В. П. ЧИСТЯКОВ

# Курс теории вероятностей

*Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов высших технических учебных заведений*

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1978

517.8
4-68
УДК 519.2

**Курс теории вероятностей.** В. П. Чистяков. Главная редакция физико-математической литературы издательства «Наука», М., 1978, 224 стр.

В основу учебного пособия положен полугодовой курс лекций по теории вероятностей, читавшийся в течение ряда лет в МИФИ. В книге дается математическое изложение разделов теории вероятностей, традиционных для полугодового курса; при этом используются только факты из обычного курса математического анализа технических вузов. В книге изложены также элементы математической статистики и рассмотрен ряд примеров случайных процессов. Приведены решения примеров и задач; имеются задачи для самостоятельного решения. В конце книги помещены таблицы основных распределений, небольшая таблица случайных чисел и ответы к задачам. Книга предназначена студентам технических вузов, преподавателям и инженерам.

Илл. 9, библ. 18.



*Владимир Павлович Чистяков*

КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ

М., 1978 г., 224 стр. с илл.
Редактор *В. В. Донченко*
Техн. редактор *Н. В. Кошелева*
Корректоры *З. В. Автонеева, Н. Б. Румянцева*

ИБ № 11007

Сдано в набор 02.09.77. Подписано к печати 17.01.78.
Бумага $84\times108^1/_{32}$, тип. № 3. Литературная гарнитура. Высокая печать.
Условн. печ. л. 11,8. Уч.-изд. л. 10,7. Тираж 80 000 экз. Заказ № 1895.
Цена книги 35 коп.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени
Первая Образцовая типография имени А. А. Жданова
Союзполиграфпрома при Государственном комитете
Совета Министров СССР по делам издательств,
полиграфии и книжной торговли.
Москва, М-54, Валовая, 28

Ч $\dfrac{20203-031}{053(02)-78}$ 17-78

# ОГЛАВЛЕНИЕ

## ПРЕДИСЛОВИЕ

В книге дается математическое изложение некоторых разделов теории вероятностей, основанное на обычном курсе математического анализа технических вузов. Теория меры, интегралы Лебега и Стилтьеса не используются. В связи с этим рассматриваются вероятностные пространства, в которых при задании вероятности используются ряды, интегралы Римана и несобственные интегралы. К таким пространствам относятся конечные и дискретные, а также вероятностные пространства, в которых вероятность события определяется как интеграл Римана от некоторой положительной функции. Последние для удобства ссылок названы абсолютно непрерывными (термин не является общепринятым). Этих пространств достаточно для изучения тем, которые обычно включаются в курс теории вероятностей: «классические» и «геометрические» вероятности; конечные последовательности испытаний (в частности, цепи Маркова, независимые испытания, схема Бернулли); случайные величины и т. д.

Однако в ряде интересных задач нельзя обойтись указанными выше вероятностными пространствами. В задаче о разорении игрока, при изучении времени до первого успеха в схеме Бернулли, в ряде задач, связанных с приложениями формулы полной вероятности, требуется введение более сложных вероятностных пространств. Такие задачи рассматриваются в предположении, что подходящие вероятностные пространства существуют. Некоторое представление об определении вероятности в более сложных пространствах дается в § 4 гл. 3. Части текста, набранные петитом, связаны главным образом с объяснением определения вероятности в более сложных случаях. При первоначальном чтении их можно опустить.

Курс предлагаемого типа автор читал много лет. В зависимости от математической подготовки слушателей менялся только объем излагаемого материала и подбор задач; аксиоматическое изложение сохранялось в любом случае.

Другой подход к способу изложения теории вероятностей в технических вузах состоит в использовании интуитивных представлений о вероятности, независимости, случайной величине вместо их точных определений. Это приводит к замене теории вероятностей решением разрозненных задач из анализа. Точную характеристику подхода такого типа дает Феллер в предисловии к своей книге [17]: «При преподавании теории вероятностей существует тенденция возможно быстрее сводить вероятностные задачи к задачам чистого анализа, избегая специфических особенностей самой теории вероятностей. Такое изложение основывается на плохо определяемом понятии случайной величины, вводимом обычно вначале. В противоположность этому настоящая книга строится на понятии пространства элементарных событий, без которого случайные величины остаются искусственной выдумкой».

Следует отметить, что аналогичные вопросы в преподавании математического анализа давно решены. Обычно не возникают сомнения в том, что нужно давать точное определение предела на языке «$\varepsilon-\delta$», хотя оно не очень легко согласуется с интуитивным представлением о пределе. Теория вероятностей даже в кратком изложении, должна оставаться математической дисциплиной.

Написанный вариант «Курса теории вероятностей» наиболее близок к полугодовому курсу, читаемому автором в последние годы в Московском инженерно-физическом институте (МИФИ).

В зависимости от профиля вуза и математической подготовки студентов разделы «Элементы математической статистики», «Элементы теории случайных процессов» можно исключить из курса, сохранить в объеме гл. 10 и 11 или расширить. Гл. 10 легко дополнить, если воспользоваться книгой Н. В. Смирнова, И. В. Дунина-Барковского [16]; сведения по теории случайных процессов можно найти в книге Ю. А. Розанова [13]. Обзор важнейших результатов, методов и направлений теории вероятностей содержится в книге Ю. В. Прохорова, Ю. А. Розанова [12].

По ходу изложения «Курса теории вероятностей» дается решение некоторого количества задач. Это не устраняет необходимости в самостоятельном решении задач, которые приведены в конце глав. Некоторые из задач являются существенным дополнением к изложенному в основном тексте. Знаком * отмечены более трудные задачи или задачи, в формулировке которых кратко вводятся понятия, не определенные в тексте. При подборе задач были использованы многие задачники, учебники и монографии (в частности, [4], [7], [9], [11], [17], [18]). Приведенный набор задач желательно пополнить задачами, отражающими специфику вуза, в котором будет читаться курс, предлагаемого типа.

В каждой главе принята своя нумерация формул, теорем, лемм. При ссылке на формулу, теорему или лемму внутри главы указывается только этот номер; при ссылке на формулу, теорему или лемму из другой главы к номеру добавляется номер главы. Например, (1.5.8) означает формулу (5.8) из § 5 гл. 1.

Я пользуюсь случаем, чтобы поблагодарить сотрудников Математического института, с которыми неоднократно обсуждались вопросы преподавания теории вероятностей. Особенно интенсивными были беседы с Л. Н. Большевым, В. К. Захаровым, О. В. Висковым во время работы по составлению новой программы курса теории вероятностей.

Мне были очень полезны критические замечания и советы Б. А. Севастьянова и В. Ф. Колчина, сделанные ими при многочисленных обсуждениях различных разделов курса лекций по теории вероятностей, а также их замечания, связанные непосредственно с рукописью данной книги.

# ВВЕДЕНИЕ

До возникновения теории вероятностей объектом исследования науки были явления или опыты, в которых условия практически однозначно определяют исход. Классическим примером является механика. Пусть, например, на материальную точку действует сила тяжести. Если в некоторый момент задать положение и скорость материальной точки, то ее дальнейшее движение определяется однозначно, так как соответствующее дифференциальное уравнение имеет единственное решение. Однако если эту механическую модель применить к реальным физическим телам, например рассматривать движение пули, то однозначно траектория уже не будет определяться. Начальная скорость заранее точно неизвестна; при различных выстрелах она может принимать различные значения. Таким образом, неопределенность здесь вносится случайностью начальной скорости. Если колебания значений начальной скорости невелики (например, меньше ошибки при численном интегрировании уравнения), то можно использовать детерминированную механическую модель, в которой движение однозначно определяется начальными условиями.

Неоднозначность исхода при сохранении основных условий опыта наблюдается для широкого круга явлений. При подбрасывании монеты мы не можем предсказать исход: выпадет монета гербом вверх или нет. Результаты нескольких измерений, полученных одним и тем же прибором в одних и тех же условиях, различны. Влияние очень большого числа разнообразных причин, каждая из которых в отдельности не может повлиять на результат опыта, приводит к тому, что результат опыта не определяется заранее однозначно; говорят, что результат такого опыта случаен.

Примеры случайных явлений можно указать во многих областях науки и техники (например, в физике, биологии, демографии, в массовом производстве, в системах автоматического управления и т. д.).

Индивидуальные результаты опытов, подбрасываний монеты, измерений, ведут себя очень «неправильно». Однако при наблюдении результатов большой последовательности опытов обнаруживаются интересные закономерности. Если обозначить $N_\Gamma$ число выпадений герба при $N$ подбрасываниях монеты, то оказывается, что частота $N_\Gamma/N$ с ростом $N$ становится близкой к $1/2$. Это явление имеет общий характер.

Частота наступления какого-либо события в серии опытов, повторяемых в одинаковых условиях, с ростом числа опытов приближается к некоторому числу $p$, $0 < p < 1$. Этот факт устойчивости частот неоднократно проверялся и может считаться экспериментально установленным.

Изучением случайных явлений занимается теория вероятностей. Каждому событию приписывается особая числовая мера объективной возможности его появления — вероятность.

Величина вероятности каждого события проявляется в той частоте, с которой оно встречается при повторениях опыта. Естественно потребовать, чтобы определение вероятности отражало свойства частот.

Пусть в каждом из $N$ опытов может наступить только одно из двух событий, $A$ или $B$. (Например, пусть $N$ раз брошена игральная кость; событие $A$ — при одном бросании игральной кости выпало четное число очков, $B$ — выпала «3».) Обозначим $C$ более сложное событие, состоящее в том, что произошло $A$ или $B$ (в примере с костью $C$ — число очков либо четно, либо равно 3). Обозначим $N_A$ — число опытов, в которых наступило $A$; аналогично определяем $N_B$ и $N_C$ для событий $B$ и $C$. Так как в каждом опыте из двух событий $A$ и $B$ может наступить только одно, то $N_C = N_A + N_B$. Нетрудно проверить следующие свойства частот:

1°. При любом $A$

$$\frac{N_A}{N} \geqslant 0. \tag{1}$$

2°. Если $\Omega$ достоверное событие, то

$$\frac{N_\Omega}{N} = 1. \tag{2}$$

3°. Если $A$ и $B$ не могут произойти одновременно и $C$ состоит в том, что произошло $A$ или $B$, то

$$\frac{N_C}{N} = \frac{N_A}{N} + \frac{N_B}{N}. \tag{3}$$

В гл. 1 после обсуждения связи реального явления с его математической моделью будет дано аксиоматическое определение вероятности.

ГЛАВА 1

# ВЕРОЯТНОСТНЫЕ ПРОСТРАНСТВА

## § 1. Математическая схематизация случайных явлений

Теория вероятностей, так же как и другие разделы математики, имеет дело не с явлениями окружающего мира непосредственно, а с их математическими моделями. В математической модели должны быть правильно переданы существенные стороны изучаемого явления, а несущественные — отброшены. Слишком подробное описание изучаемого явления приводит к усложнению математической модели и может значительно затруднить дальнейшее исследование. Излишнее упрощение модели может привести к неверным выводам. Насколько удачно введена модель, можно в каждом конкретном случае судить по согласованности теоретических выводов с опытом.

В повседневной речи мы часто используем слова «вероятность», «случай», «событие». В простейших явлениях уже интуитивное представление о вероятности позволяет давать правильные ответы. Например, хорошо известно, что вероятность выпадения герба при подбрасывании монеты равна 1/2. Объяснение приводится достаточно убедительное:

1) возможны два события — выпал герб, выпала решетка; ни одному из них нельзя отдать предпочтение, так как монета симметрична;

2) при многократном повторении опыта частота появления герба близка к 1/2.

Первая часть этого объяснения является попыткой построить модель случайного явления; вторая часть — экспериментальная проверка соответствия модели физическому явлению.

Однако при незначительном усложнении опыта «повседневная интуиция» или «здравый смысл» могут подводить. Представим себе, что при каждом из 10 подбрасываний монеты выпал герб. Предлагается угадать, какой стороной упадет монета в следующий раз. Довольно распространено мнение, что выпадение решетки более вероятно. Этот «вывод» делается из наблюдения результатов подбрасывания монеты примерно следующим образом: длинные серии подряд идущих гербов встречаются не часто, а если известно, что длинная серия начинает появляться, то она должна быстрее кончиться и, следовательно, появление решетки более вероятно. Отсутствие длинных серий гербов нетрудно заметить, даже если не записывать результаты опытов. Однако экспериментальная оценка вероятности появления решетки после 10 гербов требует уже фиксирования результатов опыта. Нужно отобрать все цепочки из 10 гербов и посмотреть, сколько раз в следующем результате будет герб и сколько раз решетка. Ответ 1/2 в этом более сложном опыте свидетельствует уже о более развитой интуиции.

В качестве примера совсем неприемлемого рассуждения можно привести следующий: я *не знаю*, пойдет сегодня дождь или нет, *следовательно*, его вероятность 1/2. Это—тоже модель, однако, к дождю она вряд ли имеет какое-нибудь отношение. Отметим, что в примере с монетой учитывалось ее свойство—симметричность. В примере с дождем ничего не учитывается. Многообразие задач, встречающихся в приложениях, требует четкого определения понятий, связанных со случайными явлениями.

Имеется широкий круг случайных явлений, когда при многократном повторении эксперимента в одних и тех же условиях доля наблюдений, в которых имеет место некоторое событие $A$, становится близкой к некоторому числу $P$. Тем самым число $P$ является объективной количественной оценкой возможности появления события $A$. На языке математической схемы, которую мы должны построить, «вероятность события $A$ равна $P$». Таким образом, мы будем рассматривать вероятность как функцию от случайного события. В курсах математического анализа прежде, чем приступить к изучению функций, довольно много внимания уделяется изучению ее аргумента—действительного числа. Аргументом вероятности является случайное событие. Прежде всего мы займемся уточнением интуитивного понятия случайного события. Далее аксиоматически будет введено понятие вероятности.

Аксиоматический подход построения теории вероятностей, предложенный А. Н. Колмогоровым в книге «Основные понятия теории вероятностей», сделал теорию вероятностей математической наукой. Ее аксиомы и теоремы в абстрактной форме отражают закономерности, присущие случайным событиям массового характера. В настоящее время аксиоматический подход является общепринятым.

## § 2. Пространство элементарных событий

Чтобы избежать неясностей при описании случайных явлений, результатов опытов или наблюдений, мы введем ряд точных понятий. Назовем произвольное множество $\Omega$ *пространством элементарных событий*. Элементы $\omega$ этого множества $\Omega$ будем называть *элементарными событиями*. Эти понятия являются первоначальными. В реальном опыте элементарным событиям соответствуют взаимно исключающие исходы. Ввиду большого разнообразия случайных явлений нельзя дать более конкретное определение множества элементарных событий. Для описания каждой реальной задачи множество $\Omega$ выбирается наиболее подходящим образом. Рассмотрим ряд примеров, поясняющих выбор множества $\Omega$.

1. Подбрасывание монеты один раз. Возможными исходами в этом опыте будут: выпадение монеты гербом вверх (или просто выпадение герба), выпадение решетки. Кроме того, монета возможно встанет на ребро, укатится куда-нибудь и т. д. Можно перечислить ряд исключающих друг друга событий, которые могут произойти с реальной монетой. При математическом описании этого опыта естественно отвлечься от ряда несущественных исходов и ограничиться только двумя: выпадение герба (можно обозначить это событие Г, $\omega_1$ или $\omega_Г$), выпадение решетки (Р, $\omega_2$

или $\omega_Р$). Таким образом, при описании этого опыта мы полагаем

$$\Omega=\{Г,\ Р\},\quad \Omega=\{\omega_1,\ \omega_2\}\quad \text{или}\quad \Omega=\{\omega_Г,\ \omega_Р\}.$$

2. Подбрасывание игральной кости один раз. В этом опыте естественно выбрать $\Omega=\{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5,\ \omega_6\}$, где $\omega_k$ обозначен исход опыта, заключающийся в выпадении $k$ очков. Имеем шесть исключающих друг друга исходов.

3. Одна монета подброшена $n$ раз. Каждому исходу опыта естественно поставить в соответствие последовательность длины $n$ по следующему правилу: если при $k$-м подбрасывании монеты выпал герб, то на $k$-м месте последовательности пишем Г, а при выпадении решетки — Р. Так, последовательность ГГГ ... ГГ обозначает исход опыта, заключающийся в том, что каждый раз выпадал герб. Элементарное событие

$$\underbrace{ГГ\ldots Г}_{m\text{ раз}}\ \underbrace{Р\ldots Р}_{n-m\text{ раз}}$$

соответствует исходу, в котором монета сначала подряд $m$ раз падала гербом вверх, а оставшиеся $n-m$ раз — решеткой вверх. Таким образом, множество $\Omega$ состоит из всевозможных последовательностей длины $n$ вида: ГРРГ ... РГ. При небольших значениях $n$ все элементарные события нетрудно выписать. Например, при $n=3$

$$\Omega=\{ГГГ,\ ГГР,\ ГРГ,\ ГРР,\ РГГ,\ РГР,\ РРГ,\ РРР\}.$$

Нетрудно проверить, что число элементарных событий при любом $n$ равно $2^n$. Действительно, все цепочки длины $n$ можно разбить на две группы цепочек вида

$$\{Г\ldots\},\quad \{Р\ldots\}.$$

Каждую из этих групп, фиксируя второй знак последовательности, можно снова разбить на 2 группы. Получим $2\cdot 2$ групп:

$$\{ГР\ldots\},\ \{ГГ\ldots\},\ \{РГ\ldots\},\ \{РР\ldots\}.$$

Фиксируя третий знак, получим $2\cdot2\cdot2$ групп. Продолжив этот процесс до фиксирования $n$ знаков, мы получим $2^n$ групп, каждая из которых состоит из одной последовательности.

