В. К. ЗАХАРОВ
Б. А. СЕВАСТЬЯНОВ
В. П. ЧИСТЯКОВ

Теория вероятностей

В. К. ЗАХАРОВ
Б. А. СЕВАСТЬЯНОВ
В. П. ЧИСТЯКОВ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебника для студентов инженерно-технических специальностей вузов

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1983

95

22.171
З-38
УДК 519.21

Захаров В. К., Севастьянов Б. А., Чистяков В. П.
Теория вероятностей.— М.: Наука, Главная редакция физико-математической литературы, 1983. — 160 с.

Учебник соответствует минимальному варианту программы по теории вероятностей, допускаемому общей программой по высшей математике для инженерно-технических специальностей технических вузов.

Книга содержит материал по следующим темам: математические модели случайных явлений; независимость событий; последовательности испытаний; случайные величины; числовые характеристики случайных величин; закон больших чисел, предельные теоремы; обработка результатов измерений; статистическая проверка гипотез.

В книге имеются задачи в количестве, достаточном для проведения упражнений, предусмотренных программой; приведены ответы.

Для студентов инженерно-технических специальностей вузов.

З $\frac{1702060000-088}{053(02)-83}$ 75-83

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ

Действующая в настоящее время программа по теории вероятностей для технических вузов допускает довольно большие колебания объема включаемого в начальный полугодовой курс материала. Предлагаемый учебник соответствует минимальному варианту программы. В связи с этим в учебник не включены разделы «Цепи Маркова», «Характеристические функции», «Элементы теории случайных процессов». Центральная предельная теорема приводится без доказательства; определение цепи Маркова дается лишь как частный случай общей последовательности зависимых испытаний.

Особое внимание авторов было обращено на четкость формулировок основных вероятностных моделей, обсуждение условий их применения, а также на методы вычисления вероятностей и математических ожиданий. Изложение этих вопросов проведено достаточно подробно с рассмотрением примеров.

Нередко при кратком изложении теории вероятностей понятие случайной величины фактически не используется, а рассматриваются лишь задачи, в которых можно обойтись только законом распределения. Это в значительной мере обедняет курс и сводит теорию вероятностей к разрозненным задачам математического анализа. В предлагаемом курсе дается математическое определение случайной величины сначала для более простых математических моделей, а затем — общее определение. Много внимания уделяется случайным величинам, представимым в виде функции от более простых величин, в частности, в виде сумм индикаторов.

Используя текст, набранный петитом, а также учебники [2], [4], [7]—[10], [12], [14], нетрудно, в случае необходимости, расширить предлагаемый данным учебником материал.

Изучение курса теории вероятностей обязательно должно сопровождаться решением задач. В конце параграфов приводятся задачи для самостоятельного решения; в конце учебника приведены ответы. Общее количество задач рассчитано на 7—8 занятий по темам, предусмотренным программой. Приведенные задачи можно дополнить, если воспользоваться задачниками [3], [11].

Изложенный в учебнике материал достаточен (в случае его усвоения) для решения многих задач, часто встречающихся в практике. Эффективность изучения теории вероятностей может быть значительно повышена регулярным и квалифицированным использованием теории вероятностей в специальных курсах, читаемых профилирующими кафедрами, а также совместной научной работой студентов и преподавателей математических и профилирующих кафедр.

§ 1. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СЛУЧАЙНЫХ ЯВЛЕНИЙ

1.1. Математические модели. Широко известна исключительно большая роль, которую играет математика при изучении закономерностей реального мира. Схема на рис. 1 показывает место математики при

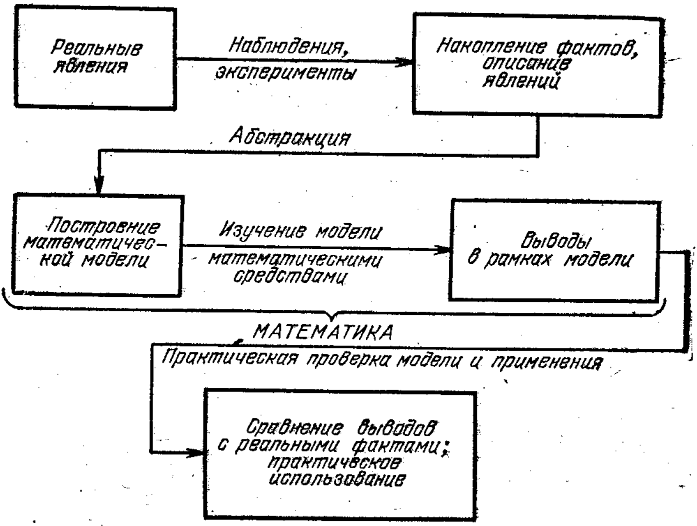


Рис. 1.

исследовании реальных явлений. Очень важно при этом подчеркнуть, что математика имеет дело не с самими реальными явлениями, а лишь с их математическими моделями. Связь математики с явлениями окружающего нас мира осуществляется в двух направлениях. Сначала, абстрагируясь от многих второстепенных фактов, мы строим математическую модель, отражающую основные закономерности изучаемого явления. В этой модели используются математические понятия, формулируются аксиомы, которым удовлетворяют эти понятия. Далее, в рамках построенной математической модели из аксиом выводится ряд следствий, сформулированных в виде теорем и лемм. И, наконец, полученные в модели новые математические факты интерпретируются в первоначальных понятиях реальных явлений. Это позволяет проверить пригодность математической модели и использовать в практике математические расчеты, произведенные в модели. Практические выводы будут достаточно надежными, если построенная модель отражает существенные стороны изучаемого явления. Примером хорошо и успешно работающей математической модели является механика, построенная на системе аксиом Ньютона.

1.2. Случайные явления. Приведенный выше пример модели относится к закономерным явлениям, т. е. к таким явлениям, исход которых однозначно определяется некоторыми условиями. Но мы знаем, что для широкого круга явлений наблюдается неоднозначность исхода при повторении опыта с сохранением основных условий его проведения. К неоднозначности исхода приводит влияние большого числа причин, каждая из которых не может заметно изменить результат опыта. События, связанные с такими явлениями, называют случайными. Случайными событиями являются, например, выпадение «герба» при подбрасывании монеты, результаты измерений, длительность телефонного разговора и т. д. Мы будем изучать *массовые* случайные события, т. е. события, возникающие в результате осуществления условий, которые можно (хотя бы в принципе) воспроизводить много раз.

Из повседневного опыта известно, что одни случайные события наступают довольно часто, другие менее часто или совсем редко. Эти характеристики событий слишком неопределенны. Более объективной экспериментальной характеристикой случайного события (обозначим его, например, A) является частота $h_n(A)$, равная отношению числа опытов n_A, в которых событие A наступило, к общему числу опытов n, т. е. $h_n(A) = n_A/n$. Экспериментально установлено, что для многих событий частота при увеличении n становится почти постоянной. Это свойство называют *статистической устойчивостью частот* случайного события. Массовые случайные события, как правило, обладают свойством устойчивости частот. Таким образом, с каждым событием A можно связать некоторое число $\mathsf{P}(A)$, с которым сближается частота, и считать это число вероятностью события A. Такое описание вероятности довольно неопределенно. Чтобы придать этому описанию точный смысл, мы должны построить математическую модель случайного явления. Для этого прежде всего надо дать математическое описание опыта, для исходов которого мы желаем находить вероятности.

Разделом математики, в котором изучаются математические модели случайных явлений, является теория вероятностей.

1.3. Пространство элементарных событий. Начнем с рассмотрения простых примеров.

Пример 1. *Подбрасывание игральной кости один раз.* В результате этого опыта могут наступать различные события: «выпало 2 очка», «выпало 6 очков», «число выпавших очков четно» и т. д. Мы будем различать элементарные (неразложимые) события и составные события (или просто события). Например, сказать, что число выпавших очков четно, все равно, что сказать, что опыт привел к выпадению двух, четырех или шести очков. Обозначим ω_k событие, состоящее в выпадении k очков. Элементарными событиями в данном опыте являются события ω_1, ω_2, ω_3, ω_4, ω_5, ω_6. Составные события, или просто события, могут быть описаны как подмножества множества элементарных событий: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. Так, событие $A = \{$выпало четное число очков$\}$ через элементарные события выражается следующим образом: $A = \{\omega_2, \omega_4, \omega_6\}$.

Пример 2. *Трехкратное подбрасывание монеты.* При каждом подбрасывании монеты будем записывать результат опыта, обозначая выпадение герба символом «1» и решетки — символом «0». Так, запись 010 будет обозначать результат опыта, в котором при первом и третьем подбрасываниях выпала решетка, а при втором — герб. Множество элементарных событий Ω состоит из 8 элементарных событий:

$$\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\}. \qquad (1)$$

Для обозначения элементарных событий можно использовать любые удобные символы. Например, вместо $000, 001, \ldots, 111$ можно использовать числа $0, 1, \ldots, 7$, двоичная запись которых соответствует (1); или более формально обозначить элементарные события: $\omega_1, \omega_2, \ldots, \omega_8$. Событие $A =$ {при первом подбрасывании выпал герб} является в данном опыте составным: $A = \{100, 101, 110, 111\}$. Любое подмножество множества Ω можно интерпретировать как некоторое событие реального опыта. Например, событие $B = \{110, 101, 011\}$ состоит в том, что выпало ровно два герба.

Пример 3. *Стрельба по плоской мишени.* Пусть по плоской мишени производится один выстрел. Элементарными событиями в этом опыте являются точки мишени. Введем в плоскости мишени прямоугольную систему координат uOv. Тогда множество элементарных событий $\omega = (u, v)$ можно записать в виде

$$\Omega = \{\omega\} = \{(u, v): -\infty < u < +\infty, \ -\infty < v < +\infty\}.$$

Событие $A =$ {попадание произошло в круг единичного радиуса} является подмножеством Ω:

$$A = \{(u, v): u^2 + v^2 \leqslant 1\}.$$

Во всех рассмотренных примерах мы ввели элементарные события, представляющие собой мыслимые исходы опыта или наблюдения. Все события, связанные с данным опытом, могут быть описаны с помощью *элементарных событий* ω. Совокупность всех элементарных событий Ω будем называть *пространством элементарных событий*.

Таким образом, в общем случае

пространством элементарных событий будем называть произвольное множество $\Omega = \{\omega\}$, а элементы ω этого множества будем называть элементарными событиями.

Ввиду большого разнообразия случайных явлений нельзя дать более конкретное общее определение пространства элементарных событий. Для описания каждого реального опыта множество Ω выбирается наиболее подходящим образом.

1.4. Алгебра событий. В приведенных выше примерах в качестве составных событий, или просто событий, мы рассматривали подмножества множества Ω. Если Ω конечно или счетно,

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\} \quad \text{или} \quad \Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\},$$

то случайным событием, или просто *событием*, назовем любое подмножество множества Ω. В случае произвольного Ω событиями будем называть только подмножества из некоторого класса $\mathscr{A}$ подмножеств Ω, который будет определен после введения операций над событиями, совпадающими с операциями над множествами.

Суммой $A+B$ (или $A \cup B$) двух событий A и B назовем событие, состоящее из всех элементарных событий, принадлежащих по крайней мере одному из событий A или B. Можно сказать, что в реальном опыте событие, соответствующее $A+B$, состоит в том, что произошло по крайней мере одно из событий A или B.

Пусть в примере 3 событиями A и B являются попадания соответственно в большой и малый круг (рис. 2). Тогда событием $A+B$ является заштрихованная область на рис. 2, *а*.

Произведением AB (или $A \cap B$) называется событие, состоящее из элементарных событий, принадлежащих и A и B. Событие AB происходит тогда и только тогда, когда происходит и A и B (рис. 2, *б*). *Разностью* $A \setminus B$ называется событие, состоящее из элементов множества A, не принадлежащих B

(рис. 2, *в*). Событие $A \setminus B$ состоит в том, что A произошло, а B не произошло.

Событие Ω назовем *достоверным*; пустое множество $\varnothing$ назовем *невозможным* событием. Событие $\bar{A} = \Omega \setminus A$ называется *противоположным* событию A (рис. 2, *г*). Событие $\bar{A}$ означает, что A не произошло.

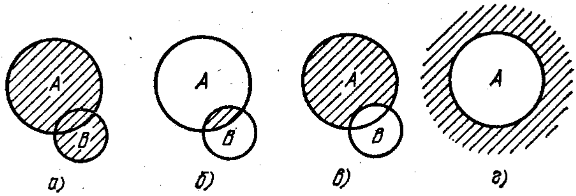


Рис. 2.

События A и B *несовместны*, если $AB = \varnothing$. Тот факт, что A является подмножеством B, будем записывать так: $A \subset B$ (или $B \supset A$). Это значит, что из наступления события A следует наступление B. В примере 3, если $A \subset B$, то попадание в область A, содержащуюся в B, означает также попадание в B. В случае $A \subset B$ мы будем говорить, что событие A влечет за собой событие B или что событие B следует из события A. Если $A \subset B$ и $B \subset A$, то мы будем говорить, что события A и B *равносильны* или *эквивалентны* и писать $A = B$. Принадлежность элемента множеству обозначается символом $\in$. Например, $\omega \in \Omega$.

Понятия произведения и суммы событий переносятся на бесконечные последовательности событий. Событие

$$A_1 + A_2 + \ldots + A_n + \ldots = \bigcup_{n=1}^{\infty} A_n$$

состоит из элементарных событий, принадлежащих хотя бы одному из событий A_n, $n = 1, 2, \ldots$. Событие $\bigcap_{n=1}^{\infty} A_n = A_1 A_2 \ldots A_n \ldots$ состоит из элементарных событий, принадлежащих каждому событию A_n, $n = 1, 2, \ldots$

Для произвольных событий непосредственно из определения легко проверить, что

$$AA = A, \quad A + A = A, \quad A\bar{A} = \varnothing.$$

Часто оказываются полезными следующие равенства:

$$\overline{A+B} = \bar{A}\bar{B}, \quad (A+B)C = AC + BC.$$

Докажем, например, второе. Нужно убедиться, что множества, стоящие в обеих частях равенства, состоят из одних и тех же элементов. Пусть произвольное $\omega \in (A+B)C$. Тогда $\omega \in A+B$ и $\omega \in C$. Из $\omega \in A+B$ следует, что ω принадлежит хотя бы одному слагаемому. Пусть, например, $\omega \in A$. Из $\omega \in A$ и $\omega \in C$ следует по определению произведения событий, что $\omega \in AC$, и, следовательно, $\omega \in AC + BC$. Таким образом, любой элемент множества $(A+B)C$ является элементом множества $AC+BC$, т. е. $(A+B)C \subset AC + BC$. Предположив, что $\omega \in AC + BC$, мы, используя рассуждения, аналогичные приведенным выше, покажем, что любой элемент $AC + BC$ является элементом $(A+B)C$. Отсюда следует доказываемое равенство, так как множества его левой и правой частей состоят из одних и тех же элементов. До проведения доказательства равенств полезно, считая A, B, C множествами на плоскости, сделать рисунки множеств, стоящих в левой и правой частях доказываемых равенств.

Дадим теперь определение некоторых классов подмножеств Ω. Уже было отмечено, что вероятность будет рассматриваться как функция от события. Для любых подмножеств Ω вероятность не всегда удается определить. В тех случаях, когда класс подмножеств приходится ограничивать, будем предполагать, что в результате любых введенных выше операций вновь получится множество из данного класса.

Пусть Ω — произвольное пространство элементарных событий, а $\mathscr{A}$ — некоторый класс подмножеств Ω.

Класс подмножеств $\mathscr{A}$ называется *алгеброй событий*, если $\Omega \in \mathscr{A}$ и если $AB \in \mathscr{A}$, $A+B \in \mathscr{A}$, $A \setminus B \in \mathscr{A}$ при любых $A \in \mathscr{A}$, $B \in \mathscr{A}$.

1.5. Вероятность. Перейдем к определению вероятности. Отметим сразу, что не существует общего определения вероятности, позволяющего сразу находить ее числовые значения. В качестве общего определения формулируется ряд аксиом, которым должна удовлетворять вероятность, определяемая для каждого конкретного опыта или случайного явления.

Числовая функция P, определенная на алгебре событий $\mathscr{A}$, называется вероятностью, если выполнены следующие аксиомы:

А1. Аксиома неотрицательности. *Для любого* $A \in \mathscr{A}$ $\mathsf{P}(A) \geqslant 0$.

А2. Аксиома нормированности.

$$P(\Omega) = 1.$$

А3. Аксиома аддитивности. *Если A и B несовместны, то*

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

Отметим, что аксиомы А1—А3 совершенно необходимы, если мы хотим, чтобы вероятность $\mathsf{P}(A)$ события A была числом, около которого колеблются частоты $h_n(A)$ при большом числе испытаний n, так как сами частоты $h_n(A)$ удовлетворяют А1—А3. Действительно, пусть некоторый опыт повторен n раз. Обозначим n_A число опытов, в которых осуществилось реальное событие A. Очевидно, что

$$h_n(A) = \frac{n_A}{n} \geqslant 0, \quad h_n(\Omega) = \frac{n_\Omega}{n} = \frac{n}{n} = 1.$$

Если реальные события A и B несовместны, то они осуществлялись при разных опытах и, следовательно, $n_{A+B} = n_A + n_B$. Отсюда

$$h_n(A+B) = \frac{n_{A+B}}{n} = \frac{n_A}{n} + \frac{n_B}{n} = h_n(A) + h_n(B),$$

что соответствует А3.

Для решения задач, связанных с бесконечными последовательностями событий, требуется дополнить приведенные аксиомы следующей аксиомой.

А4. Расширенная аксиома аддитивности. *Если в последовательности* $A_1, A_2, \ldots, A_n, \ldots$ *события попарно несовместны* (т. е. $A_iA_j = \varnothing$ *при* $i \neq j$) *и* $A = \bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, *то*

$$\mathsf{P}(A) = \sum_{n=1}^{\infty} \mathsf{P}(A_n).$$

Тройку $(\Omega, \mathscr{A}, \mathsf{P})$, в которой P удовлетворяет А1 — А4 и множество $\mathscr{A}$ не только является алгеброй событий, но и еще содержит счетные суммы и произведения событий, называют *вероятностным пространством*.

Система аксиом А1 — А4 вероятностного пространства дает самую общую математическую модель случайных явлений.

Приведем теперь несколько важных частных случаев вероятностных пространств. В дальнейшем вновь вводимые теоретико-вероятностные понятия позволят нам расширить набор частных случаев и приемов их построения.

1.6. Конечное вероятностное пространство. Пусть $\Omega = \{\omega\}$ — конечное множество (например, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$, где N — натуральное число); $\{p(\omega)\colon \omega \in \Omega\}$ — набор чисел, удовлетворяющий условиям

$$p(\omega) \geqslant 0, \quad \omega \in \Omega; \qquad \sum_{\omega \in \Omega} p(\omega) = 1. \tag{2}$$

Обозначим $\mathscr{A}$ множество всех подмножеств Ω. Вероятностью события $A \in \mathscr{A}$ назовем число $\mathsf{P}(A)$, определенное формулой

$$\mathsf{P}(A) = \sum_{\omega \in A} p(\omega) = \sum_{l=1}^{k} p(\omega_{i_l}), \tag{3}$$

где событие $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$. Если $A = \varnothing$, то по определению полагаем, что $\mathsf{P}(A) = 0$. Числа $\{p(\omega)\}$ являются вероятностями элементарных событий; мы их будем называть просто *элементарными вероятностями*. Таким образом, *вероятность события A равна сумме тех элементарных вероятностей $p(\omega)$, у которых ω входят в A.*

Непосредственно из определения $\mathsf{P}(A)$ следует, что выполняются аксиомы А1, А2: $\mathsf{P}(A) \geqslant 0$, $\mathsf{P}(\Omega) = 1$. Если $AB = \varnothing$, то

$$\mathsf{P}(A+B) = \sum_{\omega \in A+B} p(\omega) = \sum_{\omega \in A} p(\omega) + \sum_{\omega \in B} p(\omega) = \mathsf{P}(A) + \mathsf{P}(B),$$

и, следовательно, аксиома А3 тоже выполняется.

Определенное нами конечное вероятностное пространство будем иногда называть *конечной схемой*. В конечной схеме вероятность однозначно определяется элементарными вероятностями. Конечная схема во многих случаях служит хорошей математической моделью случайных явлений.

В дальнейшем различные частные случаи общей математической модели случайных явлений мы часто будем называть схемами, указывая их характерные особенности (конечная схема, схема независимых испытаний (§ 3) и т. д.).

Обсуждение приложений конечной схемы к описанию реальных явлений мы начнем с одного частного случая этой схемы, в котором вероятности $p(\omega)$, $\omega \in \Omega$ одинаковы. Если условиться число элементов множества M обозначать $|M|$, то из формулы (3), полагая $p(\omega) = 1/N$, получим

$$\mathsf{P}(A) = \frac{|A|}{|\Omega|} = \frac{k}{N}, \tag{4}$$

где $\Omega = \{\omega_1, \ldots, \omega_N\}$, $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$. Определение (4) называют *классическим определением вероятности*. Таким образом, согласно (4),

вероятность случайного события A равна отношению числа элементарных событий, при которых событие A происходит, к общему числу элементарных событий.

Классическое определение вероятности служит хорошей математической моделью тех случайных явлений, для которых исходы опыта в каком-либо смысле симметричны, и поэтому представляется естественным предположение об их равновозможности. Обычно это предположение оправдано в задачах из области азартных игр, лотерей и т. д. Это объясняется тем, что при изготовлении игральных костей, карт и организации лотерей заботятся о соблюдении равновозможности различных исходов. Такие же требования предъявляются к организации выборочного контроля и выборочных статистических исследований.

При использовании формулы (4) часто оказываются полезными различные комбинаторные формулы. Приведем наиболее распространенные из них.

Из конечного множества $\{a_1, a_2, \ldots, a_n\}$, состоящего из n различных элементов, можно образовывать различные наборы, состоящие из m $(m < n)$ элементов.

Упорядоченные наборы называют *размещениями*, а неупорядоченные — *сочетаниями*. Например, из множества $\{1, 2, 3\}$, выбирая по два элемента ($n = 3$, $m = 2$), можно образовать 6 размещений ((1,2), (2,1), (1,3), (3,1), (2,3), (3,2)) и 3 сочетания ((1,2), (1,3), (2,3)). Сочетание (1,2) можно записать в виде (2,1), так как при образовании сочетаний по определению порядок элементов не учитывается. Размещения из n элементов по n называют *перестановками*. Различные перестановки содержат одни и те же элементы, расположенные в разном порядке.

Число размещений, которые можно образовать, выбирая различными способами m элементов из n, обозначают A_n^m, а число сочетаний — обозначают символами C_n^m или $\binom{n}{m}$. Числа A_n^m, C_n^m можно найти по формулам

$$A_n^m = n^{[m]}, \qquad C_n^m = \frac{n^{[m]}}{m!},$$

где $m! = 1 \cdot 2 \cdot \ldots \cdot m$, а *обобщенная степень* $n^{[m]}$ определяется формулой $n^{[m]} = n(n-1)\ldots(n-m+1)$. Часто оказываются полезными следующие формулы:

$$C_n^m = \frac{n!}{m!\,(n-m)!}, \qquad C_n^m = C_n^{n-m}.$$

Пример 4. Из урны, содержащей M белых и $N - M$ черных шаров, наудачу извлекается n шаров. Найдем вероятность того, что среди выбранных n шаров окажется ровно m белых.

Слово «наудачу» в описаниях опыта встречается довольно часто. В данной задаче предполагается, что шары были хорошо перемешаны, что все они одного радиуса, одинаково гладкие и отличаются только цветом; выбирающий шаров не видит. В таком случае разумно предположить равновероятность элементарных событий и воспользоваться классическим определением вероятности.

За элементарные события естественно принять любые подмножества по n элементов, выбранные из множества N шаров. Из школьного курса математики известно, что число таких подмножеств равно C_N^n Таким образом, в формуле (4) нужно положить $|\Omega| = C_N^n$. Каждый набор шаров, входящий в интересующее нас событие (обозначим его A_m), состоит из двух частей: 1) m белых шаров и 2) $n-m$ черных шаров. Все такие наборы можно получить следующим образом. Сначала выберем части наборов из белых шаров; число таких частей C_M^m; затем отдельно составим части наборов из черных шаров; число таких частей C_{N-M}^{n-m}. Объединение любой части набора из белых шаров с любой частью набора из черных шаров дает полный набор шаров, принадлежащий A_m. Следовательно, $|A_m| = C_M^m C_{N-M}^{n-m}$ и по формуле (4)

$$\mathbf{P}(A_m) = P_n(m, N, M) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}. \qquad (5)$$

Здесь и в дальнейшем предполагается, что $C_n^m = 0$ при $m > n$. Набор чисел $P_n(0, N, M)$, $P_n(1, N, M)$, ... называют *гипергеометрическим распределением.*

В приложениях к выборочному контролю роль шаров играют N изделий проверяемой партии. Число M бракованных изделий (белых шаров) неизвестно. Может оказаться, что сплошь все изделия проверить нельзя: их слишком много или проверка приводит к уничтожению изделия (например, потребуется при проверке установить срок службы лампочки). Тогда из всей партии изделий отбирают для проверки небольшую часть из n изделий. Если среди выбранных изделий оказалось m бракованных, то полагают $\frac{M}{N} \approx \frac{m}{n}$. (Такой выбор приближенного значения будет обоснован в § 12, пример 5.) Иногда требуется оценить неизвестный параметр N. Пусть N — неизвестное число рыб в некотором водоеме. Можно провести отлов M рыб, пометить их и пустить обратно. По числу m помеченных рыб в повторном отлове из n рыб можно из приближенного равенства $\frac{M}{N} \approx \frac{m}{n}$ делать заключения о величине N: $N \approx M \cdot \frac{n}{m}$.

В теории вероятностей часто математические модели, имеющие приложения в самых различных областях, формулируются в терминах урновых схем, в терминах размещений дробинок или частиц по ящикам и т. д. Такая форма описания обычно не приводит к затруднениям в практическом их использовании. Формулировка математической модели в терминах, используемых в определенной узкой области приложений, вызовет у других потребителей большие затруднения, чем традиционные формулировки со знакомыми всем шарами и урнами. Для специалистов по выборочному контролю «белые шары», по-видимому, более привлекательны, чем «меченые рыбы».

В дальнейшем мы, как правило, будем придерживаться традиционных для теории вероятностей формулировок.

Пример 5. Дадим более детальное описание исходов опыта, описанного в предыдущем примере. Пусть N шаров, имеющихся в урне, занумерованы числами 1, 2, ..., N. Из урны наудачу по одному извлекается n шаров, причем вынутые шары обратно не возвращаются. В данном опыте элементарными событиями ω являются цепочки $\omega = (x_1, x_2, \ldots, x_n)$, составленные из чисел 1, 2, ..., N, причем среди $x_1, \ldots, x_n$ нет одинаковых чисел. Число элементов $|\Omega|$ множества Ω можно найти непосредственно или воспользоваться формулой для числа размещений из N элементов по n:

$$|\Omega| = A_N^n = N(N-1)\ldots(N-n+1).$$

Выражение в правой части этого равенства называют *обобщенной степенью числа* N. Мы будем использовать следующее обозначение:

$$m^{[k]} = m(m-1)\ldots(m-k+1). \qquad (6)$$

При подходящей организации описанного опыта можно считать элементарные события равновероятными и использовать классическое определение вероятности. Описанную схему выбора называют *схемой случайного выбора без возвращения*.

Элементарные события предыдущего примера по отношению к данному являются более «крупными». В предыдущем примере рассматривались неупорядоченные наборы номеров, и, следовательно, из каждого такого набора можно с учетом расположения получить одинаковое числое $n!$ наборов, являющихся элементарными событиями в рассматриваемой схеме. Таким образом, каждому элементарному событию примера 4 соответствует $n!$ элементарных событий примера 5. Это приводит к совпадению вероятностей событий, определенных в обоих примерах.

Пример 6. Из N одинаковых карточек, занумерованных числами $0, 1\ 2, \ldots, N-1$, извлекается по одной n карточек. После каждого извлечения вынутая карточка возвращается. Так же, как и в примере 1, при подходящей организации опыта все возможные исходы естественно считать равновероятными и использовать классическое определение.

Если номера извлекаемых карточек записывать, то в результате получится цепочка ω из n номеров: $\omega = (x_1, x_2, \ldots, x_k)$, где x_i могут принимать любое значение из $\{0, 1, \ldots, N-1\}$ независимо друг от друга. Множество всех цепочек ω является в данной задаче пространством элементарных событий Ω. Будем предполагать, что все элементарные события равновероятны. Описанную схему выбора называют *схемой случайного выбора с возвращением*.

Найдем число элементарных событий $|\Omega|$ множества Ω. Разобьем множество цепочек $\omega = (x_1, \ldots, x_n)$ на N непересекающихся множеств по значению x_1:

$$\{\omega: x_1 = 0\}, \quad \{\omega: x_1 = 1\}, \quad \ldots, \quad \{\omega: x_1 = N-1\}.$$

Каждое из этих множеств по x_2 можно в свою очередь разбить на N групп, и, следовательно, по значениям пары $\{x_1 x_2\}$ образуется N^2 групп; разбиений по значениям тройки $\{x_1, x_2, x_3\}$ получим N^3 и т. д. В результате n разбиений получим N^n групп, каждая из которых содержит один элемент. Таким образом, $|\Omega| = N^n$, $p(\omega) = N^{-n}$ и вероятность определяется формулой (4). Вычислим вероятность события $A_m =$ $= \{$среди n вынутых карточек ровно у m карточек номера не превосходят $M - 1\}$. Среди n мест цепочки $\omega = (x_1, \ldots, x_n)$ можно C_n^m способами выбрать места, на которых будем помещать номера карточек, не превосходящие $M - 1$. Эти места можно заполнить M^m способами. Оставшиеся $n - m$ мест $(N - M)^{n-m}$ способами заполним карточками с номерами, большими $M - 1$. Таким образом, $|A_m| = C_n^m M^m (N - M)^{n-m}$, и по формуле (4) находим

$$\mathbf{P}(A_m) = \frac{C_n^m M^m (N - M)^{n-m}}{N^n} = C_n^m \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{n-m},$$

$$m = 0, 1, 2, \ldots, n.$$

Числа, полученные в результате достаточно хорошей реализации случайного выбора с возвращением при $N = 10$, называют *случайными равномерно распределенными числами* (или просто случайными числами). Таблицы таких чисел (см. [1], [11], [14]) можно использовать для моделирования случайных явлений. Пусть, например, нужно моделировать результаты n опытов, заключающихся в подбрасывании симметричной монеты. Условимся считать, что выпадению герба соответствует четная цифра (0, 2, 4, 6, 8) в последовательности случайных чисел. Заменив четные цифры буквой «Г» (герб), а нечетные — «Р» (решетка), получим последовательность из двух символов: Г, Р. Очевидно, что равновероятны все возможные последовательности, полученные указанным способом. Обозначим $n_\Gamma(n)$ число «гербов» в последовательности длины n. В качестве примера по таблице случайных чисел найдены значения $n_\Gamma = n_\Gamma(n)$ и вычислены частоты $h_\Gamma(n) = n_\Gamma/n$. Результаты приведены в следующей таблице:

n	5	10	20	30	40	50
$n_г$	3	7	13	15	21	25
$h_г(n)$	0,60	0,70	0,65	0,50	0,53	0,50

n	60	70	80	90	100
$n_г$	32	34	37	43	47
$h_г(n)$	0,53	0,49	0,46	0,48	0,47

График функции $h_г(n)$ приведен на рис. 3.

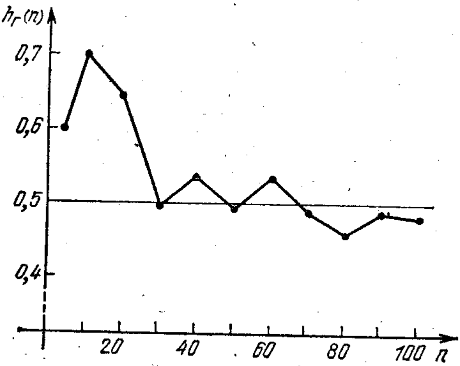


Рис. 3.

Пример 7. Пусть один раз брошена игральная кость. Естественно в этом случае положить $\Omega = \{1, 2, 3, 4, 5, 6\}$. Если кость симметрична, то вероятности элементарных событий выбираем одинаковыми: $p(1) = p(2) = \ldots = p(6) = 1/6$. В этом случае получаем классическое определение вероятности. Если кость не симметрична, то приходится использовать конечную схему с различными $p(i)$, $i = 1, \ldots, 6$, которые можно оценить экспериментально. Построение оценок неизвестных параметров схемы является одной из задач математической статистики (см. § 12).

1.7. Счетное вероятностное пространство. Пусть $\Omega = \{\omega\}$ — счетное множество ($\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$); $\mathscr{A}$ — множество всех подмножеств Ω; $\{p(\omega)\}$ — набор чисел, удовлетворяющих условиям

$$p(\omega) \geqslant 0, \quad \omega \in \Omega; \quad \sum_{n=1}^{\infty} p(\omega_n) = 1.$$

Вероятность события $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_l}, \ldots\}$ назовем число $\mathsf{P}(A)$, определяемое формулами

$$\mathsf{P}(A) = \sum_{\omega \in A} p(\omega) = \sum_{l=1}^{\infty} p(\omega_{i_l}), \quad A \neq \varnothing, \tag{7}$$

и $\mathsf{P}(\varnothing) = 0$. Таким образом,

вероятность события A равна сумме ряда, составленного из элементарных вероятностей $p(\omega)$, у которых ω входят в A.

Аксиомы А1 — А4 проверяются так же, как и в конечной схеме. Порядок нумерации элементарных событий не влияет на определение, так как $p(\omega) \geqslant 0$ и сумма ряда, входящая в (7), не изменяется при изменении порядка суммирования. Построенное *счетное вероятностное пространство* будем иногда называть *счетной схемой.*

Отметим, что вероятность $\mathsf{P}(A)$ так же, как и в конечной схеме, однозначно определяется вероятностями элементарных событий $p(\omega)$. Конечная схема является частным случаем счетной схемы с $p(\omega_k) = 0$, $k \geqslant N + 1$. Счетную и конечную схемы называют *дискретной схемой* или *дискретным вероятностным пространством.*

Пример 8. Монета подбрасывается до тех пор, пока не выпадет два раза подряд одной и той же стороной. Предположим, что нам требуется определить вероятность событий: $B = \{$опыт закончится за четное число подбрасываний$\}$, $A = \{$опыт продлится не дольше 5 бросаний$\}$. Положим $\Omega \triangleq \{n\colon\ n \geqslant 2\}$, где натуральные числа n будем интерпретировать как продолжительность опыта. Попытаемся естественно ввести вероятности элементарных событий $p(n)$. При n подбрасываниях монеты возможно 2^n различных исходов опыта. Будем прдполагать, что эти исходы равновероятны. Среди 2^n исходов есть два исхода, которые соответствуют выпадению монеты впервые два раза подряд одной стороной на n-м испытании. Таким образом, естественно предположить, что $p(n) = 2/2^n = 2^{-n+1}$ $(n \geqslant 2)$.

Приведенные рассуждения о вероятности $p(n)$ не являются математическим доказательством; их можно рассматривать только как некоторые наводящие соображения. Нетрудно проверить, что выбранные нами числа удовлетворяют условиям

$$p(n) = 2^{-n+1} > 0, \quad n \geqslant 2; \qquad \sum_{n=2}^{\infty} p(n) = 2 \sum_{n=2}^{\infty} \frac{1}{2^n} = 1.$$

Теперь мы можем определить вероятность по формуле (7). Рассматриваемые в данном примере события A и B можно представить в виде

$$A = \{2, 3, 4, 5\}, \quad B = \{2, 4, 6, 8, \ldots\}.$$

Следовательно,

$$\mathsf{P}(A) = p(2) + p(3) + p(4) + p(5) = \\ = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} = \frac{15}{16},$$

$$\mathsf{P}(B) = \sum_{k=1}^{\infty} p(2k) = 2 \sum_{k=1}^{\infty} \frac{1}{2^{2k}} = \frac{2}{3}.$$

1.8. Непрерывное вероятностное пространство. Положим $\Omega = \{u = (u_1, \ldots, u_n)\colon u \in G\}$, где G — квадрируемая область n-мерного евклидова пространства. Обозначим $\mathscr{A}$ систему квадрируемых подмножеств области G. Из курса анализа известно, что сумма, произведение и разность квадрируемых фигур являются квадрируемыми фигурами. Таким образом, $\mathscr{A}$ является алгеброй событий. Пусть $\pi(u_1, \ldots, u_n) \geqslant 0$ — интегрируемая на области G функция и интеграл от нее по области G равен 1.

Вероятностью события A назовем число $\mathsf{P}(A)$, определяемое формулой

$$\mathsf{P}(A) = \int \ldots \int_A \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n, \tag{8}$$

где интеграл в правой части (8) является n-кратным интегралом Римана. Используя свойства интегралов, нетрудно проверить, что функция $\mathsf{P}(A)$ удовлетворяет аксиомам А1 — А3.

Отметим, что вместо конечной области G можно рассматривать n-мерное пространство, а интеграл в (8) понимать как несобственный. Функцию $\mathsf{P}(A)$, определенную на алгебре $\mathscr{A}$, можно продолжить на более широкую систему множеств, содержащую счетные суммы и произведения событий.

Построенное вероятностное пространство мы будем называть иногда *непрерывной вероятностной моделью* или просто *непрерывной схемой.*

Рассмотрим частный случай общей непрерывной схемы, положив $\pi(u_1, \ldots, u_n) = 1/m(G)$, если $u \in G$, и $\pi(u_1, \ldots, u_n) = 0$, если $u \notin G$. Здесь $m(G)$ — площадь или объем области G. При таком выборе функции $\pi(u)$ формула (8) запишется в следующем виде:

$$\mathsf{P}(A) = m(A)/m(G). \tag{9}$$

Такое определение вероятности называют *геометрическим.* Геометрическое определение вероятности можно рассматривать как обобщение классического определения на случай опытов с бесконечным числом исходов. В дальнейшем мы иногда будем употреблять выражение «точка равномерно распределена на множестве G». Это означает, что вероятность надо вычислять по формуле (9).

В качестве иллюстрации использования геометрического определения рассмотрим следующий пример.

Пример 9. На обслуживающее устройство в промежуток времени $[0, T]$ должны поступить две заявки. Если разность между моментами поступления заявок меньше t, то вторая заявка теряется. Требуется найти вероятность потери заявки.

Обозначим t_1, t_2 моменты поступления заявок. Положим $\Omega = \{(t_1, t_2)\colon\ 0 \leqslant t_1 \leqslant T,\ 0 \leqslant t_2 \leqslant T\}$; $A = \{$заявка будет потеряна$\}$. Тогда $A = \{(t_1, t_2)\colon |t_1 - t_2| < t\}$ (рис. 4). Если воспользоваться геометрическим определением, то потребуется вычислить площадь $m(\Omega) = T^2$ и $m(A) = T^2 - (T - t)^2$. По формуле (9) находим $\mathsf{P}(A) = 1 - \left(1 - \frac{t}{T}\right)^2$. Использование геометрического определения вероятности содержит в себе некоторые предположения о законе поступления заявок. Соответствие выбранной модели случайного явления действительности может быть оценено на основе экспериментов.

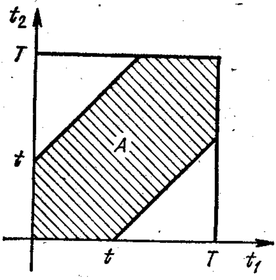


Рис. 4.

Получим теперь некоторые простейшие следствия из аксиом A1 — A4. Из аксиом A2, A3 и равенства $A + \bar{A} = \Omega$ следует, что

$$\mathsf{P}(\bar{A}) = 1 - \mathsf{P}(A). \quad (10)$$

Полагая здесь $A = \Omega$, получим $\mathsf{P}(\varnothing) = 0$. Для любых событий A и B имеет место формула

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB). \quad (11)$$

Действительно, представим события $A + B$ и B в виде $A + B = A + B\bar{A}$ и $B = B\bar{A} + BA$. События в правых частях этих равенств несовместны и, следовательно,

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B\bar{A}), \quad \mathsf{P}(B) = \mathsf{P}(B\bar{A}) + \mathsf{P}(AB).$$

Отсюда легко следует (11). Используя формулу (11), получим формулу для $\mathsf{P}(A_1 + A_2 + A_3)$. Положим $A = A_1 + A_2$, $B = A_3$. Тогда из формулы (11) и формулы $AB = A_1A_3 + A_2A_3$ следует, что

$$\mathsf{P}(A_1 + A_2 + A_3) = \mathsf{P}(A_1 + A_2) + \mathsf{P}(A_3) - \mathsf{P}(A_1A_3 + A_2A_3).$$

Отсюда, применив к вероятностям $\mathsf{P}(A_1 + A_2)$ и $\mathsf{P}(A_1A_3 + A_2A_3)$ формулу (11) и воспользовавшись равенством $(A_1A_3)(A_2A_3) = A_1A_2A_3$, получим

$$\mathsf{P}(A_1 + A_2 + A_3) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \mathsf{P}(A_3) -$$
$$- \mathsf{P}(A_1A_2) - \mathsf{P}(A_1A_3) - \mathsf{P}(A_2A_3) + \mathsf{P}(A_1A_2A_3).$$

Общая формула для вероятности суммы n событий имеет следующий вид:

$$\mathsf{P}(A_1+A_2+\ldots+A_n)=\sum_{k=1}^{n}\mathsf{P}(A_k)-\sum_{1\leqslant k<l\leqslant n}\mathsf{P}(A_kA_l)+$$
$$+\sum_{1\leqslant k<l<m\leqslant n}\mathsf{P}(A_kA_lA_m)-\ldots+(-1)^{n+1}\mathsf{P}(A_1\ldots A_n)=$$
$$=\sum_{r=1}^{n}(-1)^{r+1}\sum_{1\leqslant l_1<\ldots<l_r\leqslant n}\mathsf{P}(A_{l_1}\ldots A_{l_r}). \qquad (12)$$

В доказательстве по индукции формулы (12) переход от $n-1$ к n проводится аналогично приведенному выше переходу от 2 к 3.

