring{X}(t)$.

Решение. Значение $s_X(\omega)$ при $\omega_1 < \omega < \omega_2$ определяем из условия, что площадь, ограниченная кривой $s_X(\omega)$, равна единице:

$$s_X(\omega)(\omega_2 - \omega_1) = 1, \quad s_X(\omega) = \frac{1}{\omega_2 - \omega_1}.$$

Из (17.3.12) имеем:

$$\rho_X(\tau) = \int_{\omega_1}^{\omega_2} s_X(\omega) \cos \omega\tau \, d\omega = \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} \cos \omega\tau \, d\omega =$$

$$= \frac{1}{\tau(\omega_2 - \omega_1)} (\sin \omega_2\tau - \sin \omega_1\tau) = \frac{2}{\tau(\omega_2 - \omega_1)} \cos\left(\frac{\omega_2 + \omega_1}{2}\tau\right) \sin\left(\frac{\omega_2 - \omega_1}{2}\tau\right).$$

Общий вид функции $\rho_X(\tau)$ изображен на рис. 17.3.6. Она носит характер убывающих по амплитуде колебаний с рядом узлов, в которых функция обращается в нуль. Конкретный вид графика, очевидно, зависит от значений ω_1, ω_2.

Представляет интерес предельный вид функции $\rho_X(\tau)$ при $\omega_1 \to \omega_2$. Очевидно, при $\omega_2 = \omega_1 = \omega$ спектр случайной функции обращается в дискретный с одной-единственной линией, соответствующей частоте ω; при этом корреляционная функция обращается в простую косинусоиду:

$$\rho_X(\tau) = \cos \omega\tau.$$

Посмотрим, какой вид в этом случае имеет сама случайная функция $\mathring{X}(t)$. При дискретном спектре с одной-единственной линией спектральное разложение стационарной случайной функции $\mathring{X}(t)$ имеет вид:

$$\mathring{X}(t) = U \cos \omega t + V \sin \omega t, \tag{17.3.14}$$

где U и V — некоррелированные случайные величины с математическими ожиданиями, равными нулю, и равными дисперсиями:

$$D[U] = D[V] = D.$$

Покажем, что случайная функция типа (17.3.14) может быть представлена как одно гармоническое колебание частоты ω со случайной амплитудой и случайной фазой. Обозначая

$$\cos\Phi = \frac{U}{\sqrt{U^2+V^2}}, \quad \sin\Phi = \frac{V}{\sqrt{U^2+V^2}},$$

приводим выражение (17.3.14) к виду:

$$\mathring{X}(t) = \sqrt{U^2+V^2}\,(\cos\Phi\cos\omega t + \sin\Phi\sin\omega t) = \sqrt{U^2+V^2}\cos(\omega t - \Phi).$$

В этом выражении $\sqrt{U^2+V^2}$ — случайная амплитуда; Φ — случайная фаза гармонического колебания.

До сих пор мы рассматривали только тот случай, когда распределение дисперсий по частотам является непрерывным, т. е. когда на бесконечно малый участок частот приходится бесконечно малая дисперсия. На практике иногда встречаются случаи, когда случайная

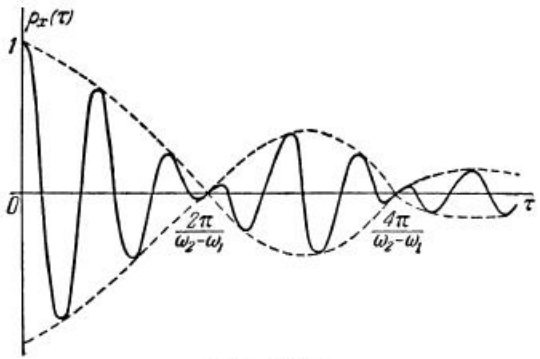


Рис. 17.3.6.

функция имеет в своем составе чисто периодическую составляющую частоты ω_k со случайной амплитудой. Тогда в спектральном разложении случайной функции, помимо непрерывного спектра частот, будет фигурировать еще отдельная частота ω_k с конечной дисперсией D_k. В общем случае таких периодических составляющих может быть несколько. Тогда спектральное разложение корреляционной функции будет состоять из двух частей: дискретного и непрерывного спектра:

$$k_x(\tau) = \sum_k D_k \cos\omega_k\tau + \int_0^\infty S_x(\omega)\cos\omega\tau\, d\omega. \qquad (17.3.15)$$

Случаи стационарных случайных функций с таким «смешанным» спектром на практике встречаются довольно редко. В этих случаях всегда имеет смысл разделить случайную функцию на два слагаемых — с непрерывным и дискретным спектром — и исследовать эти слагаемые в отдельности.

Относительно часто приходится иметь дело с частным случаем, когда конечная дисперсия в спектральном разложении случайной функции приходится на нулевую частоту ($\omega = 0$). Это значит, что в состав случайной функции в качестве слагаемого входит обычная случайная величина с дисперсией D_0. В подобных случаях также имеет смысл выделить это случайное слагаемое и оперировать с ним отдельно.

17.4. Спектральное разложение случайной функции в комплексной форме

В ряде случаев с точки зрения простоты математических преобразований оказывается удобным пользоваться не действительной, а комплексной формой записи как спектрального разложения случайной функции, так и ее характеристик: спектральной плотности и корреляционной функции. Комплексная форма записи удобна, в частности, потому, что всевозможные линейные операции над функциями, имеющими вид гармонических колебаний (дифференцирование, интегрирование, решение линейных дифференциальных уравнений и т. д.), осуществляются гораздо проще, когда эти гармонические колебания записаны не в виде синусов и косинусов, а в комплексной форме, в виде показательных функций. Комплексная форма записи корреляционной функции и спектральной плотности применяется и в тех случаях, когда сама случайная функция (а следовательно, и ее корреляционная функция и спектральная плотность) действительна.

Покажем, как можно в спектральном разложении случайной функции чисто формально перейти от действительной формы к комплексной.

Рассмотрим спектральное разложение (17.2.8) случайной функции $\mathring{X}(t)$ на участке $(0, T)$:

$$\mathring{X}(t) = \sum_{k=0}^{\infty} (U_k \cos \omega_k t + V_k \sin \omega_k t), \qquad (17.4.1)$$

где U_k, V_k — некоррелированные случайные величины, причем для каждой пары U_k, V_k с одинаковыми индексами дисперсии равны:

$$D[U_k] = D[V_k] = D_k.$$

Учитывая, что $\omega_k = k\omega_1$; $\omega_0 = 0$, перепишем выражение (17.4.1) в виде:

$$\mathring{X}(t) = U_0 + \sum_{k=1}^{\infty} (U_k \cos \omega_k t + V_k \sin \omega_k t). \qquad (17.4.2)$$

Придадим спектральному разложению (17.4.2) комплексную форму. Для этого воспользуемся известными формулами Эйлера:

$$\cos \omega_k t = \frac{e^{i\omega_k t} + e^{-i\omega_k t}}{2};$$

$$\sin \omega_k t = \frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2i} = -i\frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2}.$$

Подставляя эти выражения в формулу (17.4.2), имеем:

$$\mathring{X}(t) = U_0 + \sum_{k=1}^{\infty}\left(U_k \frac{e^{i\omega_k t} + e^{-i\omega_k t}}{2} - iV_k \frac{e^{i\omega_k t} - e^{-i\omega_k t}}{2}\right), \quad (17.4.3)$$

т. е. разложение с координатными функциями $e^{i\omega_k t}$, $e^{-i\omega_k t}$.

Преобразуем разложение (17.4.3) так, чтобы в нем в качестве координатных функций фигурировали только функции $e^{i\omega_k t}$; для этого распространим условно область частот ω_k на отрицательные значения ω и в качестве частот спектрального разложения будем рассматривать значения

$$\omega_k = k\omega_1 \quad (k = \pm 1, \pm 2, \ldots),$$

т. е. будем считать, что k принимает не только положительные, но и отрицательные значения. Тогда формулу (17.4.3) можно переписать в виде:

$$\mathring{X}(t) = U_0 + \sum_{k=1}^{\infty} \frac{U_k - iV_k}{2} e^{i\omega_k t} + \sum_{k=-1}^{-\infty} \frac{U_k + iV_k}{2} e^{i\omega_k t}, \quad (17.4.4)$$

если положить

$$U_{-k} = U_k; \; V_{-k} = V_k.$$

Формула (17.4.4) представляет собой разложение случайной функции $\mathring{X}(t)$, в котором в качестве координатных функций фигурируют комплексные функции $e^{i\omega_k t}$, а коэффициенты представляют собой комплексные случайные величины. Обозначая эти комплексные случайные величины W_k ($k = \pm 1, \pm 2, \ldots$), придадим разложению (17.4.4) форму:

$$\mathring{X}(t) = \sum_{k=-\infty}^{\infty} W_k e^{i\omega_k t}, \quad (17.4.5)$$

где

$$\left.\begin{aligned} W_k &= U_0 && \text{при } k = 0, \\ W_k &= \frac{U_k - iV_k}{2} && \text{при } k > 0, \\ W_k &= \frac{U_k + iV_k}{2} && \text{при } k < 0. \end{aligned}\right\} \quad (17.4.6)$$

Докажем, что разложение (17.4.5) является *каноническим* разложением случайной функции $\mathring{X}(t)$. Для этого достаточно показать, что случайные коэффициенты этого разложения не коррелированы между собой.

Рассмотрим сначала коэффициенты двух различных членов разложения в положительной части спектра W_k и W_l при $k \neq l$, $k > 0$, $l > 0$ и определим корреляционный момент этих величин. Согласно определению корреляционного момента для комплексных случайных величин (см. п° 15.9) имеем:

$$K_{kl} = M[W_k \overline{W}_l],$$

где $\overline{W}_l$ — комплексная сопряженная величина для W_l.

При $k > 0$, $l > 0$

$$K_{kl} = M\left[\frac{U_k - iV_k}{2}\,\overline{\frac{U_l - iV_l}{2}}\right] = M\left[\frac{U_k - iV_k}{2}\,\frac{U_l + iV_l}{2}\right] =$$

$$= \frac{1}{4}\{M[U_k U_l] + iM[U_k V_l] - iM[U_l V_k] + M[V_k V_l]\} = 0,$$

так как случайные величины U_k, V_k, фигурирующие в разложении (17.4.1), все не коррелированы между собой.

Совершенно так же докажем некоррелированность величин W_k, W_l при любых знаках индексов k и l, если $k \neq \pm l$.

Остается доказать только некоррелированность коэффициентов при симметричных членах разложения, т. е. величин W_k и W_{-k} при любом k. Имеем:

$$K_{k,-k} = M[W_k W_{-k}] =$$

$$= M\left[\frac{U_k - iV_k}{2}\,\overline{\frac{U_k + iV_k}{2}}\right] = \frac{1}{4}M[(U_k - iV_k)^2] =$$

$$= \frac{1}{4}\{M[U_k^2] - M[V_k^2] - 2iM[U_k V_k]\}.$$

Учитывая, что величины U_k, V_k, входящие в один и тот же член разложения (17.4.1), не коррелированы и имеют одинаковые дисперсии D_k, получим:

$$K_{k,-k} = \frac{1}{4}\{D_k - D_k - 2i \cdot 0\} = 0.$$

Таким образом, доказано, что разложение (17.4.5) представляет собой не что иное, как каноническое разложение случайной функции $\mathring{X}(t)$ с комплексными координатными функциями $e^{i\omega_k t}$ и комплексными коэффициентами W_k.

Найдем дисперсии этих коэффициентов. При $k = 0$ дисперсия D_0 осталась, очевидно, такой же, как была при действительной форме

спектрального разложения. Дисперсия каждой из комплексных величин W_k (при $k \neq 0$) равна сумме дисперсий ее действительной и мнимой частей:

$$D[W_k] = \frac{D[U_k]}{4} + \frac{D[V_k]}{4} = \frac{2D[U_k]}{4} = \frac{D_k}{2}.$$

Введем обозначение:

$$D_k^* = \frac{D_k}{2} \text{ при } k \neq 0; \; D_k^* = D_0 \text{ при } k = 0$$

и построим дискретный спектр случайной функции $\mathring{X}(t)$, распространенный на частоты от $-\infty$ до $+\infty$ (рис. 17.4.1).

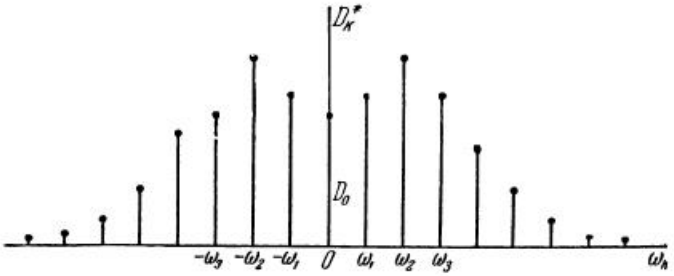


Рис. 17.4.1.

Этот спектр симметричен относительно оси ординат; от ранее построенного спектра (рис. 17.2.3) он отличается тем, что определен не только для положительных, но и для отрицательных частот, но зато его ординаты при $k \neq 0$ вдвое меньше соответствующих ординат прежнего спектра; сумма всех ординат по-прежнему равна дисперсии случайной функции $\mathring{X}(t)$:

$$D_x = \sum_{k=-\infty}^{\infty} D_k. \qquad (17.4.7)$$

Определим корреляционную функцию случайной функции $\mathring{X}(t)$, представленной в виде комплексного спектрального разложения (17.4.5). Применяя формулу (16.2.15) для корреляционной функции комплексной случайной функции, заданной каноническим разложением, имеем:

$$k_x(t, t') = \sum_{k=-\infty}^{\infty} D_k^* e^{i\omega_k t'} \overline{e^{i\omega_k t}} =$$

$$= \sum_{k=-\infty}^{\infty} D_k^* e^{i\omega_k t'} e^{-i\omega_k t} = \sum_{k=-\infty}^{\infty} D_k^* e^{i\omega_k (t'-t)},$$

или, переходя к аргументу $\tau = t' - t$,

$$k_x(\tau) = \sum_{k=-\infty}^{\infty} D_k^* e^{i\omega_k \tau}, \tag{17.4.8}$$

где

$$D_k^* = \frac{1}{2} D_k = \frac{1}{T} \int_0^T k_x(\tau) \cos \omega_k \tau \, d\tau \text{ при } k \neq 0. \tag{17.4.9}$$

Придадим выражению (17.4.9) также комплексную форму. Полагая

$$\cos \omega_k \tau = \frac{e^{i\omega_k \tau} + e^{-i\omega_k \tau}}{2},$$

получим:

$$D_k^* = \frac{1}{2T} \int_0^T k_x(\tau) \left(e^{i\omega_k \tau} + e^{-i\omega_k \tau}\right) d\tau =$$

$$= \frac{1}{2T} \left\{ \int_0^T k_x(\tau) e^{-i\omega_k \tau} d\tau + \int_0^T k_x(\tau) e^{i\omega_k \tau} d\tau \right\}.$$

Полагая во втором интеграле $\tau = -u$, имеем:

$$\int_0^T k_x(\tau) e^{i\omega_k \tau} d\tau = -\int_0^{-T} k_x(u) e^{-i\omega_k u} du = \int_{-T}^0 k_x(\tau) e^{-i\omega_k \tau} d\tau,$$

откуда

$$D_k^* = \frac{1}{2T} \int_{-T}^T k_x(\tau) e^{-i\omega_k \tau} d\tau. \tag{17.4.10}$$

Таким образом, мы построили комплексную форму спектрального разложения случайной функции на конечном интервале $(0, T)$. Далее естественно перейти к пределу при $T \to \infty$, как мы делали для действительной формы, т. е. ввести в рассмотрение спектральную плотность

$$S_x^*(\omega) = \lim_{\Delta\omega \to 0} \frac{D_k^*}{\Delta\omega}$$

и получить в пределе из формул (17.4.8), (17.4.10) интегральные соотношения, связывающие корреляционную функцию и спектральную плотность в комплексной форме. В пределе при $T \to \infty$ формулы (17.4.8)

и (17.4.10) принимают вид:

$$k_x(\tau) = \int_{-\infty}^{\infty} S_x^*(\omega)\, e^{i\omega\tau}\, d\omega, \tag{17.4.11}$$

$$S_x^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_x(\tau)\, e^{-i\omega\tau}\, d\tau. \tag{17.4.12}$$

Формулы (17.4.11) и (17.4.12) представляют собой комплексную форму преобразований Фурье, связывающих корреляционную функцию и спектральную плотность [1]).

Формулы (17.4.11) и (17.4.12) могут быть и непосредственно получены из формул (17.3.9) и (17.3.10), если произвести в них замену

$$\cos\omega\tau = \frac{e^{i\omega\tau} + e^{-i\omega\tau}}{2},$$

положить $S_x(\omega) = 2S_x^*(\omega)$ и расширить область интегрирования на интервал от $-\infty$ до $+\infty$.

Полагая в формуле (17.4.11) $\tau = 0$, получим выражение дисперсии случайной функции $\overset{\circ}{X}(t)$:

$$D_x = \int_{-\infty}^{\infty} S_x^*(\omega)\, d\omega. \tag{17.4.13}$$

Формула (17.4.13) выражает дисперсию случайной функции в виде суммы элементарных дисперсий, распределенных с некоторой плотностью по всему диапазону частот от $-\infty$ до $+\infty$.

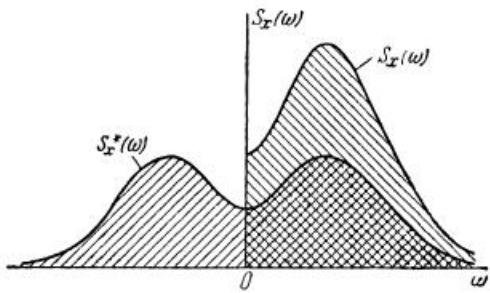


Рис. 17.4.2.

Сравнивая формулу (17.4.13) и ранее выведенную (для действительной формы спектрального разложения) формулу (17.3.2), мы видим, что они различаются лишь тем, что в формуле (17.4.13) стоит несколько иная функция спектральной плотности $S_x^*(\omega)$, определенная не от 0 до ∞, а от $-\infty$ до ∞, но зато с вдвое меньшими ординатами. Если изобразить обе функции спектральной плотности на

[1]) Заметим, что в главе 13, рассматривая характеристические функции, мы уже встречались с преобразованиями такого типа, а именно: характеристическая функция и плотность вероятности выражаются одна через другую с помощью преобразований Фурье.

графике, они различаются только масштабом по оси ординат и тем, что функция $S_x(\omega)$ для отрицательных частот не определена (рис. 17.4.2). На практике в качестве спектральной плотности применяются как та, так и другая функции.

Иногда в качестве аргумента спектральной плотности рассматривают не круговую частоту (ω), а частоту колебаний f, выраженную в герцах:

$$f = \frac{\omega}{2\pi}.$$

В этом случае подстановкой $\omega = 2\pi f$ формула (17.4.11) приводится к виду:

$$k_x(\tau) = 2\pi \int_{-\infty}^{\infty} S_x^*(2\pi f)\, e^{2\pi i f \tau}\, df,$$

или, вводя обозначение

$$G_x(f) = 2\pi S_x^*(2\pi f),$$

$$k_x(\tau) = \int_{-\infty}^{\infty} G_x(f)\, e^{2\pi i f \tau}\, df. \qquad (17.4.14)$$

Функция $G_x(f)$ также может применяться как спектральная плотность дисперсии. Ее выражение через корреляционную функцию, очевидно, имеет вид:

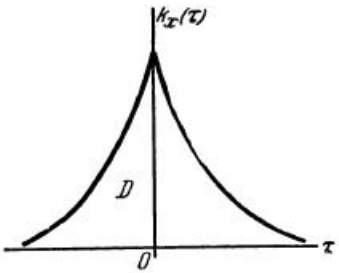


Рис. 17.4.3.

$$G_x(f) = \int_{-\infty}^{\infty} k_x(\tau)\, e^{-2\pi i f \tau}\, d\tau. \qquad (17.4.15)$$

Все приведенные нами и некоторые другие применяемые на практике выражения спектральной плотности, очевидно, отличаются друг от друга только масштабом. Каждую из них можно нормировать, деля соответствующую функцию спектральной плотности на дисперсию случайной функции.

Пример 1. Корреляционная функция случайной функции $X(t)$ задана формулой:

$$k_x(\tau) = De^{-\alpha|\tau|}, \qquad (17.4.16)$$

где $\alpha > 0$ (рис. 17.4.3).

Пользуясь комплексной формой преобразования Фурье, определить спектральную плотность $S_x^*(\omega)$.

Р е ш е н и е. По формуле (17.4.12) находим:

$$S_x^*(\omega) = \frac{D}{2\pi}\int_{-\infty}^{\infty} e^{-\alpha|\tau|}e^{-i\omega\tau}\,d\tau =$$

$$= \frac{D}{2\pi}\left\{\int_{-\infty}^{0} e^{\alpha\tau}e^{-i\omega\tau}\,d\tau + \int_{0}^{\infty} e^{-\alpha\tau}e^{-i\omega\tau}\,d\tau\right\} =$$

$$= \frac{D}{2\pi}\left\{\int_{0}^{\infty} e^{-(\alpha-i\omega)\tau}\,d\tau + \int_{0}^{\infty} e^{-(\alpha+i\omega)\tau}\,d\tau\right\} =$$

$$= \frac{D}{2\pi}\left[\frac{-1}{\alpha-i\omega}e^{-(\alpha-i\omega)\tau} + \frac{-1}{\alpha+i\omega}e^{-(\alpha+i\omega)\tau}\right]_0^{\infty} =$$

$$= \frac{D}{2\pi}\left[\frac{1}{\alpha-i\omega} + \frac{1}{\alpha+i\omega}\right] = \frac{D\alpha}{\pi(\alpha^2+\omega^2)}.$$

График спектральной плотности

$$S_x^*(\omega) = \frac{D\alpha}{\pi(\alpha^2+\omega^2)}$$

представлен на рис. 17.4.4.

Посмотрим, как будут вести себя корреляционная функция и спектральная плотность при изменении α.

При уменьшении α корреляционная функция будет убывать медленнее; характер изменения случайной функции становится более плавным; соответственно в спектре случайной функции бо́льший удельный вес приобретают малые частоты: кривая спектральной плотности вытягивается вверх, одновременно сжимаясь с боков; в пределе при $\alpha \to 0$ случайная функция выродится в обычную случайную величину с дискретным спектром, состоящим из единственной линии с частотой $\omega_0 = 0$.

При увеличении α корреляционная функция убывает быстрее, характер колебаний случайной функции становится более резким и беспорядочным; соответственно этому в спектре случайной функции преобладание малых частот становится все менее выраженным; при $\alpha \to \infty$ спектр случайной функции приближается к равномерному (так называемому «белому») спектру, в котором нет преобладания каких-либо частот.

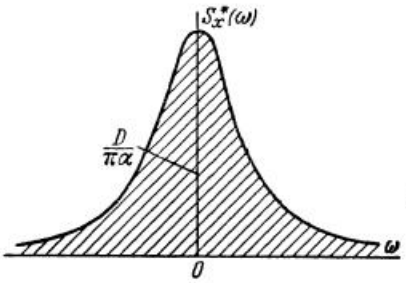


Рис. 17.4.4.

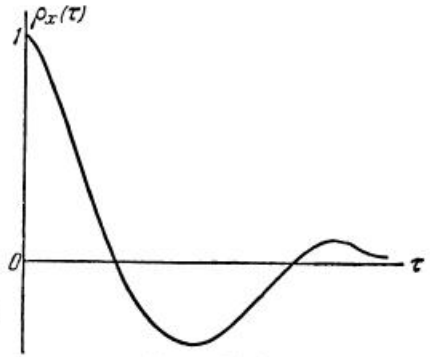


Рис. 17.4.5.

Пример 2. Нормированная корреляционная функция случайной функции $\mathring{X}(t)$ имеет вид:

$$\rho_x(\tau) = e^{-\alpha|\tau|}\cos\beta\tau$$

(рис. 17.4.5).

Определить нормированную спектральную плотность.

Решение. Представляем $\rho_x(\tau)$ в комплексной форме:

$$\rho_x(\tau) = e^{-\alpha|\tau|}\frac{e^{i\beta\tau}+e^{-i\beta\tau}}{2}.$$

Нормированную спектральную плотность $s_x^*(\omega)$ находим по формуле (17.4.12), подставляя в нее $\rho_x(\tau)$ вместо $k_x(\tau)$:

$$s_x^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\alpha|\tau|}\frac{e^{i\beta\tau}+e^{-i\beta\tau}}{2}e^{-i\omega\tau}\,d\tau =$$

$$= \frac{1}{4\pi}\left\{\int_{-\infty}^{0} e^{\alpha\tau}(e^{i\beta\tau}+e^{-i\beta\tau})e^{-i\omega\tau}\,d\tau + \int_{0}^{\infty} e^{-\alpha\tau}(e^{i\beta\tau}+e^{-i\beta\tau})e^{-i\omega\tau}\,d\tau\right\},$$

откуда после элементарных преобразований получаем:

$$s_x^*(\omega) = \frac{1}{2\pi}\left\{\frac{\alpha}{\alpha^2+(\omega+\beta)^2}+\frac{\alpha}{\alpha^2+(\omega-\beta)^2}\right\}.$$

Вид графика спектральной плотности зависит от соотношения параметров α и β, т. е. от того, что преобладает в корреляционной функции: убывание по закону $e^{-\alpha|\tau|}$ или колебание по закону $\cos\beta\tau$. Очевидно, при сравнительно малых α преобладает колебание, при сравнительно больших α — убывание. В первом случае случайная функция близка к периодическим колебаниям частоты β со случайной амплитудой и фазой; соответственно в спектре случайной функции преобладают частоты, близкие к частоте β. Во втором случае спектральный состав случайной функции более равномерен, преобладания тех или иных частот не наблюдается; в пределе при $\alpha\to\infty$ спектр случайной функции приближается к «белому» спектру.

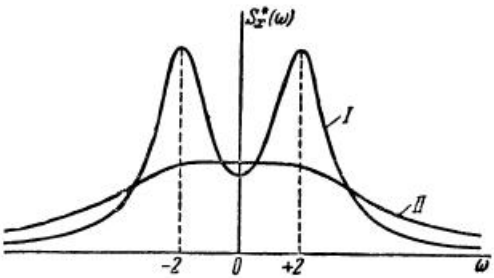


Рис. 17.4.6.

В качестве иллюстрации на рис. 17.4.6 изображены нормированные спектральные плотности для случаев:

1) $\beta=2$, $\alpha=1$ (кривая *I*); 2) $\beta=2$, $\alpha=3$ (кривая *II*). Как видно из чертежа, при $\alpha=1$ спектр случайной функции обнаруживает ярко выраженный максимум в области частот $\omega=\pm\beta$. При $\alpha=3$ (кривая *II*) спектральная плотность в значительном диапазоне частот остается почти постоянной.

17.5. Преобразование стационарной случайной функции стационарной линейной системой

В главе 16 мы познакомились с общими правилами линейных преобразований случайных функций, представленных в виде канонических разложений. Эти правила сводятся к тому, что при линейных преобразованиях случайных функций их математические ожидания и координатные функции подвергаются тем же линейным преобразованиям. Таким образом, задача линейного преобразования случайной функции сводится к задаче такого же линейного преобразования нескольких неслучайных функций.

В случае, когда речь идет о линейных преобразованиях стационарных случайных функций, задачу удается упростить еще больше. Если и входное воздействие $\mathring{X}(t)$ и реакция системы $\mathring{Y}(t)$ стационарны, задачу преобразования случайной функции можно свести к преобразованию одной-единственной неслучайной функции — спектральной плотности $S_x(\omega)$.

Для того чтобы при стационарном воздействии реакция системы могла быть тоже стационарной, очевидно необходимо, чтобы параметры системы (например, входящие в нее сопротивления, емкости, индуктивности и т. п.) были постоянными, а не переменными. Условимся называть линейную систему с постоянными параметрами *стационарной линейной системой*. Обычно работа стационарной линейной системы описывается линейными дифференциальными уравнениями с постоянными коэффициентами.

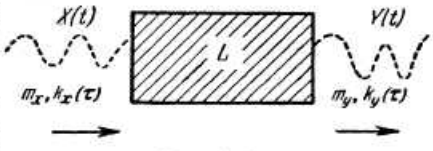


Рис. 17.5.1.

Рассмотрим задачу о преобразовании стационарной случайной функции стационарной линейной системой. Пусть на вход линейной системы L поступает стационарная случайная функция $X(t)$; реакция системы есть случайная функция $Y(t)$ рис. (17.5.1). Известны характеристики случайной функции $X(t)$: математическое ожидание m_x и корреляционная функция $k_x(\tau)$. Требуется определить характеристики случайной функции $Y(t)$ на выходе линейной системы.

Так как для решения задачи нам придется преобразовывать неслучайные функции — математическое ожидание и координатные функции, рассмотрим прежде всего задачу об определении реакции системы L на неслучайное воздействие $x(t)$.

Напишем в операторной форме линейное дифференциальное уравнение с постоянными коэффициентами, связывающее реакцию

системы $y(t)$ с воздействием $x(t)$:

$$(a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_0)\, y(t) = \\ = (b_m p^m + b_{m-1} p^{m-1} + \ldots + b_1 p + b_0)\, x(t), \quad (17.5.1)$$

где $p = \frac{d}{dt}$ — оператор дифференцирования.

Уравнение (17.5.1) короче можно записать в виде:

$$A_n(p)\, y(t) = B_m(p)\, x(t), \quad (17.5.2)$$

или, наконец, условно разрешая уравнение (17.5.2) относительно $y(t)$, записать оператор системы в «явном» виде:

$$y(t) = \frac{B_m(p)}{A_n(p)}\, x(t). \quad (17.5.3)$$

Реакцию системы L на воздействие $x(t)$ можно найти путем решения линейного дифференциального уравнения (17.5.1). Как известно из теории дифференциальных уравнений, это решение состоит из двух слагаемых: $y_{\text{I}}(t)$ и $y_{\text{II}}(t)$. Слагаемое $y_{\text{II}}(t)$ представляет собой решение уравнения без правой части и определяет так называемые *свободные* или *собственные* колебания системы. Это — колебания, совершаемые системой при отсутствии входного воздействия, если система в начальный момент как-то была выведена из состояния равновесия. На практике чаще всего встречаются так называемые *устойчивые системы*; в этих системах свободные колебания с течением времени затухают.

Если ограничиться рассмотрением участков времени, достаточно удаленных от начала процесса, когда все переходные процессы в системе можно считать законченными, и система работает в установившемся режиме, можно отбросить второе слагаемое $y_{\text{II}}(t)$ и ограничиться рассмотрением только первого слагаемого $y_{\text{I}}(t)$. Это первое слагаемое определяет так называемые *вынужденные* колебания системы под влиянием воздействия на нее заданной функции $x(t)$.

В случае, когда воздействие $x(t)$ представляет собой достаточно простую аналитическую функцию, часто удается найти реакцию системы также в виде простой аналитической функции. В частности, когда воздействие представляет собой гармоническое колебание определенной частоты, система отвечает на него также гармоническим колебанием той же частоты, но измененным по амплитуде и фазе.

Так как координатные функции спектрального разложения стационарной случайной функции $X(t)$ представляют собой гармонические колебания, то нам прежде всего необходимо научиться определять реакцию системы на гармоническое колебание заданной частоты ω. Эта задача решается очень просто, особенно если гармоническое колебание представлено в комплексной форме.

Пусть на вход системы поступает гармоническое колебание вида:

$$x(t) = e^{i\omega t}. \tag{17.5.4}$$

Будем искать реакцию системы $y(t)$ также в виде гармонического колебания частоты ω, но умноженного на некоторый комплексный множитель $\Phi(i\omega)$:

$$y(t) = \Phi(i\omega)\, e^{i\omega t}. \tag{17.5.5}$$

Множитель $\Phi(i\omega)$ найдем следующим образом. Подставим функцию (17.5.4) в правую, а функцию (17.5.5) в левую часть уравнения (17.5.1). Получим:

$$a_n \frac{d^n}{dt^n}[\Phi(i\omega)\, e^{i\omega t}] + a_{n-1}\frac{d^{n-1}}{dt^{n-1}}[\Phi(i\omega)\, e^{i\omega t}] + \ldots$$

$$\ldots + a_1 \frac{d}{dt}[\Phi(i\omega)\, e^{i\omega t}] + a_0[\Phi(i\omega)\, e^{i\omega t}] =$$

$$= b_m \frac{d^m}{dt^m} e^{i\omega t} + b_{m-1}\frac{d^{m-1}}{dt^{m-1}} e^{i\omega t} + \ldots + b_1 \frac{d}{dt} e^{i\omega t} + b_0 e^{i\omega t}. \tag{17.5.6}$$

Имея в виду, что при любом k

$$\frac{d^k}{dt^k} e^{i\omega t} = (i\omega)^k e^{i\omega t}, \qquad \frac{d^k}{dt^k}\left[\Phi(i\omega)\, e^{i\omega t}\right] = (i\omega)^k e^{i\omega t}\Phi(i\omega),$$

и деля обе части уравнения (17.5.6) на $e^{i\omega t}$, получим:

$$\Phi(i\omega)\left[a_n(i\omega)^n + a_{n-1}(i\omega)^{n-1} + \ldots + a_1(i\omega) + a_0\right] =$$

$$= b_m(i\omega)^m + b_{m-1}(i\omega)^{m-1} + \ldots + b_1(i\omega) + b_0. \tag{17.5.7}$$

Мы видим, что множитель при $\Phi(i\omega)$ представляет собой не что иное, как многочлен $A_n(p) = a_n p^n + a_{n-1}p^{n-1} + \ldots + a_1 p + a_0$, в который вместо оператора дифференцирования p подставлено $(i\omega)$; аналогично правая часть равенства (17.5.7) есть не что иное, как $B_m(i\omega)$. Уравнение (17.5.7) можно записать в виде:

$$\Phi(i\omega)\, A_n(i\omega) = B_m(i\omega),$$

откуда

$$\Phi(i\omega) = \frac{B_m(i\omega)}{A_n(i\omega)}. \tag{17.5.8}$$

Функция $\Phi(i\omega)$ носит специальное название *частотной характеристики* линейной системы. Для определения частотной характеристики достаточно в оператор системы, записанный в явной форме (17.5.3), вместо оператора дифференцирования p подставить $i\omega$.

Таким образом, *если на вход линейной системы с постоянными параметрами поступает гармоническое колебание вида* $e^{i\omega t}$, *то реакция системы представляется в виде того же гармонического колебания, умноженного на частотную характеристику системы* $\Phi(i\omega)$. Пусть на вход системы поступает воздействие вида

$$x(t) = Ue^{i\omega t}, \tag{17.5.9}$$

где U — некоторая величина, не зависящая от t. В силу линейности системы величина U выходит за знак оператора, и реакция системы на воздействие (17.5.9) будет равна:

$$y(t) = U\Phi(i\omega)\, e^{i\omega t}. \tag{17.5.10}$$

Очевидно, это свойство сохранится и в том случае, когда величина U будет случайной (лишь бы она не зависела от t).

Применим изложенные приемы преобразования гармонических колебаний линейной системой к математическому ожиданию случайной функции $X(t)$ и координатным функциям ее спектрального разложения.

Представим математическое ожидание m_x стационарной случайной функции $X(t)$ как гармоническое колебание нулевой частоты $\omega = 0$ и положим в формуле (17.5.8) $\omega = 0$:

$$\Phi(0) = \frac{B_m(0)}{A_n(0)} = \frac{b_0}{a_0}, \tag{17.5.11}$$

откуда получаем математическое ожидание на выходе системы:

$$m_y = \frac{b_0}{a_0} m_x. \tag{17.5.12}$$

Перейдем к преобразованию линейной системой существенно случайной части функции $X(t)$, а именно функции

$$\mathring{X}(t) = X(t) - m_x. \tag{17.5.13}$$

Для этого представим функцию $\mathring{X}(t)$ на участке $(0, T)$ в виде спектрального разложения:

$$\mathring{X}(t) = \sum_{k=-\infty}^{\infty} U_k\, e^{i\omega_k t}, \tag{17.5.14}$$

где U_k — некоррелированные случайные величины, дисперсии которых образуют спектр случайной функции $X(t)$.

Рассмотрим отдельное слагаемое этой суммы:

$$X_k(t) = U_k\, e^{i\omega_k t}. \tag{17.5.15}$$

Реакция системы на это воздействие будет иметь вид:

$$Y_k(t) = U_k\, \Phi(i\omega_k)\, e^{i\omega_k t}. \tag{17.5.16}$$

Согласно принципу суперпозиции реакция системы на сумму воздействий равна сумме реакций на отдельные воздействия. Следовательно, реакцию системы на воздействие (17.5.14) можно представить в виде спектрального разложения:

$$\mathring{Y}(t) = \sum_{k=-\infty}^{\infty} U_k\, \Phi(i\omega_k)\, e^{i\omega_k t},$$

или, обозначая $U_k\Phi(i\omega_k) = W_k$,

$$\overset{\circ}{Y}(t) = \sum_{k=-\infty}^{\infty} W_k e^{i\omega_k t}, \qquad (17.5.17)$$

где W_k — некоррелированные случайные величины с математическими ожиданиями, равными нулю.

Определим спектр этого разложения. Для этого найдем дисперсию комплексной случайной величины W_k в разложении (17.5.17). Имея в виду, что дисперсия комплексной случайной величины равна математическому ожиданию квадрата ее модуля, имеем:

$$D[W_k] = M[|U_k\Phi(i\omega_k)|^2] = M[|U_k|^2|\Phi(i\omega_k)|^2] =$$
$$= |\Phi(i\omega_k)|^2 M[|U_k|^2] = |\Phi(i\omega_k)|^2 D_k. \qquad (17.5.18)$$

Мы приходим к следующему выводу: *при преобразовании стационарной случайной функции стационарной линейной системой каждая из ординат ее спектра умножается на квадрат модуля частотной характеристики системы для соответствующей частоты.*

Таким образом, при прохождении стационарной случайной функции через линейную стационарную систему ее спектр определенным образом перестраивается: некоторые частоты усиливаются, некоторые, напротив, ослабляются (фильтруются). Квадрат модуля частотной характеристики (в зависимости от ω_k) и показывает, как реагирует система на колебания той или иной частоты.

Аналогично тому, как это делалось раньше, перейдем в спектральном представлении случайной функции к пределу при $T \to \infty$ и от дискретного спектра — к спектральной плотности. Очевидно, спектральная плотность на выходе линейной системы получается из спектральной плотности на входе тем же умножением на $|\Phi(i\omega)|^2$, как и ординаты дискретного спектра:

$$S_y(\omega) = |\Phi(i\omega)|^2 S_x(\omega). \qquad (17.5.19)$$

Таким образом, получено весьма простое правило:

При преобразовании стационарной случайной функции стационарной линейной системой ее спектральная плотность умножается на квадрат модуля частотной характеристики системы.

Пользуясь этим правилом, мы легко можем решить поставленную выше задачу: по характеристикам случайной функции на входе линейной системы найти характеристики случайной функции на ее выходе.

Пусть на вход стационарной линейной системы с оператором (17.5.3) поступает стационарная случайная функция $X(t)$ с математическим ожиданием m_x и корреляционной функцией $k_x(\tau)$. Требуется найти математическое ожидание m_y и корреляционную функцию $k_y(\tau)$ случайной функции $Y(t)$ на выходе системы.

Задачу будем решать в следующем порядке.

1. Находим математическое ожидание на выходе:

$$m_y = \frac{b_0}{a_0} m_x. \tag{17.5.20}$$

2. По корреляционной функции $k_x(\tau)$ находим спектральную плотность на входе (см. формулу (17.4.12)):

$$S_x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_x(\tau) e^{-i\omega\tau} d\tau \text{ [1]}. \tag{17.5.21}$$

3. По формуле (17.5.8) находим частотную характеристику системы и квадрат ее модуля:

$$|\Phi(i\omega)|^2 = \frac{|B_m(i\omega)|^2}{|A_n(i\omega)|^2}. \tag{17.5.22}$$

4. Умножая спектральную плотность на входе на квадрат модуля частотной характеристики, находим спектральную плотность на выходе:

$$S_y(\omega) = |\Phi(i\omega)|^2 S_x(\omega). \tag{17.5.23}$$

5. По спектральной плотности $S_y(\omega)$ находим корреляционную функцию $k_y(\tau)$ на выходе системы:

$$k_y(\tau) = \int_{-\infty}^{\infty} S_y(\omega) e^{i\omega\tau} d\omega. \tag{17.5.24}$$

Таким образом, поставленная задача решена.