4. Работа телефонной станции. Предположим, что мы наблюдаем работу телефонной станции в течение четверти часа и нас интересует число поступивших вызовов. Если телефонная станция обслуживает незначительное количество абонентов, то за время наблюдения может не поступить ни одного вызова, может поступить один вызов, два вызова и т. д. Достаточно ясно, что число вызовов будет всегда конечно. Однако разумно установить верхнюю границу числа вызовов довольно затруднительно. Проще не ограничивать возможное число вызовов и считать возможными исходами 0 и все натуральные числа: $\Omega = \{0, 1, 2, \ldots\}$. Это предположение проще, чем довольно искусственный подбор верхней границы числа вызовов. Предположение о возможности любого числа вызовов кажется абсурдным. Однако если окажется, что очень большое число вызовов происходит с очень малой вероятностью, то это будет совместимо с нашим практическим понятием невозможности. Подобные ситуации возникают довольно часто. Например, в страховом деле при расчетах не ограничивается максимальный возраст; при измерении довольно часто предполагается, что величина ошибки может принимать любые значения. Если в рассматриваемой задаче с работой телефонной станции нас интересует не только общее число поступивших вызовов, но и моменты их поступления, то пространство элементарных событий нужно выбрать более детализированным. Можно, например, исход нашего наблюдения описать ступенчатой функцией, постоянной между моментами поступления вызовов и имеющей скачки в моменты поступления вызовов. В этом случае $\Omega$ — множество ступенчатых функций.

5. Стрельба по плоской мишени. Введем в плоскости мишени прямоугольную систему координат $uOv$ и каждому исходу опыта, попадание в определенную точку плоскости, поставим в соответствие координаты этой точки. Тогда множеством $\Omega$ является

вся плоскость или множество всех упорядоченных пар действительных чисел. Этот факт мы запишем следующим образом:

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

где $u$, $v$ — действительные числа. В дальнейшем мы будем часто пользоваться такими обозначениями для различных множеств. За скобкой $\{$ будем писать обозначение элемента множества, а после двоеточия указывать, какие элементы включаются в данное множество. Например, множество точек замкнутого единичного круга с центром в начале координат будем записывать в виде

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

6. Броуновское движение. В микроскоп наблюдается движение малой частицы, подвергающейся большому числу ударов со стороны молекул. Наблюдение проводится в промежутке времени $[0, T]$. Исходом этого опыта будет определенная траектория движения частицы. Если интересоваться перемещением частицы вдоль заданного направления, то в каждый момент времени $t$, $t \in [0, T]$, положение ее проекции на заданное направление будет определяться координатой $x(t)$. В этом случае $\Omega = \{x(t)\}$ — множество непрерывных действительных функций, определенных на отрезке $[0, T]$.

Введем еще два пространства элементарных событий, которые будут в дальнейшем часто использоваться. Назовем *$n$-мерным дискретным пространством элементарных событий* множество $\Omega$, состоящее из векторов $(u_1, u_2, \ldots, u_n)$, координаты которых $u_k$ принимают конечное или счетное множество значений

$$u_k = u_k(1),\ u_k(2),\ \ldots$$

Пространство элементарных событий примера 3 является дискретным. Каждая координата вектора принимает значение Г или Р. В примере 4 — одномерное дискретное пространство; множество значений его координат является счетным.

Назовем *$n$-мерным непрерывным пространством элементарных событий* множество $\Omega$ векторов $(u_1, u_2, \ldots, u_n)$, где $u_k$ — любые действительные числа. В примере 5 пространство элементарных событий является двумерным и непрерывным.

## § 3. Случайные события

В примере 2 из § 2 мы перечислили все исключающие друг друга исходы, которые могут произойти при бросании игральной кости. Однако в этом же опыте можно говорить, например, о случайном событии, состоящем в том, что выпало четное число очков. Очевидно, что это событие происходит только в том случае, когда осуществилось одно из трех элементарных событий: $\omega_2$, $\omega_4$, $\omega_6$. Если в этом примере взять другое подмножество множества $\Omega$, например $A = \{\omega_1, \omega_3, \omega_5\}$, то можно сказать, что при наступлении любого элементарного события из $A$, и только в этом случае, выпадает нечетное число очков. Пусть теперь $\Omega$ — произвольное пространство элементарных событий. *Случайными событиями* или просто *событиями* будем называть подмножества $A$ множества $\Omega$. Случайные события обычно обозначают большими латинскими буквами $A$, $B$, $C$ и т. д. Событие в физическом опыте, которое описывается подмножеством $A$, происходит только в тех случаях, когда происходит какое-либо элементарное событие из $A$.

В примере 5 событием является, например, какая-либо область $A$ или $B$ в плоскости $uOv$; событие в реальном опыте, соответствующее $A$, происходит, если произошло попадание в точку $(u, v)$, принадлежащую множеству $A$. В примере 3 при $n = 3$ подмножество {ГГГ, ГГР, ГРГ, ГРР} множества $\Omega$ можно интерпретировать как выпадение герба при первом бросании монеты.

*Суммой* двух событий $A$ и $B$ назовем событие $A + B$ (или $A \cup B$), 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 $A$ или $B$. Пусть в примере 5 событиями $A$ и $B$ являются попадания соответственно в большой и малый круг (рис. 1). Тогда событием $A+B$ является заштрихованная область на рис. 1, *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 $B$. Для примера 5 событие $AB$ изображено заштрихованной фигурой на рис. 1, *б*.

*Разностью* $A \setminus B$ называется событие, состоящее из элементов множества $A$, не принадлежащих $B$ (см. рис. 1, *в*). Событие $A \setminus B$ состоит в том, что $A$ произошло, а $B$ не произошло.

Событие $\Omega$ назовем *достоверным*; пустое множество $\varnothing$ назовем невозможным событием. Событие $\overline{A} = \Omega \setminus A$ называется *противоположным* событию $A$ (см. рис. 1, *г*). Событие $\overline{A}$ означает, что $A$ не произошло.

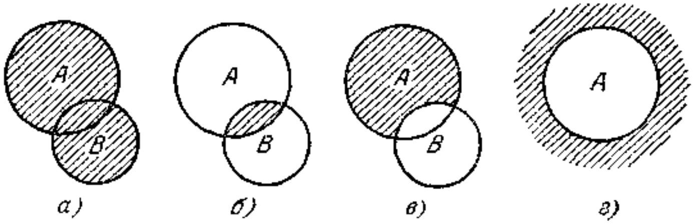


Рис. 1.

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 5, если $A \subset B$, то попадание в область $A$, содержащуюся в $B$, означает также попадание в $B$. 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

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

Часто оказываются полезными следующие равенства:

$$\overline{A+B} = \overline{A}\,\overline{B}, \tag{3.1}$$

$$(A+B)C = AC + BC. \tag{3.2}$$

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 $\omega$ 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.

## § 4. Аксиомы теории вероятностей

В конце § 1 было отмечено, что вероятность будет рассматриваться как функция от случайного события. В § 3 было дано точное определение случайного события. Функции действительной переменной определяются обычно не для всех действительных чисел, а только для их части, называемой областью определения. Вероятность тоже не всегда удается определить для любых подмножеств (случайных событий) множества $\Omega$. Приходится ограничиваться некоторым классом подмножеств, от которого естественно требовать замкнутость относительно операций, введенных в предыдущем параграфе.

Пусть $\Omega$ — произвольное пространство элементарных событий, а $\mathfrak{F}$ — некоторая система случайных событий.

Система событий $\mathfrak{F}$ называется *алгеброй событий*, если

1. $\Omega \in \mathfrak{F}$.
2. Из того, что $A \in \mathfrak{F}$ и $B \in \mathfrak{F}$, следует, что

$$AB \in \mathfrak{F}, \quad A+B \in \mathfrak{F}, \quad A \setminus B \in \mathfrak{F}.$$

Из 1 и 2 следует, что $\varnothing = \Omega \setminus \Omega \in \mathfrak{F}$. Наименьшей системой подмножеств, являющейся алгеброй, очевидно является система $\mathfrak{F} = \{\varnothing, \Omega\}$. Нетрудно проверить следующее утверждение. Если $\mathfrak{F}$ — система всех подмножеств множества $\Omega$, то $\mathfrak{F}$ — алгебра. Если $\Omega$ — конечное множество, то система всех подмножеств будет также конечным множеством. Для $\Omega$ из примера 2 § 2 можно выписать все события алгебры $\mathfrak{F}$, состоящей из всех подмножеств $\Omega$:

$$\varnothing; \ \{\omega_1\}, \{\omega_2\}, \ \ldots, \{\omega_6\};$$
$$\{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \ \ldots, \ \{\omega_5, \omega_6\}; \ \{\omega_1, \omega_2, \omega_3\}, \ \ldots;$$
$$\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} = \Omega.$$

В этом примере алгебра $\mathfrak{F}$ состоит из $2^6 = 64$ событий. Если множество $\Omega$ состоит из $N$ элементов, то число всех подмножеств равно $2^N$. Действительно, число последовательностей из 0 и 1 длины $N$ равно $2^N$, а между такими последовательностями и подмножествами $\Omega$ можно установить взаимно однозначное соответствие по следующему правилу: элемент с номером $i$ из множества $\Omega$ включается в подмножество, соответствующее данной последовательности, если на $i$-м месте последовательности стоит 1.

Приведем еще один пример алгебры событий. Пусть $\Omega = \{(u, v): 0 < u < 1, 0 < v < 1\}$ — единичный квадрат в плоскости. В курсе анализа доказано, что объединение, пересечение и разность квадрируемых фигур является квадрируемой фигурой. Следовательно, система $\mathfrak{F}$ квадрируемых подмножеств квадрата $\Omega$ образует алгебру событий.

Алгебра событий $\mathfrak{F}$ называется *σ-алгеброй* или *борелевской алгеброй*, если из того, что $A_n \in \mathfrak{F}$, $n=1, 2, \ldots$, следует

$$\bigcup_{n=1}^{\infty} A_n \in \mathfrak{F}, \qquad \bigcap_{n=1}^{\infty} A_n \in \mathfrak{F}.$$

Рассмотрим следующий пример. Пусть $\Omega=\{u: -\infty < u < \infty\}$ — числовая прямая. Определим систему множеств $\mathfrak{F}_0$, состоящую из конечных и бесконечных отрезков, интервалов и полуинтервалов:

$$[u_1, u_2]=\{u: u_1 \leqslant u \leqslant u_2\}, \qquad (u_1, u_2]=\{u: u_1 < u \leqslant u_2\},$$
$$[u_1, u_2)=\{u: u_1 \leqslant u < u_2\}, \qquad (u_1, u_2)=\{u: u_1 < u < u_2\},$$

где $u_1$, $u_2$ — действительные числа ($u_1 \leqslant u_2$) и, кроме того, в строгих неравенствах $u_1$, $u_2$ можно заменять на $-\infty$, $+\infty$. Эта система $\mathfrak{F}_0$ не является алгеброй; например, сумма множеств $(-\infty, -1)+(1, +\infty)$ не входит в $\mathfrak{F}_0$, хотя каждое слагаемое принадлежит $\mathfrak{F}_0$. Дополним $\mathfrak{F}_0$ всеми конечными суммами множеств из $\mathfrak{F}_0$. Новая более широкая система множеств $\mathfrak{F}$ является алгеброй.

Назовем σ-алгебру $\mathfrak{F}^*$ минимальной σ-алгеброй, порожденной $\mathfrak{F}$, если любая другая σ-алгебра, содержащая множества из $\mathfrak{F}$, содержит также все множества из $\mathfrak{F}^*$. Множества из минимальной σ-алгебры рассматриваемого примера называются *борелевскими*. В частности, борелевским множеством является каждая точка числовой прямой (вырожденный отрезок $\{u_1\} = [u_1, u_1]$). Так как множество рациональных чисел $R$ счетно, то их можно перенумеровать: $R=\{r_1, r_2, \ldots, r_m, \ldots\}$. Тогда $R=\bigcup\limits_{n=1}^{\infty}\{r_m\}$ и, следовательно, $R$ является борелевским множеством. Из $\Omega \in \mathfrak{F}^*$ и $R \in \mathfrak{F}^*$ следует, что множество иррациональных чисел $\Omega \setminus R \in \mathfrak{F}^*$, т. е. тоже является борелевским.

В дальнейшем мы обычно будем рассматривать те модели случайных явлений, в которых можно ограничиться алгеброй событий и не переходить к σ-алгебре.

Теперь мы можем ввести понятие вероятности события. Числовая функция $\mathsf{P}$, определенная на системе событий $\mathfrak{F}$, называется вероятностью, если

A1. *$\mathfrak{F}$ — является алгеброй событий.*

A2. $\mathsf{P}(A) \geqslant 0$ *для любого* $A \in \mathfrak{F}$.

A3. $\mathsf{P}(\Omega)=1$.

A4 (*аксиома конечной аддитивности*). *Если $A$ и $B$ несовместны, то*

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

Не вдаваясь в детальное обсуждение связей определенной таким образом вероятности с действительностью, отметим только, что аксиомы A2, A3, A4 полностью аналогичны свойствам частот (1)—(3) (см. введение).

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

A5. *Для любой убывающей последовательности*

$$A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots \tag{4.1}$$

*событий из $\mathfrak{F}$ такой, что*

$$\bigcap_{n=1}^{\infty} A_n = \varnothing, \tag{4.2}$$

*имеет место равенство*

$$\lim_{n \to \infty} \mathsf{P}(A_n) = 0. \tag{4.3}$$

Тройку $(\Omega, \mathfrak{F}, \mathsf{P})$, в которой $\mathfrak{F}$ является $\sigma$-алгеброй и $\mathsf{P}$ удовлетворяет A2—A5, называют *вероятностным пространством*.

Если вероятность $\mathsf{P}$ удовлетворяет A1—A5 и алгебра $\mathfrak{F}$ не является $\sigma$-алгеброй, то функцию $\mathsf{P}$ можно доопределить на множествах из минимальной $\sigma$-алгебры $\mathfrak{F}^*$, порожденной $\mathfrak{F}$. Приведем без доказательства следующую теорему.

Теорема о продолжении вероятности. *Пусть вероятность $\mathsf{P}$ удовлетворяет* A1—A5 *и* $\mathfrak{F}^*$ — минимальная $\sigma$-*алгебра, порожденная $\mathfrak{F}$. Тогда существует и притом единственная функция $\mathsf{P}^*(A)$, которая определена на $\mathfrak{F}^*$, удовлетворяет* A2—A5 *и совпадает с $\mathsf{P}(A)$, когда $A \in \mathfrak{F}$.*

Таким образом, можно сразу считать, что в A1 система $\mathfrak{F}$ является $\sigma$-алгеброй.

Простое вероятностное пространство соответствует примеру 2 из § 2. Множество всех возможных исходов, при бросании игральной кости состоит из шести элементов $\Omega=\{\omega_1, \omega_2, \ldots, \omega_6\}$. Пусть алгебра событий $\mathfrak{F}$

состоит из всех подмножеств $\Omega$. Произвольное случайное событие $A\in\mathfrak{F}$ можно записать в виде $A=\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$, где индексы $i_1, i_2, \ldots, i_k$ все различны и каждый из них может быть любой цифрой от 1 до 6, $k$ — число элементов в $A$. Событие $B$, состоящее в том, что выпало четное число очков, запишется в виде $B=\{\omega_2, \omega_4, \omega_6\}$. Здесь $k=3$, $i_1=2$, $i_2=4$, $i_3=6$. Для пустого множества $\varnothing$ положим $k=0$. Определим на множестве $\mathfrak{F}$ функцию $\mathsf{P}(A)$ формулой

$$\mathsf{P}(A)=\frac{k}{6}, \tag{4.4}$$

где $k$ — число элементов в $A$. Эта функция определена на алгебре $\mathfrak{F}$ (аксиома A1 выполнена); из определения следует, что $\mathsf{P}(A)\geqslant 0$ (A2), $\mathsf{P}(\Omega)=1$ (A3), так как $\Omega$ состоит из 6 элементов. Пусть теперь $A=\{\omega_{i_1}, \ldots, \omega_{i_k}\}$, $B=\{\omega_{l_1}, \ldots, \omega_{l_s}\}$ и $AB=\varnothing$ (например, $A=\{\omega_1\}$, $B=\{\omega_3, \omega_5\}$). Тогда $A+B=\{\omega_{l_1}, \omega_{l_2}, \ldots, \omega_{l_{s+k}}\}$ ($A+B=\{\omega_1, \omega_3, \omega_5\}$). По формуле (4.4) находим

$$\mathsf{P}(A)=\frac{k}{6} \qquad \left(\mathsf{P}(A)=\frac{1}{6}\right),$$
$$\mathsf{P}(B)=\frac{s}{6} \qquad \left(\mathsf{P}(B)=\frac{2}{6}\right),$$
$$\mathsf{P}(A+B)=\frac{k+s}{6} \qquad \left(\mathsf{P}(A+B)=\frac{3}{6}\right).$$

Отсюда следует, что выполнена аксиома A4. Аксиома A5 в случае конечных $\mathfrak{F}$ следует из предыдущих аксиом.

Действительно, в этом случае в (4.1), начиная с некоторого $n$, все события совпадут с $\varnothing$, а $\mathsf{P}(\varnothing)=0$ (согласно A4 имеем $\mathsf{P}(A)=\mathsf{P}(A+\varnothing)=\mathsf{P}(A)+\mathsf{P}(\varnothing)$, отсюда $\mathsf{P}(\varnothing)=0$).

Таким образом, введенная формулой (4.4) функция $\mathsf{P}(A)$ является вероятностью.