Непосредственно из (11) для любых A и B получаем

$$\mathsf{P}(A+B)\leqslant\mathsf{P}(A)+\mathsf{P}(B). \qquad (13)$$

Если $A\subset B$, то

$$\mathsf{P}(A)\leqslant\mathsf{P}(B), \qquad (14)$$

так как $B=A+B\bar{A}$, $\mathsf{P}(B)=\mathsf{P}(A)+\mathsf{P}(B\bar{A})$ и $\mathsf{P}(B\bar{A})\geqslant 0$. Если события $A_1, A_2, \ldots, A_n$ попарно несовместны, т. е. $A_iA_j=\varnothing$ при любых $i\neq j$, то

$$\mathsf{P}(A_1+A_2+\ldots+A_n)=$$
$$=\mathsf{P}(A_1)+\mathsf{P}(A_2)+\ldots+\mathsf{P}(A_n). \qquad (15)$$

Эта формула доказывается по индукции при помощи аксиомы аддитивности А3; кроме того, (15) можно получить как частный случай (12). По индукции из (13) следует, что

$$\mathsf{P}(A_1+\ldots+A_n)\leqslant\mathsf{P}(A_1)+\ldots+\mathsf{P}(A_n).$$

Рассмотрим несколько примеров, иллюстрирующих использование полученных из аксиом следствий.

Пример 10. Оценим вероятность получения какого-либо выигрыша в «Спортлото» по одной карточке. Участник лотереи из 49 номеров отмечает 6. После того как участник сдал карточку, проводится выборка $n=6$ номеров. Пусть $A_m=\{m$ номеров, отмеченных участником, попало в выборку$\}$. Если число номеров, отмеченных участником и попавших в выборку, оказалось больше 2, то участник получает какой-либо выигрыш. Таким образом, $A=$ {получение какого-либо выигрыша} $=A_3+A_4+A_5+A_6$ и, следовательно, $\bar{A}=A_0+A_1+A_2$. По формуле (5) с $N=49$, $M=6$, $n=6$ находим

$$\mathsf{P}(A_0)=0{,}435965,\ \mathsf{P}(A_1)=0{,}413019,\ \mathsf{P}(A_2)=0{,}132378,$$
$$\mathsf{P}(A_3)=0{,}017650,\ \mathsf{P}(A_4)=0{,}000969,\ \mathsf{P}(A_5)=0{,}000018,$$
$$\mathsf{P}(A_6)=0{,}00000007151.$$

Отсюда, используя (15), получим

$$\mathsf{P}(\bar{A})=\mathsf{P}(A_0)+\mathsf{P}(A_1)+\mathsf{P}(A_2)=0{,}981362$$

и

$$\mathsf{P}(A)=1-\mathsf{P}(\bar{A})=0{,}018638.$$

Пример 11. Найдем вероятность того, что среди n случайных чисел (см. пример 6) цифра «1» встретится хотя бы один раз.

Множеством элементарных событий Ω являются последовательности, которые можно составить из всех 10 цифр; $\Omega=\{(x_1, x_2, \ldots, x_n)\}$, где $x_k=0, 1, \ldots, 9$, $k=1, \ldots, n$. В примере 6 было показано, что $|\Omega|=10^n$. Положим $A=$ {среди n случайных чисел встретилась хотя бы одна «1»}. Тогда $\bar{A}=$ {среди n случайных чисел «1» не встретилась ни разу}. Таким образом, в $\bar{A}$ входят цепочки, составленные только из 9 цифр, так как «1» не может быть использована при составлении цепочек. Рассуждая так же, как в примере 6, при вычислении $|\Omega|$, получим $|\bar{A}|=9^n$. Следовательно, по формуле (4) имеем $\mathsf{P}(\bar{A})=9^n/10^n$. Отсюда, воспользовавшись формулой (10), найдем

$$\mathsf{P}(A)=1-\mathsf{P}(\bar{A})=1-(9/10)^n.$$

Пример 12. Брошено две игральных кости. Найдем вероятность события $A=$ {хотя бы на одной из костей выпало не больше двух очков}.

Множество всех возможных исходов можно записать в виде

$$\Omega=\{(i, j)\colon i, j=1, 2, \ldots, 6\},$$

где элементарное событие (i, j) соответствует выпадению i очков на первой кости и j очков на второй. Нетрудно проверить, что $|\Omega| = 6 \cdot 6 = 36$. Обозначим B_1 событие, состоящее в том, что на 1-й кости выпало не больше двух очков; аналогичное событие для 2-й кости обозначим B_2. Тогда

$$B_1 = \{(i, j)\}: i = 1, 2;\ j = 1, 2, \ldots, 6\},$$
$$B_2 = \{(i, j): i = 1, \ldots, 6;\ j = 1, 2\}$$

и, следовательно, $|B_1| = |B_2| = 12$. Так как $B_1B_2 = \{(i, j): i, j = 1, 2\}$, то $|B_1B_2| = 4$. Таким образом,

$$\mathsf{P}(B_1) = \mathsf{P}(B_2) = \frac{12}{36} = \frac{1}{3}, \quad \mathsf{P}(B_1B_2) = \frac{1}{9}.$$

Отсюда по формуле (11) находим

$$\mathsf{P}(A) = \mathsf{P}(B_1 + B_2) = \mathsf{P}(B_1) + \mathsf{P}(B_2) - \mathsf{P}(B_1B_2) = \frac{1}{3} + \frac{1}{3} - \frac{1}{9} = \frac{5}{9}.$$

Пример 13. Рассмотрим некоторую сложную систему, состоящую из n блоков. Пусть в некоторой вероятностной модели определены события $A_k = \{k\text{-й}$ блок не выйдет из строя за рассматриваемый период$\}$ и их вероятности $\mathsf{P}(A_k)$, $k = 1, 2, \ldots, n$. Например, вероятности $\mathsf{P}(A_k)$ найдены приближенно по результатам испытаний отдельных блоков. Оценим вероятность события $A = \{$система не выйдет из строя за рассматриваемый период$\} = A_1A_2 \ldots A_n$. Так как $\bar{A} = \bigcup_{k=1}^{n} \bar{A}_k$, то

$$1 - \mathsf{P}(A) = \mathsf{P}(\bar{A}) = \mathsf{P}\left(\bigcup_{k=1}^{n} \bar{A}_k\right) \leqslant \sum_{k=1}^{n} \mathsf{P}(\bar{A}_k).$$

Кроме того, при любом k

$$A_1 \ldots A_n \subset A_k, \quad \mathsf{P}(A) \leqslant \mathsf{P}(A_k).$$

Окончательно получаем

$$1 - \sum_{k=1}^{n} (1 - \mathsf{P}(A_k)) \leqslant \mathsf{P}(A) \leqslant \min_k \mathsf{P}(A_k).$$

Задачи

1. Показать, что $A \setminus (B \setminus C) \neq A \setminus B + C$. Найти более простое выражение для $A \setminus (B \setminus C)$.

2. Пусть A, B, C — три произвольных события. Найти выражения для событий, состоящих в том, что из A, B, C

1) произошло только A,
2) произошли A и B, а C не произошло,
3) все три события произошли,
4) произошло по крайней мере одно из событий,
5) ни одно событие не произошло.

3. Среди 100 изделий данной партии имеется 5 бракованных. Найти вероятность того, что среди 10 случайно отобранных изделий не больше одного бракованного.

4. Найти вероятность того, что дни рождения 12 человек приходятся на разные месяцы года.

5. Участник лотереи «Спортлото» заполнил две карточки так, что все зачеркнутые им номера на обеих карточках — разные. Найти вероятность того, что участник не угадал ни одного номера.

6. Трем радиостанциям разрешена работа на любой из трех заданных частот. Найти вероятности событий: $A = \{$все радиостанции работают на разных частотах$\}$; $B = \{$две радиостанции работают на одинаковых частотах, а 3-я — на другой частоте$\}$. Предположить, что все возможные выборы частот радиостанциями равновероятны.

7. По схеме случайного выбора с возвращением из множества целых чисел $\{1, 2, \ldots, N\}$ выбираются два числа ξ и η. Обозначим p_N вероятность события $\xi^2 + \eta^2 \leqslant N^2$. Найти $\lim\limits_{N \to \infty} p_N$.

8. Точка A равномерно распределена в квадрате со стороной 1. Найти вероятности событий:

а) расстояние от точки A до фиксированной стороны квадрата не превосходит x;

б) расстояние от точки A до ближайшей стороны квадрата не превосходит x.

9. На паркет, составленный из правильных треугольников со стороной a, случайно бросается монета радиуса $r \left(r < a\dfrac{\sqrt{3}}{6} \right)$. Найти вероятность того, что монета не заденет границу ни одного из треугольников.

10. Стержень длины l разломан в двух наудачу выбранных точках. Чему равна вероятность того, что из полученных отрезков можно составить треугольник?

11. В n конвертов разложено по одному письму n адресатам. На каждом конверте наудачу написан один из n адресов. Найти вероятность p_n того, что хотя бы одно письмо попадет по назначению; вычислить $\lim\limits_{n \to \infty} p_n$.

Указание: использовать формулу (12).

§ 2. УСЛОВНЫЕ ВЕРОЯТНОСТИ. НЕЗАВИСИМОСТЬ СОБЫТИЙ

2.1. Условные вероятности. Прежде чем переходить к формальному определению, приведем некоторые соображения, которые помогут понять естественность определения.

Пусть проводится n опытов. Обозначим n_A число опытов, в которых наступило реальное событие A. Предположим, что среди n_A опытов вместе с A наступило n_{AB} раз событие B. Частотой события B по отношению к опытам с исходом A является отношение $h_n(B|A) = n_{AB}/n_A$. Если числитель и знаменатель в этом отношении поделить на n, то получим выражение, связывающее $h_n(B|A)$ с безусловными частотами $h_n(A)$ и $h_n(AB)$:

$$h_n(B\,|\,A) = \frac{n_{AB}/n}{n_A/n} = \frac{h_n(AB)}{h_n(A)}. \tag{1}$$

При больших n частоты $h_n(A)$, $h_n(AB)$ близки к вероятностям $\mathsf{P}(A)$, $\mathsf{P}(AB)$. Равенство (1) можно положить в основу определения условной вероятности. Пусть $\mathsf{P}(A) > 0$.

Число $\mathsf{P}(B|A)$, *определяемое равенством*

$$\mathsf{P}(B\,|\,A) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A)}, \tag{2}$$

называется условной вероятностью события B *при условии* A.

Условную вероятность $\mathsf{P}(B|A)$ будем иногда обозначать $\mathsf{P}_A(B)$.

Пример 1. На складе собрано N изделий от двух мастерских, причем от 1-й мастерской получено N_1 изделий, среди которых M_1 изделий первого сорта, а от 2-й мастерской — N_2 изделий и среди них M_2 изделий первого сорта. Полученное со склада изделие выбрано наудачу. Найдем вероятность события $B =$ $= \{$получено первосортное изделие$\}$. По классическому определению вероятности очевидно имеем $\mathsf{P}(B) = (M_1 + M_2)/N$. Предположим, что удалось установить, какой мастерской произведено полученное изделие; например, выяснили, что изделие произведено 1-й мастерской. Обозначим это событие A. При этой дополнительной информации естественно искать условную вероятность $\mathsf{P}(B|A)$. По формуле (2) находим

$$\mathsf{P}(B\mid A)=\frac{\mathsf{P}(AB)}{\mathsf{P}(A)}=\frac{M_1/N}{N_1/N}=\frac{M_1}{N_1},$$

т. е. условная вероятность является в данном случае долей первосортной продукции среди продукции, произведенной 1-й мастерской.

Пусть фиксировано некоторое событие A с $\mathsf{P}(A)>0$. Функция $\mathsf{P}_A(B)=\mathsf{P}(B\mid A)=\mathsf{P}(AB)/\mathsf{P}(A)$, определенная для всех $B\in\mathscr{A}$, удовлетворяет аксиомам A1—A4. Таким образом, для $\mathsf{P}_A(B)$ справедливы все следствия (10)—(15), полученные в § 1 непосредственно из аксиом.

Равенство (2) можно записать в виде

$$\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B\mid A). \tag{3}$$

Отсюда по индукции легко получить более общую формулу

$$(\mathsf{P}\,A_1A_2\ldots A_n)=$$
$$=\mathsf{P}(A_1)\,\mathsf{P}(A_2\mid A_1)\ldots\mathsf{P}(A_n\mid A_1A_2\ldots A_{n-1}). \tag{4}$$

Действительно, при $n=2$ формула (4) совпадает с (3). Пусть формула (4) доказана для $n-1$ сомножителей. Тогда для n сомножителей формула (4) следует из предположения индукции и равенства (3), в котором нужно положить $A=A_1A_2\ldots A_{n-1}$, $B=A_n$.

Формулы (3), (4) очевидно не пригодны для вычисления вероятностей произведений событий, так как правые части этих формул содержат условные вероятности, для вычисления которых нужно знать вероятности произведений. Полезность формул (3) и (4) обнаруживается при построении математических моделей серий опытов, которые будут рассмотрены в § 3. Здесь мы ограничимся разбором отдельных примеров.

В классическом и геометрическом определении вероятности мы исходили из «равновозможности» исходов опыта. Теперь мы обратимся к примерам вероятностных пространств, построение которых основано на знании значений условных вероятностей.

В приложениях часто оказывается, что условные вероятности естественно задаются условиями опыта или определяются приближенно из дополнительных опытов. Например, если при проведении серии из n опытов реальное событие A наступило n_A раз и n_{AB} раз событие A наступило вместе с реальным событием B, то в качестве приближенного значения условной вероятности $\mathsf{P}(B|A)$ можно взять частоту $h_n(B|A) = n_{AB}/n_A$. В сериях опытов с перекладыванием или извлечением шаров из урн можно естественно задать условные вероятности получения выборки определенного типа из данной урны, если из предыдущих опытов известен ее состав. Отправляясь от заданных значений условных вероятностей, можно вычислить вероятности элементарных событий.

Пример 2. Из первой урны, содержащей 5 белых и 3 черных шара, наудачу переложили один шар неизвестного цвета во вторую урну, содержащую 2 белых и 3 черных шара. После этого из второй урны вынули один шар. Построим математическую модель описанного опыта. Обозначим A_1 и A_2 события, состоящие в том, что из первой и второй урн были извлечены белые шары. Из соображений равновозможности исходов естественно считать, что $\mathsf{P}(A_1) = 5/8$. Если из первой урны во вторую был переложен белый шар, то вторая урна стала содержать 3 белых и 3 черных. Следовательно, при этом условии нужно положить $\mathsf{P}(A_2|A_1) = 3/6 = 1/2$. Аналогично определяем условную вероятность события A_2 при перекладывании черного шара: $\mathsf{P}(A_2|\bar{A}_1) = 4/6 = 2/3$. Из приведенных рассуждений (не являющихся математическим доказательством) можно сделать вывод, что в математической модели, которую мы хотим построить, должны быть выполнены равенства

$$\mathsf{P}(A_1) = 5/8, \quad \mathsf{P}(A_2|A_1) = 1/2, \quad \mathsf{P}(A_2|\bar{A}_1) = 2/3. \qquad (5)$$

Таким образом, нужно найти подходящее вероятностное пространство $(\Omega, \mathscr{A}, \mathsf{P})$, алгебра $\mathscr{A}$ которого содержит события A_1, A_2, а вероятность P определена так, что выполняются равенства (5).

Положим $\Omega=\{A_1A_2,\ \bar{A}_1A_2,\ A_1\bar{A}_2,\ \bar{A}_1\bar{A}_2\}$, $\mathscr{A}$ — множество всех подмножеств Ω и

$$\begin{aligned}\mathsf{P}(A_1A_2)&=\mathsf{P}(A_1)\,\mathsf{P}(A_2\mid A_1)=5/8\cdot 1/2,\\ \mathsf{P}(\bar{A}_1A_2)&=\mathsf{P}(\bar{A}_1)\,\mathsf{P}(A_2\mid \bar{A}_1)=3/8\cdot 2/3,\\ \mathsf{P}(A_1\bar{A}_2)&=\mathsf{P}(A_1)\,\mathsf{P}(\bar{A}_2\mid A_1)=5/8\cdot 1/2,\\ \mathsf{P}(\bar{A}_1\bar{A}_2)&=\mathsf{P}(\bar{A}_1)\,\mathsf{P}(\bar{A}_2\mid \bar{A}_1)=3/8\cdot 1/3.\end{aligned}\tag{6}$$

Вероятности (6) однозначно определяют вероятность любого подмножества Ω.

Построение вероятностных пространств в более общих случаях с использованием условных вероятностей и независимости событий будет рассмотрено в § 3.

2.2. Формула полной вероятности. Формула Байеса. Будем говорить, что события B_1, B_2, ..., B_n образуют *разбиение*, если

$$\Omega=B_1+B_2+\ldots+B_n \tag{7}$$

и $B_iB_j=\varnothing$ при любых $i\neq j$.

Для любого разбиения (7) имеет место следующая формула (*формула полной вероятности*):

$$\mathsf{P}(A)=\sum_{k=1}^{n}\mathsf{P}(B_k)\,\mathsf{P}(A\mid B_k). \tag{8}$$

Для доказательства этой формулы заметим, что A можно представить в виде следующей суммы попарно несовместных событий:

$$A=AB_1+AB_2+\ldots+AB_n.$$

Отсюда, воспользовавшись аксиомой А3 и формулой (3), получим формулу (8):

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

Заменив в равенстве

$$\mathsf{P}(B_k \mid A) = \frac{\mathsf{P}(AB_k)}{\mathsf{P}(A)} = \frac{\mathsf{P}(B_k)\,\mathsf{P}(A \mid B_k)}{\mathsf{P}(A)}$$

вероятность $\mathsf{P}(A)$ по формуле (8), получим *формулы Байеса*

$$\mathsf{P}(B_k \mid A) = \frac{\mathsf{P}(B_k)\,\mathsf{P}(A \mid B_k)}{\sum\limits_{i=1}^{n} \mathsf{P}(B_i)\,\mathsf{P}(A \mid B_i)}. \tag{9}$$

Формулы Байеса часто интерпретируются как формулы, позволяющие по *априорным* (известным предварительно, до проведения опыта) вероятностям $\mathsf{P}(B_k)$ и по результатам опыта (наступления события A) найти *апостериорные* (вычисленные после опыта) вероятности $\mathsf{P}(B_k|A)$. Довольно содержательный пример, указывающий возможные направления применений формул Байеса, содержится в задаче 10.

Формула полной вероятности, так же как формулы (3) и (4), может быть использована для построения подходящего вероятностного пространства по заданным условным вероятностям. Ниже рассматриваются два примера, в которых условные вероятности естественно считать заданными.

Пример 3. На фабрике, изготовляющей болты, первая машина производит 25%, вторая — 35%, третья — 40% всех изделий. Брак в их продукции составляет соответственно 5%, 4%, 2%. Найти

а) вероятность того, что случайно выбранный болт оказался дефектным,

б) вероятность того, что если случайно выбранный болт оказался дефектным, то он произведен первой, второй и третьей машинами.

Обозначим A событие, состоящее в том, что случайно выбранный болт — дефектный, а B_1, B_2, B_3 — события, состоящие в том, что этот болт произведен соответственно первой, второй и третьей машинами. Очевидно, что формула (8) применима. Таким образом, используя условие задачи, получим

$$\mathsf{P}(A) = \mathsf{P}(B_1)\,\mathsf{P}(A \mid B_1) + \mathsf{P}(B_2)\,\mathsf{P}(A \mid B_2) + \mathsf{P}(B_3)\,\mathsf{P}(A \mid B_3) = 0{,}25 \cdot 0{,}05 + 0{,}35 \cdot 0{,}04 + 0{,}40 \cdot 0{,}02 = 0{,}0345.$$

К тем же событиям можно применить формулы Байеса (9) при $n=3$ для $k=1, 2, 3$:

$$\mathsf{P}(B_1 \mid A) = \frac{0{,}25 \cdot 0{,}05}{0{,}0345} = \frac{125}{345},$$

$$\mathsf{P}(B_2 \mid A) = \frac{0{,}35 \cdot 0{,}04}{0{,}0345} = \frac{140}{345},$$

$$\mathsf{P}(B_3 \mid A) = \frac{0{,}40 \cdot 0{,}02}{0{,}0345} = \frac{80}{345}.$$

Пример 4. Из урны, содержавшей M белых и $N-M$ черных шаров, один шар неизвестного цвета утерян. Найдем вероятность того, что шар, извлеченный из урны после утери, окажется белым.

Пусть B_k — событие, состоящее в том, что утеряно k белых шаров ($k=0, 1$); A — событие, состоящее в том, что шар, извлеченный из оставшихся шаров, оказался белым. Положим

$$\mathsf{P}(B_0) = \frac{N-M}{N}, \quad \mathsf{P}(B_1) = \frac{M}{N},$$

$$\mathsf{P}(A \mid B_0) = \frac{M}{N-1}, \quad \mathsf{P}(A \mid B_1) = \frac{M-1}{N-1}.$$

По формуле полной вероятности

$$\mathsf{P}(A) = \frac{N-M}{N} \frac{M}{N-1} + \frac{M}{N} \frac{M-1}{N-1} = \frac{M}{N}.$$

Отметим, что вероятность извлечения белого шара из урны до утери тоже равна M/N.

2.3. Независимость событий. В математической модели понятие независимости удобно ввести с помощью понятия условной вероятности. Мы будем говорить, что событие B не зависит от события A, если

$$\mathsf{P}(B \mid A) = \mathsf{P}(B), \quad \mathsf{P}(A) > 0. \tag{10}$$

Если $\mathsf{P}(B) > 0$, то из равенств (10) и (2) следует, что $\mathsf{P}(A \mid B) = \mathsf{P}(A)$. Таким образом, если $\mathsf{P}(A) > 0$ и $\mathsf{P}(B) > 0$, то понятие независимости двух событий симметрично, т. е. если B не зависит от A, то и A не зависит от B. Более удобным определением независимости по сравнению с (10) является следующее определение.

События A и B независимы, если

$$\mathsf{P}(AB)=\mathsf{P}(A)\mathsf{P}(B). \qquad (11)$$

Это определение эквивалентно (10), если $\mathsf{P}(A)>0$. Дадим обобщение определения (11) на несколько событий.

События $A_1, A_2, \ldots, A_n$ будем называть взаимно независимыми (или независимыми в совокупности, или просто независимыми), если для всех комбинаций индексов $1\leqslant i_1<i_2<\ldots$ $\ldots<i_k\leqslant n$ $(k=2, \ldots, n)$ имеем

$$\mathsf{P}(A_{i_1}A_{i_2}\ldots A_{i_k})=\mathsf{P}(A_{i_1})\mathsf{P}(A_{i_2})\ldots\mathsf{P}(A_{i_k}). \qquad (12)$$

Если (12) выполнено только при $k=2$, то события называют *попарно независимыми.* Отметим, что из попарной независимости не следует взаимная независимость (см. задачу 4).

Рассмотрим теперь связь теоретико-вероятностной независимости, введенной равенствами (11) и (12), с причинной независимостью реальных событий. Пусть при n наблюдениях события A, B и AB произошли соответственно n_A, n_B и n_{AB} раз. Из свойства устойчивости частот следуют приближенные равенства:

$$\frac{n_A}{n}\approx\mathsf{P}(A), \quad \frac{n_B}{n}\approx\mathsf{P}(B), \quad \frac{n_{AB}}{n}\approx\mathsf{P}(AB),$$

$$\frac{n_{AB}}{n_A}\approx\mathsf{P}(A\mid B)=\frac{\mathsf{P}(AB)}{\mathsf{P}(B)}.$$

Следовательно, для событий A и B независимых в теоретико-вероятностном смысле, из равенства $\mathsf{P}(A\mid B)=\mathsf{P}(A)$ следует ожидать выполнения приближенного равенства $n_{AB}/n_B\approx n_A/n$ или, что эквивалентно,

$$\frac{n_{AB}}{n}\approx\frac{n_A}{n}\,\frac{n_B}{n}.$$

Это свойство для причинно-независимых событий A и B установлено длительной практикой.

При построении математической модели часто используют следующий принцип: *причинно-независи-*

мые события независимы в теоретико-вероятностном смысле. Отметим, что этот принцип не может являться теоремой, так как он сформулирован не в терминах математической модели. Равенства (11) и (12) используются обычно не для установления независимости, а для определения вероятностей произведений событий при построении математической модели опытов, исходами которых являются причинно-независимые события. Такие построения проводятся в § 3.

2.4. Применение формулы полной вероятности. Рассмотрим пример, который дает некоторое представление об использовании формулы полной вероятности при исследовании случайных процессов с непрерывным временем.

Пример 5. Вероятность поступления на телефонную линию одного вызова за время $(t, t+h)$ равна $\alpha h + o(h)$, $h \to 0$; вероятность того, что ни один вызов за время $(t, t+h)$ не поступит, равна $1-\alpha h + o(h)$. Если линия занята, то вызов теряется. Если в момент t еще продолжается разговор, то за время $(t, t+h)$ он окончится с вероятностью $\beta h + o(h)$, $h \to 0$. Вызовы поступают независимо друг от друга. Найдем вероятность $P_0(t)$ того, что линия в момент t свободна, и вероятность $P_1(t)$ того, что линия занята.

Предположим сразу, что подходящее вероятностное пространство существует и что для интересующих нас событий

$$B_t^{(0)} = \{\text{линия в момент } t \text{ свободна}\},$$

$$B_t^{(1)} = \{\text{линия в момент } t \text{ занята}\}$$

при каждом t вероятности $P_0(t) = \mathsf{P}(B_t^{(0)})$, $P_1(t) = \mathsf{P}(B_t^{(1)})$ определены и непрерывны по t. Так как

$$B_t^{(0)} + B_t^{(1)} = \Omega, \quad B_t^{(0)} B_t^{(1)} = \varnothing,$$

то по формуле полной вероятности (11)

$$\mathsf{P}(B_{t+h}^{(0)}) = \mathsf{P}(B_t^{(0)})\,\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(0)}) + \mathsf{P}(B_t^{(1)})\,\mathsf{P}(B_{t+h}^{(0)} \mid B_t^{(1)}). \tag{13}$$

Свободная в момент времени t линия останется свободной в момент $t+h$, если за время $(t, t+h)$ вызовов не будет. Так как другие события, при которых линия останется свободной в момент $t+h$, имеют вероятность $o(h)$, то

$$\mathsf{P}\left(B_{t+h}^{(0)} \mid B_t^{(0)}\right) = 1 - \alpha h + o(h).$$

Занятая в момент t линия будет свободной к моменту $t+h$, если закончится разговор и за время $(t, t+h)$ новых вызовов не будет. Вероятность этого события равна

$$(\beta h + o(h))(1 - \alpha h + o(h)) = \beta h + o(h).$$

Эта вероятность вносит основной вклад в $\mathsf{P}\left(B_{t+h}^{(0)} \mid B_t^{(1)}\right)$. Сумма остальных слагаемых равна $o(h)$. Таким образом,

$$\mathsf{P}\left(B_{t+h}^{(0)} \mid B_t^{(1)}\right) = \beta h + o(h).$$

Подставляя найденные условные вероятности в (13) и используя обозначения $P_0(t)$, $P_1(t)$, получим

$$P_0(t+h) = (1 - \alpha h) P_0(t) + \beta h P_1(t) + o(h).$$

Отсюда

$$\frac{P_0(t+h) - P_0(t)}{h} = -\alpha P_0(t) + \beta P_1(t) + o(1).$$

Переходя к пределу при $h \to 0$, получим

$$\frac{dP_0(t)}{dt} = -\alpha P_0(t) + \beta P_1(t). \tag{14}$$

Аналогично найдем уравнение для $P_1(t)$:

$$\frac{dP_1(t)}{dt} = \alpha P_0(t) - \beta P_1(t). \tag{15}$$

Полагая $P_0(0) = 1$, $P_1(0) = 0$, найдем решение системы (14)—(15):

$$\begin{aligned} P_0(t) &= \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}, \\ P_1(t) &= \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}. \end{aligned} \tag{16}$$

При $t \to \infty$ вероятность $P_1(t)$ стремится к величине $\alpha/(\alpha+\beta)$. Естественно, что с ростом интенсивности поступления вызовов α вероятность занятости линии увеличивается.

Задачи

1. Брошено две игральных кости. Какова вероятность того, что выпало две «3», если известно, что сумма выпавших очков делится на три?

2. Из множества чисел $\{1, 2, \ldots, N\}$ по схеме случайного выбора без возвращения выбираются три числа. Найти условную вероятность того, что третье число попадет в интервал, образованный первыми двумя, если известно, что первое число меньше второго.

3. Доказать, что события $\bar{A}$ и $\bar{B}$ независимы, если независимы события A и B.

4. Случайная точка (ξ_1, ξ_2) имеет равномерное распределение в квадрате $\{(x_1, x_2)\colon\ 0 \leqslant x_1 \leqslant 1,\ 0 \leqslant x_2 \leqslant 1\}$. Пусть

$$C_1 = \left\{ \xi_1 \leqslant \frac{1}{2} \right\},\quad C_2 = \left\{ \xi_2 \leqslant \frac{1}{2} \right\},\quad C_3 = \left\{ \left(\xi_1 - \frac{1}{2}\right)\left(\xi_2 - \frac{1}{2}\right) < 0 \right\}.$$

Показать, что события C_1, C_2, C_3 попарно независимы. Являются ли события C_1, C_2, C_3 взаимно независимыми? Зависимы ли события C_1C_2 и C_3?

5. События A_1, A_2, A_3, A_4 взаимно независимы; $\mathsf{P}(A_k) = p_k$, $k = 1, 2, 3, 4$. Найти вероятности событий: 1) $A_1\bar{A}_3A_4$; 2) $A_1 + A_2$; 3) $(A_1 + A_2)(A_3 + A_4)$.

6. Электрическая цепь составлена из элементов A_k, $k = 1, 2, 3$, по схеме, приведенной на рис. 5. При выходе из строя лю-

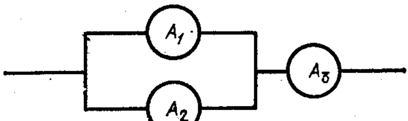


Рис. 5.

бого элемента цепь в месте его включения разрывается. Вероятность выхода из строя за данный период элемента A_k равна p_k, $k = 1, 2, 3$. Предполагается, что элементы выходят или не выходят из строя независимо друг от друга. Найти вероятность того, что за рассматриваемый период по цепи будет проходить ток.

7. Упрощенная система контроля изделий состоит из двух независимых проверок. В результате k-й проверки ($k = 1,2$) изделие, удовлетворяющее стандарту, отбраковывается с вероятностью β_k, а бракованное изделие принимается с вероятностью α_k. Изделие принимается, если оно прошло обе проверки. Найти вероятности событий:

1) бракованное изделие будет принято,

2) изделие, удовлетворяющее стандарту, будет отбраковано.

8. В первой урне 2 белых и 4 черных шара, а во второй — 3 белых и 1 черный шар. Из первой урны переложили во вторую два шара. Найти вероятность того, что шар, вынутый из 2-й урны после перекладывания, окажется белым.

9. Изделия поступают на проверку, описанную в задаче 7. Предполагая, что каждое изделие удовлетворяет стандарту с вероятностью p, найти следующие вероятности:

1) вероятность того, что поступившее на проверку изделие не будет отбраковано;

2) вероятность того, что неотбракованное изделие удовлетворяет стандарту.

10. Предположим, что 5% всех мужчин и 0,25% всех женщин дальтоники. Наугад выбранное лицо оказалось дальтоником. Какова вероятность, что это мужчина? (Считать, что мужчин и женщин одинаковое число.)

11. По каналу связи может быть передана одна из трех последовательностей букв: AAAA, BBBB, CCCC; известно, что вероятности каждой из последовательностей равны соответственно 0,3; 0,4; 0,3. В результате шумов буква принимается правильно с вероятностью 0,6. Вероятности приема переданной буквы за две другие равны 0,2 и 0,2. Предполагается, что буквы искажаются независимо друг от друга. Найти вероятность того, что передано AAAA, если на приемном устройстве получено ABCA.

12. Вероятность того, что молекула, испытавшая в момент $t=0$ столкновение с другой молекулой и не имевшая других столкновений до момента t, испытает столкновение в промежуток времени $(t, t+h)$, равна $\lambda h+o(h)$, $h\to 0$. Найти вероятность того, что время свободного пробега будет больше t.

13. Изменим условия работы телефонной линии, описанной в примере 5. Будем считать, что при занятой линии вызовы не теряются, а становятся в очередь. Обозначим $P_k(t)$ вероятность того, что в момент t один вызов обслуживается и $k-1$ образуют очередь $(k\geqslant 1)$; $P_0(t)$ — вероятность того, что линия свободна. Составить для $P_k(t)$ дифференциальные уравнения. Найти $\lim\limits_{t\to\infty} P_k(t)=\pi_k$, если $\theta=\alpha/\beta<1$.

§ 3. ПОСЛЕДОВАТЕЛЬНОСТЬ ИСПЫТАНИЙ

3.1. Определение последовательности независимых испытаний. Испытанием (наблюдением, экспериментом) обычно называют совокупность определенных условий, при которых могут осуществиться некоторые случайные события. В результате каждого испытания появляется одно из нескольких попарно несовместных событий, которые мы назовем исходами. Например, испытание, состоящее в контроле готового изделия, может окончиться одним из двух исходов: либо изделие окажется годным, либо — дефектным.

Дадим описание теоретико-вероятностной модели последовательности независимых испытаний. Пусть в каждом испытании может наступить один из r исходов, $1, 2, \ldots, r$, и события, связанные с различными испытаниями, — причинно-независимы. Тогда в математической модели, которую мы построим, соответствующие события должны быть независимы в теоретико-вероятностном смысле (см. § 12, (2)). Результат n испытаний можно записать в виде цепочки $x_1, x_2, \ldots, x_n$, где x_t — исход t-го испытания. За множество Ω можно принять множество всех возможных цепочек $\omega = (x_1, x_2, \ldots, x_n)$. Таким образом,

$$\Omega = \{\omega\}, \quad \omega = (x_1, x_2, \ldots, x_n),$$
$$x_t \in \{1, 2, \ldots, r\}, \quad t = 1, 2, \ldots n. \tag{1}$$

Событие $A_i(t) = \{$в t-м испытании наступил исход $i\}$ можно теперь выразить через элементарные события как подмножество Ω:

$$A_i(t) = \{\omega = (x_1, x_2, \ldots, x_n) \colon x_t = i\},$$
$$t = 1, \ldots, n, \quad i = 1, \ldots, r.$$

С другой стороны, элементарное событие $\omega = (x_1, \ldots, x_n)$ представляется как произведение событий $A_i(t)$:

$$\omega = (x_1, \ldots, x_n) = A_{x_1}(1) A_{x_2}(2) \ldots A_{x_n}(n). \tag{2}$$

Элементарные вероятности $p(\omega)$, исходя из формул (2) и (12) из § 2, определим равенством

$$p(\omega) = p_{x_1} p_{x_2} \ldots p_{x_n}, \tag{3}$$

где вероятности исходов отдельных испытаний $p_k = \mathsf{P}(A_k(t))$, $t = 1, \ldots, n$, $k = 1, \ldots, r$, удовлетворяют условиям

$$p_k \geqslant 0, \quad k = 1, \ldots, r; \quad \sum_{k=1}^{r} p_k = 1. \tag{4}$$

Последовательностью независимых испытаний называется конечная вероятностная схема, в которой вероятности элементарных событий определяются формулой (3) *как произведения вероятностей исходов отдельных испытаний.*

Последовательность независимых испытаний мы будем иногда называть *схемой независимых испытаний* или *полиномиальной схемой.*

Часто приходится рассматривать последовательности испытаний с двумя исходами: изделие оказалось годным или дефектным; на лотерейный билет получен выигрыш или нет; прибор за рассматриваемый период времени работал нормально или отказал и т. д.

Частный случай схемы независимых испытаний, в котором каждое испытание может закончиться только одним из двух исходов, называют схемой Бернулли.

Обычно эти исходы называют «успехом» и «неудачей», а их вероятности обозначают p и $q=1-p$ $(0 \leqslant p \leqslant 1)$ соответственно. «Успехи» и «неудачи» для краткости будем обозначать символами $У$ и $Н$, или 1 и 0 соответственно. В схеме Бернулли с n испытаниями имеем

$$\Omega=\{\omega\}, \quad \omega=(x_1, \ldots, x_n), \quad x_t \in \{0;\ 1\},\ t=1, \ldots, n. \tag{5}$$

Очевидно, что число «успехов» (или число «1») в цепочке $\omega=(x_1, x_2, \ldots, x_n)$ равно сумме $x_1+x_2+\ldots+x_n$. Элементарные вероятности, определенные формулой (3), для схемы Бернулли имеют следующий вид:

$$p(\omega)=p^{x_1+x_2+\ldots+x_n}q^{n-(x_1+\ldots+x_n)}, \quad \omega=(x_1, \ldots, x_n). \tag{6}$$

Для схемы Бернулли часто представляет интерес событие

$B_m=\{$в n испытаниях наступило ровно m успехов$\}$.

Теорема 1. *Вероятность* $\mathsf{P}(B_m)$ *того, что в* n *испытаниях схемы Бернулли наступили ровно* m *успехов, определяется формулой*

$$\mathsf{P}(B_m)=C_n^m p^m q^{n-m}, \quad m=0, 1, \ldots, n, \tag{7}$$

где p — *вероятность успеха в отдельном испытании.*

Доказательство. Событие B_m определяется как подмножество Ω:

$$B_m=\{\omega=(x_1, \ldots, x_n):\ x_1+x_2+\ldots+x_n=m\}.$$

Следовательно, $\mathsf{P}(B_m)=\sum\limits_{\omega\in B_m} p(\omega)$. Так как согласно формуле (6) элементарные вероятности $p(\omega)=p^m q^{n-m}$ при любом $\omega\in B_m$, то для вычисления вероятности осталось определить число $|B_m|$ элементарных событий, входящих в B_m. Число $|B_m|$ совпадает, очевидно, с числом способов выбора m мест для «1» в цепочке ω, так как оставшиеся места однозначно заполняются «0». Отсюда $|B_m|=C_n^m$ и $\mathsf{P}(B_m)=C_n^m p^m q^{n-m}$.

Пример 1. Система, составленная из n блоков, работает исправно, если за рассматриваемый период выйдет из строя не более двух блоков. Найдем вероятность безотказной работы системы в предположении, что отказы блоков являются независимыми событиями и вероятность отказа каждого блока равна p.

В качестве модели воспользуемся схемой Бернулли с n испытаниями. Каждое испытание заключается в работе одного из блоков за рассматриваемый период. Для удобства использования формулы (7) назовем «успехом» выход блока из строя. Пусть $A=\{$система работает безотказно$\}$. Тогда $A=B_0+B_1+B_2$, где $B_m=\{$выход из строя m блоков$\}$. Отсюда, используя формулу (7), получим

$$\mathsf{P}(A)=\mathsf{P}(B_0)+\mathsf{P}(B_1)+\mathsf{P}(B_2)=q^n+npq^{n-1}+\frac{n(n-1)}{2}p^2q^{n-2}.$$

Пример 2. Игральная кость подброшена 10 раз. Найдем вероятность $\mathsf{P}(B_3)$ того, что выпало ровно три «6». В формуле (7) нужно положить $n=10$, $m=3$, $p=1/6$, $q=5/6$. Тогда

$$\mathsf{P}(B_3)=C_{10}^3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7=\frac{10\cdot 9\cdot 8\cdot 5^7}{1\cdot 2\cdot 3\cdot 6^{10}}=0{,}155045\ \ldots$$

Получим обобщение формулы (7) на случай полиномиальной схемы.

Теорема 2. *Вероятность $P_n(m_1, \ldots, m_r)$ того, что в испытаниях полиномиальной схемы исход «1»* *наступил m_1 раз, исход «2» — m_2 раз, ..., исход «r» — m_r раз, определяется равенством*

$$P_n(m_1, \ldots, m_r) = \frac{n!}{m_1! m_2! \ldots m_r!} p_1^{m_1} p_2^{m_2} \ldots p_r^{m_r}, \tag{8}$$

где p_k — вероятность исхода k в отдельном испытании ($k = 1, \ldots, r$); $m_1, \ldots, m_r$ — целые неотрицательные числа, удовлетворяющие равенству $m_1 + m_2 + \ldots + m_r = n$.

Доказательство. Пусть B = {среди n испытаний исход «1» появился m_1 раз, ..., исход «r» появился m_r раз}. Для любого элементарного события $\omega \in B$ имеем $p(\omega) = p_1^{m_1} p_2^{m_2} \ldots p_r^{m_r}$. Число $|B|$ элементарных событий, входящих в B, нетрудно подсчитать. Исход «1» на m_1 местах цепочки $\omega = (x_1, \ldots, x_r)$ можно расположить $C_n^{m_1}$ способами; исход «2» на оставшихся $n - m_1$ местах можно расположить $C_{n-m_1}^{m_2}$ способами и т. д. Таким образом,

$$|B| = C_n^{m_1} C_{n-m_1}^{m_2} \ldots C_{n-m_1-\ldots-m_{r-1}}^{m_r} = \frac{n!}{m_1! m_2! \ldots m_r!}.$$

Следовательно,

$$\mathbf{P}(B) = \sum_{\omega \in B} p(\omega) = |B| p_1^{m_1} \ldots p_r^{m_r} = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r}.$$

Теорема доказана.

Формула (7) получается из формулы (8), если положить $r = 2$, $p_1 = p$, $p_2 = 1 - p = q$.

Рассмотрим несколько важных частных случаев схемы независимых испытаний.