Во многих задачах практики нас интересует не вся корреляционная функция $k_y(\tau)$ на выходе системы, а только дисперсия D_y, равная

$$D_y = k_y(0).$$

Тогда из формулы (17.5.24) получаем при $\tau = 0$ гораздо более простую формулу:

$$D_y = \int_{-\infty}^{\infty} S_y(\omega) d\omega,$$

или, учитывая четность функции $S_y(\omega)$,

$$D_y = 2\int_0^{\infty} S_y(\omega) d\omega. \tag{17.5.25}$$

---

[1]) Для простоты записи мы здесь опускаем знак * в обозначении спектральной плотности.

Пример. Работа линейной динамической системы описывается линейным дифференциальным уравнением первого порядка:

$$(a_1 p + a_0)\, y(t) = (b_1 p + b_0)\, x(t), \tag{17.5.26}$$

или

$$y(t) = \frac{b_1 p + b_0}{a_1 p + a_0}\, x(t).$$

На вход системы поступает стационарная случайная функция $X(t)$ с математическим ожиданием m_x и корреляционной функцией

$$k_x(\tau) = D_x e^{-\alpha|\tau|}, \tag{17.5.27}$$

где α — положительный коэффициент (см. пример 1 n° 17.4). Найти математическое ожидание m_y и дисперсию D_y на выходе системы [1]).

Решение. По формуле (17.5.20) имеем:

$$m_y = \frac{b_0}{a_0} m_x.$$

Очевидно, величина m_y не зависит от параметра α, растет при возрастании b_0 и убывает при возрастании a_0.

Спектральную плотность на входе определяем как в примере 1 n° 17.4:

$$S_x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_x(\tau)\, e^{-i\omega\tau}\, d\tau = \frac{D_x \alpha}{\pi(\alpha^2 + \omega^2)}$$

(см. рис. 17.4.4).

По формуле (17.5.8) находим частотную характеристику системы:

$$\Phi(i\omega) = \frac{b_1 i\omega + b_0}{a_1 i\omega + a_0}$$

и квадрат ее модуля:

$$|\Phi(i\omega)|^2 = \frac{|b_1 i\omega + b_0|^2}{|a_1 i\omega + a_0|^2} = \frac{b_1^2\omega^2 + b_0^2}{a_1^2\omega^2 + a_0^2}.$$

Затем определяем спектральную плотность на выходе системы:

$$S_y(\omega) = |\Phi(i\omega)|^2 S_x(\omega) = \frac{D_x}{\pi} \frac{b_1^2\omega^2 + b_0^2}{a_1^2\omega^2 + a_0^2} \frac{\alpha}{\alpha^2 + \omega^2}.$$

Далее по формуле (17.5.25) определяем дисперсию на выходе:

$$D_y = \frac{2D_x}{\pi} \int_0^{\infty} \frac{b_1^2\omega^2 + b_0^2}{a_1^2\omega^2 + a_0^2} \frac{\alpha}{\alpha^2 + \omega^2}\, d\omega.$$

---

[1]) Выбирая для корреляционной функции случайной функции $X(t)$ выражение типа (17.5.27), широко применяемое на практике ввиду его простоты, необходимо иметь в виду следующее. Строго говоря, случайная функция $X(t)$, имеющая корреляционную функцию такого вида, недифференцируема, и, следовательно, для нее нельзя писать дифференциальные уравнения в обычном смысле слова. Это затруднение можно обойти, если рассматривать выражение (17.5.27) для корреляционной функции только как приближенное.

Для вычисления интеграла разложим подынтегральное выражение на простые дроби:

$$\frac{b_1^2\omega^2+b_0^2}{a_1^2\omega^2+a_0^2}\,\frac{\alpha}{\alpha^2+\omega^2}=\frac{A}{a_0^2+a_1^2\omega^2}+\frac{B}{\alpha^2+\omega^2}$$

и определим коэффициенты:

$$A=\alpha\frac{a_1^2b_0^2-a_0^2b_1^2}{a_1^2\alpha^2-a_0^2};$$

$$B=\alpha\frac{b_1^2\alpha^2-b_0^2}{a_1^2\alpha^2-a_0^2}.$$

После интегрирования получим:

$$D_y=D_x\frac{a_1b_0^2+a_0b_1^2\alpha}{a_0a_1(a_1\alpha+a_0)}.$$

В заключение данного п° упомянем о том, как преобразуется линейной системой стационарная случайная функция, содержащая в качестве слагаемого обычную случайную величину:

$$X_1(t)=U_0+X(t), \tag{17.5.28}$$

где U_0 — случайная величина с дисперсией D_0, $X(t)$ — стационарная случайная функция.

Реакция системы на воздействие $X_1(t)$ найдется как сумма реакций на отдельные воздействия в правой части (17.5.28). Реакцию на воздействие $X(t)$ мы уже умеем находить. Воздействие U_0 мы рассмотрим как гармоническое колебание нулевой частоты $\omega=0$; согласно формуле (17.5.11) реакция на него будет равна

$$V_0=\frac{b_0}{a_0}U_0. \tag{17.5.29}$$

Слагаемое V_0 просто прибавится к реакции системы на воздействие $X(t)$.

17.6. Применения теории стационарных случайных процессов к решению задач, связанных с анализом и синтезом динамических систем

В предыдущем п° был рассмотрен вопрос о преобразовании стационарной случайной функции стационарной линейной системой и получены простые математические приемы решения этой задачи. Преобразование случайной функции свелось к простейшему преобразованию (умножению на квадрат модуля частотной характеристики) одной-единственной функции: спектральной плотности. Такая простота спектральной теории стационарных случайных процессов делает ее незаменимым аппаратом при исследовании линейных динамических систем, работающих в условиях наличия случайных возмущений (помех).

Обычно при решении практических задач нас интересует не сама по себе корреляционная функция $k_y(\tau)$ на выходе системы, а связанная с нею дисперсия

$$D_y = k_y(0),$$

которая характеризует ошибки системы, вызванные поступающими на нее случайными возмущениями, и во многих случаях может служить критерием точности работы системы.

При исследовании динамических систем методами теории случайных функций решаются два вида задач, которые можно назвать «прямыми» и «обратными».

Прямая задача состоит в следующем. Анализируется заданная линейная динамическая система с вполне определенными параметрами, работа которой описывается линейным дифференциальным уравнением:

$$(a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_0)\, y(t) = \\ = (b_m p^m + b_{m-1} p^{m-1} + \ldots + b_1 p + b_0)\, x(t). \qquad (17.6.1)$$

Требуется исследовать точность работы системы при наличии на ее входе стационарного случайного воздействия — так называемой «стационарной помехи». Для этого прежде всего исследуется случайная помеха, определяются ее корреляционная функция и спектральный состав. Далее, описанными выше методами находятся спектр и дисперсия случайной функции на выходе системы. Дисперсия на выходе, очевидно, зависит как от характеристик случайного воздействия на входе, так и от коэффициентов уравнения. Решая такую задачу, можно оценить точность работы заданной системы в условиях различного рода помех.

Обратная задача состоит в том, чтобы так выбрать коэффициенты уравнения (17.6.1), чтобы при заданном спектральном составе помехи ошибки на выходе системы были минимальными. При заданных характеристиках случайной функции (помехи) на входе системы дисперсия на выходе зависит от всей совокупности коэффициентов уравнения:

$$D_y = D_y(a_n,\ a_{n-1},\ \ldots,\ a_1,\ a_0,\ b_m,\ b_{m-1},\ \ldots,\ b_1,\ b_0).$$

Коэффициенты уравнения зависят от конструктивных параметров системы, и некоторыми из них при проектировании системы можно в достаточно широких пределах распоряжаться. Задача выбора рациональных значений этих параметров может быть решена исходя из того требования, чтобы дисперсия D_y была минимальна.

Следует оговориться, что на практике часто не удается полностью удовлетворить этому требованию. Действительно, выведенные нами выражения для корреляционной функции и дисперсии на выходе системы справедливы только для значений времени t, достаточно удаленных от начала случайного процесса, когда все переходные процессы в системе, связанные с ее свободными колебаниями, успели

уже затухнуть. В действительности же часто приходится применять линейные динамические системы (прицелы, счетно-решающие механизмы, следящие системы и т. п.) на ограниченном участке времени; при этом быстрота затухания переходных процессов в системе существенно зависит от ее конструктивных параметров, т. е. от тех же коэффициентов уравнения (17.6.1). Если выбрать эти коэффициенты так, чтобы они обращали в минимум дисперсию на выходе (для достаточно удаленных моментов времени), это, как правило, приводит к тому, что на выходе системы появляются другие ошибки, связанные с тем, что переходные процессы в системе еще не успели затухнуть. Эти ошибки обычно называют динамическими ошибками.

В связи с ограниченностью времени применения линейных систем и наличием динамических ошибок на практике обычно приходится решать задачу о рациональном выборе параметров системы не на чистом принципе минимума дисперсии, а с учетом динамических ошибок. Рациональное решение задачи находится как компромиссное, при котором, с одной стороны, дисперсия на выходе системы достаточно мала, с другой стороны — динамические ошибки не слишком велики.

В случае, когда ищутся оптимальные параметры системы с учетом как дисперсии, так и систематических динамических ошибок, часто в качестве критерия точности работы системы выбирают второй начальный момент α_2 на выходе системы:

$$\alpha_2 = D_y + m_y^2, \tag{17.6.2}$$

где D_y — дисперсия, m_y — систематическая ошибка на выходе системы. При этом параметры системы выбирают так, чтобы они обращали в минимум величину α_2.

Иногда в качестве критерия для оценки системы выбирают не дисперсию и не второй начальный момент, а какую-либо другую величину, связанную с целевым назначением системы. Например, при исследовании прицельных устройств и систем управления, предназначеннных для стрельбы, к выбору их параметров часто подходят, исходя из максимума вероятности поражения цели.

Упомянем еще об одной типичной задаче, связанной с рациональным конструированием динамических систем. До сих пор мы рассматривали только задачу о рациональном выборе коэффициентов уравнения (17.6.1), самый же вид уравнения считался заданным. При решении задач, связанных с так называемым *синтезом динамических систем*, задача ставится более широко. В частности, ставится вопрос о рациональном выборе самого вида уравнения или, еще шире, задача об определении *оптимального оператора* динамической системы. Такого рода задачи в настоящее время успешно решаются методами теории случайных функций.

При решении практических задач, связанных с анализом и синтезом динамических систем, часто не удается ограничиться кругом

стационарных случайных процессов и относящимся к нему аппаратом спектральной теории. Однако в ряде случаев, несколько видоизменив этот аппарат, можно применить его и для нестационарных процессов. На практике часто встречаются так называемые «квазистационарные» случайные функции и «квазистационарные» динамические системы; они характерны тем, что изменения характеристик случайных функций и параметров системы со временем протекают сравнительно медленно. Для таких случайных процессов В. С. Пугачевым разработан метод, по структуре мало отличающийся от спектрального, но применимый в более широком диапазоне условий [1]).

17.7. Эргодическое свойство стационарных случайных функций

Рассмотрим некоторую стационарную случайную функцию $X(t)$ и предположим, что требуется оценить ее характеристики: математическое ожидание m_x и корреляционную функцию $k_x(\tau)$. Выше (см. п° 15.4) были изложены способы получения этих характеристик из опыта. Для этого нужно располагать известным числом реализаций случайной функции $X(t)$. Обрабатывая эти реализации, можно найти оценки для математического ожидания $\tilde{m}_x(t)$ и корреляционной функции $\tilde{K}_x(t, t')$. В связи с ограниченностью числа наблюдений функция $\tilde{m}_x(t)$ не будет строго постоянной; ее придется осреднить и заменить некоторым постоянным $\tilde{m}_x$; аналогично, осредняя значения $\tilde{K}_x(t, t')$ для разных $\tau = t' - t$, получим корреляционную функцию $\tilde{k}_x(\tau)$.

Этот метод обработки, очевидно, является довольно сложным и громоздким и к тому же состоит из двух этапов: приближенного определения характеристик случайной функции и также приближенного осреднения этих характеристик. Естественно возникает вопрос: нельзя ли для стационарной случайной функции этот сложный, двухступенчатый процесс обработки заменить более простым, который заранее базируется на предположении, что математическое ожидание не зависит от времени, а корреляционная функция — от начала отсчета?

Кроме того, возникает вопрос: при обработке наблюдений над стационарной случайной функцией является ли существенно необходимым располагать *несколькими* реализациями? Поскольку случайный процесс является стационарным и протекает однородно по времени, естественно предположить, что *одна-единственная реализация* достаточной продолжительности может служить достаточным опытным материалом для получения характеристик случайной функции.

---

[1]) См. В. С. Пугачев. Теория случайных функций и ее применение к задачам автоматического управления, Физматгиз, 1962.

При более подробном рассмотрении этого вопроса оказывается, что такая возможность существует не для всех случайных процессов: не всегда одна реализация достаточной продолжительности оказывается эквивалентной множеству отдельных реализаций.

Для примера рассмотрим две стационарные случайные функции $X_1(t)$ и $X_2(t)$, представленные совокупностью своих реализаций на рис. 17.7.1 и 17.7.2.

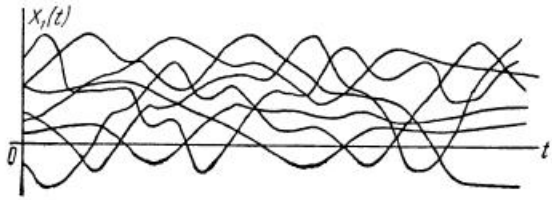


Рис. 17.7.1.

Для случайной функции $X_1(t)$ характерна следующая особенность: каждая из ее реализаций обладает одними и теми же характерными признаками: средним значением, вокруг которого происходят колебания, и средним размахом этих колебаний. Выберем произвольно одну из таких реализаций и продолжим мысленно опыт, в результате которого она получена, на некоторый участок времени T. Очевидно,

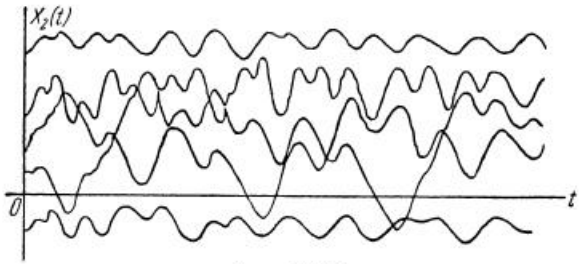


Рис. 17.7.2.

при достаточно большом T эта одна реализация сможет дать нам достаточно хорошее представление о свойствах случайной функции в целом. В частности, осредняя значения этой реализации вдоль оси абсцисс — по времени, мы должны получить приближенное значение математического ожидания случайной функции; осредняя квадраты отклонений от этого среднего, мы должны получить приближенное значение дисперсии, и т. д.

Про такую случайную функцию говорят, что она обладает *эргодическим свойством*. Эргодическое свойство состоит в том, что каждая отдельная реализация случайной функции является как бы «полномочным представителем» всей совокупности возможных реализаций;

одна реализация достаточной продолжительности может заменить при обработке множество реализаций той же общей продолжительности [1]).

Рассмотрим теперь случайную функцию $X_2(t)$. Выберем произвольно одну из ее реализаций, продолжим ее мысленно на достаточно большой участок времени и вычислим ее среднее значение по времени на всем участке наблюдения. Очевидно, это среднее значение для каждой реализации будет свое и может существенно отличаться от математического ожидания случайной функции, построенного как среднее из множества реализаций. Про такую случайную функцию говорят, что она не обладает эргодическим свойством.

Если случайная функция $X(t)$ обладает эргодическим свойством, то для нее *среднее по времени* (на достаточно большом участке наблюдения) *приближенно равно среднему по множеству наблюдений*. То же будет верно и для $X^2(t)$, $X(t) \cdot X(t+\tau)$ и т. д. Следовательно, все характеристики случайной функции (математическое ожидание, дисперсию, корреляционную функцию) можно будет приближенно определять по одной достаточно длинной реализации.

Какие же стационарные случайные функции обладают, а какие не обладают эргодическим свойством?

Поясним этот вопрос наглядно, исходя из примера. Рассмотрим случайную функцию $\alpha(t)$ — колебания угла атаки самолета на установившемся режиме горизонтального полета. Предположим, что полет происходит в каких-то типичных средних метеорологических условиях. Колебания угла атаки вызваны случайными возмущениями, связанными с турбулентностью атмосферы. Среднее значение угла атаки, около которого происходят колебания, зависит от высоты полета H. Зависит от этой высоты и размах колебаний: известно, что в нижних слоях атмосферы турбулентность сказывается сильнее, чем в верхних.

Рассмотрим случайную функцию $\alpha(t)$ — колебания угла атаки на заданной высоте H. Каждая из реализаций этой случайной функции осуществляется в результате воздействия одной и той же группы случайных факторов и обладает одними и теми же вероятностными характеристиками; случайная функция $\alpha(t)$ обладает эргодическим свойством (рис. 17.7.3).

Представим себе теперь, что рассматривается случайная функция $\alpha(t)$ не для одной высоты H, а для целого диапазона, внутри которого задан какой-то закон распределения высот (например, закон равномерной плотности). Такая случайная функция, оставаясь стационарной, очевидно, уже не будет обладать эргодическим свойством; ее возможные реализации, осуществляющиеся с какими-то вероятностями, имеют различный характер (рис. 17.7.4).

[1]) Строго говоря, следовало бы сказать не «каждая отдельная реализация», а «почти каждая». Дело в том, что в отдельных случаях могут появляться реализации, не обладающие таким свойством, но вероятность появления такой реализации равна нулю.

Для этого случайного процесса характерно то, что он как бы «разложим» на более элементарные случайные процессы; каждый из них осуществляется с некоторой вероятностью и имеет свои индивидуальные характеристики. Таким образом, разложимость,

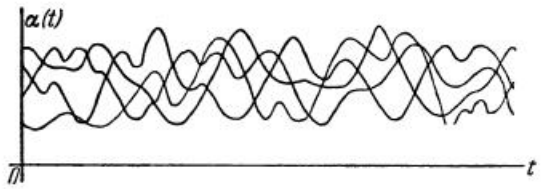


Рис. 17.7.3.

внутренняя неоднородность случайного процесса, протекающего с некоторой вероятностью по тому или другому типу, есть физическая причина неэргодичности этого процесса.

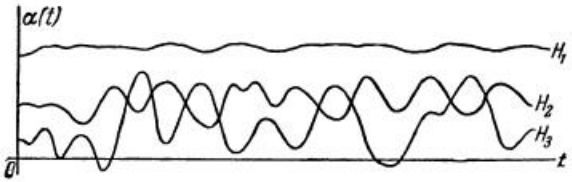


Рис. 17.7.4.

В частности, неэргодичность случайного процесса может быть связана с наличием в его составе слагаемого в виде обычной случайной величины (т. е. наличие в спектре случайного процесса, помимо непрерывной части, конечной дисперсии при частоте 0).

Действительно, рассмотрим случайную функцию

$$Z(t) = X(t) + Y, \tag{17.7.1}$$

где $X(t)$ — эргодическая стационарная случайная функция с характеристиками m_x, $k_x(\tau)$, Y — случайная величина с характеристиками m_y и D_y; предположим к тому же, что $X(t)$ и Y некоррелированны.

Определим характеристики случайной функции $Z(t)$. Согласно общим правилам сложения случайных функций (см. п° 15.8) имеем:

$$m_z = m_x + m_y, \tag{17.7.2}$$

$$k_z(\tau) = k_x(\tau) + D_y. \tag{17.7.3}$$

Из формул (17.7.2) и (17.7.3) видно, что случайная функция $Z(t)$ является стационарной. Но обладает ли она эргодическим свойством? Очевидно, нет. Каждая ее реализация будет по характеру отличаться от других, будет обладать тем или иным средним по времени зна-

чением в зависимости от того, какое значение приняла случайная величина Y (рис. 17.7.5).

Об эргодичности или неэргодичности случайного процесса может непосредственно свидетельствовать вид его корреляционной функции. Действительно, рассмотрим корреляционную функцию неэргодической случайной функции (17.7.1). Она отличается от корреляционной функции случайной функции $X(t)$ наличием постоянного слагаемого D_y

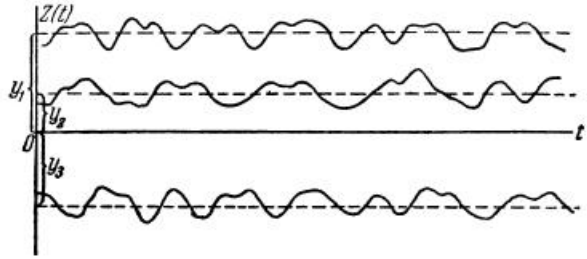


Рис. 17.7.5.

(рис. 17.7.6). В то время как корреляционная функция $k_x(\tau)$ стремится к нулю при $\tau \to \infty$ (корреляционная связь между значениями случайной функции неограниченно убывает по мере увеличения расстояния между ними), функция $k_z(\tau)$ уже не стремится к нулю при $t \to \infty$, а приближается к постоянному значению D_y.

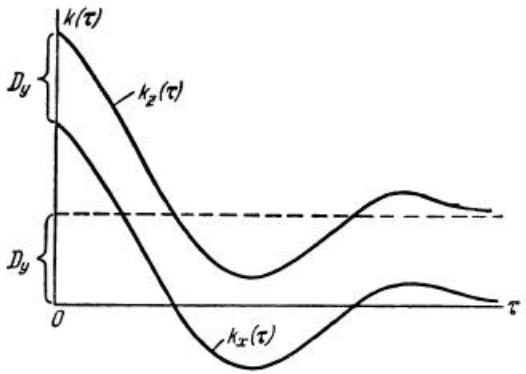


Рис. 17.7.6.

На практике мы не имеем возможности исследовать случайный процесс и его корреляционную функцию на бесконечном участке времени; участок значений τ, с которым мы имеем дело, всегда ограничен. Если при этом корреляционная функция стационарного случайного процесса при увеличении τ не убывает, а, начиная с некоторого τ, остается приблизительно постоянной, это обычно есть признак того, что в составе случайной функции имеется слагаемое в виде обычной случайной величины и что процесс не является эргодическим. Стремление же корреляционной функции к нулю при $t \to \infty$ говорит в пользу эргодичности процесса. Во всяком случае оно

достаточно для того, чтобы математическое ожидание функции можно было определять как среднее по времени.

При решении практических задач часто суждение об эргодичности случайного процесса выносится не на основе исследования поведения корреляционной функции при $\tau \to \infty$, а на основании физических соображений, связанных с существом процесса (его предположительной «разложимостью» или «неразложимостью» на элементарные процессы различного типа, появляющиеся с некоторыми вероятностями).

17.8. Определение характеристик эргодической стационарной случайной функции по одной реализации

Рассмотрим стационарную случайную функцию $X(t)$, обладающую эргодическим свойством, и предположим, что в нашем распоряжении имеется всего одна реализация этой случайной функции, но зато на достаточно большом участке времени T. Для эргодической стационарной случайной функции одна реализация достаточно большой продолжительности практически эквивалентна (в смысле объема сведений о случайной функции) множеству реализаций той же общей продолжительности; характеристики случайной функции могут быть приближенно определены не как средние по множеству наблюдений, а как *средние по времени t*. В частности, при достаточно большом T математическое ожидание m_x может быть приближенно вычислено по формуле

$$m_x \approx \frac{1}{T}\int_0^T x(t)\,dt \text{ [1]}. \qquad (17.8.1)$$

Аналогично может быть приближенно найдена корреляционная функция $k_x(\tau)$ при любом τ. Действительно, корреляционная функция, по определению, представляет собой не что иное, как *математическое ожидание* случайной функции $\mathring{X}(t)\,\mathring{X}(t+\tau)$:

$$k_x(\tau) = M[\mathring{X}(t)\,\mathring{X}(t+\tau)]. \qquad (17.8.2)$$

Это математическое ожидание также, очевидно, может быть приближенно вычислено как среднее по времени.

Фиксируем некоторое значение τ и вычислим указанным способом корреляционную функцию $k_x(\tau)$. Для этого удобно предварительно «центрировать» данную реализацию $x(t)$, т. е. вычесть из нее математическое ожидание (17.8.1):

$$\mathring{x}(t) = x(t) - m_x. \qquad (17.8.3)$$

---

[1]) Для простоты записи мы здесь опускаем знак $\sim$ при характеристиках случайной функции, означающий, что мы имеем дело не с самими характеристиками, а с их оценками.

Вычислим при заданном τ математическое ожидание случайной функции $\mathring{X}(t)\,\mathring{X}(t+\tau)$ как среднее по времени. При этом, очевидно, нам придется учитывать не весь участок времени от 0 до T, а несколько меньший, так как второй сомножитель $\mathring{X}(t+\tau)$ известен нам не для всех t, а только для тех, для которых $t+\tau \leqslant T$.

Вычисляя среднее по времени указанным выше способом, получим:

$$k_x(\tau) \approx \frac{1}{T-\tau}\int_0^{T-\tau} \mathring{x}(t)\,\mathring{x}(t+\tau)\,dt. \qquad (17.8.4)$$

Вычислив интеграл (17.8.4) для ряда значений τ, можно приближенно воспроизвести по точкам весь ход корреляционной функции.

На практике обычно интегралы (17.8.1) и (17.8.4) заменяют конечными суммами. Покажем, как это делается. Разобьем интервал

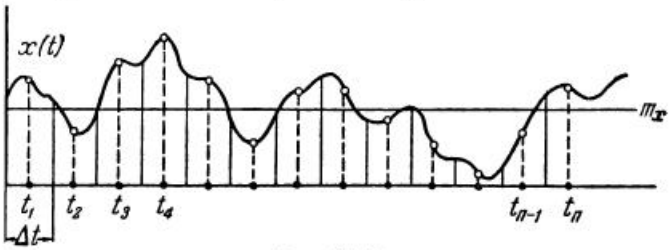


Рис. 17.8.1.

записи случайной функции на n равных частей длиной Δt и обозначим середины полученных участков $t_1, t_2, \ldots, t_n$ (рис. 17.8.1). Представим интеграл (17.8.1) как сумму интегралов по элементарным участкам Δt и на каждом из них вынесем функцию $x(t)$ из-под знака интеграла средним значением, соответствующим центру интервала $x(t_i)$. Получим приближенно:

$$m_x = \frac{1}{T}\frac{T}{n}\sum_{i=1}^{n} x(t_i),$$

или

$$m_x = \frac{1}{n}\sum_{i=1}^{n} x(t_i). \qquad (17.8.5)$$

Аналогично можно вычислить корреляционную функцию для значений τ, равных $0, \Delta t, 2\Delta t, \ldots$. Придадим, например, величине τ значение

$$\tau = m\,\Delta t = \frac{mT}{n}$$

и вычислим интеграл (17.8.4), деля интервал интегрирования

$$T-\tau = T - \frac{mT}{n} = \frac{n-m}{n}T$$

на $n-m$ равных участков длиной Δt и вынося на каждом из них функцию $\mathring{x}(t)\,\mathring{x}(t+\tau)$ за знак интеграла средним значением. Получим:

$$k_x\left(\frac{mT}{n}\right)=\frac{n}{(n-m)\,T}\,\frac{T}{n}\sum_{i=1}^{n-m}\mathring{x}(t_i)\,\mathring{x}(t_{i+m}),$$

или окончательно

$$k_x\left(\frac{mT}{n}\right)=\frac{1}{n-m}\sum_{i=1}^{n-m}\mathring{x}(t_i)\,\mathring{x}(t_{i+m}). \qquad (17.8.6)$$

Вычисление корреляционной функции по формуле (17.8.6) производят для $m=0,\ 1,\ 2,\ \ldots$ последовательно вплоть до таких значений m, при которых корреляционная функция становится практически равной нулю или начинает совершать небольшие нерегулярные колебания около нуля. Общий ход функции $k_x(\tau)$ воспроизводится по отдельным точкам (рис. 17.8.2).

Для того чтобы математическое ожидание m_x и корреляционная функция $k_x(\tau)$ были определены с удовлетворительной точностью, нужно, чтобы число точек n было достаточно велико (порядка сотни, а в некоторых случаях даже нескольких сотен). Выбор длины элементарного участка Δt определяется характером изменения случайной функции. Если случайная функция изменяется сравнительно плавно, участок Δt можно выбирать бóльшим, чем когда она совершает резкие и частые колебания. Чем более высокочастотный состав имеют колебания, образующие случайную функцию, тем чаще нужно располагать опорные точки при обработке. Ориентировочно можно рекомендовать выбирать элементарный участок Δt так, чтобы на полный период самой высокочастотной гармоники в составе случайной функции приходилось порядка 5—10 опорных точек.

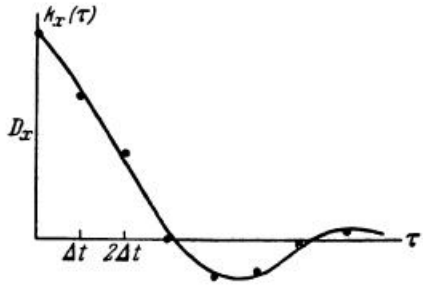


Рис. 17.8.2.

Часто выбор опорных точек вообще не зависит от обрабатывающего, а диктуется темпом работы записывающей аппаратуры. В этом случае следует вести обработку непосредственно полученного из опыта материала, не пытаясь вставить между наблюденными значениями промежуточные, так как это не может повысить точности результата, а излишне осложнит обработку.

Пример. В условиях горизонтального полета самолета произведена запись вертикальной перегрузки, действующей на самолет. Перегрузка регистрировалась на участке времени 200 *сек* с интервалом 2 *сек*. Результаты

приведены в таблице 17.8.1. Считая процесс изменения перегрузки стационарным, определить приближенно математическое ожидание перегрузки m_N, дисперсию D_N и нормированную корреляционную функцию $\rho_N(\tau)$. Аппроксимировать $\rho_N(\tau)$ какой-либо аналитической функцией, найти и построить спектральную плотность случайного процесса.

Таблица 17.8.1

t (сек)	Перегрузка $N(t)$	t (сек)	Перегрузка $N(t)$	t (сек)	Перегрузка $N(t)$	t (сек)	Перегрузка $N(t)$
0	1,0	50	1,0	100	1,2	150	0,8
2	1,3	52	1,1	102	1,4	152	0,6
4	1,1	54	1,5	104	0,8	154	0,9
6	0,7	56	1,0	106	0,9	156	1,2
8	0,7	58	0,8	108	1,0	158	1,3
10	1,1	60	1,1	110	0,8	160	0,9
12	1,3	62	1,1	112	0,8	162	1,3
14	0,8	64	1,2	114	1,4	164	1,5
16	0,8	66	1,0	116	1,6	166	1,2
18	0,4	68	0,8	118	1,7	168	1,4
20	0,3	70	0,8	120	1,3	170	1,4
22	0,3	72	1,2	122	1,6	172	0,8
24	0,6	74	0,7	124	0,8	174	0,8
26	0,3	76	0,7	126	1,2	176	1,3
28	0,5	78	1,1	128	0,6	178	1,0
30	0,5	80	1,5	130	1,0	180	0,7
32	0,7	82	1,0	132	0,6	182	1,1
34	0,8	84	0,6	134	0,8	184	0,9
36	0,6	86	0,9	136	0,7	186	0,9
38	1,0	88	0,8	138	0,9	188	1,1
40	0,5	90	0,8	140	1,3	190	1,2
42	1,0	92	0,9	142	1,5	192	1,3
44	0,9	94	0,9	144	1,1	194	1,3
46	1,4	96	0,6	146	0,7	196	1,6
48	1,4	98	0,4	148	1,0	198	1,5

Решение. По формуле (17.8.5) имеем:

$$m_N = \frac{\sum_{i=1}^{100} N(t_i)}{100} \approx 0{,}98.$$

Центрируем случайную функцию (табл. 17.8.2).

Возводя в квадрат все значения $\mathring{N}(t)$ и деля сумму на $n = 100$, получим приближенно дисперсию случайной функции $N(t)$:

$$D_N = \frac{\sum_{i=1}^{100} [\mathring{N}(t_i)]^2}{100} \approx 0{,}1045$$

Таблица 17.8.2

t (сек)	$\mathring{N}(t)$	t (сек)	$\mathring{N}(t)$	t (сек)	$\mathring{N}(t)$	t (сек)	$\mathring{N}(t)$
0	0,02	50	0,02	100	0,22	150	— 0,18
2	0,32	52	0,12	102	0,42	152	— 0,38
4	0,12	54	0,52	104	— 0,18	154	— 0,08
6	— 0,28	56	0,02	106	— 0,08	156	0,22
8	— 0,28	58	— 0,18	108	0,02	158	0,32
10	0,12	60	0,12	110	— 0,18	160	— 0,08
12	0,32	62	0,12	112	— 0,18	162	0,32
14	— 0,18	64	0,22	114	0,42	164	0,52
16	— 0,18	66	0,02	116	0,62	166	0,22
18	— 0,58	68	— 0,18	118	0,72	168	0,42
20	— 0,68	70	— 0,18	120	0,32	170	0,42
22	— 0,68	72	0,22	122	0,62	172	— 0,18
24	— 0,38	74	— 0,28	124	— 0,18	174	— 0,18
26	— 0,68	76	— 0,28	126	0,22	176	0,32
28	— 0,48	78	0,12	128	— 0,38	178	0,02
30	— 0,48	80	0,52	130	0,02	180	— 0,28
32	— 0,28	82	0,02	132	— 0,38	182	0,12
34	— 0,18	84	— 0,38	134	— 0,18	184	— 0,08
36	— 0,38	86	— 0,08	136	— 0,28	186	— 0,08
38	0,02	88	— 0,18	138	— 0,08	188	0,12
40	— 0,48	90	— 0,18	140	0,32	190	0,22
42	0,02	92	— 0,08	142	0,52	192	0,32
44	— 0,08	94	— 0,08	144	0,12	194	0,32
46	0,42	96	— 0,38	146	— 0,28	196	0,62
48	0,42	98	— 0,58	148	0,02	198	0,52

и среднее квадратическое отклонение:

$$\sigma_N \approx 0{,}323.$$

Перемножая значения $\mathring{N}(t)$, разделенные интервалом $\tau = 2, 4, 6, \ldots$, и деля сумму произведений соответственно на $n-1=99$; $n-2=98$; $n-3=97, \ldots$, получим значения корреляционной функции $k_N(\tau)$. Нормируя корреляционную функцию делением на $D_N = 0{,}1045$, получим таблицу значений функции $\rho_N(\tau)$ (табл. 17.8.3).

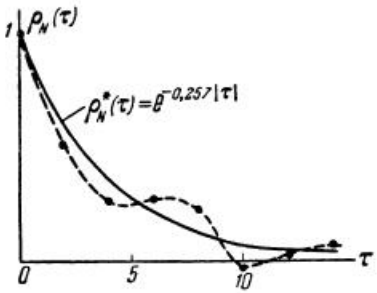


Рис. 17.8.3.

График функции $\rho_N(\tau)$ представлен на рис. 17.8.3 в виде точек, соединенных пунктиром. Не вполне гладкий ход корреляционной функции может быть объяснен недостаточным объемом экспериментальных данных (недостаточной продолжительностью опыта), в связи с чем случайные неровности в ходе функции не успевают сгладиться. Вычисление $\rho_N(\tau)$ продолжено до таких значений τ, при которых фактически корреляционная связь пропадает.

Для того чтобы сгладить явно незакономерные колебания экспериментально найденной функции $\rho_N(\tau)$, заменим ее приближенно функцией вида:

$$\rho_N^*(\tau) = e^{-\alpha|\tau|},$$

где параметр α подберем методом наименьших квадратов (см. n° 14.5).

Таблица 17.8.3

τ	$\rho_N(\tau)$	$\rho_N^*(\tau)=e^{-\alpha\|\tau\|}$
0	1,000	1,000
2	0,505	0,598
4	0,276	0,358
6	0,277	0,214
8	0,231	0,128
10	—0,015	0,077
12	0,014	0,046
14	0,071	0,027

Применяя этот метод, находим $\alpha = 0{,}257$. Вычисляя значения функции $\rho_N^*(\tau)$ при $\tau = 0, 2, 4, \ldots$, построим график сглаживающей кривой. На рис. 17.8.3 он проведен сплошной линией. В последнем столбце таблицы 17.8.3 приведены значения функции $\rho_N^*(\tau)$.

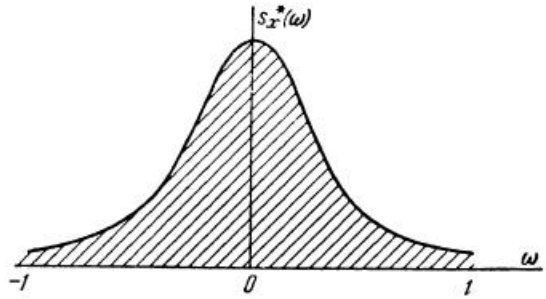


Рис. 17.8.4.

Пользуясь приближенным выражением корреляционной функции (17.8.6), получим (см. n° 17.4, пример 1) нормированную спектральную плотность случайного процесса в виде:

$$S_x^*(\omega) = \frac{\alpha}{\pi(\alpha^2+\omega^2)} = \frac{0{,}257}{\pi(0{,}257^2+\omega^2)}.$$

График нормированной спектральной плотности представлен на рис. 17.8.4.

ГЛАВА 18

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ИНФОРМАЦИИ

18.1. Предмет и задачи теории информации

Теорией информации называется наука, изучающая количественные закономерности, связанные с получением, передачей, обработкой и хранением информации. Возникнув в 40-х годах нашего века из практических задач теории связи, теория информации в настоящее время становится необходимым математическим аппаратом при изучении всевозможных процессов управления.

Черты случайности, присущие процессам передачи информации, заставляют обратиться при изучении этих процессов к вероятностным методам. При этом не удается ограничиться классическими методами теории вероятностей, и возникает необходимость в создании новых вероятностных категорий. Поэтому теория информации представляет собой не просто прикладную науку, в которой применяются вероятностные методы исследования, а должна рассматриваться как раздел теории вероятностей.

Получение, обработка, передача и хранение различного рода информации — непременное условие работы любой управляющей системы. В этом процессе всегда происходит обмен информацией между различными звеньями системы. Простейший случай — передача информации от управляющего устройства к исполнительному органу (передача команд). Более сложный случай — замкнутый контур управления, в котором информация о результатах выполнения команд передается управляющему устройству с помощью так называемой «обратной связи».

Любая информация для того, чтобы быть переданной, должна быть соответственным образом «закодирована», т. е. переведена на язык специальных символов или сигналов. Сигналами, передающими информацию, могут быть электрические импульсы, световые или звуковые колебания, механические перемещения и т. д.

Одной из задач теории информации является отыскание наиболее экономных методов кодирования, позволяющих передать заданную информацию с помощью минимального количества символов. Эта задача решается как при отсутствии, так и при наличии искажений (помех) в канале связи.

Другая типичная задача теории информации ставится следующим образом: имеется источник информации (передатчик), непрерывно вырабатывающий информацию, и канал связи, по которому эта информация передается в другую инстанцию (приемник). Какова должна быть пропускная способность канала связи для того, чтобы канал «справлялся» со своей задачей, т. е. передавал всю поступающую информацию без задержек и искажений?

Ряд задач теории информации относится к определению объема запоминающих устройств, предназначенных для хранения информации, к способам ввода информации в эти запоминающие устройства и вывода ее для непосредственного использования.

Чтобы решать подобные задачи, нужно прежде всего научиться измерять количественно объем передаваемой или хранимой информации, пропускную способность каналов связи и их чувствительность к помехам (искажениям). Основные понятия теории информации, излагаемые в настоящей главе, позволяют дать количественное описание процессов передачи информации и наметить некоторые математические закономерности, относящиеся к этим процессам.