По формуле (4.4) вероятность выпадения $k$ очков (событие $\{\omega_k\}$) равна 1/6. Этот результат согласуется с интуитивными представлениями. В рассматриваемом примере в случае симметричной кости естественно считать все возможные исходы равновероятными.

Нетрудно задать на $\mathfrak{F}$ в рассматриваемом примере функции, отличные от функции (4.4). Пусть задано 6 произвольных чисел $p_i \geqslant 0$, $i=1, \ldots, 6$, удовлетворяющих условию $p_1+p_2+\ldots+p_6=1$. Определим на $\mathfrak{F}$ функцию $\mathsf{P}(A)$, поставив событию $A=\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$ в соответствие число

$$\mathsf{P}(A)=p_{i_1}+p_{i_2}+\ldots+p_{i_k}. \tag{4.5}$$

При различных наборах $p_1, p_2, \ldots, p_6$ будем получать разные функции $\mathsf{P}(A)$. Нетрудно проверить, что все они удовлетворяют аксиомам. В частности, при $p_1=p_2=\ldots=p_6=\frac{1}{6}$ формула (4.5) совпадает с (4.4). Приведенные примеры показывают, что вероятность не определяется однозначно системой аксиом.

В рассматриваемом опыте с костью любая его модель, определяемая формулой (4.5), является с точки зрения математики безупречной. Однако конкретный опыт может хорошо описывать только одна модель, или, точнее, несколько довольно близких. Разумеется, не может быть математического доказательства соответствия данной модели данному явлению. Выбор модели осуществляется на основе не вошедших в систему аксиом дополнительных соображений с привлечением проверки практикой и опытом. Так, выбор формулы (4.4) связан с соображениями симметричности кости. Если кость несимметрична, то подходящую модель нужно выбирать среди вероятностей (4.5), когда $p_i$ отличны от 1/6. Пусть, например, кубик игральной кости склеен из плотной бумаги и к грани противоположной грани с «6» прикреплена свинцовая пластинка. В этом случае будет выпадать всегда «6». В формуле (4.5) нужно положить $p_6=1$, $p_1=p_2=\ldots=p_5=0$.

Рассмотрим еще две математические модели опыта, заключающегося в подбрасывании двух монет одного достоинства.

*Модель 1.* Положим $\Omega=\{\omega_1, \omega_2, \omega_3\}$, где элементарное событие $\omega_1$ означает, что обе монеты выпали гербами вверх, $\omega_2$ — обе монеты выпали решетками вверх, $\omega_3$ — монеты выпали разными сторонами. Если считать элементарные события равновероятными, то вероятность события $A=\{\omega_1, \omega_2\}$, состоящего в том, что монеты выпали одинаковыми сторонами, равна $P_1(A)=2/3$.

*Модель* 2. Пусть $\Omega=\{ГГ, ГР, РГ, РР\}$, где знак Г на первом или втором месте в обозначении элементарного события означает, что на первой или на второй монете выпал герб; Р аналогично связано с выпадением решетки. Считая элементарные события равновероятными, для события $A=\{ГГ, РР\}$ получим $P_2(A)=2/4=1/2$.

Получили две разные вероятности одного и того же события. Экспериментальный выбор подходящей модели может быть проведен следующим образом. Подбросим $N$ раз ($N$ — достаточно большое) пару монет; пусть $N_A$ раз произошло событие $A$. Если частота $N_A/N\approx 2/3$, то считаем подходящей модель 1; если $N_A/N\approx 1/2$, то выбираем модель 2. Эта задача выбора из двух моделей (или двух гипотез) является одной из задач математической статистики. Более подробно она будет рассмотрена в гл. 10.

В задачах, подобных задаче о подбрасывании двух монет, практически всегда правильный выбор может быть сделан из соображений симметрии. Подбрасываются две физически различные монеты, поэтому естественно считать множество $\Omega$ состоящим из 4 элементарных событий, которым естественно приписать одинаковые вероятности. Модель 1 казалась бы более естественной, если бы можно было себе представить монеты физически неразличимыми. Опыт показывает, что в действительности монеты ведут себя как различимые. Однако, этот достаточно очевидный факт для монет, оказывается неверным для некоторых типов частиц. Бозе и Эйнштейн показали *), что некоторые типы частиц ведут себя как неразличимые.

Рассмотренный пример показывает, что нельзя слишком полагаться на интуицию; подходящей моделью может оказаться совершенно неожиданная.

---

*) Подробнее см. Феллер [17], гл. II, § 5, стр. 32.

## § 5. Следствия из аксиом

Укажем сначала несколько простых свойств вероятности, которые непосредственно вытекают из аксиом А2—А4. Из равенства $A+\overline{A}=\Omega$ и аксиом А3 и А4 следует, что $\mathsf{P}(A)+\mathsf{P}(\overline{A})=1$ и

$$\mathsf{P}(\overline{A})=1-\mathsf{P}(A). \qquad (5.1)$$

Отсюда и из аксиомы А3, полагая $A=\Omega$, получим

$$\mathsf{P}(\varnothing)=0. \qquad (5.2)$$

Если $A_1, A_2, \ldots, A_n$ *попарно несовместны* (т. е. $A_iA_j=\varnothing$ при любых $i\neq j$, $i, j=1, \ldots, n$), то

$$\mathsf{P}(A_1+A_2+\ldots+A_n)=\mathsf{P}(A_1)+\mathsf{P}(A_2)+\ldots+\mathsf{P}(A_n). \qquad (5.3)$$

Формула (5.3) следует по индукции из аксиомы А4. Для любых событий $A$ и $B$ имеет место формула

$$\mathsf{P}(A+B)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(AB). \qquad (5.4)$$

Действительно, представим события $A+B$ и $B$ в виде $A+B=A+B\overline{A}$ и $B=B\overline{A}+BA$. В правых частях этих равенств события, являющиеся слагаемыми, несовместны, и, следовательно, к правым частям применима аксиома А4. Таким образом, находим

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

Из двух последних равенств сразу следует (5.4). Из (5.4) и неравенства $\mathsf{P}(AB)\geqslant 0$ для любых $A$ и $B$ получаем

$$\mathsf{P}(A+B)\leqslant\mathsf{P}(A)+\mathsf{P}(B). \qquad (5.5)$$

Отсюда по индукции для *любых* $A_1, A_2, \ldots, A_n$ находим

$$\mathsf{P}(A_1+A_2+\ldots+A_n)\leqslant\mathsf{P}(A_1)+\mathsf{P}(A_2)+\ldots+\mathsf{P}(A_n).$$

Если событие $A$ *влечет* за собой $B$ $(A\subset B)$, то

$$\mathsf{P}(A)\leqslant\mathsf{P}(B). \qquad (5.6)$$

Действительно, в этом случае $B=A+\bar{A}B$ и, следовательно,

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

Так как $\varnothing\subset A\subset\Omega$ и $\mathsf{P}(\varnothing)=0$, $\mathsf{P}(\Omega)=1$, то отсюда и из (5.6) для любого события $A$ получим

$$0\leqslant\mathsf{P}(A)\leqslant 1. \tag{5.7}$$

Рассмотренные свойства были получены как следствия А2—А4. Можно показать, что А4 и А5 эквивалентны следующему свойству *σ-аддитивности* (или *счетной аддитивности*) вероятности: *если в последовательности*

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

*события попарно несовместны и* $A=\bigcup\limits_{n=1}^{\infty}A_n\in\mathfrak{F}$, *то*

$$\mathsf{P}(A)=\sum_{n=1}^{\infty}\mathsf{P}(A_n). \tag{5.8}$$

Отметим без доказательства еще два свойства, связанных с аксиомой непрерывности А5.

*Если* $A_1\subset A_2\subset\ldots\subset A_n\subset A_{n+1}\subset\ldots$ *и* $A=\bigcup\limits_{n=1}^{\infty}A_n$ *или* $A_1\supset A_2\supset\ldots\supset A_n\supset A_{n+1}\supset\ldots$ *и* $A=\bigcap\limits_{n=1}^{\infty}A_n$, *то*

$$\mathsf{P}(A)=\lim_{n\to\infty}\mathsf{P}(A_n). \tag{5.9}$$

В явном виде вероятность довольно просто задается в случае, когда $\Omega$ является счетным множеством; обычно нетрудно рассмотреть случай несчетного $\Omega$ (например, прямая или плоскость), но для системы $\mathfrak{F}$, являющейся алгеброй. Более сложные случаи рассматриваются с помощью теоремы о продолжении вероятности, приведенной в § 4. Если вероятность задана на алгебре $\mathfrak{F}$, то, зная, что ее единственным образом можно продолжить на минимальную σ-алгебру $\mathfrak{F}^*$, порожденную $\mathfrak{F}$, мы можем воспользоваться формулами (5.8) и (5.9) для вычисления вероятностей событий из $\mathfrak{F}^*$.

## § 6. Примеры вероятностных пространств

**6.1. Классическая схема.** Пусть $\Omega=\{\omega_1, \omega_2, \ldots, \omega_s\}$, алгебра событий $\mathfrak{F}$ состоит из всех подмножеств $A=\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$, $1\leqslant i_1<i_2<\ldots<i_k\leqslant s$, $k=0, 1, 2, \ldots, s$, множества $\Omega$ и

$$\mathsf{P}(A)=\frac{k}{s}. \tag{6.1}$$

Определенная этим равенством функция $\mathsf{P}(A)$ удовлетворяет всем аксиомам А1—А5. Это классическое определение вероятности. Событие $A$ происходит, если произошло одно из $k$ элементарных событий, которые иногда называют благоприятствующими $A$. Распространена следующая формулировка классического определения вероятности: вероятностью события $A$ называется отношение числа исходов, благоприятствующих $A$, к общему числу исходов.

Из формулы (6.1) следует, что события, состоящие из одного элементарного события, равновероятны:

$$\mathsf{P}(\{\omega_1\})=\ldots=\mathsf{P}(\{\omega_s\})=\frac{1}{s}. \tag{6.2}$$

Если заданы вероятности (6.2), то из формулы (5.3) следует, что вероятность любого события $A=\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}=\{\omega_{i_1}\}+\{\omega_{i_2}\}+\ldots+\{\omega_{i_k}\}$ определяется формулой (6.1). Таким образом, классическая схема может служить моделью тех случайных явлений, для которых представляется естественным предположение (6.2). 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. В задачах по теории вероятностей, предназначенных для упражнений, довольно часто приводится только описание опыта или явления и не дается математическая формулировка. Предполагается, что решение должно состоять из двух частей:

1) выбор подходящей модели для описания данного в условии задачи опыта и математическая формулировка задачи;

2) решение математической задачи.

Пример 1. Из урны, содержащей $M$ белых и $N-M$ черных шаров, наудачу извлекается сразу $n$ шаров. Какова вероятность того, что среди выбранных $n$ шаров окажется ровно $m$ белых?

Решение. 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 естественно предположить, что все элементарные исходы опыта равновероятны и применима классическая схема. 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 (6.1) нужно положить $s=C_N^n$. 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, число элементарных событий в $A_m$ равно $k=C_M^m C_{N-M}^{n-m}$ и по формуле (6.1)

$$\mathsf{P}(A_m)=P_n(m, N, M)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad m=0, 1, 2, \ldots \tag{6.3}$$

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 $\frac{M}{N} \approx \frac{m}{n}$. Формула (6.3) используется при оценке отклонения $\frac{m}{n}$ от $\frac{M}{N}$. В рассмотренной ситуации неизвестным параметром является число $M$.

Рассмотрим пример, когда 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 делать заключения о величине $N$.

Оценим по формуле (6.3) вероятность получения какого-либо выигрыша в «Спортлото» по одной карточке. Участник лотереи из 49 номеров отмечает 6 (49 шаров, среди которых 6 белых; $N = 49$, $M = 6$). После того как участник сдал карточку, проводится выборка $n = 6$ номеров. Если число $m$ номеров, отмеченных участником и попавших в выборку, оказалось больше 2, то участник получает какой-либо выигрыш. Если событие $A$ — получение выигрыша, то $\bar{A} = A_0 + A_1 + A_2$. По формуле (6.3) можно составить следующую таблицу:

| $m$ | 0 | 1 | 2 |
|---|---|---|---|
| $\mathsf{P}(A_m)$ | 0,435965 | 0,413019 | 0,132378 |
| $m$ | 3 | 4 | 5 |
| $\mathsf{P}(A_m)$ | 0,017650 | 0,000969 | 0,000018 |

$$\mathsf{P}(A_6) = 0{,}00000007151$$

Отсюда

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

Искомую вероятность находим по формуле (5.1)

$$\mathsf{P}(A) = 1 - \mathsf{P}(\bar{A}) = 0{,}018\,638.$$

Пример 2. Ребенок, играя десятью кубиками, на которых написаны буквы М, М, Т, Т, А, А, А, К, И, Е, сложил слово «МАТЕМАТИКА». Можно ли считать, что ребенок грамотный?

Решение. Сначала дадим математическую формулировку задачи. Если предположить, что ребенок неграмотный и родители не научили его складывать единственное слово, то естественно предположить, что расположение кубиков «математика» не более привлекательно по сравнению с остальными. В таком случае можно ожидать, что классическая схема окажется достаточно подходящей. Оценим вероятность события $A$, состоящего в расположении кубиков «математика». Пространством элементарных событий являются всевозможные перестановки 10 кубиков. Таких перестановок $10!$. При этом кубики с одинаковыми буквами мы считали различными. Подсчитаем, сколько элементарных событий входит в $A$. Кубики с буквами, отличными от М, Т, А, должны стоять на определенных местах, 3 кубика с буквой Т можно расположить на трех местах $3! = 6$ способами; кубики с буквой М располагаются двумя способами и с буквой Т тоже двумя. Сочетая каждое расположение с каждым, получим, что $A$ состоит из $2 \cdot 2 \cdot 6 = 24$ элементарных событий. По формуле (6.1)

$$\mathsf{P}(A) = \frac{24}{10!} = \frac{1}{15\,120};$$

Эта вероятность очень мала, и событие $A$ можно считать практически невозможным. Если же оно осуществилось, то следует считать нашу гипотезу о неграмотности ребенка неверной. Однако при большом числе испытаний по данной «программе» даже неграмотный может сложить слово «математика»; в среднем из

15 120 неграмотных 1 будет ошибочно считаться грамотным.

Пример 3. У человека в кармане $n$ ключей, из которых только один подходит к его двери. Ключи последовательно извлекаются (без возвращения) до тех пор, пока не появится нужный ключ. Найти вероятность того, что нужный ключ появится при $k$-м извлечении.

Решение. Будем продолжать извлекать ключи, после появления нужного ключа, до конца. За множество элементарных событий примем всевозможные последовательности из $n$ ключей. Таких последовательностей $n!$. Последовательностей, у которых нужный ключ находится на определенном месте, очевидно, $(n-1)!$, так как одно место занято нужным ключом, а остальные $n-1$ ключей можно на $n-1$ месте расставить $(n-1)!$ способами. Таким образом, искомая вероятность равна $(n-1)!/n! = 1/n$.

**6.2. Дискретное вероятностное пространство.** Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$ — счетное множество, а $\mathfrak{F}$ — набор всех его подмножеств. Очевидно, что $\mathfrak{F}$ является $\sigma$-алгеброй. Пусть $p_n$, $n = 1, 2, \ldots$, — последовательность неотрицательных чисел, удовлетворяющая условию

$$\sum_{n=1}^{\infty} p_n = 1. \tag{6.4}$$

Для любого $A \in \mathfrak{F}$ положим

$$\mathsf{P}(A) = \sum_{n \in \{n:\ \omega_n \in A\}} p_n. \tag{6.5}$$

Если (6.5) — числовой ряд, то сходимость его следует из (6.4).

Так же, как в § 4 в примере с костью, можно показать, что функция (6.5) удовлетворяет аксиомам А2 — А4.

Проверим А5.

Событие $A_n$ в последовательности (4.1) запишем в виде

$$A_n = \{\omega_{i_1(n)}, \omega_{i_2(n)}, \ldots\},$$

где $l_1(n) < l_2(n) < l_3(n) < \ldots$, причем при любом $n > 1$ последовательность индексов элементарных событий из $A_n$ является подпоследовательностью соответствующей последовательности из $A_{n-1}$. Из условия (4.2) следует, что $l_1(n) \longrightarrow \infty$ при $n \longrightarrow \infty$. Если бы это было не так, то произведение в (4.2) не было бы пусто. По формуле (6.5) находим

$$\mathsf{P}(A_n) = \sum_{k=1}^{\infty} p_{l_k(n)} \leqslant \sum_{s=l_1(n)}^{\infty} p_s.$$

Отсюда, при $n \longrightarrow \infty$ следует (4.3), так как правая часть последнего неравенства является остатком сходящегося ряда.

Нетрудно проверить, что все распределения вероятностей на множестве $\mathfrak{F}$ можно получить в виде (6.5).

Заметим, что при $p_n = 0$, $n > N$, можно ограничиться конечным пространством элементарных событий $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$. Тогда формула (6.5) задает распределение вероятностей на конечном пространстве элементарных событий. В случае $p_1 = p_2 = \ldots = p_N = \frac{1}{N}$, $p_n = 0$ при $n > N$ получаем классическое определение вероятности.