Случайные числа. Рассмотрим последовательность независимых испытаний с $r = 10$ исходами, которые мы обозначим 1, 2, ..., 9, 0. Вероятности этих исходов в отдельных испытаниях обозначим соответственно $p_1, p_2, \ldots, p_{10}$.

Если положить $p_1 = p_2 = \ldots = p_{10} = 1/10$, то полученная вероятностная схема совпадет с вероятностной схемой, введенной в § 1 для определения случайных чисел на основе классического определения вероятности.

Пример 3. Найдем вероятность $P_{20}(10, 2, 3, 5)$ того, что среди 20 случайных чисел имеется ровно

10 четных цифр, две «тройки» и три «семерки». Для вычисления указанной вероятности конечную схему, определяющую случайные числа, удобно свести к последовательности независимых испытаний, в каждом из которых возможно появление одного из четырех исходов: 1) четная цифра, 2) тройка, 3) семерка, 4) все остальное. Вероятности этих исходов равны соответственно $p_1 = 5/10$, $p_2 = p_3 = 1/10$, $p_4 = 3/10$. По формуле (8) получим

$$P_{20}(10, 2, 3, 5) = \frac{20!}{10!2!3!5!}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{10}\right)^{5}\left(\frac{3}{10}\right)^{5}.$$

В последовательности независимых испытаний иногда n испытаний интерпретируют как размещение n частиц по $r = N$ ячейкам; вероятность p_i $(i = 1, \ldots, N)$ является вероятностью попадания частицы в i-ю ячейку. В случае $p_1 = p_2 = \ldots = p_N = 1/N$ схему размещения частиц по ячейкам называют *равновероятной*. (Иногда говорят о размещении «дробинок по ящикам» вместо — «частиц по ячейкам».)

К задачам о размещении относятся также различные задачи о «днях рождения».

Пример 4. В группе 25 студентов. Вычислим вероятность того, что найдутся хотя бы два студента с общим днем рождения. Будем предполагать, что дни рождения независимы и попадание дня рождения на определенный день года равна 1/365. Пусть событие A = {найдутся хотя бы два студента с общим днем рождения}. Тогда $\bar{A}$ = {все дни рождения приходятся на разные дни года}. Элементарными событиями ω в данной задаче являются цепочки $\omega = (x_1, x_2, \ldots, x_{25})$, где x_1 — день года, являющийся днем рождения 1-го студента; x_2 — второго и т. д. Тогда для любого ω по определению схемы независимых испытаний $p(\omega) = (1/365)^{25}$, а $\mathsf{P}(\bar{A}) = |\bar{A}|/365^{25}$, где $|\bar{A}|$ — число элементарных событий, входящих в $\bar{A}$. Так как для ω, входящих в $\bar{A}$, среди $x_1, \ldots, x_{25}$ нет совпадающих, то значение x_1 можно выбрать 365 способами; значение x_2 — 364 способами (любой день кроме выбранного для рождения 1-го студента); значение x_3 — выбирается 363 способами и т. д. В результате находим $|\bar{A}| = 365 \cdot 364 \cdot \ldots \cdot 341$. Отсюда

$\mathsf{P}(\bar{A}) = (365 \cdot 364 \cdot \ldots \cdot 341)/365^{25} = 0{,}43130\ldots$ и, следовательно,

$$\mathsf{P}(A) = 1 - \mathsf{P}(\bar{A}) \approx 0{,}56869\ldots$$

3.2. Общее определение последовательности испытаний. Общее описание теоретико-вероятностной модели последовательности испытаний (не обязательно независимых) дадим в терминах изменения состояний некоторой «частицы» (или физической системы). Пусть частица может находиться в одном из r состояний. Рассмотрим n последовательных состояний частицы. За множество $\Omega = \{\omega\}$ примем множество всех возможных «траекторий» частицы $\omega = (x_1, x_2, \ldots, x_n)$, где x_t — состояние частицы в момент t, $t = 1, 2, \ldots, n$. Таким образом, имеем

$$\Omega = \{\omega\}, \quad \omega = (x_1, \ldots, x_n), \quad x_t \in \{1, 2, \ldots, r\}, \quad t = 1, \ldots, n. \tag{9}$$

Если известны условные вероятности $P_{t+1}(x_{t+1} \mid x_1, \ldots, x_t)$ того, что частица в момент $t+1$ находилась в состоянии x_{t+1}, при условии, что в моменты $1, 2, \ldots, t$ ее состояниями являлись $x_1, x_2, \ldots, x_t$, то согласно формуле (6) § 2 вероятность элементарного события $\omega = (x_1, \ldots, x_n)$ определяется однозначно:

$$p(\omega) = p_1(x_1)\, p_2(x_2 \mid x_1)\, p_3(x_3 \mid x_1, x_2) \ldots \\ \ldots p_n(x_n \mid x_1, \ldots, x_{n-1}). \tag{10}$$

Здесь $p_1(x_1)$ — вероятность того, что частица в момент $t = 1$ находится в состоянии x_1.

Конечная вероятностная схема, в которой множество Ω определяется равенством (9), а элементарные вероятности формулой (10), называется последовательностью n испытаний с r исходами.

Описание множества Ω для схемы последовательных испытаний в виде множества траекторий $\omega = (x_1, \ldots, x_n)$ некоторой частицы удобно, так как обладает известной наглядностью. Это образное описание можно применять к самым различным схемам, не имеющим никакого отношения к «частицам» и их «состояниям». Например, в терминах условных вероятностей естественно определяется схема случайного выбора без возвращения n элементов из N элементов (см. § 1, пример 5). В этом случае элементарные события $\omega=(x_1, x_2, \ldots, x_n)$ определяются как размещения из N элементов по n, а условные вероятности $p_t(x_t|x_1, \ldots, x_{t-1})$ определяются следующим образом: если извлечено t шаров, то при следующем испытании любой из оставшихся $N-t$ шаров извлекается с вероятностью $1/(N-t)$. Таким образом, в этом случае $p_1(x_1)=\frac{1}{N}$, $p_t(x_t|x_1, \ldots, x_{t-1})=\frac{1}{N-t}$ и определение элементарных вероятностей

$$p(\omega)=$$
$$=p_1(x_1)p_2(x_2|x_1)p_3(x_3|x_1, x_2) \ldots p_n(x_n|x_1, \ldots, x_{n-1})=$$
$$=\frac{1}{N}\frac{1}{N-1}\ldots\frac{1}{N-n+1}=\frac{1}{N^{[n]}}$$

совпадает с тем, которое дано в § 1 (пример 5).

Пример 5. Среди n ключей к двери подходит только один. По схеме случайного выбора без возвращения ключи извлекаются до тех пор, пока не появится нужный ключ. Найдем вероятности событий $A_k=$ {нужный ключ впервые появился при k-м испытании}.

Воспользуемся схемой случайного выбора без возвращения, определенной в терминах условных вероятностей. Так как $A_k\subset\bar{A}_i$ при $i<k$, то событие A_k можно представить в виде $A_k=\bar{A}_1\bar{A}_2\ldots\bar{A}_{k-1}A_k$. Тогда

$$\mathsf{P}(A_k)=\mathsf{P}(\bar{A}_1\ldots\bar{A}_{k-1}A_k)=\mathsf{P}(\bar{A}_1)\mathsf{P}(\bar{A}_2|\bar{A}_1)\mathsf{P}(\bar{A}_3|\bar{A}_1\bar{A}_2)\ldots$$
$$\ldots\mathsf{P}(\bar{A}_{k-1}|\bar{A}_1\ldots\bar{A}_{k-2})\mathsf{P}(A_k|\bar{A}_1\bar{A}_2\ldots\bar{A}_{k-1}).$$

Входящие в эту формулу условные вероятности нетрудно найти, так как по условию задачи вероятности появления любого из оставшихся к данному испытанию шаров одинаковы. Таким образом,

$$\mathsf{P}(\bar{A}_1)=\frac{n-1}{n},\quad \mathsf{P}(\bar{A}_2|\bar{A}_1)=\frac{n-2}{n-1},\ \ldots$$
$$\ldots,\ \mathsf{P}(\bar{A}_{k-1}|\bar{A}_1\ldots\bar{A}_{k-2})=\frac{n-k+1}{n-k+2},$$
$$\mathsf{P}(A_k|\bar{A}_1\ldots\bar{A}_{k-1})=\frac{1}{n-k+1}$$

и, следовательно,

$$\mathsf{P}(A_k)=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\frac{n-3}{n-2}\cdot\ldots\cdot\frac{n-k+1}{n-k+2}\cdot\frac{1}{n-k+1}=\frac{1}{n}.$$

Общая схема испытаний в качестве частного случая содержит схему независимых испытаний, в которой

$$p_t(x_t\,|\,x_1, x_2, \ldots, x_{t-1})=p_t(x_t), \qquad t=1, \ldots, n. \tag{11}$$

Эту схему независимых испытаний естественно назвать *неоднородной* в отличие от схемы независимых испытаний, введенной в п. 3.1, так как вероятности исходов в п. 3.1 не зависели от номера испытания t.

Отметим в заключение другой важный частный случай общей схемы последовательности испытаний. Если условные вероятности $p_t(x_t\,|\,x_1, \ldots, x_{t-1})$ зависят только от результата последнего испытания, т. е.

$$p_t(x_t\,|\,x_1, \ldots, x_{t-1})=p_t(x_t\,|\,x_{t-1}),$$

то последовательность испытаний называют *цепью Маркова*.

Задачи

1. Два игрока поочередно извлекают шары (без возвращения) из урны, содержащей 2 белых и 4 черных шара. Выигрывает тот, кто первым вынет белый шар. Найти вероятности выигрыша участников.

2. При передаче сообщения вероятность искажения одного знака равна 1/100. В предположении независимости искажения знаков найти вероятность того, что сообщение из 5 знаков: а) не будет искажено; б) содержит ровно одно искажение; в) содержит хотя бы два искажения.

3. В некоторой лотерее вероятность выигрыша на один билет равна 1/5. Предполагая, что выигрыши на различные билеты независимы, определить число билетов, которые нужно купить, чтобы вероятность $Q(n)$ получения хотя бы одного выигрыша была не меньше 0,9. Найти $Q(n)$ для этого n.

4. Каждое из 10 независимых испытаний заключается в подбрасывании трех игральных костей. Найти вероятность того, что в четырех испытаниях появятся в точности по две «6».

§ 4. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ В СХЕМЕ БЕРНУЛЛИ

4.1. Теорема Пуассона. В приложениях часто приходится вычислять вероятности различных событий, связанных с числом успехов μ_n в n испытаниях

Бернулли при больших значениях n. В этом случае вычисления по формуле (7) § 3 становятся затруднительными. Трудности возникают также и в том случае, когда приходится суммировать вероятности событий $\{\mu_n = m\}$. К суммированию сводится вычисление вероятностей событий $\{a < \mu_n < b\}$. Действительно,

$$\mathsf{P}\{a < \mu_n < b\} = \sum_{a<m<b} C_n^m p^m q^{n-m}. \tag{1}$$

Затруднения при вычислениях возникают также при малых значениях p или q.

Иногда при больших n удается заменить формулу (7) § 3 какой-либо приближенной асимптотической формулой. Приведем три предельные теоремы, содержащие асимптотические формулы для вероятностей $\mathsf{P}\{\mu_n = m\}$ и $\mathsf{P}\{a < \mu_n < b\}$ при $n \to \infty$.

Теорема 1 (теорема Пуассона). *Если $n \to \infty$ и $p \to 0$ так, что $np \to \lambda$, $0 < \lambda < \infty$, то*

$$\mathsf{P}\{\mu_n = m\} = P_n(m) = C_n^m p^m q^{n-m} \to p_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda}$$

при любом постоянном m, $m = 0, 1, 2, \ldots$

Доказательство. Положив $np = \lambda_n$, представим вероятность $P_n(m)$ в виде

$$\mathsf{P}\{\mu_n = m\} = \frac{n(n-1)\ldots(n-m+1)}{m!} \left(\frac{\lambda_n}{n}\right)^m \left(1 - \frac{\lambda_n}{n}\right)^{n-m} =$$
$$= \frac{\lambda_n^m}{m!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots$$
$$\ldots \left(1 - \frac{m-1}{n}\right)\left(1 - \frac{\lambda_n}{n}\right)^{-m}.$$

Отсюда при $n \to \infty$ получим утверждение теоремы.

Таким образом, при больших n и малых p мы можем воспользоваться приближенной формулой

$$\mathsf{P}\{\mu_n = m\} \approx \frac{\lambda_n^m}{m!} e^{-\lambda_n}, \quad \lambda_n = np. \tag{2}$$

Для сравнения точных и приближенных значений приведем следующую таблицу:

m	0	1	2	3	4	5
$p_m(np)$	0,1353	0,2707	0,2707	0,1805	0,0902	0,0361
$P_{10}(m)$	0,1074	0,2684	0,3020	0,2013	0,0880	0,0264
$P_{100}(m)$	0,1326	0,2707	0,2734	0,1823	0,0902	0,0353

При $n=10$ вероятности $P_n(m)$ вычислены для $p=1/5$, а при $n=100$ — для $p=1/50$.

Если мало значение q, то пуассоновским приближением можно воспользоваться для числа неудач.

Пример 1. В таблице случайных чисел цифры сгруппированы по две. Найдем приближенное значение вероятности того, что среди 100 пар пара 00 встретится не более двух раз.

Воспользуемся схемой Бернулли с $n=100$ испытаниями. Если появление пары 00 назвать «успехом», то нужно положить $p=0{,}01$, $q=0{,}99$. Точное значение вероятности события $\{\mu_{100}\leqslant 2\}$, где μ_{100} — число пар 00, находится по формуле (7) § 2

$$\mathsf{P}\{\mu_{100}\leqslant 2\}=\mathsf{P}\{\mu_{100}=0\}+\mathsf{P}\{\mu_{100}=1\}+\mathsf{P}\{\mu_{100}=2\}=$$
$$=0{,}99^{100}+100\cdot 0{,}01\cdot 0{,}99^{99}+\frac{100\cdot 99}{2}\cdot 0{,}01^2\cdot 0{,}99^{98}.$$

В этом случае можно воспользоваться теоремой 1, формулой (2) с $\lambda_n=0{,}01\cdot 100=1$ и таблицей 3. Получим

$$\mathsf{P}\{\mu_{100}\leqslant 2\}\approx 0{,}3679+0{,}3679+0{,}1839=0{,}9197.$$

4.2. Теоремы Муавра — Лапласа.

Теорема 2 (локальная теорема Муавра — Лапласа). *Пусть* p $(0<p<1)$ *постоянно*, $x_m=(m-np)/\sqrt{npq}$, $C>0$ — *любая постоянная. Если* $|x_m|\leqslant C$, *то при* $n\to\infty$

$$\mathsf{P}\{\mu_n=m\}=C_n^m p^m q^{n-m}=$$
$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{x_m^2}{2}}\Delta x_m(1+\varepsilon_{n,m}),\qquad(3)$$

где $\Delta x_m = 1/\sqrt{npq}$, $|\varepsilon_{n,m}| < C_1/\sqrt{n}$, $C_1 > 0$ — *некоторая постоянная.*

Таким образом, в качестве приближенных значений вероятностей $P_n(m) = P\{\mu_n = m\}$ можно взять величины

$$\varphi\left(\frac{m-np}{\sqrt{npq}}\right)\frac{1}{\sqrt{npq}},$$

где

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

На рис. 6 приведены графики функций

$$y = \varphi\left(\frac{u-np}{\sqrt{npq}}\right)\frac{1}{\sqrt{npq}}$$

и в точках $u = m$, $m = 0, 1, 2, \ldots$ отложены вертикальные отрезки, длины которых равны величинам $P_n(m)$:

а) Значения $P_{10}(m)$ при $m = 0$; 1; 2; 3; 4; 5; 6 равны 0,1074; 0,2684; 0,3020; 0,2013; 0,0881; 0,0264; 0,0055; соответствующие приближенные значения равны 0,0905; 0,2308; 0,3153; 0,2308; 0,0905; 0,0191; 0,0021.

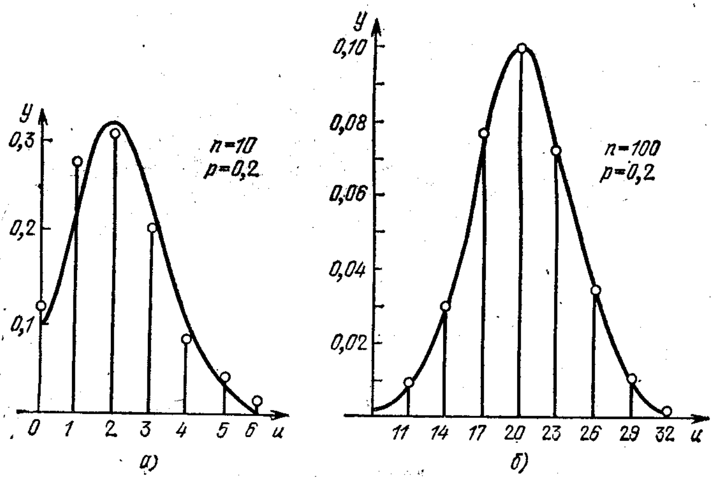


Рис. 6.

б) Значения $P_{100}(m)$ при $m = 11; 14; 17; 20; 23; 26; 29; 32$ равны 0,0069; 0,0335; 0,0789; 0,0993; 0,0720; 0,0316; 0,0088; 0,0016; соответствующие приближенные значения равны 0,0079; 0,0324; 0,0742; 0,0997; 0,0742; 0,0324; 0,0079; 0,0011.

Доказательство локальной теоремы приведено в приложении.

Отметим, что в условиях теоремы 2 из того, что $n \to \infty$, следует стремление к бесконечности m.

Теорема 3 (интегральная теорема Муавра — Лапласа). *Если вероятность успеха в каждом испытании p, $0 < p < 1$, постоянна, то при $n \to \infty$*

$$\mathsf{P}\left\{a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right\} - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx \to 0 \tag{4}$$

равномерно по a, b, $-\infty \leqslant a \leqslant b \leqslant +\infty$.

Доказательство. Приведем доказательство этой теоремы при фиксированных a, b, $-\infty < a \leqslant \leqslant b < +\infty$ *). Очевидно, что вероятность события $\left\{a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right\}$ можно представить в виде

$$\mathsf{P}\left\{a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right\} = P_n(a, b) = \sum_{a \leqslant x_m \leqslant b} P_n(m), \tag{5}$$

где $x_m = \frac{m - np}{\sqrt{npq}}$ и суммирование ведется по тем значениям m, для которых $x_m \in [a, b]$. Применяя к слагаемым (5) локальную теорему, получим

$$P_n(a, b) = S_n + T_n, \tag{6}$$

где

$$S_n = \sum_{x_m \in [a,\, b]} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_m^2}{2}} \Delta x_m,$$

$$T_n = \sum_{x_m \in [a,\, b]} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_m^2}{2}} \Delta x_m \varepsilon_{n,\, m}.$$

---

*) Доказательство равномерной сходимости по a и b дано, например, в [4].

Заметим, что S_n отличается не более чем двумя слагаемыми от подходящим образом выбранной интегральной суммы, соответствующей интегралу $\frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}\,dx$. Следовательно,

$$\lim_{n\to\infty} S_n = \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}\,dx. \tag{7}$$

Для T_n имеем

$$|T_n| = \sum_{x_m\in[a,\,b]} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_m^2}{2}}\Delta x_m \varepsilon_{n,\,m} \leqslant$$

$$\leqslant \sum_{x_m\in[a,\,b]} \Delta x_m \varepsilon_{n,\,m} = O\left(\frac{1}{\sqrt{n}}\right).$$

Таким образом, при $n\to\infty$

$$T_n \to 0. \tag{8}$$

Утверждение теоремы при постоянных a, b, $-\infty < a \leqslant b < +\infty$, следует из формул (5)—(8).

Приближенная формула

$$\mathsf{P}\left\{a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right\} \approx \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}\,dx = \Phi(b) - \Phi(a), \tag{9}$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt, \tag{10}$$

дает хорошее приближение, когда n достаточно велико, а p и q не очень близки к нулю. Часто нормальным приближением пользуются при $npq \geqslant 20$.

Численное значение интеграла (10) можно найти, воспользовавшись таблицами для функции

$$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}\,dt. \tag{11}$$

При небольших значениях npq приближенную формулу (9) нужно заменить следующей формулой (см. [13] гл. 7, § 2 стр. 189):

$$\mathsf{P}\left\{a\leqslant\frac{\mu_n-np}{\sqrt{npq}}\leqslant b\right\}\approx\Phi\left(b+\frac{1}{2\sqrt{npq}}\right)-\Phi\left(a-\frac{1}{2\sqrt{npq}}\right). \tag{12}$$

Приведем таблички точных и приближенных значений вероятностей событий $\{m_1\leqslant\mu_n\leqslant m_2\}$. Положим

$$P(m_1, m_2)=\mathsf{P}\{m_1\leqslant\mu_n\leqslant m_2\}=\mathsf{P}\left\{x_1\leqslant\frac{\mu_n-np}{\sqrt{npq}}\leqslant x_2\right\},$$

$$P_0(m_1, m_2)=\Phi(x_2)-\Phi(x_1),$$

$$P_1(m_1, m_2)=\Phi\left(x_2+\frac{h}{2}\right)-\Phi\left(x_1-\frac{h}{2}\right),$$

где $h=1/\sqrt{npq}$, $x_1=(m_1-np)h$, $x_2=(m_2-np)h$.

а) $n=10$

p	$m_1;\ m_2$	0; 2	3; 5	6; 8	9; 10
0,2	$P(m_1, m_2)$	0,6778	0,3158	0,0064	0,0000
	$P_0(m_1, m_2)$	0,4429	0,2059	0,0007	0,0000
	$P_1(m_1, m_2)$	0,6316	0,3455	0,0029	0,0000
0,5	$P(m_1, m_2)$	0,0547	0,5684	0,3662	0,0107
	$P_0(m_1, m_2)$	0,0277	0,3980	0,2356	0,0057
	$P_1(m_1, m_2)$	0,0558	0,5696	0,3651	0,0133

б) $n = 100,\ p = 0{,}2$

$m_1;\ m_2$	0; 9	10; 15	16; 20	21; 25	26; 30	31; 40
$P(m_1,\ m_2)$	0,0023	0,1262	0,4310	0,3531	0,0814	0,0061
$P_0(m_1,\ m_2)$	0,0030	0,0994	0,3413	0,2957	0,0606	0,0030
$P_1(m_1,\ m_2)$	0,0041	0,1250	0,4223	0,3480	0,0795	0,0043

в) $n = 100,\ p = 0{,}5$

$m_1;\ m_2$	30; 35	36; 40	41; 45	46; 50
$P(m_1,\ m_2)$	0,0017	0,0267	0,1557	0,3557
$P_0(m_1,\ m_2)$	0,0013	0,0202	0,1227	0,2881
$P_1(m_1,\ m_2)$	0,0018	0,0269	0,1563	0,3558

Формула (9) позволяет также оценить близость частоты и вероятности. Пусть p — вероятность успеха в схеме Бернулли и μ_n — общее число успехов. Частотой успеха называют отношение μ_n/n. Оценим вероятность события $\left\{\left|\frac{\mu_n}{n} - p\right| < \Delta\right\}$. Если n достаточно велико, то можно воспользоваться формулой (9). Тогда

$$\mathsf{P}\left\{\left|\frac{\mu_n}{n} - p\right| < \Delta\right\} = \mathsf{P}\left\{-\Delta\sqrt{\frac{n}{pq}} < \frac{\mu_n - np}{\sqrt{npq}} < \Delta\sqrt{\frac{n}{pq}}\right\} \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-\Delta\sqrt{n/pq}}^{\Delta\sqrt{n/pq}} e^{-\frac{x^2}{2}}\,dx = 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right), \tag{13}$$

так как функция $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ — четная. Значение $\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right)$ находится по таблице 1 в конце книги.

Пример 2. Две монеты подбрасываются $n = 1000$ раз. Пусть μ_n — число выпадений комбинации «герб — герб». Найдем приближенное значение вероятности того, что число комбинаций «герб — герб» заключено между 236 и 264, т. е. оценим вероятность $\mathsf{P}\{236 < \mu_n < 264\}$.

В качестве математической модели воспользуемся схемой Бернулли. Появление комбинации «герб — герб» назовем «успехом». Тогда $p = 1/4$, $q = 3/4$. Применим интегральную теорему Муавра — Лапласа. Используя значения $np = 250$, $\sqrt{npq} \approx 13{,}7$, запишем событие $\{236 < \mu_n < 264\}$ в форме, используемой в теореме 3:

$$\{236 < \mu_n < 264\} = \left\{ \frac{236 - 250}{13{,}7} < \frac{\mu_n - np}{\sqrt{npq}} < \frac{264 - 250}{13{,}7} \right\} = \left\{ \left| \frac{\mu_n - np}{\sqrt{npq}} \right| < 1{,}02 \right\}.$$

По формуле (9) находим

$$\mathsf{P}\{236 < \mu_n < 264\} = \mathsf{P}\left\{ -1{,}02 < \frac{\mu_n - np}{\sqrt{npq}} < 1{,}02 \right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-1{,}02}^{1{,}02} e^{-\frac{x^2}{2}}\, dx = 2\Phi_0(1{,}02) = 2 \cdot 0{,}3461 = 0{,}6922.$$

Значение Φ_0 (1,02) найдено по таблице 1.

Часто возникает обратная задача: сколько нужно провести испытаний, чтобы частота μ_n/n отличалась от вероятности p не больше, чем на Δ, с вероятностью $1 - 2\alpha$ (α — мало)? В таких задачах естественно считать p неизвестным. Тогда, чтобы подобрать наименьшее n, при котором вероятность отклонения будет равна $1 - 2\alpha$, нужно согласно (13) решить уравнение

$$2\Phi_0\left(\Delta \sqrt{\frac{n}{pq}}\right) = 1 - 2\alpha.$$

Решение будет зависеть от неизвестного p. От этой зависимости можно избавиться, если потребовать, чтобы

$$\mathsf{P}\left\{ \left| \frac{\mu_n}{n} - p \right| < \Delta \right\} \geqslant 1 - 2\alpha.$$

Тогда из (13), используя неравенство $pq \leqslant 1/4$, получим

$$\mathsf{P}\left\{\left|\frac{\mu_n}{n} - p\right| < \Delta\right\} \approx 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right) \geqslant$$

$$\geqslant 2\Phi_0\left(2\Delta\sqrt{n}\right) = 1 - 2\alpha$$

и для определения n имеем уравнение $\Phi_0\left(2\Delta\sqrt{n}\right) = \frac{1-2\alpha}{2}$. По таблице можно найти u_α, для которых $\Phi_0(u_\alpha) = \frac{1-2\alpha}{2}$. Тогда $2\Delta\sqrt{n} = u_\alpha$ и $n \geqslant u_\alpha^2/4\Delta^2$. Довольно часто используются значения 2α, равные 0,05 и 0,01. Для этих значений соответствующие u_α равны 1,960 и 2,576.

Задачи

1. Найти приближеннное выражение вероятности того, что число выпадений «1» при 12 000 бросаний правильной игральной кости заключено между 1900 и 2150.

2. Полагая вероятность рождения мальчика равной 0,515, найти вероятность того, что среди 10 000 новорожденных мальчиков будет не больше, чем девочек.

3. Книга в 500 страниц содержит 50 опечаток. Используя схему Бернулли в качестве математической модели распределения опечаток по страницам, оценить вероятность того, что на определенной странице будет не менее трех опечаток. Вычислить точные значения вероятности, а также пуассоновское приближение. Сравнить результаты.

4. Вероятность выпуска сверла повышенной хрупкости (брак) равна 0,02. Сверла укладываются в коробки по 100 штук. Пользуясь законом Пуассона, определить вероятность того, что
а) в коробке не окажется бракованных сверл;
б) число бракованных сверл не превосходит двух.

5. Сколько нужно взять случайных цифр, чтобы среди них с вероятностью, не меньшей 0,9, цифра «6» появилась хотя бы один раз?

6. Сколько раз нужно бросить монету, чтобы частота выпадения герба отличалась от 1/2 не более чем на Δ с вероятностью, не меньшей $1-2\alpha$ $(2\alpha = 0{,}05,\ \Delta = 0{,}1)$?

§ 5. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

5.1. Случайные величины в конечной схеме. Рассмотрим конечное вероятностное пространство $(\Omega, \mathscr{A}, \mathsf{P})$.

Числовую функцию $\xi=\xi(\omega)$ от элементарного события $\omega\in\Omega$ назовем случайной величиной.

Будем обозначать случайные величины греческими буквами ξ, η, ζ, μ, ν, ..., иногда случайные величины обозначаются прописными латинскими буквами X, Y, Z, ...

В модели конечного вероятностного пространства каждое элементарное событие $\omega\in\Omega$ сопоставляется случайному исходу некоторого случайного явления, а случайная величина $\xi=\xi(\omega)$ приписывает этим случайным исходам числовые значения. Таким образом, случайная величина в зависимости от случая принимает разные числовые значения.

Пример 1. При бросании игральной кости возможны шесть элементарных исходов. Обозначим элементарное событие ω_i, если на кости выпало i очков. Тогда $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, а случайная величина $\xi(\omega_i)=i$ равна числу очков, выпавших на кости при бросании.

Пример 2. В схеме независимых испытаний Бернулли (см. § 3) множество Ω состоит из элементарных событий $\omega=(x_1, x_2, \ldots, x_n)$, где $x_i=1$, если при i-м испытании произошел успех, и $x_i=0$ в случае неуспеха. Случайная величина $\mu=\mu(\omega)=x_1+x_2+\ldots+x_n$ равна числу успехов при n испытаниях в схеме Бернулли.

Любую константу C можно рассматривать как частный случай случайной величины $\xi=\xi(\omega)\equiv C$. Такие случайные величины будем называть *вырожденными*. Простейшими случайными величинами, отличными от вырожденных, являются индикаторы. С каждым событием $A\in\mathscr{A}$ можно связать случайную величину

$$\chi_A=\chi_A(\omega)=\begin{cases}1, & \text{если } \omega\in A,\\ 0, & \text{если } \omega\notin A,\end{cases}$$

называемую *индикатором* события A.

Пример 3. В n испытаниях схемы Бернулли (см. пример 2) определим события $A_i=\{$в i-м испытании наступил «успех»$\}$, $i=1, 2, \ldots, n$. Случайную величину μ_n, равную числу успехов, можно представить в виде суммы индикаторов:

$$\mu_n = \chi_{A_1} + \chi_{A_2} + \ldots + \chi_{A_n}. \tag{1}$$

Действительно, $\mu_n(x_1, \ldots, x_n) = x_1 + x_2 + \ldots + x_n$, а, с другой стороны, $\chi_{A_i}(x_1, \ldots, x_n) = x_i$. Представление (1) часто используется при исследовании величины μ_n.

Индикаторы удовлетворяют следующим легко проверяемым свойствам:

$$\chi_\varnothing \equiv 0, \quad \chi_\Omega \equiv 1, \quad \chi_{AB} = \chi_A \chi_B, \quad \chi_{\bar{A}} = 1 - \chi_A.$$

Если события A и B несовместны, то

$$\chi_{A+B} = \chi_A + \chi_B.$$

В самом деле, $\chi_{A+B}(\omega) = 1$ равносильно условию $\omega \in A + B$. Это значит, что либо $\omega \in A$ и $\omega \notin B$, либо $\omega \notin A$ и $\omega \in B$, т. е. либо $\chi_A = 1$ и $\chi_B = 0$, либо $\chi_A = 0$ и $\chi_B = 1$.

Нетрудно по индукции установить, что

$$\chi_{A_1 A_2 \ldots A_n} = \chi_{A_1} \chi_{A_2} \cdots \chi_{A_n}$$

для любых событий $A_1, \ldots, A_n$, и

$$\chi_{A_1 + \ldots + A_n} = \chi_{A_1} + \ldots + \chi_{A_n}$$

для попарно несовместных событий $A_1, A_2, \ldots, A_n$.

Пусть $x_1 < x_2 < \ldots < x_n$ — всевозможные значения случайной величины ξ. Обозначим событие $A_i = \{\omega: \xi(\omega) = x_i\}$, состоящее из всех тех элементарных событий ω, для которых $\xi(\omega) = x_i$. Так как x_i все различны, то события A_i и A_j при $i \neq j$ несовместны. Поскольку $x_1, x_2, \ldots, x_k$ исчерпывают все возможные значения ξ, то $\sum_{i=1}^{k} A_i = \Omega$. Таким образом, со случайной величиной ξ можно связать разбиение α_ξ, состоящее из множеств

$$A_i = \{\omega: \xi(\omega) = x_i\}, \qquad i = 1, \ldots, k, \tag{2}$$

на которых $\xi(\omega)$ постоянна. Разбиение α_ξ мы будем называть *порожденным случайной величиной* ξ.

Случайную величину ξ можно с помощью разбиения (2) представить в виде линейной комбинации индикаторов

$$\xi = \xi(\omega) = \sum_{i=1}^{k} x_i \chi_{A_i}(\omega), \tag{3}$$

так как левая и правая части (3) принимают одно и то же значение x_i при $\omega \in A_i$.

Пример 4. В примерах 2 и 3 положим $n = 3$. Тогда $\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\}$. Случайная величина μ_3 (число успехов в трех испытаниях) определяет следующее разбиение α_{μ_3}:

$$A_0 = \{\omega\colon \mu_3 = 0\} = \{000\},$$
$$A_1 = \{\omega\colon \mu_3 = 1\} = \{001,\ 010,\ 100\},$$
$$A_2 = \{\omega\colon \mu_3 = 2\} = \{011,\ 101,\ 110\},$$
$$A_3 = \{\omega\colon \mu_3 = 3\} = \{111\}.$$

Представление (2) для μ_3 имеет вид

$$\mu_3 = 0 \cdot \chi_{A_0} + \chi_{A_1} + 2\chi_{A_2} + 3\chi_{A_3}.$$

При любом $\omega \in \Omega$ ровно одна из величин χ_{A_0}, χ_{A_1}, χ_{A_2}, χ_{A_3} отлична от 0.

Законом распределения $\mathsf{P}_\xi(B)$ случайной величины ξ мы будем называть вероятность $\mathsf{P}_\xi(B) = \mathsf{P}\{\xi \in B\}$, определенную для каждого числового множества B. Закон распределения случайной величины ξ определяется ее значениями $x_1, x_2, \ldots, x_k$ и вероятностями $\mathsf{P}\{\xi = x_i\}$ этих значений. Обозначим $\mathsf{P}\{\xi = x_i\} = p_i$. Тогда закон распределения можно определить с помощью таблицы

x_1	x_2	...	x_k
p_1	p_2	...	p_k

, (4)

верхний ряд которой состоит из различных чисел x_i, а числа нижнего ряда удовлетворяют условиям $p_i \geqslant 0$, $\sum_{i=1}^{k} p_i = 1$. Эти условия вытекают из того, что

$p_i = \mathsf{P}(A_i)$ — вероятности разбиения $\sum_{i=1}^{k} A_i = \Omega$, откуда из аксиомы аддитивности следует $\sum_{i=1}^{k} \mathsf{P}(A_i) = \mathsf{P}(\Omega) = 1$. С помощью таблицы (4) можно определить вероятность

$$\mathsf{P}\{\xi \in B\} = \sum_{i:\ x_i \in B} p_i \tag{5}$$

для любого числового множества B.

В теории вероятностей часто рассматривают случайные величины ξ с законом распределения (5), не указывая ни вероятностного пространства $(\Omega, \mathscr{A}, \mathsf{P})$, ни функции $\xi(\omega)$, которая задает случайную величину. В этом случае предполагается, что существует какое-то вероятностное пространство $(\Omega, \mathscr{A}, \mathsf{P})$, на котором можно определить функцию $\xi = \xi(\omega)$ так, что таблица (3) будет задавать ее закон распределения. Выбор вероятностного пространства каждый раз определяется существом задачи или простотой получающейся схемы. Простейшим вероятностным пространством, связанным с законом распределения (5), будет пространство в котором множеством элементарных событий является множество значений случайной величины $\Omega = \{x_1, x_2, \ldots, x_k\}$ с элементарными вероятностями $p(x_i) = p_i$. Случайная величина определяется тогда функцией $\xi(x_i) = x_i$.

Закон распределения индикатора χ_A события A определяется таблицей

0	1
$1-\mathsf{P}(A)$	$\mathsf{P}(A)$

(6)

Каждой случайной величине соответствует закон распределения. Один и тот же закон распределения могут иметь разные случайные величины. Обратимся к примеру 1 с бросанием игральной кости. Обозначим события $A = \{\omega_1, \omega_3, \omega_5\}$, $B = \{\omega_2, \omega_4, \omega_6\}$, т. е. событие $A = \{$число выпавших очков нечетно$\}$, событие $B = \{$число выпавших очков четно$\}$. С этими событиями связаны разные случайные величины χ_A и χ_B

Однако в силу равенства $\mathsf{P}(A)=\mathsf{P}(B)=1/2$ эти разные случайные величины имеют один и тот же закон распределения (6).

Закон распределения ξ иногда называют кратко просто *законом* или *распределением*. Законом распределения часто называют задающую его таблицу (4).

Приведем несколько законов распределения.

1°. *Биномиальный закон* для числа успехов μ_n при n независимых испытаниях в схеме Бернулли

$$\mathsf{P}\{\mu_n=m\}=C_n^m p^m(1-p)^{n-m},\qquad m=0,1,\ldots,n$$

(см. (7) § 3).

2°. *Гипергеометрическое распределение* — распределение числа белых шаров ξ в выборке без возвращения объема n из урны, содержащей M белых и $N-M$ черных шаров (см. п. 1.6, пример 4)

$$\mathsf{P}\{\xi=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n},\qquad m=0,1,\ldots,\min(n,M).$$

3°. *Равномерное распределение* на $\{1,2,\ldots,N\}$:

$$\mathsf{P}(\xi=m\}=\frac{1}{N},\qquad m=1,2,\ldots,N.$$

С помощью числовой функции $y=g(x)$ и случайной величины $\xi=\xi(\omega)$ можно получить новую случайную величину $\eta=\eta(\omega)=g(\xi(\omega))$. Если функция $g(x)$ такова, что различным значениям ξ соответствуют различные значения η (т. е. $g(x_i)\neq g(x_j)$ для $x_i\neq x_j$), то разбиение (2), порожденное ξ, совпадает с разбиением, порожденным η, так как

$$A_i=\{\omega:\ \xi(\omega)=x_i\}=\{\omega:\ g(\xi(\omega))=g(x_i)\}.$$

В этом случае закон распределения $\eta=g(\xi)$ задается таблицей

$g(x_1)$	$g(x_2)$	...	$g(x_k)$
p_1	p_2	...	p_k

, (7)

а сама случайная величина η представима в виде суммы

$$\eta = g(\xi) = \sum_{i=1}^{k} g(x_i)\chi_{A_i}. \tag{8}$$

Если же при некоторых $i \neq j$ $g(x_i) = g(x_j)$, то закон распределения $\eta = g(\xi)$ будет иным, хотя его легко получить из таблицы (7). Обозначим $y_1 < y_2 < \ldots$ $\ldots < y_m$ все различные значения величины $\eta = g(\xi)$. Тогда закон распределения η определяется таблицей

y_1	y_2	...	y_m
q_1	q_2	...	q_m

(9)

где в верхней строке выписаны все различные значения $g(x_i)$ верхней строки (7), а в нижней строке q_j — это сумма тех p_i, для которых $g(x_i) = y_j$. Заметим, что равенство (8) справедливо и в том случае, когда $g(x_i)$ при некоторых x_i совпадают.

Пример 5. Пусть ξ случайная величина с законом распределения $\mathsf{P}\{\xi = k\} = \frac{1}{10}$, $k = 0, 1, 2, \ldots, 9$, и функция $g(x) = (x-5)^2$. Тогда случайная величина $\eta = (\xi - 5)^2$ будет иметь следующий закон распределения: $\mathsf{P}\{\eta = 0\} = \mathsf{P}\{\eta = 25\} = 1/10$, $\mathsf{P}\{\eta = 1\} = \mathsf{P}\{\eta = 4\} = \mathsf{P}\{\eta = 9\} = \mathsf{P}\{\eta = 16\} = 1/5$. Вычислим, например, $\mathsf{P}\{\eta = 4\}$. Имеем

$$\mathsf{P}\{\eta = 4\} = \mathsf{P}\{(\xi - 5)^2 = 4\} = \mathsf{P}\{\xi = 3\} + \mathsf{P}(\xi = 7\} = 1/10 + 1/10 = 1/5.$$

5.2. Случайные величины в счетной схеме. Пусть в вероятностном пространстве $(\Omega, \mathscr{A}, \mathsf{P})$ множество

$$\Omega = \{\omega\} = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$$

счетно, $\mathscr{A}$ состоит из всех подмножеств Ω, а вероятность P задается элементарными вероятностями $p(\omega)$. Случайная величина ξ в этом случае тоже определяется как числовая функция $\xi = \xi(\omega)$ от элементарного события. Множество различных значений x_i случайной величины ξ может быть конечным или счетным. Аналогично (2) со случайной величиной ξ можно связать разбиение α_ξ, состоящее из множеств

$$A_i = \{\omega\colon \xi(\omega) = x_i\}, \quad i = 1, 2, \ldots,$$

однако число таких множеств может быть счетно, тогда порожденное ξ разбиение α_ξ называется *счетным.* Представление ξ в виде суммы, аналогичной (3), приводит к ряду

$$\xi = \xi(\omega) = \sum_{i=1}^{\infty} x_i \chi_{A_i}, \tag{10}$$

в котором справа при каждом ω имеется не более одного ненулевого члена. Закон распределения $\mathsf{P}\{\xi \in B\}$ в этом случае задается таблицей, аналогичной (4) и состоящей из двух бесконечных строк, определяющих вероятности $\mathsf{P}\{\xi = x_i\} = p_i$, $\sum\limits_{i=1}^{\infty} p_i = 1$.