18.2. Энтропия как мера степени неопределенности состояния физической системы

Любое сообщение, с которым мы имеем дело в теории информации, представляет собой совокупность сведений о некоторой физической системе. Например, на вход автоматизированной системы управления производственным цехом может быть передано сообщение о нормальном или повышенном проценте брака, о химическом составе сырья или температуре в печи. На вход системы управления средствами противовоздушной обороны может быть передано сообщение о том, что в воздухе находятся две цели, летящие на определенной высоте, с определенной скоростью. На тот же вход может быть передано сообщение о том, что на определенном аэродроме в данный момент находится такое-то количество истребителей в боевой готовности, или что аэродром выведен из строя огневым воздействием противника, или что первая цель сбита, а вторая продолжает полет с измененным курсом. Любое из этих сообщений описывает состояние какой-то физической системы.

Очевидно, если бы состояние физической системы было известно заранее, не было бы смысла передавать сообщение. Сообщение приобретает смысл только тогда, когда состояние системы заранее неизвестно, случайно.

Поэтому в качестве объекта, о котором передается информация, мы будем рассматривать некоторую физическую систему X, которая случайным образом может оказаться в том или ином состоянии, т. е. систему, которой заведомо присуща какая-то *степень*

неопределенности. Очевидно, сведения, полученные о системе, будут, вообще говоря, тем ценнее и содержательнее, чем больше была неопределенность системы до получения этих сведений («априори»). Возникает естественный вопрос: что значит «бóльшая» или «меньшая» степень неопределенности и чем можно ее измерить?

Чтобы ответить на этот вопрос, сравним между собой две системы, каждой из которых присуща некоторая неопределенность.

В качестве первой системы возьмем монету, которая в результате бросания может оказаться в одном из двух состояний: 1) выпал герб и 2) выпала цифра. В качестве второй — игральную кость, у которой шесть возможных состояний: 1, 2, 3, 4, 5 и 6. Спрашивается, неопределенность какой системы больше? Очевидно, второй, так как у нее больше возможных состояний, в каждом из которых она может оказаться с одинаковой вероятностью.

Может показаться, что степень неопределенности определяется числом возможных состояний системы. Однако в общем случае это не так. Рассмотрим, например, техническое устройство, которое может быть в двух состояниях: 1) исправно и 2) отказало. Предположим, что до получения сведений (априори) вероятность исправной работы устройства 0,99, а вероятность отказа 0,01. Такая система обладает только очень малой степенью неопределенности: почти наверное можно предугадать, что устройство будет работать исправно. При бросании монеты тоже имеется два возможных состояния, но степень неопределенности гораздо больше. Мы видим, что степень неопределенности физической системы определяется не только числом ее возможных состояний, но и вероятностями состояний.

Перейдем к общему случаю. Рассмотрим некоторую систему X, которая может принимать конечное множество состояний: $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$, где

$$p_i = P(X \sim x_i) \tag{18.2.1}$$

— вероятность того, что система X примет состояние x_i (символом $X \sim x_i$ обозначается событие: система находится в состоянии x_i). Очевидно, $\sum_{i=1}^{n} p_i = 1$.

Запишем эти данные в виде таблицы, где в верхней строке перечислены возможные состояния системы, а в нижней — соответствующие вероятности:

x_i	x_1	x_2	$\ldots$	x_n
p_i	p_1	p_2	$\ldots$	p_n

Эта табличка по написанию сходна с рядом распределения прерывной случайной величины X с возможными значениями $x_1, x_2, \ldots, x_n$, имеющими вероятности $p_1, p_2, \ldots, p_n$. И действительно, между физической системой X с конечным множеством состояний и прерывной случайной величиной много общего; для того чтобы свести первую ко второй, достаточно приписать каждому состоянию какое-то числовое значение (скажем, номер состояния). Подчеркнем, что для описания степени неопределенности системы совершенно неважно, какие именно значения $x_1, x_2, \ldots, x_n$ записаны в верхней строке таблицы; важны только количество этих значений и их вероятности.

В качестве меры априорной неопределенности системы (или прерывной случайной величины X) в теории информации применяется специальная характеристика, называемая энтропией. Понятие об энтропии является в теории информации основным.

Энтропией системы называется сумма произведений вероятностей различных состояний системы на логарифмы этих вероятностей, взятая с обратным знаком:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i \text{ [1]}. \qquad (18.2.2)$$

Энтропия $H(X)$, как мы увидим в дальнейшем, обладает рядом свойств, оправдывающих ее выбор в качестве характеристики степени неопределенности. Во-первых, она обращается в нуль, когда одно из состояний системы достоверно, а другие — невозможны. Во-вторых, при заданном числе состояний она обращается в максимум, когда эти состояния равновероятны, а при увеличении числа состояний — увеличивается. Наконец, и это самое главное, она обладает свойством *аддитивности*, т. е. когда несколько независимых систем объединяются в одну, их энтропии складываются.

Логарифм в формуле (18.2.2) может быть взят при любом основании $a > 1$. Перемена основания равносильна простому умножению энтропии на постоянное число, а выбор основания равносилен выбору определенной *единицы измерения* энтропии. Если за основание выбрано число 10, то говорят о «десятичных единицах» энтропии, если 2 — о «двоичных единицах». На практике удобнее всего пользоваться логарифмами при основании 2 и измерять энтропию в двоичных единицах; это хорошо согласуется с применяемой в электронных цифровых вычислительных машинах двоичной системой счисления.

В дальнейшем мы будем везде, если не оговорено противное, под символом log понимать двоичный логарифм.

---

[1]) Знак минус перед суммой поставлен для того, чтобы энтропия была положительной (числа p_i меньше единицы и их логарифмы отрицательны).

В приложении (табл. 6) даны двоичные логарифмы целых чисел от 1 до 100[1]).

Легко убедиться, что при выборе 2 в качестве основания логарифмов за единицу измерения энтропии принимается энтропия простейшей системы X, которая имеет два равновозможных состояния:

x_i	x_1	x_2
p_i	$\frac{1}{2}$	$\frac{1}{2}$

Действительно, по формуле (18.2.2) имеем:

$$H(X) = -\left(\frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}\right) = 1.$$

Определенная таким образом единица энтропии называется «двоичной единицей» и иногда обозначается bit (от английского «binary digit» — двоичный знак). Это энтропия одного разряда двоичного числа, если он с одинаковой вероятностью может быть нулем или единицей.

Измерим в двоичных единицах энтропию системы X, которая имеет n равновероятных состояний:

x_i	x_1	x_2	. . .	x_n
p_i	$\frac{1}{n}$	$\frac{1}{n}$	. . .	$\frac{1}{n}$

Имеем:

$$H(X) = -n\frac{1}{n}\log\frac{1}{n} = -\log 1 + \log n$$

или

$$H(X) = \log n, \tag{18.2.3}$$

т. е. *энтропия системы с равновозможными состояниями равна логарифму числа состояний.*

Например, для системы с восемью состояниями $H(X) = \log 8 = 3$.

Докажем, что в случае, когда состояние системы в точности известно заранее, ее энтропия равна нулю. Действительно, в этом случае все вероятности $p_1, p_2, \ldots, p_n$ в формуле (18.2.2) обращаются

[1]) Логарифмы дробей находятся вычитанием, например:

$$\log 0{,}13 = \log 13 - \log 100.$$

Для нахождения логарифмов чисел с тремя значащими цифрами можно пользоваться линейной интерполяцией.

в нуль, кроме одной — например p_k, которая равна единице. Член $p_k \log p_k$ обращается в нуль, так как $\log 1 = 0$. Остальные члены тоже обращаются в нуль, так как

$$\lim_{p \to 0} p \log p = 0.$$

Докажем, что энтропия системы с конечным множеством состояний достигает максимума, когда все состояния равновероятны. Для этого рассмотрим энтропию системы (18.2.2) как функцию вероятностей $p_1, p_2, \ldots, p_n$ и найдем условный экстремум этой функции при условии:

$$\sum_{i=1}^{n} p_i = 1. \tag{18.2.4}$$

Пользуясь методом неопределенных множителей Лагранжа, будем искать экстремум функции:

$$F = -\sum_{i=1}^{n} p_i \log p_i + \lambda \sum_{i=1}^{n} p_i. \tag{18.2.5}$$

Дифференцируя (18.2.5) по $p_1, \ldots, p_n$ и приравнивая производные нулю, получим систему уравнений:

$$\log p_i + \log e + \lambda = 0 \qquad (i = 1, \ldots, n)$$

или

$$\log p_i = -\lambda - \log e \qquad (i = 1, \ldots, n), \tag{18.2.6}$$

откуда видно, что экстремум (в данном случае максимум) достигается при равных между собой значениях p_i. Из условия (18.2.4) видно, что при этом

$$p_1 = p_2 = \ldots = p_n = \frac{1}{n}, \tag{18.2.7}$$

а максимальная энтропия системы равна:

$$H_{\max}(X) = \log n, \tag{18.2.8}$$

т. е. максимальное значение энтропии системы с конечным числом состояний равно логарифму числа состояний и достигается, когда все состояния равновероятны.

Вычисление энтропии по формуле (18.2.2) можно несколько упростить, если ввести в рассмотрение специальную функцию:

$$\eta(p) = -p \log p, \tag{18.2.9}$$

где логарифм берется по основанию 2.

Формула (18.2.2) принимает вид:

$$H(X) = \sum_{i=1}^{n} \eta(p_i). \tag{18.2.10}$$

Функция $\eta(p)$ затабулирована; в приложении (табл. 7) приведены ее значения для p от 0 до 1 через 0,01.

Пример 1. Определить энтропию физической системы, состоящей из двух самолетов (истребителя и бомбардировщика), участвующих в воздушном бою. В результате боя система может оказаться в одном из четырех возможных состояний:

1) оба самолета не сбиты;
2) истребитель сбит, бомбардировщик не сбит;
3) истребитель не сбит, бомбардировщик сбит;
4) оба самолета сбиты.

Вероятности этих состояний равны соответственно 0,2; 0,3; 0,4 и 0,1.

Решение. Записываем условия в виде таблицы:

x_l	x_1	x_2	x_3	x_4
p_l	0,2	0,3	0,4	0,1

По формуле (18.2.10) имеем:

$$H(X) = \eta(0,2) + \eta(0,3) + \eta(0,4) + \eta(0,1).$$

Пользуясь таблицей 7 приложения, находим

$$H(X) = 0,4644 + 0,5211 + 0,5288 + 0,3322 \approx 1,85 \text{ (дв. ед.)}.$$

Пример 2. Определить энтропию системы, состояние которой описывается прерывной случайной величиной X с рядом распределения

x_l	x_1	x_2	x_3	x_4	x_5
p_l	0,01	0,01	0,01	0,01	0,96

Решение.

$$H(X) = 4\eta(0,01) + \eta(0,96) \approx 0,322 \text{ (дв. ед.)}.$$

Пример 3. Определить максимально возможную энтропию системы, состоящей из трех элементов, каждый из которых может быть в четырех возможных состояниях.

Решение. Общее число возможных состояний системы равно $n = 4 \cdot 4 \cdot 4 = 64$. Максимально возможная энтропия системы равна $\log 64 = 6$ (дв. ед.).

Пример 4. Определить максимально возможную энтропию сообщения, состоящего из пяти букв, причем общее число букв в алфавите равно 32.

Решение. Число возможных состояний системы $n = 32^5$. Максимально возможная энтропия равна $5 \log 32 = 25$ (дв. ед).

Формула (18.2.2) (или равносильная ей (18.2.10)) служит для непосредственного вычисления энтропии. Однако при выполнении преобразований часто более удобной оказывается другая форма записи энтропии, а именно, представление ее в виде математического ожидания:

$$H(X) = M[-\log P(X)], \qquad (18.2.11)$$

где $\log P(X)$ — логарифм вероятности любого (случайного) состояния системы, рассматриваемый как случайная величина.

Когда система X принимает состояния $x_1, \ldots, x_n$, случайная величина $\log P(X)$ принимает значения:

$$\log p_1, \quad \log p_2, \quad \ldots, \quad \log p_n. \tag{18.2.12}$$

Среднее значение (математическое ожидание) случайной величины $-\log P(X)$ и есть, как нетрудно убедиться, энтропия системы X. Для ее получения значения (18.2.12) осредняются с «весами», равными соответствующим вероятностям $p_1, p_2, \ldots, p_n$.

Формулы, подобные (18.2.11), где энтропия представляется в виде математического ожидания, позволяют упрощать преобразования, связанные с энтропией, сводя их к применению известных теорем о математических ожиданиях.

18.3. Энтропия сложной системы. Теорема сложения энтропий

На практике часто приходится определять энтропию для сложной системы, полученной объединением двух или более простых систем.

Под *объединением* двух систем X и Y с возможными состояниями $x_1, \ldots, x_n$; $y_1, \ldots, y_m$ понимается сложная система (X, Y), состояния которой (x_i, y_j) представляют собой все возможные комбинации состояний x_i, y_j систем X и Y.

Очевидно, число возможных состояний системы (X, Y) равно $n \times m$. Обозначим P_{ij} вероятность того, что система (X, Y) будет в состоянии (x_i, y_j):

$$P_{ij} = P\big((X \sim x_i)(Y \sim y_j)\big). \tag{18.3.1}$$

Вероятности P_{ij} удобно расположить в виде таблицы (матрицы)

y_j \ x_i	x_1	x_2	. . .	x_n
y_1	P_{11}	P_{21}	. . .	P_{n1}
y_2	P_{12}	P_{22}	. . .	P_{n2}
. . .	. . .	. . .	. . .	. . .
y_m	P_{1m}	P_{2m}	. . .	P_{nm}

Найдем энтропию сложной системы. По определению она равна сумме произведений вероятностей всех возможных ее состояний на их логарифмы с обратным знаком:

$$H(X, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij} \log P_{ij} \qquad (18.3.2)$$

или, в других обозначениях:

$$H(X, Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} \eta(P_{ij}). \qquad (18.3.2')$$

Энтропию сложной системы, как и энтропию простой, тоже можно записать в форме математического ожидания:

$$H(X, Y) = M[-\log P(X, Y)], \qquad (18.3.3)$$

где $\log P(X, Y)$ — логарифм вероятности состояния системы, рассматриваемый как случайная величина (функция состояния).

Предположим, что системы X и Y независимы, т. е. принимают свои состояния независимо одна от другой, и вычислим в этом предположении энтропию сложной системы. По теореме умножения вероятностей для независимых событий

$$P(X, Y) = P(X) P(Y),$$

откуда

$$\log P(X, Y) = \log P(X) + \log P(Y).$$

Подставляя в (18.3.3), получим

$$H(X, Y) = M[-\log P(X) - \log P(Y)],$$

или

$$H(X, Y) = H(X) + H(Y), \qquad (18.3.4)$$

т. е. *при объединении независимых систем их энтропии складываются*.

Доказанное положение называется теоремой сложения энтропий.

Теорема сложения энтропий может быть легко обобщена на произвольное число независимых систем:

$$H(X_1, X_2, \ldots, X_s) = \sum_{k=1}^{s} H(X_k). \qquad (18.3.5)$$

Если объединяемые системы зависимы, простое сложение энтропий уже неприменимо. В этом случае энтропия сложной системы меньше, чем сумма энтропий ее составных частей. Чтобы найти энтропию системы, составленной из зависимых элементов, нужно ввести новое понятие условной энтропии.

18.4. Условная энтропия. Объединение зависимых систем

Пусть имеются две системы X и Y, в общем случае зависимые. Предположим, что система X приняла состояние x_i. Обозначим $P(y_j|x_i)$ условную вероятность того, что система Y примет состояние y_j при условии, что система X находится в состоянии x_i:

$$P(y_j|x_i) = P(Y \sim y_j | X \sim x_i). \tag{18.4.1}$$

Определим теперь *условную энтропию системы Y при условии, что система X находится в состоянии x_i*. Обозначим ее $H(Y|x_i)$. По общему определению, имеем:

$$H(Y|x_i) = -\sum_{j=1}^{m} P(y_j|x_i) \log P(y_j|x_i) \tag{18.4.2}$$

или

$$H(Y|x_i) = \sum_{j=1}^{m} \eta\big(P(y_j|x_i)\big). \tag{18.4.2'}$$

Формулу (18.4.2) можно также записать в форме математического ожидания:

$$H(Y|x_i) = M_{x_i}[-\log P(Y|x_i)]. \tag{18.4.3}$$

где знаком M_{x_i} обозначено условное математическое ожидание величины, стоящей в скобках, при условии $X \sim x_i$.

Условная энтропия зависит от того, какое состояние x_i приняла система X; для одних состояний она будет больше, для других — меньше. Определим *среднюю*, или *полную*, энтропию системы Y с учетом того, что система может принимать разные состояния. Для этого нужно каждую условную энтропию (18.4.2) умножить на вероятность соответствующего состояния p_i и все такие произведения сложить. Обозначим полную условную энтропию $H(Y|X)$:

$$H(Y|X) = \sum_{i=1}^{n} p_i H(Y|x_i) \tag{18.4.4}$$

или, пользуясь формулой (18.4.2),

$$H(Y|X) = -\sum_{i=1}^{n} p_i \sum_{j=1}^{m} P(y_j|x_i) \log P(y_j|x_i).$$

Внося p_i под знак второй суммы, получим:

$$H(Y|X) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p_i P(y_j|x_i) \log P(y_j|x_i) \tag{18.4.5}$$

или

$$H(Y|X) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \eta\big(P(y_j|x_i)\big). \tag{18.4.5'}$$

Но по теореме умножения вероятностей $p_i P(y_j|x_i) = P_{ij}$,

следовательно,

$$H(Y|X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij}\log P(y_j|x_i). \qquad (18.4.6)$$

Выражению (18.4.6) тоже можно придать форму математического ожидания:

$$H(Y|X) = M[-\log P(Y|X)]. \qquad (18.4.7)$$

Величина $H(Y|X)$ характеризует степень неопределенности системы Y, остающуюся после того, как состояние системы X полностью определилось. Будем называть ее *полной условной энтропией системы Y относительно X.*

Пример 1. Имеются две системы X и Y, объединяемые в одну (X, Y); вероятности состояний системы (X, Y) заданы таблицей

y_j \ x_i	x_1	x_2	x_3	r_j
y_1	0,1	0,2	0	0,3
y_2	0	0,3	0	0,3
y_3	0	0,2	0,2	0,4
p_i	0,1	0,7	0,2	

Определить полные условные энтропии $H(Y|X)$ и $H(X|Y)$.

Решение. Складывая вероятности P_{ij} по столбцам, получим вероятности $p_i = P(X \sim x_i)$:

$$p_1 = 0{,}1; \quad p_2 = 0{,}7; \quad p_3 = 0{,}2.$$

Записываем их в нижней, добавочной строке таблицы. Аналогично, складывая P_{ij} по строкам, найдем:

$$r_1 = 0{,}3; \quad r_2 = 0{,}3; \quad r_3 = 0{,}4 \qquad (r_j = P(Y \sim y_j))$$

и запишем справа дополнительным столбцом. Деля P_{ij} на p_i, получим таблицу условных вероятностей $P(y_j|x_i)$:

y_j \ x_i	x_1	x_2	x_3
y_1	1	$\frac{0{,}2}{0{,}7}$	0
y_2	0	$\frac{0{,}3}{0{,}7}$	0
y_3	0	$\frac{0{,}2}{0{,}7}$	1

По формуле (18.4.5′) находим $H(Y|X)$. Так как условные энтропии при $X \sim x_1$ и $X \sim x_3$ равны нулю, то

$$H(Y|X) = 0{,}7\left[\eta\left(\frac{0{,}2}{0{,}7}\right) + \eta\left(\frac{0{,}3}{0{,}7}\right) + \eta\left(\frac{0{,}2}{0{,}7}\right)\right].$$

Пользуясь таблицей 7 приложения, находим

$$H(Y|X) \approx 1{,}09 \text{ (дв. ед.)}.$$

Аналогично определим $H(X|Y)$. Из формулы (18.4.5′), меняя местами X и Y, получим

$$H(X|Y) = \sum_{j=1}^{m} r_j \eta(P(x_i|y_j)).$$

Составим таблицу условных вероятностей $P(x_i|y_j)$. Деля P_{ij} на r_j, получим:

y_j \ x_i	x_1	x_2	x_3
y_1	$\frac{0{,}1}{0{,}3}$	$\frac{0{,}2}{0{,}3}$	0
y_2	0	1	0
y_3	0	$\frac{0{,}2}{0{,}4}$	$\frac{0{,}2}{0{,}4}$

Отсюда

$$H(X|Y) = 0{,}3\left[\eta\left(\frac{0{,}1}{0{,}3}\right) + \eta\left(\frac{0{,}2}{0{,}3}\right)\right] + 0{,}4\left[\eta\left(\frac{0{,}2}{0{,}4}\right) + \eta\left(\frac{0{,}2}{0{,}4}\right)\right] \approx 0{,}68 \text{ (дв. ед.)}.$$

Пользуясь понятием условной энтропии, можно определить энтропию объединенной системы через энтропию ее составных частей.

Докажем следующую теорему:

Если две системы X и Y объединяются в одну, то энтропия объединенной системы равна энтропии одной из ее составных частей плюс условная энтропия второй части относительно первой:

$$H(X, Y) = H(X) + H(Y|X). \qquad (18.4.8)$$

Для доказательства запишем $H(X, Y)$ в форме математического ожидания (18.3.3):

$$H(X, Y) = M[-\log P(X, Y)].$$

По теореме умножения вероятностей

$$P(X, Y) = P(X)P(Y|X),$$

следовательно,

$$\log P(X, Y) = \log P(X) + \log P(Y|X),$$

откуда

$$H(X, Y) = M[-\log P(X)] + M[-\log P(Y|X)]$$

или, по формулам (18.2.11), (18.3.3)

$$H(X, Y) = H(X) + H(Y|X),$$

что и требовалось доказать.

В частном случае, когда системы X и Y независимы, $H(Y|X) = H(Y)$, и мы получаем уже доказанную в предыдущем п° теорему сложения энтропий:

$$H(X, Y) = H(X) + H(Y).$$

В общем случае

$$H(X, Y) \leqslant H(X) + H(Y). \qquad (18.4.9)$$

Соотношение (18.4.9) следует из того, что полная условная энтропия $H(Y|X)$ не может превосходить безусловной:

$$H(Y|X) \leqslant H(Y). \qquad (18.4.10)$$

Неравенство (18.4.10) будет доказано в п° 18.6. Интуитивно оно представляется довольно очевидным: ясно, что степень неопределенности системы не может увеличиться оттого, что состояние какой-то другой системы стало известным.

Из соотношения (18.4.9) следует, что энтропия сложной системы достигает максимума в крайнем случае, когда ее составные части независимы.

Рассмотрим другой крайний случай, когда состояние одной из систем (например X) полностью определяет собой состояние другой (Y). В этом случае $H(Y|X) = 0$, и формула (18.4.7) дает

$$H(X, Y) = H(X).$$

Если состояние каждой из систем X, Y однозначно определяет состояние другой (или, как говорят, системы X и Y эквивалентны), то

$$H(X, Y) = H(X) = H(Y).$$

Теорему об энтропии сложной системы легко можно распространить на любое число объединяемых систем:

$$H(X_1, X_2, \ldots, X_s) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$
$$\ldots + H(X_s|X_1, X_2, \ldots, X_{s-1}), \qquad (18.4.11)$$

где энтропия каждой последующей системы вычисляется при условии, что состояние всех предыдущих известно.

18.5. Энтропия и информация

В предыдущих п°п° была определена энтропия как мера неопределенности состояния некоторой физической системы. Очевидно, что в результате получения сведений неопределенность системы может быть уменьшена. Чем больше объем полученных сведений, чем они более содержательны, тем больше будет информация о системе, тем менее неопределенным будет ее состояние. Естественно поэтому количество информации измерять уменьшением энтропии той системы, для уточнения состояния которой предназначены сведения.

Рассмотрим некоторую систему X, над которой производится наблюдение, и оценим информацию, получаемую в результате того, что состояние системы X становится полностью известным. До получения сведений (априори) энтропия системы была $H(X)$; после получения сведений состояние системы полностью определилось, т. е. энтропия стала равной нулю. Обозначим I_X информацию, получаемую в результате выяснения состояния системы X. Она равна уменьшению энтропии:

$$I_X = H(X) - 0$$

или

$$I_X = H(X), \tag{18.5.1}$$

т. е. ***количество информации, приобретаемое при полном выяснении состояния некоторой физической системы, равно энтропии этой системы.***

Представим формулу (18.5.1) в виде:

$$I_X = -\sum_{i=1}^{n} p_i \log p_i, \tag{18.5.2}$$

где $p_i = P(X \sim x_i)$.

Формула (18.5.2) означает, что информация I_X есть осредненное по всем состояниям системы значение логарифма вероятности состояния с обратным знаком.

Действительно, для получения I_X каждое значение $\log p_i$ (логарифм вероятности i-го состояния) со знаком минус множится на вероятность этого состояния и все такие произведения складываются. Естественно каждое отдельное слагаемое $-\log p_i$ рассматривать как частную информацию, получаемую от отдельного сообщения, состоящего в том, что система X находится в состоянии x_i. Обозначим эту информацию I_{x_i}:

$$I_{x_i} = -\log p_i. \tag{18.5.3}$$

Тогда информация I_X представится как средняя (или полная) информация, получаемая от всех возможных отдельных сообщений

с учетом их вероятностей. Формула (18.5.2) может быть переписана в форме математического ожидания:

$$I_X = M[-\log P(X)], \tag{18.5.4}$$

где буквой X обозначено любое (случайное) состояние системы X.

Так как все числа p_i не больше единицы, то как частная информация I_{x_i}, так и полная I_X не могут быть отрицательными.

Если все возможные состояния системы априори одинаково вероятны $\left(p_1 = p_2 = \ldots = p_n = \frac{1}{n}\right)$, то, естественно, частная информация I_{x_i} от каждого отдельного сообщения

$$I_{x_i} = -\log p = \log n$$

равна средней (полной) информации

$$I_X = -n\frac{1}{n}\log\frac{1}{n} = \log n.$$

В случае, когда состояния системы обладают различными вероятностями, информации от разных сообщений неодинаковы: наибольшую информацию несут сообщения о тех событиях, которые априори были наименее вероятны. Например, сообщение о том, что 31 декабря в г. Москве выпал снег, несет гораздо меньше информации, чем аналогичное по содержанию сообщение, что 31 июля в г. Москве выпал снег.

Пример 1. На шахматной доске в одной из клеток произвольным образом поставлена фигура. Априори все положения фигуры на доске одинаково вероятны. Определить информацию, получаемую от сообщения, в какой именно клетке находится фигура.

Решение. Энтропия системы X с n равновероятными состояниями равна $\log n$; в данном случае

$$I_X = H(X) = \log 64 = 6 \text{ (дв. ед.)},$$

т. е. сообщение содержит 6 двоичных единиц информации. Так как все состояния системы равновероятны, то ту же информацию несет и любое конкретное сообщение типа: фигура находится в квадрате e2.

Пример 2. В условиях примера 1 определить частную информацию от сообщения, что фигура находится в одной из угловых клеток доски.

Решение. Априорная вероятность состояния, о котором сообщается, равна

$$p = \frac{4}{64} = \frac{1}{16}.$$

Частная информация равна

$$I = -\log\frac{1}{16} = 4 \text{ (дв. ед.)}.$$

Пример 3. Определить частную информацию, содержащуюся в сообщении впервые встреченного лица A: «сегодня мой день рождения».

Решение. Априори все дни в году с одинаковой вероятностью могут быть днями рождения лица A. Вероятность полученного сообщения $p = \frac{1}{365}$. Частная информация от данного сообщения

$$I = -\log \frac{1}{365} \approx 8{,}51 \text{ (дв. ед.)}.$$

Пример 4. В условиях примера 3 определить полную информацию от сообщения, выясняющего, является ли сегодняшний день днем рождения впервые встреченного лица A.

Решение. Система, состояние которой выясняется, имеет два возможных состояния: x_1 — день рождения и x_2 — не день рождения. Вероятности этих состояний $p_1 = \frac{1}{365}$; $p_2 = \frac{364}{365}$.

Полная информация равна:

$$I_X = H(X) = \eta\left(\frac{1}{365}\right) + \eta\left(\frac{364}{365}\right) \approx 0{,}063 \text{ (дв. ед.)}.$$

Пример 5. По цели может быть произведено n независимых выстрелов; вероятность поражения цели при каждом выстреле равна p. После k-го выстрела ($1 \leqslant k < n$) производится разведка, сообщающая, поражена или не поражена цель; если она поражена, стрельба по ней прекращается. Определить k из того условия, чтобы количество информации, доставляемое разведкой, было максимально [1]).

Решение. Рассмотрим физическую систему X_k — цель после k-го выстрела. Возможные состояния системы X_k будут

x_1 — цель поражена;

x_2 — цель не поражена.

Вероятности состояний даны в таблице:

x_i	x_1	x_2
p_i	$1-(1-p)^k$	$(1-p)^k$

Очевидно, информация, доставляемая выяснением состояния системы X_k, будет максимальна, когда оба состояния x_1 и x_2 равновероятны:

$$1-(1-p)^k = (1-p)^k,$$

откуда

$$k = \frac{-1}{\log(1-p)},$$

где log — знак двоичного логарифма.

Например, при $p = 0{,}2$ получаем (округляя до ближайшего целого числа)

$$k = \frac{1}{0{,}3219} \approx 3.$$

[1]) Пример заимствован у И. Я. Динера.

Если информация выражена в двоичных единицах, то ей можно дать довольно наглядное истолкование, а именно: измеряя информацию в двоичных единицах, мы условно характеризуем ее числом ответов «да» или «нет», с помощью которых можно приобрести ту же информацию. Действительно, рассмотрим систему с двумя состояниями:

x_i	x_1	x_2
p_i	p_1	p_2

Чтобы выяснить состояние этой системы, достаточно задать один вопрос, например: находится ли система в состоянии x_1? Ответ «да» или «нет» на этот вопрос доставляет некоторую информацию, которая достигает своего максимального значения 1, когда оба состояния априори равновероятны: $p_1 = p_2 = \frac{1}{2}$. Таким образом, максимальная информация, даваемая ответом «да» или «нет», равна одной двоичной единице.

Если информация от какого-то сообщения равна n двоичным единицам, то она равносильна информации, даваемой n ответами «да» или «нет» на вопросы, поставленные так, что «да» и «нет» одинаково вероятны.

В некоторых простейших случаях для выяснения содержания сообщения действительно удается поставить несколько вопросов так, чтобы ответы «да» и «нет» на эти вопросы были равновероятны. В таких случаях полученная информация фактически измеряется числом таких вопросов.

Если же поставить вопросы точно таким образом не удается, можно утверждать только, что минимальное число вопросов, необходимое для выяснения содержания данного сообщения, *не меньше*, чем информация, заключенная в сообщении. Чтобы число вопросов было минимальным, нужно формулировать их так, чтобы вероятности ответов «да» и «нет» были как можно ближе к $\frac{1}{2}$.

Пример 6. Некто задумал любое целое число X от единицы до восьми:

$$1 \leqslant X \leqslant 8,$$

а нам предлагается угадать его, поставив минимальное число вопросов, на каждый из которых дается ответ «да» или «нет».

Решение. Определяем информацию, заключенную в сообщении, какое число задумано. Априори все значения X от 1 до 8 одинаково вероятны: $p_1 = p_2 = \ldots = p_8 = \frac{1}{8}$, и формула (18.5.2) дает

$$I_X = \log 8 = 3.$$

Минимальное число вопросов, которые нужно поставить для выяснения задуманного числа, не меньше трех.

В данном случае можно, действительно, обойтись тремя вопросами, если сформулировать их так, чтобы вероятности ответов «да» и «нет» были равны.

Пусть, например, задумано число «пять», мы этого не знаем и задаем вопросы:

Вопрос 1. Число X меньше пяти?

Ответ. Нет.

(Вывод: X — одно из чисел 5, 6, 7, 8.)

Вопрос 2. Число X меньше семи?

Ответ. Да.

(Вывод: X — одно из чисел 5, 6.)

Вопрос 3. Число X меньше шести?

Ответ. Да.

(Вывод: число X равно пяти.)

Легко убедиться, что тремя такими (или аналогичными) вопросами можно установить любое задуманное число от 1 до 8 [1]).

Таким образом, мы научились измерять информацию о системе X, содержащуюся как в отдельных сообщениях о ее состоянии, так и в самом факте выяснения состояния. При этом предполагалось, что наблюдение ведется непосредственно за самой системой X. На практике это часто бывает не так: может оказаться, что система X непосредственно недоступна для наблюдения, и выясняется состояние не самой системы X, а некоторой другой системы Y, связанной с нею. Например, вместо непосредственного наблюдения за воздушными целями на посту управления средствами противовоздушной обороны ведется наблюдение за планшетом или экраном отображения воздушной обстановки, на котором цели изображены условными значками. Вместо непосредственного наблюдения за космическим кораблем ведется наблюдение за системой сигналов, передаваемых его аппаратурой. Вместо текста X отправленной телеграммы получатель наблюдает текст Y принятой, который не всегда совпадает с X.

Различия между непосредственно интересующей нас системой X и поддающейся непосредственному наблюдению Y вообще могут быть двух типов:

1) Различия за счет того, что некоторые состояния системы X не находят отражения в системе Y, которая «беднее подробностями», чем система X.

2) Различия за счет ошибок: неточностей измерения параметров системы X и ошибок при передаче сообщений.

Примером различий первого типа могут служить различия, возникающие при округлении численных данных и вообще при грубом описании свойств системы X отображающей ее системой Y. Примерами различий второго типа могут быть искажения сигналов, возникающие за счет помех (шумов) в каналах связи, за счет

[1]) Читателю рекомендуется самостоятельно поставить шесть вопросов, необходимых для выяснения положения фигуры на шахматной доске (см. пример 1), и определить число вопросов, достаточное для выяснения задуманной карты в колоде из 36 карт.

неисправностей передающей аппаратуры, за счет рассеянности людей, участвующих в передаче информации, и т. д.

В случае, когда интересующая нас система X и наблюдаемая Y различны, возникает вопрос: *какое количество информации о системе X дает наблюдение системы Y?*

Естественно определить эту информацию как *уменьшение энтропии системы X* в результате получения сведений о состоянии системы Y:

$$I_{Y\to X}=H(X)-H(X|Y). \tag{18.5.5}$$

Действительно, до получения сведений о системе Y энтропия системы X была $H(X)$; после получения сведений «остаточная» энтропия стала $H(X|Y)$; уничтоженная сведениями энтропия и есть информация $I_{Y\to X}$.

Величину (18.5.5) мы будем называть *полной* (или средней) *информацией о системе X, содержащейся в системе Y.*

Докажем, что

$$I_{Y\to X}=I_{X\to Y},$$

т. е. из двух систем каждая содержит относительно другой одну и ту же полную информацию.

Для доказательства запишем энтропию системы (X, Y) согласно теореме на стр. 479, двумя равносильными формулами:

$$H(X, Y)=H(X)+H(Y|X),$$
$$H(X, Y)=H(Y)+H(X|Y),$$

откуда

$$H(X)+H(Y|X)=H(Y)+H(X|Y).$$
$$H(X)-H(X|Y)=H(Y)-H(Y|X),$$

или

$$I_{Y\to X}=I_{X\to Y}, \tag{18.5.6}$$

что и требовалось доказать.

Введем обозначение:

$$I_{Y\leftrightarrow X}=I_{Y\to X}=I_{X\to Y} \tag{18.5.7}$$

и будем называть информацию $I_{Y\leftrightarrow X}$ *полной взаимной информацией*, содержащейся в системах X и Y.

Посмотрим, во что обращается полная взаимная информация в крайних случаях полной независимости и полной зависимости систем. Если X и Y независимы, то $H(Y|X)=H(Y)$, и

$$I_{Y\leftrightarrow X}=0, \tag{18.5.8}$$

т. е. полная взаимная информация, содержащаяся в независимых системах, равна нулю. Это вполне естественно, так как нельзя получить сведений о системе, наблюдая вместо нее другую, никак с нею не связанную.

Рассмотрим другой крайний случай, когда состояние системы X полностью определяет состояние системы Y и наоборот (системы эквивалентны). Тогда $H(X) = H(Y)$:

$$H(X|Y) = H(Y|X) = 0$$

и

$$I_{Y \leftrightarrow X} = I_X = I_Y = H(X) = H(Y), \tag{18.5.9}$$

т. е. получается случай, уже рассмотренный нами выше (формула (18.5.2)), когда наблюдается непосредственно интересующая нас система X (или, что то же, эквивалентная ей Y).

Рассмотрим случай, когда между системами X и Y имеется жесткая зависимость, но односторонняя: состояние одной из систем полностью определяет состояние другой, но не наоборот. Условимся называть ту систему, состояние которой полностью определяется состоянием другой, «подчиненной системой». По состоянию подчиненной системы вообще нельзя однозначно определить состояние другой. Например, если система X представляет собой полный текст сообщения, составленного из ряда букв, а Y — его сокращенный текст, в котором для сокращения пропущены все гласные буквы, то, читая в сообщении Y слово «стл», нельзя в точности быть уверенным, означает оно «стол», «стул», «стал» или «устал».

Очевидно, энтропия подчиненной системы меньше, чем энтропия той системы, которой она подчинена.

Определим полную взаимную информацию, содержащуюся в системах, из которых одна является подчиненной.

Пусть из двух систем X и Y подчиненной является X. Тогда $H(X|Y) = 0$, и

$$I_{Y \leftrightarrow X} = H(X), \tag{18.5.10}$$

т. е. *полная взаимная информация, содержащаяся в системах, из которых одна является подчиненной, равна энтропии подчиненной системы.*

Выведем выражение для информации $I_{Y \leftrightarrow X}$ не через условную энтропию, а непосредственно через энтропию объединенной системы $H(X, Y)$ и энтропии ее составных частей $H(X)$ и $H(Y)$.

Пользуясь теоремой об энтропии объединенной системы (стр. 479), получим:

$$H(X|Y) = H(X, Y) - H(Y). \tag{18.5.11}$$

Подставляя это выражение в формулу (18.5.5), получим:

$$I_{Y \leftrightarrow X} = H(X) + H(Y) - H(X, Y), \tag{18.5.12}$$

т. е. *полная взаимная информация, содержащаяся в двух системах, равна сумме энтропий составляющих систем минус энтропия объединенной системы.*

На основе полученных зависимостей легко вывести общее выражение для полной взаимной информации в виде математического

ожидания. Подставляя в (18.5.12) выражения для энтропий:

$$H(X) = M[-\log P(X)], \quad H(Y) = M[-\log P(Y)],$$
$$H(X, Y) = M[-\log P(X, Y)],$$

получим

$$I_{Y \leftrightarrow X} = M[-\log P(X) - \log P(Y) + \log P(X, Y)]$$

или

$$I_{Y \leftrightarrow X} = M\left[\log \frac{P(X, Y)}{P(X) P(Y)}\right]. \qquad (18.5.13)$$

Для непосредственного вычисления полной взаимной информации формулу (18.5.13) удобно записать в виде

$$I_{Y \leftrightarrow X} = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} \log \frac{P_{ij}}{p_i r_j}, \qquad (18.5.14)$$

где

$$P_{ij} = P\big((X \sim x_i)(Y \sim y_j)\big),$$
$$p_i = P(X \sim x_i); \quad r_j = P(Y \sim y_j).$$

Пример 1. Найти полную взаимную информацию, содержащуюся в системах X и Y в условиях примера 1 п° 18.4.

Решение. Из примера 1 п° 18.4 с помощью таблицы 7 приложения получим:

$$H(X, Y) = 2{,}25; \quad H(X) = 1{,}16; \quad H(Y) = 1{,}57;$$
$$I_{Y \leftrightarrow X} = H(X) + H(Y) - H(X, Y) = 0{,}48 \text{ (дв. ед.)}.$$

Пример 2. Физическая система X может находиться в одном из четырех состояний x_1, x_2, x_3, x_4; соответствующие вероятности даны в таблице

x_i	x_1	x_2	x_3	x_4
p_i	0,1	0,2	0,4	0,3

При наблюдении за системой X состояния x_1 и x_2 неразличимы; состояния x_3 и x_4 также неразличимы. Сообщение о системе X указывает, находится ли она в одном из состояний x_1, x_2 или же в одном из состояний x_3, x_4. Получено сообщение, указывающее, в каком из состояний: x_1, x_2 или x_3, x_4 — находится система X. Определить информацию, заключенную в этом сообщении.