Отметим еще, что $n$-мерное дискретное пространство элементарных событий, введенное в § 2, является счетным и отличается от рассматриваемого случая только обозначениями элементарных событий. В дальнейшем будем называть *$n$-мерным дискретным вероятностным пространством* тройку $(\Omega, \mathfrak{F}, \mathsf{P})$, в которой $\Omega = \{u_1, u_2, \ldots, u_n\}$ — $n$-мерное дискретное пространство элементарных событий, $\mathfrak{F}$ — система всех подмножеств множества $\Omega$ и для любого $A \in \Omega$

$$\mathsf{P}(A) = \sum_{(u_1(l_1), \ldots, u_n(l_n)) \in A} p_{u_1(l_1)\, u_2(l_2) \ldots u_n(l_n)}, \qquad (6.6)$$

где $p_{u_1(l_1) \ldots u_n(l_n)} \geqslant 0$ и $\sum_{l_1, \ldots, l_n = 1}^{\infty} p_{u_1(l_1) \ldots u_n(l_n)} = 1$.

Заметим, что замена вектора $(u_1(l_1), \ldots, u_n(l_n))$ последовательностью $(u_1(l_1), \ldots, u_n(l_n), \ldots)$ приведет к тому, что множество $\Omega$ будет уже несчетным.

**6.3. Геометрические вероятности.** Классическое определение вероятности нельзя применить к опыту с бесконечным числом «равновероятных» исходов.

К описанию такой ситуации было приспособлено геометрическое определение вероятности. Пусть $\Omega$ является квадрируемой областью плоскости. Рассмотрим систему $\mathfrak{F}$ квадрируемых подмножеств $\Omega$. Для любого $A \in \mathfrak{F}$ положим

$$\mathsf{P}(A) = \frac{\text{пл. } A}{\text{пл. } \Omega}. \tag{6.7}$$

По теореме о продолжении вероятности функцию (6.7) можно определить на множествах, входящих в $\sigma$-алгебру $\mathfrak{F}^*$, порожденную алгеброй $\mathfrak{F}$. Во многих задачах можно ограничиться событиями из алгебры $\mathfrak{F}$. Мы не будем вычислять вероятности новых (по сравнению с $\mathfrak{F}$) событий из $\mathfrak{F}^*$. Приведем только в качестве примера событие $R$, состоящее из точек с рациональными координатами. Событие $R \in \mathfrak{F}^*$. В § 4 в одномерном случае этот факт был установлен. Рассуждение для двумерного случая проводится аналогично. Для вероятностей событий из $\mathfrak{F}^*$ сохраним старое обозначение $\mathsf{P}$ (не будем вводить $\mathsf{P}^*$, как в теореме о продолжении). Покажем сначала, что вероятность события $\{(x_0, y_0)\}$, состоящего из одной точки $(x_0, y_0)$, равна 0. Действительно, так как

$$\varnothing \subset \{(x_0, y_0)\} \subset \left\{(u, v)\colon (u-x_0)^2+(v-y_0)^2 \leqslant \frac{1}{n}\right\},$$

то, используя (5.6), получим

$$0 \leqslant \mathsf{P}(\{(x_0, y_0)\}) \leqslant \frac{\pi}{n^2 \text{ пл. } G}.$$

Отсюда при $n \to \infty$ находим, что $\mathsf{P}(\{x_0, y_0\}) = 0$. Множество $R$ счетно. Перенумеровав точки $(x_n, y_n)$, получим $R = \bigcup\limits_{n=1}^{\infty} \{(x_n, y_n)\}$. Отсюда и из формулы (5.8) следует, что $\mathsf{P}(R) = 0$.

В схеме геометрических вероятностей выбор модели, подходящей для описания реального явления, более затруднителен по сравнению с классической схемой. Если выбрать разные модели реального опыта, то для одного и того же события в разных моделях можно получить разные вероятности. На выборе разных математических моделей основан известный парадокс Бертрана (более подробно см. Б. В. Гнеденко [4], гл. I, § 6, стр. 36—37).

Решим следующую задачу.

Задача Бюффона. Плоскость расчерчена параллельными прямыми, расстояние между которыми равно $2a$. На плоскость наудачу брошена игла длины $2l$

($l < a$). Найти вероятность того, что игла пересечет какую-нибудь прямую.

Решение. Дадим сначала определение подходящего пространства элементарных событий. Пусть $u$ — расстояние от центра иглы до ближайшей прямой, а $\varphi$ — угол, составленный иглой с этой прямой. Пара чисел $(\varphi, u)$ задает положение иглы с точностью до выбора конкретной прямой. Поскольку нас интересует только взаимное расположение иглы с ближайшей прямой, то можно в качестве $\Omega$ выбрать прямоугольник

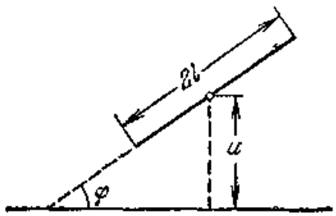


Рис. 2.

$$\Omega = \{(\varphi, u):\ 0 \leqslant \varphi \leqslant \pi,\ 0 \leqslant u \leqslant a\}.$$

Пересечение иглы с прямой происходит в том и только

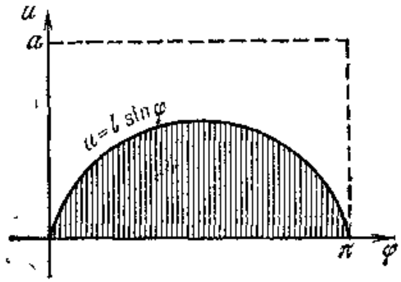


Рис. 3.

в том случае, когда (рис. 2)

$$u \leqslant l \sin \varphi.$$

Таким образом, интересующее нас событие

$$A = \{(\varphi, u):\ u \leqslant l \sin \varphi\}$$

является множеством, заштрихованным на рис. 3. Так как

$$\text{пл.}\, A = \int_0^\pi l \sin \varphi \, d\varphi = 2l, \quad \text{пл.}\, \Omega = a\pi,$$

то по формуле (6.7) находим

$$\mathsf{P}(A) = \frac{2l}{a\pi}. \tag{6.8}$$

О соответствии математической модели опыту можно судить по результатам экспериментов. Пусть игла была брошена $n$ раз и $m$ раз произошло пересечение. При больших $n$ частота $\frac{m}{n} \approx \mathsf{P}(A) = \frac{2l}{a\pi}$. Отсюда можно получить экспериментальную оценку числа $\pi \approx 2\frac{l}{a} \cdot \frac{n}{m}$. Приведем результаты некоторых экспериментов с бросанием иглы, заимствованные из книги М. Кендалла, П. Морана «Геометрические вероятности», «Наука», 1972 г. (стр. 83):

| Экспериментатор | $l/a$ | $n$ | $m$ | Оценка $\pi$ |
|---|---|---|---|---|
| Вольф, 1850 г. | 0,8 | 5000 | 2532 | 3,1596 |
| Рейна, 1925 г. | 0,5419 | 2520 | 859 | 3,1795 |

Схема геометрических вероятностей успешно применяется в астрономии, атомной физике, биологии, кристаллографии.

**6.4. Абсолютно непрерывные вероятностные пространства.** Пусть $\Omega = \{(u_1, u_2, \ldots, u_n)\}$ — $n$-мерное действительное евклидово пространство, $\pi(u_1, u_2, \ldots, u_n)$ — неотрицательная функция, интегрируемая по Риману по любой квадрируемой области из $\Omega$. Будем предполагать, что существует несобственный интеграл по $\Omega$ от функции $\pi(u_1, u_2, \ldots, u_n)$ и

$$\int \ldots \int_{\Omega} \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n = 1. \tag{6.9}$$

Обозначим $\mathfrak{F}$ алгебру, порожденную квадрируемыми областями в $\Omega$. Для любого $A \in \mathfrak{F}$ положим

$$\mathsf{P}(A) = \int \ldots \int_{A} \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n. \tag{6.10}$$

Можно рассмотреть более общий случай, когда $\pi(u_1, \ldots, u_n)$ не ограничена в конечном числе точек $\Omega$. Тогда интегралы в (6.10) по множествам $A$, содержащим такие точки, нужно понимать как несобственные. Нетрудно проверить, что функция $\mathsf{P}(A)$, определенная соотношением (6.9), удовлетворяет аксиомам А2—А5. Из теоремы о продолжении вероятности следует, что формула (6.10) позволяет определить вероятность на минимальной $\sigma$-алгебре $\mathfrak{F}^*$, порожденной алгеброй $\mathfrak{F}$. Таким образом, мы определили вероятностное пространство. Построенное вероятностное пространство будем называть $n$-мерным *абсолютно непрерывным вероятностным пространством*.

Отметим, что рассмотренная выше схема геометрических вероятностей является двумерным абсолютно непрерывным вероятностным пространством с $\pi(u_1, u_2) = \frac{1}{\text{пл.} G}$, если $(u_1, u_2) \in G$, $G$ — квадрируемая фигура, и $\pi(u_1, u_2) = 0$, если $(u_1, u_2) \notin G$.

### Задачи к главе I

**1.** Проверить следующие соотношения между случайными событиями:

1) $\overline{A+B} = \bar{A}\bar{B}$;
2) $(A+B) \setminus B = A \setminus AB = A\bar{B}$;
3) $AA = A + A = A$;
4) $(A+B)C = AC + BC$;
5) $\bar{A} + \bar{B} = \overline{AB}$.

**2.** Упростить следующие выражения:

1) $(A+B)(A+\bar{B})$;
2) $(A+B)(\bar{A}+B)(A+\bar{B})$;
3) $(A+B)(B+C)$.

**3.** Установить, какие из следующих соотношений правильны:

1) $(A+B) \setminus C = A + (B \setminus C)$;
2) $ABC = AB(C+B)$;
3) $A\bar{B}C \subseteq A + B$;
4) $(\overline{A+B})C = \bar{A}\bar{B}\bar{C}$.

**4.** 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 $B$, но $C$ не произошло;
3) все три события произошли;
4) произошло, по крайней мере, одно из событий;
5) произошло одно и только одно событие;
6) ни одно событие не произошло;

7) произошло не больше двух событий.

√ 5. *Бросаются две игральные кости. Пусть событие $A$ состоит* в том, что сумма очков четная, а $B$ заключается в том, что выпала хотя бы одна единица. Описать пространство элементарных событий, события $AB$, $A+B$, $A\overline{B}$. Найти их вероятности, если *все элементарные события равновероятны.*

6. Описать пространство элементарных событий, соответствующих трем испытаниям, в каждом из которых может появиться У (успех) или Н (неуспех). Выразить через элементарные события:

1) событие $A$ — в первом испытании произошел успех;

2) событие $B$ — произошло ровно два успеха;

3) *событие $C$ — произошло не больше двух успехов.*

7. Монета бросается до тех пор, пока два раза подряд она не выпадет одной и той же стороной. Каждому возможному исходу, *продолжавшемуся* $n$ бросаний, приписана вероятность $2^{-n}$. Описать пространство элементарных событий. Вычислить вероятности событий:

1) *опыт окончился до шестого бросания;*

2) потребуется четное число бросаний;

3) после конечного числа бросаний опыт окончится

8. Числа 1, 2, ..., $n$ расставлены случайным образом. Найти вероятность того, что 1 и 2 расположены рядом и притом в порядке возрастания.

9. *Найти вероятность того, что среди трех наугад выбранных цифр встретится 0, 1, 2 повторений.*

√ 10. Показать, что более вероятно при одновременном бросании четырех костей получить хотя бы одну единицу, чем при 24 бросаниях двух костей получить *хотя бы один раз две единицы*

11. Случайно размещается $n$ шаров по $N$ ящикам При $n=N$ найти вероятность того, что ровно один ящик останется пустым.

12. *Найти вероятность того, что дни рождения 12 человек* придутся на разные месяцы года.

13. В чулане $n$ пар ботинок. Из них случайно выбирается $2r$ ботинок ($2r<n$). *Найти вероятность того, что 1) среди выбранных ботинок нет парных;* 2) имеется ровно одна пара

14. Ящик содержит 90 годных и 10 бракованных деталей. Найти вероятность *того, что среди 10 вынутых из ящика деталей* нет бракованных.

15. Найти вероятность того, что на две карточки спортлото с отмеченными номерами (4, 12, 38, 20, 41, 46) и (4, 12, 38, 20, 41, 49) будет получено ровно два минимальных выигрыша (угадано ровно по три числа).

16. Из последовательности чисел 1, 2, ..., $N$ отобраны наудачу *$n$ чисел и расположены в порядке возрастания* $x_1 < x_2 < \ldots < x_n$. Найти вероятность того, что $x_m \leqslant M < x_{m+1}$ и вычислить ее предел при $N,\ M \longrightarrow \infty$, $M/N=\alpha$.

17. На бесконечную шахматную доску со стороной квадрата $a$ бросается наудачу *монета радиуса $r$, $2r<a$. Найти вероятность* того, что 1) монета целиком попадет внутрь одного квадрата; 2) пересечет не более одной стороны квадрата.

18. Двое договорились встретиться в отрезке времени $[0, T]$. Первый пришедший ждет второго время $t$, $t < T$. Найти вероятность того, что встреча произойдет. За множество $\Omega$ принять точки квадрата $\{(t_1, t_2): 0 \leqslant t_1 \leqslant T, 0 \leqslant t_2 \leqslant T\}$, где $t_1$ и $t_2$ — моменты прихода встречающихся.

19. На отрезок $[0, 1]$ наудачу брошена точка. Пусть $\xi$ – ее координата. Найти функции $F(x) = \mathsf{P}(\xi < x)$, $F'(x)$. Построить их графики.

20. В квадрат $\{(x_1, x_2): 0 \leqslant x_i \leqslant 1, i = 1, 2\}$ наудачу брошена точка. Пусть $(\xi_1, \xi_2)$ — ее координаты. Найти функции $F(x) = \mathsf{P}(\xi_1 + \xi_2 < x)$, $F'(x)$.

21. В куб $\{(x_1, x_2, x_3): 0 \leqslant x_i \leqslant 1, i = 1, 2, 3\}$ наудачу брошена точка. Пусть $(\xi_1, \xi_2, \xi_3)$ — ее координаты. Найти функции $F(x) = \mathsf{P}(\xi_1 + \xi_2 + \xi_3 < x)$, $F'(x)$. Построить графики. Сравнить задачи 19—21.

22. Величины $\xi_1$, $\xi_2$ определены в задаче 20. Найти вероятность того, что корни уравнения $x^2 - \xi_1 x + \xi_2 = 0$ действительны.

23. На отрезок $[0, 1]$ наудачу брошено 2 точки, разбившие его на 3 отрезка. Какова вероятность того, что из них можно построить треугольник? За множество $\Omega$ принять пары чисел, являющиеся координатами брошенных точек.

24. Обозначим $\eta_1 = \xi_1 \xi_2$, $\eta_2 = \min(\xi_1, \xi_2)$, $\eta_3 = \max(\xi_1, \xi_2)$, где $\xi_1$, $\xi_2$ определены в задаче 20. Найти

$$\mathsf{P}(\eta_1 < x), \quad \mathsf{P}(\eta_2 < x), \quad \mathsf{P}(\eta_3 < x).$$

ГЛАВА 2

# УСЛОВНЫЕ ВЕРОЯТНОСТИ. НЕЗАВИСИМОСТЬ СОБЫТИЙ

## § 1. Условные вероятности

Прежде чем переходить к формальному определению условной вероятности, рассмотрим ряд примеров.

Пример 1. Допустим, что студент выучил из 28 билетов 4 четных и 12 нечетных. По классической схеме вероятность того, что студент получит выученный им билет (событие $A$), равна $16/28 = 4/7$. Пусть теперь известно, что к моменту прихода студента осталось 14 билетов и все они нечетные (событие $B$). Какова вероятность события $A$ при этой дополнительной информации ($B$ произошло)? Естественно принять ее равной $12/14 = 6/7$.

Пример 2. Имеется $N$ карточек трех типов: черные (обе стороны черные), белые (обе стороны белые), разноцветные (одна сторона белая, а другая черная). Пусть $N_A$ карточек имеют белый цвет (белые или разноцветные), $N_B$ имеют черный цвет (черные или разноцветные), $N_{AB}$ имеют и белый, и черный цвет. Вынимается одна карточка. Какова вероятность появления белого цвета (событие $A$) в предположении, что карточки извлекаются с равными вероятностями? Наудачу выбранная карточка положена на стол, и ее верхняя сторона оказалась черной (событие $B$ — появился черный цвет). Какова в этом случае ($B$ произошло) вероятность того, что другая сторона белая (событие $A$)? Без дополнительной информации по классической схеме $\mathsf{P}(A) = N_A/N$. Если осуществилось событие $B$, то осуществился один из $N_B$ исходов. Среди этих исходов событие $A$ появляется $N_{AB}$ раз. Естественно в этом примере условной вероятностью $\mathsf{P}(A \mid B)$ события $A$ при условии, что $B$ произошло, назвать отношение

$$\mathsf{P}(A \mid B) = \frac{N_{AB}}{N_B}.$$

Правую часть этого равенства можно представить в виде отношения $\mathsf{P}(AB)/\mathsf{P}(B)$, так как $\mathsf{P}(AB)=N_{AB}/N$, $\mathsf{P}(B)=N_B/N$. Таким образом, в рассматриваемом примере

$$\mathsf{P}(A\mid B)=\frac{\mathsf{P}(AB)}{\mathsf{P}(B)}. \tag{1.1}$$

Это равенство позволяет дать общее определение условной вероятности. Пусть $(\Omega, \mathfrak{F}, \mathsf{P})$ — произвольное вероятностное пространство. Если $A$, $B\in\mathfrak{F}$ и $\mathsf{P}(B)>0$, то *условная вероятность* события $A$ при условии, что произошло событие $B$, определяется формулой (1.1), в правой части которой символ $\mathsf{P}$ понимается как вероятность в рассматриваемом вероятностном пространстве.