Приведем два распределения, часто встречающихся в различных моделях.

1. *Пуассоновское распределение* определяется вероятностями

$$\mathsf{P}\{\xi = m\} = \frac{a^m}{m!} e^{-a}, \quad m = 0, 1, \ldots,$$

зависящими от параметра $a > 0$. Легко видеть, что в этом случае выполняется условие $\sum\limits_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} = 1$.

2. *Геометрическое распределение* зависит от параметра $0 < p < 1$ и определяется вероятностями

$$\mathsf{P}\{\xi = m\} = p^m q, \quad m = 0, 1, 2, \ldots, \ q = 1 - p.$$

В этом случае также выполнено условие $\sum\limits_{m=0}^{\infty} p^m q = 1$.

5.3. Случайные величины в общей схеме. Функция распределения. В случае произвольного вероятностного пространства $(\Omega, \mathscr{A}, \mathsf{P})$ *случайной величиной* ξ называется такая функция $\xi = \xi(\omega)$ от элементарного события ω, для которой при любом численном значении x множество $\{\omega: \xi(\omega) \leqslant x\} \in \mathscr{A}$, т. е. неравенство $\{\xi \leqslant x\}$ является событием. Вероятность этого события $\mathsf{P}\{\xi \leqslant x\}$ называется *функцией распределения*. Мы будем обозначать ее $F_\xi(x)$, а иногда просто $F(x)$.

Пример 6. Пусть в вероятностном пространстве $(\Omega, \mathscr{A}, \mathsf{P})$ Ω состоит из точек (x, y) единичного квадрата плоскости $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$, $\mathscr{A}$ — квадрируемые множества этого квадрата, а вероятность $\mathsf{P}(A) = |A|$ — это площадь множества A. Определим случайную величину $\xi = \chi_A$, где χ_A — индикатор события A (т. е. $\chi_A(\omega) = 1$, если $\omega \in A$, и $\chi_A(\omega) = 0$, если $\omega \notin A$). Случайная величина ξ имеет всего два возможных значения $\xi = 0$ и $\xi = 1$. Ее функция распределения определяется равенствами

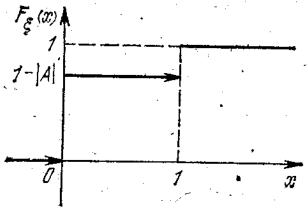


Рис. 7.

$$F_\xi(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ \mathsf{P}(\bar{A}) = 1 - |A| & \text{при} \quad 0 \leqslant x < 1, \\ 1 & \text{при} \quad x \geqslant 1; \end{cases}$$

график показан на рис. 7. Если множество B не квадрируемо, то функция элементарного события $g(\omega)$, определяемая равенствами

$$g(\omega) = \begin{cases} 0, & \text{если} \quad \omega \in B, \\ 1, & \text{если} \quad \omega \notin B, \end{cases}$$

не будет случайной величиной, так как множество $\{\omega: g(\omega) \leqslant 1/2\}$ совпадает с множеством B, которое не является событием и для которого не определена вероятность.

Свойства функции распределения. Функция распределения $F_\xi(x)$ задает закон распределения

$$\mathsf{P}_\xi(B) = \mathsf{P}\{\xi \in B\} \tag{11}$$

для достаточно богатого набора числовых множеств B. Пусть $x_1 < x_2$. Тогда событие $\{\xi \leqslant x_2\}$

можно представить в виде суммы несовместных событий $\{\xi \leqslant x_1\} + \{x_1 < \xi \leqslant x_2\}$; применяя аксиому сложения А3, имеем $\mathsf{P}\{\xi \leqslant x_2\} = \mathsf{P}\{\xi \leqslant x_1\} + \mathsf{P}\{x_1 < \xi \leqslant x_2\}$, откуда получаем

$$\mathsf{P}\{x_1 < \xi \leqslant x_2\} = F_\xi(x_2) - F_\xi(x_1). \qquad (12)$$

Таким образом, с помощью формулы (12) функция распределения $F_\xi(x)$ определяет вероятность $\mathsf{P}\{\xi \in B\}$ для $B = (x_1, x_2]$ — полуинтервала, открытого слева и замкнутого справа. С помощью функции распределения $F_\xi(x)$ можно определить вероятность попадания ξ в множества B, представимые в виде суммы непересекающихся интервалов (конечных или бесконечных). Пусть $-\infty \leqslant x_1 < y_1 \leqslant x_2 < y_2 \leqslant \ldots \leqslant x_N < y_N \leqslant +\infty$ и $B = \bigcup\limits_{n=1}^{N} (x_n, y_n]$. Тогда по аксиоме сложения А3 и формуле (12) имеем

$$\mathsf{P}_\xi(B) = \mathsf{P}\{\xi \in B\} = \sum_{n=1}^{N} [F_\xi(y_n) - F_\xi(x_n)]. \qquad (13)$$

Формулу (13) можно распространить и на сумму счетного числа интервалов $(x_k, y_k]$. Для этого надо применить расширенную аксиому сложения А4.

Если $x_1 = -\infty$ или $y_N = +\infty$, то в (13) появляются выражения $F_\xi(-\infty)$ и $F_\xi(+\infty)$. Определим их как пределы $F_\xi(+\infty) = \lim\limits_{x \to +\infty} F_\xi(x)$, $F_\xi(-\infty) = \lim\limits_{x \to -\infty} F_\xi(x)$ и докажем, что $F_\xi(+\infty) = 1$ и $F(-\infty) = 0$. Введем события $A_n = \{-n < \xi \leqslant n\}$, $n = 1, 2, \ldots, A_0 = \varnothing$, и $B_n = A_n \setminus A_{n-1} = \{n - 1 < \xi \leqslant n\} + \{-n < \xi \leqslant -(n-1)\}$. Нетрудно видеть, что события B_n попарно несовместны и $\Omega = \sum\limits_{n=1}^{\infty} B_n$, поэтому по расширенной аксиоме сложения А4

$$1 = \mathsf{P}(\Omega) = \sum_{n=1}^{\infty} \mathsf{P}(B_n) = \lim_{N \to \infty} \sum_{n=1}^{N} \mathsf{P}(B_n) =$$
$$= \lim_{N \to \infty} \mathsf{P}(A_N) = \lim_{N \to \infty} F_\xi(N) - \lim_{N \to \infty} F(-N) =$$
$$= F_\xi(+\infty) - F_\xi(-\infty), \qquad (14)$$

так как $A_N = \sum_{n=1}^{N} B_n$ и $\mathrm{P}(A_N) = F_\xi(N) - F_\xi(-N)$. Все значения $F_\xi(x)$ лежат между 0 и 1, поэтому $0 \leqslant F_\xi(\mp\infty) \leqslant 1$, и из (14) следует $F_\xi(+\infty) = 1$, $F_\xi(-\infty) = 0$.

Обозначим $F_\xi(x-0)$ предел $\lim_{n\to\infty} F_\xi\left(x - \frac{1}{n}\right)$, который называется *пределом в точке x слева*. Покажем, что $F_\xi(x-0) = \mathrm{P}\{\xi < x\}$. В самом деле, событие $\{\xi < x\}$ представимо в виде счетной суммы попарно несовместных событий $\sum_{n=1}^{\infty}\left\{x - \frac{1}{n-1} < \xi \leqslant x - \frac{1}{n}\right\}$ $\left(\text{при } n = 1 \text{ мы полагаем } \left\{-\frac{1}{n-1} < \xi \leqslant x - \frac{1}{n}\right\} = \{\xi \leqslant x - 1\}\right)$. По аксиоме А4

$$\mathrm{P}\{\xi < x\} = \sum_{n=1}^{\infty} \mathrm{P}\left\{x - \frac{1}{n-1} < \xi \leqslant x - \frac{1}{n}\right\} =$$
$$= \lim_{N\to\infty} \sum_{n=1}^{N} \mathrm{P}\left\{x - \frac{1}{n-1} < \xi \leqslant x - \frac{1}{n}\right\} =$$
$$= \lim_{N\to\infty} F_\xi\left(x - \frac{1}{N}\right) = F_\xi(x-0).$$

Теперь мы можем доказать, что

$$\mathrm{P}\{\xi = x\} = F_\xi(x) - F_\xi(x-0). \tag{15}$$

В самом деле, (15) легко следует из $\{\xi \leqslant x\} = \{\xi < x\} + \{\xi = x\}$ и аксиомы А3. С помощью формул (12) и (15) легко теперь получить вероятности попадания в открытый и замкнутый интервалы. Так как при $x_1 < x_2$ имеют место соотношения $\{x_1 \leqslant \xi \leqslant x_2\} = \{\xi = x_1\} + \{x_1 < \xi \leqslant x_2\}$ и $\{x_1 < \xi \leqslant x_2\} = \{x_1 < \xi < x_2\} + \{\xi = x_2\}$, то по аксиоме А3 имеем

$$\mathrm{P}\{x_1 \leqslant \xi \leqslant x_2\} = F_\xi(x_2) - F_\xi(x_1 - 0),$$
$$\mathrm{P}\{x_1 < \xi < x_2\} = F_\xi(x_2 - 0) - F_\xi(x_1).$$

Можно показать, что предел в точке x *справа* $F_\xi(x+0) = \lim_{n\to\infty} F_\xi\left(x + \frac{1}{n}\right)$ функции распределения равен $F_\xi(x)$ (это свойство означает, что функция распределения *непрерывна спра-*

ва). Для этого надо применить аксиому А4 к сумме $\{x<\xi\leqslant \leqslant x+1\}=\sum_{n=1}^{\infty}\left\{x+\frac{1}{n+1}<\xi\leqslant x+\frac{1}{n}\right\}$.

Таким образом, функция распределения $F(x)$ обладает следующими свойствами:

1) $0\leqslant F(x)\leqslant 1$ для всех x;
2) $F(x_1)\leqslant F(x_2)$, если $x_1\leqslant x_2$;
3) $F(-\infty)=0$, $F(+\infty)=1$;
4) $F(x+0)=F(x)$.

Можно показать, что любая функция $F(x)$, обладающая перечисленными свойствами, может быть функцией распределения некоторой случайной величины.

Дискретные и непрерывные случайные величины. Формулой (9) на общем вероятностном пространстве $(\Omega, \mathscr{A}, \mathsf{P})$ можно задать случайную величину ξ, если под A_i, $i=1, 2, \ldots$, понимать события, образующие счетное или конечное разбиение, т. е. попарно несовместные события, для которых $\sum_{i=1}^{\infty} A_i=\Omega$.

Закон распределения $\mathsf{P}_\xi(B)=\mathsf{P}\{\xi\in B\}$ в этом случае задается вероятностями $\mathsf{P}\{\xi=x_i\}$ так, что

$$\mathsf{P}_\xi(B)=\sum_{x_i\in B}\mathsf{P}\{\xi=x_i\}. \tag{16}$$

Такие случайные величины и их законы распределения (16) называют *дискретными.* Все случайные величины п. п. 5.1 и 5.2, определенные на конечном или счетном вероятностном пространстве, являются дискретными.

Другой важный класс случайных величин образуют так называемые *непрерывные* случайные величины ξ, для которых $\mathsf{P}\{\xi=x\}=0$ при любых x. В этом случае закон распределения $P_\xi(B)$ нельзя задать формулой (16). Мы будем рассматривать такие непрерывные случайные величины ξ, закон распределения которых можно задать с помощью формулы

$$\mathsf{P}\{\xi\in B\}=\int_B p_\xi(x)\,dx, \tag{17}$$

где $p_\xi(x) \geqslant 0$ — интегрируемая функция. Функция $p_\xi(x)$ называется *плотностью* распределения случайной величины ξ. Из (17) следует (при $B = (-\infty, \infty)$), что

$$\int_{-\infty}^{\infty} p_\xi(x)\,dx = 1. \tag{18}$$

В практически важных случаях функции $p_\xi(x)$ оказываются непрерывными или кусочно-непрерывными. Если $B = (-\infty, x)$, то из (16) получаем связь между плотностью $p_\xi(x)$ и функцией распределения

$$F_\xi(x) = \int_{-\infty}^{x} p_\xi(u)\,du. \tag{19}$$

Из этой формулы следует, что в точке x, в которой непрерывна плотность $p_\xi(x)$,

$$\frac{dF_\xi(x)}{dx} = p_\xi(x),$$

и при $\Delta x > 0$ и $\Delta x \to 0$

$$\mathsf{P}\{x < \xi < x + \Delta x\} = p_\xi(x)\,\Delta x + o(\Delta x).$$

Любая кусочно-непрерывная неотрицательная функция $p(x)$, обладающая свойством (17), может быть плотностью некоторого распределения вероятностей.

Приведем несколько примеров часто встречающихся законов распределения непрерывных случайных величин.

1°. *Равномерное распределение* на $[a, b]$ задается плотностью

$$p(x) = \begin{cases} \dfrac{1}{b-a}, & \text{если } 0 \leqslant x \leqslant b, \\ 0, & \text{если } x < a \text{ или } x > b. \end{cases}$$

График плотности $y = p(x)$ равномерного распределения см. на рис. 8.

2°. *Показательное распределение* задается плотностью, зависящей от параметра $\lambda > 0$,

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{если} \quad x \geqslant 0, \\ 0, & \text{если} \quad x < 0. \end{cases}$$

График $y = p(x)$ см. на рис. 9.

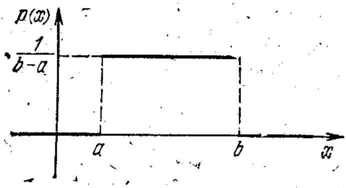


Рис. 8.

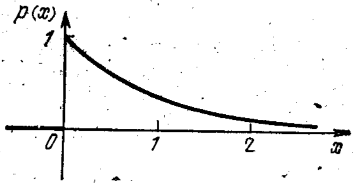


Рис. 9.

3°. *Нормальное*, или *гауссовское*, *распределение* задается плотностью, зависящей от параметров $-\infty \leq a \leq \infty$ и $\sigma > 0$:

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

График $y = p(x)$ см. на рис. 10.

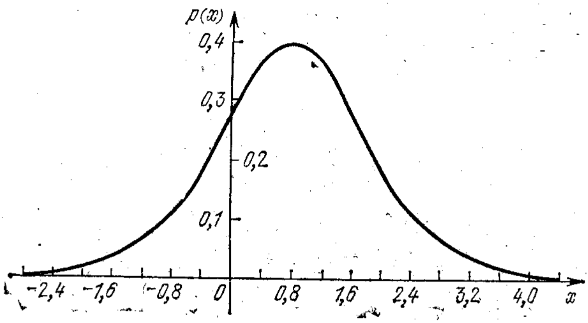


Рис. 10.

При пользовании формулами (17), (18) или (19) надо помнить, что плотность $p(x)$ часто задается разными аналитическими формулами на разных интервалах.

Пример 7. Вычислим функцию распределения случайной величины, равномерно распределенной на отрезке $[a, b]$. Для этого воспользуемся формулой

(19) и заданной выше плотностью $p(x)$ равномерного распределения. При $x < 0$ имеем $F(x) = \int\limits_{-\infty}^{x} p(u)\,du = 0$. Если $a \leqslant x \leqslant b$, то

$$F(x) = \int\limits_{-\infty}^{x} p(u)\,du = \int\limits_{-\infty}^{a} p(u)\,du + \int\limits_{a}^{x} p(u)\,du =$$
$$= 0 + \frac{1}{b-a}\int\limits_{a}^{x} du = \frac{x-a}{b-a}.$$

Если $x > b$, то

$$F(x) = \int\limits_{-\infty}^{a} p(u)\,du + \int\limits_{a}^{b} p(u)\,du + \int\limits_{b}^{x} p(u)\,du =$$
$$= 0 + \frac{1}{b-a}\int\limits_{a}^{b} du + 0 = 1.$$

Таким образом,

$$F(x) = \begin{cases} 0, & \text{если } x < a, \\ \dfrac{x-a}{b-a}, & \text{если } a \leqslant x \leqslant b, \\ 1, & \text{если } x > b. \end{cases}$$

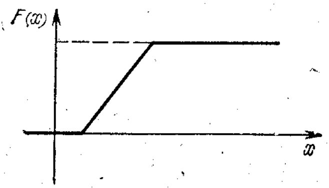


Рис. 11.

График функции распределения $y = F(x)$ показан на рис. 11.

5.4. Функции от случайных величин. Пусть ξ — непрерывная случайная величина с функцией распределения $F_\xi(x)$ и плотностью $p_\xi(x)$. Если $g(x)$ — числовая функция от действительного аргумента x, то можно образовать случайную величину $\eta = g(\xi)$, которая является функцией от ξ. Спрашивается, как по закону распределения ξ найти закон распределения η? Приведем несколько примеров, из которых станет ясно, как надо поступать в общем случае. Мы будем рассматривать непрерывные случайные величины, для которых $\mathsf{P}\{\xi = x\} = 0$, поэтому функция распределения $F_\xi(x)$ будет непрерывна во всех точках и $\mathsf{P}\{\xi \leqslant x\} = \mathsf{P}\{\xi < x\} = F_\xi(x)$.

Пример 8. Рассмотрим линейную функцию $g(x) = ax + b$. Положим $\eta = a\xi + b$. Если $a > 0$, то по определению функции распределения

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \mathsf{P}\{a\xi + b < x\} = \\ = \mathsf{P}\left\{\xi < \frac{x-b}{a}\right\} = F_\xi\left(\frac{x-b}{a}\right),$$

откуда $p_\eta(x) = F'_\eta(x) = F'_\xi\left(\frac{x-b}{a}\right)\frac{1}{a} = \frac{1}{a}p_\xi\left(\frac{x-b}{a}\right)$.

Если $a < 0$, то имеем

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \mathsf{P}\{a\xi + b < x\} = \\ = \mathsf{P}\left\{\xi > \frac{x-b}{a}\right\} = 1 - F_\xi\left(\frac{x-b}{a}\right)$$

и

$$p_\eta(x) = F'_\eta(x) = -\frac{1}{a}F'_\xi\left(\frac{x-b}{a}\right) = -\frac{1}{a}p_\xi\left(\frac{x-b}{a}\right).$$

Объединяя эти два случая, имеем

$$p_\eta(x) = \frac{1}{|a|}p_\xi\left(\frac{x-b}{a}\right).$$

Пример 9. Рассмотрим монотонно возрастающую функцию $g(x) = x^3$. Положим $\eta = \xi^3$. Имеем по определению функции распределения

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \mathsf{P}\{\xi^3 < x\} = \mathsf{P}\{\xi < x^{1/3}\} = F_\xi(x^{1/3}).$$

Дифференцируя это равенство, получаем выражение для плотности $p_\eta(x)$:

$$p_\eta(x) = F'_\eta(x) = F'_\xi(x^{1/3})\cdot\frac{1}{3}x^{-2/3} = \frac{1}{3}x^{-2/3}p_\xi(x^{1/3}).$$

Пример 10. Рассмотрим монотонно убывающую функцию $g(x) = e^{-x}$. Найдем $F_\eta(x)$ и $p_\eta(x)$ случайной величины $\eta = e^{-\xi}$. Имеем при $x > 0$

$$F_\eta(x) = \mathsf{P}\{\eta < x\} = \mathsf{P}\{e^{-\xi} < x\} = \mathsf{P}\{\xi > -\ln x\} = \\ = 1 - \mathsf{P}\{\xi \leqslant -\ln x\} = 1 - F_\xi(-\ln x),$$

откуда

$$p_\eta(x)=F'_\eta(x)=\frac{1}{x}F'_\xi(-\ln x)=\frac{1}{x}p_\xi(-\ln x).$$

При $x<0$ $F_\eta(x)=\mathsf{P}\{\eta<x\}=0$, так как $\eta=e^{-\xi}\geqslant 0$ при всех ξ. Поэтому на отрицательной полуоси $p_\eta(x)=F'_\eta(x)=0$.

Пример 11. Рассмотрим немонотонную функцию $g(x)=x^2$. Найдем $F_\eta(x)$ и $p_\eta(x)$ случайной величины $\eta=\xi^2$. Имеем при $x>0$

$$F_\eta(x)=\mathsf{P}\{\eta<x\}=\mathsf{P}\{\xi^2<x\}=\mathsf{P}\{-\sqrt{x}<\xi<\sqrt{x}\}=$$
$$=\mathsf{P}\{\xi<\sqrt{x}\}-\mathsf{P}\{\xi\leqslant-\sqrt{x}\}=F_\xi(\sqrt{x})-F_\xi(-\sqrt{x}),$$

откуда

$$p_\eta(x)=F'_\eta(x)=$$
$$=F'_\xi(\sqrt{x})\frac{1}{2\sqrt{x}}+\frac{1}{2\sqrt{x}}F'_\xi(-\sqrt{x})=\frac{1}{2\sqrt{x}}(p_\xi\sqrt{x})+$$
$$+p_\xi(-\sqrt{x})).$$

При $x\leqslant 0$ имеем $F_\eta(x)=0$ и $p_\eta(x)=0$.

Задачи

1. Бросается игральная кость. Обозначим события $A=\{$число выпавших очков четно$\}$, $B=\{$число выпавших очков делится на 3$\}$. Найти: а) закон распределения χ_A; б) закон распределения χ_B; в) разбиение, порожденное случайной величиной $\xi=\chi_A+\chi_B$, и ее закон распределения.

2. На пустую шахматную доску случайно ставится слон. Вероятности поставить слона на каждую клетку будем считать одинаковыми. Найти закон распределения ξ — числа битых полей.

3. Из 28 костей домино случайно и равновероятно выбирается одна. Найти закон распределения суммы очков ξ на ее половинах.

4. Случайная величина ξ имеет пуассоновское распределение с параметром a. Найти вероятности событий: $A=\{\xi$ — четное$\}$, $B=\{\xi$ — нечетное$\}$. Указание. Воспользоваться рядами

$$e^a=\sum_{n=0}^{\infty}\frac{a^n}{n!},\qquad e^{-a}=\sum_{n=0}^{\infty}\frac{(-1)^n a^n}{n!}.$$

5. Плотность распределения ξ задана формулой

$$p_\xi(x)=\begin{cases}Cx^{-3/2} & \text{при } x\geqslant 1,\\ 0 & \text{при } x<1.\end{cases}$$

Найти C и функцию распределения $F_\xi(x)$.

6. Производится стрельба по круглой мишени радиуса r. Найти функцию распределения $F_\xi(x)$ и плотность $p_\xi(x)$ случайной величины ξ, равной расстоянию от центра мишени точки попадания, если точка попадания равномерно распределена по мишени.

7. Мишень радиуса r десятью концентрическими окружностями радиусов $r_k = \frac{k}{10} r$, $k = 1, 2, \ldots, 10$, разделена на 10 частей, которые занумерованы, начиная от центра, числами 10, 9, ..., 3, 2, 1. При попадании в соответствующую зону, номер ν зоны дает число очков стрелка. В условиях задачи 6 найти закон распределения ν.

8. Случайная величина ξ имеет равномерное распределение на отрезке $[1, 2]$. Найти вероятность $\mathsf{P}\{2 < \xi^2 < 5\}$.

9. Случайная величина ξ имеет равномерное распределение на отрезке $[-1, 3]$. Найти вероятность $\mathsf{P}\{\xi^2 \leqslant 2\}$.

10. Случайная величина ξ имеет показательное распределение с плотностью $p_\xi(x) = e^{-x}$, $x \geqslant 0$. Найти функцию распределения и плотность случайной величины $\eta = e^{-\xi}$.

11. Случайная величина ξ равномерно распределена на отрезке $[-1, 2]$. Найти плотность $p_\eta(x)$ распределения $\eta = \xi^2$.

§ 6. СОВМЕСТНЫЕ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

6.1. Многомерные законы распределения. Очень часто в вероятностных моделях приходится рассматривать сразу несколько случайных величин. Например, при стрельбе по плоской мишени случайная точка попадания имеет две координаты ξ и η, которые являются случайными величинами; при антропометрических исследованиях основными параметрами человеческого тела считаются вес, рост и объем груди, которые при случайном выборе человека из какой-либо совокупности также случайны; при массовом изготовлении каких-либо деталей на станке различные их размеры тоже дают пример случайных величин.

В математической модели в этих случаях на вероятностном пространстве $(\Omega, \mathscr{A}, \mathsf{P})$ определены несколько случайных величин $\xi_1, \xi_2, \ldots, \xi_r$, которые иногда удобно рассматривать как координаты случайной точки, или случайного вектора $\xi = (\xi_1, \ldots, \xi_r)$ из r-мерного пространства R^r. Совместным законом распределения этих случайных величин называется вероятность попадания точки ξ в r-мерное множество B:

$$P_\xi(B) = \mathsf{P}\{\xi \in B\}, \tag{1}$$

рассматриваемая в зависимости от множества B. Закон распределения (1) называют также *многомерным*, или, более точно, *r-мерным*. Мы рассмотрим два способа задания закона (1). Пусть имеется конечный или счетный набор векторов $\boldsymbol{x}(i) = (x_1(i), x_2(i), \ldots, x_r(i))$, и пусть даны вероятности $\mathsf{P}\{\xi = \boldsymbol{x}(i)\}$, удовлетворяющие условию $\sum_{i=1}^{\infty} \mathsf{P}\{\xi = \boldsymbol{x}(i)\} = 1$. Закон распределения (1), задаваемый формулой

$$P_\xi(B) = \sum_{\boldsymbol{x}(i) \in B} \mathsf{P}\{\xi = \boldsymbol{x}(i)\}, \tag{2}$$

называется *дискретным законом распределения*.

Пример 1. В схеме n независимых испытаний с r исходами в § 3 мы ввели полиномиальное распределение. Его можно рассматривать как r-мерное распределение случайных величин $\mu_1, \mu_2, \ldots, \mu_r$, где μ_i — число i-х исходов в n испытаниях. Тогда

$$\mathsf{P}\{\mu_1 = m_1, \ldots, \mu_r = m_r\} = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r},$$

если $m_1 + \ldots + m_r = n$, и

$$\mathsf{P}\{\boldsymbol{\mu} = \boldsymbol{m}\} = 0$$

в остальных случаях (здесь $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_r)$, $\boldsymbol{m} = (m_1, \ldots, m_r)$). Другой класс законов распределения задается *r-мерной плотностью* $p_\xi(\boldsymbol{x}) = p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) \geqslant 0$, удовлетворяющей условию $\int \ldots \int_{R^r} p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r)\, dx_1 \ldots dx_r = 1$:

$$\mathsf{P}_\xi(B) = \int \ldots \int_B p_\xi(x_1, \ldots, x_r)\, dx_1 \ldots dx_r. \tag{3}$$

Если $B = \{(u_1, \ldots, u_r): x_i \leqslant u_i \leqslant x_i + \Delta x_i,\ i = 1, \ldots, r\}$, $(x_1, \ldots, x_r)$ — точка непрерывности плотности, то при $\Delta x_i \to 0$, $i = 1, \ldots, r$, применяя к инте-

гралу в правой части (3) теорему о среднем, получим
$\mathsf{P}_{\xi}(B)=p_{\xi_1\ldots\xi_r}(x_1,\ldots,x_r)\,\Delta x_1\ldots\Delta x_r+o(\Delta x_1\ldots\Delta x_r)$.

Пример 2. Пусть при стрельбе по плоской квадратной мишени $|x|\leqslant a$, $|y|\leqslant a$ точка попадания (ξ,η) имеет следующую двумерную плотность:

$$p_{\xi\eta}(x, y)=\begin{cases}\dfrac{1}{4a^2}, & \text{если } |x|\leqslant a \text{ и } |y|\leqslant a,\\ 0 & \text{в остальных случаях.}\end{cases}$$

Тогда вероятность попасть в круг $x^2+y^2\leqslant a^2$ вычисляется, согласно формуле (3), следующим образом:

$$\mathsf{P}\{\xi^2+\eta^2\leqslant a^2\}=\iint\limits_{x^2+y^2\leqslant a^2} p_{\xi\eta}(x, y)\,dx\,dy=\frac{\pi a^2}{4a^2}=\frac{\pi}{4}.$$

Полагая в формулах (2) или (3) множество B равным $\{y=(y_1, \ldots, y_r)\colon y_i\leqslant x_i,\ i=1, \ldots, r\}$, где $x_1, \ldots, x_r$ — заданные числа, получаем вероятность события $\{\xi_1\leqslant x_1, \ldots, \xi_r\leqslant x_r\}$, зависящую от x_1, $x_2, \ldots, x_r$; эта вероятность называется *r-мерной функцией распределения* случайных величин $\xi_1, \ldots, \xi_r$ и обозначается

$$F_{\xi}(x)=F_{\xi_1\ldots\xi_r}(x_1, \ldots, x_r)=\\ =\mathsf{P}\{\xi_1\leqslant x_1, \ldots, \xi_r\leqslant x_r\}. \qquad (4)$$

Можно доказать, что r-мерная функция распределения (4) однозначно задает закон распределения $\mathsf{P}_{\xi}(B)$. Например, если имеется r-мерная плотность $p_{\xi_1\ldots\xi_r}(x_1, \ldots, x_r)$, то согласно (3) и (4) функция распределения $F_{\xi_1\ldots\xi_r}(x_1,\ldots,x_r)$ задается интегралом

$$F_{\xi_1\ldots\xi_r}(x_1, \ldots, x_r)=\\ =\int\limits_{-\infty}^{x_1} du_1\int\limits_{-\infty}^{x_2} du_2\ldots\int\limits_{-\infty}^{x_r} p_{\xi_1\ldots\xi_r}(u_1, \ldots, u_r)\,du_r, \qquad (5)$$

откуда дифференцированием по $x_1, x_2, \ldots, x_r$ получаем равенство

$$\frac{\partial^r F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r)}{\partial x_1 \ldots \partial x_r} = p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r), \qquad (6)$$

справедливое во всех точках непрерывности плотности $p_{\xi_1 \ldots \xi_r}$. Таким образом, зная функцию распределения $F_{\xi_1 \ldots \xi_r}$, по формуле (6) можно найти плотность $p_{\xi_1 \ldots \xi_r}$, а по этой плотности формулой (3) определяется закон распределения $\mathrm{P}_\xi(B)$.

С помощью r-мерного закона распределения $\mathrm{P}_{\xi_1 \ldots \xi_r}$ можно определить любой m-мерный закон распределения $\mathrm{P}_{\xi_1 \ldots \xi_m}$, $m < r$. Если закон распределения задан функцией распределения (4), то, полагая $x_{m+1} \to +\infty, \ldots, x_r \to +\infty$, получаем

$$F_{\xi_1 \ldots \xi_m}(x_1, \ldots, x_m) =$$

$$= \mathrm{P}\{\xi_1 \leqslant x_1, \ldots, \xi_m \leqslant x_m\} = \lim_{\substack{x_{m+1} \to +\infty \\ \cdots\cdots\cdots \\ x_r \to +\infty}} F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r). \qquad (7)$$

Обозначая $F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_m, +\infty, \ldots, +\infty)$ предел справа в (7), можно записать (7) следующим образом:

$$F_{\xi_1 \ldots \xi_m}(x_1, \ldots, x_m) =$$

$$= F_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_m, +\infty, \ldots, +\infty). \qquad (8)$$

Полагая $x_{m+1} \to +\infty, \ldots, x_r \to +\infty$ в формуле (5) и учитывая формулу (8), мы можем получить плотность распределения

$$p_{\xi_1 \ldots \xi_m}(x_1, \ldots, x_m) =$$

$$= \int_{-\infty}^{\infty} dx_{m+1} \ldots \int_{-\infty}^{\infty} p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_m, x_{m+1}, \ldots, x_r)\, dx_r. \qquad (9)$$

В частности, с помощью формул типа (8) и (9) из r-мерного распределения можно получить одномерные распределения отдельных компонент:

$$F_{\xi_1}(x_1) = F_{\xi_1 \ldots \xi_r}(x_1, +\infty, \ldots, +\infty),$$
$$F_{\xi_2}(x_2) = F_{\xi_1 \ldots \xi_r}(+\infty, x_2, +\infty, \ldots, +\infty),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$F_{\xi_r}(x_r) = F_{\xi_1 \ldots \xi_r}(+\infty, \ldots, +\infty, x_r)$$

и

$$p_{\xi_1}(x) = \int\limits_{R^{r-1}} p_{\xi_1 \ldots \xi_r}(x, x_2, \ldots, x_r)\, dx_2 \ldots dx_r,$$
$$p_{\xi_2}(x) = \int\limits_{R^{r-1}} p_{\xi_1 \ldots \xi_r}(x_1, x, x_3, \ldots, x_r)\, dx_1\, dx_3 \ldots dx_r,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$p_{\xi_r}(x) = \int\limits_{R^{r-1}} p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_{r-1}, x)\, dx_1 \ldots dx_{r-1}.$$

6.2. Независимость случайных величин. Случайные величины $\xi_1, \xi_2, \ldots, \xi_r$ называются *независимыми*, если для любых числовых множеств $B_1, B_2, \ldots, B_r$, для которых определены вероятности событий $\{\xi_i \in B_i\}$, имеет место равенство

$$\mathsf{P}\{\xi_1 \in B_1, \xi_2 \in B_2, \ldots, \xi_r \in B_r\} = \\ = \mathsf{P}\{\xi_1 \in B_1\}\, \mathsf{P}\{\xi_2 \in B_2\} \ldots \mathsf{P}\{\xi_r \in B_r\}. \quad (10)$$

Если равенство (10) нарушается для каких-нибудь множеств B_i, то случайные величины $\xi_1, \ldots, \xi_r$ называются *зависимыми*.

В частности, если $B_i = \{y_i\colon y_i \leqslant x_i\}$, то для независимых случайных величин

$$F_{\xi_1 \ldots \xi_r}(x_1, x_2, \ldots, x_r) = F_{\xi_1}(x_1)\, F_{\xi_2}(x_2) \ldots F_{\xi_r}(x_r), \quad (11)$$

т. е. r-мерная функция распределения равна произведению одномерных функций распределения. Можно показать, что условие (11) можно принять за определение независимости в общем случае. Если имеется r-мерная плотность $p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r)$, то с помощью дифференцирования по $x_1, \ldots, x_r$ равенства (11) получаем, что для независимых случайных величин

$$p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = p_{\xi_1}(x_1)\, p_{\xi_2}(x_2) \ldots p_{\xi_r}(x_r). \quad (12)$$

В случае дискретного распределения из определения независимости (10) вытекает равенство

$$\mathrm{P}\{\xi_1 = x_1,\ \xi_2 = x_2, \ldots, \xi_r = x_r\} = \\ = \mathrm{P}\{\xi_1 = x_1\}\, \mathrm{P}\{\xi_2 = x_2\} \ldots \mathrm{P}\{\xi_r = x_r\}. \quad (13)$$

Условия (12) и (13) равносильны определению независимости (10) в случае соответственно непрерывного и дискретного распределения. В самом деле, если выполнено равенство (12), то

$$\mathrm{P}\{\xi_1 \in B_1, \ldots, \xi_r \in B_r\} = \\ = \int_{B_1} dx_1 \ldots \int_{B_r} p_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r)\, dx_1 \ldots dx_r = \\ = \int_{B_1} p_{\xi_1}(x_1)\, dx_1 \cdot \int_{B_2} p_{\xi_2}(x_2)\, dx_2 \ldots \int_{B_r} p_{\xi_r}(x_r)\, dx_r = \\ = P\{\xi_1 \in B_1\} \ldots P\{\xi_r \in B_r\}.$$

Аналогично доказывается эквивалентность условий независимости (10) и (13) для дискретных распределений.

В п. 6.1 мы показали, что многомерное распределение определяет одномерные распределения. В общем случае по одномерным распределениям $\xi_1, \ldots, \xi_r$ нельзя восстановить многомерное распределение. Такое восстановление можно сделать только в случае, когда $\xi_1, \ldots, \xi_r$ независимы. При построении моделей случайных явлений это обстоятельство используется, когда известно, что случайные явления, связанные со случайными величинами, причинно независимы. Тогда, как указывалось в § 3, при построении общего вероятностного пространства, на котором определены все ξ_i, естественно считать эти случайные величины независимыми в теоретико-вероятностном смысле.

Независимые случайные величины обладают следующим общим свойством.

Теорема 1. *Если случайные величины $\eta_1, \ldots, \eta_r$ являются функциями $\eta_i = g_i(\xi_i)$ от независимых случайных величин $\xi_1, \ldots, \xi_r$, то они независимы.*

Доказательство. Рассмотрим числовые множества B_i, для которых $\{\eta_i \in B_i\}$ являются событиями. Обозначим B'_i такое числовое множество, что $y_i = g_i(x_i) \in B'_i$ в тех и только тех точках x_i, для которых $x_i \in B'_i$. Тогда

$$\{\eta_i \in B_i\} = \{g(\xi_i) \in B_i\} = \{\xi_i \in B'_i\}.$$

Так как $\xi_1, \ldots, \xi_r$ независимы, то согласно (10) события $\{\xi_i \in B'_i\}$ независимы; следовательно, независимы события $\{\eta_i \in B_i\}$, а это означает, что случайные величины $\eta_1, \ldots, \eta_r$ — независимы. Теорема доказана.

6.3. Свертка распределений. Если задано дискретное или непрерывное многомерное распределение $\mathbf{P}_{\xi_1 \ldots \xi_r}(B)$, то с помощью формул (2) или (3) можно получить распределение случайной величины $\eta = g(\xi_1, \ldots, \xi_r)$, где $g(x_1, \ldots, x_r)$ — числовая функция от $x_1, \ldots, x_r$. Рассмотрим важный частный случай, когда случайные величины ξ_1, ξ_2 независимы, и $g(x_1, x_2) = x_1 + x_2$. Если имеются плотности $p_{\xi_1}(x_1)$ и $p_{\xi_2}(x_2)$, то в силу независимости $p_{\xi_1\xi_2}(x_1, x_2) = p_{\xi_1}(x_1)\, p_{\xi_2}(x_2)$. Вычислим по формуле (3) функцию распределения суммы $\xi_1 + \xi_2$, положив

$$F_{\xi_1+\xi_2}(z) = P\{\xi_1 + \xi_2 \leqslant z\} = \iint\limits_{x_1+x_2 \leqslant z} p_{\xi_1}(x_1)\, p_{\xi_2}(x_2)\, dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} p_{\xi_1}(x_1)\, dx_1 \int\limits_{-\infty}^{z-x_1} p_{\xi_2}(x_2)\, dx_2.$$

Производя во внутреннем интеграле замену $x_2 = u - x_1$ и меняя порядок интегрирования, получаем

$$F_{\xi_1+\xi_2}(z) = \int\limits_{-\infty}^{z} du \int\limits_{-\infty}^{\infty} p_{\xi_1}(x_1)\, p_{\xi_2}(u - x_1)\, dx_1.$$

Дифференцируя это равенство по z, приходим к формуле *свертки*, или *композиции для плотностей*:

$$p_{\xi_1+\xi_2}(z)=\int\limits_{-\infty}^{\infty} p_{\xi_1}(x_1)\,p_{\xi_2}(z-x_1)\,dx_1. \quad (14)$$

С помощью формулы (14) можно по плотностям двух независимых случайных величин ξ_1 и ξ_2 находить плотность их суммы $\xi_1+\xi_2$. При пользовании формулой (14) надо помнить, что плотности часто задаются разными аналитическими формулами на разных участках.

Пример 3. Студент при поездке в институт пользуется метро и автобусом. В метро ему приходится ожидать поезда не более двух минут, ожидание автобуса продолжается не более десяти минут. Считая времена ожидания ξ и η в метро и автобусе независимыми случайными величинами, распределенными равномерно соответственно в интервалах $[0, 2]$ и $[0, 10]$, найти плотность распределения суммарного ожидания $\xi+\eta$. Плотности $p_\xi(x)$, $p_\eta(x)$ задаются формулами

$$p_\xi(x)=\begin{cases} 1/2, & 0\leqslant x\leqslant 2,\\ 0 & \text{в остальных случаях;}\end{cases}$$

$$p_\eta(x)=\begin{cases} 1/10, & 0\leqslant x\leqslant 10,\\ 0 & \text{в остальных случаях.}\end{cases}$$

По формуле (14) имеем

$$p_{\xi+\eta}(x)=\int\limits_{-\infty}^{\infty} p_\xi(u)\,p_\eta(x-u)\,du=\frac{1}{2}\int\limits_0^2 p_\eta(x-u)\,du,$$

так как $p_\xi(u)=0$ при $u<0$ и $u>2$. При $x<0$ $p_\eta(x-u)=0$, поэтому $p_{\xi+\eta}(x)=0$. При $0\leqslant x\leqslant 2$

$$p_{\xi+\eta}(x)=\frac{1}{2}\int\limits_0^x p_\eta(x-u)\,du=\frac{1}{2}\cdot\frac{1}{10}\int\limits_0^x du=\frac{x}{20}.$$

При $2\leqslant x\leqslant 10$

$$p_{\xi+\eta}(x)=\frac{1}{2}\int\limits_0^2 p_\eta(x-u)\,du=\frac{2}{20}=\frac{1}{10}.$$

При $10 \leqslant x \leqslant 12$

$$p_{\xi+\eta}(x) = \frac{1}{2}\int\limits_{x-10}^{2} p_\eta(x-u)\,du = \frac{12-x}{20}.$$

И, наконец, при $x > 12$ получаем $p_{\xi+\eta}(x) = 0$. График плотности $p_{\xi+\eta}(x)$ изображен на рис. 12.

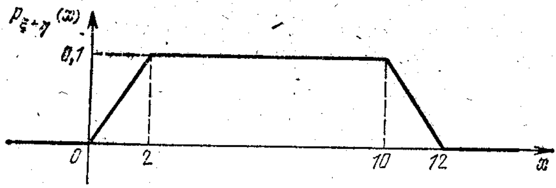


Рис. 12.