Решение. В данном примере мы наблюдаем не саму систему X, а подчиненную ей систему Y, которая принимает состояние y_1, когда система X оказывается в одном из состояний x_1, x_2, и состояние y_2, когда X оказывается в одном из состояний x_3, x_4. Имеем:

$$r_1 = P(Y \sim y_1) = 0{,}1 + 0{,}2 = 0{,}3;$$
$$r_2 = P(Y \sim y_2) = 0{,}3 + 0{,}4 = 0{,}7.$$

Находим взаимную информацию, т. е. энтропию подчиненной системы:

$$I_{Y \leftrightarrow X} = -r_1 \log r_1 - r_2 \log r_2 = \eta(0{,}3) + \eta(0{,}7) \approx 0{,}88 \text{ (дв. ед.)}.$$

18.6. Частная информация о системе, содержащаяся в сообщении о событии. Частная информация о событии, содержащаяся в сообщении о другом событии

В предыдущем п° мы рассмотрели полную (или среднюю) информацию о системе X, содержащуюся в сообщении о том, в каком состоянии находится система Y. В ряде случаев представляет интерес оценить *частную информацию* о системе X, содержащуюся в отдельном сообщении, указывающем, что система Y находится в конкретном состоянии y_j. Обозначим эту частную информацию $I_{y_j \to X}$. Заметим, что полная (или, иначе, средняя) информация $I_{Y \to X}$ должна представлять собой математическое ожидание частной информации для всех возможных состояний, о которых может быть передано сообщение:

$$I_{Y \to X} = \sum_{j=1}^{m} r_j I_{y_j \to X}. \tag{18.6.1}$$

Придадим формуле (18.5.14), по которой вычисляется $I_{Y \to X}$ (она же $I_{Y \leftrightarrow X}$), такой вид, как у формулы (18.6.1):

$$I_{Y \to X} = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} \log \frac{P_{ij}}{p_i r_j} = \sum_{i=1}^{n} \sum_{j=1}^{m} r_j P(x_i | y_j) \log \frac{r_j P(x_i | y_j)}{p_i r_j} =$$

$$= \sum_{j=1}^{m} r_j \sum_{i=1}^{n} P(x_i | y_j) \log \frac{P(x_i | y_j)}{p_i}, \tag{18.6.2}$$

откуда, сравнивая с формулой (18.6.1), получим выражение частной информации:

$$I_{y_j \to X} = \sum_{i=1}^{n} P(x_i | y_j) \log \frac{P(x_i | y_j)}{p_i}. \tag{18.6.3}$$

Выражение (18.6.3) и примем за определение частной информации. Проанализируем структуру этого выражения. Оно представляет собой не что иное, как осредненное по всем состояниям x_i значение величины

$$\log \frac{P(x_i | y_j)}{p_i}. \tag{18.6.4}$$

Осреднение происходит с учетом различных вероятностей значений $x_1, x_2, \ldots, x_n$. Так как система Y уже приняла состояние y_j, то при осреднении значения (18.6.4) множатся не на вероятности p_i состояний x_i, а на у с л о в н ы е вероятности $P(x_i | y_j)$.

Таким образом, выражение для частной информации можно записать в виде условного математического ожидания:

$$I_{y_j \to X} = M_{y_j} \left[\log \frac{P(X | y_j)}{P(X)} \right]. \tag{18.6.5}$$

Докажем, что частная информация $I_{y_j \to X}$, так же как и полная, не может быть отрицательна. Действительно, обозначим:

$$\frac{P(x_i \mid y_j)}{p_i} = q_{ij} \tag{18.6.6}$$

и рассмотрим выражение

$$\log \frac{P(x_i \mid y_j)}{p_i} = \log q_{ij}.$$

Легко убедиться (см. рис. 18.6.1), что при любом $x > 0$

$$\ln x \leqslant x - 1. \tag{18.6.7}$$

Полагая в (18.6.7) $x = \frac{1}{q_{ij}}$, получим:

$$-\ln q_{ij} \leqslant \frac{1}{q_{ij}} - 1; \quad \ln q_{ij} \geqslant 1 - \frac{1}{q_{ij}},$$

откуда

$$\log q_{ij} = \frac{\ln q_{ij}}{\ln 2} \geqslant \frac{1}{\ln 2}\left(1 - \frac{1}{q_{ij}}\right). \tag{18.6.8}$$

На основании (18.6.3) и (18.6.6) имеем:

$$I_{y_j \to X} = \sum_{i=1}^{n} P(x_i \mid y_j) \log q_{ij} \geqslant$$

$$\geqslant \frac{1}{\ln 2} \sum_{i=1}^{n} P(x_i \mid y_j)\left(1 - \frac{1}{q_{ij}}\right) = \frac{1}{\ln 2} \sum_{i=1}^{n} P(x_i \mid y_j)\left[1 - \frac{p_i}{P(x_i \mid y_j)}\right] =$$

$$= \frac{1}{\ln 2}\left\{\sum_{i=1}^{n} P(x_i \mid y_j) - \sum_{i=1}^{n} p_i\right\}.$$

Но

$$\sum_{i=1}^{n} P(x_i \mid y_j) = \sum_{i=1}^{n} p_i = 1,$$

и выражение в фигурной скобке равно нулю; следовательно $I_{y_j \to X} \geqslant 0$.

Таким образом, мы доказали, что *частная информация о системе X, заключенная в сообщении о любом состоянии y_j системы Y, не может быть отрицательной*. Отсюда следует, что неотрицательна и полная взаимная информация $I_{Y \leftrightarrow X}$ (как математическое ожидание неотрицательной случайной величины):

$$I_{Y \leftrightarrow X} \geqslant 0. \tag{18.6.9}$$

Из формулы (18.5.5) для информации: $I_{Y \leftrightarrow X} = H(X) - H(X \mid Y)$ следует, что

$$H(X) - H(X \mid Y) \geqslant 0 \tag{18.6.10}$$

или

$$H(X \mid Y) \leqslant H(X),$$

т. е. *полная условная энтропия системы не превосходит ее безусловной энтропии*[1]).

Таким образом, доказано положение, принятое нами на веру в п° 18.3.

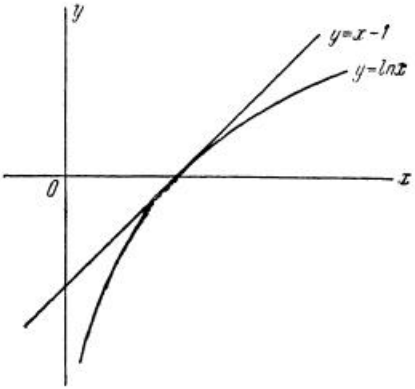


Рис. 18.6.1.

Для непосредственного вычисления частной информации формулу (18.6.3) удобно несколько преобразовать, введя в нее вместо условных вероятностей $P(x_i|y_j)$ безусловные. Действительно,

$$P(x_i|y_j) = \frac{P_{ij}}{r_j},$$

и формула (18.6.3) принимает вид

$$I_{y_j \to X} = \sum_{i=1}^{n} \frac{P_{ij}}{r_j} \log \frac{P_{ij}}{p_i r_j}. \qquad (18.6.11)$$

Пример 1. Система (X, Y) характеризуется таблицей вероятностей P_{ij}:

y_j \ x_i	x_1	x_2	r_j
y_1	0,1	0,2	0,3
y_2	0,3	0,4	0,7
p_i	0,4	0,6	

Найти частную информацию о системе X, заключенную в сообщении $Y \sim y_1$.

[1]) Заметим, что это справедливо только в отношении *полной* условной энтропии; что касается частной условной энтропии $H(X|y_j)$, то она для отдельных y_j может быть как больше, так и меньше $H(X)$.

Решение. По формуле (18.6.11) имеем:

$$I_{y_1 \to X} = \frac{0,1}{0,3} \log \frac{0,1}{0,4 \cdot 0,3} + \frac{0,2}{0,3} \log \frac{0,2}{0,6 \cdot 0,3}.$$

По таблице 6 приложения находим

$$\log \frac{0,1}{0,4 \cdot 0,3} = \log 10 - \log 12 \approx -0,263,$$

$$\log \frac{0,2}{0,6 \cdot 0,3} = \log 20 - \log 18 \approx 0,152,$$

$$I_{y_1 \to X} \approx -0,333 \cdot 0,263 + 0,667 \cdot 0,152 \approx 0,013 \text{ (дв. ед.)}.$$

Мы определили частную информацию о системе X, содержащуюся в конкретном событии $Y \sim y_j$, т. е. в сообщении «система Y находится в состоянии y_j». Возникает естественный вопрос: а нельзя ли пойти еще дальше и определить частную информацию о событии $X \sim x_i$, содержащуюся в событии $Y \sim y_j$? Оказывается, это можно сделать, только получаемая таким образом информация «от события к событию» будет обладать несколько неожиданными свойствами: она может быть как положительной, так и отрицательной.

Исходя из структуры формулы (18.6.3), естественно определить информацию «от события к событию» следующим образом:

$$I_{y_j \to x_i} = \log \frac{P(x_i \mid y_j)}{p_i}, \tag{18.6.12}$$

т. е. *частная информация о событии, получаемая в результате сообщения о другом событии, равна логарифму отношения вероятности первого события после сообщения к его же вероятности до сообщения (априори).*

Из формулы (18.6.12) видно, что если вероятность события $X \sim x_i$ в результате сообщения $Y \sim y_j$ увеличивается, т. е.

$$P(x_i \mid y_j) > p_i,$$

то информация $I_{y_j \to x_i}$ положительна; в противном случае она отрицательна. В частности, когда появление события $Y \sim y_j$ полностью исключает возможность появления события $X \sim x_i$ (т. е. когда эти события несовместны), то

$$I_{y_j \to x_i} = -\infty.$$

Информацию $I_{y_j \to x_i}$ можно записать в виде:

$$I_{y_j \to x_i} = \log \frac{P(x_i \mid y_j)}{p_i} = \log \frac{P_{ij}}{p_i r_j}, \tag{18.6.13}$$

из чего следует, что она симметрична относительно x_i и y_j, и

$$I_{y_j \to x_i} = I_{x_i \to y_j} = I_{y_j \leftrightarrow x_i}. \tag{18.6.14}$$

Таким образом, нами введены три вида информации:

1) полная информация о системе X, содержащаяся в системе Y.

$$I_{Y \to X} = I_{Y \leftrightarrow X};$$

2) частная информация о системе X, содержащаяся в событии (сообщении) $Y \sim y_j$:

$$I_{y_j \to X};$$

3) частная информация о событии $X \sim x_l$, содержащаяся в событии (сообщении) $Y \sim y_j$:

$$I_{y_j \to x_l} = I_{y_j \leftrightarrow x_l}.$$

Первые два типа информации неотрицательны; последняя может быть как положительной, так и отрицательной.

Пример 2. В урне 3 белых и 4 черных шара. Из урны вынуто 4 шара, три из них оказались черными, а один — белым. Определить информацию заключенную в наблюденном событии B по отношению к событию A — следующий вынутый из урны шар будет черным.

Решение.

$$I_{B \to A} = \log \frac{P(A \mid B)}{P(A)} = \log \frac{1/3}{4/7} \approx -0{,}779 \text{ (дв. ед.)}.$$

18.7. Энтропия и информация для систем с непрерывным множеством состояний

До сих пор мы рассматривали физические системы, различные состояния которых $x_1, x_2, \ldots, x_n$ можно было все перечислить; вероятности этих состояний были какие-то отличные от нуля величины $p_1, p_2, \ldots, p_n$. Такие системы аналогичны прерывным (дискретным) случайным величинам, принимающим значения $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$. На практике часто встречаются физические системы другого типа, аналогичные непрерывным случайным величинам. Состояния таких систем нельзя перенумеровать: они непрерывно переходят одно в другое, причем каждое отдельное состояние имеет вероятность, равную нулю, а распределение вероятностей характеризуется некоторой плотностью. Такие системы, по аналогии с непрерывными случайными величинами, мы будем называть «непрерывными», в отличие от ранее рассмотренных, которые мы будем называть «дискретными». Наиболее простой пример непрерывной системы — это система, состояние которой описывается одной непрерывной случайной величиной X с плотностью распределения $f(x)$. В более сложных случаях состояние системы описывается несколькими случайными величинами $X_1, X_2, \ldots, X_s$ с плотностью распределения $f(x_1, x_2, \ldots, x_s)$. Тогда ее можно рассматривать как объединение $(X_1, X_2, \ldots, X_s)$ простых систем $X_1, X_2, \ldots, X_s$.

Рассмотрим простую систему X, определяемую одной непрерывной случайной величиной X с плотностью распределения $f(x)$

(рис. 18.7.1). Попытаемся распространить на эту систему введенное в п° 18.1 понятие энтропии.

Прежде всего отметим, что понятие «непрерывной системы», как и понятие «непрерывной случайной величины», является некоторой идеализацией. Например, когда мы считаем величину X — рост наугад взятого человека — непрерывной случайной величиной, мы отвлекаемся от того, что фактически никто не измеряет рост точнее, чем до 1 *см*, и что различить между собой два значения роста, разнящиеся, скажем, на 1 *мм*, практически невозможно. Тем не менее данную случайную величину естественно описывать как непрерывную, хотя можно было бы описать ее и как дискретную, считая совпадающими те значения роста, которые различаются менее чем на 1 *см*.

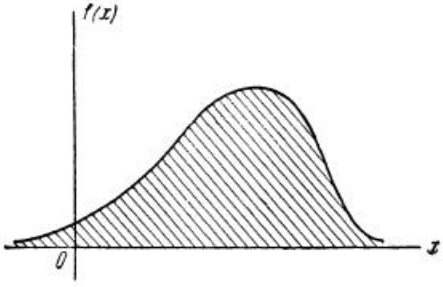


Рис. 18.7.1.

Точно таким образом, установив предел точности измерений, т. е. некоторый отрезок Δx, в пределах которого состояния системы X практически неразличимы, можно приближенно свести непрерывную систему X к дискретной. Это равносильно замене плавной кривой $f(x)$ ступенчатой, типа гистограммы (рис. 18.7.2); при этом каждый участок (разряд) длины Δx заменяется одной точкой-представителем. Площади прямоугольников изображают вероятности попадания в соответствующие разряды: $f(x_i)\Delta x$. Если условиться считать неразличимыми состояния системы, относящиеся к одному разряду, и объединить их все в одно состояние, то можно приближенно определить энтропию системы X, рассматриваемой с точностью до Δx:

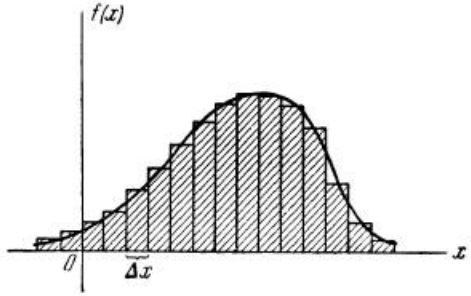


Рис. 18.7.2.

$$H_{\Delta x}(X) = -\sum_i f(x_i)\Delta x \log\{f(x_i)\Delta x\} =$$
$$= -\sum_i f(x_i)\Delta x\{\log f(x_i) + \log \Delta x\} =$$
$$= -\sum_i \{f(x_i)\log f(x_i)\}\Delta x - \log \Delta x \sum_i f(x_i)\Delta x. \quad (18.7.1)$$

При достаточно малом Δx

$$\sum_i \{f(x_i)\log f(x_i)\}\,\Delta x \approx \int_{-\infty}^{\infty} f(x)\log f(x)\,dx,$$

$$\sum_i f(x_i)\,\Delta x \approx \int_{-\infty}^{\infty} f(x)\,dx = 1,$$

и формула (18.7.1) принимает вид:

$$H_{\Delta x}(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)\,dx - \log \Delta x. \qquad (18.7.2)$$

Заметим, что в выражении (18.7.2) первый член получился совсем не зависящим от Δx — степени точности определения состояний системы. Зависит от Δx только второй член ($-\log \Delta x$), который стремится к бесконечности при $\Delta x \to 0$. Это и естественно, так как чем точнее мы хотим задать состояние системы X, тем бóльшую степень неопределенности мы должны устранить, и при неограниченном уменьшении Δx эта неопределенность растет тоже неограниченно.

Итак, задаваясь произвольно малым «участком нечувствительности» Δx наших измерительных приборов, с помощью которых определяется состояние физической системы X, можно найти энтропию $H_{\Delta x}(X)$ по формуле (18.7.2), в которой второй член неограниченно растет с уменьшением Δx. Сама энтропия $H_{\Delta x}(X)$ отличается от этого неограниченно растущего члена на независимую от Δx величину

$$H^*(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)\,dx. \qquad (18.7.3)$$

Эту величину можно назвать «приведенной энтропией» непрерывной системы X. Энтропия $H_{\Delta x}(X)$ выражается через приведенную энтропию $H^*(X)$ формулой

$$H_{\Delta x}(X) = H^*(X) - \log \Delta x. \qquad (18.7.4)$$

Соотношение (18.7.4) можно истолковать следующим образом: от точности измерения Δx зависит только начало отсчета, при котором вычисляется энтропия.

В дальнейшем для упрощения записи мы будем опускать индекс Δx в обозначении энтропии и писать просто $H(X)$; наличие Δx в правой части всегда укажет, о какой точности идет речь.

Формуле (18.7.2) для энтропии можно придать более компактный вид, если, как мы это делали для прерывных величин, записать ее в виде математического ожидания функции. Прежде всего перепишем

(18.7.2) в виде

$$H(X) = -\int_{-\infty}^{\infty} f(x) \log \{f(x)\,\Delta x\}\, dx. \qquad (18.7.5)$$

Это есть не что иное, как математическое ожидание функции $-\log\{f(X)\,\Delta x\}$ от случайной величины X с плотностью $f(x)$:

$$H(X) = M[-\log\{f(X)\,\Delta x\}]. \qquad (18.7.6)$$

Аналогичную форму можно придать величине $H^*(X)$:

$$H^*(X) = M[-\log f(X)]. \qquad (18.7.7)$$

Перейдем к определению условной энтропии. Пусть имеются две непрерывные системы: X и Y. В общем случае эти системы зависимы. Обозначим $f(x, y)$ плотность распределения для состояний объединенной системы (X, Y); $f_1(x)$ — плотность распределения системы X; $f_2(y)$ — плотность распределения системы Y; $f(y|x)$, $f(x|y)$ — условные плотности распределения.

Прежде всего определим частную условную энтропию $H(Y|x)$, т. е. энтропию системы Y при условии, что система X приняла определенное состояние x. Формула для нее будет аналогична (18.4.2), только вместо условных вероятностей $P(y_j|x_i)$ будут стоять условные законы распределения $f(y|x)$ и появится слагаемое $\log \Delta y$:

$$H(Y|x) = -\int_{-\infty}^{\infty} f(y|x) \log f(y|x)\, dy - \log \Delta y. \qquad (18.7.8)$$

Перейдем теперь к полной (средней) условной энтропии $H(Y|X)$; для этого нужно осреднить частную условную энтропию $H(Y|x)$ по всем состояниям x с учетом их вероятностей, характеризуемых плотностью $f_1(x)$:

$$H(Y|X) = -\int_{-\infty}^{\infty}\int f_1(x) f(y|x) \log f(y|x)\, dx\, dy - \log \Delta y \qquad (18.7.9)$$

или, учитывая, что

$$f(x, y) = f_1(x) f(y|x),$$

$$H(Y|X) = -\int_{-\infty}^{\infty}\int f(x, y) \log f(y|x)\, dx\, dy - \log \Delta y. \qquad (18.7.10)$$

Иначе эта формула может быть записана в виде

$$H(Y|X) = M[-\log f(Y|X)] - \log \Delta y \qquad (18.7.11)$$

или

$$H(Y|X) = M[-\log\{f(Y|X)\,\Delta y\}]. \qquad (18.7.12)$$

Определив таким образом условную энтропию, покажем, как она применяется при определении энтропии объединенной системы.

Найдем сначала энтропию объединенной системы непосредственно. Если «участками нечувствительности» для систем X и Y будут Δx и Δy, то для объединенной системы (X, Y) роль их будет играть элементарный прямоугольник $\Delta x\,\Delta y$. Энтропия системы (X, Y) будет

$$H(X, Y) = M[-\log\{f(X, Y)\,\Delta x\,\Delta y\}]. \qquad (18.7.13)$$

Так как

$$f(x, y) = f_1(x)\,f(y \mid x),$$

то и

$$f(X, Y) = f_1(X)\,f(Y \mid X). \qquad (18.7.14)$$

Подставим (18.7.14) в (18.7.13):

$$H(X, Y) = M[-\log f_1(X) - \log f(Y \mid X) - \log \Delta x - \log \Delta y] =$$
$$= M\left[-\log\left(f_1(X)\,\Delta x\right)\right] + M\left[-\log\left(f(Y \mid X)\,\Delta y\right)\right],$$

или, по формулам (18.7.6) и (18.7.12),

$$H(X, Y) = H(X) + H(Y \mid X), \qquad (18.7.15)$$

т. е. теорема об энтропии сложной системы остается в силе и для непрерывных систем.

Если X и Y независимы, то энтропия объединенной системы равна сумме энтропий составных частей:

$$H(X, Y) = H(X) + H(Y). \qquad (18.7.16)$$

Пример 1. Найти энтропию непрерывной системы X, все состояния которой на каком-то участке (α, β) одинаково вероятны:

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{при } \alpha < x < \beta, \\ 0 & \text{при } x < \alpha \text{ или } x > \beta. \end{cases}$$

Решение.

$$H^*(X) = -\int_\alpha^\beta \frac{1}{\beta - \alpha} \log \frac{1}{\beta - \alpha}\,dx = \log(\beta - \alpha);$$

$$H(X) = \log(\beta - \alpha) - \log \Delta x$$

или

$$H(X) = \log \frac{\beta - \alpha}{\Delta x}. \qquad (18.7.17)$$

Пример 2. Найти энтропию системы X, состояния которой распределены по нормальному закону:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{x^2}{2\sigma^2}}.$$

Решение.

$$H^*(X) = M[-\log f(X)] = M\left[-\log\left\{\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{X^2}{2\sigma^2}}\right\}\right] =$$

$$= M\left[\log(\sqrt{2\pi}\,\sigma) + \frac{X^2}{2\sigma^2}\log e\right] = \log(\sqrt{2\pi}\,\sigma) + \frac{M[X^2]}{2\sigma^2}\log e.$$

Но

$$M[X^2] = D[X] = \sigma^2,$$

$$H^*(X) = \log(\sqrt{2\pi}\,\sigma) + \frac{1}{2}\log e = \log(\sqrt{2\pi e}\,\sigma)$$

и

$$H(X) = \log(\sqrt{2\pi e}\,\sigma) - \log\Delta x = \log\left[\frac{\sqrt{2\pi e}\,\sigma}{\Delta x}\right]. \qquad (18.7.18)$$

Пример 3. Состояние самолета характеризуется тремя случайными величинами: высотой H, модулем скорости V и углом θ, определяющим направление полета. Высота самолета распределена с равномерной плотностью на участке (h_1, h_2); скорость V — по нормальному закону с м.о. v_0 и с.к.о. σ_v; угол θ — с равномерной плотностью на участке $(0, \pi)$. Величины H, V, θ независимы. Найти энтропию объединенной системы.

Решение.

Из примера 1 (формула (18.7.17)) имеем

$$H(H) = \log\frac{h_2 - h_1}{\Delta h},$$

где Δh — «участок нечувствительности» при определении высоты.

Так как энтропия случайной величины не зависит от ее математического ожидания, то для определения энтропии величины V воспользуемся формулой (18.7.18):

$$H(V) = \log\left[\frac{\sqrt{2\pi e}\,\sigma_v}{\Delta v}\right].$$

Энтропия величины θ:

$$H(\theta) = \log(\pi - 0) - \log\Delta\theta = \log\frac{\pi}{\Delta\theta}.$$

Окончательно имеем:

$$H(H, V, \theta) = \log\frac{h_2 - h_1}{\Delta h} + \log\frac{\sqrt{2\pi e}\,\sigma_v}{\Delta v} + \log\frac{\pi}{\Delta\theta}$$

или

$$H(H, V, \theta) = \log\left\{\frac{h_2 - h_1}{\Delta h}\cdot\frac{\sqrt{2\pi e}\,\sigma_v}{\Delta v}\cdot\frac{\pi}{\Delta\theta}\right\}. \qquad (18.7.19)$$

Заметим, что каждый из сомножителей под знаком фигурной скобки имеет один и тот же смысл: он показывает, сколько «участков нечувствительности» укладывается в некотором характерном для данной случайной величины отрезке. В случае распределения с равномерной плотностью этот участок представляет собой просто участок возможных значений случайной величины; в случае нормального распределения этот участок равен $\sqrt{2\pi e}\,\sigma$, где σ — среднее квадратическое отклонение.

Таким образом, мы распространили понятие энтропии на случай непрерывных систем. Аналогично может быть распространено и понятие информации. При этом неопределенность, связанная с наличием в выражении энтропии неограниченно возрастающего слагаемого, отпадает: при вычислении информации, как разности двух энтропий, эти члены взаимно уничтожаются. Поэтому все виды информации, связанные с непрерывными величинами, оказываются не зависящими от «участка нечувствительности» Δx.

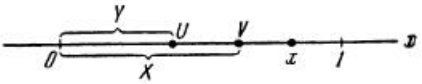


Рис. 18.7.3.

Выражение для полной взаимной информации, содержащейся в двух непрерывных системах X и Y, будет аналогично выражению (18.5.4), но с заменой вероятностей законами распределения, а сумм — интегралами:

$$I_{Y \leftrightarrow X} = \int_{-\infty}^{\infty} \int f(x, y) \log \frac{f(x, y)}{f_1(x) f_2(y)} dx\, dy \qquad (18.7.20)$$

или, применяя знак математического ожидания,

$$I_{Y \leftrightarrow X} = M\left[\log \frac{f(X, Y)}{f_1(X) f_2(Y)}\right]. \qquad (18.7.21)$$

Полная взаимная информация $I_{Y \leftrightarrow X}$, как и в случае дискретных систем, есть неотрицательная величина, обращающаяся в нуль только тогда, когда системы X и Y независимы.

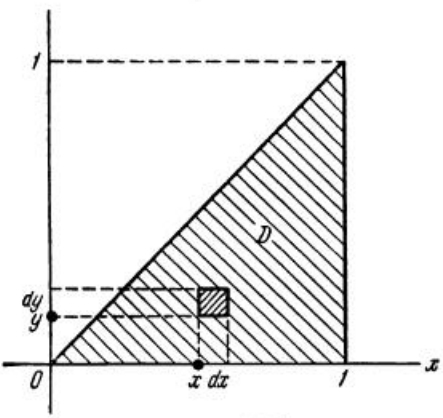


Рис. 18.7.4.

Пример 4. На отрезке (0, 1) выбираются случайным образом, независимо друг от друга, две точки U и V; каждая из них распределена на этом отрезке с равномерной плотностью. В результате опыта одна из точек легла правее, другая — левее. Сколько информации о положении правой точки дает значение положения левой?

Решение. Рассмотрим две случайные точки U и V на оси абсцисс Ox (рис. 18.7.3). Обозначим Y абсциссу той из них, которая оказалась слева, а X — абсциссу той, которая оказалась справа (на рис. 18.7.3 слева оказалась точка U, но могло быть и наоборот). Величины X и Y определяются через U и V следующим образом:

$$Y = \min\{U, V\}; \qquad X = \max\{U, V\}.$$

Найдем закон распределения системы (X, Y). Так как $Y < X$, то он будет существовать только в области D, заштрихованной на рис. 18.7.4. Обозначим

$f(x, y)$ плотность распределения системы (X, Y) и найдем элемент вероятности $f(x, y)\,dx\,dy$, т. е. вероятность того, что случайная точка (X, Y) попадет в элементарный прямоугольник $(x, x+dx; y, y+dy)$. Это событие может произойти двумя способами: либо слева окажется точка U, а справа V, либо наоборот. Следовательно,

$$f(x, y)\,dx\,dy = \varphi(x, y)\,dx\,dy + \varphi(y, x)\,dx\,dy,$$

где $\varphi(u, v)$ обозначена плотность распределения системы величин (U, V).

В данном случае

$$\varphi(u, v) = 1 \qquad \begin{pmatrix} 0 < u < 1 \\ 0 < v < 1 \end{pmatrix},$$

следовательно,

$$\varphi(x, y) = \varphi(y, x) = 1;$$

$$f(x, y)\,dx\,dy = 2\,dx\,dy$$

и

$$f(x, y) = \begin{cases} 2 & \text{при} \quad (x, y) \subset D, \\ 0 & \text{при} \quad (x, y) \not\subset D. \end{cases}$$

Найдем теперь законы распределения отдельных величин, входящих в систему:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\,dy = \int_0^x 2dy = 2x \quad \text{при} \quad 0 < x < 1;$$

аналогично

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\,dx = \int_y^1 2\,dx = 2(1-y) \quad \text{при} \quad 0 < y < 1.$$

Графики плотностей $f_1(x)$ и $f_2(y)$ изображены на рис. 18.7.5.

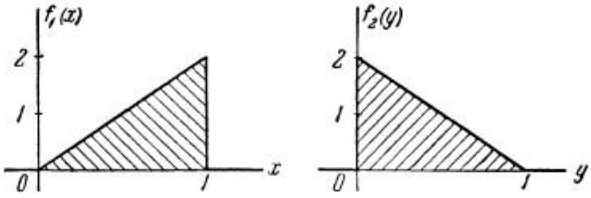


Рис. 18.7.5.

Подставляя $f(x, y)$, $f_1(x)$ и $f_2(y)$ в формулу (18.7.20), получим

$$I_{Y \leftrightarrow X} = \frac{1}{\ln 2} \int\limits_{(D)} \int \ln \frac{1}{2x(1-y)}\, 2\,dx\,dy =$$

$$= \frac{1}{\ln 2} \left\{ -\ln 2 + \int\limits_{(D)} \int 2(-\ln x)\,dx\,dy + \int\limits_{(D)} \int 2[-\ln(1-y)]\,dx\,dy \right\}.$$

В силу симметрии задачи последние два интеграла равны, и

$$I_{Y\leftrightarrow X} = -1 - \frac{4}{\ln 2}\int\limits_{(D)}\int \ln x\,dx\,dy =$$

$$= -1 - \frac{4}{\ln 2}\int\limits_0^1 \ln x\,dx \int\limits_0^x dy = -1 - \frac{2}{\ln 2}\int\limits_0^1 2x\ln x\,dx =$$

$$= \frac{1}{\ln 2} - 1 \approx 0{,}44 \text{ (дв. ед.)}.$$

Пример 5. Имеется случайная величина X, распределенная по нормальному закону с параметрами $m_x = 0$, σ_x. Величина X измеряется с ошибкой Z, тоже распределенной по нормальному закону с параметрами $m_z = 0$, σ_z. Ошибка Z не зависит от X. В нашем распоряжении — результат измерения, т. е. случайная величина

$$Y = X + Z.$$

Определить, сколько информации о величине X содержит величина Y.

Решение. Воспользуемся для вычисления информации формулой (18.7.21), т. е. найдем ее как математическое ожидание случайной величины

$$U = \log \frac{f(X, Y)}{f_1(X)\,f_2(Y)}. \qquad (18.7.22)$$

Для этого сначала преобразуем выражение

$$\log \frac{f(x, y)}{f_1(x)\,f_2(y)} = \log \frac{f_1(x)\,f(y \mid x)}{f_1(x)\,f_2(y)} = \log \frac{f(y \mid x)}{f_2(y)}.$$

В нашем случае

$$f_2(y) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2 + \sigma_z^2}}\, e^{-\frac{y^2}{2\left(\sigma_x^2+\sigma_z^2\right)}},$$

$$f(y \mid x) = \frac{1}{\sqrt{2\pi}\,\sigma_z}\, e^{-\frac{(y-x)^2}{2\sigma_z^2}} \qquad \text{(см. главу 9)}.$$

Выражение (18.7.22) равно:

$$U = \log \frac{\sqrt{\sigma_x^2+\sigma_z^2}}{\sigma_z} + \frac{1}{\ln 2}\left[\frac{(Y-X)^2}{2\sigma_z^2} - \frac{Y^2}{2\left(\sigma_x^2+\sigma_z^2\right)}\right] =$$

$$= \log \frac{\sqrt{\sigma_x^2+\sigma_z^2}}{\sigma_z} + \frac{1}{\ln 2}\left[\frac{Z^2}{2\sigma_z^2} - \frac{Y^2}{2\left(\sigma_x^2+\sigma_z^2\right)}\right].$$

Отсюда

$$I_{Y\leftrightarrow X} = M[U] = \log \frac{\sqrt{\sigma_x^2+\sigma_z^2}}{\sigma_z} + \frac{1}{\ln 2}\left[\frac{M[Z^2]}{2\sigma_z^2} - \frac{M[Y^2]}{2\left(\sigma_x^2+\sigma_z^2\right)}\right]. \qquad (18.7.23)$$

Но $m_z = m_y = 0$, следовательно,

$$\left.\begin{aligned} M[Z^2] &= D[Z] = \sigma_z^2, \\ M[Y^2] &= D[Y] = \sigma_x^2 + \sigma_z^2. \end{aligned}\right\} \qquad (18.7.24)$$

Подставляя (18.7.24) в (18.7.23), получим

$$I_{Y \leftrightarrow X} = \log \frac{\sqrt{\sigma_x^2 + \sigma_z^2}}{\sigma_z} \text{ (дв. ед.) [1].}$$

Например, при $\sigma_x = \sigma_z$

$$I_{Y \leftrightarrow X} = \log \sqrt{2} = \frac{1}{2} \text{ (дв. ед.).}$$

Если $\sigma_x = 4$; $\sigma_z = 3$, то $I_{Y \leftrightarrow X} = \log \frac{5}{3} \approx 0{,}74$ (дв. ед.).

18.8. Задачи кодирования сообщений. Код Шеннона — Фэно

При передаче сообщений по линиям связи всегда приходится пользоваться тем или иным *кодом*, т. е. представлением сообщения в виде ряда сигналов. Общеизвестным примером кода может служить принятая в телеграфии для передачи словесных сообщений азбука Морзе. С помощью этой азбуки любое сообщение представляется в виде комбинации элементарных сигналов: точка, тире, пауза (пробел между буквами), длинная пауза (пробел между словами).

Вообще *кодированием* называется отображение состояния одной физической системы с помощью состояния некоторой другой. Например, при телефонном разговоре звуковые сигналы кодируются в виде электромагнитных колебаний, а затем снова декодируются, превращаясь в звуковые сигналы на другом конце линии. Наиболее простым случаем кодирования является случай, когда обе системы X и Y (отображаемая и отображающая) имеют конечное число возможных состояний. Так обстоит дело при передаче записанных буквами сообщений, например, при телеграфировании. Мы ограничимся рассмотрением этого простейшего случая кодирования.

Пусть имеется некоторая система X (например, буква русского алфавита), которая может случайным образом принять одно из состояний $x_1, x_2, \ldots, x_n$. Мы хотим отобразить ее (закодировать) с помощью другой системы Y, возможные состояния которой $y_1, y_2, \ldots, y_m$. Если $m < n$ (число состояний системы Y меньше числа состояний системы X), то нельзя каждое состояние системы X закодировать с помощью одного-единственного состояния системы Y. В таких случаях одно состояние системы X приходится отображать с помощью определенной комбинации (последовательности) состояний системы Y. Так, в азбуке Морзе буквы отображаются различными комбинациями эле-

[1]) Применяя формулу (18.7.20), можно было бы прийти к тому же результату, но более громоздким способом.

ментарных символов (точка, тире). Выбор таких комбинаций и установление соответствия между передаваемым сообщением и этими комбинациями и называется «кодированием» в узком смысле слова.

Коды различаются по числу элементарных символов (сигналов), из которых формируются комбинации, иными словами — по числу возможных состояний системы Y. В азбуке Морзе таких элементарных символов четыре (точка, тире, короткая пауза, длинная пауза). Передача сигналов может осуществляться в различной форме: световые вспышки, посылки электрического тока различной длительности, звуковые сигналы и т. п. Код с двумя элементарными символами (0 и 1) называется *двоичным*. Двоичные коды широко применяются на практике, особенно при вводе информации в электронные цифровые вычислительные машины, работающие по двоичной системе счисления.

Одно и то же сообщение можно закодировать различными способами. Возникает вопрос об оптимальных (наивыгоднейших) способах кодирования. Естественно считать наивыгоднейшим такой код, при котором на передачу сообщений затрачивается минимальное время. Если на передачу каждого элементарного символа (например 0 или 1) тратится одно и то же время, то оптимальным будет такой код, при котором на передачу сообщения заданной длины будет затрачено минимальное количество элементарных символов.

Предположим, что перед нами поставлена задача: закодировать двоичным кодом буквы русской азбуки так, чтобы каждой букве соответствовала определенная комбинация элементарных символов 0 и 1 и чтобы среднее число этих символов на букву текста было минимальным.

Рассмотрим 32 буквы русской азбуки: а, б, в, г, д, е, ж, з, и, й, к, л, м, н, о, п, р, с, т, у, ф, х, ц, ч, ш, щ, ъ, ы, ь, э, ю, я плюс промежуток между словами, который мы будем обозначать «—». Если, как принято в телеграфии, не различать букв ъ и ь (это не приводит к разночтениям), то получится 32 буквы: а, б, в, г, д, е, ж, з, и, й, к, л, м, н, о, п, р, с, т, у, ф, х, ц, ч, ш, щ, (ъ, ь), ы, э, ю, я, «—».

Первое, что приходит в голову — это, не меняя порядка букв, занумеровать их подряд, приписав им номера от 0 до 31, и затем перевести нумерацию в двоичную систему счисления. Двоичная система — это такая, в которой единицы разных разрядов представляют собой разные степени двух. Например, десятичное число 12 изобразится в виде

$$12 = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0$$

и в двоичной системе запишется как 1100.

Десятичное число 25 —

$$25 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$$

— запишется в двоичной системе как 11001.

Каждое из чисел 0, 1, 2, ..., 31 может быть изображено пятизначным двоичным числом. Тогда получим следующий код:

а ~ 00000
б ~ 00001
в ~ 00010
г ~ 00011
....
я ~ 11110
«—» ~ 11111

В этом коде на изображение каждой буквы тратится ровно 5 элементарных символов. Возникает вопрос, является ли этот простейший код оптимальным и нельзя ли составить другой код, в котором на одну букву будет в среднем приходиться меньше элементарных символов?

Действительно, в нашем коде на изображение каждой буквы — часто встречающихся «а», «е», «о» или редко встречающихся «щ», «э», «ф» — тратится одно и то же число элементарных символов. Очевидно, разумнее было бы, чтобы часто встречающиеся буквы были закодированы меньшим числом символов, а реже встречающиеся — бóльшим.

Чтобы составить такой код, очевидно, нужно знать частоты букв в русском тексте. Эти частоты приведены в таблице 18.8.1. Буквы в таблице расположены в порядке убывания частот.