Пусть теперь некоторое событие $B$ с $\mathsf{P}(B)>0$ фиксировано. Нетрудно проверить, что функция

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

определенная для всех $A\in\mathfrak{F}$, удовлетворяет аксиомам А2—А5, в частности

$$\mathsf{P}(A\mid B)\geqslant 0, \quad \mathsf{P}(\Omega\mid B)=1,$$
$$\mathsf{P}(A_1+A_2\mid B)=\mathsf{P}(A_1\mid B)+\mathsf{P}(A_2\mid B),$$

если $A_1A_2=\varnothing$. Таким образом, для $\mathsf{P}_B(A)$ справедливы все следствия из аксиом, доказанные в § 5 гл. 1. Кроме того,

$$\mathsf{P}(B\mid B)=1, \quad \mathsf{P}_B(A\mid C)=\mathsf{P}(A\mid BC).$$

## § 2. Вероятность произведения событий

Поменяем в (1.1) местами $A$ и $B$. Полученное равенство можно записать в виде «теоремы умножения»

$$\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B\mid A). \tag{2.1}$$

По индукции из (2.1) легко получить более общую формулу

$$\mathsf{P}(A_1A_2\ldots A_n)=$$
$$=\mathsf{P}(A_1)\,\mathsf{P}(A_2\mid A_1)\ldots\mathsf{P}(A_n\mid A_1\ldots A_{n-1}). \tag{2.2}$$

Действительно, при $n=2$ (2.2) совпадает с (2.1). Пусть (2.2) доказано для $n-1$ сомножителей. Тогда для $n$ сомножителей (2.2) следует из равенства

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

которое получается из (2.1) при $A=A_1\ldots A_{n-1}$, $B=A_n$. Если заданы $\Omega$, $\mathfrak{F}$ и $\mathsf{P}$, то использование формул (2.1) и (2.2) для вычисления вероятностей произведений событий, например, для вычисления $\mathsf{P}(AB)$, бессмысленно, так как сама условная вероятность $\mathsf{P}(B\mid A)$ определяется как отношение $\mathsf{P}(AB)$ к $\mathsf{P}(A)$. Однако при решении реальных задач не заданы ни $\Omega$, ни $\mathfrak{F}$, ни $\mathsf{P}$. Классическая схема и схема геометрических вероятностей, в которых вероятность определяется в явном виде, описывают далеко не все встречающиеся в приложениях задачи. Довольно часто оказывается возможным задать ряд чисел, которые должны стать значениями условных вероятностей в математической модели. Таким образом, требуется построить вероятностное пространство, в котором вероятности некоторых событий имеют наперед заданные значения. В ряде случаев подобная задача может быть решена при помощи формул (2.1) и (2.2). Пусть, например, дано, что

$$\mathsf{P}(A)=a_1,\qquad \mathsf{P}(B\mid A)=a_{11},\qquad \mathsf{P}(B\mid \overline{A})=a_{21}. \tag{2.3}$$

Тогда должны быть выполнены равенства

$$\mathsf{P}(\overline{A})=1-a_1=a_2,\qquad \mathsf{P}(\overline{B}\mid A)=1-a_{11}=a_{12},\\ \mathsf{P}(\overline{B}\mid \overline{A})=1-a_{21}=a_{22}.$$

Положим

$$\Omega=\{\omega_1,\ \omega_2,\ \omega_3,\ \omega_4\}.$$

Событиями $A$ и $B$ назовем подмножества

$$A=\{\omega_1,\ \omega_3\},\qquad B=\{\omega_1,\ \omega_2\}.$$

Тогда $AB=\{\omega_1\}$. Аналогично можно проверить, что

$$\{\omega_2\}=\overline{A}B,\qquad \{\omega_3\}=A\overline{B},\qquad \{\omega_4\}=\overline{A}\,\overline{B}.$$

Если подходить к определению $\Omega$ менее формально, то можно сразу положить

$$\Omega=\{AB,\ \bar{A}B,\ A\bar{B},\ \overline{AB}\}.$$

Распределение вероятностей на всех подмножествах конечного множества однозначно определяется заданием вероятностей всех подмножеств, состоящих из одного элементарного события. Положим

$$\begin{aligned}
\mathsf{P}(AB)&=\mathsf{P}(A)\,\mathsf{P}(B\mid A)=a_1a_{11},\\
\mathsf{P}(\bar{A}B)&=\mathsf{P}(\bar{A})\,\mathsf{P}(B\mid\bar{A})=a_2a_{21},\\
\mathsf{P}(A\bar{B})&=\mathsf{P}(A)\,\mathsf{P}(\bar{B}\mid A)=a_1a_{12},\\
\mathsf{P}(\bar{A}\bar{B})&=\mathsf{P}(\bar{A})\,\mathsf{P}(\bar{B}\mid\bar{A})=a_2a_{22}.
\end{aligned}$$

Правые части этих равенств неотрицательны и в сумме дают единицу. Распределение вероятностей задано. Отправляясь от заданного распределения вероятностей, найдем

$$\begin{aligned}
\mathsf{P}(A)&=\mathsf{P}(AB)+\mathsf{P}(A\bar{B})=a_1a_{11}+a_1a_{12}=a_1,\\
\mathsf{P}(B\mid A)&=\frac{\mathsf{P}(AB)}{\mathsf{P}(A)}=\frac{a_1a_{11}}{a_1}=a_{11},\\
\mathsf{P}(B\mid\bar{A})&=\frac{\mathsf{P}(\bar{A}B)}{\mathsf{P}(\bar{A})}=\frac{a_2a_{21}}{1-a_1}=a_{21}.
\end{aligned}$$

Таким образом, мы нашли распределение вероятностей, удовлетворяющее условию (2.3).

Построение вероятностного пространства по условным вероятностям в более общем случае будет рассмотрено в следующей главе.

В примерах, рассматриваемых в этой главе, мы будем обычно предполагать, что соответствующие условные вероятности заданы. Построение по ним вероятностного пространства может быть проведено так же, как в рассмотренном примере, и приводиться не будет.

Решим при помощи условных вероятностей задачу о ключах из § 6 гл. 1, в котором было приведено решение с классическим определением вероятности. Предлагаемое здесь решение является частичным построением другого вероятностного пространства, описывающего последовательное извлечение ключей. Вероятностные пространства, удобные для описания последовательности испытаний, будут подробно рассмотрены в следующей главе. Событие $A_k$, состоящее в том, что нужный ключ появится в $k$-м испытании, можно представить в виде произведения

$$A_k = \overline{A}_1 \overline{A}_2 \ldots \overline{A}_{k-1} A_k.$$

Отсюда, используя формулу (2.2), получим

$$\mathsf{P}(A_k) = \mathsf{P}(\overline{A}_1)\,\mathsf{P}(\overline{A}_2 \mid \overline{A}_1) \ldots \mathsf{P}(A_k \mid \overline{A}_1 \overline{A}_2 \ldots \overline{A}_{k-1}).$$

Значения сомножителей можно считать заданными в условиях задачи. Действительно, из $n$ ключей только один подходит и $n-1$ не подходит. Следовательно, $\mathsf{P}(\overline{A}_1) = \frac{n-1}{n}$. Если произошло событие $\overline{A}_1$, то остался $n-1$ ключ, среди которых один подходит и $n-2$ не подходит. Отсюда $\mathsf{P}(\overline{A}_2 \mid \overline{A}_1) = \frac{n-2}{n-1}$ и т. д. Чтобы приписать вероятности $\mathsf{P}(A_k \mid \overline{A}_1 \ldots \overline{A}_{k-1})$ определенное значение, нужно иметь в виду, что к $k$-му извлечению осталось $n-k+1$ ключей, из которых один подходит к двери. Тогда естественно положить

$$\mathsf{P}(A_k \mid \overline{A}_1 \ldots \overline{A}_{k-1}) = \frac{1}{n-k+1}.$$

Окончательно получаем

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

Результат получился такой же, как в § 6 гл. 1. Если отправляться от распределения вероятностей, введенного для этой задачи в § 6, то можно по формуле (1.1) вычислить условные вероятности $\mathsf{P}(\overline{A}_2 \mid \overline{A}_1)$, $\mathsf{P}(\overline{A}_3 \mid \overline{A}_1 \overline{A}_2), \ldots$ Их числовые значения совпадут со значениями, которые были использованы в этом параграфе.

## § 3. Формула полной вероятности. Формулы Байеса

Пусть $A$ — произвольное событие, события $B_1, B_2, \ldots, B_n$ попарно несовместны, $\mathsf{P}(B_k) > 0$, $k = 1, \ldots, n$, и $A \subset B_1 + B_2 + \ldots + B_n$. Тогда имеет место следующая формула (*формула полной вероятности*):

$$\mathsf{P}(A) = \sum_{k=1}^{n} \mathsf{P}(B_k)\,\mathsf{P}(A \mid B_k). \tag{3.1}$$

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 (1.5.3) и (2.1), получим формулу (3.1):

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

Используя (1.5.8), формулу (3.1) можно распространить на случай счетной системы попарно несовместных событий $B_k$, $k = 1, 2, \ldots, n, \ldots$

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 (3.1), получим *формулы Байеса*

$$\mathsf{P}(B_k \mid A) = \frac{\mathsf{P}(B_k)\,\mathsf{P}(A \mid B_k)}{\sum\limits_{i=1}^{n} \mathsf{P}(B_i)\,\mathsf{P}(A \mid B_i)}. \tag{3.2}$$

Ниже приведены два примера, в которых используются формулы (3.1) и (3.2). Приложения формулы полной вероятности к задачам, связанным со случайным блужданием, и к простейшим задачам теории массового обслуживания приводятся в §§ 5, 6. Представление о некоторых направлениях приложений формул Байеса дает задача 7 этой главы.

Пример 1. На фабрике, изготовляющей болты, первая машина производит 25%, вторая — 35%, третья — 40% всех изделий. Брак в их продукции составляет соответственно 5%, 4%, 2%.

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

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

Решение. а) Обозначим через $A$ событие, состоящее в том, что случайно выбранный болт — дефектный, а через $B_1$, $B_2$, $B_3$ — события, состоящие в том, что этот болт произведен соответственно первой, второй и третьей машинами. Очевидно, что формула (3.1) применима. Таким образом, используя условие задачи, получим

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

б) К тем же событиям можно применить формулы Байеса (3.2) 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 2. Из урны, содержавшей $M$ белых и $N-M$ черных шаров, один шар неизвестного цвета утерян. Какова вероятность извлечь наудачу из урны белый шар?

Решение. 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

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

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

По формуле полной вероятности

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

Отметим, что вероятность извлечь белый шар из урны до утери шара тоже равна $M/N$.

## § 4. Независимость событий

Понятие независимости является одним из важнейших понятий теории вероятностей.

События $A$ и $B$ называются *независимыми*, если

$$\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B). \tag{4.1}$$

Из формулы (1.1) следует, что в случае $\mathsf{P}(A)>0$ и $\mathsf{P}(B)>0$ независимость $A$ и $B$ эквивалентна любому из равенств

$$\mathsf{P}(A\,|\,B)=\mathsf{P}(A),\qquad \mathsf{P}(B\,|\,A)=\mathsf{P}(B). \tag{4.2}$$

Определение независимости в форме (4.1) симметрично относительно $A$ и $B$; условие (4.1) несколько шире, чем условия (4.2).

Если математическая модель, описывающая некоторый опыт, подобрана достаточно хорошо, то независимым событиям реального опыта соответствуют события модели, независимые в смысле определения (4.1). Пусть, например, опыт заключается в том, что один раз бросают две симметричные монеты. В обозначениях § 1 гл. 1 положим $\Omega=\{\text{ГГ, РР, РГ, ГР}\}$; $A=\{\text{ГГ, ГР}\}$ — первая монета выпала гербом вверх, $B=\{\text{РГ, ГГ}\}$ — вторая монета выпала гербом вверх. Предполагая равновероятность элементарных событий, получим

$$\mathsf{P}(A)=\mathsf{P}(B)=\frac{1}{2},\qquad \mathsf{P}(AB)=\frac{1}{4}.$$

Таким образом, $\mathsf{P}(AB)=\mathsf{P}(A)\,\mathsf{P}(B)$. События $A$ и $B$ оказались независимыми в смысле определения (4.1).

Об использовании независимости (формулы (4.1)) при построении подходящего вероятностного пространства можно сделать замечания, аналогичные замечаниям о использовании формулы (2.2).

## § 5. Примеры приложений формулы полной вероятности

**5.1. Случайные блуждания.** По целым точкам отрезка $[0, n]$ движется частица. Пусть $\xi_t$ — координата частицы в момент $t$, $t=0, 1, 2, \ldots$, и $\xi_0=k$. В каждый момент времени $t$, $t=0, 1, 2, \ldots$, выбирается направление движения независимо от всех предыдущих выборов. С вероятностью $p$ частица сдвигается на единицу вправо и с вероятностью $q=1-p$ — на единицу влево *). Если частица попала в точку 0 или $n$, то она там остается в любой последующий момент времени. При $t=0$ частица находилась в точке $k$. Требуется определить вероятность $\pi_{kn}=\mathsf{P}(A_k)$ события $A_k$, состоящего в том, что частица когда-нибудь попадет в точку $n$.

Эту задачу можно интерпретировать как задачу о разорении игрока. Пусть в начале игры 1-й игрок имеет $k$ рублей, а 2-й игрок — $n-k$ рублей. Если при бросании монеты (или кости) выпал герб (или «6»), то 1-й игрок получает 1 рубль (это соответствует движению частицы вправо), а в противном случае отдает 1 рубль (движение частицы влево). Поглощению частицы на правом конце соответствует выигрыш первого игрока.

За пространство элементарных событий $\Omega$ в этой задаче можно принять бесконечные последовательности, составленные из результатов выборов направлений движения. Предположим, что можно выбрать $\sigma$-алгебру $\mathfrak{F}$ и задать на ней вероятность $\mathsf{P}$, удовлетворяющую условиям задачи (см. § 4 гл. 3). Положим $\pi_{kn}(t)=\mathsf{P}(\xi_t=n)$. Можно показать, что

$$\begin{aligned} &\mathsf{P}(\xi_{t+1}=n \mid \xi_t=k+1)=\pi_{k+1,\,n}(t), \\ &\mathsf{P}(\xi_{t+1}=n \mid \xi_1=k-1)=\pi_{k-1,\,n}(t), \qquad k=1, \ldots, n-1. \end{aligned} \tag{5.1}$$

*) Эти условия можно записать в виде

$$\mathsf{P}(\xi_{t+1}=l+1 \mid \xi_t=l,\, C_t)=p,$$
$$\mathsf{P}(\xi_{t+1}=l-1 \mid \xi_t=l,\, C_t)=q=1-p$$

при любых значениях $l=1, 2, \ldots, n-1$, $C_t$ — любое событие, относящееся к движению точки до момента $t$.

Здесь вероятности $\pi_{k+1,n}(t)$ и $\pi_{k-1,n}(t)$ являются вероятностями попадания в состояние $n$ к моменту времени $t$ в схемах блуждания, начавшегося в момент $t=0$, из точек $k+1$ и $k-1$ соответственно. Эти равенства (5.1) достаточно очевидны: если первым переходом был переход, например, из $k$ в $k+1$, то вероятность за оставшееся время $t$ попасть из $k+1$ в $n$ будет такая же, как если бы процесс начался из точки $k+1$. Равенства (5.1) будут доказаны в § 1 гл. 9, пример 1.

По формуле полной вероятности

$$\begin{aligned} \mathsf{P}(\xi_{t+1}=n)=\mathsf{P}(\xi_1=k+1)\,\mathsf{P}(\xi_{t+1}=n\mid\xi_1=k+1)+\\ +\mathsf{P}(\xi_1=k-1)\,\mathsf{P}(\xi_{t+1}=n\mid\xi_1=k-1). \end{aligned} \tag{5.2}$$

По условию задачи

$$\mathsf{P}(\xi_1=k+1)=p,\qquad \mathsf{P}(\xi_1=k-1)=q, \tag{5.3}$$

так как $\xi_0=k$. Заменив вероятности, входящие в (5.2), по формулам (5.1), (5.3), получим

$$\begin{gathered} \pi_{kn}(t+1)=p\pi_{k+1,n}(t)+q\pi_{k-1,n}(t), \\ k=1, 2, \ldots, n-1. \end{gathered} \tag{5.4}$$

Заметим, что

$$(\xi_1=n)\subset(\xi_2=n)\subset\ldots\subset(\xi_t=n)\subset(\xi_{t+1}=n)\subset\ldots$$

и $A_k=\bigcup\limits_{t=1}^{\infty}(\xi_t=n)$. Таким образом, по формуле (1.5.9)

$$\pi_k=\mathsf{P}(A_k)=\lim_{t\to\infty}\mathsf{P}(\xi_t=n)=\lim_{t\to\infty}\pi_{kn}(t).$$

Переходя к пределу при $t\to\infty$ в (5.4), получим

$$\pi_{k,n}=p\pi_{k+1,n}+q\pi_{k-1,n},\qquad k=1, 2, \ldots, n-1.$$

Так как $\pi_{k,n}$ — вероятность поглощения на правом конце в процессе, начавшемся из точки $k$, то $\pi_{0,n}=0$, $\pi_{nn}=1$. Таким образом, вероятность $\pi_{kn}$, рассматриваемая как функция от $k$, является решением линейного однородного уравнения в конечных разностях с постоянными коэффициентами

$$pf_{k+1}-f_k+qf_{k-1}=0, \tag{5.5}$$

удовлетворяющим условиям

$$f_0=0, \quad f_n=1. \tag{5.6}$$

Теория таких уравнений (см. А. О. Гельфонд [3]) во многом аналогична теории линейных дифференциальных уравнений с постоянными коэффициентами.