Для независимых случайных величин ξ_1, ξ_2 с дискретным распределением имеется формула свертки или композиции, аналогичная (14):

$$\mathsf{P}\{\xi+\eta=x\} = \sum_{x_i} \mathsf{P}\{\xi = x_i\}\,\mathsf{P}\{\eta = x - x_i\}, \quad (15)$$

где x_i — точки, в которых $\mathsf{P}\{\xi = x_i\} > 0$. Доказательство (15) просто получается из формулы (2) и условия независимости (13).

Пример 4. Пусть случайные величины ξ_1 и ξ_2 независимы и распределены нормально с параметрами (a_1, σ_1), (a_2, σ_2) соответственно. Найдем закон распределения $\xi = \xi_1 + \xi_2$. Подставляя в формулу (14) плотности распределения ξ_1 и ξ_2, получим

$$p_\xi(x) = \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(u-a_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{(x-u-a_2)^2}{2\sigma_2^2}}\,du.$$

Отсюда, так как

$$\frac{(u-a_1)^2}{\sigma_1^2} + \frac{(x-u-a_2)^2}{\sigma_2^2} =$$

$$= \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_1^2}\left(u - \frac{a_1\sigma_2^2 + (x-a_2)\sigma_1^2}{\sigma_1\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2 + \frac{(x-a_1-a_2)^2}{\sigma_1^2+\sigma_2^2},$$

следует, что

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \cdot \frac{\sigma}{\sqrt{2\pi}\,\sigma_1\sigma_2} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2(u-C)^2}{2\sigma_1^2\sigma_2^2}}\, du,$$

где $a = a_1 + a_2$, $\sigma^2 = \sigma_1^2 + \sigma_2^2$ и

$$C = \frac{a_1\sigma_2^2 + (x - a_2)\,\sigma_1^2}{\sigma_1\sigma_2\sqrt{\sigma_1^2 + \sigma_2^2}}.$$

Заменяя переменную интегрирования по формуле $y = \sigma(u - C)/\sigma_1\sigma_2$, получим

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \qquad (16)$$

так как

$$\frac{\sigma}{\sqrt{2\pi}\,\sigma_1\sigma_2} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2(u-C)^2}{2\sigma_1^2\sigma_2^2}}\, du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, du = 1.$$

Полученная формула (16) определяет плотность нормального распределения с параметрами (a, σ). Таким образом, *сумма независимых нормально распределенных величин имеет нормальное распределение.*

Пример 5. Пусть случайные величины ξ_1 и ξ_2 независимы и распределены по закону Пуассона:

$$\mathsf{P}\{\xi_1 = k\} = \frac{\lambda_1^k}{k!} e^{-\lambda_1}, \quad \mathsf{P}\{\xi_2 = l\} = \frac{\lambda_2^l}{l!} e^{-\lambda_2}. \qquad (17)$$

Подставляя в формулу (15) вероятности (17), находим

$$\mathsf{P}\{\xi_1 + \xi_2 = m\} = \sum_{n=0}^{m} \frac{\lambda_1^n}{n!} e^{-\lambda_1} \cdot \frac{\lambda_2^{m-n}}{(m-n)!} e^{-\lambda_2} =$$

$$= \frac{1}{m!} e^{-(\lambda_1+\lambda_2)} \sum_{n=0}^{m} \frac{m!}{n!\,(m-n)!} \lambda_1^n \lambda_2^{m-n} =$$

$$= \frac{1}{m!} e^{-(\lambda_1+\lambda_2)} \cdot (\lambda_1 + \lambda_2)^m.$$

Отсюда, полагая $\lambda = \lambda_1 + \lambda_2$, получим

$$\mathsf{P}\{\xi_1 + \xi_2 = m\} = \frac{\lambda^m}{m!} e^{-\lambda},$$

т. е. *сумма независимых случайных величин, распределенных по закону Пуассона, также распределена по закону Пуассона.*

Задачи

1. Двумерная плотность $p_{\xi\eta}(x, y)$ для точек, удовлетворяющих условиям $x + y \leqslant 1$, $x \geqslant 0$, $y \geqslant 0$, равна 2.

а) Найти вероятность события $\xi^2 + \eta^2 \leqslant 1/2$.

б) Показать, что ξ и η зависимы. (Указание. Найти вероятность событий $\{1/2 \leqslant \xi \leqslant 1,\ 1/2 \leqslant \eta \leqslant 1\}$, $\{1/2 \leqslant \xi \leqslant 1\}$, $\{1/2 \leqslant \eta \leqslant 1\}$).

2. Грани игральной кости занумерованы так, что сумма очков противоположных граней всегда равна 7. Пусть ξ — число очков, выпавшее на верхней грани, η — число очков на нижней грани. Найти закон распределения $\zeta = \xi\eta$.

3. Найти функцию и плотность распределения $\theta = \max(\xi, \eta)$, если ξ и η независимы и равномерно распределены на $[0,1]$.

4. Случайные величины ξ и η независимы и имеют одно и то же показательное распределение с плотностью $p_\xi(x) = p_\eta(x) = \lambda e^{-\lambda x}$; $x \geqslant 0$.

Найти плотность $p_{\xi+\eta}(x)$.

§ 7. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ

7.1. Математическое ожидание в конечной схеме. Пусть на конечном вероятностном пространстве $(\Omega, \mathscr{A}, \mathsf{P})$ задана случайная величина $\xi = \xi(\omega)$. Одна из наиболее важных числовых характеристик случайной величины ξ — это ее математическое ожидание.

Математическим ожиданием случайной величины ξ в конечной схеме называется число, обозначаемое $\mathsf{M}\xi$ и равное

$$\mathsf{M}\xi = \sum_{\omega \in \Omega} \xi(\omega)\, p(\omega), \tag{1}$$

где $p(\omega)$ — элементарные вероятности, задающие вероятность P.

Формула (1) показывает, что математическое ожидание $\mathsf{M}\xi$ определяется как число, равное сумме произведений значений $\xi(\omega)$ случайной величины ξ на соответствующие элементарные вероятности $p(\omega)$. Эта сумма берется по всем элементарным событиям ω из пространства Ω. В частности, если занумеровать элементарные события и положить $\Omega = \{\omega_1, \ldots, \omega_n\}$, то формулу (1) можно записать в виде

$$\mathsf{M}\xi = \xi(\omega_1)\,p(\omega_1) + \xi(\omega_2)\,p(\omega_2) + \ldots + \xi(\omega_n)\,p(\omega_n).$$

Вместо термина математическое ожидание случайной величины ξ употребляют иногда термин *среднее значение* ξ или, еще короче, — *среднее* ξ.

Пример 1. Пусть в лотерее разыгрываются N билетов, причем на N_1 билетов падает выигрыш a_1, на N_2 билетов — выигрыш a_2, ..., на N_r билетов — выигрыш a_r. (Например, в книжной лотерее на каждые 1000 билетов бо́льшая часть билетов имеет выигрыш $a_1 = 0$ руб., а остальные билеты дают выигрыши в размерах $a_2 = 0{,}5$ руб., $a_3 = 1$ руб., $a_4 = 5$ руб., и т. д.). Рассмотрим случайную величину ξ, равную выигрышу, который достался участнику лотереи, приобретшему один билет. Существуют разные системы разыгрывания лотереи. Например, в денежно-вещевой лотерее продаются билеты, имеющие номера, а затем в специальных тиражах разыгрываются выигрыши, падающие на те или иные номера. Мы для простоты рассмотрим другую систему, принятую в книжной лотерее и в лотерее «Спринт», в которых билеты продаются в запечатанном виде и на них заранее отмечен размер выигрыша. Поскольку внешне билеты выглядят все одинаково, мы можем в этом случае определить вероятностное пространство следующим образом. Пусть элементарное событие ω — это отдельный лотерейный билет, а множество всех элементарных событий $\Omega = \{\omega\}$ имеет мощность $|\Omega| = N$. Элементарные вероятности $p(\omega)$ в силу внешней идентичности лотерейных билетов все равны между собой $p(\omega) = 1/N$, поэтому мы приходим к классическому определению вероятности P, по которому $\mathsf{P}(A) = \frac{|A|}{|\Omega|}$. Пусть на множестве A_i лотерейных билетов, имеется выигрыш a_i и $|A_i| = N_i$. Тогда случайную величину $\xi = \xi(\omega)$ можно задать с помощью разбиения

$$\Omega = A_1 + A_2 + \ldots + A_r$$

равенствами

$$\xi(\omega) = a_i, \text{ если } \omega \in A_i.$$

В этом случае математическое ожидание $\mathsf{M}\xi$ определяется по формуле (1) следующим образом:

$$\mathsf{M}\xi = \sum_{\omega \in \Omega} \xi(\omega) p(\omega) =$$

$$= \sum_{\omega \in A_1} \xi(\omega) p(\omega) + \sum_{\omega \in A_2} \xi(\omega) p(\omega) + \ldots + \sum_{\omega \in A_r} \xi(\omega) p(\omega) =$$

$$= a_1 \sum_{\omega \in A_1} p(\omega) + a_2 \sum_{\omega \in A_2} p(\omega) + \ldots + a_r \sum_{\omega \in A_r} p(\omega) =$$

$$= a_1 \frac{N_1}{N} + a_2 \frac{N_2}{N} + \ldots + a_r \frac{N_r}{N},$$

так как $\xi(\omega) = a_i$ для всех $\omega \in A_i$. Из приведенного примера ясно происхождение термина *среднее значение* ξ, так как сумма $\sum_{k=1}^{r} a_k N_k$ — это выигрышный фонд всей лотереи, а ее отношение к общему числу билетов N — это средний выигрыш на один билет.

Из определения математического ожидания легко вытекают следующие его свойства.

1°. *Аддитивность*: $\mathsf{M}(\xi + \eta) = \mathsf{M}\xi + \mathsf{M}\eta$, т. е. *математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых.*

Доказательство. Из определения математического ожидания имеем

$$\mathsf{M}(\xi + \eta) = \sum_{\omega \in \Omega} (\xi(\omega) + \eta(\omega)) p(\omega) =$$

$$= \sum_{\omega \in \Omega} \xi(\omega) p(\omega) + \sum_{\omega \in \Omega} \eta(\omega) p(\omega) =$$

$$= \mathsf{M}\xi + \mathsf{M}\eta.$$

Это свойство распространяется на случай любого конечного числа слагаемых:

$$\mathsf{M}(\xi_1+\xi_2+\ldots+\xi_k)=\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_k.$$

2°. *Для любого числа c*

$$\mathsf{M}(c\xi)=c\mathsf{M}\xi,$$

т. е. *постоянный множитель c можно выносить за знак математического ожидания.*

Доказательство. Из определения 1 имеем

$$\mathsf{M}(c\xi)=\sum_{\omega\in\Omega}c\xi(\omega)p(\omega)=c\sum_{\omega\in\Omega}\xi(\omega)p(\omega)=c\mathsf{M}\xi.$$

Совокупность свойств 1° и 2° называется *свойством линейности* математического ожидания и выражается следующим равенством:

$$\mathsf{M}(c_1\xi_1+c_2\xi_2+\ldots+c_k\xi_k)=$$
$$=c_1\mathsf{M}\xi_1+c_2\mathsf{M}\xi_2+\ldots+c_k\mathsf{M}\xi_k, \quad (2)$$

справедливым для любых случайных величин ξ_i и любых чисел c_i.

3°. *Математическое ожидание индикатора χ_A события A равно вероятности этого события:*

$$\mathsf{M}\chi_A=\mathsf{P}(A).$$

Доказательство. Так как $\mathsf{P}(A)=\sum_{\omega\in A}p(\omega)$, то

$$\mathsf{M}\chi_A=\sum_{\omega\in\Omega}\chi_A(\omega)p(\omega)=\sum_{\omega\in A}1\cdot p(\omega)+\sum_{\omega\notin A}0\cdot p(\omega)=$$
$$=\sum_{\omega\in A}p(\omega)=\mathsf{P}(A).$$

4°. *Свойство монотонности: Если $\xi\geqslant\eta$, то $\mathsf{M}\xi\geqslant\mathsf{M}\eta$.*

Доказательство. Докажем сначала, что из $\xi\geqslant 0$ следует $\mathsf{M}\xi\geqslant 0$. В самом деле, так как для каждого $\omega\in\Omega$ $\xi(\omega)\geqslant 0$ и $p(\omega)\geqslant 0$, то и сумма $\mathsf{M}\xi=\sum_{\omega\in\Omega}\xi(\omega)p(\omega)$ будет неотрицательна. Применяя доказанное свойство к неотрицательной разности $\xi-\eta\geqslant 0$, получаем $\mathsf{M}(\xi-\eta)=\mathsf{M}\xi-\mathsf{M}\eta\geqslant 0$, что и требовалось доказать.

Формулы для вычисления математического ожидания

Данное выше определение математического ожидания $\mathsf{M}\xi$ случайной величины ξ не всегда удобно для вычисления. Как мы уже видели в § 6, для вычисления вероятностей $\mathsf{P}\{\xi \in B\}$ нам не обязательно знать вероятностное пространство $(\Omega, \mathscr{A}, \mathsf{P})$ и функцию $\xi(\omega)$, которая задает случайную величину ξ, а достаточно знать лишь закон распределения ξ, т. е. набор всех значений $x_1, \ldots, x_k$, которые принимает случайная величина ξ, и соответствующие вероятности $\mathsf{P}\{\xi = x_i\}$, $i = 1, 2, \ldots, k$. Математическое ожидание $\mathsf{M}\xi$ вычисляется с помощью закона распределения ξ по следующей формуле:

$$\mathsf{M}\xi = \sum_{i=1}^{k} x_i \mathsf{P}\{\xi = x_i\}. \tag{3}$$

Для доказательства формулы (3) воспользуемся представлением случайной величины ξ через разбиение $A_1 + A_2 + \ldots + A_k = \Omega$, где $A_i = \{\omega\colon \xi(\omega) = x_i\}$ (см. § 6):

$$\xi = \sum_{i=1}^{k} x_i \chi_{A_i}. \tag{4}$$

Применяя к сумме (4) свойство линейности и свойство 3°, имеем

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

Таким образом, мы доказали, что математическое ожидание случайной величины ξ равно сумме произведений ее значений x_i на вероятности $\mathsf{P}\{\xi = x_i\}$ того, что ξ принимает соответствующее значение.

Рассуждая аналогично, нетрудно получить следующие формулы для вычисления математического ожидания от случайных величин вида $g(\xi)$, $g(\xi, \eta)$, где $g(x)$, $g(x, y)$ — числовые функции, а ξ, η — случайные величины:

$$\mathsf{M}g(\xi) = \sum_{i=1}^{k} g(x_i) \mathsf{P}\{\xi = x_i\}, \tag{5}$$

$$\mathsf{M}g(\xi, \eta) = \sum_{i=1}^{k} \sum_{j=1}^{m} g(x_i, y_i) \mathsf{P}\{\xi = x_i, \eta = y_j\}. \qquad (6)$$

Для получения формул (5) и (6) надо воспользоваться представлениями $g(\xi)$ и $g(\xi, \eta)$ в виде следующих сумм:

$$g(\xi) = \sum_{i=1}^{k} g(x_i) \chi_{A_i}, \quad g(\xi, \eta) = \sum_{i=1}^{k} \sum_{j=1}^{m} g(x_i, y_j) \chi_{A_{ij}},$$

где $\{A_i\}$ — разбиение, введенное выше, а $\{A_{ij}\}$ — разбиение, образуемое множествами $A_{ij} = \{\omega: \xi(\omega) = x_i, \eta(\omega) = y_j\}$. Формулу, аналогичную (6), нетрудно получить и для $\mathsf{M}g(\xi_1, \ldots, \xi_r)$. Поскольку согласно формуле (3) математическое ожидание $\mathsf{M}\xi$ случайной величины ξ однозначно определяется ее законом распределения, то иногда вместо термина математическое ожидание случайной величины ξ употребляют термин *математическое ожидание закона распределения* ξ.

Пример 2. Найдем математическое ожидание биномиального закона распределения

$$\mathsf{P}\{\mu = k\} = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n, \quad q = 1 - p, \ 0 \leqslant p \leqslant 1.$$

Согласно (3), имеем

$$\mathsf{M}\mu = \sum_{k=0}^{n} k \mathsf{P}\{\mu = k\} = \sum_{k=0}^{n} k C_n^k p^k q^{n-k}.$$

В силу равенства $kC_n^k = k \dfrac{n!}{k!\,(n-k)!} = nC_{n-1}^{k-1}$, сумма справа вычисляется:

$$\sum_{k=0}^{n} kC_n^k p^k q^{n-k} = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-k} = np,$$

так как $\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-k} = (p+q)^{n-1} = 1$. Таким образом, $\mathsf{M}\mu = np$.

Мультипликативное свойство. Для независимых случайных величин ξ, η имеет место мультипликативное свойство математического ожидания:

$$\mathsf{M}\xi\eta = \mathsf{M}\xi \cdot \mathsf{M}\eta, \qquad (7)$$

т. е. *математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий.*

Доказательство. Воспользуемся формулой (6):

$$\mathsf{M}\xi\eta = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i y_j \mathsf{P}\{\xi = x_i,\ \eta = y_j\}.$$

Для независимых ξ и η имеет место равенство $\mathsf{P}\{\xi = x_i, \eta = y_j\} = \mathsf{P}\{\xi = x_i\}\mathsf{P}\{\eta = y_j\}$, поэтому

$$\mathsf{M}\xi\eta = \sum_{i=1}^{k} \sum_{j=1}^{m} x_i x_j \mathsf{P}\{\xi = x_i\}\, \mathsf{P}\{\eta = y_j\} =$$

$$= \sum_{i=1}^{k} x_i \mathsf{P}\{\xi = x_i\} \sum_{j=1}^{m} y_j \mathsf{P}\{\eta = y_j\} = \mathsf{M}\xi\mathsf{M}\eta.$$

Свойство мультипликативности естественно распространяется на случай произвольного конечного числа независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_r$:

$$\mathsf{M}\xi_1\xi_2 \ldots \xi_r = \mathsf{M}\xi_1\mathsf{M}\xi_2 \ldots \mathsf{M}\xi_r. \qquad (8)$$

Следует отметить, что если свойство аддитивности математического ожидания справедливо для любых случайных величин, то свойство мультипликативности математического ожидания справедливо для независимых случайных величин.

Способы вычисления математических ожиданий. При вычислении $\mathsf{M}\xi$ или $\mathsf{M}g(\xi)$ можно пользоваться либо формулой (1), дающей определение математического ожидания, либо формулами (3), (5) и (6), выражающими математическое ожидание через закон распределения. Однако иногда случайная величина ξ или функция от нее $g(\xi)$ заданы таким образом, что соответствующий закон распределения либо неизвестен, либо очень сложен. Иногда закон распределения $\mathsf{P}\{\xi = x_i\}$ задается хорошей формулой, однако сумма в (3), через которую задается $\mathsf{M}\xi$, очень сложна и с трудом поддается упрощению. В этих случаях часто можно быстро и просто вычислить $\mathsf{M}\xi$ с помощью свойства аддитивности. Для этого надо представить ξ в виде суммы более простых слагаемых $\xi_1 + \xi_2 + \ldots + \xi_r$, например индикаторов некоторых событий, вероятности которых известны. Тогда $\mathsf{M}\xi = \mathsf{M}\xi_1 + \mathsf{M}\xi_2 + \ldots + \mathsf{M}\xi_r$.

Пример 3. Вычислим математическое ожидание биномиального распределения предложенным выше способом. Для этого воспользуемся тем, что биномиальное распределение имеет случайная величина μ, равная числу успехов в n испытаниях в схеме Бернулли с вероятностью успеха p в каждом испытании. Обозначим A_i событие, состоящее в том, что в i-м испытании произошел успех. Тогда

$$\mu = \chi_{A_1} + \chi_{A_2} + \ldots + \chi_{A_n},$$

и

$$\mathsf{M}\mu = \sum_{i=1}^{n} \mathsf{M}\chi_{A_i} = \sum_{i=1}^{n} \mathsf{P}(A_i) = np,$$

т. е. мы получили тот же результат, что и в примере 2, но гораздо проще и прозрачней.

Пример 4. Вычислим $\mathsf{M}\xi$ в гипергеометрическом распределении

$$\mathsf{P}\{\xi = k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad k = 0, 1, \ldots, n.$$

Воспользуемся тем же приемом, что и в примере 3. Гипергеометрическое распределение появляется в урновой схеме, в которой из урны, содержащей M белых шаров и $(N - M)$ — черных, производится бесповторная выборка объема n. Случайная величина ξ — это число белых шаров в выборке. Будем представлять себе, что шары вынимаются из урны последовательно. Обозначим событие $A_i = \{i\text{-й шар выборки белый}\}$. Тогда

$$\xi = \chi_{A_1} + \chi_{A_2} + \ldots + \chi_{A_n},$$

и по свойству аддитивности $\mathsf{M}\xi = \sum_{i=1}^{n} \mathsf{P}(A_i)$. Так как $\mathsf{P}(A_i) = \frac{M}{N}$ при любом i (см. пример 5 в § 1), то $\mathsf{M}\xi = n\frac{M}{N}$.

Иногда при вычислении математических ожиданий полезно использовать и свойство мультипликативности.

Пример 5. В схеме Бернулли с n испытаниями и с вероятностью успеха p будем полагать, что в i-м испытании произошел двойной успех, если успехи были в $(i-1)$-м и в i-м испытаниях. Обозначим ν_n число двойных успехов во всей серии испытаний. Найдем $\mathsf{M}\nu_n$. Представим $\nu_n = \eta_2 + \eta_3 + \ldots + \eta_n$, где $\eta_i = 1$, если в i-м испытании был двойной успех и $\eta_i = 0$ в противоположном случае. Нетрудно видеть, что $\eta_i = \chi_{A_{i-1}} \cdot \chi_{A_i}$, где A_i — события из примера 3. Имеем из свойства аддитивности

$$\mathsf{M}\nu_n = \sum_{i=2}^{n} \mathsf{M}\chi_{A_{i-1}}\chi_{A_i}.$$

Так как $\chi_{A_{i-1}}$ и χ_{A_i} независимы, то из свойства мультипликативности $\mathsf{M}\chi_{A_{i-1}}\chi_{A_i} = \mathsf{M}\chi_{A_{i-1}} \cdot \mathsf{M}\chi_{A_i} = p^2$. Окончательно имеем $\mathsf{M}\nu_n = (n-1)\,p^2$.

7.2. Математическое ожидание в счетной схеме. Рассмотрим вероятностное пространство $(\Omega, \mathcal{A}, \mathsf{P})$ со счетным множеством Ω элементарных событий. Занумеруем элементарные события ω_n, $n = 1, 2, \ldots$ Пусть вероятность P определяется через элементарные вероятности $p(\omega_n)$.

Математическим ожиданием $\mathsf{M}\xi$ *случайной величины* ξ *в счетной схеме называется сумма ряда*

$$\mathsf{M}\xi = \sum_{n=1}^{\infty} \xi(\omega_n)\, p(\omega_n), \tag{9}$$

если этот ряд сходится абсолютно, т. е. если

$$\sum_{n=1}^{\infty} |\xi(\omega_n)| \cdot p(\omega_n) < \infty. \tag{10}$$

В противном случае мы будем говорить, что $\mathsf{M}\xi$ не существует.

Условие (10) очень важно. В самом деле, если ряд (9) сходится неабсолютно, то, как известно из математического анализа, ряд, полученный перестановкой членов первоначального ряда, может сходиться к любому числу, а так как нумерация элементарных событий ω_n произвольна и никак не связана с существом рассматриваемой модели, то и сумма (9) в этом случае ничего не выражает и теряет объективный смысл. Поэтому мы полагаем, что в этом случае математическое ожидание не существует.

Сумма абсолютно сходящегося ряда (9) по своим свойствам аналогична конечной сумме (1), поэтому свойства математического ожидания $\mathsf{M}\xi$, установленные в п. 7.1 для случайных величин ξ в конечной схеме, остаются справедливыми и в случае счетной схемы. В частности, справедлива формула, аналогичная (3):

$$\mathsf{M}\xi = \sum_{n=1}^{\infty} x_n \mathsf{P}\{\xi = x_n\}, \tag{11}$$

если ряд справа абсолютно сходится. Формула (11) выражает $\mathsf{M}\xi$ через закон распределения ξ.

Пример 6. Вычислим $\mathsf{M}\xi$ случайной величины ξ, распределенной по закону Пуассона $\mathsf{P}\{\xi = k\} = \frac{a^k}{k!} e^{-a}$:

$$\mathsf{M}\xi = \sum_{k=0}^{\infty} k\mathsf{P}\{\xi = k\} = \sum_{k=1}^{\infty} k \frac{a^k e^{-a}}{k!} = a \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} e^{-a} = a.$$

Формулы (5) и (6) также распространяются на счетный случай:

$$\mathsf{M}g(\xi) = \sum_{n=1}^{\infty} g(x_n) \mathsf{P}\{\xi = x_n\}, \tag{12}$$

$$\mathsf{M}g(\xi, \eta) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} g(x_n, y_m) \mathsf{P}\{\xi = x_n, \eta = y_m\}; \tag{13}$$

для их справедливости требуется, чтобы ряды сходились абсолютно.

7.3. Математическое ожидание в общем случае. Данные выше определения математического ожидания $\mathsf{M}\xi$ для случайных величин в конечной или счетной схеме легко переносятся на случайные величины, принимающие конечное или счетное число значений и определенные в любом вероятностном пространстве $(\Omega, \mathcal{A}, \mathsf{P})$. Пусть $A_1, \ldots, A_k$ — конечное разбиение, т. е. события A_i попарно несовместны и $\sum_{i=1}^{k} A_i = \Omega$. Тогда случайная величина

$$\xi = \sum_{i=1}^{k} x_i \chi_{A_i} \tag{14}$$

принимает лишь конечное число значений. Математическое ожидание $\mathsf{M}\xi$ такой случайной величины определяется как сумма

$$\mathsf{M}\xi = \sum_{i=1}^{k} x_i \mathsf{P}(A_i). \tag{15}$$

В формулах (14) и (15) мы не налагаем никаких ограничений на значения x_i. Если все x_i различны, то формула (15) переходит в формулу (3), так как в этом случае $A_i = \{\xi = x_i\}$.

Аналогично любая дискретная случайная величина записывается в виде

$$\xi = \sum_{i=1}^{\infty} x_i \chi_{A_i}, \tag{16}$$

где $\{A_i\}$ — счетное разбиение, т. е. A_i — попарно несовместны и $\sum_{i=1}^{\infty} A_i = \Omega$. В этом случае математическое ожидание $\mathsf{M}\xi$ определяется как сумма ряда

$$\mathsf{M}\xi = \sum_{i=1}^{\infty} x_i \mathsf{P}(A_i), \tag{17}$$

если этот ряд сходится абсолютно. В противном случае математическое ожидание не существует. Если все x_i различны, то определение (17) превращается в определение (11), позволяющее вычислять $\mathsf{M}\xi$ по закону распределения. Аналогично можно показать, что для дискретных случайных величин справедливы формулы (12) и (13). Все свойства математических ожиданий, установленные в п. 7.1, остаются справедливыми и в общем случае определения $\mathsf{M}\xi$ формулами (15) и (17).

Если случайная величина ξ непрерывна и ее закон распределения задается плотностью $p_\xi(x)$, то ее математическое ожидание вычисляется по формуле

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\,dx, \tag{18}$$

если интеграл сходится абсолютно. Эта формула аналогична формуле (11) для $\mathsf{M}\xi$ в дискретном случае. Формулы (12) и (13) для вычисления $\mathsf{M}g(\xi)$ и $\mathsf{M}g(\xi, \eta)$ переходят в случае непрерывных случайных величин в формулы

$$\mathsf{M}g(\xi) = \int_{-\infty}^{\infty} g(x)\,p_\xi(x)\,dx, \tag{19}$$

$$\mathsf{M}g(\xi, \eta) = \iint_{-\infty}^{\infty} g(x, y)\,p_{\xi\eta}(x, y)\,dx\,dy, \tag{20}$$

если интегралы справа абсолютно сходятся. В (20) $p_{\xi\eta}(x, y)$ — совместная плотность случайных величин ξ и η.

З а м е ч а н и е. Формулы (6), (13) и (20), выписанные для двумерных распределений, имеют аналоги для любых многомерных распределений.

Определение $\mathsf{M}\xi$ для непрерывной случайной величины формулой (18) неудобно, так как в этом случае нельзя доказать свойства математического ожидания. Более логично определить $\mathsf{M}\xi$ следующим образом. С каждой случайной величиной ξ можно связать последовательность дискретных случайных величин ξ_n, полагая

$$\xi_n = \frac{k}{2^n}, \quad \text{если} \quad \frac{k}{2^n} < \xi \leqslant \frac{k+1}{2^n};$$

нетрудно видеть, что $\xi \leqslant \xi_n \leqslant \xi + \frac{1}{2^n}$, поэтому $\lim\limits_{n\to\infty} \xi_n = \xi$. Случайные величины ξ_n можно записать в виде, подобном (16):

$$\xi_n = \sum_{k=-\infty}^{\infty} \frac{k}{2^n}\,\chi_{\left\{\frac{k}{2^n} < \xi \leqslant \frac{k+1}{2^n}\right\}}.$$

Применяя к ξ_n формулу (17), получаем

$$\mathsf{M}\xi_n = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} P\left\{\frac{k}{2^n} < \xi \leqslant \frac{k+1}{2^n}\right\}, \tag{21}$$

если ряд справа абсолютно сходится. Нетрудно видеть, что $\xi_n \leqslant \xi_{n+1}$ при любом n, поэтому $\mathsf{M}\xi_n \leqslant \mathsf{M}\xi_{n+1}$ и существует предел $\mathsf{M}\xi_n$, значение которого мы и принимаем за определение математического ожидания:

$$\mathsf{M}\xi = \lim_{n\to\infty} \mathsf{M}\xi_n. \tag{22}$$

Из определения (22) нетрудно доказать формулу (18). В самом деле, полагая $x_{nk} = \frac{k}{2^n}$ можно записать (21) следующим образом:

$$\mathsf{M}\xi_n = \sum_{k=-\infty}^{\infty} \frac{k}{2^n} \int\limits_{x_{nk}}^{x_{n,\,k+1}} p_\xi(u)\,du,$$

откуда

$$0 \leqslant \int\limits_{-\infty}^{\infty} up_\xi(u)\,du - \mathsf{M}\xi_n =$$

$$= \sum_{k=-\infty}^{\infty} \int\limits_{x_{nk}}^{x_{n,\,k+1}} \left(u - \frac{k}{2^n}\right) p_\xi(u)\,du \leqslant \frac{1}{2^n} \int\limits_{-\infty}^{\infty} p_\xi(u)\,du = \frac{1}{2^n},$$

и $\lim\limits_{n\to\infty} \mathsf{M}\xi_n = \int\limits_{-\infty}^{\infty} up_\xi(u)\,du$. Такое определение математического ожидания $\mathsf{M}\xi$ позволяет перенести все его свойства, доказанные в п. 7.1, и на случай непрерывных случайных величин, так как справедливость этих свойств для дискретных случайных величин ξ_n устанавливается просто и переход к пределу при $n\to\infty$ эти свойства не нарушает.

Пример 7. Вычислим математическое ожидание случайной величины ξ, равномерно распределенной в $[a, b]$:

$$\mathsf{M}\xi = \int\limits_{-\infty}^{\infty} xp_\xi(x)\,dx = \frac{1}{b-a}\int\limits_a^b x\,dx = \frac{a+b}{2}.$$

Пример 8. Математическое ожидание случайной величины ξ, имеющей нормальное распределение с плотностью $p_\xi(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}$, вычисляется следующим образом:

$$M\xi=\int_{-\infty}^{\infty}xp_\xi(x)\,dx=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a+a)e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx=$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a)e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx+$$

$$+a\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx=a.$$

Пример 9. Пусть случайная величина τ — время исправной работы детали, $p(x)$ — плотность распределения τ. Вычислим среднее время работы детали, если известно, что в любом случае деталь заменяется по прошествии времени T. Время работы детали определяется значением функции $g(\tau)$, где

$$g(x)=\begin{cases}x & \text{при } x\leqslant T,\\ T & \text{при } x>T.\end{cases}$$

Применим формулу (19):

$$Mg(\tau)=\int_0^{\infty}g(x)\,p(x)\,dx=$$

$$=\int_0^T g(x)\,p(x)\,dx+T\int_T^{\infty}p(x)\,dx=$$

$$=\int_0^T g(x)\,p(x)\,dx+T\cdot P\{\tau>T\}.$$

7.4. Неравенство Чебышёва. Докажем первую форму неравенства Чебышёва для неотрицательных случайных величин ξ и для $x>0$:

$$P\{\xi>x\}\leqslant\frac{M\xi}{x}. \tag{23}$$

Для этого представим неотрицательную случайную величину ξ в виде суммы двух неотрицательных случайных величин

$$\xi = \xi \cdot \chi_{\{\xi > x\}} + \xi\chi_{\{\xi \leqslant x\}}.$$

По свойству аддитивности монотонности имеем

$$\mathsf{M}\xi = \mathsf{M}\xi\chi_{\{\xi > x\}} + \mathsf{M}\xi\chi_{\{\xi \leqslant x\}} \geqslant \mathsf{M}\xi\chi_{\{\xi > x\}}.$$

Так как $\xi\chi_{\{\xi > x\}} \geqslant x\chi_{\{\xi > x\}}$, то, применяя еще раз свойство монотонности, получаем

$$\mathsf{M}\xi \geqslant x\mathsf{M}\chi_{\{\xi > x\}} = x \cdot \mathsf{P}\{\xi > x\},$$

что и доказывает (23).

С помощью неравенства Чебышёва можно усилить свойство 4° п. 7.1. Справедливо следующее утверждение.

Теорема 1. *Если* $\mathsf{P}\{\xi \geqslant 0\} = 1$ *и* $\mathsf{M}\xi = 0$, *то* $\mathsf{P}\{\xi = 0\} = 1$.

Доказательство. Представим событие $\{\xi > 0\}$ в виде суммы попарно несовместных событий

$$\{\xi > 0\} = \{\xi > 1\} + \sum_{n=1}^{\infty} \left\{\frac{1}{n+1} < \xi \leqslant \frac{1}{n}\right\}.$$

По аксиоме А4

$$\mathsf{P}\{\xi > 0\} = \mathsf{P}\{\xi > 1\} + \sum_{n=1}^{\infty} \mathsf{P}\left\{\frac{1}{n+1} < \xi \leqslant \frac{1}{n}\right\}.$$

По неравенству Чебышёва при $\mathsf{M}\xi = 0$ имеем

$$0 \leqslant \mathsf{P}\{\xi > 1\} \leqslant \mathsf{M}\xi = 0$$

и

$$0 \leqslant \mathsf{P}\left\{\frac{1}{n+1} < \xi \leqslant \frac{1}{n}\right\} \leqslant$$
$$\leqslant \mathsf{P}\left\{\frac{1}{n+1} < \xi\right\} \leqslant (n+1) \cdot \mathsf{M}\xi = 0,$$

откуда следует $\mathsf{P}\{\xi > 1\} = \mathsf{P}\left\{\frac{1}{n+1} < \xi \leqslant \frac{1}{n}\right\} = 0$, и $\mathsf{P}\{\xi > 0\} = 0$. Поэтому $1 = \mathsf{P}\{\xi = 0\} + \mathsf{P}\{\xi > 0\} = \mathsf{P}\{\xi = 0\}$, что и требовалось доказать.

Задачи

1. Найти математическое ожидание ξ — числа очков, выпавших на игральной кости.

2. Случайно берется одна кость домино. Найти математическое ожидание суммы очков на ее половинах ξ и η. Найти $\mathsf{M}(\xi+\eta)$.

3. Время работы τ электрической лампочки до перегорания имеет показательное распределение с плотностью $p(x) = \lambda e^{-\lambda x}$ при $x \geqslant 0$ и $p(x) = 0$ при $x < 0$. Найти $\mathsf{M}\tau$.

4. Найти среднее время работы электрической лампочки, если дополнительно к условиям задачи 3 предположить, что с вероятностью p лампочка может перегореть при включении (лампочка включается один раз).

Указание. Представить время горения лампочки в виде произведения $\tau\cdot\xi$ двух независимых случайных величин, где τ — то же, что и в задаче 3, а $\xi = 0$ с вероятностью p и $\xi = 1$ с вероятностью $1 - p$.

5. Точка попадания в круглую мишень радиуса R имеет равномерное распределение. Найти математическое ожидание расстояния ρ точки попадания от центра.

6. В группе учится 25 студентов. Предполагая, что дни рождения студентов независимы и равномерно распределены по 12 месяцам года, найти математическое ожидание числа месяцев, на которые не приходится ни один день рождения.

7. Деталь имеет форму цилиндра радиуса r и высоты h. При изготовлении цилиндра на эти размеры разрешается допуск, так что радиус равен $r+\xi$, а высота $h+\eta$, причем ξ и η независимы и равномерно распределены соответственно в интервалах $[-\Delta, \Delta]$ и $[-\delta, \delta]$. Найти математическое ожидание объема цилиндра.

§ 8. ДИСПЕРСИЯ. МОМЕНТЫ

8.1. Определение дисперсии. *Дисперсией* случайной величины $\xi = \xi(\omega)$, $\omega \in \Omega$, называется число

$$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2, \tag{1}$$

т. е. дисперсия равна математическому ожиданию квадрата отклонения случайной величины от своего математического ожидания.

На практике в качестве меры разброса значений случайной величины иногда пользуются величиной $\sqrt{\mathsf{D}\xi}$, называемой *среднеквадратичным отклонением*. Заметим, что дисперсия $\mathsf{D}\xi$ определяется формулой (1), если $\mathsf{M}(\xi - \mathsf{M}\xi)^2$ существует; в противном случае мы будем говорить, что дисперсия не существует.

Из свойств линейности математического ожидания следует, что

$$\mathsf{M}(\xi-\mathsf{M}\xi)^2=\mathsf{M}[\xi^2-2\xi\cdot\mathsf{M}\xi+(\mathsf{M}\xi)^2]=$$
$$=\mathsf{M}\xi^2-2\mathsf{M}\xi\cdot\mathsf{M}\xi+(\mathsf{M}\xi)^2=\mathsf{M}\xi^2-(\mathsf{M}\xi)^2.$$

Отсюда и из (1) получается еще одна формула для вычисления дисперсии

$$\mathsf{D}\xi=\mathsf{M}\xi^2-(\mathsf{M}\xi)^2. \qquad (2)$$

Если случайная величина ξ дискретна и имеет закон распределения вероятностей $\mathsf{P}\{\xi=x_k\}$, $k=1, 2, \ldots, \sum\limits_{k=1}^{\infty}\mathsf{P}\{\xi=x_k\}=1$, то ее дисперсия вычисляется по формуле

$$\mathsf{D}\xi=\sum_{k=1}^{\infty}(x_k-\mathsf{M}\xi)^2\,\mathsf{P}\{\xi=x_k\}; \qquad (3)$$

для дисперсии непрерывной случайной величины ξ с плотностью распределения $p_\xi(x)$ имеем следующую формулу:

$$\mathsf{D}\xi=\int_{-\infty}^{+\infty}(x-\mathsf{M}\xi)^2\,p_\xi(x)\,dx. \qquad (4)$$

Формулы (3) и (4) следуют из формул (5), (12) § 7. Дадим механическую интерпретацию математического ожидания и дисперсии. Будем интерпретировать закон распределения вероятностей $p_k=\mathsf{P}\{\xi=x_k\}$, $k=1, 2, \ldots, n$, $\sum\limits_{k=1}^{n}p_k=1$, случайной величины ξ как закон распределения единичной массы на прямой: в точках x_k сосредоточены массы p_k, $k=1, 2, \ldots, n$. Тогда $\mathsf{M}\xi=\sum\limits_{k=1}^{n}x_k\mathsf{P}\{\xi=x_k\}=\sum\limits_{k=1}^{n}x_kp_k$ есть центр тяжести, а $\mathsf{D}\xi=\sum\limits_{k=1}^{n}(x_k-\mathsf{M}\xi)^2\,\mathsf{P}\{\xi=x_k\}=\sum\limits_{k=1}^{n}(x_k-\mathsf{M}\xi)^2\,p_k$ — момент инерции.

Пример 1. *Пуассоновское распределение.* Для случайной величины ξ, распределенной по закону

Пуассона с параметром a,

$$\mathsf{P}\{\xi=k\}=e^{-a}\frac{a^k}{k!},\qquad k=0,\ 1,\ 2,\ \ldots,\quad a>0,$$

известно (см. пример 6 из § 7), что $\mathsf{M}\xi=a$. Вычислим теперь $\mathsf{M}\xi^2$. Используя равенство $\xi=\xi(\xi-1)+\xi$ и свойство линейности математического ожидания, получим

$$\mathsf{M}\xi^2=\mathsf{M}[\xi(\xi-1)+\xi]=\mathsf{M}[\xi(\xi-1)]+\mathsf{M}\xi=a^2+a,$$

так как $\mathsf{M}\xi=a$ и

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

(здесь мы воспользовались равенством $\sum\limits_{l=2}^{\infty}\frac{a^{l-2}}{(l-2)!}=\sum\limits_{l=0}^{\infty}\frac{a^l}{l!}=e^a$). Подставляя $\mathsf{M}\xi^2=a^2+a$ и $\mathsf{M}\xi=a$ в формулу (2), получим $\mathsf{D}\xi=a^2+a-a^2=a$. Таким образом, для случайной величины, распределенной по закону Пуассона с параметром a, имеем

$$\mathsf{M}\xi=\mathsf{D}\xi=a.$$

Пример 2. *Равномерное распределение.* Плотность распределения случайной величины ξ, равномерно распределенной на отрезке $[a,b]$, имеет вид

$$p_\xi(x)=\begin{cases}\dfrac{1}{b-a}, & \text{если}\quad x\in[a,\ b],\\ 0, & \text{если}\quad x\notin[a,\ b].\end{cases}$$

В этом случае (см. пример 7 из § 7) $\mathsf{M}\xi=(a+b)/2$. По формуле (19) § 7, полагая $g(x)=x^2$, находим

$$\mathsf{M}\xi^2=\int_a^b\frac{x^2\,dx}{b-a}=\frac{x^3}{3(b-a)}\bigg|_a^b=\frac{b^3-a^3}{3(b-a)}=\frac{b^2+ab+a^2}{3}.$$

Подставляя значения $\mathsf{M}\xi$ и $\mathsf{M}\xi^2$ в (2), получим

$$\mathsf{D}\xi=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{(b-a)^2}{12}.$$

Пример 3. *Нормальное распределение.* Плотность распределения случайной величины ξ, распределенной нормально с параметрами (a, σ), имеет вид

$$p_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad -\infty < a < +\infty, \quad \sigma > 0.$$

В этом случае (см. пример 9 из § 7) $\mathsf{M}\xi = a$ и
$\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 =$

$$= \int_{-\infty}^{+\infty} (x-a)^2 p_\xi(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Отсюда, производя замену переменной интегрирования $(x-a)/\sigma = y$, $x = \sigma y + a$, получим

$$\mathsf{D}\xi = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}} dy = \frac{\sigma^2}{\sqrt{2\pi}} \left(- \int_{-\infty}^{+\infty} y\, de^{-\frac{y^2}{2}} \right) =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(-ye^{-\frac{y^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \right) =$$

$$= \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = \sigma^2.$$

Итак, математическое ожидание и дисперсия случайной величины ξ, распределенной нормально с параметрами (a, σ^2), связаны с этими параметрами следующим образом:

$$\mathsf{M}\xi = a, \quad \mathsf{D}\xi = \sigma^2.$$

8.2. Свойства дисперсии. Отметим основные свойства дисперсии.