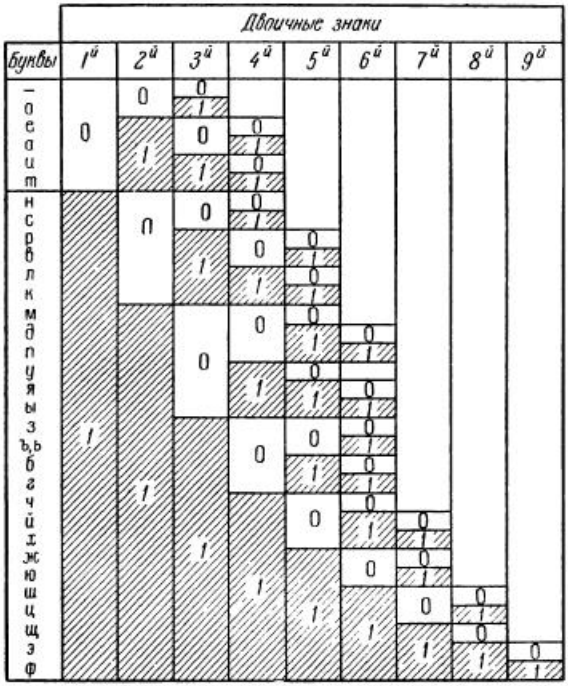
Таблица 18.8.1

Буква	Частота	Буква	Частота	Буква	Частота	Буква	Частота
«—»	0,145	р	0,041	я	0,019	х	0,009
о	0,095	в	0,039	ы	0,016	ж	0,008
е	0,074	л	0,036	з	0,015	ю	0,007
а	0,064	к	0,029	ъ, ь	0,015	ш	0,006
и	0,064	м	0,026	б	0,015	ц	0,004
т	0,056	д	0,026	г	0,014	щ	0,003
н	0,056	п	0,024	ч	0,013	э	0,003
с	0,047	у	0,021	й	0,010	ф	0,002

Пользуясь такой таблицей, можно составить наиболее экономичный код на основе соображений, связанных с количеством информации. Очевидно, код будет самым экономичным, когда каждый элементарный символ будет передавать максимальную информацию. Рассмотрим элементарный символ (т. е. изображающий его сигнал) как физическую систему с двумя возможными состояниями: 0 и 1.

Информация, которую дает этот символ, равна энтропии системы и максимальна в случае, когда оба состояния равновероятны; в этом случае элементарный символ передает информацию 1 (дв. ед.). Поэтому основой оптимального кодирования будет требование, чтобы элементарные символы в закодированном тексте встречались в среднем одинаково часто.

Таблица 18.8.2



Изложим здесь способ построения кода, удовлетворяющего поставленному условию; этот способ известен под названием «кода Шеннона — Фэно». Идея его состоит в том, что кодируемые символы (буквы или комбинации букв) разделяются на две приблизительно равновероятные группы: для первой группы символов на первом месте комбинации ставится 0 (первый знак двоичного числа, изображающего символ); для второй группы — 1. Далее каждая группа снова делится на две приблизительно равновероятные подгруппы; для символов первой подгруппы на втором месте ставится нуль; для второй подгруппы — единица и т. д.

Продемонстрируем принцип построения кода Шеннона — Фэно на материале русского алфавита (табл. 18.8.1). Отсчитаем первые шесть

букв (от «—» до «т»); суммируя их вероятности (частоты), получим 0,498; на все остальные буквы (от «н» до «ф») придется приблизительно такая же вероятность 0,502. Первые шесть букв (от «—» до «т») будут иметь на первом месте двоичный знак 0. Остальные буквы (от «н» до «ф») будут иметь на первом месте единицу. Далее снова разделим первую группу на две приблизительно равновероятные подгруппы: от «—» до «о» и от «е» до «т»; для всех букв первой подгруппы на втором месте поставим нуль, а второй подгруппы — единицу. Процесс будем продолжать до тех пор, пока в каждом подразделении не останется ровно одна буква, которая и будет закодирована определенным двоичным числом. Механизм построения кода показан на таблице 18.8.2, а сам код приведен в таблице 18.8.3.

Таблица 18.8.3

Буква	Двоичное число	Буква	Двоичное число	Буква	Двоичное число
«—»	000	к	10111	ч	111100
о	001	м	11000	й	1111010
е	0100	д	110010	х	1111011
а	0101	п	110011	ж	1111100
и	0110	у	110100	ю	1111101
т	0111	я	110110	ш	11111100
н	1000	ы	110111	ц	11111101
с	1001	з	111000	щ	11111110
р	10100	ъ, ь	111001	э	111111110
в	10101	б	111010	ф	111111111
л	10110	г	111011		

С помощью таблицы 18.8.3 можно закодировать и декодировать любое сообщение.

В виде примера запишем двоичным кодом фразу: «теория информации»

01110100001101000110110110000
01101000111111111100110100
11000010111111101011001l0

Заметим, что здесь нет необходимости отделять друг от друга буквы специальным знаком, так как и без этого декодирование выполняется однозначно. В этом можно убедиться, декодируя с помощью таблицы 18.8.2 следующую фразу:

10011100110011001001111010000
10111001110010011010100001l0101
010110000110110110

(«способ кодирования»).

Однако необходимо отметить, что любая ошибка при кодировании (случайное перепутывание знаков 0 и 1) при таком коде губительна, так как декодирование всего следующего за ошибкой текста становится невозможным. Поэтому данный принцип кодирования может быть рекомендован только в случае, когда ошибки при кодировании и передаче сообщения практически исключены.

Возникает естественный вопрос: а является ли составленный нами код при отсутствии ошибок действительно оптимальным? Для того чтобы ответить на этот вопрос, найдем среднюю информацию, приходящуюся на один элементарный символ (0 или 1), и сравним ее с максимально возможной информацией, которая равна одной двоичной единице. Для этого найдем сначала среднюю информацию, содержащуюся в одной букве передаваемого текста, т. е. энтропию на одну букву:

$$H(\text{б}) = -\sum_{i=1}^{32} p_i \log p_i = \sum_{i=1}^{32} \eta(p_i),$$

где p_i — вероятность того, что буква примет определенное состояние («—», о, е, а, ..., ф).

Из табл. 18.8.1 имеем

$$H(\text{б}) = \eta(0{,}145) + \eta(0{,}095) + \ldots + \eta(0{,}003) + \eta(0{,}002) \approx 4{,}42$$

(дв. единиц на букву текста).

По таблице 18.8.2 определяем среднее число элементарных символов на букву

$$n_{\text{ср}} = 3 \cdot 0{,}145 + 3 \cdot 0{,}095 + 4 \cdot 0{,}074 + \ldots$$
$$\ldots + 9 \cdot 0{,}003 + 9 \cdot 0{,}002 = 4{,}45.$$

Деля энтропию $H(\text{б})$ на $n_{\text{ср}}$, получаем информацию на один элементарный символ

$$I_{1\text{с}} = \frac{4{,}42}{4{,}45} \approx 0{,}994 \text{ (дв. ед.)}.$$

Таким образом, информация на один символ весьма близка к своему верхнему пределу 1, а выбранный нами код весьма близок к оптимальному. Оставаясь в пределах задачи кодирования по буквам, мы ничего лучшего получить не сможем.

Заметим, что в случае кодирования просто двоичных номеров букв мы имели бы изображение каждой буквы пятью двоичными знаками и информация на один символ была бы

$$I_{1\text{с}} = \frac{4{,}42}{5{,}00} = 0{,}884 \text{ (дв. ед.)},$$

т. е. заметно меньше, чем при оптимальном буквенном кодировании.

Однако надо заметить, что кодирование «по буквам» вообще не является экономичным. Дело в том, что между соседними буквами

любого осмысленного текста всегда имеется з а в и с и м о с т ь. Например, после гласной буквы в русском языке не может стоять «ъ» или «ь»; после шипящих не могут стоять «я» или «ю»; после нескольких согласных подряд увеличивается вероятность гласной и т. д.

Мы знаем, что при объединении зависимых систем суммарная энтропия меньше суммы энтропий отдельных систем; следовательно, информация, передаваемая отрезком связного текста, всегда меньше, чем информация на один символ, умноженная на число символов. С учетом этого обстоятельства более экономный код можно построить, если кодировать не каждую букву в отдельности, а целые «блоки» из букв. Например, в русском тексте имеет смысл кодировать целиком некоторые часто встречающиеся комбинации букв, как «тся», «ает», «ние» и т. п. Кодируемые блоки располагаются в порядке убывания частот, как буквы в табл. 18.8.1, а двоичное кодирование осуществляется по тому же принципу.

В ряде случаев оказывается разумным кодировать даже не блоки из букв, а целые осмысленные куски текста. Например, для разгрузки телеграфа в предпраздничные дни целесообразно кодировать условными номерами целые стандартные тексты, вроде:

«поздравляю новым годом желаю здоровья успехов работе».

Не останавливаясь специально на методах кодирования блоками, ограничимся тем, что сформулируем относящуюся сюда теорему Шеннона.

Пусть имеется источник информации X и приемник Y, связанные каналом связи K (рис. 18.8.1).

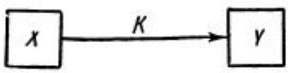


Рис. 18.8.1.

Известна *производительность* источника информации $H_1(X)$, т. е. среднее количество двоичных единиц информации, поступающее от источника в единицу времени (численно оно равно средней энтропии сообщения, производимого источником в единицу времени). Пусть, кроме того, известна *пропускная способность* канала C_1, т. е. максимальное количество информации (например, двоичных знаков 0 или 1), которое способен передать канал в ту же единицу времени. Возникает вопрос: какова должна быть пропускная способность канала, чтобы он «справлялся» со своей задачей, т. е. чтобы информация от источника X к приемнику Y поступала без задержки?

Ответ на этот вопрос дает первая теорема Шеннона. Сформулируем ее здесь без доказательства.

1-я т е о р е м а Ш е н н о н а

Если пропускная способность канала связи C_1 больше энтропии источника информации в единицу времени

$$C_1 > H_1(X),$$

то всегда можно закодировать достаточно длинное сообщение так, чтобы оно передавалось каналом связи без задержки. Если же, напротив,

$$C_1 < H_1(X),$$

то передача информации без задержек невозможна.

18.9. Передача информации с искажениями. Пропускная способность канала с помехами

В предыдущем п° мы рассмотрели вопросы, связанные с кодированием и передачей информации по каналу связи в идеальном случае, когда процесс передачи информации осуществляется без ошибок. В действительности этот процесс неизбежно сопровождается ошибками (искажениями). Канал передачи, в котором возможны искажения, называется *каналом с помехами* (или шумами). В частном случае ошибки возникают в процессе самого кодирования, и тогда кодирующее устройство может рассматриваться как канал с помехами.

Совершенно очевидно, что наличие помех приводит к потере информации. Чтобы в условиях наличия помех получить на приемнике требуемый объем информации, необходимо принимать специальные меры. Одной из таких мер является введение так называемой «избыточности» в передаваемые сообщения; при этом источник информации выдает заведомо больше символов, чем это было бы нужно при отсутствии помех. Одна из форм введения избыточности — простое повторение сообщения. Таким приемом пользуются, например, при плохой слышимости по телефону, повторяя каждое сообщение дважды. Другой общеизвестный способ повышения надежности передачи состоит в передаче слова «по буквам» — когда вместо каждой буквы передается хорошо знакомое слово (имя), начинающееся с этой буквы.

Заметим, что все живые языки естественно обладают некоторой избыточностью. Эта избыточность часто помогает восстановить правильный текст «по смыслу» сообщения. Вот почему встречающиеся вообще нередко искажения отдельных букв телеграмм довольно редко приводят к действительной потере информации: обычно удается исправить искаженное слово, пользуясь одними только свойствами языка. Этого не было бы при отсутствии избыточности. Мерой избыточности языка служит величина

$$U = 1 - \frac{H_{1c}}{\log n}, \tag{18.9.1}$$

где H_{1c} — средняя фактическая энтропия, приходящаяся на один передаваемый символ (букву), рассчитанная для достаточно длинных отрывков текста, с учетом зависимости между символами, n — число

применяемых символов (букв), $\log n$ — максимально возможная в данных условиях энтропия на один передаваемый символ, которая была бы, если бы все символы были равновероятны и независимы.

Расчеты, проведенные на материале наиболее распространенных европейских языков, показывают, что их избыточность достигает 50% и более (т. е., грубо говоря, 50% передаваемых символов являются лишними и могли бы не передаваться, если бы не опасность искажений).

Однако для безошибочной передачи сведений естественная избыточность языка может оказаться как чрезмерной, так и недостаточной: все зависит от того, как велика опасность искажений («уровень помех») в канале связи.

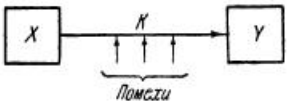


Рис. 18.9.1.

С помощью методов теории информации можно для каждого уровня помех найти нужную степень избыточности источника информации. Те же методы помогают разрабатывать специальные помехоустойчивые коды (в частности, так называемые «самокорректирующиеся» коды). Для решения этих задач нужно уметь учитывать потерю информации в канале, связанную с наличием помех.

Рассмотрим сложную систему, состоящую из источника информации X, канала связи K и приемника Y (рис. 18.9.1). Источник информации представляет собой физическую систему X, которая имеет n возможных состояний

$$x_1, \ x_2, \ \ldots, \ x_n$$

с вероятностями

$$p_1, \ p_2, \ \ldots, \ p_n.$$

Будем рассматривать эти состояния как элементарные символы, которые может передавать источник X через канал K к приемнику Y. Количество информации на один символ, которое дает источник, будет равно энтропии на один символ:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i.$$

Если бы передача сообщений не сопровождалась ошибками, то количество информации, содержащееся в системе Y относительно X, было бы равно самой энтропии системы X. При наличии ошибок оно будет меньше:

$$I_{Y \leftrightarrow X} = H(X) - H(X \mid Y).$$

Естественно рассматривать условную энтропию $H(X \mid Y)$ как *потерю информации* на один элементарный символ, связанную с наличием помех.

Умея определять потерю информации в канале, приходящуюся на один элементарный символ, переданный источником информации, можно определить пропускную способность канала с помехами, т. е. максимальное количество информации, которое способен передать канал в единицу времени.

Предположим, что канал может передавать в единицу времени k элементарных символов. В отсутствие помех пропускная способность канала была бы равна

$$C = k \log n, \tag{18.9.2}$$

так как максимальное количество информации, которое может содержать один символ, равно $\log n$, а максимальное количество информации, которое могут содержать k символов, равно $k \log n$, и оно достигается, когда символы появляются независимо друг от друга.

Теперь рассмотрим канал с помехами. Его пропускная способность определится как

$$C = k \max I^{(1)}_{Y \to X}, \tag{18.9.3}$$

где $\max I^{(1)}_{Y \to X}$ — максимальная информация на один символ, которую может передать канал при наличии помех.

Определение этой максимальной информации в общем случае — дело довольно сложное, так как она зависит от того, как и с какими вероятностями искажаются символы; происходит ли их перепутывание, или же простое выпадение некоторых символов; происходят ли искажения символов независимо друг от друга и т. д.

Однако для простейших случаев пропускную способность канала удается сравнительно легко рассчитать.

Рассмотрим, например, такую задачу. Канал связи K передает от источника информации X к приемнику Y элементарные символы 0 и 1 в количестве k символов в единицу времени. В процессе передачи каждый символ, независимо от других, с вероятностью μ может быть искажен (т. е. заменен противоположным). Требуется найти пропускную способность канала.

Определим сначала максимальную информацию на один символ, которую может передавать канал. Пусть источник производит символы 0 и 1 с вероятностями p и $1-p$.

Тогда энтропия источника будет

$$H(X) = -p \log p - (1-p) \log (1-p).$$

Определим информацию $I^{(1)}_{Y \to X}$ на один элементарный символ:

$$I^{(1)}_{Y \to X} = H(Y) - H(Y|X).$$

Чтобы найти полную условную энтропию $H(Y|X)$, найдем сначала частные условные энтропии: $H(Y|x_1)$ (энтропию системы Y при условии, что система X приняла состояние x_1) и $H(Y|x_j$

(энтропию системы Y при условии, что система X приняла состояние x_2). Вычислим $H(Y|x_1)$; для этого предположим, что передан элементарный символ 0. Найдем условные вероятности того, что при этом система Y находится в состоянии $y_1=0$ и в состоянии $y_2=1$. Первая из них равна вероятности того, что сигнал не перепутан:

$$P(y_1|x_1)=1-\mu;$$

вторая — вероятности того, что сигнал перепутан:

$$P(y_2|x_1)=\mu.$$

Условная энтропия $H(Y|x_1)$ будет:

$$H(Y|x_1)=-\sum_{i=1}^{2}P(y_i|x_1)\log P(y_i|x_1)=$$

$$=-(1-\mu)\log(1-\mu)-\mu\log\mu.$$

Найдем теперь условную энтропию системы Y при условии, что $X\sim x_2$ (передан сигнал единица):

$$P(y_1|x_2)=\mu;\qquad P(y_2|x_2)=1-\mu,$$

откуда

$$H(Y|x_2)=-\mu\log\mu-(1-\mu)\log(1-\mu).$$

Таким образом,

$$H(Y|x_1)=H(Y|x_2)=-\mu\log\mu-(1-\mu)\log(1-\mu). \quad (18.9.4)$$

Полная условная энтропия $H(Y|X)$ получится, если осреднить условные энтропии $H(Y|x_1)$ и $H(Y|x_2)$ с учетом вероятностей p и $1-p$ значений x_1, x_2. Так как частные условные энтропии равны, то

$$H(Y|X)=-\mu\log\mu-(1-\mu)\log(1-\mu).$$

Мы получили следующий вывод: *условная энтропия $H(Y|X)$ совсем не зависит от того, с какими вероятностями p, $1-p$ встречаются символы* 0; 1 *в передаваемом сообщении, а зависит только от вероятности ошибки* μ.

Вычислим полную информацию, передаваемую одним символом:

$$I^{(1)}_{Y\to X}=H(Y)-H(Y|X)=$$

$$=[-r\log r-(1-r)\log(1-r)]-$$

$$-[-\mu\log\mu-(1-\mu)\log(1-\mu)]=$$

$$=[\eta(r)+\eta(1-r)]-[\eta(\mu)+\eta(1-\mu)],$$

где r — вероятность того, что на выходе появится символ 0. Очевидно, при заданных свойствах канала информация на один символ $I^{(1)}_{Y\to X}$ достигает максимума, когда $\eta(r)+\eta(1-r)$ максимально. Мы знаем, что такая функция достигает максимума при $r=1/2$, т. е. когда на приемнике оба сигнала равновероятны. Легко убедиться, что это достигается, когда источник передает оба символа с одинаковой вероятностью $p=1/2$. При том же значении $p=1/2$ достигает максимума и информация на один символ. Максимальное значение равно

$$I^{(1)}_{Y\to X}=H(Y)-[\eta(\mu)+\eta(1-\mu)]=1-[\eta(\mu)+\eta(1-\mu)].$$

Следовательно, в нашем случае

$$\max I^{(1)}_{Y\to X}=1-[\eta(\mu)+\eta(1-\mu)],$$

и пропускная способность канала связи будет равна

$$C=k\{1-[\eta(\mu)+\eta(1-\mu)]\}. \qquad (18.9.5)$$

Заметим, что $\eta(\mu)+\eta(1-\mu)$ есть не что иное, как энтропия системы, имеющей два возможных состояния с вероятностями μ и $1-\mu$. Она характеризует потерю информации на один символ, связанную с наличием помех в канале.

Пример. 1. Определить пропускную способность канала связи, способного передавать 100 символов 0 или 1 в единицу времени, причем каждый из символов искажается (заменяется противоположным) с вероятностью $\mu=0{,}01$.

Решение. По таблице 7 приложения находим

$$\eta(\mu)=0{,}0664,$$
$$\eta(1-\mu)=0{,}0144,$$
$$\eta(\mu)+\eta(1-\mu)=0{,}0808.$$

На один символ теряется информация 0,0808 (дв. ед). Пропускная способность канала равна

$$C=100\,(1-0{,}0808)=91{,}92\approx 92$$

двоичные единицы в единицу времени.

С помощью аналогичных расчетов может быть определена пропускная способность канала и в более сложных случаях: когда число элементарных символов более двух и когда искажения отдельных символов зависимы. Зная пропускную способность канала, можно определить верхний предел скорости передачи информации по каналу с помехами. Сформулируем (без доказательства) относящуюся к этому случаю вторую теорему Шеннона.

2-я теорема Шеннона

Пусть имеется источник информации X, энтропия которого в единицу времени равна $\tilde{H}(X)$, и канал с пропускной способностью C. Тогда если

$$\tilde{H}(X)>C,$$

то при любом кодировании передача сообщений без задержек и искажений невозможна. Если же

$$\tilde{H}(X) < C,$$

то всегда можно достаточно длинное сообщение закодировать так, чтобы оно было передано без задержек и искажений с вероятностью, сколь угодно близкой к единице.

Пример 2. Имеются источник информации с энтропией в единицу времени $\tilde{H}(X) = 100$ (дв. ед.) и два канала связи; каждый из них может передавать в единицу времени 70 двоичных знаков (0 или 1); каждый двоичный знак заменяется противоположным с вероятностью $\mu = 0{,}1$. Требуется выяснить: достаточна ли пропускная способность этих каналов для передачи информации, поставляемой источником?

Решение. Определяем потерю информации на один символ:

$$\eta(\mu) + \eta(1-\mu) = 0{,}332 + 0{,}137 = 0{,}469 \text{ (дв. ед.)}.$$

Максимальное количество информации, передаваемое по одному каналу в единицу времени:

$$C = 70(1 - 0{,}469) = 37{,}2.$$

Максимальное количество информации, которое может быть передано по двум каналам в единицу времени:

$$37{,}2 \cdot 2 = 74{,}4 \text{ (дв. ед.)},$$

чего недостаточно для обеспечения передачи информации от источника.

ГЛАВА 19
ЭЛЕМЕНТЫ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ

19.1. Предмет теории массового обслуживания

За последние десятилетия в самых разных областях практики возникла необходимость в решении своеобразных вероятностных задач, связанных с работой так называемых *систем массового обслуживания*. Примерами таких систем могут служить: телефонные станции, ремонтные мастерские, билетные кассы, справочные бюро, парикмахерские и т. п. Каждая такая система состоит из какого-то числа обслуживающих единиц, которые мы будем называть «каналами» обслуживания. В качестве каналов могут фигурировать: линии связи; лица, выполняющие те ли иные операции; различные приборы и т. п. Системы массового обслуживания могут быть как одно-,так и многоканальными.

Работа любой системы массового обслуживания состоит в выполнении поступающего на нее потока требований или з а я в о к. Заявки поступают одна за другой в некоторые, вообще говоря, случайные, моменты времени. Обслуживание поступившей заявки продолжается какое-то время, после чего канал освобождается и снова готов для приема следующей заявки. Каждая система массового обслуживания, в зависимости от числа каналов и их производительности, обладает какой-то пропускной способностью, позволяющей ей более или менее успешно справляться с потоком заявок. Предмет теории массового обслуживания — установление зависимости между характером потока заявок, производительностью отдельного канала, числом каналов и успешностью (эффективностью) обслуживания. В качестве характеристик эффективности обслуживания — в зависимости от условий задачи и целей исследования — могут применяться различные величины и функции, например: средний процент заявок, получающих отказ и покидающих систему необслуженными; среднее время «простоя» отдельных каналов и системы в целом; среднее время ожидания в очереди; вероятность того, что поступившая заявка немедленно будет принята к обслуживанию; закон распределения длины очереди и т. д. Каждая из этих характеристик описывает, с той или другой стороны, степень приспособленности системы к выполнению потока заявок, иными словами — ее пропускную способность.

Под «пропускной способностью» в узком смысле слова обычно понимают среднее число заявок, которое система может обслужить в единицу времени. Наряду с нею часто рассматривают относительную пропускную способность — среднее отношение числа обслуженных заявок к числу поданных. Пропускная способность (как абсолютная, так и относительная) в общем случае зависит не только от параметров системы, но и от характера потока заявок. Если бы заявки поступали регулярно, через точно определенные промежутки времени, и обслуживание каждой заявки тоже имело строго определенную длительность, расчет пропускной способности системы не представлял бы никакой трудности. На практике обычно моменты поступления заявок случайны; по большей части случайна и длительность обслуживания заявки. В связи с этим процесс работы системы протекает нерегулярно: в потоке заявок образуются местные сгущения и разрежения. Сгущения могут привести либо к отказам в обслуживании, либо к образованию очередей. Разрежения могут привести к непроизводительным простоям отдельных каналов или системы в целом. На эти случайности, связанные с неоднородностью потока заявок, накладываются еще случайности, связанные с задержками обслуживания отдельных заявок. Таким образом, процесс функционирования системы массового обслуживания представляет собой *случайный процесс*. Чтобы дать рекомендации по рациональной организации системы, выяснить ее пропускную способность и предъявить к ней требования, необходимо изучить случайный процесс, протекающий в системе, и описать его математически. Этим и занимается теория массового обслуживания.

Заметим, что за последние годы область применения математических методов теории массового обслуживания непрерывно расширяется и все больше выходит за пределы задач, связанных с «обслуживающими организациями» в буквальном смысле слова. Многие задачи автоматизации производства оказываются близкими к теории массового обслуживания: потоки деталей, поступающих для выполнения над ними различных операций, могут рассматриваться как «потоки заявок», ритмичность поступления которых нарушается за счет случайных причин. Своеобразные задачи теории массового обслуживания возникают в связи с проблемой организации транспорта и системы сообщений. Близкими к теории массового обслуживания оказываются и задачи, относящиеся к надежности технических устройств: такие их характеристики, как среднее время безотказной работы, потребное количество запасных деталей, среднее время простоя в связи с ремонтом и т. д., определяются методами, непосредственно заимствованными из теории массового обслуживания.

Проблемы, родственные задачам массового обслуживания, постоянно возникают в военном деле. Каналы наведения, линии связи, аэродромы, устройства для сбора и обработки информации представляют собой

своеобразные системы массового обслуживания со своим режимом работы и пропускной способностью.

Трудно даже перечислить все области практики, в которых находят применение методы теории массового обслуживания. За последние годы она стала одной из самых быстро развивающихся ветвей теории вероятностей.

В настоящей главе будут изложены некоторые элементарные сведения по теории массового обслуживания, знание которых необходимо любому инженеру, занимающемуся вопросами организации в области промышленности, народного хозяйства, связи, а также в военном деле.

19.2. Случайный процесс со счетным множеством состояний

Случайный процесс, протекающий в системе массового обслуживания, состоит в том, что система в случайные моменты времени переходит из одного состояния в другое: меняется число занятых каналов, число заявок, стоящих в очереди, и т. п. Такой процесс существенно отличается от случайных процессов, которые мы рассматривали в главах 15—17. Дело в том, что система массового обслуживания представляет собой физическую систему дискретного типа с конечным (или счетным) множеством состояний [1]), а переход системы из одного состояния в другое происходит скачком, в момент, когда осуществляется какое-то событие (приход новой заявки, освобождение канала, уход заявки из очереди и т. п.).

Рассмотрим физическую систему X со счетным множеством состояний

$$x_1, x_2, \ldots, x_n, \ldots$$

В любой момент времени t система X может быть в одном из этих состояний. Обозначим $p_k(t)$ $(k = 1, 2, \ldots, n, \ldots)$ вероятность того, что в момент t система будет находиться в состоянии x_k. Очевидно, для любого t

$$\sum_k p_k(t) = 1. \qquad (19.2.1)$$

Совокупность вероятностей $p_k(t)$ для каждого момента времени t характеризует данное *сечение* случайного процесса, протекающего в системе. Эта совокупность не является исчерпывающей характеристикой процесса (она, например, совсем не отражает зависимости между сечениями), но все же достаточно хорошо описывает процесс и для ряда практических применений оказывается достаточной.

Случайные процессы со счетным множеством состояний бывают двух типов: с дискретным или непрерывным временем. Первые отличаются тем, что переходы из состояния в состояние могут происходить

[1]) В математике «счетным» называется конечное или бесконечное множество, члены которого можно перенумеровать, т. е. записать в виде последовательности $a_1, a_2, \ldots, a_n, \ldots$

только в строго определенные, разделенные конечными интервалами моменты времени $t_1, t_2, \ldots$ Случайные процессы с непрерывным временем отличаются тем, что переход системы из состояния в состояние возможен в любой момент времени t.

В качестве примера дискретной системы X, в которой протекает случайный процесс с непрерывным временем, рассмотрим группу из n самолетов, совершающих налет на территорию противника, обороняемую истребительной авиацией. Ни момент обнаружения группы, ни моменты подъема по ней истребителей заранее не известны. Различные состояния системы соответствуют различному числу пораженных самолетов в составе группы:

x_0 — не поражено ни одного самолета,
x_1 — поражен ровно один самолет,
. . . . .
x_k — поражено ровно k самолетов,
. . . . .
x_n — поражены все n самолетов.

Схема возможных состояний системы и возможных переходов из состояния в состояние показана на рис. 19.2.1.

Стрелками показаны возможные переходы системы из состояния в состояние. Закругленная стрелка, направленная из состояния x_k в него же, означает, что система может не только перейти в соседнее

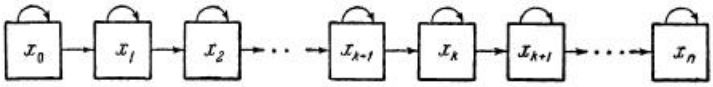


Рис. 19.2.1.

состояние x_{k+1}, но и остаться в прежнем. Для данной системы характерны необратимые переходы (пораженные самолеты не восстанавливаются); в связи с этим из состояния x_n никакие переходы в другие состояния уже невозможны.

Отметим, что на схеме возможных переходов (рис. 19.2.1) показаны только переходы из состояния в соседнее состояние и не показаны «перескоки» через состояние: эти перескоки отброшены как практически невозможные. Действительно, для того чтобы система «перескочила» через состояние, нужно, чтобы строго одновременно были поражены два или более самолета, а вероятность такого события равна нулю.

Случайные процессы, протекающие в системах массового обслуживания, как правило, представляют собой процессы с непрерывным временем. Это связано со случайностью потока заявок. В противоположность системе с необратимыми переходами, рассмотренной в предыдущем примере, для системы массового обслуживания характерны

обратимые переходы: занятый канал может освободиться, очередь может «рассосаться».

В качестве примера рассмотрим одноканальную систему массового обслуживания (например, одну телефонную линию), в которой заявка, заставшая канал занятым, не становится в очередь, а покидает систему (получает «отказ»). Это — дискретная система с непрерывным временем и двумя возможными состояниями:

x_0 — канал свободен,
x_1 — канал занят.

Переходы из состояния в состояние обратимы. Схема возможных переходов показана на рис. 19.2.2.

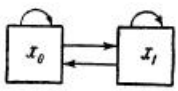


Рис. 19.2.2.

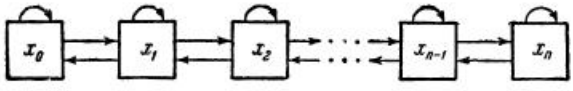


Рис. 19.2.3.

Для n-канальной системы такого же типа схема возможных переходов показана на рис. 19.2.3. Состояние x_0 — все каналы свободны; x_1 — занят ровно один канал, x_2 — занято ровно два канала и т. д.

Рассмотрим еще один пример дискретной системы с непрерывным временем: одноканальную систему массового обслуживания, которая может находиться в четырех состояниях:

x_0 — канал исправен и свободен,
x_1 — канал исправен и занят,
x_2 — канал неисправен и ждет ремонта,
x_3 — канал неисправен и ремонтируется.

Схема возможных переходов для этого случая показана на рис. 19.2.4 [1]). Переход системы из x_3 непосредственно в x_1, минуя x_0, можно считать практически невозможным, так как для этого нужно, чтобы окончание ремонта и приход очередной заявки произошли строго в один и тот же момент времени.

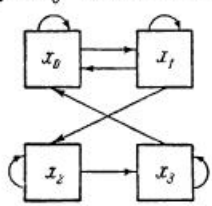


Рис. 19.2.4.

Для того чтобы описать случайный процесс, протекающий в дискретной системе с непрерывным временем, прежде всего нужно проанализировать причины, вызывающие переход системы из состояния в состояние. Для системы массового обслуживания основным фактором, обусловливающим протекающие в ней процессы, является поток заявок. Поэтому математическое описание любой системы массового обслуживания начинается с описания потока заявок.

[1]) Схема составлена в предположении, что неработающий канал выйти из строя не может.

19.3. Поток событий. Простейший поток и его свойства

Под *потоком событий* в теории вероятностей понимается последовательность событий, происходящих одно за другим в какие-то моменты времени. Примерами могут служить: поток вызовов на телефонной станции; поток включений приборов в бытовой электросети; поток заказных писем, поступающих в почтовое отделение; поток сбоев (неисправностей) электронной вычислительной машины; поток выстрелов, направляемых на цель во время обстрела, и т. п. События, образующие поток, в общем случае могут быть различными, но здесь мы будем рассматривать лишь поток *однородных событий*, различающихся только моментами появления. Такой поток можно изобразить как последовательность точек t_1, t_2, ..., t_k, ... на числовой оси (рис. 19.3.1), соответствующих моментам появления событий.



Рис. 19.3.1.

Поток событий называется *регулярным*, если события следуют одно за другим через строго определенные промежутки времени. Такой поток сравнительно редко встречается в реальных системах, но представляет интерес как предельный случай. Типичным для системы массового обслуживания является случайный поток заявок.

В настоящем п° мы рассмотрим потоки событий, обладающие некоторыми особенно простыми свойствами. Для этого введем ряд определений.

1. Поток событий называется *стационарным*, если вероятность попадания того или иного числа событий на участок времени длиной τ (рис. 19.3.1) зависит только от длины участка и не зависит от того, где именно на оси Ot расположен этот участок.

2. Поток событий называется *потоком без последействия*, если для любых неперекрывающихся участков времени число событий, попадающих на один из них, не зависит от числа событий, попадающих на другие.

3. Поток событий называется *ординарным*, если вероятность попадания на элементарный участок Δt двух или более событий пренебрежимо мала по сравнению с вероятностью попадания одного события.

Если поток событий обладает всеми тремя свойствами (т. е. стационарен, ординарен и не имеет последействия), то он называется *простейшим* (или стационарным пуассоновским) потоком. Название «пуассоновский» связано с тем, что при соблюдении условий 1 ÷ 3 число событий, попадающих на любой фиксированный интервал времени, будет распределено по закону Пуассона (см. п° 5.9).

Рассмотрим подробнее условия 1 ÷ 3, посмотрим, чему они соответствуют для потока заявок и за счет чего они могут нарушаться.

1. Условию стационарности удовлетворяет поток заявок, вероятностные характеристики которого не зависят от времени. В частности, для стационарного потока характерна постоянная плотность (среднее число заявок в единицу времени). На практике часто встречаются потоки заявок, которые (по крайней мере, на ограниченном отрезке времени) могут рассматриваться как стационарные. Например, поток вызовов на городской телефонной станции на участке времени от 12 до 13 часов может считаться стационарным. Тот же поток в течение целых суток уже не может считаться стационарным (ночью плотность вызовов значительно меньше, чем днем). Заметим, что так обстоит дело и со всеми физическими процессами, которые мы называем «стационарными»: в действительности все они стационарны лишь на ограниченном участке времени, а распространение этого участка до бесконечности — лишь удобный прием, применяемый в целях упрощения анализа. Во многих задачах теории массового обслуживания представляет интерес проанализировать работу системы при постоянных условиях; тогда задача решается для стационарного потока заявок.

2. Условие отсутствия последействия — наиболее существенное для простейшего потока — означает, что заявки поступают в систему независимо друг от друга. Например, поток пассажиров, входящих на станцию метро, можно считать потоком без последействия потому, что причины, обусловившие приход отдельного пассажира именно в тот, а не другой момент, как правило, не связаны с аналогичными причинами для других пассажиров. Однако условие отсутствия последействия может быть легко нарушено за счет появления такой зависимости. Например, поток пассажиров, покидающих станцию метро, уже не может считаться потоком без последействия, так как моменты выхода пассажиров, прибывших одним и тем же поездом, зависимы между собой.

Вообще нужно заметить, что выходной поток (или поток обслуженных заявок), покидающий систему массового обслуживания, обычно имеет последействие, даже если входной поток его не имеет. Чтобы убедиться в этом, рассмотрим одноканальную систему массового обслуживания, для которой время обслуживания одной заявки вполне определено и равно $t_{об}$. Тогда в потоке обслуженных заявок минимальный интервал времени между заявками, покидающими систему, будет равен $t_{об}$. Нетрудно убедиться, что наличие такого минимального интервала неизбежно приводит к последействию. Действительно, пусть стало известно, что в какой-то момент t_1 систему покинула обслуженная заявка. Тогда можно утверждать с достоверностью, что на любом участке времени τ, лежащем в пределах $(t_1,\ t_1 + t_{об})$, обслуженной

заявки не появится; значит, будет иметь место зависимость между числами событий на неперекрывающихся участках.

Последействие, присущее выходному потоку, необходимо учитывать, если этот поток является входным для какой-либо другой системы массового обслуживания (так называемое «многофазовое обслуживание», когда одна и та же заявка последовательно переходит из системы в систему).

Отметим, между прочим, что самый простой на первый взгляд регулярный поток, в котором события отделены друг от друга равными интервалами, отнюдь не является «простейшим» в нашем смысле слова, так как в нем имеется ярко выраженное последействие: моменты появления следующих друг за другом событий связаны жесткой, функциональной зависимостью. Именно из-за наличия последействия анализ процессов, протекающих в системе массового обслуживания при регулярном потоке заявок, гораздо сложнее, чем при простейшем.

3. Условие ординарности означает, что заявки приходят поодиночке, а не парами, тройками и т. д. Например, поток атак, которому подвергается воздушная цель в зоне действия истребительной авиации, будет ординарным, если истребители атакуют цель поодиночке, и не будет ординарным, если истребители идут в атаку парами. Поток клиентов, входящих в парикмахерскую, может считаться практически ординарным, чего нельзя сказать о потоке клиентов, направляющихся в ЗАГС для регистрации брака.

Если в неординарном потоке заявки поступают только парами, только тройками и т. д., то неординарный поток легко свести к ординарному; для этого достаточно вместо потока отдельных заявок рассмотреть поток пар, троек и т. д. Сложнее будет, если каждая заявка случайным образом может оказаться двойной, тройной и т. д. Тогда уже приходится иметь дело с потоком не однородных, а разнородных событий.

В дальнейшем мы для простоты ограничимся рассмотрением ординарных потоков.

Простейший поток играет среди потоков событий вообще особую роль, до некоторой степени аналогичную роли нормального закона среди других законов распределения. Мы знаем, что при суммировании большого числа независимых случайных величин, подчиненных практически любым законам распределения, получается величина, приближенно распределенная по нормальному закону. Аналогично можно доказать, что при суммировании (взаимном наложении) большого числа ординарных, стационарных потоков с практически любым последействием получается поток, сколь угодно близкий к простейшему. Условия, которые должны для этого соблюдаться, аналогичны условиям центральной предельной теоремы, а именно — складываемые потоки должны оказывать на сумму приблизительно равномерно малое влияние.

Не доказывая этого положения и даже не формулируя математически условия, которым должны удовлетворять потоки [1]), проиллюстрируем его элементарными рассуждениями. Пусть имеется ряд независимых потоков $\Pi_1, \Pi_2, \ldots, \Pi_n$ (рис. 19.3.2). «Суммирование» потоков состоит в том, что все моменты появления событий сносятся на одну и ту же ось Ot, как показано на рис. 19.3.2.

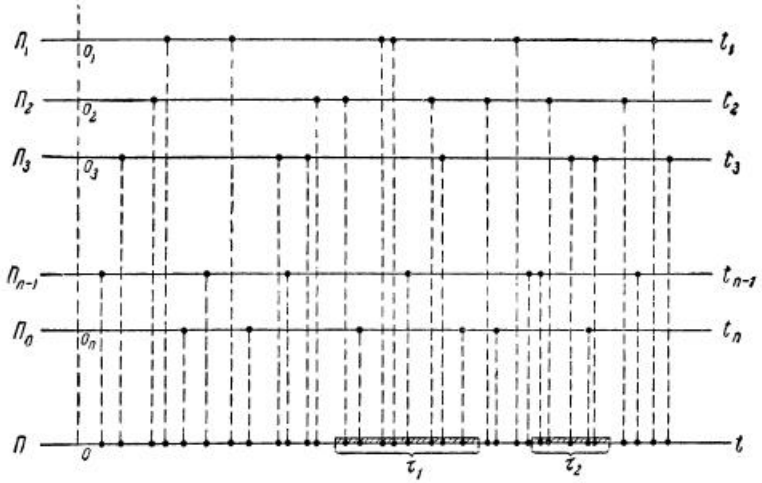


Рис. 19.3.2.