Пусть сначала $p \neq q$. Подставим $f_k=\lambda^k$ в уравнение (5.5):

$$p\lambda^{k+1}-\lambda^k+q\lambda^{k-1}=0$$

или

$$p\lambda^2-\lambda+q=0.$$

Корни этого уравнения равны

$$\lambda_1=1, \quad \lambda_2=\frac{q}{p}.$$

Следовательно, функции $\lambda_1^k$ и $\lambda_2^k$ удовлетворяют уравнению (5.5). Линейная комбинация решений

$$f_k=C_1\lambda_1^k+C_2\lambda_2^k=C_1+\left(\frac{q}{p}\right)^k C_2 \tag{5.7}$$

при любых значениях постоянных $C_1$, $C_2$ тоже является решением. Подставив в (5.6) правую часть (5.7) при $k=0$ и $k=n$, получим для определения $C_1$, $C_2$ два уравнения

$$C_1+C_2=0, \quad C_1+\left(\frac{q}{p}\right)^n C_2=1.$$

Отсюда и из (5.7) окончательно находим

$$\pi_{kn}=\frac{1-\left(\frac{q}{p}\right)^k}{1-\left(\frac{q}{p}\right)^n}, \quad k=0,\ 1,\ 2,\ \ldots,\ n. \tag{5.8}$$

Вероятности поглощения на левом конце $\pi_{k0}$ удовлетворяют тому же уравнению (5.5), но решать его нужно с условиями

$$f_0=1, \quad f_n=0 \tag{5.9}$$

Подобрав в (5.7) постоянные $C_1$ и $C_2$ согласно (5.9), получим

$$\pi_{k0}=\frac{\left(\frac{q}{p}\right)^k-\left(\frac{q}{p}\right)^n}{1-\left(\frac{q}{p}\right)^n}. \tag{5.10}$$

Так как $\pi_{kn}+\pi_{k0}=1$, то блуждание закончится с вероятностью 1.

Пусть теперь $p=q=\frac{1}{2}$. В этом случае $\lambda_1=\lambda_2=1$, и решение (5.5) нужно искать в виде

$$f_k=C_1+kC_2.$$

Используя (5.6) и (5.9), получим

$$\pi_{kn}=\frac{k}{n}, \quad \pi_{k0}=1-\frac{k}{n}.$$

В схеме блуждания по целым точкам прямой с поглощением только в нуле (в $n$ поглощение не происходит) вероятность попасть когда-нибудь в $n$ равна вероятности $\pi_{kn}$, вычисленной для схемы с поглощением в нуле и в $n$. Вероятность того, что частица побывает во всех точках правее $k$, равна

$$\pi_k=\lim_{n\to\infty}\pi_{kn}=\begin{cases} 0, & \text{если} \quad q\geqslant p, \\ 1-\frac{q}{p}, & \text{если} \quad q<p. \end{cases}$$

Этот результат не противоречит интуитивным представлениям. Если движение вправо более вероятно, чем влево, то с положительной вероятностью частица может уйти вправо; в противном случае с вероятностью 1 колебания ограничены и происходит поглощение в 0.

**5.2. Работа телефонной линии.** 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.

Пусть $P_0(t)$, $P_1(t)$ — вероятность того, что линия в момент $t$ соответственно свободна и занята.

Предположим сразу, что построение подходящего вероятностного пространства возможно и что для интересующих нас событий вероятности $P_0(t)$ и $P_1(t)$ определены и непрерывны по $t$. Используя формулу полной вероятности, получим для $P_0(t)$, $P_1(t)$ дифференциальные уравнения.

Пусть $B_0(t)$, $B_1(t)$ — события, состоящие в том, что в момент $t$ линия соответственно свободна и занята. Тогда $B_0(t)+B_1(t)=\Omega$ и $B_0(t)B_1(t)=\varnothing$. Применим формулу полной вероятности к вычислению $\mathsf{P}(B_0(t+h))$, положив в ней $n=2$, $A=B_0(t+h)$, $B_1=B_0(t)$, $B_2=B_1(t)$:

$$\mathsf{P}(B_0(t+h))=\mathsf{P}(B_0(t))\mathsf{P}(B_0(t+h)\,|\,B_0(t))+\\+\mathsf{P}(B_1(t))\mathsf{P}(B_0(t+h)\,|\,B_1(t)). \quad (5.11)$$

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

$$\mathsf{P}(B_0(t+h)\,|\,B_0(t))=1-\alpha h+o(h).$$

Занятая в момент $t$ линия будет свободной к моменту $t+h$, если закончится разговор и не поступит ни один вызов. Вероятность этого события равна

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

Эта вероятность вносит основной вклад в $\mathsf{P}(B_0(t+h)\,|\,B_1(t))$. Сумма остальных слагаемых равна $o(h)$. Таким образом,

$$\mathsf{P}(B_0(t+h)\,|\,B_1(t))=\beta h+o(h).$$

Подставляя найденные условные вероятности в (5.11) и используя обозначения $P_0(t)$, $P_1(t)$, получим

$$P_0(t+h)=(1-\alpha h)P_0(t)+\beta hP_1(t)+o(h).$$

Отсюда

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

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

$$\frac{dP_0(t)}{dt} = -\alpha P_0(t) + \beta P_1(t). \tag{5.12}$$

Аналогично найдем уравнение для $P_1(t)$:

$$\frac{dP_1(t)}{dt} = \alpha P_0(t) - \beta P_1(t). \tag{5.13}$$

Полагая $P_0(0) = 1$, $P_1(0) = 0$, найдем решение системы (5.12)—(5.13):

$$\begin{aligned} P_0(t) &= \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}, \\ P_1(t) &= \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-(\alpha+\beta)t}. \end{aligned} \tag{5.14}$$

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

Из формул (5.14) следует, что решение $(P_0(t), P_1(t))$ системы (5.12), (5.13) удовлетворяет равенству $P_0(t) + P_1(t) = 1$. Это же равенство следует из предположения о существовании вероятностного пространства, в котором определены вероятности $P_0(t)$ и $P_1(t)$ противоположных событий. Если подставить $P_1(t) = 1 - P_0(t)$ в уравнение (5.12), то вместо системы можно решить одно уравнение для $P_0(t)$.

К предположениям о существовании подходящего вероятностного пространства, в котором определены вероятности нужных нам событий, нужно относиться с известной осторожностью. Рассмотрим следующий пример. Пусть состояниями частицы являются числа 0, 1, 2, ... При $t = 0$ частица находится в состоянии 0. Если частица в момент $t$ находится в состоянии $k$, то к моменту $t+h$ останется в этом состоянии с вероятностью $1 - \alpha_k h + o(h)$, $h \to 0$, и перейдет в состояние $k+1$ с вероятностью $\alpha_k h + o(h)$. Пусть

*) Из существования предела при $h \to 0$ следует существование только левосторонних производных в (5.12) и (5.13). Заменив в предыдущих рассуждениях $t$ и $t+h$ на $t-h$ и $t$, получим выражения для правосторонних производных, совпадающие с (5.12) и (5.13).

$P_k(t)$ — вероятность того, что частица в момент $t$ находится в состоянии $h$. Рассуждая так же, как в рассмотренном выше примере, можно для $P_k(t)$, $k=0, 1, 2, \ldots$, получить бесконечную систему дифференциальных уравнений. Можно было ожидать, что $\sum\limits_{k=1}^{\infty} P_k(t)=1$. Однако если $\alpha_k$ быстро увеличиваются с ростом $k$, то оказывается, что $\sum\limits_{k=0}^{\infty} P_k(t)<1$. Более подробно этот пример рассмотрен в книге Феллера [17], § 4 гл. 17. Оказывается, что частица за конечное время с положительной вероятностью может пройти по всем состояниям. В этом случае набор чисел $\{\alpha_k\}$ уже не определяет процесс при всех $t>0$.

### Задачи к главе 2

1. Бросаются три кости. Какова вероятность того, что хотя бы на одной из них выпала единица, если на трех костях выпали разные грани?

2. Известно, что при бросании 10 костей появилась, по крайней мере, одна единица. Какова вероятность того, что появилось две единицы или более?

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

4. В первой урне 2 белых и 3 черных шара, а во второй 1 белый и 4 черных. Из первой урны во вторую переложили два шара. Найти вероятность того, что вынутый из второй урны шар окажется белым.

5. Среди 64 клеток шахматной доски выбирают наудачу две клетки и ставят на них двух слонов. Какова вероятность того, что они не будут бить друг друга?

6. 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.)

7. 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.

8. 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 $\lambda h+o(h)$, $h\longrightarrow 0$. Найти вероятность того, что время свободного пробега будет больше $t$.

9. На одну телефонную линию могут поступать вызовы двух типов: срочные и простые. При поступлении срочного вызова разговор по простому вызову прекращается. Вероятности поступления за время $(t,\ t+h)$ срочного и обычного вызовов равны соответственно $\alpha_1 h+o(h)$, $\alpha_2 h+o(h)$, $h\longrightarrow 0$; вероятность прекращения любого разговора за время $(t,\ t+h)$ равна $\beta h+o(h)$. Пусть $P_0(t)$, $P_1(t)$, $P_2(t)$ — вероятности того, что в момент $t$ линия свободна, занята срочным вызовом, занята простым вызовом. Написать для $P_k(t)$ дифференциальные уравнения и найти $\lim\limits_{t\to\infty} P_k(t)=\pi_k$, $k=0,\ 1,\ 2$.

10*. Изменим условия работы телефонной линии, описанной в п. 2 § 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 $\theta=\dfrac{\alpha}{\beta}<1$.

ГЛАВА 3

# ПОСЛЕДОВАТЕЛЬНОСТИ ИСПЫТАНИЙ

## § 1. Конечные последовательности испытаний

Рассмотрим следующую задачу.

В первой урне 2 белых и 3 черных шара; во второй — 2 белых и 2 черных; в третьей — 3 белых и 1 черный. Из первой урны наугад один шар переложен во вторую. После этого из второй урны также наугад переложен один шар в третью. Наконец, из третьей какой-то из шаров переложен в первую. Какой состав шаров в первой урне является наиболее вероятным? Что более вероятно: изменение состава шаров первой урны или сохранение?

Согласно условию задачи мы имеем три испытания (три перекладывания шаров). Нетрудно выписать все 8 возможных результатов испытаний. Введем сначала следующие обозначения. Пусть $A_k$ — событие, состоящее в том, что при $k$-м перекладывании ($k=1,\ 2,\ 3$) был переложен белый шар. Положим

$$\Omega=\{A_1A_2A_3,\ \overline{A}_1A_2,\ A_3,\ A_1\overline{A}_2A_3,\ A_1A_2\overline{A}_3,\ \overline{A}_1\overline{A}_2\overline{A}_3,\ A_1\overline{A}_2\overline{A}_3,\ \overline{A}_1A_2\overline{A}_3,\ \overline{A}_1\overline{A}_2A_3\}.$$

Здесь, например, элементарное событие $A_1\overline{A}_2\overline{A}_3$ означает, что первый переложенный шар был белый, а два других — черные. Событием $A_1$ является следующее подмножество $\Omega$:

$$A_1=\{A_1A_2A_3,\ A_1\overline{A}_2A_3,\ A_1A_2\overline{A}_3,\ A_1\overline{A}_2\overline{A}_3\}.$$

Нетрудно проверить, что на введенные формально обозначения для элементарных событий можно смотреть как на произведения случайных событий. Если состав шаров в данной урне известен, то мы легко можем определить вероятность события, состоящего в извлечении из этой урны шара определенного цвета. В рассматриваемой задаче состав шаров в урне становится известным, если известно, какой шар в данную урну был переложен. Таким образом, мы можем считать, что заданы вероятности $\mathsf{P}(A_1), \mathsf{P}(\overline{A}_1), \mathsf{P}(A_2 \mid A_1), \mathsf{P}(A_3 \mid A_1\overline{A}_2)$ и т. д. Например,

$$\mathsf{P}(A_1)=\frac{2}{5}, \quad \mathsf{P}(A_2 \mid A_1)=\frac{3}{5}, \quad \mathsf{P}(A_3 \mid A_1A_2)=\frac{4}{5}. \tag{1.1}$$

При определении второй и третьей вероятностей мы использовали то, что во вторую и третью урны был переложен белый шар и их состав стал соответственно: 3 белых, 2 черных; 4 белых, 1 черный. Аналогично можно приписать значения другим условным вероятностям. По заданным условным вероятностям так же, как в § 2 гл. 2, стр. 43, мы можем полностью восстановить распределение вероятностей на подмножествах $\Omega$. По формулам (2.2.2) и (1.1) находим

$$\mathsf{P}(A_1A_2A_3)=\mathsf{P}(A_1)\mathsf{P}(A_2 \mid A_1)\mathsf{P}(A_3 \mid A_1A_2)= \\ =\frac{2}{5}\cdot\frac{3}{5}\cdot\frac{4}{5}=\frac{24}{125}.$$

Аналогично приписываются вероятности другим элементарным событиям. Окончательно получим

$$\begin{aligned}
&\mathsf{P}(A_1A_2A_3)=\frac{24}{125}, && \mathsf{P}(\overline{A}_1\overline{A}_2\overline{A}_3)=\frac{18}{125},\\
&\mathsf{P}(\overline{A}_1A_2A_3)=\frac{24}{125}, && \mathsf{P}(A_1\overline{A}_2\overline{A}_3)=\frac{8}{125},\\
&\mathsf{P}(A_1\overline{A}_2A_3)=\frac{12}{125}, && \mathsf{P}(\overline{A}_1A_2\overline{A}_3)=\frac{6}{125},\\
&\mathsf{P}(A_1A_2\overline{A}_3)=\frac{6}{125}, && \mathsf{P}(\overline{A}_1\overline{A}_2A_3)=\frac{27}{125}.
\end{aligned} \tag{1.2}$$

По этим вероятностям однозначно определяется вероятность любого случайного события. Пусть $B_l$, $l=1, 2, 3$,— событие, состоящее в том, что после перекладываний в первой урне оказалось $l$ белых шаров.

Тогда

$$B_1=\{A_1A_2\bar{A}_3,\quad A_1\bar{A}_2\bar{A}_3\},$$
$$B_2=\{A_1A_2A_3,\quad A_1\bar{A}_2A_3,\quad \bar{A}_1A_2\bar{A}_3,\quad \bar{A}_1\bar{A}_2A_3\},$$
$$B_3=\{\bar{A}_1A_2A_3,\quad \bar{A}_1\bar{A}_2A_3\}$$

и

$$\mathsf{P}(B_1)=\frac{14}{125},\qquad \mathsf{P}(B_2)=\frac{60}{125},\qquad \mathsf{P}(B_3)=\frac{51}{125}.$$

Вероятность сохранения состава шаров — 60/125. Таким образом, более вероятно изменение состава шаров в первой урне, но наиболее вероятным составом является первоначальный состав. В решенной задаче легко проверить, что (1.2) определяют распределения вероятностей. Дадим теперь общее определение последовательности из $n$ испытаний, в каждом из которых может произойти один из $N$ исходов. Исходы в каждом испытании занумеруем числами: $1, 2, \ldots, N$. Под *последовательностью из $n$ испытаний* мы будем понимать (см. гл. 1 § 6.2) дискретное вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$, в котором

$$\Omega=\{(l_1l_2\ldots l_n)\},\quad l_k=1, 2, \ldots, N;\quad k=1, \ldots, n, \tag{1.3}$$

и вероятности $p(l_1l_2\ldots l_k)$, приписываемые элементарным событиям $\omega=(l_1l_2\ldots l_n)$, задаются формулой

$$p(l_1l_2\ldots l_n)=p(l_1)\,p(l_2\,|\,l_1)\,p(l_3\,|\,l_1l_2)\ldots p(l_n\,|\,l_1\ldots l_{n-1}), \tag{1.4}$$

где числа $p(l_1)$, $p(l_2\,|\,l_1)$, $\ldots$, $p(l_n\,|\,l_1l_2\ldots l_{n-1})$ удовлетворяют условиям:

$$\begin{aligned}&1)\ p(l_1)\geqslant 0,\ \sum_{l_1=1}^{N}p(l_1)=1;\\ &2)\ p(l_2\,|\,l_1)\geqslant 0,\ \sum_{l_2=1}^{N}p(l_2\,|\,l_1)=1 \text{ при любом } l_1;\end{aligned} \tag{1.5}$$

3) $p(l_3 \mid l_1 l_2) \geqslant 0, \quad \sum_{l_3=1}^{N} p(l_3 \mid l_1 l_2) = 1$ при любых $l_1, l_2$;

. . . . . . . . . . . . . . . . . . . . . . . . . . .

$n$) $p(l_n \mid l_1, \ldots, l_{n-1}) \geqslant 0, \quad \sum_{l_n=1}^{N} p(l_n \mid l_1, \ldots, l_{n-1}) = 1$

при любых $l_1, l_2, \ldots, l_{n-1}$.