1°. *Дисперсия любой случайной величины ξ неотрицательна, причем $\mathsf{D}\xi = 0$ тогда и только тогда, когда ξ — постоянная.*

Свойство неотрицательности следует из неравенства $(\xi - \mathsf{M}\xi)^2 \geqslant 0$ и свойства монотонности математического ожидания: $\mathsf{D}\xi = \mathsf{M}(\xi - \mathsf{M}\xi)^2 \geqslant 0$. Если $\xi = c$, то $\mathsf{D}c = \mathsf{M}(c - \mathsf{M}c)^2 = 0$. В силу теоремы 1 § 7 из

$D\xi = M(\xi - M\xi)^2 = 0$ вытекает, что $(\xi - M\xi)^2 = 0$, а следовательно, и $\xi = M\xi$ с вероятностью 1.

2°. *Если a — постоянная, то*

$$D(a\xi) = a^2 D\xi. \tag{5}$$

Действительно, $D(a\xi) = M(a\xi - M(a\xi))^2 = M[a(\xi - M\xi)]^2 = a^2 M(\xi - M\xi)^2 = a^2 D\xi$.

3°. *Если случайные величины ξ и η независимы, то*

$$D(\xi + \eta) = D\xi + D\eta. \tag{6}$$

Доказательство. Используя определение дисперсии (1) и свойство линейности математического ожидания, получим

$$D(\xi + \eta) = M[(\xi + \eta) - M(\xi + \eta)]^2 = M[(\xi - M\xi) + (\eta - M\eta)]^2 = M(\xi - M\xi)^2 + 2M(\xi - M\xi)(\eta - M\eta) + M(\eta - M\eta)^2.$$

Отсюда следует формула (6), так как согласно свойству мультипликативности математического ожидания

$$M(\xi - M\xi)(\eta - M\eta) = M(\xi - M\xi)\,M(\eta - M\eta) = (M\xi - M\xi)(M\eta - M\eta) = 0.$$

Формула (6) по индукции распространяется на сумму n независимых случайных величин. Если $\xi_1, \xi_2, \ldots, \xi_n$ независимы, то

$$D(\xi_1 + \xi_2 + \ldots + \xi_n) = D\xi_1 + D\xi_2 + \ldots + D\xi_n. \tag{7}$$

Вычислим дисперсии некоторых случайных величин, используя доказанные свойства дисперсии.

Пример 4. *Биномиальное распределение.* Число успехов μ_n в n испытаниях Бернулли имеет биномиальное распределение (см. (7) § 3). Воспользуемся представлением μ_n в виде суммы $\mu_n = \sum_{k=1}^{n} \xi_k$, где ξ_k, $k = 1, 2, \ldots, n$, — следующие индикаторы: $\xi_k = 1$, если в k-м испытании был успех, и $\xi_k = 0$ в противном случае. Таким образом, $P\{\xi_k = 1\} = p$, $P\{\xi_k = 0\} = q$, $p + q = 1$, $k = 1, 2, \ldots, n$. Индикаторы ξ_k независимы в силу независимости испытаний. Для независимых случайных величин согласно формуле (7) находим

$$\mathsf{D}\mu_n = \mathsf{D}\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n} \mathsf{D}\xi_k.$$

Отсюда, так как $\mathsf{M}\xi_k = 1 \cdot p + 0 \cdot q = p$, $\mathsf{M}\xi_k^2 = p$, $\mathsf{D}\xi_k = \mathsf{M}\xi_k^2 - (\mathsf{M}\xi_k)^2 = p - p^2 = p(1-p) = pq$, получим

$$\mathsf{D}\mu_n = npq.$$

Пример 5. *Гипергеометрическое распределение.* Пусть из урны, содержащей M белых и $N-M$ черных шаров, по схеме случайного выбора без возвращения (см. пример 4 из § 1) извлечено n шаров. Случайная величина η_n, равная числу белых шаров в выборке, имеет гипергеометрическое распределение (см. (5) § 1). В § 1 формула (5) была выведена для модели, в которой в качестве элементарных событий рассматривались неупорядоченные наборы извлекаемых шаров. Здесь при вычислении $\mathsf{M}\eta_n$ и $\mathsf{D}\eta_n$ удобно рассматривать более детальное вероятностное пространство с упорядоченными наборами, введенное в примере 5 § 1. При таком выборе вероятностной модели формула (5) § 1 сохранится. Введем индикаторы ξ_k, $k = 1, \ldots, n$, положив $\xi_k = 1$, если k-й шар оказался белым, и $\xi_k = 0$ — в противном случае. Очевидно, что $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Отметим, что индикаторы ξ_k в представлении μ_n (пример 4) независимы. Здесь в представлении η_n индикаторы ξ_k не являются даже попарно независимыми и, следовательно, для вычисления $\mathsf{D}\eta_n$ нельзя воспользоваться формулой (7).

Воспользуемся формулой (2). Найдем сначала $\mathsf{M}\xi_k$, $\mathsf{M}\xi_k\xi_l$. По формуле (3) § 7

$$\mathsf{M}\xi_k = 1 \cdot \mathsf{P}\{\xi_k = 1\} + 0 \cdot \mathsf{P}\{\xi_k = 0\} = \mathsf{P}\{\xi_k = 1\} = \frac{M}{N},$$

$$\mathsf{M}\xi_k^2 = \mathsf{M}\xi_k.$$

При $k \neq l$, так как

$$\mathsf{P}\{\xi_k\xi_l = 1\} = \frac{M(M-1)}{N(N-1)},$$

$$\mathsf{P}\{\xi_k\xi_l = 0\} = 1 - \frac{M(M-1)}{N(N-1)},$$

получим

$$\mathsf{M}\xi_k\xi_l=\frac{M(M-1)}{N(N-1)}.$$

Отсюда, используя свойство линейности математического ожидания, находим

$$\mathsf{M}\eta_n=\mathsf{M}\left(\sum_{k=1}^{n}\xi_k\right)=\sum_{k=1}^{n}\mathsf{M}\xi_k=n\frac{M}{N},$$

$$\mathsf{M}\eta_n^2=\mathsf{M}\left(\sum_{k=1}^{n}\xi_k\right)^2=\mathsf{M}\left[\sum_{k=1}^{n}\xi_k^2+\sum_{k\neq l}\xi_k\xi_l\right]=$$

$$=\sum_{k=1}^{n}\mathsf{M}\xi_k^2+\sum_{k\neq l}\mathsf{M}\xi_k\xi_l=n\frac{M}{N}+n(n-1)\frac{M(M-1)}{N(N-1)}$$

и, следовательно,

$$\mathsf{D}\eta_n=M\eta_n^2-(M\eta_n)^2=n\frac{M}{N}+$$
$$+n(n-1)\frac{M(M-1)}{N(N-1)}-n^2\frac{M^2}{N^2}.$$

После элементарных преобразований приходим к окончательной формуле

$$\mathsf{D}\eta_n=n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}.$$

Число белых шаров (обозначим его μ_n) в выборке, полученной по схеме случайного выбора с возвращением, имеет биномиальное распределение с параметром $p=\frac{M}{N}$. Отметим, что

$$\mathsf{M}\mu_n=\mathsf{M}\eta_n=\frac{M}{N};\quad \mathsf{D}\eta_n=\frac{N-n}{N-1}\mathsf{D}\mu_n.$$

8.3. Моменты высшего порядка. Наряду с рассмотренными выше числовыми характеристиками случайных величин (математическим ожиданием и дисперсией), часто используются и другие характеристики, называемые моментами. *Моментом порядка k, или k-м моментом случайной величины* ξ, называется число $m_k = \mathsf{M}\xi^k$, $k = 1, 2, \ldots$ Число $\mathsf{M}|\xi|^k$ называется *абсолютным моментом порядка* k, $k = 1, 2, \ldots$; числа $\mathsf{M}(\xi - \mathsf{M}\xi)^k$ и $\mathsf{M}|\xi - \mathsf{M}\xi|^k$ называются соответственно *центральным* и *абсолютным центральным моментами порядка* k, $k = 1, 2, \ldots$ (или соответствующими k-ми моментами).

Заметим, что из существования момента $\mathsf{M}\xi^m$ следует существование моментов более низких порядков $\mathsf{M}\xi^k$, $k = 1, 2, \ldots, m-1$. Это утверждение следует из неравенств

$$|\xi(\omega)|^k \leqslant |\xi(\omega)|^m + 1, \quad \omega \in \Omega, \quad k = 1, 2, \ldots, m-1.$$

Отметим, что момент первого порядка $m_1 = \mathsf{M}\xi$ является математическим ожиданием, а центральный момент второго порядка $\mathsf{M}(\xi - \mathsf{M}\xi)^2$ — это дисперсия.

В приложениях часто пользуются еще двумя характеристиками — *асимметрией* S и *эксцессом* E. Они определяются следующим образом:

$$S = \frac{1}{2}\frac{\mathsf{M}(\xi - \mathsf{M}\xi)^3}{(\mathsf{D}\xi)^{3/2}}, \quad E = \frac{1}{8}\left[\frac{\mathsf{M}(\xi - \mathsf{M}\xi)^4}{(\mathsf{D}\xi)^2} - 3\right]. \tag{8}$$

Асимметрия S и эксцесс E нормально распределенной величины равны нулю. Отличие от 0 асимметрии и эксцесса свидетельствует о том, что распределение случайной величины отлично от нормального.

Пример 6. Найдем моменты всех порядков для случайной величины ξ, распределенной нормально с параметрами $(0, \sigma^2)$. В этом случае плотность распределения вероятностей $p_\xi(x)$ имеет вид

$$p_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad -\infty < x < +\infty.$$

Все моменты нечетного порядка, как интегралы от нечетных функций, равны нулю:

$$m_{2k+1} = \mathsf{M}\xi^{2k+1} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2k+1} e^{-\frac{x^2}{2\sigma^2}}\, dx = 0,$$

$$k = 0, 1, 2, \ldots$$

Вычислим теперь моменты четного порядка. Интегрируя по частям, получим

$$m_{2k} = M\xi^{2k} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2k} e^{-\frac{x^2}{2\sigma^2}}\, dx =$$

$$= -\frac{\sigma}{\sqrt{2\pi}} x^{2k-1} e^{-\frac{x^2}{2\sigma^2}} \Big|_{-\infty}^{+\infty} +$$

$$+ (2k-1)\frac{\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2k-2} e^{-\frac{x^2}{2\sigma^2}}\, dx =$$

$$= (2k-1)\,\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2k-2} e^{-\frac{x^2}{2\sigma^2}}\, dx = (2k-1)\,\sigma^2 m_{2k-2}.$$

Таким образом, установлена следующая рекуррентная формула:

$$m_{2k} = (2k-1)\,\sigma^2 m_{2k-2}.$$

Отсюда, так как

$$m_0 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}}\, dx = 1,$$

находим, что $m_2 = \sigma^2 m_0 = \sigma^2$, $m_4 = 3\sigma^2 m_2$, и т. д.

По индукции получаем

$$m_{2k} = (2k-1)(2k-3)\ldots 3\cdot 1\cdot \sigma^{2k} = (2k-1)!!\,\sigma^{2k}.$$

Итак, моменты случайной величины ξ, распределенной нормально с параметрами $(0, \sigma^2)$, определяются формулами

$$M\xi^{2k-1} = 0, \qquad M\xi^{2k} = (2k-1)!!\,\sigma^{2k}, \qquad k = 1, 2, \ldots$$

Задачи

1. Длина диаметра круга равномерно распределена в отрезке [0,1]. Найти математическое ожидание и дисперсию площади круга.

2. Плотность распределения ξ равна

$$p_\xi(x)=\begin{cases}\dfrac{c}{x}, & x\in[1,2],\\ 0, & x\notin[1,2].\end{cases}$$

Найти c, $\mathsf{M}\xi$, $\mathsf{D}\xi$.

3. Координаты двух случайных точек на прямой независимы и равномерно распределены на отрезке $[0,1]$. Найти математическое ожидание и дисперсию расстояния между ними.

4. Из урны, содержащей m белых и n черных шаров, по схеме случайного выбора с возвращением извлекают шары до первого появления белого шара. Найти математическое ожидание и дисперсию числа вынутых шаров.

5. Доказать, что асимметрия и эксцесс для нормального распределения равны нулю.

6. В k-м испытании схемы Бернулли произошел двойной успех, если успех наступил в k-м и $k+1$-м испытаниях. Обозначим η_n — число двойных успехов в n испытаниях схемы Бернулли. Найти $\lim\limits_{n\to\infty}\dfrac{\mathsf{D}\eta_n}{n}$, если вероятность успеха в отдельном испытании равна p $(0<p<1)$.

§ 9. КОВАРИАЦИЯ. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Ковариацией случайных величин $\xi_1=\xi_1(\omega)$ и $\xi_2=\xi_2(\omega)$ называется число

$$\operatorname{cov}(\xi_1,\xi_2)=\mathsf{M}[(\xi_1-\mathsf{M}\xi_1)(\xi_2-\mathsf{M}\xi_2)]. \tag{1}$$

Из определения ковариации и свойств линейности математического ожидания следует, что

$$\operatorname{cov}(\xi_1,\xi_2)=\mathsf{M}[(\xi_1-\mathsf{M}\xi_1)(\xi_2-\mathsf{M}\xi_2)]=$$
$$=\mathsf{M}[\xi_1\xi_2-\xi_1\mathsf{M}\xi_2-\xi_2\mathsf{M}\xi_1+\mathsf{M}\xi_1\cdot\mathsf{M}\xi_2]=$$
$$=\mathsf{M}\xi_1\xi_2-\mathsf{M}\xi_1\cdot\mathsf{M}\xi_2.$$

Таким образом, ковариацию можно вычислить также по следующей формуле:

$$\operatorname{cov}(\xi_1,\xi_2)=\mathsf{M}\xi_1\xi_2-\mathsf{M}\xi_1\cdot\mathsf{M}\xi_2. \tag{2}$$

Если ξ_1 и ξ_2 независимы, то

$$\operatorname{cov}(\xi_1,\xi_2)=0. \tag{3}$$

Это равенство следует из определения (1) и свойства мультипликативности математического ожидания произведения независимых величин $\xi_1 - M\xi_1$ и $\xi_2 - M\xi_2$. На основании свойств математического ожидания также легко проверяется, что

$$\operatorname{cov}(\xi, \xi) = D\xi, \quad \operatorname{cov}(\xi_1, \xi_2) = \operatorname{cov}(\xi_2, \xi_1),$$
$$\operatorname{cov}(c_1\xi_1 + c_2\xi_2, \xi_3) = c_1 \operatorname{cov}(\xi_1, \xi_3) + c_2 \operatorname{cov}(\xi_2, \xi_3), \tag{4}$$

где c_1, c_2 — постоянные.

Теорема 1. *Если для случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ существуют $\sigma_{ij} = \operatorname{cov}(\xi_i, \xi_j)$, $i, j = 1, 2, \ldots, n$, то при любых постоянных $c_1, c_2, \ldots, c_n$ существует $D\left(\sum_{i=1}^{n} c_i\xi_i\right)$ и*

$$D\left(\sum_{i=1}^{n} c_i\xi_i\right) = \sum_{i,j=1}^{n} c_i c_j \sigma_{ij}. \tag{5}$$

Действительно, пользуясь свойством линейности математического ожидания, получим

$$D\left(\sum_{k=1}^{n} c_k\xi_k\right) = M\left[\sum_{k=1}^{n} c_k\xi_k - M\left(\sum_{k=1}^{n} c_k\xi_k\right)\right]^2 =$$
$$= M\left[\sum_{k=1}^{n} c_k(\xi_k - M\xi_k)\right]^2 =$$
$$= M\left[\sum_{i,j=1}^{n} c_i c_j (\xi_i - M\xi_i)(\xi_j - M\xi_j)\right] =$$
$$= \sum_{i,j=1}^{n} c_i c_j \operatorname{cov}(\xi_i, \xi_j) = \sum_{i,j=1}^{n} c_i c_j \sigma_{ij}.$$

Замечание 1. Если случайные величины ξ_i, $i = 1, 2, \ldots, n$, попарно независимы, то из равенства (3) следует, что $\sigma_{ij} = \operatorname{cov}(\xi_i, \xi_j) = 0$ при $i \neq j$. Отсюда и из равенства (5), получим формулу

$$D\left(\sum_{k=1}^{n} c_k\xi_k\right) = \sum_{k=1}^{n} c_k^2 D\xi_k,$$

являющуюся обобщением формулы (7) § 8.

Замечание 2. Используя свойство неотрицательности дисперсии, из формулы (5), положив в ней

$n=2$, $c_1=X$, $c_2=1$, получим, что

$$f(X)=\mathsf{D}(X\xi_1+\xi_2)=X^2\sigma_{11}+2X\sigma_{12}+\sigma_{22}\geqslant 0$$

при любом действительном X. Следовательно, $\sigma_{12}^2-\sigma_{11}\cdot\sigma_{22}\leqslant 0$ или $|\sigma_{12}|<\sqrt{\sigma_{11}\sigma_{22}}$. Отсюда, заменяя σ_{12}, σ_{11}, σ_{22} на $\operatorname{cov}(\xi_1,\xi_2)$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$, получим неравенство.

$$|\operatorname{cov}(\xi_1,\xi_2)|\leqslant\sqrt{\mathsf{D}\xi_1\cdot\mathsf{D}\xi_2}. \tag{6}$$

Теорема 2. *Если для случайных величин ξ_1, $\xi_2,\ldots,\xi_n$ существуют $\sigma_{ij}=\operatorname{cov}(\xi_i,\xi_j)$, $i, j=1, 2,\ldots, n$, то при любых постоянных c_{ij}, $i, j=1, 2,\ldots, n$, для случайных величин $\eta_k=c_{k1}\xi_1+c_{k2}\xi_2+\ldots+c_{kn}\xi_n$, $k=1, 2, \ldots, n$, существуют ковариации $h_{ij}=\operatorname{cov}(\eta_i,\eta_j)$, и ковариационные матрицы $H=\|h_{ij}\|$ и $S=\|\sigma_{ij}\|$ связаны равенством*

$$H=CSC',$$

где $C=\|c_{ij}\|$, а C' — матрица, транспонированная к C.

Доказательство. Используя свойство (4), получим

$$h_{ij}=\operatorname{cov}(\eta_i,\eta_j)=\operatorname{cov}\left(\sum_{k=1}^n c_{ik}\xi_k,\ \sum_{r=1}^n c_{jr}\xi_r\right)=$$

$$=\sum_{k,\ r=1}^n c_{ik}c_{jr}\sigma_{kr}=\sum_{k,\ r=1}^n c_{ik}\sigma_{kr}c'_{rj},$$

где $c'_{rj}=c_{jr}$ — элемент матрицы C'. Таким образом, элемент h_{ij} совпадает с (i,j)-м элементом матрицы CSC'. Теорема доказана.

Равенство (3) верно для независимых случайных величин. Следовательно, если $\operatorname{cov}(\xi_1,\xi_2)\neq 0$, то случайные величины ξ_1 и ξ_2 зависимы. В качестве количественной характеристики степени зависимости случайных величин ξ_1 и ξ_2 вводится *коэффициент корреляции* $\rho(\xi_1,\xi_2)$, определяемый следующим равенством:

$$\rho(\xi_1,\xi_2)=\frac{\operatorname{cov}(\xi_1,\xi_2)}{\sqrt{\mathsf{D}\xi_1\mathsf{D}\xi_2}}.$$

Основные свойства коэффициента корреляции

1°. *Коэффициент корреляции удовлетворяет неравенству*

$$|\rho(\xi_1, \xi_2)| \leqslant 1; \tag{8}$$

причем, если ξ_2 *линейно выражается через* ξ_1, $\xi_2 = a\xi_1 + b$, *где* a *и* b *— постоянные, то*

$$|\rho(\xi_1, \xi_2)| = 1. \tag{9}$$

Доказательство. Поделив обе части неравенства (6) на $\sqrt{\mathsf{D}\xi_1 \cdot \mathsf{D}\xi_2}$, получим неравенство (8). Докажем равенство (9). Пусть $\xi_2 = a\xi_1 + b$. Тогда $\mathsf{D}\xi_2 = \mathsf{D}(a\xi_1) = a^2\mathsf{D}\xi_1$ и $\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)] = \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)(a\xi_1 + b - \mathsf{M}(a\xi_1 + b))] = \mathsf{M}[a(\xi_1 - \mathsf{M}\xi_1)(\xi_1 - \mathsf{M}\xi_1)] = a\mathsf{D}\xi_1$. Отсюда и из формулы (7) получим

$$\rho(\xi_1, \xi_2) = \rho(\xi_1, a\xi_1 + b) = \frac{a\mathsf{D}\xi_1}{\sqrt{\mathsf{D}\xi_1 \cdot a^2\mathsf{D}\xi_1}} = \frac{a}{|a|}.$$

Следовательно, $|\rho(\xi_1, \xi_2)| = 1$.

Имеет место обратное утверждение: если коэффициент корреляции $|\rho(\xi_1, \xi_2)| = 1$, то между ξ_1 и ξ_2 имеется линейная зависимость.

2°. *Если случайные величины* ξ_1 *и* ξ_2 *независимы, то* $\rho(\xi_1, \xi_2) = 0$.

Это утверждение следует из определения коэффициента корреляции (7) и из того, что $\operatorname{cov}(\xi_1, \xi_2) = 0$ для независимых случайных величин ξ_1 и ξ_2.

Обратное утверждение неверно. Из равенства $\rho(\xi_1, \xi_2) = 0$ не следует независимость случайных величин ξ_1 и ξ_2. Рассмотрим соответствующий пример.

Пример. Пусть случайная величина ξ_1 распределена нормально с параметрами (0, 1), и пусть $\xi_2 = \xi_1^2 - 1$. Случайные величины ξ_1 и ξ_2 зависимы. Покажем, что $\rho(\xi_1, \xi_2) = 0$. В самом деле,

$$\mathsf{M}\xi_1 = 0, \quad \mathsf{M}\xi_1^2 = \mathsf{D}\xi_1 = 1, \quad \mathsf{M}\xi_2 = \mathsf{M}(\xi_1^2 - 1) = 0,$$

$$\mathsf{D}\xi_2 = \mathsf{M}\xi_2^2 = \mathsf{M}(\xi_1^2 - 1)^2 = \mathsf{M}(\xi_1^4 - 2\xi_1^2 + 1) = 3 - 2 + 1 = 2,$$

$$\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}\xi_1\xi_2 = \mathsf{M}\xi_1(\xi_1^2 - 1) = \mathsf{M}\xi_1^3 - \mathsf{M}\xi = 0.$$

Следовательно, $\rho(\xi_1, \xi_2) = (\operatorname{cov}(\xi_1, \xi_2))/\sqrt{D\xi_1 D\xi_2} = 0$, хотя ξ_1 и ξ_2 зависимы.

Случайные величины ξ_1 и ξ_2 называются некоррелированными, если $\rho(\xi_1, \xi_2) = 0$.

Из вышеизложенного следует, что независимые случайные величины некоррелированы, а некоррелированные случайные величины не обязательно независимы. Рассмотрим случайные величины ξ, η с плотностью распределения

$$p_{\xi,\eta}(x, y) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{1-R^2}} e^{-(x^2+y^2-2Rxy)/(2(1-R^2))}, \tag{10}$$

где $|R| < 1$. В этом случае говорят, что случайные величины ξ, η имеют двумерное нормальное распределение.

Найдем одномерные плотности и коэффициент корреляции случайных величин ξ, η.

Воспользовавшись равенством

$$x^2 + y^2 - 2Rxy = (x - Ry)^2 + y^2(1 - R^2),$$

получим

$$p_\eta(y) = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sqrt{1-R^2}} e^{-\frac{x^2+y^2-2Rxy}{2(1-R^2)}} dx =$$

$$= \frac{1}{2\pi\sqrt{1-R^2}} e^{-\frac{y^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-Ry)^2}{2(1-R^2)}} dx.$$

Отсюда, полагая $\frac{x-Ry}{\sqrt{1-R^2}} = z$, $dx = \sqrt{1-R^2}\,dz$, найдем

$$p_\eta(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}},$$

или

$$p_\eta(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad -\infty < y < +\infty.$$

Аналогично

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad -\infty < x < +\infty.$$

Вычислим теперь коэффициент корреляции $\rho(\xi, \eta)$ для случайных величин ξ, η с плотностью двумерного нормального распределения (10). Так как $\mathsf{M}\xi = \mathsf{M}\eta = 0$, $\mathsf{D}\xi = \mathsf{D}\eta = 1$, то

$$\operatorname{cov}(\xi, \eta) = \mathsf{M}\xi\eta = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyp_{\xi\eta}(x, y)\,dx\,dy =$$

$$= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} ye^{-\frac{y^2}{2}}\left(\int\limits_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi(1-R^2)}}\,xe^{-\frac{(x-Ry)^2}{2(1-R^2)}}\,dx\right)dy.$$

Внутренний интеграл можно рассматривать как математическое ожидание случайной величины, распределенной нормально с параметрами $a = Ry$ и $\sigma^2 = 1 - R^2$. Поэтому он равен Ry и

$$\mathsf{M}\xi\eta = \frac{R}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}}\,dy = R.$$

Подставляя найденные значения $\mathsf{M}\xi$, $\mathsf{M}\eta$, $\mathsf{D}\xi$, $\mathsf{D}\eta$, $\mathsf{M}\xi\eta$ в формулы (2) и (7), получим, что коэффициент корреляции случайных величин ξ, η с нормальной плотностью совместного распределения (10) равен R, т. е. $\rho(\xi, \eta) = R$.

Отметим, что для случайных величин ξ, η, распределенных нормально с плотностью (10), независимость равносильна некоррелированности.

Действительно, если ξ и η независимы, то $\rho(\xi, \eta) = 0$ и, следовательно, случайные величины ξ, η некоррелированы. Если же случайные величины ξ, η имеют совместное нормальное распределение с плотностью (10) и некоррелированы, т. е. $\rho(\xi, \eta) = R = 0$, то

$$p_{\xi,\eta}(x, y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2} =$$
$$= \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,\frac{1}{\sqrt{2\pi}}e^{-y^2/2} = p_\xi(x)\,p_\eta(y),$$

что и означает независимость случайных величин ξ и η.

Задачи

1. Совместное распределение случайных величин ξ_1, ξ_2 задано формулами

$$\mathsf{P}\{\xi_1=-1,\ \xi_2=-1\}=\mathsf{P}\{\xi_1=0,\ \xi_2=-1\}=$$
$$=\mathsf{P}\{\xi_1=1,\ \xi_2=-1\}=\frac{1}{6},$$
$$\mathsf{P}\{\xi_1=-1,\xi_2=1\}=\frac{1}{4},\quad \mathsf{P}\{\xi_1=0,\ \xi_2=1\}=$$
$$=\mathsf{P}\{\xi_1=1,\ \xi_2=1\}=\frac{1}{8}.$$

Вычислить математические ожидания и дисперсии ξ_1 и ξ_2, а также ковариацию и коэффициент корреляции.

2. Совместное распределение случайных величин ξ_1, ξ_2 определяется формулами $\mathsf{P}\{\xi_1=0,\ \xi_2=1\}=\mathsf{P}\{\xi_1=0,\ \xi_2=-1\}=$ $=\mathsf{P}\{\xi_1=1,\ \xi_2=0\}=\mathsf{P}\{\xi_1=-1,\ \xi_2=0\}=\frac{1}{4}$. а) Найти $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$; б) вычислить ковариацию и коэффициент корреляции ξ_1, ξ_2. Являются ли случайные величины независимыми?

3. Плотность совместного распределения случайных величин ξ_1, ξ_2 задана формулой

$$p_{\xi_1,\xi_2}(x,\ y)=\begin{cases}\dfrac{c}{x^3y^3} & \text{при } x\geqslant 1,\ y\geqslant 1,\\ 0 & \text{в остальных случаях.}\end{cases}$$

Найти постоянную c и ковариацию ξ_1, ξ_2.

4. Плотность совместного распределения случайных величин ξ_1, ξ_2 задана формулой

$$p_{\xi_1,\xi_2}(x,\ y)=\begin{cases} c\,(1-xy^3) & \text{при } |x|\leqslant 1,\quad |y|\leqslant 1,\\ 0 & \text{в остальных случаях.}\end{cases}$$

Найти постоянную c и коэффициент корреляции ξ_1, ξ_2.

5. Случайные величины ξ_1, ξ_2, ξ_3, ξ_4, ξ_5 независимы; $\mathsf{D}\xi_i=\sigma^2$. Найти коэффициент корреляции а) величин $\xi_1+\xi_2$, $\xi_3+\xi_4+\xi_5$; б) величин $\xi_1+\xi_2+\xi_3$, $\xi_3+\xi_4+\xi_5$.

6. Бросаются две игральные кости. Пусть ξ — число очков на первой кости, а η — максимальное из двух выпавших чисел очков.

а) Записать совместное распределение ξ и η.

б) Найти математические ожидания и дисперсии ξ и η.

в) Вычислить ковариацию и коэффициент корреляции ξ и η.

§ 10. ЗАКОН БОЛЬШИХ ЧИСЕЛ

10.1. Неравенство Чебышёва. Числовые характеристики случайных величин позволяют давать некоторые оценки распределений вероятностей случайных величин.

Теорема 1 (вторая форма неравенства Чебышёва). *Если случайная величина имеет дисперсию* $\mathsf{D}\xi$, *а* Δ — *произвольное положительное число, то*

$$\mathsf{P}\{|\xi-\mathsf{M}\xi|\geqslant\Delta\}\leqslant\frac{\mathsf{D}\xi}{\Delta^2}. \qquad (1)$$

Доказательство. В § 7 была доказана первая форма неравенства Чебышёва, $\mathsf{P}\{\eta\geqslant\varepsilon\}\leqslant\frac{\mathsf{M}\eta}{\varepsilon}$, для неотрицательных случайных величин η. Полагая $\eta=|\xi-\mathsf{M}\xi|^2$ и $\varepsilon=\Delta^2$, получаем (1).

Если в (1) величину Δ заменить на $t\sqrt{\mathsf{D}\xi}$ и перейти к противоположному событию, то неравенство Чебышёва (1) можно записать в следующем виде:

$$\mathsf{P}\{|\xi-\mathsf{M}\xi|<t\sqrt{\mathsf{D}\xi}\}\geqslant 1-\frac{1}{t^2}. \qquad (2)$$

Неравенство Чебышёва позволяет оценивать вероятности отклонений значений случайной величины от своего математического ожидания. Пусть проводится n независимых измерений некоторой неизвестной величины a. Ошибки измерения $\delta_1, \delta_2, \ldots, \delta_n$ будем считать случайными величинами. Предположим, что $\mathsf{M}\delta_k=0$, $k=1, 2, \ldots, n$. Это условие можно рассматривать как отсутствие систематической ошибки. Предположим также, что $\mathsf{D}\delta_k=b^2$. За значение неизвестной величины a принимают обычно среднее арифметическое результатов измерений. Тогда ошибка в определении числа a будет равна $\eta_n=(\delta_1+\delta_2+\ldots+\delta_n)/n$ и

$$\mathsf{D}\eta_n=\frac{1}{n^2}(\mathsf{D}\delta_1+\ldots+\mathsf{D}\delta_n)=\frac{b^2}{n}, \quad \mathsf{M}\eta_n=0.$$

Оценим число измерений n, при котором ошибка η_n не превосходит Δ с достаточно большой вероятностью. Например, $\mathsf{P}\{|\eta_n|<\Delta\}>0{,}99$ или

$$\mathsf{P}\{|\eta_n|\geqslant\Delta\}\leqslant 0{,}01. \qquad (3)$$

По неравенству Чебышёва (1) имеем

$$\mathsf{P}\{|\eta_n|\geqslant\Delta\}\leqslant\frac{\mathsf{D}\eta_n}{\Delta^2}=\frac{b^2}{n\Delta^2}.$$

Следовательно, (3) будет выполнено, если

$$\frac{b^2}{n\Delta^2} \leqslant 0{,}01 \quad \text{или} \quad n > 100\frac{b^2}{\Delta^2}.$$

Таким образом, мы получили оценку числа измерений, необходимого для получения заданной точности.

10.2. Закон больших чисел. Говорят, что к случайным величинам ξ_k, $k = 1, 2, \ldots$, имеющим математические ожидания $\mathsf{M}\xi_k$, $k = 1, 2, \ldots$, применим закон больших чисел, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} - \frac{\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_n}{n}\right| < \varepsilon\right\} = 1. \tag{4}$$

Теорема 2 (теорема Маркова). *Если у случайных величин ξ_k, $k = 1, 2, \ldots$, существуют дисперсии и если при $n\to\infty$*

$$\frac{1}{n^2}\mathsf{D}\left(\sum_{k=1}^{n}\xi_k\right) \to 0, \tag{5}$$

то к случайным величинам ξ_k, $k = 1, 2, \ldots$, применим закон больших чисел.

Доказательство. Обозначим $\eta_n = (\xi_1 + \ldots + \xi_n)/n$. Пользуясь (1) с $\Delta = \varepsilon$, $\xi = \eta_n$, получим

$$\mathsf{P}\{|\eta_n - \mathsf{M}\eta_n| \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^2}\mathsf{D}\eta_n.$$

Отсюда, так как

$$\mathsf{M}\eta_n = \frac{\mathsf{M}\xi_1+\ldots+\mathsf{M}\xi_n}{n}, \quad \mathsf{D}\eta_n = \frac{1}{n^2}\mathsf{D}\left(\sum_{k=1}^{n}\xi_k\right),$$

для вероятности противоположного события находим оценку

$$\mathsf{P}\left\{\left|\frac{\xi_1+\ldots+\xi_n}{n} - \frac{\mathsf{M}\xi_1+\ldots+\mathsf{M}\xi_n}{n}\right| < \varepsilon\right\} \geqslant 1 - \frac{1}{n^2}\mathsf{D}\left(\sum_{k=1}^{n}\xi_k\right).$$

Из этого неравенства и условия (5) следует (4). Теорема доказана.

Укажем некоторые частные случаи этой теоремы.

Теорема 3 (теорема Чебышёва). *Если случайные величины* ξ_k, $k = 1, 2, \ldots$, *попарно независимы, имеют равномерно ограниченные дисперсии* (*т. е. существует постоянная* c *такая, что* $\mathsf{D}\xi_k < c$ *при всех* $k = 1, 2, \ldots$), *то к случайным величинам* ξ_k, $k = 1, 2, \ldots$, *применим закон больших чисел.*

В самом деле, для доказательства теоремы достаточно проверить условие (5). Из неравенств $\mathsf{D}\xi_k < c$, $k = 1, 2, \ldots$, и попарной независимости случайных величин ξ_k, $k = 1, 2, \ldots$, следует, что

$$\mathsf{D}\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n} \mathsf{D}\xi_k < nc.$$

Отсюда получаем условие (5):

$$\frac{1}{n^2}\mathsf{D}\left(\sum_{k=1}^{n} \xi_k\right) < \frac{c}{n} \to 0$$

при $n \to \infty$.

Теорема 4. *Если случайные величины* ξ_k, $k = 1, 2, \ldots$, *одинаково распределены, попарно независимы и имеют конечные дисперсии, то к этим случайным величинам применим закон больших чисел.*

Теорема 4 следует из теоремы 3. Действительно, дисперсии $\mathsf{D}\xi_k$, $k = 1, 2, \ldots$, существуют и равны между собой; следовательно, они равномерно ограничены и мы находимся в условиях теоремы 3.

Утверждение теоремы 4 означает, что для любых $\varepsilon > 0$, $\delta > 0$, найдется такое N, что при $n > N$ верно неравенство

$$\mathsf{P}\left\{\left|\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - a\right| < \varepsilon\right\} \geqslant 1 - \delta, \qquad (6)$$

где $a = \mathsf{M}\xi_k$, $k = 1, 2, \ldots$

Теорема 5 (теорема Бернулли). *Пусть* μ_n — *число успехов в* n *испытаниях Бернулли и* p — *вероятность успеха в каждом отдельном испытании. Тогда* *для любого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1. \tag{7}$$

Для доказательства этой теоремы воспользуемся представлением μ_n в виде суммы n индикаторов: $\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n$, где $\xi_k = 1$, если в k-м испытании был успех, и $\xi_k = 0$ в противном случае.

Так как ξ_k, $k = 1, 2, \ldots, n$, независимы, одинаково распределены ($\mathsf{P}\{\xi_k = 1\} = p$, $\mathsf{P}\{\xi_k = 0\} = 1 - p = q$), дисперсии случайных величин ξ_k существуют и $\mathsf{M}\xi_k = p$, то теорема 5 сразу следует из теоремы 4.

Замечание. Доказательство теоремы 5 можно провести и непосредственно. Действительно, так как $\mathsf{M}\frac{\mu_n}{n} = p$, $\mathsf{D}\xi_k = pq$, $\mathsf{D}\frac{\mu_n}{n} = \frac{pq}{n}$, то по неравенству Чебышёва (1)

$$\mathsf{P}\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} \geqslant 1 - \frac{1}{\varepsilon^2}\frac{pq}{n}. \tag{8}$$

Это неравенство позволяет оценить вероятность того, что отклонение частоты $\frac{\mu_n}{n}$ в схеме Бернулли от вероятности p не больше ε. Если выбрать n_0 таким, что при $n \geqslant n_0$ $\frac{pq}{\varepsilon^2 n} < \delta$, где δ — малое число, то согласно (8) при $n \geqslant n_0$ с вероятностью, не меньшей $1 - \delta$, частота $\frac{\mu_n}{n}$ будет находиться в пределах

$$p - \varepsilon \leqslant \frac{\mu_n}{n} \leqslant p + \varepsilon. \tag{9}$$

Исходя из свойства устойчивости частот, о котором говорилось в § 1, можно сформулировать принцип практической достоверности: события, вероятность которых близка к единице, практически достоверны, т. е. при единичном испытании, как правило, осуществляются. По этому принципу в серии из n независимых испытаний неравенство (9) должно выполняться, если $n \geqslant n_0$.

Задачи

1. Случайная величина ξ имеет нормальное распределение с параметрами (a, σ). Оценить по неравенству Чебышёва $\mathsf{P}\{|\xi - a| \geqslant 2\sigma\}$. Сравнить с точным значением этой вероятности.

2. Закон распределения случайной величины ξ определяется формулами

$$\mathsf{P}\{\xi = 0\} = 1 - \frac{\sigma^2}{\Delta^2}, \quad \mathsf{P}\{\xi = -\Delta\} = \mathsf{P}\{\xi = \Delta\} = \frac{\sigma^2}{2\Delta^2}.$$

Сравнить точное значение вероятности $\mathsf{P}\{|\xi| \geqslant \Delta\}$ с оценкой, полученной по неравенству Чебышёва.

3. Предполагается провести 10 измерений $x_1, x_2, \ldots, x_{10}$ неизвестной величины a. Считая $x_1, \ldots, x_{10}$ независимыми нормально распределенными случайными величинами с $\mathsf{M}x_k = a$, $\mathsf{D}x_k = 0{,}01$, найти Δ, если

$$\mathsf{P}\left\{\left|\frac{x_1 + x_2 + \ldots + x_{10}}{10} - a\right| < \Delta\right\} = 0{,}99.$$

Оценить Δ, используя неравенство Чебышёва. Сравнить полученные результаты.

4. Применим ли закон больших чисел к последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_k, \ldots$, если

$$\mathsf{P}\{\xi_k = \sqrt{k}\} = \mathsf{P}\{\xi_k = -\sqrt{k}\} = \frac{1}{2\sqrt{k}}, \quad \mathsf{P}\{\xi_k = 0\} = 1 - \frac{1}{\sqrt{k}}?$$

§ 11. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

В § 4 была доказана теорема Муавра — Лапласа, согласно которой число успехов μ_n в n испытаниях Бернулли имеет распределение, близкое к нормальному. Представим μ_n в виде суммы независимых индикаторов $\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n$, где $\xi_k = 1$, если в k-м испытании был успех и $\xi_k = 0$ в противном случае.

Теорему Муавра — Лапласа можно сформулировать в следующем виде.

Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, *независимы,* $\mathsf{P}\{\xi_k = 1\} = 1 - \mathsf{P}\{\xi_k = 0\} = p$, $k = 1, 2, \ldots$, *то при* $n \to \infty$, $0 < p < 1$,

$$\mathsf{P}\left\{\frac{\xi_1 + \xi_2 + \ldots + \xi_n - n\mathsf{M}\xi_1}{\sqrt{n\mathsf{D}\xi_1}} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du \quad (1)$$

равномерно по $x \in (-\infty, +\infty)$.