Предположим, что потоки $\Pi_1, \Pi_2, \ldots$ сравнимы по своему влиянию на суммарный поток (т. е. имеют плотности одного порядка), а число их достаточно велико. Предположим, кроме того, что эти потоки стационарны и ординарны, но каждый из них может иметь последействие, и рассмотрим суммарный поток

$$\Pi = \sum_{k=1}^{n} \Pi_k \qquad (19.3.1)$$

на оси Ot (рис. 19.3.2). Очевидно, что поток Π должен быть стационарным и ординарным, так как каждое слагаемое обладает этим свойством и они независимы. Кроме того, достаточно ясно, что при увеличении числа слагаемых последействие в суммарном потоке, даже если оно значительно в отдельных потоках, должно постепенно слабеть. Действительно, рассмотрим на оси Ot два неперекрывающихся отрезка τ_1 и τ_2 (рис. 19.3.2). Каждая из точек, попадающих в эти

[1]) См. А. Я. Хинчин, «Математические методы теории массового обслуживания», 1955.

отрезки, случайным образом может оказаться принадлежащей тому или иному потоку, и по мере увеличения n удельный вес точек, принадлежащих одному и тому же потоку (и, значит, зависимых), должен уменьшаться, а остальные точки принадлежат разным потокам и появляются на отрезках τ_1, τ_2 независимо друг от друга. Достаточно естественно ожидать, что при увеличении n суммарный поток будет терять последействие и приближаться к простейшему.



Рис. 19.3.3.

На практике оказывается обычно достаточно сложить $4 \div 5$ потоков, чтобы получить поток, с которым можно оперировать как с простейшим.

Простейший поток играет в теории массового обслуживания особенно важную роль. Во-первых, простейшие и близкие к простейшим потоки заявок часто встречаются на практике (причины этого изложены выше). Во-вторых, даже при потоке заявок, отличающемся от простейшего, часто можно получить удовлетворительные по точности результаты, заменив поток любой структуры простейшим с той же плотностью. Поэтому займемся подробнее простейшим потоком и его свойствами.

Рассмотрим на оси Ot простейший поток событий Π (рис. 19.3.3) как неограниченную последовательность случайных точек.

Выделим произвольный участок времени длиной τ. В главе 5 (п° 5.9) мы доказали, что при условиях 1, 2 и 3 (стационарность, отсутствие последействия и ординарность) число точек, попадающих на участок τ, распределено по закону Пуассона с математическим ожиданием

$$a = \lambda\tau, \tag{19.3.2}$$

где λ — плотность потока (среднее число событий, приходящееся на единицу времени).

Вероятность того, что за время τ произойдет ровно m событий, равна

$$P_m(\tau) = \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau}. \tag{19.3.3}$$

В частности, вероятность того, что участок окажется пустым (не произойдет ни одного события), будет

$$P_0(\tau) = e^{-\lambda\tau}. \tag{19.3.4}$$

Важной характеристикой потока является закон распределения длины промежутка между соседними событиями. Рассмотрим случай-

ную величину T — промежуток времени между произвольными двумя соседними событиями в простейшем потоке (рис. 19.3.3) и найдем ее функцию распределения

$$F(t) = P(T < t).$$

Перейдем к вероятности противоположного события

$$1 - F(t) = P(T \geqslant t).$$

Это есть вероятность того, что на участке времени длиной t, начинающемся в момент t_k появления одного из событий потока, не появится ни одного из последующих событий. Так как простейший поток не обладает последействием, то наличие в начале участка (в точке t_k) какого-то события никак не влияет на вероятность появления тех или других событий в дальнейшем. Поэтому вероятность $P(T \geqslant t)$ можно вычислить по формуле (19.3.4)

$$P_0(t) = e^{-\lambda t},$$

откуда

$$F(t) = 1 - e^{-\lambda t} \quad (t > 0). \quad (19.3.5)$$

Дифференцируя, найдем плотность распределения

$$f(t) = \lambda e^{-\lambda t} \quad (t > 0). \quad (19.3.6)$$

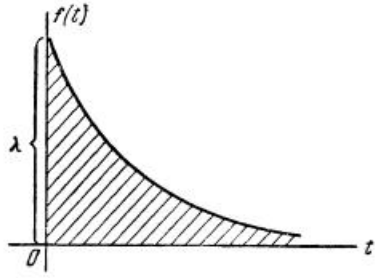


Рис. 19.3.4.

Закон распределения с плотностью (19.3.6) называется *показательным законом*, а величина τ — его *параметром*. График плотности $f(t)$ представлен на рис. 19.3.4.

Показательный закон, как мы увидим в дальнейшем, играет большую роль в теории дискретных случайных процессов с непрерывным временем. Поэтому рассмотрим его подробнее.

Найдем математическое ожидание величины T, распределенной по показательному закону:

$$m_t = M[T] = \int_0^\infty t f(t)\, dt = \lambda \int_0^\infty t e^{-\lambda t}\, dt$$

или, интегрируя по частям,

$$m_t = \frac{1}{\lambda}. \quad (19.3.7)$$

Дисперсия величины T равна

$$D_t = D(T) = \int_0^\infty t^2 f(t)\, dt - \frac{1}{\lambda^2} = \int_0^\infty \lambda t^2 e^{-\lambda t}\, dt - \frac{1}{\lambda^2},$$

откуда

$$D_t = \frac{1}{\lambda^2}, \tag{19.3.8}$$

$$\sigma_t = \frac{1}{\lambda}. \tag{19.3.9}$$

Докажем одно замечательное свойство показательного закона. Оно состоит в следующем: если промежуток времени, распределенный по показательному закону, *уже* длился некоторое время τ, то это никак не влияет на закон распределения *оставшейся* части промежутка: он будет таким же, как и закон распределения всего промежутка T.

Для доказательства рассмотрим случайный промежуток времени T с функцией распределения

$$F(t) = 1 - e^{-\lambda t} \tag{19.3.10}$$

и предположим, что этот промежуток уже продолжается некоторое время τ, т. е. произошло событие $T > \tau$. Найдем при этом предположении условный закон распределения оставшейся части промежутка $T_1 = T - \tau$; обозначим его $F^{(\tau)}(t)$

$$F^{(\tau)}(t) = P(T - \tau < t \mid T > \tau). \tag{19.3.11}$$

Докажем, что условный закон распределения $F^{(\tau)}(t)$ не зависит от τ и равен $F(t)$. Для того чтобы вычислить $F^{(\tau)}(t)$, найдем сначала вероятность произведения двух событий

$$T > \tau \quad \text{и} \quad T - \tau < t.$$

По теореме умножения вероятностей

$$P\big((T > \tau)(T - \tau < t)\big) = P(T > \tau)\, P(T - \tau < t \mid T > \tau) = \\ = P(T > \tau)\, F^{(\tau)}(t),$$

откуда

$$F^{(\tau)}(t) = \frac{P\big((T > \tau)(T - \tau < t)\big)}{P(T > \tau)}.$$

Но событие $(T > \tau)(T - \tau < t)$ равносильно событию $\tau < T < t + \tau$, вероятность которого равна

$$P(\tau < T < t + \tau) = F(t + \tau) - F(\tau).$$

С другой стороны,

$$P(T > \tau) = 1 - F(\tau),$$

следовательно,

$$F^{(\tau)}(t) = \frac{F(t + \tau) - F(\tau)}{1 - F(\tau)},$$

откуда, согласно формуле (19.3.10), получим

$$F^{(\tau)}(t) = \frac{e^{-\lambda\tau} - e^{-\lambda(t+\tau)}}{e^{-\lambda\tau}} = 1 - e^{-\lambda t} = F(t),$$

что и требовалось доказать.

Таким образом, мы доказали, что если промежуток времени T распределен по показательному закону, то любые сведения о том, сколько времени уже протекал этот промежуток, не влияют на закон распределения оставшегося времени. Можно доказать, что показательный закон — единственный, обладающий таким свойством. Это свойство показательного закона представляет собой, в сущности, другую формулировку для «отсутствия последействия», которое является основным свойством простейшего потока.

19.4 Нестационарный пуассоновский поток

Если поток событий нестационарен, то его основной характеристикой является мгновенная плотность $\lambda(t)$. *Мгновенной плотностью* потока называется предел отношения среднего числа событий, приходящегося на элементарный участок времени $(t, t+\Delta t)$, к длине этого участка, когда последняя стремится к нулю:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{m(t+\Delta t) - m(t)}{\Delta t} = m'(t), \qquad (19.4.1)$$

где $m(t)$ — математическое ожидание числа событий на участке $(0, t)$.

Рассмотрим поток однородных событий, ординарный и без последействия, но не стационарный, с переменной плотностью $\lambda(t)$. Такой поток называется *нестационарным пуассоновским потоком*. Это — первая ступень обобщения по сравнению с простейшим потоком. Легко показать методом, аналогичным примененному в п° 5.9, что для такого потока число событий, попадающих на участок длины τ, начинающийся в точке t_0, подчиняется закону Пуассона

$$P_m(\tau, t_0) = \frac{a^m}{m!} e^{-a} \qquad (m = 0, 1, 2, \ldots), \qquad (19.4.2)$$

где a — математическое ожидание числа событий на участке от t_0 до $t_0+\tau$, равное

$$a = \int_{t_0}^{t_0+\tau} \lambda(t)\, dt. \qquad (19.4.3)$$

Здесь величина a зависит не только от длины τ участка, но и от его положения на оси Ot.

Найдем для нестационарного потока закон распределения промежутка времени T между соседними событиями. Ввиду нестационарности потока этот закон будет зависеть от того, где на оси Ot

расположено первое из событий. Кроме того, он будет зависеть от вида функции $\lambda(t)$. Предположим, что первое из двух соседних

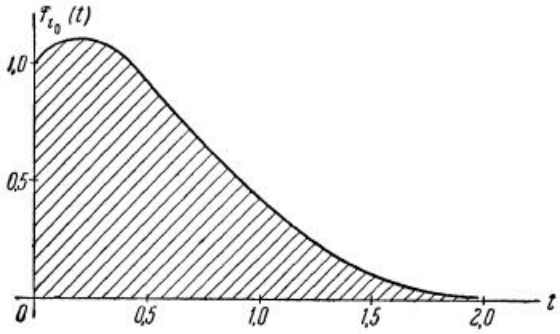


Рис. 19.4.1.

событий появилось в момент t_0, и найдем при этом условии закон распределения времени T между этим событием и последующим:

$$F_{t_0}(t) = P(T < t) = 1 - P(T \geqslant t).$$

Найдем $P(T \geqslant t)$ — вероятность того, что на участке от t_0 до $t_0 + t$ не появится ни одного события:

$$P(T \geqslant t) = e^{-a} = e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt},$$

откуда

$$F_{t_0}(t) = 1 - e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt}. \qquad (19.4.4)$$

Дифференцируя, найдем плотность распределения

$$f_{t_0}(t) = \lambda(t_0+t)\, e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt} \qquad (t > 0). \qquad (19.4.5)$$

Этот закон распределения уже не будет показательным. Вид его зависит от параметра t_0 и вида функции $\lambda(t)$. Например, при линейном изменении $\lambda(t)$

$$\lambda(t) = a + bt$$

плотность (19.4.5) имеет вид

$$f_{t_0}(t) = [a + b(t_0+t)]\, e^{-at - bt_0 t - \frac{bt^2}{2}}. \qquad (19.4.6)$$

График этого закона при $a = 0{,}4$, $b = 2$ и $t_0 = 0{,}3$ представлен на рис. 19.4.1.

Несмотря на то, что структура нестационарного пуассоновского потока несколько сложнее, чем простейшего, он очень удобен в практических применениях: главное свойство простейшего потока — отсутствие последействия — в нем сохранено. А именно, если мы зафиксируем на оси Ot произвольную точку t_0, то закон распределения $f_{t_0}(t)$ времени T, отделяющего эту точку от ближайшего по времени будущего события, не зависит от того, что происходило на участке времени, предшествующем t_0, и в самой точке t_0 (т. е. появлялись ли ранее другие события и когда именно).

19.5. Поток с ограниченным последействием (поток Пальма)

В предыдущем n° мы познакомились с естественным обобщением простейшего потока — с нестационарным пуассоновским потоком. Обобщением простейшего потока в другом направлении является поток с ограниченным последействием.

Рассмотрим ординарный поток однородных событий (рис. 19.5.1). Этот поток называется *потоком с ограниченным последействием*

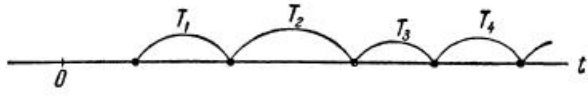


Рис. 19.5.1.

(или *потоком Пальма*), если промежутки времени между последовательными событиями T_1, T_2, ... представляют собой независимые случайные величины.

Очевидно, простейший поток является частным случаем потока Пальма: в нем расстояния T_1, T_2, ... представляют собой независимые случайные величины, распределенные по показательному закону. Что касается нестационарного пуассоновского потока, то он не является потоком Пальма. Действительно, рассмотрим два соседних промежутка T_k и T_{k+1} в нестационарном пуассоновском потоке. Как мы видели в предыдущем n°, закон распределения промежутка между событиями в нестационарном потоке зависит от того, где этот промежуток начинается, а начало промежутка T_{k+1} совпадает с концом промежутка T_k; значит, длины этих промежутков зависимы.

Рассмотрим примеры потоков Пальма.

1. Некоторая деталь технического устройства (например, радиолампа) работает непрерывно до своего отказа (выхода из строя), после чего она мгновенно заменяется новой. Срок безотказной работы детали случаен; отдельные экземпляры выходят из строя независимо

друг от друга. При этих условиях поток отказов (или поток «восстановлений») представляет собой поток Пальма. Если, к тому же, срок работы детали распределен по показательному закону, то поток Пальма превращается в простейший.

2. Группа самолетов идет в боевом порядке «колонна» (рис. 19.5.2) с одинаковой для всех самолетов скоростью V. Каждый самолет, кроме ведущего, обязан выдерживать строй, т. е. держаться на заданном расстоянии L от впереди идущего. Это расстояние, вследствие погрешностей радиодальномера, выдерживается с ошибками. Моменты

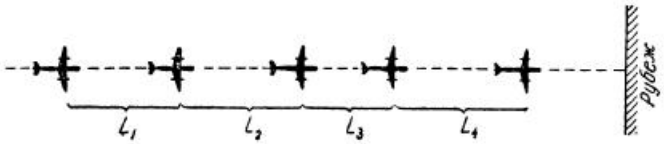


Рис. 19.5.2.

пересечения самолетами заданного рубежа образуют поток Пальма, так как случайные величины $T_1 = \frac{L_1}{V}$; $T_2 = \frac{L_2}{V}$; ... независимы. Заметим, что тот же поток не будет потоком Пальма, если каждый из самолетов стремится выдержать заданное расстояние не от соседа, а от ведущего.

Потоки Пальма часто получаются в виде выходных потоков систем массового обслуживания. Если на какую-либо систему поступает какой-то поток заявок, то он этой системой разделяется на два: поток обслуженных и поток необслуженных заявок.

Поток необслуженных заявок часто поступает на какую-либо другую систему массового обслуживания, поэтому представляет интерес изучить его свойства.

Основной в теории выходных потоков является *теорема Пальма*, которую мы сформулируем без доказательства.

Пусть на систему массового обслуживания поступает поток заявок типа Пальма, причем заявка, заставшая все каналы занятыми, получает отказ (не обслуживается). Если при этом время обслуживания имеет показательный закон распределения, то поток необслуженных заявок является также потоком типа Пальма.

В частности, если входной поток заявок будет простейшим, то поток необслуженных заявок, не будучи простейшим, будет все же иметь ограниченное последействие.

Интересным примером потоков с ограниченным последействием являются так называемые *потоки Эрланга*. Они образуются «просеиванием» простейшего потока.

Рассмотрим простейший поток (рис. 19.5.3) и выбросим из него каждую вторую точку (на рисунке выброшенные точки отмечены крестами). Оставшиеся точки образуют поток; этот поток называется *потоком Эрланга первого порядка* (Θ_1). Очевидно, этот поток



Рис. 19.5.3.

есть поток Пальма: поскольку независимы промежутки между событиями в простейшем потоке, то независимы и величины T_1, T_2, ..., получающиеся суммированием таких промежутков по два.

Поток Эрланга второго порядка получится, если сохранить в простейшем потоке каждую третью точку, а две промежуточные выбросить (рис. 19.5.4).

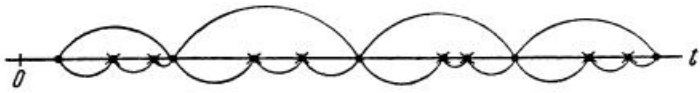


Рис. 19.5.4.

Вообще, *потоком Эрланга k-го порядка* (Θ_k) называется поток, получаемый из простейшего, если сохранить каждую $(k+1)$-ю точку, а остальные выбросить. Очевидно, простейший поток можно рассматривать как поток Эрланга нулевого порядка (Θ_0).

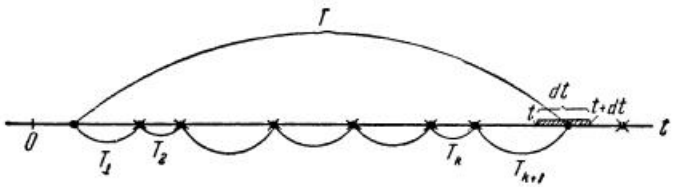


Рис. 19.5.5.

Найдем закон распределения промежутка времени T между соседними событиями в потоке Эрланга k-го порядка (Θ_k). Рассмотрим на оси Ot (рис. 19.5.5) простейший поток с интервалами $T_1, T_2, \ldots$ Величина T представляет собой сумму $k+1$ независимых случайных величин

$$T = \sum_{i=1}^{k+1} T_i, \qquad (19.5.1)$$

где T_1, T_2, ..., T_{k+1} — независимые случайные величины, подчиненные одному и тому же показательному закону

$$f(t) = \lambda e^{-\lambda t} \qquad (t > 0). \tag{19.5.2}$$

Можно было бы найти закон распределения величины T как композицию $(k+1)$ законов (19.5.2). Однако проще вывести его элементарными рассуждениями.

Обозначим $f_k(t)$ плотность распределения величины T для потока Θ_k; $f_k(t)\,dt$ есть вероятность того, что величина T примет значение между t и $t+dt$ (рис. 19.5.5). Это значит, что последняя точка промежутка T должна попасть на элементарный участок $(t, t+dt)$, а предыдущие k точек простейшего потока — на участок $(0, t)$. Вероятность первого события равна $\lambda\,dt$; вероятность второго, на основании формулы (19.3.2), будет

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Перемножая эти вероятности, получим

$$f_k(t)\,dt = \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t}\,dt,$$

откуда

$$f_k(t) = \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t} \qquad (t > 0). \tag{19.5.3}$$

Закон распределения с плотностью (19.5.3) называется *законом Эрланга k-го порядка*. Очевидно, при $k = 0$ он обращается в показательный

$$f_0(t) = \lambda e^{-\lambda t} \qquad (t > 0). \tag{19.5.4}$$

Найдем характеристики закона Эрланга $f_k(t)$: математическое ожидание m_k и дисперсию D_k. По теореме сложения математических ожиданий

$$m_k = \sum_{i=1}^{k+1} m_0 = (k+1)\,m_0,$$

где $m_0 = \frac{1}{\lambda}$ — математическое ожидание промежутка между событиями в простейшем потоке.

Отсюда

$$m_k = \frac{k+1}{\lambda}. \tag{19.5.5}$$

Аналогично по теореме сложения дисперсий

$$D_k = \frac{k+1}{\lambda^2}, \quad \sigma_k = \frac{\sqrt{k+1}}{\lambda}. \tag{19.5.6}$$

Плотность Λ_k потока Θ_k будет обратна величине m_k

$$\Lambda_k = \frac{\lambda}{k+1}. \tag{19.5.7}$$

Таким образом, при увеличении порядка потока Эрланга увеличиваются как математическое ожидание, так и дисперсия промежутка времени между событиями, а плотность потока падает.

Выясним, как будет изменяться поток Эрланга при $k \to \infty$, если его плотность будет сохраняться постоянной? Пронормируем величину T так, чтобы ее математическое ожидание (и, следовательно, плотность потока) оставалось неизменным. Для этого изменим масштаб по оси времени и вместо T рассмотрим величину

$$\tilde{T} = \frac{T}{k+1}. \tag{19.5.8}$$

Назовем такой поток *нормированным потоком Эрланга k-го порядка*. Закон распределения промежутка $\tilde{T}$ между событиями этого потока будет

$$\tilde{f}_k(t) = \frac{\Lambda_k (\Lambda_k t)^k}{k!} e^{-\Lambda_k t} \qquad (t > 0), \tag{19.5.9}$$

где $\Lambda_k = \lambda(k+1)$, или

$$\tilde{f}_k(t) = \frac{\lambda(k+1)}{k!} (\lambda(k+1)t)^k e^{-\lambda(k+1)t} \quad (t > 0). \tag{19.5.10}$$

Математическое ожидание величины $\tilde{T}$, распределенной по закону (19.5.10), не зависит от k и равно

$$\tilde{m}_k = m_0 = \frac{1}{\lambda},$$

где λ — плотность потока, совпадающая при любом k с плотностью исходного простейшего потока. Дисперсия величины $\tilde{T}$ равна

$$\tilde{D}_k = \frac{D_k}{(k+1)^2} = \frac{1}{\lambda^2 (k+1)} \tag{19.5.11}$$

и неограниченно убывает с возрастанием k.

Таким образом, мы приходим к выводу: *при неограниченном увеличении k нормированный поток Эрланга приближается к регулярному потоку с постоянными интервалами, равными* $\frac{1}{\lambda}$.

Это свойство потоков Эрланга удобно в практических применениях: оно дает возможность, задаваясь различными k, получить любую степень последействия: от полного отсутствия ($k = 0$) до жесткой функциональной связи между моментами появления событий ($k = \infty$). Таким образом, порядок потока Эрланга может служить как бы «мерой последействия», имеющегося в потоке. В целях упрощения часто

бывает удобно заменить реальный поток заявок, имеющий последействие, нормированным потоком Эрланга с примерно теми же характеристиками промежутка между заявками: математическим ожиданием и дисперсией.

Пример. В результате статистической обработки промежутков между заявками в потоке получены оценки для математического ожидания и дисперсии величины T:

$$m_t = 2\ (мин),\quad D_t = 0{,}8\ (мин^2).$$

Заменить этот поток нормированным потоком Эрланга с теми же характеристиками.

Решение. Имеем

$$\lambda = \frac{1}{m_t} = 0{,}5.$$

Из формулы (19.5.11) получим

$$k+1 \approx \frac{1}{D_t\lambda^2} = \frac{1}{0{,}2} = 5,\quad k = 4.$$

Поток можно приближенно заменить нормированным потоком Эрланга четвертого порядка.

19.6. Время обслуживания

Кроме характеристик входного потока заявок, режим работы системы зависит еще от характеристик производительности самой системы: числа каналов n и быстродействия каждого канала. Одной из важнейших величин, связанных с системой, является время обслуживания одной заявки $T_{об}$. Эта величина может быть как неслучайной, так и случайной. Очевидно, более общим является случайное время обслуживания.

Рассмотрим случайную величину $T_{об}$ и обозначим $G(t)$ ее функцию распределения:

$$G(t) = P(T_{об} < t), \tag{19.6.1}$$

а $g(t)$ — плотность распределения:

$$g(t) = G'(t). \tag{19.6.2}$$

Для практики особый интерес представляет случай, когда величина $T_{об}$ имеет показательное распределение

$$g(t) = \mu e^{-\mu t} \qquad (t > 0), \tag{19.6.3}$$

где параметр μ — величина, обратная среднему времени обслуживания одной заявки:

$$\mu = \frac{1}{m_{t_{об}}},\quad m_{t_{об}} = M[T_{об}]. \tag{19.6.4}$$

Особая роль, которую играет в теории массового обслуживания показательный закон распределения величины $T_{об}$, связана с тем

свойством этого закона, которое было доказано в п° 19.4. В применении к данному случаю оно формулируется так: если в какой-то момент t_0 происходит обслуживание заявки, то закон распределения *оставшегося* времени обслуживания не зависит от того, сколько времени обслуживание *уже продолжалось*.

На первый взгляд допущение о том, что время обслуживания распределено по показательному закону, представляется довольно искусственным. В ряде практических задач кажется естественнее предположить его либо совсем не случайным, либо распределенным по нормальному закону. Однако существуют условия, в которых время обслуживания действительно распределяется по закону, близкому к показательному.

Это, прежде всего, все задачи, в которых обслуживание сводится к ряду «попыток», каждая из которых приводит к необходимому результату с какой-то вероятностью p.

Пусть, например, «обслуживание» состоит в обстреле какой-то цели и заканчивается в момент ее поражения. Обстрел ведется независимыми выстрелами с некоторой средней скорострельностью λ выстрелов в единицу времени. Каждый выстрел поражает цель с вероятностью p. Чтобы не связывать себя необходимостью точного учета момента каждого выстрела, предположим, что они происходят в случайные моменты времени и образуют простейший поток Π с плотностью λ (рис. 19.6.1).

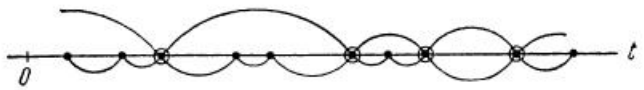


Рис. 19.6.1.

Выделим мысленно из этого потока другой — поток «успешных», или «поражающих», выстрелов (они отмечены кружками на рис. 19.6.1). Выстрел будем называть «успешным», если он приводит к поражению цели (если только цель не была поражена ранее). Нетрудно убедиться, что успешные выстрелы тоже образуют простейший поток Π^* с плотностью $\Lambda = \lambda p$ (исходный поток Π — простейший, а каждый выстрел может стать поражающим, независимо от других, с вероятностью p). Вероятность того, что цель будет поражена до момента t, будет равна

$$G(t) = P(T_{\text{об}} < t) = 1 - e^{-\Lambda t},$$

откуда плотность распределения времени «обслуживания»

$$g(t) = \Lambda e^{-\Lambda t},$$

а это есть показательный закон с параметром $\mu = \Lambda$.

Показательным законом распределения времени обстрела до поражения цели можно приближенно пользоваться и в случае, когда выстрелы не образуют простейшего потока, а отделены, например, строго определенными промежутками времени t_1, если только вероятность поражения одним выстрелом p не очень велика. Для иллюстрации приведем на одном и том же графике (рис. 19.6.2) функцию распределения момента поражающего выстрела (ступенчатая линия) для случая $p=0{,}4$, $t_1=1$ и функцию распределения показательного закона с параметром $\mu=p=0{,}4$ (плавная кривая). Как видно на рис. 19.6.2, непрерывное показательное распределение хорошо соответствует характеру нарастания функции распределения для дискретного случая. Естественно, если моменты выстрелов не будут строго определенными, соответствие с показательным законом будет еще лучше.

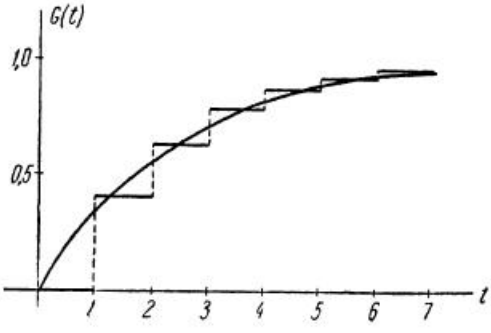


Рис. 19.6.2.

Случай стрельбы — не единственный, когда обслуживание осуществляется рядом «попыток». К такому типу часто можно отнести обслуживание по устранению неисправностей технических устройств, когда поиски неисправной детали или узла осуществляются рядом тестов или проверок. К такому же типу можно отнести задачи, где «обслуживание» заключается в обнаружении какого-либо объекта радиолокатором, если объект с какой-то вероятностью может быть обнаружен при каждом цикле обзора.

Показательным законом хорошо описываются и те случаи, когда плотность распределения времени обслуживания по тем или иным причинам убывает при возрастании аргумента t. Это бывает, когда основная масса заявок обслуживается очень быстро, а значительные задержки в обслуживании наблюдаются редко. Рассмотрим, например, окно почтового отделения, где продаются марки и конверты, а также принимаются почтовые отправления и переводы. Основная масса посетителей покупает марки или конверты и обслуживается очень быстро. Реже встречаются заявки на отправление заказных писем, они обслуживаются несколько дольше. Переводы посылаются еще реже и обслуживаются еще дольше. Наконец, в самых редких случаях представители организаций отправляют сразу большое количество писем. Гистограмма распределения времени обслуживания имеет вид, представленный на рис. 19.6.3. Так как плотность рас-

пределения убывает с возрастанием t, можно без особой погрешности выровнять распределение с помощью показательного закона, подобрав соответствующим образом его параметр μ.

Разумеется, показательный закон не является универсальным законом распределения времени обслуживания. Часто время обслуживания лучше описывается, например, законом Эрланга. Однако, к счастью, пропускная способность и другие характеристики системы массового обслуживания сравнительно мало зависят от вида закона распределения времени обслуживания, а зависят, главным образом, от его среднего значения $m_{t_{об}}$. Поэтому в теории массового обслуживания чаще всего пользуются допущением, что время обслуживания распределено по показательному закону. Эта гипотеза позволяет сильно упростить математический аппарат, применяемый для решения задач массового обслуживания, и, в ряде случаев, получить простые аналитические формулы для характеристик пропускной способности системы.

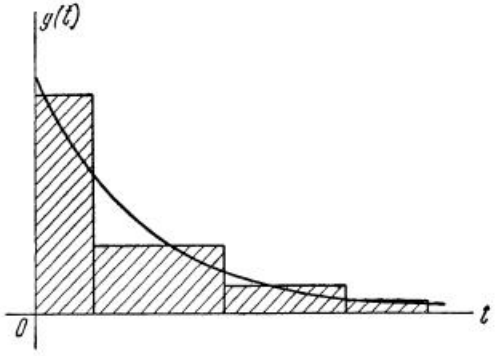


Рис. 19.6.3.

19.7. Марковский случайный процесс

Допущения о пуассоновском характере потока заявок и о показательном распределении времени обслуживания ценны тем, что позволяют применить в теории массового обслуживания аппарат так называемых марковских случайных процессов.

Процесс, протекающий в физической системе, называется *марковским* (или процессом без последействия), если для каждого момента времени *вероятность любого состояния системы в будущем зависит только от состояния системы в настоящий момент* (t_0) *и не зависит от того, каким образом система пришла в это состояние.*

Рассмотрим элементарный пример марковского случайного процесса. По оси абсцисс Ox случайным образом перемещается точка X. В момент времени $t=0$ точка X находится в начале координат ($x=0$) и остается там в течение одной секунды. Через секунду бросается монета; если выпал герб — точка X перемещается на одну единицу длины вправо, если цифра — влево. Через секунду снова бросается монета и производится такое же случайное перемещение,

и т. д. Процесс изменения положения точки (или, как говорят, «блуждания») представляет собой случайный процесс с дискретным временем ($t = 0, 1, 2, \ldots$,) и счетным множеством состояний

$$x_0 = 0;\ x_1 = 1;\ x_{-1} = -1;\ x_2 = 2;\ x_{-2} = -2;\ \ldots$$

Схема возможных переходов для этого процесса показана на рис. 19.7.1.

Покажем, что этот процесс — марковский. Действительно, представим себе, что в какой-то момент времени t_0 система находится, например, в состоянии x_1 — на одну единицу правее начала координат. Возможные положения точки через единицу времени будут x_0 и x_2 с вероятностями 1/2 и 1/2; через две единицы — x_{-1}, x_1, x_3 с вероятностями 1/4, 1/2, 1/4 и так далее. Очевидно, все эти вероятности зависят только от того, где находится точка в данный момент t_0, и совершенно не зависят от того, как она пришла туда.

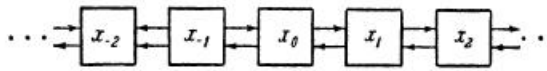


Рис. 19.7.1.

Рассмотрим другой пример. Имеется техническое устройство X, состоящее из элементов (деталей) типов a и b, обладающих разной долговечностью. Эти элементы в случайные моменты времени и независимо друг от друга могут выходить из строя. Исправная работа каждого элемента безусловно необходима для работы устройства в целом. Время безотказной работы элемента — случайная величина, распределенная по показательному закону; для элементов типа a и b параметры этого закона различны и равны соответственно λ_a и λ_b. В случае отказа устройства немедленно принимаются меры для выявления причин и обнаруженный неисправный элемент немедленно заменяется новым. Время, потребное для восстановления (ремонта) устройства, распределено по показательному закону с параметром μ_a (если вышел из строя элемент типа a) и μ_b (если вышел из строя элемент типа b).

В данном примере случайный процесс, протекающий в системе, есть марковский процесс с непрерывным временем и конечным множеством состояний:

x_0 — все элементы исправны, система работает,

x_1 — неисправен элемент типа a, система ремонтируется [1]),

x_2 — неисправен элемент типа b, система ремонтируется.

Схема возможных переходов дана на рис. 19.7.2.

[1]) Предполагается, что детали могут выходить из строя только во время работы системы; выход из строя одновременно двух или более деталей практически невозможен.

Действительно, процесс обладает марковским свойством. Пусть например, в момент t_0 система находится в состоянии x_0 (исправна). Так как время безотказной работы каждого элемента — показательное [1]), то момент отказа каждого элемента в будущем не зависит от того, сколько времени он уже работал (когда поставлен). Поэтому вероятность того, что в будущем система останется в состоянии x_0 или уйдет из него, не зависит от «предыстории» процесса. Предположим теперь, что в момент t_0 система находится в состоянии x_1 (неисправен элемент типа a). Так как время ремонта тоже показательное, вероятность окончания ремонта в любое время после t_0 не зависит от того, когда начался ремонт и когда были поставлены остальные (исправные) элементы. Таким образом, процесс является марковским.

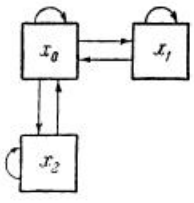


Рис. 19.7.2.

Заметим, что показательное распределение времени работы элемента и показательное распределение времени ремонта — существенные условия, без которых процесс не был бы марковским. Действительно, предположим, что время исправной работы элемента распределено не по показательному закону, а по какому-нибудь другому — например, по закону равномерной плотности на участке (t_1, t_2). Это значит, что каждый элемент с гарантией работает время t_1, а на участке от t_1 до t_2 может выйти из строя в любой момент с одинаковой плотностью вероятности. Предположим, что в какой-то момент времени t_0 элемент работает исправно. Очевидно, вероятность того, что элемент выйдет из строя на каком-то участке времени в будущем, зависит от того, насколько давно поставлен элемент, т. е. зависит от предыстории, и процесс не будет марковским.

Аналогично обстоит дело и с временем ремонта T_p; если оно не показательное и элемент в момент t_0 ремонтируется, то оставшееся время ремонта зависит от того, когда он начался; процесс снова не будет марковским.

Вообще показательное распределение играет особую роль в теории марковских случайных процессов с непрерывным временем. Легко убедиться, что в стационарном марковском процессе время, в течение которого система остается в каком-либо состоянии, распределено всегда по показательному закону (с параметром, зависящим, вообще говоря, от этого состояния). Действительно, предположим, что в момент t_0 система находится в состоянии x_k и до этого уже находилась в нем какое-то время. Согласно определению марковского процесса, вероятность любого события в будущем не зависит от предыстории; в частности, вероятность того, что система уйдет из состояния x_k

[1]) Так говорят кратко вместо «имеет показательный закон распределения».

в течение времени t, не должна зависеть от того, сколько времени система уже провела в этом состоянии. Следовательно, время пребывания системы в состоянии x_k должно быть распределено по показательному закону.

В случае, когда процесс, протекающий в физической системе со счетным множеством состояний и непрерывным временем, является марковским, можно описать этот процесс с помощью обыкновенных дифференциальных уравнений, в которых неизвестными функциями являются вероятности состояний $p_1(t)$, $p_2(t)$, ... Составление и решение таких уравнений мы продемонстрируем в следующем п° на примере простейшей системы массового обслуживания.

19.8. Система массового обслуживания с отказами. Уравнения Эрланга

Системы массового обслуживания делятся на два основных типа: а) системы с отказами, б) системы с ожиданием.

В системах с отказами заявка, поступившая в момент, когда все каналы обслуживания заняты, немедленно получает отказ, покидает систему и в дальнейшем процессе обслуживания не участвует.

В системах с ожиданием заявка, заставшая все каналы занятыми, не покидает систему, а становится в очередь и ожидает, пока не освободится какой-нибудь канал. В настоящем п° мы рассмотрим систему с отказами как наиболее простую.

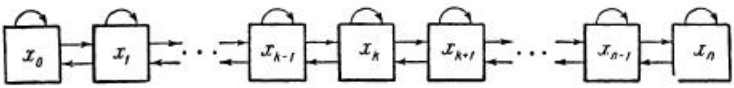


Рис. 19.8.1.

Пусть имеется n-канальная система массового обслуживания с отказами. Рассмотрим ее как физическую систему X с конечным множеством состояний:

x_0 — свободны все каналы,
x_1 — занят ровно один канал,
. . . . .
x_k — занято ровно k каналов,
. . . . .
x_n — заняты все n каналов.

Схема возможных переходов дана на рис. 19.8.1.

Поставим задачу: определить вероятности состояний системы $p_k(t)$ $(k=0, 1, \ldots, n)$ для любого момента времени t. Задачу будем решать при следующих допущениях:

1) поток заявок — простейший, с плотностью λ;

2) время обслуживания $T_{об}$ — показательное, с параметром $\mu = \frac{1}{m_{t_{об}}}$

$$g(t) = \mu e^{-\mu t} \quad (t > 0). \tag{19.8.1}$$

Заметим, что параметр μ в формуле (19.8.1) полностью аналогичен параметру λ показательного закона распределения промежутка T между соседними событиями простейшего потока:

$$f(t) = \lambda e^{-\lambda t} \quad (t > 0). \tag{19.8.2}$$

Параметр λ имеет смысл «плотности потока заявок». Аналогично, величину μ можно истолковать как «плотность потока освобождений» занятого канала. Действительно, представим себе канал, непрерывно занятый (бесперебойно снабжаемый заявками); тогда, очевидно, в этом канале будет идти простейший поток освобождений с плотностью μ.

Так как оба потока — заявок и освобождений — простейшие, процесс, протекающий в системе, будет марковским.

Рассмотрим возможные состояния системы и их вероятности

$$p_0(t),\ p_1(t),\ \ldots,\ p_n(t). \tag{19.8.3}$$

Очевидно, для любого момента времени

$$\sum_{k=0}^{n} p_k(t) = 1. \tag{19.8.4}$$

Составим дифференциальные уравнения для всех вероятностей (19.8.3), начиная с $p_0(t)$. Зафиксируем момент времени t и найдем вероятность $p_0(t+\Delta t)$ того, что в момент $t+\Delta t$ система будет находиться в состоянии x_0 (все каналы свободны). Это может произойти двумя способами (рис. 19.8.2):

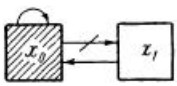


Рис. 19.8.2.

A — в момент t система находилась в состоянии x_0, а за время Δt не перешла из нее в x_1 (не пришло ни одной заявки),

B — в момент t система находилась в состоянии x_1, а за время Δt канал освободился, и система перешла в состояние x_0 [1]).

Возможностью «перескока» системы через состояние (например, из x_2 в x_0 через x_1) за малый промежуток времени можно пренебречь, как величиной высшего порядка малости по сравнению с $P(A)$ и $P(B)$ [2]).

По теореме сложения вероятностей имеем

$$p_0(t+\Delta t) \approx P(A) + P(B). \tag{19.8.5}$$

---

[1]) На рис. 19.8.2 возможные способы появления состояния x_0 в момент $t+\Delta t$ показаны стрелками, направленными в x_0; стрелка, направленная из x_0 в x_1, перечеркнута в знак того, что система не должна выйти из состояния x_0.

[2]) В дальнейшем мы все время будем пренебрегать слагаемыми высших порядков малости по сравнению с Δt. В пределе при $\Delta t \to 0$ приближенные равенства перейдут в точные.