Символ $l_k$, $k = 1, 2, \ldots, n$, в обозначении элементарного события $(l_1 l_2 \ldots l_n)$ будем интерпретировать как наступление исхода $l_k$ в испытании с номером $k$. Формула (1.6.6) с числами

$$p_{u_1(l_1)\ldots u_n(l_n)} = p(l_1 l_2 \ldots l_n)$$

задает распределение вероятностей, если

$$\sum_{l_1, \ldots, l_n=1}^{N} p(l_1 l_2 \ldots l_n) = 1. \tag{1.6}$$

Для доказательства (1.6) заметим, что

$$\sum_{l_1, \ldots, l_n=1}^{N} p(l_1 l_2 \ldots l_n) = \sum_{l_1 \ldots l_{n-1}=1}^{N} \left( \sum_{l_n=1}^{N} p(l_1 l_2 \ldots l_n) \right) =$$
$$= \sum_{l_1 \ldots l_{n-1}=1}^{N} p(l_1) \ldots p(l_{n-1} \mid l_1 \ldots l_{n-2}) \times$$
$$\times \left( \sum_{l_n=1}^{N} p(l_n \mid l_1 \ldots l_{n-1}) \right) =$$
$$= \sum_{l_1 \ldots l_{n-1}=1}^{N} p(l_1) p(l_2 \mid l_1) \ldots p(l_{n-1} \mid l_1 \ldots l_{n-2}),$$

так как $\sum_{l_n=1}^{N} p(l_n \mid l_1 \ldots l_{n-1}) = 1$ согласно (1.5). Далее мы можем выделить сумму по $l_{n-1}$ и вновь воспользоваться (1.5). Повторив этот процесс $n$ раз, получим (1.6). Таким образом, мы определили вероятностное пространство, являющееся математической моделью последовательности из $n$ испытаний. Нетрудно проверить, что число $p(l_k \mid l_1 l_2 \ldots l_{k-1})$ является условной вероятностью появления в $k$-м испытании исхода $l_k$ при условии, что до этого была получена цепочка исходов $(l_1 l_2 \ldots l_{k-1})$.

Более общая схема получится, если считать, что множество исходов в отдельном испытании — счетно.

Решенная в начале параграфа задача является следующим частным случаем рассмотренной общей схемы испытаний: $n=3$, $N=2$.

$$1)\ p(1)=\frac{2}{3},\quad p(2)=\frac{3}{5};$$

$$2)\ p(1|1)=\frac{3}{5},\ p(2|1)=\frac{2}{5},\ p(1|2)=\frac{2}{5},\ p(2|2)=\frac{3}{5};$$

$$3)\ p(1|11)=p(1|21)=\frac{4}{5},\quad p(2|11)=p(2|21)=\frac{1}{5},$$

$$p(1|12)=p(1|22)=\frac{3}{5},\quad p(2|12)=p(2|22)=\frac{2}{5},$$

где исходы «1» и «2» соответствуют извлечению белого и черного шара.

В данной задаче оказалось, что вероятности $p(i|jk)$ не зависят от $j$. Это естественно, так как на состав последней урны влияет только цвет шара, переложенного из второй урны, и если этот цвет уже известен, то неважно, что было переложено во вторую урну из первой. Последовательность испытаний, в которой условные вероятности $p(l_t|l_1 \ldots l_{t-1})$ не зависят от $l_1, \ldots, l_{t-2}$,

$$p(l_t|l_1 \ldots l_{t-1})=p^{(t)}_{l_{t-1} l_t}$$

называется *цепью Маркова*. Более подробно такие испытания будут рассмотрены в гл. 9. В случае, когда $p(l_t|l_1 \ldots l_{t-1})$ не зависят от $l_1, \ldots, l_{t-1}$, последовательность испытаний называется *последовательностью независимых испытаний*.

## § 2. Последовательность независимых испытаний

В § 1 было дано определение последовательности независимых испытаний как частного случая общей схемы испытаний. Дадим прямое определение. Под *последовательностью $n$ независимых испытаний*, в каждом из которых может осуществиться один из $N$ исходов (обозначим исходы $1, 2, \ldots, N$), мы будем понимать вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$, в котором

$$\Omega = \{(l_1 l_2 \ldots l_n)\}, \quad l_k = 1, 2, \ldots, N, \quad k = 1, 2, \ldots, n, \tag{2.1}$$

и вероятности $p(l_1 l_2 \ldots l_n)$, приписываемые цепочкам из результатов отдельных испытаний, задаются формулой

$$p(l_1 l_2 \ldots l_n) = p_{l_1} p_{l_2} \ldots p_{l_n}, \tag{2.2}$$

где $p_k \geqslant 0$, $k = 1, \ldots, N$, $\sum_{k=1}^{N} p_k = 1$.

В данном случае равенство

$$\sum_{l_1, \ldots, l_n = 1}^{N} p(l_1 l_2 \ldots l_n) = 1$$

проверяется аналогично равенству (1.6).

Число $p_k$, $k = 1, 2, \ldots, N$, входящее в (2.2), является вероятностью появления исхода $k$ в фиксированном испытании. Действительно, если событие $A_1(k)$ заключается в том, что в первом испытании наступил исход $k$, то

$$A_1(k) = \{(l_1 l_2 \ldots l_n) \colon l_1 = k\}$$

и

$$\mathsf{P}(A_1(k)) = \sum_{l_2, \ldots, l_n = 1}^{N} p_k p_{l_2} \ldots p_{l_n} = p_k.$$

Аналогично вычисляются вероятности событий более общего вида. Например, для

$$A_{i_1 \ldots i_s}(L_1, \ldots, L_s) = \{(l_1 l_2 \ldots l_n) \colon l_{i_1} = L_1, \ldots, l_{i_s} = L_s\}$$

имеем

$$\mathsf{P}(A_{i_1 \ldots i_s}(L_1, \ldots, L_s)) = p_{L_1} p_{L_2} \ldots p_{L_s}.$$

Верна следующая теорема.

Теорема 2.1. *События*

$$A = A_{i_1 \ldots i_s}(L_1, \ldots, L_s), \qquad B = B_{j_1 \ldots j_l}(L_1', \ldots, L_l')$$

*независимы, если* $(i_1 \ldots i_s) \cap (j_1 \ldots j_l) = \varnothing$.

Проведем доказательство в частном случае. Пусть

$$A = \{(l_1 \ldots l_n)\colon\ l_1 = 2\}, \qquad B = \{(l_1 \ldots l_n)\colon\ l_2 = 1,\ l_4 = 3\}.$$

Тогда

$$AB = \{(l_1 \ldots l_n)\colon\ l_1 = 2,\ l_2 = 1,\ l_4 = 3\}.$$

По формулам (1.6.6) и (2.2) находим

$$\mathsf{P}(A) = \sum_{l_2 \ldots l_n = 1}^{N} p_2 p_{l_2} \ldots p_{l_n} = p_2,$$

$$\mathsf{P}(B) = \sum_{l_1, l_3, l_5, l_6, \ldots, l_n = 1}^{N} p_{l_1} p_1 p_{l_3} p_3 p_{l_5} p_{l_6} \ldots p_{l_n} = p_1 p_3,$$

$$\mathsf{P}(AB) = \sum_{l_3, l_5, l_6, \ldots, l_n = 1}^{N} p_2 p_1 p_{l_3} p_3 p_{l_5} \ldots p_{l_n} = p_2 p_1 p_3.$$

Отсюда

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

Следовательно, события $A$ и $B$ согласно определению (2.4.1) независимы. Общий случай рассматривается аналогично.

Схема независимых испытаний является математической моделью серии испытаний, повторяющихся при неизменных условиях. Примеры таких испытаний можно найти в задачах к данной главе.

Независимые испытания при $N = 2$ называют *испытаниями Бернулли*. В этом случае исходы 1 и 2 называют соответственно «успехом» и «неудачей», и их вероятности $p_1$ и $p_2$ полагают равными $p$ и $q = 1 - p$. Элементарные события в этом случае естественно обозначать цепочками вида

$$\underbrace{\text{УУНННУ} \ldots \text{У}}_{n}. \tag{2.3}$$

По формуле (2.2) имеем

$$\mathsf{P}(\text{УУНННУ} \ldots \text{У}) = p^m q^{n-m}, \tag{2.4}$$

где $m$ — число «успехов» (У) в последовательности (2.3). Во многих задачах мы часто каждому элементарному событию ставим в соответствие определенное действительное число, например, размер выигрыша, число успехов в (2.3) и т. д. Пусть задано вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$. Назовем случайной величиной действительную функцию от элементарного события: $\xi = \xi(\omega)$. Для дискретных вероятностных пространств случайной величиной мы будем называть произвольную функцию от $\omega$, а в общем случае, рассматриваемом в гл. 4, на функцию $\xi(\omega)$ будут налагаться дополнительные условия. Случайные величины часто обозначают греческими буквами.

Обозначим $\mu_n = \mu_n$ (УУНННУ ... У) случайную величину, равную числу успехов в первых $n$ испытаниях схемы Бернулли. Тогда, например, при $n = 4$ имеем

$$\mu_4 = \mu_4(\text{УУНУ}) = 3, \qquad \mu_4 = (\text{НННН}) = 0$$

и т. д. Найдем вероятности событий

$$\{\mu_n = m\} = \{(\text{УУНННУ}\ldots\text{У}):\ \mu_n(\text{УУНННУ}\ldots\text{У}) = m\}. \tag{2.5}$$

Теорема 2.2. *Если $\mu_n$ — число успехов в $n$ испытаниях Бернулли, то*

$$\mathsf{P}(\mu_n = m) = p_n(m) = C_n^m p^m q^{n-m}, \qquad m = 0, 1, 2, \ldots, n. \tag{2.6}$$

Доказательство. Каждая цепочка вида (2.3), обозначающая элементарное событие из множества (2.5), содержит ровно $m$ успехов, и, следовательно, вероятности всех элементарных событий из (2.5) одинаковы и определяются формулой (2.4). Найдем число элементарных событий, удовлетворяющих условию $(\mu_n = m)$. Число успехов и неудач задано. Таким образом, можно только менять их расположение в цепочках вида (2.3). Расположение однозначно определяется выбором из $n$ мест $m$ мест для успехов. Это можно сделать $C_n^m$ способами. Отсюда и из (2.4) следует утверждение теоремы.

Для схемы испытаний с произвольным $N$ (эту схему часто называют полиномиальной) введем случайные величины $\xi_k$, равные числам исходов $k$, $k=1, 2, \ldots, N$. Покажем, что

$$\mathsf{P}(\xi_1=m_1, \xi_2=m_2, \ldots, \xi_N=m_N)=\frac{n!}{m_1!\ldots m_N!}p_1^{m_1}\ldots p_N^{m_N}, \tag{2.7}$$

где $m_1, m_2, \ldots, m_N$ — неотрицательные целые числа, удовлетворяющие условию $m_1+m_2+\ldots+m_N=n$. Для остальных значений $m_k$ левая часть (2.7) равна 0. В каждой цепочке $(l_1 l_2 \ldots l_n)$, удовлетворяющей условию $(\xi_1=m_1, \xi_2=m_2, \ldots, \xi_N=m_N)$, «1» встречается $m_1$ раз, «2» — $m_2$ раз, ..., «$N$» — $m_N$ раз. Следовательно, вероятность каждой такой цепочки равна $p_1^{m_1}p_2^{m_2}\ldots p_N^{m_N}$. Число цепочек заданного вида можно найти следующим образом: $C_n^{m_1}$ способами можно выбрать из $n$ мест $m_1$ мест для «1»; из оставшихся $n-m_1$ мест можно выбрать $C_{n-m_1}^{m_2}$ способами места для «2» и т. д. Отсюда общее число цепочек нужного нам вида равно

$$C_n^{m_1}C_{n-m_1}^{m_2}\ldots C_{n-m_1-\ldots-m_{N-1}}^{m_N}=\frac{n!}{m_1!m_2!\ldots m_N!}.$$

Из приведенных рассуждений следует формула (2.7). При $N=2$ формула (2.7) совпадает с (2.6).

Рассмотрим в полиномиальной схеме случайные события, полученные следующим образом. С элементарным событием $\omega_0=(l_1 l_2 \ldots l_n)$, содержащим ровно $m$ исходов «1» $(m=0, 1, 2, \ldots, n)$, мы объединим все элементарные события $\omega$, у которых «единицы» стоят на тех же местах и нет «единиц» на других местах. Полученное таким образом случайное событие обозначим $\omega^*$. Нетрудно найти вероятность события $\omega^*$. Пусть, например, в $\omega^*$ объединены все элементарные события, у которых $m$ единиц на первых местах. Тогда

$$p(\omega^*)=p_1^m\sum_{l_{m+1}, \ldots, l_n}^{*} p_{l_{m+1}}\ldots p_{l_n},$$

где знаком $\sum^*$ обозначено суммирование по всем возможным наборам $l_{m+1}, \ldots, l_n$, в которых нет «1».

Следовательно,

$$\sum_{i_{m+1},\ldots,i_n}^{*} p_{i_{m+1}}\cdots p_{i_n} =$$
$$= \left(\sum_{i_{m+1}=2}^{N} p_{i_{m+1}}\right)\left(\sum_{i_{m+2}=2}^{N} p_{i_{m+2}}\right)\cdots\left(\sum_{i_N=2}^{N} p_{i_N}\right) =$$
$$= (1-p_1)^{n-m}$$

и

$$p(\omega^*) = p_1^m (1-p_1)^{n-m}. \qquad (2.8)$$

Эта формула совпадает с (2.4). Для задания события $\omega^*$ достаточно указать для каждого испытания, наступил исход «1» или не наступил. Если принять события $\omega^*$ за элементарные события в новой схеме и приписать им вероятности (2.8), то получится схема испытаний Бернулли с $p=p_1$ и $q=1-p_1=p_2+\ldots+p_N$.

При $n$ подбрасываниях игральной кости мы имеем полиномиальную схему с $N=6$ и $p_1=p_2=\ldots=p_6=\frac{1}{6}$. Если теперь «укрупнить» элементарные события, например, различать только «6» и «не шестерка», то получим схему Бернулли с $p=\frac{1}{6}$ и $q=\frac{5}{6}$.

## § 3. Предельные теоремы в схеме Бернулли

Пусть монета брошена 5 раз. Требуется найти вероятность того, что выпало ровно 3 «герба». Воспользовавшись формулой (2.6) при $n=5$, $p=q=\frac{1}{2}$, $m=3$, получим

$$P_5(3) = C_5^3 \frac{1}{2^5} = \frac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3}\cdot\frac{1}{2^5} = \frac{5}{16}.$$

В этом примере вычисления проводятся по формуле (2.6) легко. Однако часто приходится вычислять вероятности (2.6) при больших значениях $n$ и $m$. Например, при $n=1000$, $m=500$ и $p=q=\frac{1}{2}$ найти числовое значение вероятности (2.6) довольно трудно. При вычислении вероятности события, состоящего в том, что число успехов $\mu_n$ лежит, например, между $a$ и $b$, приходится находить числовые значения сумм вероятностей $P_m(b)$ вида (2.6). Действительно,

$$\mathsf{P}(a<\mu_n<b)=\sum_{a<m<b}\mathsf{P}(\mu_n=m)=\sum_{a<m<b}P_n(m). \tag{3.1}$$

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

Иногда при больших $n$ удается заменить формулу (2.6) какой-либо приближенной асимптотической формулой. Приведем три предельные теоремы, содержащие асимптотические формулы для вероятностей (2.6) и (3.1) при $n\to\infty$.

Теорема 3.1 (теорема Пуассона). *Если $n\to\infty$ и $p\to 0$ так, что $np\to\lambda$, $0<\lambda<\infty$, то*

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

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

Доказательство. Положив $np=\lambda_n$, представим вероятность $\mathsf{P}(\mu_n=m)$ в виде

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

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

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

$$\mathsf{P}(\mu_n=m)\approx\frac{\lambda_n^m}{m!}e^{-\lambda_n},\qquad \lambda_n=np. \tag{3.2}$$

Часто эта формула используется при $n>100$ и $np<30$.

Теорема 3.2 (локальная теорема Муавра—Лапласа). *Если $n\to\infty$, $p$ $(0<p<1)$ постоянно, величина $x_m=\dfrac{m-np}{\sqrt{npq}}$ ограничена равномерно по $m$ и $n$,*

$(-\infty < a \leqslant x_m \leqslant b < +\infty)$, *то*

$$\mathsf{P}(\mu_n = m) = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{x_m^2}{2}} (1 + \alpha_n(m)),$$

*где* $|\alpha_n| < C/\sqrt{n}$ *при* $x_m \in [a, b]$, $C > 0$ — *постоянная.*

Доказательство. Коэффициент $C_n^m$ в (2.6) запишем в виде

$$C_n^m = \frac{n!}{m!\,(n-m)!}. \tag{3.3}$$

Так как

$$m = np + x_m \sqrt{npq}, \qquad n - m = nq - x_m \sqrt{npq},$$

то, воспользовавшись формулой Стирлинга

$$\ln n! = \ln \sqrt{2\pi n} + n \ln n - n + O\left(\frac{1}{n}\right), \tag{3.4}$$

получим

$$\begin{aligned} \ln m! &= \ln \sqrt{2\pi m} + (np + x_m \sqrt{npq}) \ln (np + x_m \sqrt{npq}) - \\ &\qquad - np - x_m \sqrt{npq} + O\left(\frac{1}{n}\right), \\ \ln (n-m)! &= \ln \sqrt{2\pi (n-m)} + \\ &\qquad + (nq - x_m \sqrt{npq}) \ln (nq - x_m \sqrt{npq}) - \\ &\qquad - nq + x_m \sqrt{npq} + O\left(\frac{1}{n}\right). \end{aligned} \tag{3.5}$$

Логарифмы в (3.5) при помощи формулы $\ln(1+x) = x - x^2/2 + O(x^3)$ запишем в виде

$$\begin{aligned} \ln (np + x_m \sqrt{npq}) &= \ln np + \ln\left(1 + x_m \sqrt{\frac{q}{np}}\right) = \\ &= \ln np + x_m \sqrt{\frac{q}{np}} - \frac{x_m^2}{2} \frac{q}{np} + O\left(\frac{q^3}{n^{3/2} p^3}\right), \\ \ln (nq - x_m \sqrt{npq}) &= \\ &= \ln nq - x_m \sqrt{\frac{p}{nq}} - \frac{x_m^2}{2} \frac{p}{nq} + O\left(\frac{p^3}{n^{3/2} q^3}\right). \end{aligned} \tag{3.6}$$

Утверждение теоремы можно получить из (2.6), (3.3), (3.4), (3.5), (3.6).