Утверждение (1) сохраняется при достаточно общих предположениях о законе распределения случайных величин ξ_k. Найдем плотности распределений сумм $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$, где $n = 1, 2, 3$, ξ_k $(k = 1, 2, 3)$ — независимые случайные величины, равномерно распределенные на отрезке $[-1, 1]$. Таким образом, плотности $p_{\xi_k}(x) = p(x)$ $(k = 1, 2, 3)$, где

$$p(x) = \begin{cases} 1/2, & \text{если } x \in [-1, 1], \\ 0, & \text{если } x \notin [-1, 1]. \end{cases} \tag{2}$$

Плотность распределения суммы $\eta_2 = \xi_1 + \xi_2$ находится по формуле (см. (14) § 6)

$$p_{\eta_2}(x) = \int_{-\infty}^{+\infty} p(u)\, p(x-u)\, du = \frac{1}{2} \int_{-1}^{1} p(x-u)\, du.$$

Из формулы (2) следует, что подынтегральное выражение $p(x-u) = 1/2$, если $-1 < x - u < 1$ или $x - 1 < u < x + 1$, и $p(x-u) = 0$ в остальных случаях. Общую часть отрезка интегрирования $[-1, 1]$ и отрезка $[x-1, x+1]$, на котором $p(x-u) = 1/2$, можно записать в виде $[\max(-1, x-1), \min(1, x+1)]$. Следовательно, в случае $-2 < x < 0$

$$p_{\eta_2}(x) = \frac{1}{2} \int_{-1}^{x+1} p(x-u)\, du = \frac{1}{4} \int_{-1}^{x+1} du = \frac{x+2}{4},$$

а если $0 < x < 2$, то

$$p_{\eta_2}(x) = \frac{1}{2} \int_{x-1}^{1} p(x-u)\, du = \frac{1}{4} \int_{x-1}^{1} du = \frac{2-x}{4};$$

при $|x| > 2$ имеем $p_{\eta_2}(x) = 0$. Таким образом,

$$p_{\eta_2}(x) = \begin{cases} 0, & \text{если } |x| \geqslant 2, \\ \dfrac{x+2}{4}, & \text{если } -2 \leqslant x \leqslant 0, \\ \dfrac{2-x}{4}, & \text{если } 0 \leqslant x \leqslant 2. \end{cases} \tag{3}$$

Распределение, задаваемое плотностью (3), называется *распределением Симпсона.*

Плотность распределения суммы трех величин может быть определена как плотность распределения суммы двух величин $\eta_3 = \eta_2 + \xi_3$:

$$p_{\eta_3}(x) = \int_{-\infty}^{+\infty} p_{\eta_2}(u)\, p(x-u)\, du =$$

$$= \int_{-2}^{0} \frac{u+2}{4} p(x-u)\, du + \int_{0}^{2} \frac{2-u}{4} p(x-u)\, du.$$

Выяснив интервалы, на которых подынтегральные функции положительны, после вычислений, аналогичных вычислениям для двух слагаемых, получим

$$p_{\eta_3}(x) = \begin{cases} 0, & \text{если } |x| \geqslant 3, \\ (x+3)^2/16, & \text{если } -3 \leqslant x \leqslant -1, \\ (3-x)^2/8, & \text{если } -1 \leqslant x \leqslant 1, \\ (x-3)^2/16, & \text{если } 1 \leqslant x \leqslant 3. \end{cases} \qquad (4)$$

Графики плотностей, $p_{\eta_1}(x) = p(x)$, $p_{\eta_2}(x)$, $p_{\eta_3}(x)$ приведены на рис. 13, 14, 15. Отметим, что плотность распределения суммы с ростом числа слагаемых становится более гладкой: $p_{\eta_1}(x)$ — разрывна; $p_{\eta_2}(x)$ — непрерывна; $p_{\eta_3}(x)$ — дифференцируема. Отметим также, что плотность $p_{\eta_3}(x)$ более (по сравнению с $p_{\eta_1}(x)$, $p_{\eta_2}(x)$) напоминает плотность нормального распределения. Приведем без доказательства условия, при которых закон распределения суммы независимых случайных величин сближается с нормальным распределением.

Теорема 1. *Если случайные величины $\xi_1, \xi_2, \ldots$ $\ldots, \xi_n, \ldots$ независимы, одинаково распределены и имеют конечную дисперсию, то при $n \to \infty$ равномерно по $x \in (-\infty; +\infty)$*

$$\mathsf{P}\left\{\frac{\xi_1 + \xi_2 + \ldots + \xi_n - na}{\sigma\sqrt{n}} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du,$$

где $a = \mathsf{M}\xi_k$, $\sigma^2 = \mathsf{D}\xi_k$, $k = 1, 2, \ldots, n$.

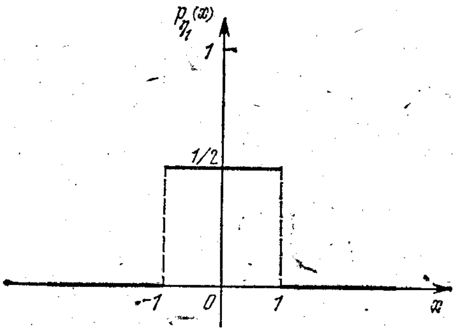


Рис. 13.

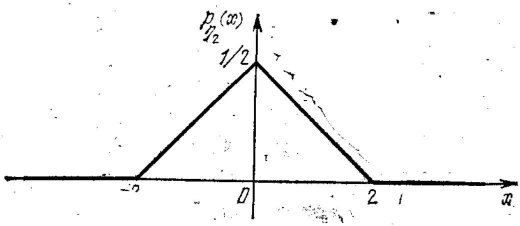


Рис. 14.

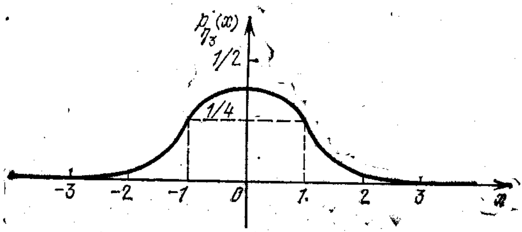


Рис. 15.

Условия сходимости функций распределения независимых случайных величин, имеющих различные распределения, содержатся в теореме Ляпунова. Приведем без доказательства ее формулировку.

Теорема 2. *Пусть* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — *независимые случайные величины, имеющие конечный* *третий абсолютный момент. Положим*

$$a_k = \mathsf{M}\xi_k, \quad b_k^2 = \mathsf{D}\xi_k, \quad c_k^3 = \mathsf{M}\,|\xi_k - a_k|^3, \ k = 1, 2, \ldots, n,$$

$$A_n = \sum_{k=1}^{n} a_k, \quad B_n^2 = \sum_{k=1}^{n} b_k^2, \quad C_n^3 = \sum_{k=1}^{n} c_k^3.$$

Если

$$\lim_{n\to\infty} \frac{C_n}{B_n} = 0, \tag{5}$$

то при $n \to \infty$ равномерно по x $(-\infty < x < +\infty)$

$$\mathsf{P}\left\{\frac{\xi_1 + \xi_2 + \ldots + \xi_n - A_n}{B_n} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du.$$

При ссылках на теоремы о сходимости распределений сумм случайных величин к нормальному закону удобно использовать понятие асимптотической нормальности. Если функции распределения последовательности случайных величин $(\eta_n - A_n)/B_n$ сходятся при $n \to \infty$ к функции $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du$, то говорят, что случайная величина η_n при $n \to \infty$ *асимптотически нормальна с параметрами* (A_n, B_n).

Теоремы 1 и 2 позволяют объяснить частую встречаемость нормального распределения. Например, ошибки при измерениях часто оказываются нормально распределенными. Этот факт может объясняться тем, что ошибка складывается из большого числа слагаемых, вызываемых независимыми факторами.

Достаточно общее утверждение о сходимости распределения сумм случайных величин к нормальному закону называют *центральной предельной теоремой*.

В качестве примера применения теорем 1 и 2 оценим число испытаний в методе Монте-Карло, необходимое для вычисления кратного интеграла с заданной точностью. Пусть функция $f(\boldsymbol{x}) = f(x_1, \ldots, x_r)$ определена на r-мерном единичном кубе V. Требуется вычислить интеграл $a = \int \ldots \int_V f(\boldsymbol{x})\,d\boldsymbol{x}$.

Пусть известна постоянная C такая, что $|f(x)|\leqslant C$, $x\in V$. Обозначим $\xi=(\xi_1, \ldots, \xi_r)$ случайный вектор, равномерно распределенный на V. Тогда $p_\xi(x_1, \ldots, x_r)=1$, если $x\in V$, и $p_\xi(x_1, \ldots, x_r)=0$ в противном случае. Математическое ожидание случайной величины $\eta=f(\xi)$ найдем по формуле

$$\mathsf{M}\eta=\int\limits_V\ldots\int f(x_1, \ldots x_r)\,p_\xi(x_1, \ldots, x_r)\,dx_1\ldots dx_r=$$
$$=\int\limits_V\ldots\int f(x)\,dx=a.$$

Таким образом, $\mathsf{M}\eta$ совпадает со значением вычисляемого интеграла. Так как $|f(x)|<C$, то

$$\sigma^2=\mathsf{D}\eta=\int\limits_V\ldots\int (f(x)-a)^2\,dx\leqslant 4C^2.$$

Пусть теперь случайные векторы $\xi_k=(\xi_{k1}, \ldots, \xi_{kr})$, $k=1, \ldots, n$, независимы*) и распределены равномерно на единичном кубе V. Тогда случайные величины $\eta_k=f(\xi_k)$, $k=1, \ldots, n$, независимы и одинаково распределены. По закону больших чисел случайная величина $\zeta_n=(\eta_1+\ldots+\eta_n)/n$ при больших n близка к постоянной $a=\mathsf{M}\eta_k$. Предположим, что нужно вычислить a с точностью Δ. Оценим вероятность

$$\mathsf{P}\{|\zeta_n-a|<\Delta\}=\mathsf{P}\left\{\left|\frac{\zeta_n-a}{\sigma/\sqrt{n}}\right|<\frac{\Delta\sqrt{n}}{\sigma}\right\}.$$

Так как $\sigma<2C$, то

$$\mathsf{P}\{|\zeta_n-a|<\Delta\}\geqslant\mathsf{P}\left\{\left|\frac{\zeta_n-a}{\sigma/\sqrt{n}}\right|<\frac{\Delta\sqrt{n}}{2C}\right\},$$

и при больших n

$$\mathsf{P}\left\{\left|\frac{\zeta_n-a}{\sigma/\sqrt{n}}\right|<\frac{\Delta\sqrt{n}}{2C}\right\}\approx 2\Phi_0\left(\frac{\Delta\sqrt{n}}{2C}\right).$$

---

*) Случайные векторы ξ_k, $k=1, \ldots, n$, независимы, **если** для любых r-мерных прямоугольников B_k, $k=1, 2, \ldots, n$,

$$\mathsf{P}\{\xi_1\in B_1, \ldots, \xi_n\in B_n\}=\mathsf{P}\{\xi_1\in B_1\}\ldots\mathsf{P}\{\xi_n\in B_n\}.$$

По заданной малой вероятности нежелательного события $|\zeta_n - a| \geqslant \Delta$ можно так же, как в § 4, найти n.

Задачи

1. Складываются 10^4 чисел, каждое из которых округлено с точностью до 10^{-m}. Предполагая, что ошибки от округления независимы и равномерно распределены в интервале $\left(-\frac{1}{2}10^{-m}, \frac{1}{2}10^{m}\right)$, найти пределы, в которых с вероятностью, не меньшей 0.99, будет лежать суммарная ошибка.

2. Получить теорему Муавра — Лапласа в качестве следствия теорем 1 и 2.

3. Пусть случайная величина ξ_λ распределена по закону Пуассона с параметром λ. Найти

$$\lim_{\lambda\to\infty} \mathsf{P}\left\{\frac{\xi_\lambda - \lambda}{\sqrt{\lambda}} < x\right\}.$$

Указание. Воспользоваться тем, что случайная величина ξ_{λ_n} ($\lambda_n = n \cdot h$) при любом $h > 0$ представляется в виде $\xi_{\lambda_n} = \xi_h^{(1)} + \xi_h^{(2)} + \ldots + \xi_h^{(n)}$, где $\xi_h^{(k)}$ — независимые случайные величины, распределенные по закону Пуассона с параметром h.

§ 12. ОБРАБОТКА РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

Измерения неизвестных величин или параметров приходится проводить в самых различных областях научной и технической деятельности. Результаты измерений одной и той же величины несколько отличаются друг от друга, так как невозможно полностью сохранить условия проведения различных измерений. Рассмотрим простейшую математическую модель процесса измерения.

12.1. Выборка. Обычно измерения стараются проводить независимо одно от другого и примерно в одинаковых условиях. Пусть проводится n измерений. В результате измерений будет получено n чисел $x_1, x_2, \ldots, x_n$. Если повторить еще раз n измерений, то получатся другие n чисел, отличные от первого набора. Процесс из n независимых измерений естественно описать как n независимых случайных величин, числовые значения которых при различных ω соответствуют результатам измерений в различных опытах.

Выборкой объема n назовем n независимых случайных величин $X_1, X_2, \ldots, X_n$, каждая из которых распределена так же, как некоторая случайная величина ξ с функцией распределения $\mathsf{P}\{\xi \leqslant x\} = F_\xi(x)$.

Таким образом, $\mathsf{P}\{X_i \leqslant x\} = F_\xi(x)$ при любом $i = 1, 2, \ldots, n$.

Выборка является математической моделью n независимых измерений, проводимых в одинаковых условиях. Случайная величина ξ является характеристикой прибора или метода измерений. Величины $X_1, \ldots, X_n$ можно рассматривать как n независимых «экземпляров» величины ξ.

12.2. Оценка. По результатам измерений требуется найти число, близкое к неизвестному значению измеряемого параметра. Пусть, например, по значениям выборки объема n требуется оценить неизвестный параметр θ закона распределения $\mathsf{P}\{\xi \leqslant x\} = F_\xi(\theta, x)$. *Оценкой* неизвестного параметра θ назовем произвольную функцию $\hat\theta_n = \hat\theta_n(X_1, X_2, \ldots, X_n)$. Значения этой функции $\hat\theta_n(x_1, x_2, \ldots, x_n)$ при полученных в результате измерений значениях $X_i = x_i$, $i = 1, \ldots, n$, будем рассматривать как приближенное значение параметра θ. Приведенное определение оценки отражает только самое общее требование, что оценка должна определяться по значениям выборки. Очевидно, что любая оценка не обязательно будет близкой к оцениваемому параметру θ. Введем два свойства оценок, которые обеспечивают их близость к соответствующим параметрам.

Оценка $\hat\theta_n$ параметра θ называется несмещенной, если $\mathsf{M}\hat\theta_n = \theta$.

Оценка $\hat\theta_n$ параметра θ называется состоятельной, если при $n \to \infty$

$$\mathsf{P}\{|\hat\theta_n - \theta| < \varepsilon\} \to 1$$

для любого $\varepsilon > 0$.

Свойство несмещенности означает, что оценка не имеет систематической ошибки. Свойство состоятельности обеспечивает сближение оценки с измеряемым параметром при увеличении числа измерений.

Пример 1. Рассмотрим выборку $X_1, X_2, \ldots, X_n$, соответствующую случайной величине ξ с $F_\xi(x) = \mathsf{P}\{\xi \leqslant x\}$. Пусть нас интересует параметр $a = \mathsf{M}\xi$. Предположим, что существует $\mathsf{D}\xi = \sigma^2$. Обычно в качестве оценки $\hat{a}_n$ параметра a выбирают среднее арифметическое $\bar{X}$ значений выборки, т. е. полагают

$$\hat{a}_n = \bar{X} = \frac{X_1 + \ldots + X_n}{n}. \tag{1}$$

Согласно определению выборки величины $X_1, \ldots, X_n$ независимы и каждая имеет распределение, совпадающее с ξ. Поэтому

$$\mathsf{M}X_i = \mathsf{M}\xi = a, \quad \mathsf{D}X_i = \mathsf{D}\xi = \sigma^2.$$

Используя свойства линейности математического ожидания и свойства дисперсии, из формулы (1) получим

$$\mathsf{M}\hat{a}_n = \mathsf{M}\bar{X} = \frac{1}{n}(\mathsf{M}X_1 + \ldots + \mathsf{M}X_n) = \frac{na}{n} = a, \tag{2}$$

$$\mathsf{D}\hat{a}_n = \frac{1}{n^2}(\mathsf{D}X_1 + \ldots + \mathsf{D}X_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}. \tag{3}$$

Равенство (2) означает, что оценка $\hat{a}_n = \bar{X}$ является несмещенной оценкой параметра a. Из (3) следует, что среднее квадратичное отклонение оценки $\bar{X}$ от неизвестного параметра a, равное $\sigma/\sqrt{n}$, убывает с увеличением числа измерений n.

Воспользовавшись неравенством Чебышёва и равенством (3), получим, что для любого $\varepsilon > 0$

$$\mathsf{P}\{|\hat{a}_n - a| > \varepsilon\} \leqslant \frac{\mathsf{D}\hat{a}_n}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \to 0$$

при $n \to \infty$. Следовательно, вероятность противоположного события $\mathsf{P}\{|\hat{a}_n - a| < \varepsilon\} \to 1$ при $n \to \infty$, т. е. оценка $\hat{a}_n = \bar{X}$ является состоятельной. Таким образом, мы показали, что

среднее арифметическое $\bar{X}$ является несмещенной и состоятельной оценкой измеряемого параметра, если измерения X_i имеют конечную дисперсию.

Может существовать несколько различных несмещенных и состоятельных оценок параметра a. Среди всех таких оценок естественно выбрать оценку с наименьшей дисперсией.

Пример 2. Пусть теперь по выборке $X_1, \ldots, X_n$, соответствующей распределению $F_\xi(x)$, нужно оценить два параметра $a = \mathsf{M}\xi$ и $\sigma^2 = \mathsf{D}\xi$. В качестве оценки a мы выберем $\bar{X}$. Оценку $\hat{\sigma}_n^2$ параметра σ^2 определим формулой $\hat{\sigma}_n^2 = s^2$, где

$$s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})^2. \tag{4}$$

Покажем, что оценка $\hat{\sigma}_n^2 = s^2$ является несмещенной оценкой σ^2. Преобразуем сначала формулу (4), введя случайные величины $Y_k = X_k - c$. Так как

$$X_k - \bar{X} = (Y_k + c) - \frac{1}{n}\sum_{k=1}^{n}(Y_k + c) = Y_k - \bar{Y},$$

где $\bar{Y} = (Y_1 + \ldots + Y_n)/n$, то из формулы (4) получим

$$s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(Y_k - \bar{Y})^2, \tag{5}$$

т. е. величина s^2 не изменится, если все величины X_k сдвинуть на одну и ту же величину. Положим теперь $c = a = \mathsf{M}X_k$. Тогда $\mathsf{M}Y_k = 0$, $\mathsf{D}Y_k = \mathsf{M}Y_k^2 = \mathsf{D}X_k = \sigma^2$ и

$$\mathsf{M}(Y_k - \bar{Y})^2 = \mathsf{M}Y_k^2 - \frac{2}{n}\mathsf{M}\left(Y_k\sum_{l=1}^{n}Y_l\right) + \frac{1}{n^2}\sum_{l_1,\, l_2=1}^{n}\mathsf{M}Y_{l_1}Y_{l_2}.$$

Так как Y_{l_1}, Y_{l_2} независимы при $l_1 \neq l_2$ и $\mathsf{M}Y_{l_1}Y_{l_2} = \mathsf{M}Y_{l_1}\mathsf{M}Y_{l_2} = 0$ $(l_1 \neq l_2)$, то

$$\mathsf{M}\left(Y_k\sum_{l=1}^{n}Y_l\right) = \mathsf{M}\left(Y_k^2 + \sum_{k:\, k\neq l}Y_kY_l\right) = \mathsf{M}Y_k^2 + \sum_{k:\, k\neq l}\mathsf{M}Y_kY_l = \sigma^2,$$

$$\sum_{l_1,\, l_2=1}^{n}\mathsf{M}Y_{l_1}Y_{l_2} = \sum_{l=1}^{n}\mathsf{M}Y_l^2 + \sum_{l_1\neq l_2}\mathsf{M}Y_{l_1}Y_{l_2} = \sum_{l=1}^{n}\mathsf{M}Y_l^2 = n\sigma^2.$$

Следовательно,

$$\mathsf{M}(Y_k-\overline{Y})^2=\sigma^2-\frac{2}{n}\sigma^2+\frac{n\sigma^2}{n^2}=\frac{n-1}{n}\sigma^2.$$

Отсюда и из формулы (5) находим

$$\mathsf{M}s^2=\frac{1}{n-1}\sum_{k=1}^{n}\mathsf{M}(Y_k-\overline{Y})^2=\frac{1}{n-1}\left(n\cdot\frac{n-1}{n}\sigma^2\right)=\sigma^2.$$

Таким образом, несмещенность оценки s^2 доказана. Если предположить, что $\mathsf{M}\xi^4$ существует, то можно показать, что при $n\to\infty$

$$\mathsf{D}s^2=O\left(\frac{1}{n}\right).$$

Отсюда, используя неравенство Чебышёва

$$\mathsf{P}\{|s^2-\sigma^2|>\varepsilon\}\leqslant\frac{\mathsf{D}s^2}{\varepsilon^2}=O\left(\frac{1}{n}\right),$$

получим, так же, как в примере 1 для $\overline{X}$, состоятельность оценки s^2.

Пример 3. Пусть выборка $X_1, \ldots, X_n$ соответствует испытаниям схемы Бернулли, т. е.

$$\mathsf{P}\{X_i=1\}=p,\quad \mathsf{P}\{X_i=0\}=q=1-p.$$

Тогда частота «успеха» $h_n=\frac{1}{n}(X_1+\ldots+X_n)$. Покажем, что h_n является состоятельной и несмещенной оценкой p. Так как $\mathsf{M}X_k=p$ и $\mathsf{D}X_k=pq$ конечна, то рассматриваемая схема является частным случаем общей схемы примера 1. Таким образом, можно воспользоваться выводами примера 1, и, следовательно, частота h_n является состоятельной и несмещенной оценкой p.

Пример 4. Найдем по выборке $X_1, \ldots, X_n$, соответствующей распределению $F_\xi(x)$, оценку $\hat{F}_n(x)$ функции распределения $F_\xi(x)$ при каждом x. Положим

$$\hat{F}_n(x)=\frac{1}{n}\mu_n(x),$$

где $\mu_n(x)$ — число случайных величин среди $X_1, X_2, \ldots, X_n$, не превосходящих x. Функция $\hat{F}_n(x)$ называется *эмпирической функцией распределения.* С каждой величиной X_k можно связать два события $\{X_k \leqslant x\}$ и $\{X_k > x\}$. Вероятности этих событий очевидно равны $p = \mathsf{P}\{X_k \leqslant x\} = F_\xi(x)$, $q = \mathsf{P}\{X_k > x\} = 1 - F(x)$. Если событие $\{X_k \leqslant x\}$ назвать «успехом», то $\mu_n(x)$ является числом успехов в серии из n независимых испытаний, а $\hat{F}_n(x)$ является частотой успеха. Следовательно, согласно предыдущему примеру $\hat{F}_n(x)$ является несмещенной и состоятельной оценкой параметра $p = F_\xi(x)$. На рис. 16 приведены

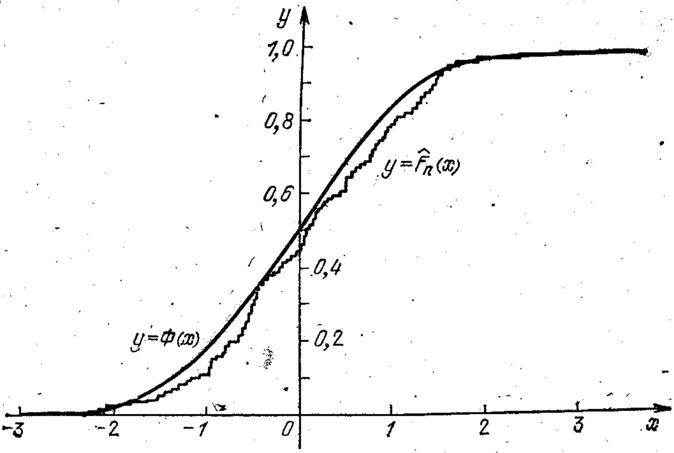


Рис. 16.

график функции $F_\xi(x) = \Phi(x)$ и график $\hat{F}_n(x)$ (при $n = 100$), вычисленной по реализации, взятой из нормально распределенных случайных чисел (см. [1], [11], [14]).

Иногда для наглядного представления выборки используется *гистограмма*, получаемая следующим образом. Числовая ось разбивается на несколько непересекающихся полуинтервалов: $(-\infty, \infty) = \bigcup_{k=0}^{r} (z_k, z_{k+1}]$, где $-\infty = z_0 < z_1 < z_2 < \ldots < z_r < z_{r+1} < \infty$. Далее вычисляются частоты $\hat{p}_k = \hat{F}_n(z_{k+1}) - \hat{F}_n(z_k)$ попадания элементов выборки в эти интервалы. Затем над каждым отрезком $[z_k, z_{k+1}]$ строится прямоугольник, площадь которого пропорциональна частоте $\hat{p}_k$. Частоты $\hat{p}_k$ при больших n близки к $p_k = \int\limits_{z_k}^{z_{k+1}} p(x)\,dx$, где $p(x)$ — плотность распределения $F_\xi(x)$. Таким образом, при удачном выборе ширины интервалов гистограмма может напоминать график плотности распределения $p(x)$.

Рассмотрение графиков эмпирической функции распределения и гистограммы может дать некоторое предварительное представление о неизвестной функции распределения $F_\xi(x)$.

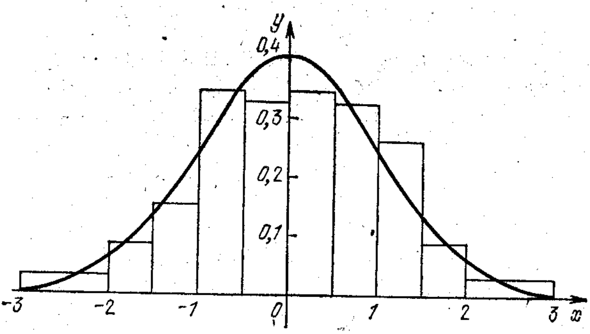


Рис. 17.

На рис. 17 приведены график плотности нормального распределения с параметрами (0, 1) и гистограмма, соответствующая $F_n(x)$ (см. рис. 16).

Пример 5. Из урны, содержащей M белых и $N-M$ черных шаров по схеме случайного выбора без возвращения, извлекается n шаров. Определим индикаторы X_k, $k=1, \ldots, n$, положив $X_k = 1$, если k-й шар (по порядку извлечения) оказался белым и $X_k = 0$ в противном случае. Случайные величины $X_1, \ldots, X_n$ очевидно зависимы. В этом случае так же, как и в случае независимых выборок, определенных в п. 12.1, можно искать оценки неизвестных параметров схемы. Оценим, например, долю $p = M/N$ белых шаров в урне. Положим

$$\hat{p}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n).$$

Математическое ожидание и дисперсия числа белых шаров в выборке $\eta_n = X_1 + \ldots + X_n$ были определены в § 8 (пример 5). Используя эти формулы, получим

$$\mathbf{M}\hat{p}_n = p, \quad \mathbf{D}\hat{p}_n = \frac{1}{n} p(1-p) \cdot \frac{N-n}{N-1}.$$

Таким образом, $\hat{p}_n$ является несмещенной оценкой p. Состоятельность $\hat{p}_n$ следует из неравенства Чебышёва

$$\mathbf{P}\{|\hat{p}_n - p| \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^2 n} p(1-p) \frac{N-n}{N-1}.$$

12.3. Интервальные оценки. Если известен закон распределения оценки или ее дисперсия, то можно указать границы, в которых с большой вероятностью находится неизвестное значение параметра.

Эти границы зависят от значений неизвестного параметра, и тогда пользоваться ими нельзя.

Рассмотрим выборку $X_1, X_2, \ldots, X_n$ с $\mathbf{P}\{X_i \leqslant x\} = F_\xi(\theta, x)$, где θ — неизвестный параметр. Предположим, что нам удалось найти две функции $\underline{\theta} = \underline{\theta}(X_1, X_2, \ldots, X_n)$ и $\overline{\theta} = \overline{\theta}(X_1, \ldots, X_n)$ такие, что $\underline{\theta} < \overline{\theta}$ при всех значениях $X_1, \ldots, X_n$ и при любых значениях θ

$$\mathbf{P}\{\underline{\theta}(X_1, \ldots, X_n) < \theta < \overline{\theta}(X_1, \ldots, X_n)\} = 1 - 2\alpha,$$

т. е. вероятность того, что случайный интервал $(\underline{\theta}, \overline{\theta})$ накроет неизвестный параметр θ, не зависит от параметра. В этом случае интервал $(\underline{\theta}, \overline{\theta})$ называют *доверительным интервалом* для неизвестного параметра θ, соответствующим *доверительной вероятности* $1-2\alpha$. В ряде случаев функции $\underline{\theta}$ и $\overline{\theta}$, обладающие указанными свойствами, можно найти.

Пусть имеется выборка $X_1, \ldots, X_n$, причем величины X_k распределены нормально с параметрами (a, σ), а параметр σ известен. Найдем доверительный интервал для a. Случайная величина $\overline{X}=(X_1+\ldots$ $\ldots+X_n)/n$ имееет нормальное распределение и $\mathsf{M}\overline{X}=a$, $\mathsf{D}\overline{X}=\sigma^2/n$. Следовательно, величина $\dfrac{\overline{X}-a}{\sigma\sqrt{n}}$ распределена нормально с параметрами (0, 1), и ее распределение не зависит от a. Отсюда

$$\mathsf{P}\left\{\left|\frac{\overline{X}-a}{\sigma/\sqrt{n}}\right|<u_\alpha\right\}=1-2\alpha$$

или

$$\mathsf{P}\left\{\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}}<a<\overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right\}=1-2\alpha, \tag{6}$$

где u_α определяется как решение уравнения

$$\frac{1}{\sqrt{2\pi}}\int_{u_\alpha}^{\infty}e^{-\frac{x^2}{2}}\,dx=\alpha. \tag{7}$$

Таким образом, мы нашли доверительный интервал $\left(\overline{X}-u_\alpha\dfrac{\sigma}{\sqrt{n}},\ \overline{X}+u_\alpha\dfrac{\sigma}{\sqrt{n}}\right)$ для параметра a.

Более естественной является ситуация, когда оба параметра a и σ неизвестны. В этом случае для построения доверительного интервала используется величина $\tau_{n-1}=\dfrac{\overline{X}-a}{s}\sqrt{n}$, где s^2 определено формулой (4). В курсах математической статистики доказывается (см., например, [7]), что распределение величины τ_{n-1} не зависит от параметров a и σ нормального распределения, соответствующего выборке; плотность распределения τ_{n-1} определяется формулой

$$p_{\tau_{n-1}}(x)=C\cdot\left(1+\frac{x^2}{n-1}\right)^{-\frac{n}{2}}, \qquad n\geqslant 2,$$

где нормирующая постоянная $C=\Gamma\left(\frac{n}{2}\right)/\Gamma\left(\frac{n-1}{2}\right)\cdot$ $\cdot\sqrt{\pi(n-1)}$. Распределение величины τ_{n-1} называют *распределением Стьюдента с $n-1$ степенями свободы*. Определим величину $t_{\alpha,\,n-1}$ как решение уравнения

$$\mathsf{P}\{|\tau_{n-1}|<t_{\alpha,\,n-1}\}=1-2\alpha. \tag{8}$$

Отсюда, положив $\tau_{n-1}(\bar{X}-a)\sqrt{n}/s$, получим доверительный интервал для a:

$$\mathsf{P}\left\{\bar{X}-t_{\alpha,\,n-1}\frac{s}{\sqrt{n}}<a<\bar{X}+t_{\alpha,\,n-1}\frac{s}{\sqrt{n}}\right\}=1-2\alpha. \quad (9)$$

Приведем небольшую табличку значений $t_{\alpha,\,n}$:

2α \ n	5	10	20	30	∞
0,05	2,571	2,228	2,086	2,042	1,960
0,01	4,032	3,169	2,845	2,750	2,576

При $n\to\infty$ плотность $p_{\tau_{n-1}}(x)$ сходится к плотности нормального распределения с параметрами (0, 1). Так как при любом конечном n плотность $p_{\tau_{n-1}}(x)$ при $x\to\infty$ убывает медленнее плотности нормального распределения, то $t_{\alpha,\,n-1}>t_{\alpha,\,\infty}=u_\alpha$. Это приводит к тому, что средняя длина доверительного интервала (9) больше длины доверительного интервала (6).

При построении доверительных интервалов (6), (9), предполагалось, что элементы выборок имеют нормальное распределение. Пусть теперь $X_1, X_2, \ldots, X_n$ — произвольная выборка. При больших n можно построить простые приближенные доверительные интервалы без предположения нормальности X_k. Найдем, например, доверительный интервал для параметра $a=\mathsf{M}X_k$. Пусть $\sigma^2=\mathsf{D}X_k$. Тогда по центральной предельной теореме распределение величины

$$\frac{X_1+X_2+\ldots+X_n-na}{\sqrt{\sigma^2 n}}=\frac{\bar{X}-a}{\sigma/\sqrt{n}}$$

близко к нормальному распределению с параметрами (0, 1). Отсюда, используя состоятельность оценки s^2 параметра σ^2, можно показать, что

$$\lim_{n\to\infty}\mathsf{P}\left\{\left|\frac{\bar{X}-a}{s/\sqrt{n}}\right|<u_\alpha\right\}=1-2\alpha,$$

или

$$\mathsf{P}\left\{\bar{X}-u_\alpha\frac{s}{\sqrt{n}}<a<\bar{X}+u_\alpha\frac{s}{\sqrt{n}}\right\}\approx 1-2\alpha \quad (10)$$

при $n \to \infty$. Доказательство этого утверждения мы приводить не будем.

12.4. Метод наибольшего правдоподобия для нахождения оценок параметров. Метод моментов. Если функция распределения $F(\theta, x) = \mathsf{P}\{X_k \leqslant x\}$, $k = 1, \ldots, n$, имеет плотность $p(\theta, x)$, то функция

$$L = L(X_1, \ldots, X_n, \theta) = p(\theta, X_1) \ldots p(\theta, X_n) \qquad (11)$$

называется *функцией правдоподобия*. Для выборки, состоящей из дискретных величин X_k, $k = 1, \ldots, n$, с распределением $\mathsf{P}\{X_k = x\} = p_x(\theta)$ функция правдоподобия определяется равенством

$$L = L(X_1, \ldots, X_n, \theta) = p_{X_1}(\theta) \cdot p_{X_2}(\theta) \cdot \ldots \cdot p_{X_n}(\theta). \qquad (12)$$

При фиксированных $X_1, \ldots, X_n$ функцию L будем рассматривать как функцию параметра θ.

По методу наибольшего правдоподобия за оценку параметра θ принимается значение аргумента $\theta = \theta(X_1, \ldots, X_n)$, при котором L имеет максимальное значение, т. е. выбирается значение θ, при котором вероятность получения данных значений выборки максимальна.

Поскольку $\ln L$ при фиксированных $(X_1, \ldots, X_n)$ достигает максимума при том же значении θ, что и L, то для нахождения оценки можно решать уравнение правдоподобия

$$\frac{\partial \ln L}{\partial \theta} = 0. \qquad (13)$$

Решением (13) будем называть только корни, зависящие от $X_1, X_2, \ldots, X_n$. Каждое решение уравнения правдоподобия (13) будем называть *оценкой наибольшего правдоподобия для* θ.

При некоторых достаточно общих условиях (см. [7]) уравнение (13) имеет решение $\widehat{\theta}_n$, являющееся состоятельной оценкой параметра θ; кроме того, при $n \to \infty$ оценка θ_n асимптотически нормальна.

Если $\theta = (\theta_1, \ldots, \theta_s)$, то оценками наибольшего правдоподобия параметров $\theta_1, \ldots, \theta_s$ являются зависящие от $(X_1, \ldots, X_n)$ решения системы уравнений:

$$\frac{\partial \ln L}{\partial \theta_k} = 0, \qquad k = 1, 2, \ldots, s.$$

Пример 6. Пусть величины X_k, $k = 1, 2, \ldots, n$, имеют нормальное распределение. Неизвестными параметрами являются $a = \mathrm{M}X_k$, $b = \sigma^2 = \mathrm{D}X_k$. Найдем их оценки наибольшего правдоподобия. По формуле (11)

$$L = L(X_1, \ldots, X_n, a, b) = \left(\frac{1}{\sqrt{2\pi b}}\right)^n \exp\left(-\sum_{k=1}^{n} \frac{(X_k - a)^2}{2b}\right)$$

и

$$\ln L = -\frac{n}{2}(\ln 2\pi + \ln b) - \frac{1}{2b}\sum_{k=1}^{n}(X_k - a)^2.$$

Отсюда для оценок $\hat{a}$ и $\hat{b}$ получим систему

$$\begin{cases} \dfrac{\partial \ln L}{\partial a} = \dfrac{1}{\hat{b}} \displaystyle\sum_{k=1}^{n} (X_k - \hat{a}) = 0, \\ \dfrac{\partial \ln L}{\partial b} = -\dfrac{n}{2\hat{b}} + \dfrac{1}{2\hat{b}^2} \displaystyle\sum_{k=1}^{n} (X_k - \hat{a})^2 = 0. \end{cases}$$

Из первого уравнения $\hat{a} = \frac{1}{n}\sum_{k=1}^{n} X_k = \bar{X}$. Подставляя это значение во второе уравнение, найдем

$$\hat{b} = m_2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X})^2.$$

Пример 7. Найдем оценку наибольшего правдоподобия для вероятности успеха p в схеме Бернулли. По формуле (12) $L = L(X_1, \ldots, X_n, p) = \prod_{k=1}^{n} p^{X_k}(1-p)^{1-X_k}$, $X_k = 0$, 1 ($X_k = 1$, если в испытании k был успех, и $X_k = 0$ в противном случае) и

$$\ln L = \sum_{k=1}^{n} (X_k \ln p + (1 - X_k) \ln (1 - p)).$$

Отсюда

$$\frac{\partial \ln L}{\partial p} = \sum_{k=1}^{n} \left(\frac{X_k}{\hat{p}} - \frac{1 - X_k}{1 - \hat{p}} \right) = \frac{\bar{X} n}{\hat{p}} - \frac{n}{1 - \hat{p}} + \frac{n\bar{X}}{1 - \hat{p}} = 0$$

и $\hat{p} = \bar{X}$. Так как для числа успехов μ_n имеем равенство $\mu_n = X_1 + X_2 + \ldots + X_n$, то $\hat{p} = \mu_n/n$.

§ 13. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Неизвестные функциональные зависимости на практике часто устанавливают посредством измерений. Пусть, например, исследуется функциональная зависимость

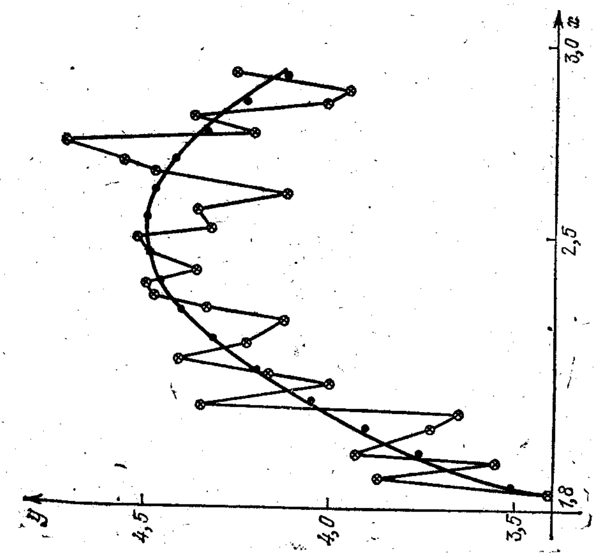
$$y = \varphi(x), \tag{1}$$

где $\varphi(x)$ неизвестна. При различных значениях x_1, x_2, ..., x_n, известных точно, измерены с ошибками значения соответствующих им y_1, y_2, ..., y_n. Модель таких измерений можно определить равенствами

$$y_i = \varphi(x_i) + \delta_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

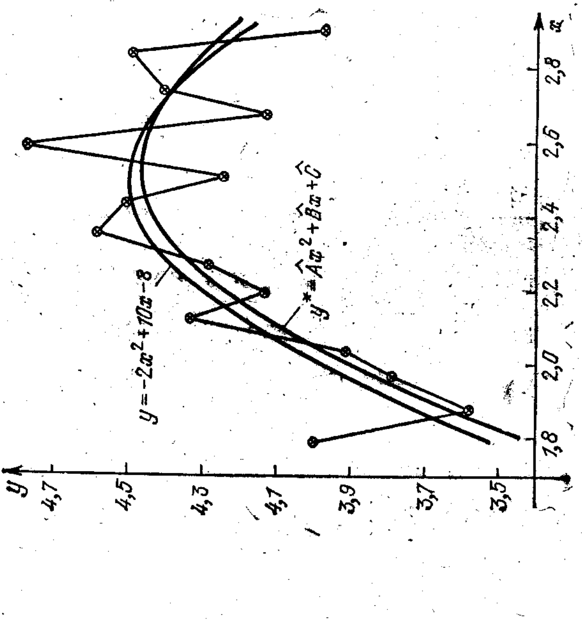
где δ_i — независимые случайные величины $\mathsf{M}\delta_i = 0$. При полном отсутствии информации о характере функции $\varphi(x)$ невозможно получить достаточно удовлетворительное ее приближение. В этом случае, соединив пары (x_i, y_i), $i = 1, 2, \ldots, n$, отрезками прямых, получим график, мало напоминающий исходную функцию $y = \varphi(x)$ (см. рис. 18, 19 и пример в конце § 13).