Найдем вероятность события A по теореме умножения. Вероятность того, что в момент t система была в состоянии x_0, равна $p_0(t)$. Вероятность того, что за время Δt не придет ни одной заявки, равна $e^{-\lambda \Delta t}$. С точностью до величин высшего порядка малости

$$e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t. \tag{19.8.6}$$

Следовательно,

$$P(A) \approx p_0(t)(1 - \lambda \Delta t).$$

Найдем $P(B)$. Вероятность того, что в момент t система была в состоянии x_1, равна $p_1(t)$. Вероятность того, что за время Δt канал освободится, равна $1 - e^{-\mu \Delta t}$; с точностью до малых величин высшего порядка

$$1 - e^{-\mu \Delta t} \approx \mu \Delta t.$$

Следовательно,

$$P(B) \approx p_1(t) \mu \Delta t.$$

Отсюда

$$p_0(t + \Delta t) \approx p_0(t)(1 - \lambda \Delta t) + \mu p_1(t) \Delta t.$$

Перенося $p_0(t)$ в левую часть, деля на Δt и переходя к пределу при $\Delta t \to 0$, получим дифференциальное уравнение для $p_0(t)$:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t). \tag{19.8.7}$$

Аналогичные дифференциальные уравнения могут быть составлены и для других вероятностей состояний.

Возьмем любое $k (0 < k < n)$ и найдем вероятность $p_k(t + \Delta t)$ того, что в момент $t + \Delta t$ система будет в состоянии x_k (рис. 19.8.3). Эта вероятность вычисляется как вероятность суммы уже не двух, а трех событий (по числу стрелок, направленных в состояние x_k):

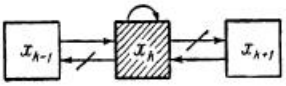


Рис. 19.8.3.

A — в момент t система была в состоянии x_k (занято k каналов), а за время Δt не перешла из него ни в x_{k+1}, ни в x_{k-1} (ни одна заявка не поступила, ни один канал не освободился);

B — в момент t система была в состоянии x_{k-1} (занято $k-1$ каналов), а за время Δt перешла в x_k (пришла одна заявка);

C — в момент t система была в состоянии x_{k+1} (занято $k+1$ каналов), а за время Δt один из каналов освободился.

Найдем $P(A)$. Вычислим сначала вероятность того, что за время Δt не придет ни одна заявка и не освободится ни один из каналов:

$$e^{-\lambda \Delta t}(e^{-\mu \Delta t})^k = e^{-(\lambda + k\mu)\Delta t}.$$

Пренебрегая малыми величинами высших порядков, имеем

$$e^{-(\lambda + k\mu)\Delta t} \approx 1 - (\lambda + k\mu)\Delta t,$$

откуда

$$P(A) \approx p_k(t)\,[1 - (\lambda + k\mu)\,\Delta t].$$

Аналогично

$$P(B) \approx p_{k-1}(t)\,\lambda\,\Delta t,$$

$$P(C) \approx p_{k+1}(t)\,(k+1)\,\mu\,\Delta t$$

и

$$p_k(t+\Delta t) \approx p_k(t)\,[1 - (\lambda + k\mu)\,\Delta t] + p_{k-1}(t)\,\lambda\,\Delta t + \\ + p_{k+1}(t)\,(k+1)\,\mu\,\Delta t.$$

Отсюда получаем дифференциальное уравнение для $p_k(t)\,(0 < k < n)$:

$$\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)\,p_k(t) + (k+1)\,\mu p_{k+1}(t).$$

Составим уравнение для последней вероятности $p_n(t)$ (рис. 19.8.4). Имеем

$$p_n(t+\Delta t) \approx p_n(t)\,(1 - n\mu\,\Delta t) + p_{n-1}(t)\,\lambda\,\Delta t,$$

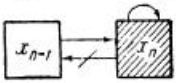


Рис. 19.8.4.

где $1 - n\mu\,\Delta t$ — вероятность того, что за время Δt не освободится ни один канал; $\lambda\,\Delta t$ — вероятность того, что за время Δt придет одна заявка. Получаем дифференциальное уравнение для $p_n(t)$:

$$\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - n\mu p_n(t).$$

Таким образом, получена система дифференциальных уравнений для вероятностей $p_1(t),\ p_2(t),\ \ldots,\ p_n(t)$:

$$\left.\begin{aligned}
&\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),\\
&\ldots\ldots\\
&\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)\,p_k(t) + (k+1)\,\mu p_{k+1}(t)\\
&\qquad\qquad (0 < k < n),\\
&\ldots\ldots\\
&\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - n\mu p_n(t).
\end{aligned}\right\} \quad (19.8.8)$$

Уравнения (19.8.8) называются *уравнениями Эрланга*. Интегрирование системы уравнений (19.8.8) при начальных условиях

$$p_0(0) = 1; \qquad p_1(0) = \ldots = p_n(0) = 0$$

(в начальный момент все каналы свободны) дает зависимость $p_k(t)$ для любого k. Вероятности $p_k(t)$ характеризуют среднюю загрузку системы и ее изменение с течением времени. В частности, $p_n(t)$ есть вероятность того, что заявка, пришедшая в момент t, застанет все каналы занятыми (получит отказ):

$$P_{\text{отк}} = p_n(t).$$

Величина $q(t) = 1 - p_n(t)$ называется *относительной пропускной способностью* системы. Для данного момента t это есть отношение среднего числа обслуженных за единицу времени заявок к среднему числу поданных.

Система линейных дифференциальных уравнений (19.8.8) сравнительно легко может быть проинтегрирована при любом конкретном числе каналов n.

Заметим, что при выводе уравнений (19.8.8) мы нигде не пользовались допущением о том, что величины λ и μ (плотности потока заявок и «потока освобождений») постоянны. Поэтому уравнения (19.8.8) остаются справедливыми и для зависящих от времени $\lambda(t)$, $\mu(t)$, лишь бы потоки событий, переводящих систему из состояния в состояние, оставались пуассоновскими (без этого процесс не будет марковским).

19.9. Установившийся режим обслуживания. Формулы Эрланга

Рассмотрим n-канальную систему массового обслуживания с отказами, на вход которой поступает простейший поток заявок с плотностью λ; время обслуживания — показательное, с параметром μ. Возникает вопрос: будет ли стационарным случайный процесс, протекающий в системе? Очевидно, что в начале, сразу после включения системы в работу, протекающий в ней процесс еще не будет стационарным: в системе массового обслуживания (как и в любой динамической системе) возникнет так называемый «переходный», нестационарный процесс. Однако, спустя некоторое время, этот переходный процесс затухнет, и система перейдет на стационарный, так называемый «установившийся» режим, вероятностные характеристики которого уже не будут зависеть от времени.

Во многих задачах практики нас интересуют именно характеристики предельного установившегося режима обслуживания.

Можно доказать, что для любой системы с отказами такой предельный режим существует, т. е. что при $t \to \infty$ все вероятности $p_0(t)$, $p_1(t)$, ..., $p_n(t)$ стремятся к постоянным пределам p_0, p_1, ..., p_n, а все их производные — к нулю.

Чтобы найти предельные вероятности p_0, p_1, ..., p_n (вероятности состояний системы в установившемся режиме), заменим в уравнениях (19.8.8) все вероятности $p_k(t)$ $(0 \leqslant k \leqslant n)$ их пределами p_k, а все

производные положим равными нулю. Получим систему уже не дифференциальных, а алгебраических уравнений

$$\left.\begin{aligned}
&-\lambda p_0 + \mu p_1 = 0,\\
&\lambda p_0 - (\lambda+\mu)\, p_1 + 2\mu p_2 = 0,\\
&\cdots\cdots\\
&\lambda p_{k-1} - (\lambda + k\mu)\, p_k + (k+1)\,\mu p_{k+1} = 0\\
&\qquad\qquad (0 < k < n),\\
&\cdots\cdots\\
&\lambda p_{n-2} - [\lambda + (n-1)\,\mu]\, p_{n-1} + n\mu p_n = 0,\\
&\lambda p_{n-1} - n\mu p_n = 0.
\end{aligned}\right\} \qquad (19.9.1)$$

К этим уравнениям необходимо добавить условие

$$\sum_{k=0}^{n} p_k = 1. \qquad (19.9.2)$$

Разрешим систему (19.9.1) относительно неизвестных p_0, p_1, ..., p_n. Из первого уравнения имеем

$$p_1 = \frac{\lambda}{\mu}\, p_0. \qquad (19.9.3)$$

Из второго, с учетом (19.9.3),

$$p_2 = \frac{1}{2\mu}[-\lambda p_0 + (\lambda+\mu)\, p_1] = \\ = \frac{1}{2\mu}\left[-\lambda p_0 + \frac{\lambda^2}{\mu}\, p_0 + \lambda p_0\right] = \frac{\lambda^2}{2\mu^2}\, p_0; \qquad (19.9.4)$$

аналогично из третьего, с учетом (19.9.3) и (19.9.4),

$$p_3 = \frac{1}{3\mu}\left(-\frac{\lambda^2}{\mu}\, p_0 + \frac{\lambda^3}{2\mu^2}\, p_0 + \frac{\lambda^2}{2\mu^2}\cdot 2\mu p_0\right) = \frac{\lambda^3}{3!\,\mu^3}\, p_0,$$

и вообще, для любого $k \leqslant n$

$$p_k = \frac{\lambda^k}{k!\,\mu^k}\, p_0. \qquad (19.9.5)$$

Введем обозначение

$$\frac{\lambda}{\mu} = \alpha \qquad (19.9.6)$$

и назовем величину α *приведенной плотностью потока заявок*. Это есть не что иное, как среднее число заявок, приходящееся на среднее время обслуживания одной заявки. Действительно,

$$\alpha = \frac{\lambda}{\mu} = \lambda m_{t_{\text{об}}},$$

где $m_{t_{об}} = M[T_{об}]$ — среднее время обслуживания одной заявки. В новых обозначениях формула (19.9.5) примет вид

$$p_k = \frac{\alpha^k}{k!} p_0. \tag{19.9.7}$$

Формула (19.9.7) выражает все вероятности p_k через p_0. Чтобы выразить их непосредственно через α и n, воспользуемся условием (19.9.2). Подставив в него (19.9.7), получим

$$\sum_{k=0}^{n} p_k = p_0 \sum_{k=0}^{n} \frac{\alpha^k}{k!} = 1,$$

откуда

$$p_0 = \frac{1}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!}}. \tag{19.9.8}$$

Подставляя (19.9.8) в (19.9.7), получим окончательно

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!}} \qquad (0 \leqslant k \leqslant n). \tag{19.9.9}$$

Формулы (19.9.9) называются *формулами Эрланга*. Они дают предельный закон распределения числа занятых каналов в зависимости от характеристик потока заявок и производительности системы обслуживания. Полагая в формуле (19.9.9) $k = n$, получим *вероятность отказа* (вероятность того, что поступившая заявка найдет все каналы занятыми):

$$P_{отк} = p_n = \frac{\frac{\alpha^n}{n!}}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!}}. \tag{19.9.10}$$

В частности, для одноканальной системы ($n = 1$)

$$P_{отк} = p_1 = \frac{\alpha}{1+\alpha}, \tag{19.9.11}$$

а относительная пропускная способность

$$q = 1 - P_{отк} = \frac{1}{1+\alpha}. \tag{19.9.12}$$

Формулы Эрланга (19.9.9) и их следствия (19.9.10) — (19.9.12) выведены нами для случая показательного распределения времени

обслуживания. Однако исследования последних лет [1]) показали, что эти формулы остаются справедливыми и при любом законе распределения времени обслуживания, лишь бы входной поток был простейшим.

Пример 1. Автоматическая телефонная станция имеет 4 линии связи. На станцию поступает простейший поток заявок с плотностью $\lambda = 3$ (вызова в минуту). Вызов, поступивший в момент, когда все линии заняты, получает отказ [2]). Средняя длительность разговора 2 минуты. Найти: а) вероятность отказа; б) среднюю долю времени, в течение которой телефонная станция вообще не загружена.

Решение. Имеем $m_{t_{об}} = 2$ (мин);

$$\mu = 0{,}5 \text{ (разг/мин)}, \quad \alpha = \frac{\lambda}{\mu} = 6.$$

а) По формуле (19.9.10) получаем

$$P_{отк} = p_4 = \frac{\frac{\alpha^4}{4!}}{1 + \frac{\alpha^1}{1!} + \frac{\alpha^2}{2!} + \frac{\alpha^3}{3!} + \frac{\alpha^4}{4!}} \approx 0{,}47.$$

б) По формуле (19.9.8)

$$p_0 = \frac{1}{1 + \frac{\alpha^1}{1!} + \frac{\alpha^2}{2!} + \frac{\alpha^3}{3!} + \frac{\alpha^4}{4!}} \approx 0{,}0087.$$

Несмотря на то, что формулы Эрланга в точности справедливы только при простейшем потоке заявок, ими можно с известным приближением пользоваться и в случае, когда поток заявок отличается от простейшего (например, является стационарным потоком с ограниченным последействием). Расчеты показывают, что замена произвольного стационарного потока с не очень большим последействием простейшим потоком той же плотности λ, как правило, мало влияет на характеристики пропускной способности системы.

Наконец, можно заметить, что формулами Эрланга можно приближенно пользоваться и в случае, когда система массового обслуживания допускает ожидание заявки в очереди, но когда срок ожидания мал по сравнению со средним временем обслуживания одной заявки.

Пример 2. Станция наведения истребителей имеет 3 канала. Каждый канал может одновременно наводить один истребитель на одну цель. Среднее время наведения истребителя на цель $m_{t_{об}} = 2$ *мин*. Поток целей — простейший, с плотностью $\lambda = 1{,}5$ (самолетов в минуту). Станцию можно считать «системой с отказами», так как цель, по которой наведение не началось в момент, когда она вошла в зону действия истребителей, вообще остается не

---

[1]) См. Б. А. Севастьянов, Эргодическая теорема для марковских процессов и ее приложение к телефонным системам с отказами. Теория вероятностей и ее применения, т. 2, вып. 1, 1957.

[2]) Предполагается, что вторичные вызовы абонентов, получивших отказ, не нарушают пуассоновского характера потока заявок.

атакованной. Найти среднюю долю целей, проходящих через зону действия не обстрелянными.

Решение. Имеем $\mu = \frac{1}{2} = 0{,}5$; $\lambda = 1{,}5$; $\alpha = \frac{\lambda}{\mu} = 3$.

По формуле (19.9.10)

$$P_{\text{отк}} = p_3 = \frac{\frac{3^3}{3!}}{1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}} \approx 0{,}346.$$

Вероятность отказа $\approx 0{,}346$; она же выражает среднюю долю необстрелянных целей.

Заметим, что в данном примере плотность потока целей выбрана такой, что при их регулярном следовании одна за другой через определенные интервалы и при точно фиксированном времени наведения $T_{\text{об}} = 2$ *мин* номинальная пропускная способность системы достаточна для того, чтобы обстрелять все без исключения цели. Снижение пропускной способности происходит из-за наличия случайных сгущений и разрежений в потоке целей, которые нельзя предвидеть заранее.

19.10. Система массового обслуживания с ожиданием

Система массового обслуживания называется *системой с ожиданием*, если заявка, заставшая все каналы занятыми, становится в очередь и ждет, пока не освободится какой-нибудь канал.

Если время ожидания заявки в очереди ничем не ограничено, то система называется «чистой системой с ожиданием». Если оно ограничено какими-то условиями, то система называется «системой смешанного типа». Это промежуточный случай между чистой системой с отказами и чистой системой с ожиданием.

Для практики наибольший интерес представляют именно системы смешанного типа.

Ограничения, наложенные на ожидание, могут быть различного типа. Часто бывает, что ограничение накладывается на время ожидания заявки в очереди; считается, что оно ограничено сверху каким-то сроком $T_{\text{ож}}$, который может быть как строго определенным, так и случайным. При этом ограничивается только срок ожидания в очереди а начатое обслуживание доводится до конца, независимо от того, сколько времени продолжалось ожидание (например, клиент в парикмахерской, сев в кресло, обычно уже не уходит до конца обслуживания). В других задачах естественнее наложить ограничение не на время ожидания в очереди, а на общее время пребывания заявки в системе (например, воздушная цель может пробыть в зоне стрельбы лишь ограниченное время и покидает ее независимо от того, кончился обстрел или нет). Наконец, можно рассмотреть и такую смешанную систему (она ближе всего к типу торговых предприятий, торгующих предметами не первой необходимости), когда заявка становится в очередь только в том случае, если длина

очереди не слишком велика. Здесь ограничение накладывается на число заявок в очереди.

В системах с ожиданием существенную роль играет так называемая «дисциплина очереди». Ожидающие заявки могут вызываться на обслуживание как в порядке очереди (раньше прибывший раньше и обслуживается), так и в случайном, неорганизованном порядке. Существуют системы массового обслуживания «с преимуществами», где некоторые заявки обслуживаются предпочтительно перед другими («генералы и полковники вне очереди»).

Каждый тип системы с ожиданием имеет свои особенности и свою математическую теорию. Многие из них описаны, например, в книге Б. В. Гнеденко «Лекции по теории массового обслуживания».

Здесь мы остановимся только на простейшем случае смешанной системы, являющемся естественным обобщением задачи Эрланга для системы с отказами. Для этого случая мы выведем дифференциальные уравнения, аналогичные уравнениям Эрланга, и формулы для вероятностей состояний в установившемся режиме, аналогичные формулам Эрланга.

Рассмотрим смешанную систему массового обслуживания X с n каналами при следующих условиях. На вход системы поступает простейший поток заявок с плотностью λ. Время обслуживания одной заявки $T_{об}$ — показательное, с параметром $\mu = \frac{1}{m_{t_{об}}}$. Заявка, заставшая все каналы занятыми, становится в очередь и ожидает обслуживания; время ожидания ограничено некоторым сроком $T_{ож}$; если до истечения этого срока заявка не будет принята к обслуживанию, то она покидает очередь и остается необслуженной. Срок ожидания $T_{ож}$ будем считать случайным и распределенным по показательному закону

$$h(t) = \nu e^{-\nu t} \qquad (t > 0),$$

где параметр ν — величина, обратная среднему сроку ожидания:

$$\nu = \frac{1}{m_{t_{ож}}}; \quad m_{t_{ож}} = M[T_{ож}].$$

Параметр ν полностью аналогичен параметрам λ и μ потока заявок и «потока освобождений». Его можно интерпретировать, как плотность «потока уходов» заявки, стоящей в очереди. Действительно, представим себе заявку, которая только и делает, что становится в очередь и ждет в ней, пока не кончится срок ожидания $T_{ож}$, после чего уходит и сразу же снова становится в очередь. Тогда «поток уходов» такой заявки из очереди будет иметь плотность ν.

Очевидно, при $\nu \to \infty$ система смешанного типа превращается в чистую систему с отказами; при $\nu \to 0$ она превращается в чистую систему с ожиданием.

Заметим, что при показательном законе распределения срока ожидания пропускная способность системы не зависит от того, обслуживаются ли заявки в порядке очереди или в случайном порядке: для каждой заявки закон распределения оставшегося времени ожидания не зависит от того, сколько времени заявка уже стояла в очереди.

Благодаря допущению о пуассоновском характере всех потоков событий, приводящих к изменениям состояний системы, процесс, протекающий в ней, будет марковским. Напишем уравнения для вероятностей состояний системы. Для этого, прежде всего, перечислим эти состояния. Будем их нумеровать не по числу занятых каналов, а по числу *связанных с системой заявок*. Заявку будем называть «связанной с системой», если она либо находится в состоянии обслуживания, либо ожидает очереди. Возможные состояния системы будут:

x_0 — ни один канал не занят (очереди нет),
x_1 — занят ровно один канал (очереди нет),
. . . . .
x_k — занято ровно k каналов (очереди нет),
. . . . .
x_n — заняты все n каналов (очереди нет),
x_{n+1} — заняты все n каналов, одна заявка стоит в очереди,
. . . . .
x_{n+s} — заняты все n каналов, s заявок стоят в очереди,
. . . . .

Число заявок s, стоящих в очереди, в наших условиях может быть сколь угодно большим. Таким образом, система X имеет бесконечное (хотя и счетное) множество состояний. Соответственно, число описывающих ее дифференциальных уравнений тоже будет бесконечным.

Очевидно, первые n дифференциальных уравнений ничем не будут отличаться от соответствующих уравнений Эрланга:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),$$

. . . . .

$$\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)\, p_k(t) + (k+1)\mu p_{k+1}(t),$$

. . . . .

$$\frac{dp_{n-1}(t)}{dt} = \lambda p_{n-2}(t) - [\lambda + (n-1)\mu]\, p_{n-1}(t) + n\mu p_n(t).$$

Отличие новых уравнений от уравнений Эрланга начнется при $k = n$. Действительно, в состояние x_n система с отказами может

перейти только из состояния x_{n-1}; что касается системы с ожиданием, то она может перейти в состояние x_n не только из x_{n-1}, но и из x_{n+1} (все каналы заняты, одна заявка стоит в очереди).

Составим дифференциальное уравнение для $p_n(t)$. Зафиксируем момент t и найдем $p_n(t+\Delta t)$ — вероятность того, что система в момент $t+\Delta t$ будет в состоянии x_n. Это может осуществиться тремя способами:

1) в момент t система уже была в состоянии x_n, а за время Δt не вышла из него (не пришла ни одна заявка и ни один из каналов не освободился);

2) в момент t система была в состоянии x_{n-1}, а за время Δt перешла в состояние x_n (пришла одна заявка);

3) в момент t система была в состоянии x_{n+1} (все каналы заняты, одна заявка стоит в очереди), а за время Δt перешла в x_n (либо освободился один канал и стоящая в очереди заявка заняла его, либо стоящая в очереди заявка ушла в связи с окончанием срока).

Имеем:

$$p_n(t+\Delta t) \approx p_n(t)(1-\lambda\Delta t - n\mu\Delta t) + \\ + p_{n-1}(t)\lambda\Delta t + p_{n+1}(t)(n\mu+\nu)\Delta t,$$

откуда

$$\frac{dp_n(t)}{dt} = -(\lambda+n\mu)p_n(t) + \lambda p_{n-1}(t) + (n\mu+\nu)p_{n+1}(t).$$

Вычислим теперь $p_{n+s}(t+\Delta t)$ при любом $s>0$ — вероятность того, что в момент $t+\Delta t$ все n каналов будут заняты и ровно s заявок будут стоять в очереди. Это событие снова может осуществиться тремя способами:

1) в момент t система уже была в состоянии x_{n+s}, а за время Δt это состояние не изменилось (значит, ни одна заявка не пришла, ни один канал не освободился и ни одна из s стоящих в очереди заявок не ушла);

2) в момент t система была в состоянии x_{n+s-1}, а за время Δt перешла в состояние x_{n+s} (т. е. пришла одна заявка);

3) в момент t система была в состоянии x_{n+s+1}, а за время Δt перешла в состояние x_{n+s} (для этого либо один из каналов должен освободиться, и тогда одна из $s+1$ стоящих в очереди заявок займет его, либо одна из стоящих в очереди заявок должна уйти в связи с окончанием срока).

Следовательно:

$$p_{n+s}(t+\Delta t) \approx p_{n+s}(t)(1-\lambda\Delta t - n\mu\Delta t - s\nu\Delta t) + \\ + p_{n+s-1}(t)\lambda\Delta t + p_{n+s+1}(t)[n\mu+(s+1)\nu]\Delta t,$$

откуда

$$\frac{dp_{n+s}(t)}{dt} = -(\lambda+n\mu+s\nu)p_{n+s}(t) + \\ + \lambda p_{n+s-1}(t) + [n\mu+(s+1)\nu]p_{n+s+1}(t).$$

Таким образом, мы получили для вероятностей состояний систему бесконечного числа дифференциальных уравнений:

$$\left.\begin{aligned}
&\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),\\
&\frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda+\mu)\, p_1(t) + 2\mu p_2(t),\\
&\cdots\cdots\\
&\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)\, p_k(t) + (k+1)\,\mu p_{k+1}(t)\\
&\qquad\qquad (1 \leqslant k \leqslant n-1),\\
&\cdots\cdots\\
&\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - (\lambda + n\mu)\, p_n(t) + (n\mu+\nu)\, p_{n+1}(t),\\
&\cdots\cdots\\
&\frac{dp_{n+s}(t)}{dt} = \lambda p_{n+s-1}(t) - (\lambda + n\mu + s\nu)\, p_{n+s}(t) +\\
&\qquad\qquad + [n\mu + (s+1)\nu]\, p_{n+s+1}(t),\\
&\cdots\cdots
\end{aligned}\right\} \qquad (19.10.1)$$

Уравнения (19.10.1) являются естественным обобщением уравнений Эрланга на случай системы смешанного типа с ограниченным временем ожидания. Параметры λ, μ, ν в этих уравнениях могут быть как постоянными, так и переменными. При интегрировании системы (19.10.1) нужно учитывать, что хотя теоретически число возможных состояний системы бесконечно, но на практике вероятности $p_{n+s}(t)$ при возрастании s становятся пренебрежимо малыми, и соответствующие уравнения могут быть отброшены.

Выведем формулы, аналогичные формулам Эрланга, для вероятностей состояний системы при установившемся режиме обслуживания (при $t \to \infty$). Из уравнений (19.10.1), полагая все p_k ($k = 0, \ldots, n, \ldots$) постоянными, а все производные — равными нулю, получим систему алгебраических уравнений:

$$\left.\begin{aligned}
&-\lambda p_0 + \mu p_1 = 0,\\
&\lambda p_0 - (\lambda+\mu)\, p_1 + 2\mu p_2 = 0,\\
&\cdots\cdots\\
&\lambda p_{k-1} - (\lambda + k\mu)\, p_k + (k+1)\,\mu p_{k+1} = 0\\
&\qquad (1 \leqslant k \leqslant n-1),\\
&\cdots\cdots\\
&\lambda p_{n-1} - (\lambda + n\mu)\, p_n + (n\mu+\nu)\, p_{n+1} = 0,\\
&\cdots\cdots\\
&\lambda p_{n+s-1} - (\lambda + n\mu + s\nu)\, p_{n+s} +\\
&\qquad + [n\mu + (s+1)\nu]\, p_{n+s+1} = 0,\\
&\cdots\cdots
\end{aligned}\right\} \qquad (19.10.2)$$

К ним нужно присоединить условие:

$$\sum_{k=0}^{\infty} p_k = 1. \tag{19.10.3}$$

Найдем решение системы (19.10.2).

Для этого применим тот же прием, которым мы пользовались в случае системы с отказами: разрешим первое уравнение относительно p_1, подставим во второе, и т. д. Для любого $k \leqslant n$, как и в случае системы с отказами, получим:

$$p_k = \frac{\lambda^k}{k!\mu^k} p_0. \tag{19.10.4}$$

Перейдем к уравнениям для $k > n$ ($k = n + s$, $s \geqslant 1$). Тем же способом получим:

$$p_{n+1} = \frac{\lambda^{n+1} p_0}{n!\mu^n (n\mu + \nu)},$$

$$p_{n+2} = \frac{\lambda^{n+2} p_0}{n!\mu^n (n\mu + \nu)(n\mu + 2\nu)},$$

и вообще при любом $s \geqslant 1$

$$p_{n+s} = \frac{\lambda^{n+s} p_0}{n!\mu^n \prod_{m=1}^{s} (n\mu + m\nu)}. \tag{19.10.5}$$

В обе формулы (19.10.4) и (19.10.5) в качестве сомножителя входит вероятность p_0. Определим ее из условия (19.10.3). Подставляя в него выражения (19.10.4) и (19.10.5) для $k \leqslant n$ и $s \geqslant 1$, получим:

$$p_0 \left\{ \sum_{k=0}^{n} \frac{\lambda^k}{k!\mu^k} + \sum_{s=1}^{\infty} \frac{\lambda^{n+s}}{n!\mu^n \prod_{m=1}^{s} (n\mu + m\nu)} \right\} = 1,$$

откуда

$$p_0 = \frac{1}{\sum_{k=0}^{n} \frac{\lambda^k}{k!\mu^k} + \sum_{s=1}^{\infty} \frac{\lambda^{n+s}}{n!\mu^n \prod_{m=1}^{s} (n\mu + m\nu)}}. \tag{19.10.6}$$

Преобразуем выражения (19.10.4), (19.10.5) и (19.10.6), вводя в них вместо плотностей λ и ν «приведенные» плотности:

$$\left.\begin{aligned} \frac{\lambda}{\mu} &= \lambda m_{t_{об}} = \alpha, \\ \frac{\nu}{\mu} &= \nu m_{t_{об}} = \beta. \end{aligned}\right\} \tag{19.10.7}$$

Параметры α и β выражают соответственно среднее число заявок и среднее число уходов заявки, стоящей в очереди, приходящиеся на среднее время обслуживания одной заявки.

В новых обозначениях формулы (19.10.4), (19.10.5) и (19.10.6) примут вид:

$$p_k = \frac{\alpha^k}{k!} p_0 \qquad (0 < k \leqslant n); \tag{19.10.8}$$

$$p_{n+s} = \frac{\frac{\alpha^{n+s}}{n!} p_0}{\prod\limits_{m=1}^{s} (n + m\beta)} \qquad (s \geqslant 1); \tag{19.10.9}$$

$$p_0 = \frac{1}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum\limits_{s=1}^{\infty} \frac{\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)}}. \tag{19.10.10}$$

Подставляя (19.10.10) в (19.10.8) и (19.10.9), получим окончательные выражения для вероятностей состояний системы:

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum\limits_{s=1}^{\infty} \frac{\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)}} \qquad (0 \leqslant k \leqslant n); \tag{19.10.11}$$

$$p_{n+s} = \frac{\frac{\alpha^n}{n!} \frac{\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)}}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum\limits_{s=1}^{\infty} \frac{\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)}} \qquad (s \geqslant 1). \tag{19.10.12}$$

Зная вероятности всех состояний системы, можно легко определить другие интересующие нас характеристики, в частности, вероятность $P_н$ того, что заявка покинет систему необслуженной. Определим ее из следующих соображений: при установившемся режиме вероятность $P_н$ того, что заявка покинет систему необслуженной, есть не что иное, как отношение среднего числа заявок, уходящих из очереди в единицу времени, к среднему числу заявок, поступающих в единицу времени. Найдем среднее число заявок, уходящих из очереди в единицу времени. Для этого сначала вычислим математическое ожидание m_s числа заявок, находящихся в очереди:

$$m_s = M[s] = \sum_{s=1}^{\infty} s p_{n+s} = \frac{\dfrac{\alpha^n}{n!}\displaystyle\sum_{s=1}^{\infty}\frac{s\alpha^s}{\displaystyle\prod_{m=1}^{s}(n+m\beta)}}{\displaystyle\sum_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum_{s=1}^{\infty}\frac{\alpha^s}{\displaystyle\prod_{m=1}^{s}(n+m\beta)}}. \qquad (19.10.13)$$

Чтобы получить $P_н$, нужно m_s умножить на среднюю «плотность уходов» одной заявки ν и разделить на среднюю плотность заявок λ, т. е. умножить на коэффициент

$$\frac{\nu}{\lambda} = \frac{\dfrac{\nu}{\mu}}{\dfrac{\lambda}{\mu}} = \frac{\beta}{\alpha}.$$

Получим:

$$P_н = \frac{\beta}{\alpha}\cdot\frac{\dfrac{\alpha^n}{n!}\displaystyle\sum_{s=1}^{\infty}\frac{s\alpha^s}{\displaystyle\prod_{m=1}^{s}(n+m\beta)}}{\displaystyle\sum_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum_{s=1}^{\infty}\frac{\alpha^s}{\displaystyle\prod_{m=1}^{s}(n+m\beta)}}. \qquad (19.10.14)$$

Относительная пропускная способность системы характеризуется вероятностью того, что заявка, попавшая в систему, будет обслужена:

$$q = 1 - P_н.$$

Очевидно, что пропускная способность системы с ожиданием, при тех же λ и μ, будет всегда выше, чем пропускная способность системы с отказами: в случае наличия ожидания необслуженными

уходят не все заявки, заставшие n каналов занятыми, а только некоторые. Пропускная способность увеличивается при увеличении среднего времени ожидания $m_{t_{ож}} = \frac{1}{\nu}$.

Непосредственное пользование формулами (19.10.11), (19.10.12) и (19.10.14) несколько затруднено тем, что в них входят бесконечные суммы. Однако члены этих сумм быстро убывают [1]).

Посмотрим, во что превратятся формулы (19.10.11) и (19.10.12) при $\beta \to \infty$ и $\beta \to 0$. Очевидно, что при $\beta \to \infty$ система с ожиданием должна превратиться в систему с отказами (заявка мгновенно уходит из очереди). Действительно, при $\beta \to \infty$ формулы (19.10.12) дадут нули, а формулы (19.10.11) превратятся в формулы Эрланга для системы с отказами.

Рассмотрим другой крайний случай: чистую систему с ожиданием ($\beta \to 0$). В такой системе заявки вообще не уходят из очереди, и поэтому $P_н = 0$: каждая заявка рано или поздно дождется обслуживания. Зато в чистой системе с ожиданием не всегда имеется предельный стационарный режим при $t \to \infty$. Можно доказать, что такой режим существует только при $\alpha < n$, т. е. когда среднее число заявок, приходящееся на время обслуживания одной заявки, не выходит за пределы возможностей n-канальной системы. Если же $\alpha \geqslant n$, число заявок, стоящих в очереди, будет с течением времени неограниченно возрастать.

Предположим, что $\alpha < n$, и найдем предельные вероятности p_k ($0 \leqslant k \leqslant n$) для чистой системы с ожиданием. Для этого положим в формулах (19.9.10), (19.9.11) и (19.9.12) $\beta \to 0$. Получим:

$$p_0 = \frac{1}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum\limits_{s=1}^{\infty} \frac{\alpha^s}{n^s}},$$

или, суммируя прогрессию (что возможно только при $\alpha < n$),

$$p_0 = \frac{1}{\sum\limits_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!\,(n-\alpha)}}. \qquad (19.10.15)$$

---

[1]) Для грубой оценки ошибки, происходящей от отбрасывания всех членов сумм, начиная с r-го, можно воспользоваться формулами:

$$\sum_{s=r}^{\infty} \frac{\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)} < \frac{\left(\frac{\alpha}{\beta}\right)^r}{r!} e^{\frac{\alpha}{\beta}}, \quad \sum_{s=r}^{\infty} \frac{s\alpha^s}{\prod\limits_{m=1}^{s} (n + m\beta)} < \frac{\left(\frac{\alpha}{\beta}\right)^r}{(r-1)!} e^{\frac{\alpha}{\beta}}.$$

Отсюда, пользуясь формулами (19.10.8) и (19.10.9), найдем

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=1}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}} \qquad (0 \leqslant k \leqslant n), \qquad (19.10.16)$$

и аналогично для $k = n + s$ $(s \geqslant 0)$

$$p_{n+s} = \frac{\frac{\alpha^{n+s}}{n!\, n^s}}{\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}}. \qquad (19.10.17)$$

Среднее число заявок, находящихся в очереди, определяется из формулы (19.10.13) при $\beta \to 0$:

$$m_s = \frac{\frac{\alpha^{n+1}}{n \cdot n!\left(1-\frac{\alpha}{n}\right)^2}}{\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}}. \qquad (19.10.18)$$

Пример 1. На вход трехканальной системы с неограниченным временем ожидания поступает простейший поток заявок с плотностью $\lambda = 4$ (заявки в час). Среднее время обслуживания одной заявки $m_{t_{об}} = 30$ *мин.* Определить, существует ли установившийся режим обслуживания; если да, то найти вероятности p_0, p_1, p_2, p_3, вероятность наличия очереди и среднюю длину очереди m_s.

Решение. Имеем $\mu = \frac{1}{m_{t_{об}}} = 2$; $\alpha = \frac{\lambda}{\mu} = 2$. Так как $\alpha < n$, установившийся режим существует. По формуле (19.10.16) находим

$$p_0 = \frac{1}{9} \approx 0{,}111; \qquad p_1 = \frac{2}{9} \approx 0{,}222; \qquad p_2 = \frac{2}{9} \approx 0{,}222; \qquad p_3 = \frac{8}{54} \approx 0{,}148.$$

Вероятность наличия очереди:

$$P_{оч} = 1 - (p_0 + p_1 + p_2 + p_3) = 0{,}297.$$

Средняя длина очереди по формуле (19.10.18) будет

$$m_s \approx 0{,}89 \text{ (заявки)}.$$

19.11. Система смешанного типа с ограничением по длине очереди

В предыдущем п° мы рассмотрели систему массового обслуживания с ограничением по времени пребывания в очереди. Здесь мы рассмотрим систему смешанного типа с другим видом ограничения ожидания — по числу заявок, стоящих в очереди. Предположим, что

заявка, заставшая все каналы занятыми, становится в очередь, только если в ней находится менее m заявок; если же число заявок в очереди равно m (больше m оно быть не может), то последняя прибывшая заявка в очередь не становится и покидает систему необслуженной. Остальные допущения — о простейшем потоке заявок и о показательном распределении времени обслуживания — оставим прежними.

Итак, имеется n-канальная система с ожиданием, в которой количество заявок, стоящих в очереди, ограничено числом m. Составим дифференциальные уравнения для вероятностей состояний системы. Заметим, что в данном случае число состояний системы будет конечно, так как общее число заявок, связанных с системой, не может превышать $n+m$ (n обслуживаемых и m стоящих в очереди). Перечислим состояния системы:

x_0 — все каналы свободны, очереди нет,
x_1 — занят один канал, » » ,
. . . . .
x_k — занято k каналов, » » ,
. . . . .
x_{n-1} — занято $n-1$ каналов, » » ,
x_n — заняты все n каналов, » » ,
x_{n+1} — заняты все n каналов, одна заявка стоит в очереди,
. . . . .
x_{n+m} — заняты все n каналов, m заявок стоит в очереди.

Очевидно, первые n уравнений для вероятностей $p_0(t), \ldots p_{n-1}(t)$ будут совпадать с уравнениями Эрланга (19.8.8). Выведем остальные уравнения. Имеем

$$p_n(t+\Delta t) = p_{n-1}(t)\lambda\Delta t + p_n(t)(1-\lambda\Delta t - n\mu\Delta t) + p_{n+1}(t)n\mu\Delta t,$$

откуда

$$\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + n\mu p_{n+1}(t).$$

Далее выведем уравнение для $p_{n+s}(t)$ $(1 \leqslant s < m)$

$$p_{n+s}(t+\Delta t) = p_{n+s-1}(t)\lambda\Delta t + p_{n+s}(t)(1-\lambda\Delta t - n\mu\Delta t) + p_{n+s+1}(t)n\mu\Delta t,$$

откуда

$$\frac{dp_{n+s}(t)}{dt} = \lambda p_{n+s-1}(t) - (\lambda + n\mu)p_{n+s}(t) + n\mu p_{n+s+1}(t).$$

Последнее уравнение будет

$$\frac{dp_{n+m}(t)}{dt} = \lambda p_{n+m-1}(t) - n\mu p_{n+m}(t).$$

Таким образом, получена система $(n+m+1)$ дифференциальных уравнений:

$$\left.\begin{aligned}
&\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),\\
&\cdots\cdots\\
&\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)\, p_k(t) + \\
&\qquad + (k+1)\,\mu p_{k+1}(t) \qquad (0 < k \leqslant n-1),\\
&\cdots\cdots\\
&\frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - (\lambda + n\mu)\, p_n(t) + n\mu p_{n+1}(t),\\
&\cdots\cdots\\
&\frac{dp_{n+s}(t)}{dt} = \lambda p_{n+s-1}(t) - (\lambda + n\mu)\, p_{n+s}(t) + \\
&\qquad + n\mu p_{n+s+1}(t) \qquad (1 \leqslant s < m),\\
&\cdots\cdots\\
&\frac{dp_{n+m}(t)}{dt} = \lambda p_{n+m-1} - n\mu p_{n+m}(t).
\end{aligned}\right\} \qquad (19.11.1)$$

Рассмотрим предельный случай при $t \to \infty$. Приравнивая все производные нулю, а все вероятности считая постоянными, получим систему алгебраических уравнений

$$\left.\begin{aligned}
&-\lambda p_0 + \mu p_1 = 0,\\
&\cdots\cdots\\
&\lambda p_{k-1} - (\lambda + k\mu)\, p_k + (k+1)\,\mu p_{k+1} = 0\\
&\qquad\qquad (0 < k \leqslant n-1),\\
&\cdots\cdots\\
&\lambda p_{n-1} - (\lambda + n\mu)\, p_n + n\mu p_{n+1} = 0,\\
&\cdots\cdots\\
&\lambda p_{n+s-1} - (\lambda + n\mu)\, p_{n+s} + n\mu p_{n+s+1} = 0\\
&\qquad\qquad (1 \leqslant s < m),\\
&\cdots\cdots\\
&\lambda p_{n+m-1} - n\mu p_{n+m} = 0
\end{aligned}\right\} \qquad (19.11.2)$$

и добавочное условие:

$$\sum_{k=0}^{n+m} p_k = 1. \qquad (19.11.3)$$

Уравнения (19.11.2) могут быть решены так же, как мы решили аналогичные алгебраические уравнения в предыдущих п°п°. Не оста-

навливаясь на этом решении, приведем только окончательные формулы:

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{m} \left(\frac{\alpha}{n}\right)^s} \qquad (0 \leqslant k \leqslant n), \qquad (19.11.4)$$

$$p_{n+s} = \frac{\frac{\alpha^n}{n!}\left(\frac{\alpha}{n}\right)^s}{\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{m} \left(\frac{\alpha}{n}\right)^s} \qquad (1 \leqslant s \leqslant m). \qquad (19.11.5)$$

Вероятность того, что заявка покинет систему необслуженной, равна вероятности p_{n+m} того, что в очереди уже стоят m заявок.

Нетрудно заметить, что формулы (19.11.4) и (19.11.5) получаются из формул (19.10.11), (19.10.12), если положить в них $\beta = 0$ и ограничить суммирование по s верхней границей m.

Пример. На станцию текущего ремонта автомашин поступает простейший поток заявок с плотностью $\lambda = 0{,}5$ (машины в час). Имеется одно помещение для ремонта. Во дворе станции могут одновременно находиться, ожидая очереди, не более трех машин. Среднее время ремонта одной машины $m_{t_{об}} = \frac{1}{\mu} = 2$ (*часа*). Определить: а) пропускную способность системы; б) среднее время простоя станции; в) определить, насколько изменятся эти характеристики, если оборудовать второе помещение для ремонта.

Решение. Имеем: $\lambda = 0{,}5$; $\mu = 0{,}5$; $\alpha = 1$; $m = 3$.

а) По формуле (19.11.5), полагая $n = 1$, находим вероятность того, что пришедшая заявка покинет систему необслуженной:

$$P_н = p_{1+3} = \frac{1}{1+1+3} = 0{,}20.$$

Относительная пропускная способность системы $q = 1 - P_н = 0{,}80$. Абсолютная пропускная способность: $Q = \lambda q = 0{,}4$ (машины в час).

б) Средняя доля времени, которое система будет простаивать, найдем по формуле (19.11.4): $p_0 = \frac{1}{5} = 0{,}20$.

в) Полагая $n = 2$, найдем:

$$P_н = p_{2+3} = \frac{\frac{1}{16}}{1+1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}} = \frac{1}{47} \approx 0{,}021,$$

$q = 1 - P_н \approx 0{,}979$ (т. е. удовлетворяться будет около 98% всех заявок).

$$Q = \lambda q \approx 0{,}49 \text{ (машины в час).}$$

Относительное время простоя: $p_0 = \frac{16}{47} \approx 0{,}34$,

т. е. оборудование будет простаивать полностью около 34% всего времени.

ПРИЛОЖЕНИЕ

Таблица 1

Значения нормальной функции распределения:

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
—0,00	0,5000	40	—0,30	0,3821	38	—0,60	0,2743	34
—0,01	4960	40	—0,31	3783	38	—0,61	2709	33
—0,02	4920	40	—0,32	3745	38	—0,62	2676	33
—0,03	4880	40	—0,33	3707	38	—0,63	2643	32
—0,04	4840	39	—0,34	3669	37	—0,64	2611	33
—0,05	4801	40	—0,35	3632	38	—0,65	2578	32
—0,06	4761	40	—0,36	3594	37	—0,66	2546	32
—0,07	4721	40	—0,37	3557	37	—0,67	2514	31
—0,08	4681	40	—0,38	3520	37	—0,68	2483	32
—0,09	4641	39	—0,39	3483	37	—0,69	2451	31
—0,10	0,4602	40	—0,40	0,3446	37	—0,70	0,2420	31
—0,11	4562	40	—0,41	3409	37	—0,71	2389	31
—0,12	4522	39	—0,42	3372	36	—0,72	2358	31
—0,13	4483	40	—0,43	3336	36	—0,73	2327	30
—0,14	4443	39	—0,44	3300	36	—0,74	2297	31
—0,15	4404	40	—0,45	3264	36	—0,75	2266	30
—0,16	4364	39	—0,46	3228	36	—0,76	2236	30
—0,17	4325	39	—0,47	3192	36	—0,77	2206	29
—0,18	4286	39	—0,48	3156	35	—0,78	2177	29
—0,19	4247	40	—0,49	3121	36	—0,79	2148	29
—0,20	0,4207	39	—0,50	0,3085	35	—0,80	0,2119	29
—0,21	4168	39	—0,51	3050	35	—0,81	2090	29
—0,22	4129	39	—0,52	3015	34	—0,82	2061	28
—0,23	4090	38	—0,53	2981	35	—0,83	2033	28
—0,24	4052	39	—0,54	2946	34	—0,84	2005	28
—0,25	4013	39	—0,55	2912	35	—0,85	1977	28
—0,26	3974	38	—0,56	2877	34	—0,86	1949	27
—0,27	3936	39	—0,57	2843	33	—0,87	1922	28
—0,28	3897	38	—0,58	2810	34	—0,88	1894	27
—0,29	3859	38	—0,59	2776	33	—0,89	1867	26

Продолжение табл. 1

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
—0,90	0,1841	27	—1,30	0,0968	17	—1,70	0,0446	10
—0,91	1814	26	—1,31	0951	17	—1,71	0436	9
—0,92	1788	26	—1,32	0934	16	—1,72	0427	9
—0,93	1762	26	—1,33	0918	17	—1,73	0418	9
—0,94	1736	25	—1,34	0901	16	—1,74	0409	8
—0,95	1711	26	—1,35	0885	16	—1,75	0401	9
—0,96	1685	25	—1,36	0869	16	—1,76	0392	8
—0,97	1660	25	—1,37	0853	15	—1,77	0384	9
—0,98	1635	24	—1,38	0838	15	—1,78	0375	8
—0,99	1611	24	—1,39	0823	15	—1,79	0367	8
—1,00	0,1587	24	—1,40	0,0808	15	—1,80	0,0359	8
—1,01	1563	24	—1,41	0793	15	—1,81	0351	7
—1,02	1539	24	—1,42	0778	14	—1,82	0344	8
—1,03	1515	23	—1,43	0764	15	—1,83	0336	7
—1,04	1492	23	—1,44	0749	14	—1,84	0329	7
—1,05	1469	23	—1,45	0735	14	—1,85	0322	8
—1,06	1446	23	—1,46	0721	13	—1,86	0314	7
—1,07	1423	22	—1,47	0708	14	—1,87	0307	6
—1,08	1401	22	—1,48	0694	13	—1,88	0301	7
—1,09	1379	22	—1,49	0681	13	—1,89	0294	6
—1,10	0,1357	22	—1,50	0,0668	13	—1,90	0,0288	7
—1,11	1335	21	—1,51	0655	12	—1,91	0281	7
—1,12	1314	22	—1,52	0643	13	—1,92	0274	6
—1,13	1292	21	—1,53	0630	12	—1,93	0268	6
—1,14	1271	20	—1,54	0618	12	—1,94	0262	6
—1,15	1251	21	—1,55	0606	12	—1,95	0256	6
—1,16	1230	20	—1,56	0594	12	—1,96	0250	6
—1,17	1210	20	—1,57	0582	11	—1,97	0244	5
—1,18	1190	20	—1,58	0571	12	—1,98	0239	6
—1,19	1170	19	—1,59	0559	11	—1,99	0233	5
—1,20	0,1151	20	—1,60	0,0548	11	—2,00	0,0228	49
—1,21	1131	19	—1,61	0537	11	—2,10	0179	40
—1,22	1112	19	—1,62	0526	10	—2,20	0139	32
—1,23	1093	18	—1,63	0516	11	—2,30	0107	25
—1,24	1075	19	—1,64	0505	10	—2,40	0082	20
—1,25	1056	18	—1,65	0495	10	—2,50	0062	15
—1,26	1038	18	—1,66	0485	10	—2,60	0047	12
—1,27	1020	17	—1,67	0475	10	—2,70	0035	9
—1,28	1003	18	—1,68	0465	10	—2,80	0026	7
—1,29	0985	17	—1,69	0455	9	—2,90	0019	5

Продолжение табл. 1

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
—3,00	0,0014	4	0,30	0,6179	38	0,70	0,7580	31
—3,10	0010	3	0,31	6217	38	0,71	7611	31
—3,20	0007	2	0,32	6255	38	0,72	7642	31
—3,30	0005	2	0,33	6293	38	0,73	7673	30
—3,40	0003	1	0,34	6331	37	0,74	7703	31
—3,50	0002	0	0,35	6368	38	0,75	7734	30
—3,60	0002	1	0,36	6406	37	0,76	7764	30
—3,70	0001	0	0,37	6443	37	0,77	7794	29
—3,80	0001	1	0,38	6480	37	0,78	7823	29
—3,90	0000		0,39	6517	37	0,79	7852	29
0,00	0,5000	40	0,40	0,6554	37	0,80	0,7881	29
0,01	5040	40	0,41	6591	37	0,81	7910	29
0,02	5080	40	0,42	6628	36	0,82	7939	28
0,03	5120	40	0,43	6664	36	0,83	7967	28
0,04	5160	39	0,44	6700	36	0,84	7995	28
0,05	5199	40	0,45	6736	36	0,85	8023	28
0,06	5239	40	0,46	6772	36	0,86	8051	27
0,07	5279	40	0,47	6808	36	0,87	8078	28
0,08	5319	40	0,48	6844	35	0,88	8106	27
0,09	5359	39	0,49	6879	36	0,89	8133	26
0,10	0,5398	40	0,50	0,6915	35	0,90	0,8159	27
0,11	5438	40	0,51	6950	35	0,91	8186	26
0,12	5478	39	0,52	6985	34	0,92	8212	26
0,13	5517	40	0,53	7019	35	0,93	8238	26
0,14	5557	39	0,54	7054	34	0,94	8264	25
0,15	5596	40	0,55	7088	35	0,95	8289	26
0,16	5636	39	0,56	7123	34	0,96	8315	25
0,17	5675	39	0,57	7157	33	0,97	8340	25
0,18	5714	39	0,58	7190	34	0,98	8365	24
0,19	5753	40	0,59	7224	33	0,99	8389	24
0,20	0,5793	39	0,60	0,7257	34	1,00	0,8413	24
0,21	5832	39	0,61	7291	33	1,01	8437	24
0,22	5871	39	0,62	7324	33	1,02	8461	24
0,23	5910	38	0,63	7357	32	1,03	8485	23
0,24	5948	39	0,64	7389	33	1,04	8508	23
0,25	5987	39	0,65	7422	32	1,05	8531	23
0,26	6026	38	0,66	7454	32	1,06	8554	23
0,27	6064	39	0,67	7486	31	1,07	8577	22
0,28	6103	38	0,68	7517	32	1,08	8599	22
0,29	6141	38	0,69	7549	31	1,09	8621	22

Продолжение табл 1

x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ	x	$\Phi^*(x)$	Δ
1,10	0,8643	22	1,50	0,9332	13	1,90	0,9713	6
1,11	8665	21	1,51	9345	12	1,91	9719	7
1,12	8686	22	1,52	9357	13	1,92	9726	6
1,13	8708	21	1,53	9370	12	1,93	9732	6
1,14	8729	20	1,54	9382	12	1,94	9738	6
1,15	8749	21	1,55	9394	12	1,95	9744	6
1,16	8770	20	1,56	9406	12	1,96	9750	6
1,17	8790	20	1,57	9418	11	1,97	9756	5
1,18	8810	20	1,58	9429	12	1,98	9761	6
1,19	8830	19	1,59	9441	11	1,99	9767	5
1,20	0,8849	20	1,60	0,9452	11	2,00	0,9772	49
1,21	8869	19	1,61	9463	11	2,10	9821	40
1,22	8888	19	1,62	9474	10	2,20	9861	32
1,23	8907	18	1,63	9484	11	2,30	9893	25
1,24	8925	19	1,64	9495	10	2,40	9918	20
1,25	8944	18	1,65	9505	10	2,50	9938	15
1,26	8962	18	1,66	9515	10	2,60	9953	12
1,27	8980	17	1,67	9525	10	2,70	9965	9
1,28	8997	18	1,68	9535	10	2,80	9974	7
1,29	9015	17	1,69	9545	9	2,90	9981	5
1,30	0,9032	17	1,70	0,9554	10	3,00	0,9986	4
1,31	9049	17	1,71	9564	9	3,10	9990	3
1,32	9066	16	1,72	9573	9	3,20	9993	2
1,33	9082	17	1,73	9582	9	3,30	9995	2
1,34	9099	16	1,74	9591	8	3,40	9997	1
1,35	9115	16	1,75	9599	9	3,50	9998	0
1,36	9131	16	1,76	9608	8	3,60	9998	1
1,37	9147	15	1,77	9616	9	3,70	9999	0
1,38	9162	15	1,78	9625	8	3,80	9999	1
1,39	9177	15	1,79	9633	8	3,90	1,0000	
1,40	0,9192	15	1,80	0,9641	8			
1,41	9207	15	1,81	9649	7			
1,42	9222	14	1,82	9656	8			
1,43	9236	15	1,83	9664	7			
1,44	9251	14	1,84	9671	7			
1,45	9265	14	1,85	9678	8			
1,46	9279	13	1,86	9686	7			
1,47	9292	14	1,87	9693	6			
1,48	9306	13	1,88	9699	7			
1,49	9319	13	1,89	9706	7			

Таблица 2

Значения функции e^{-x}

x	e^{-x}	Δ	x	e^{-x}	Δ	x	e^{-x}	Δ	x	e^{-x}	Δ
0,00	1,000	10	0,40	0,670	7	0,80	0,449	4	3,00	0,050	5
0,01	0,990	10	0,41	0,664	7	0,81	0,445	5	3,10	0,045	4
02	980	10	42	657	7	82	440	4	3,20	41	4
03	970	9	43	650	6	83	436	4	3,30	37	4
04	961	10	44	644	6	84	432	5	3,40	33	3
05	951	9	45	638	7	85	427	4	3,50	30	3
06	942	10	46	631	6	86	423	4	3,60	27	2
07	932	9	47	625	6	87	419	4	3,70	25	3
08	923	9	48	619	6	88	415	4	3,80	22	2
09	914	9	49	613	7	89	411	4	3,90	20	2
0,10	0,905	9	0,50	0,606	6	0,90	0,407	4	4,00	0,0183	17
11	896	9	51	600	5	91	403	4	4,10	166	16
12	887	9	52	595	6	92	399	4	4,20	150	14
13	878	9	53	589	6	93	395	4	4,30	136	13
14	869	8	54	583	6	94	391	4	4,40	123	12
15	861	9	55	577	6	95	387	4	4,50	111	10
16	852	8	56	571	6	96	383	4	4,60	101	10
17	844	9	57	565	5	97	379	4	4,70	0,0091	9
18	835	8	58	560	6	98	375	3	4,80	82	8
19	827	8	59	554	5	99	372	4	4,90	74	7
0,20	0,819	8	0,60	0,549	6	1,00	0,368	35	5,00	0,0067	6
21	811	8	61	543	5	1,10	333	31	5,10	61	6
22	803	8	62	538	5	1,20	302	29	5,20	55	5
23	795	8	63	533	6	1,30	273	26	5,30	50	5
24	787	8	64	527	5	1,40	247	24	5,40	45	4
25	779	8	65	522	5	1,50	223	21	5,50	41	4
26	771	8	66	517	5	1,60	202	19	5,60	37	4
27	763	7	67	512	5	1,70	183	18	5,70	33	3
28	756	8	68	507	5	1,80	165	15	5,80	30	3
29	748	7	69	502	5	1,90	150	15	5,90	27	2
0,30	0,741	8	0,70	0,497	5	2,00	0,135	13	6,00	0,0025	3
31	733	7	71	492	5	2,10	122	11	6,10	22	2
32	726	7	72	487	5	2,20	111	11	6,20	20	2
33	719	7	73	482	5	2,30	100	9	6,30	18	1
34	712	7	74	477	5	2,40	0,091	9	6,40	17	2
35	705	7	75	472	4	2,50	82	8	6,50	15	1
36	698	7	76	468	5	2,60	74	7	6,60	14	2
37	691	7	77	463	5	2,70	67	6	6,70	12	1
38	684	7	78	458	4	2,80	61	6	6,80	11	1
39	677	7	79	454	5	2,90	55	5	6,90	10	1
0,40	0,670		0,80	0,449		3,00	0,050		7,00	0,0009	

Таблица 3

Значения функции $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9	x
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973	0,0
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918	0,1
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825	0,2
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697	0,3
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538	0,4
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352	0,5
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144	0,6
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920	0,7
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685	0,8
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444	0,9
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203	1,0
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965	1,1
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736	1,2
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518	1,3
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315	1,4
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127	1,5
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957	1,6
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804	1,7
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669	1,8
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551	1,9
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449	2,0
2,1	0440	0431	0422	0413	0404	0396	0388	0379	0371	0363	2,1
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290	2,2
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229	2,3
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180	2,4
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139	2,5
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107	2,6
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081	2,7
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061	2,8
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046	2,9
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034	3,0
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025	3,1
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018	3,2
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013	3,3
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009	3,4
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006	3,5
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004	3,6
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003	3,7
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002	3,8
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001	3,9
x	0	1	2	3	4	5	6	7	8	9	x

Таблица 4

Значения χ^2 в зависимости от r и p

r \ p	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,000	0,001	0,004	0,016	0,064	0,148	0,455	1,074	1,642	2,71	3,84	5,41	6,64	10,83
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41	3,22	4,60	5,99	7,82	9,21	13,82
3	0,115	0,185	0,352	0,584	1,005	1,424	2,37	3,66	4,64	6,25	7,82	9,84	11,34	16,27
4	0,297	0,429	0,711	1,064	1,649	2,20	3,36	4,88	5,99	7,78	9,49	11,67	13,28	18,46
5	0,554	0,752	1,145	1,610	2,34	3,00	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,5
6	0,872	1,134	1,635	2,20	3,07	3,83	5,35	7,23	8,56	10,64	12,59	15,03	16,81	22,5
7	1,239	1,564	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,3
8	1,646	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51	18,17	20,1	26,1
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,68	21,7	27,9
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,78	13,44	15,99	18,31	21,2	23,2	29,6
11	3,05	3,61	4,58	5,58	6,99	8,15	10,34	12,90	14,63	17,28	19,68	22,6	24,7	31,3
12	3,57	4,18	5,23	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,0	24,1	26,2	32,9
13	4,11	4,76	5,89	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,4	25,5	27,7	34,6
14	4,66	5,37	6,57	7,79	9,47	10,82	13,34	16,22	18,15	21,1	23,7	26,9	29,1	36,1
15	5,23	5,98	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,3	25,0	28,3	30,6	37,7
16	5,81	6,61	7,96	9,31	11,15	12,62	15,34	18,42	20,5	23,5	26,3	29,6	32,0	39,3
17	6,41	7,26	8,67	10,08	12,00	13,53	16,34	19,51	21,6	24,8	27,6	31,0	33,4	40,8
18	7,02	7,91	9,39	10,86	12,86	14,44	17,34	20,6	22,8	26,0	28,9	32,3	34,8	42,3
19	7,63	8,57	10,11	11,65	13,72	15,35	18,34	21,7	23,9	27,2	30,1	33,7	36,2	43,8
20	8,26	9,24	10,85	12,44	14,58	16,27	19,34	22,8	25,0	28,4	31,4	35,0	37,6	45,3
21	8,90	9,92	11,59	13,24	15,44	17,18	20,3	23,9	26,2	29,6	32,7	36,3	38,9	46,8
22	9,54	10,60	12,34	14,04	16,31	18,10	21,3	24,9	27,3	30,8	33,9	37,7	40,3	48,3
23	10,20	11,29	13,09	14,85	17,19	19,02	22,3	26,0	28,4	32,0	35,2	39,0	41,6	49,7
24	10,86	11,99	13,85	15,66	18,06	19,94	23,3	27,1	29,6	33,2	36,4	40,3	43,0	51,2
25	11,52	12,70	14,61	16,47	18,94	20,9	24,3	28,2	30,7	34,4	37,7	41,7	44,3	52,6
26	12,20	13,41	15,38	17,29	19,82	21,8	25,3	29,2	31,8	35,6	38,9	42,9	45,6	54,1
27	12,88	14,12	16,15	18,11	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0	55,5
28	13,56	14,85	16,93	18,94	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3	56,9
29	14,26	15,57	17,71	19,77	22,5	24,6	28,3	32,5	35,1	39,1	42,6	46,7	49,6	58,3
30	14,95	16,31	18,49	20,6	23,4	25,5	29,3	33,5	36,2	40,3	43,8	48,0	50,9	59,7

Значения t_β, удовлетворяющие равенству

$n-1$ \ β	0,1	0,2	0,3	0,4	0,5	0,6	0,7
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963
2	142	289	445	617	0,816	1,061	1,336
3	137	277	424	584	765	0,978	1,250
4	134	271	414	569	741	941	1,190
5	132	267	408	559	727	920	1,156
6	131	265	404	553	718	906	1,134
7	130	263	402	549	711	896	1,119
8	130	262	399	546	706	889	1,108
9	129	261	398	543	703	883	1,100
10	129	260	397	542	700	879	1,093
11	129	260	396	540	697	876	1,088
12	128	259	395	539	695	873	1,083
13	128	259	394	538	694	870	1,079
14	128	258	393	537	692	868	1,076
15	128	258	393	536	691	866	1,074
16	128	258	392	535	690	865	1,071
17	128	257	392	534	689	863	1,069
18	127	257	392	534	688	862	1,067
19	127	257	391	533	688	861	1,066
20	127	257	391	533	687	860	1,064
21	127	257	391	532	686	859	1,063
22	127	256	390	532	686	858	1,061
23	127	256	390	532	685	858	1,060
24	127	256	390	531	685	857	1,059
25	127	256	390	531	684	856	1,058
26	127	256	390	531	684	856	1,058
27	127	256	389	531	684	855	1,057
28	127	256	389	530	683	855	1,056
29	127	256	389	530	683	854	1,055
30	127	256	389	530	683	854	1,055
40	126	255	388	529	681	851	1,050
60	126	254	387	527	679	848	1,046
120	126	254	386	526	677	845	1,041
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036

| $n-1$ \ β | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,7 |

Таблица 5

$2\int_0^{t_\beta} S_{n-1}(t)\,dt = \beta$, в зависимости от β и $n-1$

0,8	0,9	0,95	0,98	0,99	0,999	β / $n-1$
3,08	6,31	12,71	31,8	63,7	636,6	1
1,886	2,92	4,30	6,96	9,92	31,6	2
1,638	2,35	3,18	4,54	5,84	12,94	3
1,533	2,13	2,77	3,75	4,60	8,61	4
1,476	2,02	2,57	3,36	4,03	6,86	5
1,440	1,943	2,45	3,14	3,71	5,96	6
1,415	1,895	2,36	3,00	3,50	5,40	7
1,397	1,860	2,31	2,90	3,36	5,04	8
1,383	1,833	2,26	2,82	3,25	4,78	9
1,372	1,812	2,23	2,76	3,17	4,59	10
1,363	1,796	2,20	2,72	3,11	4,49	11
1,356	1,782	2,18	2,68	3,06	4,32	12
1,350	1,771	2,16	2,65	3,01	4,22	13
1,345	1,761	2,14	2,62	2,98	4,14	14
1,341	1,753	2,13	2,60	2,95	4,07	15
1,337	1,746	2,12	2,58	2,92	4,02	16
1,333	1,740	2,11	2,57	2,90	3,96	17
1,330	1,734	2,10	2,55	2,88	3,92	18
1,328	1,729	2,09	2,54	2,86	3,88	19
1,325	1,725	2,09	2,53	2,84	3,85	20
1,323	1,721	2,08	2,52	2,83	3,82	21
1,321	1,717	2,07	2,51	2,82	3,79	22
1,319	1,714	2,07	2,50	2,81	3,77	23
1,318	1,711	2,06	2,49	2,80	3,74	24
1,316	1,708	2,06	2,48	2,79	3,72	25
1,315	1,706	2,06	2,48	2,78	3,71	26
1,314	1,703	2,05	2,47	2,77	3,69	27
1,313	1,701	2,05	2,47	2,76	3,67	28
1,311	1,699	2,04	2,46	2,76	3,66	29
1,310	1,697	2,04	2,46	2,75	3,65	30
1,303	1,684	2,02	2,42	2,70	3,55	40
1,296	1,671	2,00	2,39	2,66	3,46	60
1,289	1,658	1,980	2,36	2,62	3,37	120
1,282	1,645	1,960	2,33	2,58	3,29	∞
0,8	0,9	0,95	0,98	0,99	0,999	$n-1$ / β

Таблица 6

Таблица двоичных логарифмов целых чисел от 1 до 100

x	$\log x$	x	$\log x$	x	$\log x$
1	0,00000	36	5,16993	71	6,14975
2	1,00000	37	5,20945	72	6,16992
3	1,58496	38	5,24793	73	6,18982
4	2,00000	39	5,28540	74	6,20945
5	2,32193	40	5,32193	75	6,22882
6	2,58496	41	5,35755		
7	2,80735	42	5,39232	76	6,24793
8	3,00000	43	5,42626	77	6,26679
9	3,16993	44	5,45943	78	6,28540
10	3,32193	45	5,49185	79	6,30378
11	3,45943	46	5,52356	80	6,32193
12	3,58496	47	5,55459		
13	3,70044	48	5,58496	81	6,33985
14	3,80735	49	5,61471	82	6,35755
15	3,90689	50	5,64386	83	6,37504
16	4,00000	51	5,67242	84	6,39232
17	4,08746	52	5,70044	85	6,40939
18	4,16993	53	5,72792		
19	4,24793	54	5,75489	86	6,42626
20	4,32193	55	5,78136	87	6,44294
21	4,39232	56	5,80735	88	6,45943
22	4,45943	57	5,83289	89	6,47573
23	4,52356	58	5,85798	90	6,49185
24	4,58496	59	5,88264		
25	4,64386	60	5,90689	91	6,50779
26	4,70044	61	5,93074	92	6,52356
27	4,75489	62	5,95420	93	6,53916
28	4,80735	63	5,97728	94	6,55459
29	4,85798	64	6,00000	95	6,56986
30	4,90689	65	6,02237		
31	4,95420	66	6,04439	96	6,58496
32	5,00000	67	6,06609	97	6,59991
33	5,04439	68	6,08746	98	6,61471
34	5,08746	69	6,10852	99	6,62936
35	5,12928	70	6,12928	100	6,64386

Таблица значений функции $\eta(p) = -p \log_2 p$ Таблица 7

p	$\eta(p)$	Δ	p	$\eta(p)$	Δ
0	0	664	0,50	0,5000	—46
0,01	0,0664	464	0,51	4954	—48
0,02	1128	390	0,52	4906	—52
0,03	1518	340	0,53	4854	—54
0,04	1858	303	0,54	4800	—56
0,05	2161	274	0,55	4744	—59
0,06	2435	251	0,56	4685	—62
0,07	2686	229	0,57	4623	—65
0,08	2915	211	0,58	4558	—67
0,09	3126	196	0,59	4491	—69
0,10	3322	181	0,60	4422	—72
0,11	3503	168	0,61	4350	—74
0,12	3671	155	0,62	4276	—77
0,13	3826	145	0,63	4199	—78
0,14	3971	134	0,64	4121	—81
0,15	4105	125	0,65	4040	—83
0,16	4230	116	0,66	3957	—86
0,17	4346	107	0,67	3871	—87
0,18	4453	99	0,68	3784	—90
0,19	4552	92	0,69	3694	—92
0,20	4644	84	0,70	3602	—94
0,21	4728	78	0,71	3508	—96
0,22	4806	71	0,72	3412	—98
0,23	4877	67	0,73	3314	—99
0,24	4941	59	0,74	3215	—102
0,25	5000	53	0,75	3113	—104
0,26	5053	47	0,76	3009	—106
0,27	5100	42	0,77	2903	—107
0,28	5142	37	0,78	2796	—109
0,29	5179	32	0,79	2687	—112
0,30	5211	27	0,80	2575	—113
0,31	5238	22	0,81	2462	—114
0,32	5260	18	0,82	2348	—117
0,33	5278	14	0,83	2231	—119
0,34	5292	9	0,84	2112	—120
0,35	5301	5	0,85	1992	—121
0,36	5306	1	0,86	1871	—123
0,37	5307	—2	0,87	1748	—125
0,38	5305	—7	0,88	1623	—127
0,39	5298	—10	0,89	1496	—128
0,40	5288	—14	0,90	1368	—130
0,41	5274	—18	0,91	1238	—131
0,42	5256	—20	0,92	0,1107	—133
0,43	5236	—24	0,93	0,0974	—135
0,44	5210	—26	0,94	839	—136
0,45	5184	—29	0,95	703	—138
0,46	5153	—33	0,96	565	—139
0,47	5120	—37	0,97	426	—140
0,48	5083	—40	0,98	286	—142
0,49	5043	—43	0,99	144	—144
0,50	0,5000		1,00	0	

Таблица 8

Значения $P_m = \frac{a^m}{m!} e^{-a}$ (распределение Пуассона)

m	$a=0{,}1$	$a=0{,}2$	$a=0{,}3$	$a=0{,}4$	$a=0{,}5$	$a=0{,}6$	$a=0{,}7$	$a=0{,}8$	$a=0{,}9$	
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	
1	0,0905	0,1638	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659	
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647	
3	0,0002	0,0019	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494	
4		0,0001	0,0002	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111	
5				0,0001	0,0002	0,0004	0,0007	0,0012	0,0020	
6							0,0001	0,0002	0,0003	

m	$a=1$	$a=2$	$a=3$	$a=4$	$a=5$	$a=6$	$a=7$	$a=8$	$a=9$	$a=10$
0	0,3679	0,1353	0,0498	0,0183	0,0067	0,0025	0,0009	0,0003	0,0001	0,0000
1	0,3679	0,2707	0,1494	0,0733	0,0337	0,0149	0,0064	0,0027	0,0011	0,0005
2	0,1839	0,2707	0,2240	0,1465	0,0842	0,0446	0,0223	0,0107	0,0050	0,0023
3	0,0613	0,1804	0,2240	0,1954	0,1404	0,0892	0,0521	0,0286	0,0150	0,0076
4	0,0153	0,0902	0,1680	0,1954	0,1755	0,1339	0,0912	0,0572	0,0337	0,0189
5	0,0031	0,0361	0,1008	0,1563	0,1755	0,1606	0,1277	0,0916	0,0607	0,0378
6	0,0005	0,0120	0,0504	0,1042	0,1462	0,1606	0,1490	0,1221	0,0911	0,0631
7	0,0001	0,0037	0,0216	0,0595	0,1044	0,1377	0,1490	0,1396	0,1171	0,0901
8		0,0009	0,0081	0,0298	0,0653	0,1033	0,1304	0,1396	0,1318	0,1126
9		0,0002	0,0027	0,0132	0,0363	0,0688	0,1014	0,1241	0,1318	0,1251
10			0,0008	0,0053	0,0181	0,0413	0,0710	0,0993	0,1186	0,1251
11			0,0002	0,0019	0,0082	0,0225	0,0452	0,0722	0,0970	0,1137
12			0,0001	0,0006	0,0034	0,0126	0,0263	0,0481	0,0728	0,0948
13				0,0002	0,0013	0,0052	0,0142	0,0296	0,0504	0,0729
14				0,0001	0,0005	0,0022	0,0071	0,0169	0,0324	0,0521
15					0,0002	0,0009	0,0033	0,0090	0,0194	0,0347
16						0,0003	0,0014	0,0045	0,0109	0,0217
17						0,0001	0,0006	0,0021	0,0058	0,0128
18							0,0002	0,0009	0,0029	0,0071
19							0,0001	0,0004	0,0014	0,0037
20								0,0002	0,0006	0,0019
21								0,0001	0,0003	0,0009
22									0,0001	0,0004
23										0,0002
24										0,0001

ЛИТЕРАТУРА

1. Б. В. Гнеденко, Курс теории вероятностей, Физматгиз, 1961.
2. В. С. Пугачев, Теория случайных функций и ее применение к задачам автоматического управления, Физматгиз, 1962.
3. В. С. Пугачев, Статистическая теория систем автоматического управления (лекции), вып. 1 и 2, изд. ВВИА, 1961.
4. Г. Крамер, Математические методы статистики, ИЛ, 1948.
5. И. В. Дунин-Барковский и Н. В. Смирнов, Теория вероятностей и математическая статистика в технике, Гостехиздат, 1955.
6. А. М. Яглом и И. М. Яглом, Вероятность и информация, Гостехиздат, 1960.
7. С. Н. Бернштейн, Теория вероятностей, Гостехиздат, 1946.
8. А. А. Свешников, Прикладные методы теории случайных функций, Судпромиздат, 1961.
9. А. М. Яглом, Введение в теорию стационарных случайных функций, Успехи матем. наук, т. VII, вып. 5, 1952.
10. Б. В. Гнеденко, Лекции по теории массового обслуживания, изд. КВИРТУ, 1960.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Асимметрия 97

Вероятностная зависимость 173
Вероятность доверительная 318
— события 24
— — условная 46
— статистическая 29
Выборка 142
Выравнивание статистических рядов 143

Генеральная совокупность 142
Гистограмма 137
Главные оси рассеивания 194
— средние квадратические отклонения 194
Группа событий полная 25

Дисперсия 96, 403
— случайной функции 378
Доверительные границы 318
Доверительный интервал 318

Зависимость случайных величин линейная 179
— стохастическая 173
Закон больших чисел 88, 286, 290
— Гаусса 116
— Пуассона 106
— равномерной плотности 103
— распределения 68, 182, 376
— — дифференциальный 80
— — интегральный 73
— — предельный 287
— — случайной функции 375
— — суммы 271
— — условный 169
— — устойчивый 278
— редких явлений 113
— Симпсона 274
— треугольника 274
— Эрланга 532

Интеграл вероятностей 122
— Фурье 434
Информация 481
— полная взаимная 486
— — (средняя) 486
— частная 489

Код 502
— двоичный 503
— Шеннона — Фено 505
Кодирование 502
Комплекс 159
Композиция законов распределения 272
— нормальных законов 275
— — — на плоскости 280
Корреляция отрицательная 180
— положительная 180
Коэффициент корреляции 178
Кривая распределения 80
Кривые Пирсона 145
— равной плотности 165
Критерии согласия 149
Критерий Колмогорова 156
— χ^2 151
Круговое рассеивание 196

Линеаризация функций 252
Линия регрессии 192

Математическая статистика 131
Математическое ожидание 86, 218, 402
— — случайной функции 377
— — условное 192
Матрица корреляционная 185
— — нормированная 186
Медиана 92
Мера точности 120
Метод моментов 145
— наименьших квадратов 354
Многоугольник распределения 69
Мода 90

Момент 92
— абсолютный 99
— корреляционный 176
— начальный 92, 93, 175
— нормального распределения 120
— связи 176
— системы случайных величин 175
— статистический 140
— центральный 94, 175, 176

Непосредственный подсчет вероятностей 25
Неравенство Чебышева 287
Нормальный закон 116, 280, 297
— — в пространстве 205
— — — теории стрельбы 297
— —, каноническая форма 194
— — на плоскости 188
— —, параметры 117

Обработка стрельб 347
Объединение двух систем 475
Опыты независимые 59
Отклонение вероятное 127
— — главное 195
— срединное 127
— среднее арифметическое 99
— — квадратическое 96, 97
Оценка 313
— несмещенная 314
— состоятельная 313
— эффективная 314
Оценки для характеристик системы 339

Плотность вероятности 80
— дисперсии средняя 431
— потока мгновенная 527
— — заявок приведенная 545
— распределения 80, 164
— — системы двух величин 163
— — — нескольких величин 182
— спектральная 432
— — нормированная 434
Поверхность распределения 165
Поток Пальма 529
— событий 520
— — без последействия 520, 521
— — нестационарный пуассоновский 527
— — однородных 520
— — ординарный 520, 522
— — простейший 520
— — регулярный 520
— — с ограниченным последействием 529
— — стационарный 520, 521
— Эрланга 530—533
Правило трех сигма 125, 290
Преобразование Фурье 300, 434
— — обратное 300
Принцип практической уверенности 35
Произведение событий 39
Производительность источника информации 508
Пропускная способность канала 508, 511
— — системы абсолютная 516
— — — относительная 516, 544
Простая статистическая совокупность 133, 134
Процесс без последействия 537
— марковский 537
— стохастический 20

Разложение корреляционной функции каноническое 412
— случайной функции 409
— — — каноническое 412
— спектральное 430, 431
— — в комплексной форме 438
Распределение антимодальное 91
— биномиальное 61
— полимодальное 91
— Пуассона 106
— Релея 200
— Стьюдента 325
Рассеивание 96, 176
Реализация случайной функции 371
Ряд распределения 69
— статистический 136
— — простой 134

Свойство суперпозиции 409
Сетка рассеивания 202
Сечение случайной функции 374
Система случайных величин 159
Системы величин некоррелированные 187
— массового обслуживания 515
Скошенность 97
Случаи 26
— благоприятные 26
Случайная величина 32, 67
— — дискретная 32
— — комплексная 402
— — непрерывная 32
— — прерывная 32
— — смешанная 77
— — характеристическая 33
— — центрированная 93
— точка 159
— функция 370, 371
— — комплексная 404

Случайная функция нормированная 383
— —, разложение 409
— — центрированная 382
— — элементарная 407
Случайное явление 11
Случайные векторы некоррелированные 187
— величины зависимые 172
— — независимые 172, 183
— — некоррелированные 178
— — несвязанные 178
— функции, сложение 399
Случайный вектор 160
— процесс нестационарный 419
— — стационарный 419
Событие 23
— достоверное 24
— невозможное 24
События зависимые 46
— независимые 45, 48
— несовместные 26
— практически достоверные 34
— — невозможные 34
— противоположные 42
— равновозможные 26
Спектр дисперсий 431
— стационарной случайной функции 431
Стандарт 96
Статистическое среднее 140
Степень неопределенности системы 469, 470
Сумма событий 38
Сходимость по вероятности 31

Теорема Бернулли 31, 36, 286, 295
— гипотез 56
— Лапласа 308
— Маркова 294
— о повторении опытов 61, 63
— Пальма 530
— Пуассона 296
— сложения вероятностей 40
— — дисперсий 224
— — корреляционных матриц 230
— — — моментов 229
— — математических ожиданий 221
— — энтропий 476
— умножения вероятностей 46
— — законов распределения 170
— — математических ожиданий 226
— центральная предельная 287, 297, 302, 303
— Чебышева обобщенная 293
Теорема Шеннона вторая 513
— — первая 508
Теоремы предельные 287
Теория информации 468

Уравнения Эрланга 543
Условие Линдеберга 306
Устойчивость средних 88, 286
— частот 28, 29

Формула Бейеса 56
— полной вероятности 54
— — — интегральная 217, 218
Формулы Эрланга 546
Функция автокорреляционная 379
— корреляционная 379
— — взаимная 399
— — нормированная 381, 422
— Лапласа 128
— — приведенная 129
— производящая 62
— распределения 73
— — нормальная 123
— — системы двух величин 160
— — — нескольких величин 182
— — статистическая 134
— связи корреляционная 399
— характеристическая 299, 301

Характеристика частотная 449
Характеристики выборочные 142
— комплексной случайной величины 402
— положения 85
— рассеивания 96
— системы величин 184
— случайной функции 377
— числовые 85
— — функций случайных величин 210

Центр рассеивания 118

Частота события 24, 28

Эксцесс 98
Элемент вероятности 81, 165
Эллипс равной плотности 194
— рассеивания 194
— единичный 195
— — полный 195
Энтропия 471, 481
— приведенная 495
— условная 477
— частная условная 496
Эргодическое свойство 458