Обычно предполагается, что известны свойства гладкости функции $\varphi(x)$ или известен ее вид с точностью до некоторых параметров. Пусть, например, $y = \varphi(x, A, B)$, где функция $\varphi(x, A, B)$ известна, x — независимая переменная; A и B — неизвестные параметры. Предположим также, что в равенстве (2) величины δ_i имеют одинаковые дисперсии $\mathsf{D}\delta_i = \sigma^2$. По



y
4,5
4,0
3,5
1,8
2,5
3,0
x

Рис. 19.



$y=-2x^2+10x-8$
$y^*=\hat{A}x^2+\hat{B}x+\hat{C}$
y
4,7
4,5
4,3
4,1
3,9
3,7
3,5
1,8
2,0
2,2
2,4
2,6
2,8
x

Рис. 18.

значениям пар (x_i, y_i) требуется определить неизвестные коэффициенты A и B. Для определения оценок A и B используется *метод наименьших квадратов*, состоящий в том, что в качестве оценок A и B выбираются значения $\hat{A}$ и $\hat{B}$, минимизирующие сумму квадратов отклонений измеренных y_i от соответствующих вычисленных значений $\varphi(x_i, A, B)$:

$$Q(A, B) = \sum_{k=1}^{n} (y_k - \varphi(x_k, A, B))^2. \tag{3}$$

Таким образом, в качестве оценок $\hat{A}$, $\hat{B}$ параметров A и B предлагается брать решение системы уравнений

$$\frac{\partial Q}{\partial A} = 0, \quad \frac{\partial Q}{\partial B} = 0. \tag{4}$$

Рассмотрим свойства оценок, полученных методом наименьших квадратов, в простейшем случае. Пусть требуется найти коэффициенты A и B линейной функции $y = Ax + B$. Будем предполагать, что в выборке $(y_1, \ldots, y_n)$ величины y_i представимы в виде

$$y_i = Ax_i + B + \delta_i, \tag{5}$$

где δ_i независимы, $\mathsf{M}\delta_i = 0$, $\mathsf{D}\delta_i = \sigma^2$, x_i известны без ошибки. Уравнения (4) для функции $\varphi(x) = Ax + B$ имеют вид

$$\begin{cases} -\dfrac{\partial Q}{\partial A} = 2\displaystyle\sum_{k=1}^{n} x_k (y_k - Ax_k - B) = 0, \\ -\dfrac{\partial Q}{\partial B} = 2\displaystyle\sum_{k=1}^{n} (y_k - Ax_k - B) = 0. \end{cases} \tag{6}$$

Сдвигом начала отсчета на оси x можно всегда добиться, чтобы было выполнено равенство $\sum_{k=1}^{n} x_k = 0$. В этом случае система (6) приводится к следующему виду:

$$\begin{cases} \displaystyle\sum_{k=1}^{n} x_k y_k - A \sum_{k=1}^{n} x_k^2 = 0, \\ \displaystyle\sum_{k=1}^{n} y_k - B \cdot n = 0. \end{cases}$$

Отсюда получим

$$\hat{A}=\left(\sum_{k=1}^{n} x_k y_k\right)\Big/\left(\sum_{k=1}^{n} x_k^2\right), \quad \hat{B}=\frac{1}{n}\sum_{k=1}^{n} y_k. \tag{7}$$

Формулы (7) могут быть использованы для вычисления оценок. Для исследования свойств оценок $\hat{A}$ и $\hat{B}$ воспользуемся формулами

$$\hat{A}=A+\frac{\sum_{k=1}^{n} x_k\delta_k}{\sum_{k=1}^{n} x_k^2}, \quad \hat{B}=B+\frac{1}{n}\sum_{k=1}^{n}\delta_k, \tag{8}$$

которые следуют из формул (7), если в них y_k заменить по формуле (5). Так как величины δ_k независимы и $\mathsf{M}\delta_k=0$, $\mathsf{D}\delta_k=\sigma^2$, то

$$\mathsf{M}\left(\frac{\sum_{k=1}^{n} x_k\delta_k}{\sum_{k=1}^{n} x_k^2}\right)=\frac{1}{\sum_{k=1}^{n} x_k^2}\sum_{k=1}^{n} x_k\mathsf{M}\delta_k=0,$$

$$\mathsf{M}\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k\right)=\frac{1}{n}\sum_{k=1}^{n}\mathsf{M}\delta_k=0,$$

$$\mathsf{D}\left(\frac{\sum_{k=1}^{n} x_k\delta_k}{\sum_{k=1}^{n} x_k^2}\right)=\frac{1}{\left(\sum_{k=1}^{n} x_k^2\right)^2}\sum_{k=1}^{n} x_k^2\mathsf{D}\delta_k=\frac{\sigma^2}{\sum_{k=1}^{n} x_k^2},$$

$$\mathsf{D}\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}\mathsf{D}\delta_k=\frac{\sigma^2}{n}.$$

Отсюда и из формул (8) находим

$$\mathsf{M}\hat{A}=A, \quad \mathsf{M}\hat{B}=B, \quad \mathsf{D}\hat{A}=\frac{\sigma^2}{\sum_{k=1}^{n} x_k^2}, \quad \mathsf{D}\hat{B}=\frac{\sigma^2}{n}.$$

Таким образом, оценки $\hat{A}$ и $\hat{B}$ являются несмещенными оценками параметров A и B. Оценки $\hat{A}$ и $\hat{B}$ состоятельны, если $\sum_{k=1}^{n} x_k^2 \to \infty$ при $n \to \infty$, так как в этом случае $D\hat{A} \to 0$, $D\hat{B} \to 0$.

Если предположить, что δ_i, а следовательно, и y_i распределены нормально, то можно показать, что оценки, полученные методом наименьших квадратов, совпадут с оценками метода наибольшего правдоподобия.

Действительно, функция правдоподобия выборки $y_1, \ldots, y_n$ в этом случае имеет вид

$$L(y, A, B) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - Ax_k - B)^2\right\}.$$

Отсюда для оценок наибольшего правдоподобия параметров A и B получаем систему уравнений

$$\left\{\begin{aligned} -\frac{\partial \ln L}{\partial A} &= \frac{1}{\sigma^2} \sum_{k=1}^{n} x_k (y_k - Ax_k - B) = 0, \\ -\frac{\partial \ln L}{\partial B} &= \frac{1}{\sigma^2} \sum_{k=1}^{n} (y_k - Ax_k - B) = 0, \end{aligned}\right. \tag{9}$$

отличающуюся от системы (6) только положительным постоянным множителем $1/2\sigma^2$. Метод наименьших квадратов можно применять и для нелинейных функций.

Пример. Пусть неизвестная функция $y = Ax^2 + Bx + C$. Для значений $x_i = \frac{2i}{n}$, $i = 0, 1, \ldots, n-1$, $A = -2$, $B = 10$, $C = -8$ найдены значения $Ax_i^2 + Bx_i + C$ и при помощи датчика случайных чисел получены две реализации ($n = 25$, $n = 100$) $\tilde{\delta}_1, \tilde{\delta}_2, \ldots, \tilde{\delta}_n$ случайных ошибок $\delta_1, \ldots, \delta_n$ в предположении, что δ_i независимы и распределены нормально с $M\delta_i = 0$, $D\delta_i = 0{,}04$.

Таким образом, смоделированы измерения

$$\tilde{y}_i = -2x_i^2 + 10x_i - 8 + \tilde{\delta}_i, \qquad i = 1, \ldots, n. \tag{10}$$

Оценим, используя метод наименьших квадратов, параметры A, B, C, истинные значения которых нам известны. Будем минимизировать сумму

$$Q(A, B, C)=\sum_{k=1}^{n}(\tilde{y}_k-Ax_k^2-Bx_k-C)^2.$$

Оценки $\hat{A}$, $\hat{B}$, $\hat{C}$ удовлетворяют системе уравнений

$$\begin{cases} -\dfrac{\partial Q}{\partial A}=2\sum\limits_{k=1}^{n}(\tilde{y}_k-Ax_k^2-Bx_k-C)\,x_k^2=0, \\ -\dfrac{\partial Q}{\partial B}=2\sum\limits_{k=1}^{n}(\tilde{y}_k-Ax_k^2-Bx_k-C)\,x_k=0, \\ -\dfrac{\partial Q}{\partial C}=2\sum\limits_{k=1}^{n}(\tilde{y}_k-Ax_k^2-Bx_k-C)=0. \end{cases}$$

Отсюда находим

а) $n=25$: $\hat{A}=-1{,}871$, $\hat{B}=9{,}510$, $\hat{C}=-7{,}627$,

б) $n=100$: $\hat{A}=-2{,}046$, $\hat{B}=10{,}200$, $\hat{C}=-8{,}218$.

На рисунках 18 и 19, относящихся к случаям $n=25$ и $n=100$ соответственно, приведен график точной функциональной зависимости $y=-2x^2+10x-8$; знаком $\otimes$ отмечены «результаты измерений» $(x_i, \tilde{y}_i)$ (на рис. 18 отмечены все точки с $x_i \in (1{,}8;\ 3{,}0)$, а на рис. 19 отмечены точки с $x_i \in (1{,}8;\ 3{,}0)$ и $x_{i+1}-x_i=0{,}04$)) и соединены ломаной линией. На рис. 18 приведен график найденной по оценкам функциональной зависимости $y^*=\hat{A}x^2+\hat{B}x+\hat{C}$ при $n=25$. График функциональной зависимости, найденной по 100 измерениям, почти не отличается от $y=-2x^2+10x-8$, и поэтому на рис. 19 приведены только точки (x_i, y_i^*), где $y_i^*=\hat{A}x_i^2+\hat{B}x_i+\hat{C}$ $(n=100)$. Эти точки отмечены знаком $\cdot$. Приведем табличку значений $y_i=-2x_i^2+10x_i-8$, $y_i^*=\hat{A}x_i^2+\hat{B}x_i+\hat{C}$:

x_i	1,800	1,880	1,960	2,040	2,120	2,200
y_i	3,520	3,731	3,917	4,077	4,211	4,320
y_i^* $(n=25)$	3,430	3,640	3,826	3,988	4,126	4,241
y_i^* $(n=100)$	3,511	3,725	3,912	4,073	4,208	4,317
x_i	2,280	2,360	2,440	2,520	2,600	2,680
y_i	4,403	4,461	4,493	4,499	4,480	4,435
y_i^* $(n=25)$	4,331	4,397	4,440	4,458	4,453	4,423
y_i^* $(n=100)$	4,399	4,456	4,486	4,490	4,467	4,419

§ 14. СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

14.1. Критерий χ^2. В приложениях часто возникает следующая задача. Пусть в результате какого-либо эксперимента получена выборка

$$X_1, X_2, \ldots, X_n \tag{1}$$

объема n с функцией распределения $F(x) = \mathsf{P}\{X_i \leqslant x\}$. Какие-либо предположения о виде распределения $F(x)$ будем называть *статистическими гипотезами* или просто *гипотезами*. Иногда нас интересует гипотеза, состоящая в том, что функция распределения $F(x)$ совпадает с некоторой фиксированной функцией $F_0(x)$. Задача проверки статистической гипотезы $F(x) = F_0(x)$ состоит в том, чтобы решить, согласуются ли с ней значения X_i выборки (1). Для решения этой задачи поступим следующим образом. Разделим точками $z_0 = -\infty < z_1 < z_2 < \ldots < z_{r-1} < z_r = \infty$ всю прямую на r интервалов $(z_{k-1}, z_k]$. Обозначим $p_k = F_0(z_k) - F_0(z_{k-1}) = \mathsf{P}\{z_{k-1} < X_i \leqslant z_k\}$ — вероятность попадания X_i в интервал $(z_{k-1}, z_k]$ в случае, когда наша гипотеза справедлива. По выборке (1) определим числа ν_k, $k = 1, \ldots, r$, где ν_k — число элементов X_i выборки (1), попавших в интервал $(z_{k-1}, z_k]$. Таким образом, мы свели задачу к более простой. Имеется n независимых испытаний с r исходами. Вероятность k-го исхода равна p_k. Набор вероятностей исходов

$$p_1,\ p_2,\ \ldots,\ p_r,\quad \sum_{k=1}^{r} p_k = 1, \tag{2}$$

определяется первоначальной статистической гипотезой. Случайные величины

$$\nu_1,\ \nu_2,\ \ldots,\ \nu_r,\quad \sum_{k=1}^{r} \nu_k = n,$$

определяемые по выборке (1), имеют полиномиальное распределение (см. § 3 (8)) с вероятностями исходов (2):

$$\mathsf{P}\{\nu_k = m_k,\ k = 1, \ldots, r\} = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r}, \quad \sum_{l=1}^{r} m_l = n. \tag{3}$$

Если значения ν_k соответствуют вероятностям p_k, то разности $\frac{\nu_k}{n} - p_k$ должны быть малы. Рассмотрим случайную величину

$$\chi^2 = \sum_{k=1}^{r} \frac{n}{p_k}\left(\frac{\nu_k}{n} - p_k\right)^2 = \sum_{k=1}^{r} \frac{(\nu_k - np_k)^2}{np_k}, \tag{4}$$

которую будем называть χ^2-статистикой Пирсона. Если случайные величины ν_k имеют полиномиальное распределение (3), то справедливо следующее утверждение: *при любом* $x > 0$

$$\lim_{n\to\infty} \mathsf{P}\{\chi^2 \leqslant x\} = \int_0^x k_{r-1}(u)\,du; \tag{5}$$

плотность $k_{r-1}(x)$ *называется плотностью* χ^2-*распределения с* $r-1$ *степенью свободы и имеет* *следующий вид*:

$$k_{r-1}(x)=\begin{cases}\dfrac{x^{\frac{r-3}{2}}}{2^{\frac{r-1}{2}}\Gamma\left(\frac{r-1}{2}\right)}e^{-\frac{x}{2}} & \text{при}\quad x\geqslant 0,\\ 0 & \text{при}\quad x<0.\end{cases}$$

Этот результат используется следующим образом. Зададимся каким-либо малым значением вероятности α, которое будем называть *уровнем значимости* критерия. Заменим при больших n предельное соотношение (5) приближенным равенством

$$\mathsf{P}\{\chi^2\leqslant x\}\approx\int_0^k k_{r-1}(u)\,du. \qquad (6)$$

Выбирая $x=x_{\alpha,\,r-1}$ таким, чтобы $\int_0^{x_{\alpha,\,r-1}} k_{r-1}(u)\,du = 1-\alpha$ ($x_{\alpha,\,r-1}$ называется *α-квантилью* распределения χ^2), получаем, что в случае, когда проверяемая гипотеза справедлива, событие

$$\{\chi^2>x_{\alpha,\,r-1}\} \qquad (7)$$

может произойти лишь с малой вероятностью, которая приближенно равна α. Обычно полагают $\alpha=0{,}05$ или $\alpha=0{,}01$. Если гипотеза верна, то маловероятное событие (7) практически невозможно. Если оно произошло, то будем считать, что гипотеза неверна. Если же $\chi^2\leqslant x_{\alpha,\,r-1}$, то будем говорить, что выборка (1) не противоречит гипотезе $F=F_0$.

На практике считают, что приближенным равенством (6) можно пользоваться, если все произведения $np_i\geqslant 10$ (кроме, может быть, крайних). Если выбор точек деления z_k и число интервалов r зависит от нас, то мы должны, с одной стороны, обеспечить неравенство $np_i\geqslant 10$, и, с другой стороны, не брать очень крупные интервалы $(z_{k-1},\ z_k]$, чтобы вероятности $p_k=F_0(z_k)-F_0(z_{k-1})$ достаточно хорошо отражали вид функции распределения $F_0(x)$.

Следует заметить, что пользоваться предельным соотношением (5) и вытекающим из него приближенным равенством (6) можно лишь в том случае, когда точки деления z_k выбираются независимо от выборки (1).

Если функция распределения $F(x; \theta_1, \ldots, \theta_s)$ элементов выборки (1) зависит от неизвестных параметров $\theta_1, \ldots, \theta_s$, то мы можем подобрать значения этих параметров такими, чтобы статистика (4) обратилась в минимум. Обозначим

$$\tilde{\chi}^2 = \min_{\theta_1, \ldots, \theta_s} \sum_{k=1}^{s} \frac{(\nu_k - np_k(\theta_1, \ldots, \theta_s))^2}{np_k(\theta_1, \ldots, \theta_s)},$$

где $p_k(\theta_1, \ldots, \theta_s) = F(z_k; \theta_1, \ldots, \theta_s) - F(z_{k-1}; \theta_1, \ldots, \theta_s)$. Если минимум в (8) достигается при $\theta_l = \hat{\theta}_l$, то

$$\tilde{\chi}^2 = \sum_{k=1}^{r} \frac{(\nu_k - np_k(\hat{\theta}_1, \ldots, \hat{\theta}_s))^2}{np_k(\hat{\theta}_1, \ldots, \hat{\theta}_s)}, \tag{8}$$

где $\hat{\theta}_l$ — оценки параметров θ_l, т. е. $\hat{\theta}_l$ — функции от выборки (1), которые можно принять за приближенные значения θ_l. В этом случае вместо предельного соотношения (5) имеет место следующее утверждение:

$$\lim_{n\to\infty} \mathbf{P}\{\tilde{\chi}^2 \leqslant x\} = \int_0^x k_{r-s-1}(u)\,du,$$

т. е. *функция распределения* $\mathbf{P}\{\tilde{\chi}^2 \leqslant x\}$ *приближенно равна функции распределения* χ^2 *с* $r-s-1$ *степенями свободы.* Это же самое утверждение справедливо, если в равенстве (8) оценки $\hat{\theta}_l$ получены каким-либо другим способом, но достаточно быстро при $n\to\infty$ приближаются к θ_l.

Пример 1. Пусть выборка (1) получена из нормального распределения с параметрами (a, σ), т. е. ее элементы имеют функцию распределения

$$\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{(u-a)^2}{2\sigma^2}}\,du = \Phi\left(\frac{x-a}{\sigma}\right).$$

Оценим параметры a и σ следующим образом:

$$\hat{a}=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad \hat{\sigma}^2=s^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

Положим $F_0(x)=\Phi\left(\frac{x-\overline{X}}{s}\right)$ и далее действуем так же, как описано выше: определяем по z_k вероятности p_k, вычисляем величину χ^2. Только теперь, поскольку мы по выборке определили $s=2$ параметра, вместо приближенного равенства (6), надо воспользоваться приближенным равенством

$$\mathsf{P}\{\chi^2\leqslant x\}\approx\int_0^x k_{r-3}(u)\,du.$$

В конце книги приведена таблица значений $\chi_{\alpha,m}$ для $m=1\div 25$, $\alpha=0{,}05$, $\alpha=0{,}01$.

14.2. Выбор из двух гипотез. Часто возникает следующая задача. Пусть имеется выборка (1). Относительно функции распределения $F(x)$ этой выборки имеются две гипотезы: основная гипотеза H_0: $F(x)=F_0(x)$, и конкурирующая гипотеза H_1: $F(x)=F_1(x)$. По выборке (1) надо определить, какая гипотеза имеет место на самом деле. Для решения этой задачи строят *статистический критерий*, который состоит в следующем: гипотеза H_0 отвергается (и, следовательно, принимается конкурирующая гипотеза H_1), если выборка $(X_1, \ldots, X_n)$ попадает в некоторое заранее выбранное *критическое* множество S, т. е.

$$(X_1, \ldots, X_n)\in S;$$

гипотеза H_0 принимается, если

$$(X_1, \ldots, X_n)\notin S.$$

Обозначим $\mathsf{P}_i(S)$ вероятность события $(X_1, \ldots, X_n) \in S$, если верна гипотеза H_i. Каждый критерий характеризуется вероятностями ошибок первого и второго родов. *Вероятность* α *ошибки первого рода* определяется равенством $\alpha = \mathsf{P}_0(S)$ и равна вероятности отвергнуть основную гипотезу, если она верна. *Вероятность* β ошибки второго рода определяется равенством $\beta = \mathsf{P}_1(\bar{S})$, где $\bar{S} = R^n \setminus S$ — дополнение множества S, и равна вероятности принять основную гипотезу H_0, если верна конкурирующая гипотеза H_1. Вероятность $1 - \beta = \mathsf{P}_1(S)$ называется *мощностью* критерия.

С помощью вероятностей α и β ошибок первого и второго рода критерии можно сравнивать между собой и решать задачу о наилучшем критерии. Обычно фиксируют какое-либо малое значение α вероятности ошибки первого рода (эту вероятность часто называют *уровнем значимости* критерия) и ищут критерий с уровнем значимости α и с наибольшей мощностью $1 - \beta$ (или, что равносильно, с наименьшей вероятностью β ошибки второго рода). Такой критерий, если он существует, называется *наиболее мощным*, или *оптимальным*, критерием.

Оптимальным является так называемый *критерий Неймана — Пирсона*, который состоит в следующем. Пусть $p_0(x) = F_0'(x)$, $p_1(x) = F_1'(x)$ — плотности распределения выборки (1) соответственно при основной H_0 и конкурирующей H_1 гипотезах. Тогда критическое множество оптимального критерия надо искать среди множеств S следующего вида:

$(X_1, \ldots, X_n) \in S$ тогда и только тогда, когда

$$\frac{p_1(X_1)\, p_1(X_2) \ldots p_1(X_n)}{p_0(X_1)\, p_0(X_2) \ldots p_0(X_n)} > C \tag{9}$$

при некотором числе $C > 0$.

Пример 2. Пусть выборка (1) взята из нормального распределения с параметрами (a, σ), σ — известно и одно и то же для обеих гипотез. Гипотезы H_0 и H_1 определяются равенствами:

$$H_0\colon a = a_0; \quad H_1\colon a = a_1 > a_0.$$

В этом случае неравенство (9) превращается в следующее:

$$\frac{\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i-a_1)^2}}{\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i-a_0)^2}} > C,$$

которое равносильно неравенству

$$\overline{X} > C_1 \tag{10}$$

при некотором C_1, зависящем от C, a_0, a_1 и σ. Поскольку при гипотезе H_i случайная величина $\overline{X}$ имеет нормальное распределение с параметрами $(a_i, \sigma/\sqrt{n})$, то ошибки первого и второго рода равны

$$\alpha = \mathsf{P}_0\left(\frac{\overline{X}-a_0}{\sigma/\sqrt{n}} > \frac{C_1-a_0}{\sigma/\sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{C_1-a_0}{\sigma/\sqrt{n}}}^{\infty} e^{-\frac{u^2}{2}}\,du = 1-\Phi\left(\frac{C_1-a_0}{\sigma/\sqrt{n}}\right), \tag{11}$$

$$\beta = \mathsf{P}_1\left(\frac{\overline{X}-a_1}{\sigma/\sqrt{n}} \leqslant \frac{C_1-a_1}{\sigma/\sqrt{n}}\right) = \Phi\left(\frac{C_1-a_1}{\sigma/\sqrt{n}}\right).$$

Определим u_γ как решение уравнения

$$\Phi(u_\gamma) = 1-\gamma.$$

В силу симметрии плотности нормального распределения с параметрами (0, 1) $u_{1-\gamma} = -u_\gamma$. Тогда из (11) следует

$$\frac{C_1-a_0}{\sigma/\sqrt{n}} = u_\alpha, \quad \frac{C_1-a_1}{\sigma/\sqrt{n}} = -u_\beta$$

или

$$C_1 = a_0 + u_\alpha \frac{\sigma}{\sqrt{n}}, \quad C_1 = a_1 - u_\beta \frac{\sigma}{\sqrt{n}}. \tag{12}$$

Таким образом, критерий (10) с константой C_1, определенной первым равенством (12), дает при заданном значении α наименьшую вероятность ошибки второго рода β.

Из равенств (12) мы можем определить объем выборки n, при котором оптимальный критерий имеет вероятности ошибок первого и второго рода α и β:

$$n = \sigma^2 \frac{(u_\alpha + u_\beta)^2}{(a_1 - a_0)^2}.$$

ТАБЛИЦЫ

НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Значения функции $\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{u^2}{2}}\,du$

Таблица 1

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0200	0,0239	0,0279	0,0319	0,0359
0,1	398	438	478	517	557	596	636	675	714	753
0,2	793	832	871	910	948	987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	406	443	480	517
0,4	554	591	628	664	700	736	772	808	844	879
0,5	915	950	985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	357	389	422	454	486	517	549
0,7	580	611	642	673	703	734	764	794	823	852
0,8	881	910	939	967	995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	289	315	340	365	389
1,0	413	437	461	485	508	581	554	577	599	621
1,1	643	665	686	708	729	749	770	790	810	830
1,2	849	869	888	907	925	944	962	980	997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	192	207	222	236	251	265	279	292	306	319
1,5	332	345	357	370	382	394	406	418	429	441

1,6	452	463	474	484	495	505	515	525	535	545
1,7	554	564	573	582	591	599	608	616	625	633
1,8	641	649	656	664	671	678	686	693	699	706
1,9	713	719	726	732	738	744	750	756	761	767
2,0	772	778	783	788	793	798	803	808	812	817
2,1	821	826	830	834	838	842	846	850	854	857
2,2	861	864	868	871	875	878	881	884	887	890
2,3	893	896	898	901	904	906	909	911	913	916
2,4	918	920	922	925	927	929	931	932	934	936
2,5	938	940	941	943	945	946	948	949	951	952
2,6	953	955	956	957	959	960	961	962	963	964
2,7	965	966	967	968	969	970	971	972	973	974
2,8	974	975	976	977	977	978	979	979	980	981
2,9	981	982	982	983	984	984	985	985	985	986
3,0	987	987	987	988	988	989	989	989	990	990

Значения функции u_α

Функция u_α определяется равенством $\alpha = \int\limits_{u_\alpha}^{\infty} e^{-\frac{x^2}{2}}\, dx$

Таблица 2

α	0,001	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040	0,045	0,050
u_α	3,0902	2,5758	2,3263	2,1701	2,0537	1,9600	1,8808	1,8119	1,7507	1,6954	1,6449

РАСПРЕДЕЛЕНИЕ ПУАССОНА

Значения функции $p_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$

Таблица 3

k \ λ	0,1	0,2	0,3	0,4	0,5
0	0,90484	0,81873	0,74082	0,67032	0,60653
1	0,09048	0,16375	0,22225	0,26813	0,30327
2	0,00452	0,01638	0,03334	0,05363	0,07582
3	0,00015	0,00109	0,00333	0,00715	0,01264
4		0,00006	0,00025	0,00072	0,00158
5			0,00002	0,00006	0,00016
6					0,00001

k \ λ	0,6	0,7	0,8	0,9
0	0,54881	0,49659	0,44933	0,40657
1	0,32929	0,34761	0,35946	0,36591
2	0,09879	0,12166	0,14379	0,16466
3	0,01976	0,02839	0,03834	0,04940
4	0,00296	0,00497	0,00767	0,01112
5	0,00036	0,00070	0,00123	0,00200
6	0,00004	0,00008	0,00016	0,00030
7		0,00001	0,00002	0,00004

k \ λ	1,0	2,0	3,0	4,0	5,0
0	0,36788	0,13534	0,04979	0,01832	0,00674
1	0,36788	0,27067	0,14936	0,07326	0,03369
2	0,18394	0,27067	0,22404	0,14653	0,08422
3	0,06131	0,18045	0,22404	0,19537	0,14037
4	0,01533	0,09022	0,16803	0,19537	0,17547
5	0,00307	0,03609	0,10082	0,15629	0,17547
6	0,00051	0,01203	0,05041	0,10419	0,14622
7	0,00007	0,00344	0,02160	0,05954	0,10445
8	0,00001	0,00086	0,00810	0,02977	0,06528
9		0,00019	0,00270	0,01323	0,03627
10		0,00004	0,00081	0,00529	0,01813
11		0,00001	0,00022	0,00193	0,00824
12			0,00006	0,00064	0,00343
13			0,00001	0,00020	0,00132
14				0,00006	0,00047
15				0,00002	0,00016
16					0,00005
17					0,00001

χ^2-РАСПРЕДЕЛЕНИЕ

Значения функции $\chi^2_{\alpha,\, m}$

Функция $\chi^2_{\alpha,\, m}$ определяется равенством $P\{\chi^2_m > \chi^2_{\alpha,\, m}\} = \alpha$, где случайная величина χ^2_m имеет χ^2-распределение с m степенями свободы. Плотность распределения χ^2_m равна

$$p_{\chi^2_m}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}}\, x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \qquad x > 0.$$

Таблица 4

α \ m	1	2	3	4	5	6	7	8	9
0,05	3,8	6,0	7,8	9,5	11,1	12,6	14,1	15,5	16,9
0,01	6,6	9,2	11,3	13,3	15,1	16,8	18,5	20,1	21,7

α \ m	10	11	12	13	14	15	16	17	18
0,05	18,3	19,7	21,0	22,4	23,7	25,0	26,3	27,6	28,9
0,01	23,2	24,7	26,2	27,7	29,1	30,6	32,0	33,4	34,8

α \ m	19	20	21	22	23	24	25
0,05	30,1	31,4	32,7	33,9	35,2	36,4	37,7
0,01	36,2	37,6	38,9	40,3	41,6	43,0	44,3

ПРИЛОЖЕНИЕ

ДОКАЗАТЕЛЬСТВО ЛОКАЛЬНОЙ ТЕОРЕМЫ МУАВРА-ЛАПЛАСА

Оценим логарифм вероятности

$$P_n(m) = \mathsf{P}\{\mu_n = m\} = \frac{n!}{m!\,(n-m)!}\, p^m q^{n-m},$$

равный

$$\ln P_n(m) = \ln n! - \ln m! - \ln(n-m)! + m \ln p + (n-m) \ln q.$$

Так как

$$m = np + x_m \sqrt{npq}, \qquad n - m = nq - x_m \sqrt{npq},$$

то, воспользовавшись формулой Стирлинга

$$\ln n! = \ln \sqrt{2\pi n} + n \ln n - n + O\left(\frac{1}{n}\right),$$

получим

$$\ln m! = \frac{1}{2} \ln(2\pi m) +$$
$$+ (np + x_m \sqrt{npq}) \ln(np + x_m \sqrt{npq}) - np - x_m \sqrt{npq} + O\left(\frac{1}{n}\right)$$

и

$$\ln(n-m)! = \frac{1}{2} \ln(2\pi(n-m)) +$$
$$+ (nq - x_m \sqrt{npq}) \ln(nq - x_m \sqrt{npq}) - nq + x_m \sqrt{npq} + O\left(\frac{1}{n}\right).$$

Используя эти формулы, а также формулы

$$\ln(np + x_m \sqrt{npq}) = \ln np + x_m \sqrt{\frac{q}{np}} - \frac{x_m^2}{2} \frac{q}{np} + O\left(\frac{q^3}{n\sqrt{n}\,p^3}\right),$$

$$\ln(nq - x_m \sqrt{npq}) = \ln nq - x_m \sqrt{\frac{p}{nq}} - \frac{x_m^2}{2} \frac{p}{nq} + O\left(\frac{p^3}{n\sqrt{n}\,q^3}\right),$$

нетрудно получить утверждение доказываемой теоремы.

ОТВЕТЫ К ЗАДАЧАМ

§ 1

1. $A(\overline{B}+C)$. 2. 1) $A\overline{BC}$, 2) ABC, 3) ABC, 4) $A+B+C$ 5) $\overline{ABC}$. 3. $\left(C_{95}^{10}+5C_{95}^{9}\right)/C_{100}^{10}=0{,}923143\ldots$ 4. $\frac{12!}{12^{12}}=0{,}0000537\ldots$ 5. $C_{37}^{6}/C_{49}^{6}=0{,}1662481\ldots$ 6. $P(A)=3!/3^3=\frac{2}{9}$; $P(B)=3\cdot2\cdot3/3^3=2/3$. 7. $\pi/4$. 8. а) $\min(x,\ 1)$, $x\geqslant 1$; б) $4x(1-x)$, $0\leqslant x\leqslant 1/2$; $x\geqslant 1/2$. 9. $\left(1-2\frac{r\sqrt{3}}{3a}\right)^2$. 10. 1/4. 11. $p_n=\sum_{k=1}^{n}(-1)^{k+1}\frac{1}{k!}\to 1-e^{-1}$, $n\to\infty$.

§ 2

1. 1/12. 2. 1/3. 4. C_1, C_2, C_3 не являются взаимно независимыми; C_1 C_2 и C_3 — зависимы. 5. 1) $p_1(1-p_3)p_4$, 2) $p_1+p_2-p_1p_2$, 3) $(p_1+p_2-p_1p_2)\cdot(p_3+p_4-p_3p_4)$. 6. $(1-p_3)(1-p_1p_2)$. 7. 1) $\alpha_1\alpha_2$, 2) $\beta_1+\beta_2-\beta_1\beta_2$. 8. 11/18. 9. 1) $p(1-\beta_1)(1-\beta_2)+(1-p)\alpha_1\alpha_2$, 2) $p(1-\beta_1)(1-\beta_2)/[p(1-\beta_1)(1-\beta_2)+(1-p)\alpha_1\alpha_2]$. 10. 20/21. 11. 9/16. 12. $e^{-\lambda t}$. 13. $\pi_0=1-\theta$, $\pi_k=(1-\theta)\theta^k$, $k\geqslant 1$.

§ 3

1. 3/5; 2/5. 2. а) 0,95009... б) 0,0480298... в) 0,0009782... 3. 21; $Q(21)=0{,}9907768\ldots$ 4. $C_{10}^{4}(5/72)^4(1-5/72)^6=0{,}00317\ldots$

§ 4

1. 0,99. 2. 0,001. 3. 0,00015; 0,00016. 4. а) 0,13534, б) 0,67668. 5. $(0{,}9)^n\leqslant 0{,}1$; $n\geqslant 22$. 6. 96.

§ 5

1. а) $\mathsf{P}\{\chi_A=1\}=1/2$; б) $\mathsf{P}\{\chi_B=1\}=1/3$; в) $\mathsf{P}\{\xi=0\}=1/3$ $\mathsf{P}\{\xi=1\}=1/2$, $\mathsf{P}\{\xi=2\}=1/6$. 2. $\mathsf{P}\{\xi=7\}=7/16$, $\mathsf{P}\{\xi=9\}=5/16$, $\mathsf{P}\{\xi=11\}=3/16$, $\mathsf{P}\{\xi=13\}=1/16$. 3. $\mathsf{P}\{\xi=k\}=1/28$, $k=0, 1, 11, 12$; $\mathsf{P}=\{\xi=k\}=1/14$, $k=2, 3, 9, 10$; $\mathsf{P}\{\xi=k\}=$

$=3/28$, $k=4, 5, 7, 8$; $\mathsf{P}=\{\xi=6\}=1/7$. 4. $\mathsf{P}(A)=(1+e^{-2a})/2$, $\mathsf{P}(B)=(1-e^{-2a})/2$. 5. $C=1/2$; $F_\xi(x)=0$ при $x\leqslant 1$; $F_\xi(x)=1-\frac{1}{\sqrt{x}}$ при $x\geqslant 1$. 6. $F_\xi(x)=x^2/r^2$, $0\leqslant x\leqslant r$; $p_\xi(x)=2x/r^2$, $0\leqslant x\leqslant r$; $p_\xi(x)=0$ в остальных случаях. 7. $\mathsf{P}\{\nu=k\}=(21-2k)/100$, $k=1, 2, \ldots, 10$. 8. $2-\sqrt{2}$. 9. $(1+\sqrt{2})/4$. 10. $F_\eta(x)=0$ при $x<0$, $F_\eta(x)=x$ при $0\leqslant x\leqslant 1$; $F_\eta(x)=1$ при $x\geqslant 1$, $p_\eta(x)=1$ при $0\leqslant x\leqslant 1$ и $p_\eta(x)=0$ в остальных случаях. 11. $p_\eta(x)=1/3\sqrt{x}$ при $0\leqslant x\leqslant 1$, $p_\eta(x)=1/6\sqrt{x}$ при $1\leqslant x\leqslant 4$; $p_\eta(x)=0$ в остальных случаях.

§ 6

1. а) $\pi/4$; б) $0=\mathsf{P}\{1/2\leqslant\xi\leqslant 1, 1/2\leqslant\eta\leqslant 1\}\neq\mathsf{P}\{1/2\leqslant\xi\leqslant 1\}\cdot\mathsf{P}\{1/2\leqslant\eta\leqslant 1\}=\frac{1}{4^2}$. 2. $\mathsf{P}\{\zeta=6\}=1/3$, $\mathsf{P}\{\zeta=10\}=1/3$, $\mathsf{P}\{\zeta=12\}=1/3$. 3. $F_\theta(x)=x^2$, $0\leqslant x\leqslant 1$; $p_\theta(x)=2x$, $0\leqslant x\leqslant 1$. 4. $p_{\xi+\eta}(x)=\lambda^2xe^{-\lambda x}$, $x\geqslant 0$.

§ 7

1. 3,5. 2. 6. 3. $1/\lambda$. 4. $(1-p)/\lambda$. 5. $R/3$. 6. $12(1-1/12)^{25}=1{,}36292\ldots$ 7. $\pi hr^2+\frac{1}{3}\pi\delta^2h$.

§ 8

1. $\pi/12$, $\pi^2/180$. 2. $1/\ln 2$, $1/\ln 2$, $(3\ln 2-2)/\ln^2 2$. 3. $1/3$; $1/18$. 4. $(m+n)/n$; $n(m+n)/m^2$. 6. $p^2(1+2p-3p^2)$.

§ 9

1. $\mathsf{M}\xi_1=-1/8$, $\mathsf{M}\xi_2=0$, $\mathsf{D}\xi_1=133/192$, $\mathsf{D}\xi_2=1$, $\operatorname{cov}(\xi_1, \xi_2)=-1/8$. 2. а) 0; 0; 0,5; 0,5. б) 0; 0; случайные величины зависимы. 3. $C=4$; $\operatorname{cov}(\xi_1, \xi_2)=0$; случайные величины независимы. 4. $C=1/4$, $\rho(\xi_1, \xi_2)=-1/5$. 5. а) 0, б) 1/3.
6. а) $\mathsf{P}\{\xi=\eta=i\}=i/36$; $\mathsf{P}\{\xi=i, \eta=j\}=1/36$, $i<j$; $\mathsf{P}\{\xi=i, \eta=j\}=0$, $i>j$. б) $\mathsf{M}\xi=7/2$, $\mathsf{M}\eta=161/36$, $\mathsf{D}\xi=35/12$, $\mathsf{D}\eta=2555/1296$; в) $\operatorname{cov}(\xi, \eta)=105/72$.

§ 10

1. $\mathsf{P}\{|\xi-a|\geqslant 2\sigma\}\leqslant 0{,}25$; $\mathsf{P}\{|\xi-a|\geqslant 2\sigma\}=0{,}0456$. 2. $\mathsf{P}\{|\xi|\geqslant\Delta\}=\sigma^2/\Delta^2$; $\mathsf{P}\{|\xi|\geqslant\Delta\}\leqslant\mathsf{D}\xi/\Delta^2=\sigma^2/\Delta^2$. 3. $\Delta=0{,}0813$; по неравенству Чебышёва $\Delta\leqslant 1/\sqrt{10}=0{,}3162\ldots$ 4. Да.

§ 11

1. $0{,}75\cdot 10^{-m+2}$. 3. $\Phi(x)$.

ЛИТЕРАТУРА

[1] Большев Л. Н. Смирнов Н. В. Таблицы математической статистики. — М.: Наука, 1965.

[2] Боровков А. А. Теория вероятностей. — М.: Наука, 1976.

[3] Володин Б. Г. и др. Сборник задач по теории вероятностей, математической статистике и теории случайных функций/Под ред. А. А. Свешникова. — М.: Наука, 1965.

[4] Гнеденко Б. В. Курс теории вероятностей. — 5-е изд. — М.: Наука, 1971.

[5] Ермаков С. М. Метод Монте-Карло и смежные вопросы. — М.: Наука, 1971.

[6] Колмогоров А. Н.*) Основные понятия теории вероятностей. — М.: Наука, 1974.

[7] Крамер Г. Математические методы статистики. — 2-е изд./Перев. с англ. — М.: Мир, 1975.

[8] Прохоров Ю. В., Розанов Ю. А. Теория вероятностей (Основные понятия, предельные теоремы, случайные процессы). — 2-е изд. — М.: Наука, 1973.

[9] Розанов Ю. А. Случайные процессы. — 2-е изд. — М.: Наука, 1979.

[10] Севастьянов Б. А. Курс теории вероятностей и математической статистики. — М.: Наука, 1982.

[11] Севастьянов Б. А., Чистяков В. П., Зубков А. М. Сборник задач по теории вероятностей. — М.: Наука, 1980.

[12] Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. — 3-е изд. — М.: Наука, 1969.

[13] Феллер В. Введение в теорию вероятностей и ее приложения, т. 1/Перев. с англ. — М.: Мир, 1967.

[14] Чистяков В. П. Курс теории вероятностей. — 2-е изд. — М.: Наука, 1982.

*) Книга А. Н. Колмогорова «Основные понятия теории вероятностей» была впервые издана в 1933 г. на немецком языке.

Владимир Константинович Захаров
Борис Александрович Севастьянов
Владимир Павлович Чистяков

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Редактор *И. Е. Морозова*
Технический редактор *Е. В. Морозова*
Корректор *Н. Б. Румянцева*

ИБ № 12308

Сдано в набор 30.09.82. Подписано к печати 28.04.83. Формат 84×108¹/₃₂. Бумага типографская № 3. Литературная гарнитура. Высокая печать. Условн. печ. л. 8,4. Уч.-изд. л. 8,1. Тираж 84 000 экз. Заказ № 382. Цена 25 коп.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

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