При качественной оценке условий применимости приближенной формулы

$$\mathsf{P}(\mu_n = m) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{x_m^2}{2}} \tag{3.7}$$

нужно оценить величину остаточных членов в (3.6). При $n \to \infty$ сумма остаточных членов стремится к 0 при любых фиксированных $p$ и $q$, $0 < p < 1$. Однако при конечных значениях $n$ сумма остаточных членов может быть очень большой, если $p$ или $q$ малы. Хорошие приближения формула (3.7) дает при $p = q = 1/2$. Если в этом случае провести более точную оценку остаточного члена, то в формулировке теоремы можно заменить $|\alpha_n| < C/\sqrt{n}$ на $|\alpha_n| < C/n$. Формулу (3.7) часто используют при $n > 100$ и $npq > 20$. Указания о границах применимости формул (3.2) и (3.7) являются очень приближенными и носят скорее качественный характер; к ним следует относиться с осторожностью.

Отметим еще, что в условиях теоремы 3.2 из того, что $n \to \infty$, следует стремление к бесконечности $m$. Это значит, что $m$ и $n$ в (3.7) должны отличаться друг от друга не очень сильно; например, для $\mathsf{P}(\mu_n = 0)$ локальная теорема дает плохое приближение.

Теорема 3.3 (интегральная теорема Муавра — Лапласа). *Если $p$, $0 < p < 1$, постоянно, то при $n \to \infty$*

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

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

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

$$\mathsf{P}\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) = P_n(a, b) = \sum_{a \leqslant x_m \leqslant b} \mathsf{P}(\mu_n = m), \tag{3.8}$$

где $x_m = \frac{m - np}{\sqrt{npq}}$ и суммирование распространяется на значения $m$, для которых $x_m \in [a, b]$. Применяя к слагаемым (3.8) локальную теорему, получим

$$P_n(a, b) = S_n + T_n, \tag{3.9}$$

где

$$S_n = \sum_{x_m \in [a, b]} \varphi(x_m) \frac{1}{\sqrt{npq}},$$

$$T_n = \sum_{x_m \in [a, b]} \alpha_n(m) \varphi(x_m) \frac{1}{\sqrt{npq}},$$

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

Так как

$$\Delta x_m = x_{m+1} - x_m = \frac{m+1-np}{\sqrt{npq}} - \frac{m-np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}},$$

то $S_n$ можно записать в виде суммы

$$S_n = \sum_{x_m \in [a, b]} \varphi(x_m) \Delta x_m,$$

которая отличается не более чем двумя слагаемыми от подходящим образом выбранной интегральной суммы, соответствующей интегралу $\int\limits_a^b \varphi(x)\,dx$. Следовательно,

$$\lim_{n\to\infty} S_n = \int\limits_a^b \varphi(x)\,dx. \tag{3.10}$$

Используя оценку для $\alpha_n(m)$ из локальной теоремы, получим

$$|T_n| \leqslant \sum_{x_m \in [a, b]} \varphi(x_m) \Delta x_m |\alpha_n(m)| \leqslant \frac{C}{\sqrt{n}} S_n.$$

Таким образом, при $n \to \infty$

$$T_n \to 0. \tag{3.11}$$

Утверждение теоремы для постоянных $a$, $b$ следует из формул (3.8), (3.9), (3.10), (3.11).

Из доказанного и из теоремы 2.4 гл. 7 следует равномерная сходимость. Приближенная формула

$$\mathsf{P}\left(a \leqslant \frac{\mu_n - np}{\sqrt{npq}} \leqslant b\right) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx \qquad (3.12)$$

используется в тех случаях, когда возможно использование (3.7). Численное значение интеграла можно найти, если воспользоваться таблицами для функции

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}}\, du. \qquad (3.13)$$

Формула (3.12) позволяет оценить близость частоты и вероятности. Пусть $p$ — вероятность успеха в схеме Бернулли и $\mu_n$ — общее число успехов. Частотой успеха называют отношение $\mu_n/n$. Оценим вероятность события $\left(\left|\frac{\mu_n}{n} - p\right| < \Delta\right)$. Если $n$ достаточно велико, то можно воспользоваться формулой (3.12). Тогда

$$\mathsf{P}\left(\left|\frac{\mu_n}{n} - p\right| < \Delta\right) = \mathsf{P}\left(-\Delta\sqrt{\frac{n}{pq}} < \frac{\mu_n - np}{\sqrt{npq}} < \Delta\sqrt{\frac{n}{pq}}\right) \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int_{-\Delta\sqrt{\frac{n}{pq}}}^{\Delta\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}}\, dx = 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right), \qquad (3.14)$$

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 в конце книги.

Часто возникает обратная задача: сколько нужно провести испытаний, чтобы частота $\mu_n/n$ отличалась от вероятности $p$ не больше, чем на $\Delta$, с вероятностью $1-2\alpha$ ($\alpha$ мало)? Такого типа задачи возникают при использовании метода Монте-Карло (метод статистических испытаний). Идея метода заключается в моделировании случайного процесса или последовательности испытаний, вероятностные характеристики которых просто связаны с подлежащими вычислению величинами. Если много раз бросать иглу, как описано в задаче Бюффона (см. § 6, гл. 1), то частота $\mu_n/n$ будет мало отличаться от вероятности $p$ пересечения иглой какой-либо линии. Зная величину отклонения $\mu_n/n$ от $p$, можно оценить ошибку в определении числа $\pi$. В таких задачах естественно считать $p$ неизвестным. Тогда, чтобы подобрать наименьшее $n$, при котором вероятность отклонения будет равна $1-2\alpha$, нужно согласно (3.14) 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 (3.14), используя неравенство $pq\leqslant 1/4$, получим

$$\mathsf{P}\left(\left|\frac{\mu_n}{n}-p\right|<\Delta\right)\approx 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right)\geqslant$$
$$\geqslant 2\Phi_0(2\Delta\sqrt{n})=1-2\alpha$$

и для определения $n$ имеем уравнение $\Phi_0(2\Delta\sqrt{n})=\frac{1-2\alpha}{2}$. По таблице можно найти $u_\alpha$, для которых $\Phi_0(u_\alpha)=\frac{1-2\alpha}{2}$. Тогда $2\Delta\sqrt{n}=u_\alpha$ и

$$n\geqslant\frac{u_\alpha^2}{4\Delta^2}.$$

Довольно часто используются значения $2\alpha$, равные 0,05 и 0,01. Для этих значений имеем

| $2\alpha$ | 0,01 | 0,05 |
|---|---|---|
| $u_\alpha$ | 2,58 | 1,96 |

В задаче с иглой вовсе не обязательно проводить реальные подбрасывания иглы. Можно этот процесс моделировать; например, получить последовательность величин $(\varphi_k, x_k)$, $k=1, 2, \ldots, n$ (см. § 6 гл. 1), задающих положение иглы. Если эти величины можно быстро ввести в ЭВМ, то даже при больших $n$ будет легко определена частота пересечений. При моделировании различных процессов используются последовательности «случайных» или «псевдослучайных» чисел. Случайными числами называют числа, полученные при реализации последовательности независимых испытаний, в каждом из которых любая фиксированная цифра появляется с вероятностью 1/10. Псевдослучайные последовательности вырабатываются при помощи какого-либо алгоритма и имеют структуру, близкую к случайным последовательностям.

## § 4. Бесконечные последовательности испытаний

В ряде интересных задач приходится рассматривать бесконечные последовательности испытаний. В § 5.1 гл. 2 введение бесконечной последовательности испытаний потребовалось при рассмотрении задачи о случайном блуждании.

Введем бесконечную последовательность независимых испытаний с двумя исходами в каждом. Пусть $\Omega$ состоит из бесконечных последовательностей

$$\omega=(e_1, e_2, e_3, \ldots, e_n, \ldots), \qquad e_n=0;\ 1, \qquad n=1, 2, \ldots$$

Рассмотрим систему событий $\mathfrak{F}_0$, состоящую из событий вида

$$A_{n_1 n_2 \ldots n_k}(E_k)=\{\omega\colon (e_{n_1}, e_{n_2}, \ldots, e_{n_k})\in E_k\},$$

где $E_k$ — произвольное множество $k$-мерных векторов; $n_1<n_2<\ldots<n_k$. В $\mathfrak{F}_0$ входят, например, множества

$$A_1(0)=\{\omega\colon e_1=0\}=\{(0e_2e_3\ldots e_n\ldots)\},$$
$$A_1(1)=\{\omega\colon e_1=1\}=\{(1e_2e_3\ldots e_n\ldots)\},$$
$$A_{12}(\{(0, 0), (0, 1), (1, 0)\})=$$
$$=\{\omega\colon (e_1, e_2)\in\{(0, 0), (0, 1), (1, 0)\}\}=$$
$$=\{(00e_3\ldots e_n\ldots), (01e_3\ldots e_n\ldots), (10e_3\ldots e_n\ldots)\}.$$

Нетрудно проверить, что $\mathfrak{F}_0$ является алгеброй событий.

Определим на $\mathfrak{F}_0$ вероятность следующим образом. Любое событие $A_{n_1 n_2 \ldots n_k}(E_k)$ можно представить в виде суммы несовместных событий вида

$$B_{n_k}(\varepsilon_1\varepsilon_2\ldots\varepsilon_{n_k})=\{\omega\colon e_1=\varepsilon_1, e_2=\varepsilon_2, \ldots, e_{n_k}=\varepsilon_{n_k}\}.$$

Вероятности этих событий $B_{n_k}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n_k})$ определим так же, как в схеме Бернулли с $n_k$ испытаниями определялись вероятности элементарных событий. Например, положим $\mathsf{P}(B_4(0111)) = qppp = qp^3$; в общем случае

$$\mathsf{P}(B_{n_k}(\varepsilon_1\varepsilon_2\ldots\varepsilon_{n_k})) = p^{\varepsilon_1}q^{1-\varepsilon_1}p^{\varepsilon_2}q^{1-\varepsilon_2}\ldots p^{\varepsilon_{n_k}}q^{1-\varepsilon_{n_k}}. \tag{4.1}$$

Тогда

$$\mathsf{P}(A_{n_1n_2\ldots n_k}(E_k)) = \sum \mathsf{P}(B_{n_k}(\varepsilon_1\varepsilon_2\ldots\varepsilon_{n_k})), \tag{4.2}$$

где суммирование проводится по всем значениям вектора $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n_k})$, удовлетворяющим условию $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n_k}) \in E_k$. Одно и то же событие из алгебры $\mathfrak{F}_0$ можно записать несколькими способами; можно формально увеличить число координат, на которые накладываются ограничения, а в множестве $E_k$ вновь введенным координатам разрешить принимать любые значения. Например, события

$$\begin{gathered} A_1(1) = \{(1e_2e_3\ldots e_n\ldots)\}, \\ A_{12}(\{(1, 0), (1, 1)\}) = \{(10e_3\ldots), (11e_3\ldots)\} \end{gathered} \tag{4.3}$$

совпадают. В связи с этим нужно проверить, что вероятность определяется формулой (4.2) однозначно. Для событий (4.3) имеем

$$A_1(1) = B_1(1), \quad A_{12}(\{(1, 0), (1, 1)\}) = B_{12}(1, 0) + B_{12}(1, 1)$$

и, следовательно, по формулам (4.1), (4.2)

$$\mathsf{P}(A_1(1)) = p, \quad \mathsf{P}(A_{12}\{(1, 0), (1, 1)\}) = pq + p^2 = p.$$

Таким образом, при различной форме записи одного и того же события получили одинаковые значения вероятности. Аналогично проверяется общий случай.

Покажем теперь, что для вероятности, определенной формулами (4.1), (4.2), выполняется аксиома A4. Пусть $C_1, C_2 \in \mathfrak{F}_0$, $C_1C_2 = \varnothing$ и $n$ — наибольший номер координаты, на которую накладываются ограничения в какой-либо фиксированной форме записи событий $C_1$ и $C_2$. Тогда $C_1$ и $C_2$ можно записать в виде

$$C_1 = A_{12\ldots n}(E_n'), \qquad C_2 = A_{12\ldots n}(E_n''), \qquad E_n'E_n'' = \varnothing$$

и, следовательно,

$$\begin{aligned} \mathsf{P}(C_1 + C_2) &= \sum_{\varepsilon \in E_n' + E_n''} \mathsf{P}(B_n(\varepsilon)) = \\ &= \sum_{\varepsilon \in E_n'} \mathsf{P}(B_n(\varepsilon)) + \sum_{\varepsilon \in E_n''} \mathsf{P}(B_n(\varepsilon)) = \mathsf{P}(C_1) + \mathsf{P}(C_2), \\ &\qquad \varepsilon = (\varepsilon_1\varepsilon_2\ldots\varepsilon_n). \end{aligned}$$

Проверим свойство непрерывности вероятности $\mathsf{P}$, определенной на $\mathfrak{F}_0$. Если $A_1 \supset A_2 \supset \ldots \supset A_n \supset A_{n+1} \supset \ldots$ и $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$, то, начиная с некоторого $n$, имеем $A_n = \varnothing$ (если бы $A_n \neq \varnothing$ при всех $n$, то событие $\bigcap\limits_{n=1}^{\infty} A_n$, определяемое непротиворечивыми условиями на счетное множество координат последовательности $\omega$, не могло бы быть невозможным$\Big)$. Таким образом, аксиома А5 доказана, так как $\mathsf{P}(A_n)=0$ для всех $n$, начиная с некоторого. По теореме о продолжении вероятности (см. § 4 гл. 1) определенную на $\mathfrak{F}_0$ вероятность можно продолжить на минимальную $\sigma$-алгебру $\mathfrak{F}$, порожденную $\mathfrak{F}_0$. Таким образом, мы завершим построение вероятностного пространства, являющегося моделью бесконечной последовательности испытаний.

Вернемся теперь к задаче о случайном блуждании из п. 5.1 § 5 гл. 2. С последовательностью $\omega=(e_1e_2\ldots)\in\Omega$ движение частицы связано следующим образом: если $\xi_t=l$, то $\xi_{t+1}=\xi_t+2e_t-1$, $0<l<n$. События $\{\xi_t=n\}$ определяются условиями на конечное число координат и, следовательно, $\{\xi_t=n\}\in\mathfrak{F}_0$ при любом фиксированном $t$. Так как $A_k=\bigcup\limits_{t=1}^{\infty}\{\xi_t=n\}$ и $\{\xi_t=n\}$, $t=1, 2, \ldots,$ — монотонная последовательность событий, то $A_k\in\mathfrak{F}$ и $\lim\limits_{t\to\infty}\mathsf{P}(\xi_t=n)$ является вероятностью события $A_k$.

## Задачи к главе 3

**1.** Урна содержит $b$ черных и $r$ красных шаров. Наудачу извлекается шар. Вынутый шар возвращается обратно и добавляется $c$ шаров того же цвета. Найти вероятности: $\mathsf{P}(A_k)$, $k=1, 2, 3$, $\mathsf{P}(A_1\mid A_2)$, $\mathsf{P}(A_1A_2\overline{A}_3)$, $\mathsf{P}(\overline{A}_1\overline{A}_2A_3)$, где $A_k$ — событие, состоящее в том, что в $k$-м испытании появился черный шар.

**2.** В урне 2 белых и 4 черных шара. Два игрока поочередно извлекают шар (без возвращения). Выигрывает тот, кто первым вынет белый шар. Вычислить вероятность выигрыша для каждого участника.

**3.** Пусть в общей схеме последовательности испытаний (см. § 1, (1.3), (1.4), (1.5)) $n=3$. Доказать, что $p(l_1)$ — вероятность наступления исхода $l_1$ в первом испытании; $p(l_2\mid l_1)$ — условная вероятность наступления $l_2$ во втором испытании при условии, что в первом было $l_1$; $p(l_3\mid l_1l_2)$ — условная вероятность наступления $l_3$ в третьем испытании при условии, что в первом и втором были $l_1$ и $l_2$.

**4.** Что вероятнее, выиграть у равносильного противника 3 партии из 4 или 5 из 8 (ничьи не бывают)?

**5.** Испытание заключается в бросании 3 игральных костей. Найти вероятность того, что при 10 испытаниях ровно в 4 испытаниях появится в точности по две «6».

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

**7.** Сколько раз нужно бросить монету, чтобы частота выпадения герба отличалась от 1/2 не более чем на $\Delta$ с вероятностью, не меньшей $1-2\alpha$ $(2\alpha=0{,}05,\ \Delta=0{,}1)$? Использовать таблицу случайных чисел для моделирования бросании симметричной монеты. Выписать реализацию необходимой длины; найти частоту выпадения герба.

**8.** На отрезок $[0, 10]$ наудачу брошено 5 точек. Найти вероятность того, что две точки попадут в $[0, 2]$, одна в $[2, 3]$, две в $[3, 10]$.

**9.** В круг вписан квадрат. Какова вероятность того, что из 10 точек, брошенных наудачу в круг, четыре попадут в квадрат, три — в один сегмент и по одной в оставшиеся три сегмента?

**10.** Двое бросают правильную монету по $n$ раз каждый. Найти вероятность того, что у них выпадет одинаковое число гербов.

**11.** Найти вероятность того, что в $2n$ испытаниях по схеме Бернулли в первых $n$ испытаниях было $m$ успехов, а в последних $n$ испытаниях $l$ успехов и один из них наступил в испытании с номером $2n$; $p$ — вероятность успеха в каждом испытании.

**12.** Брошено 6 правильных игральных костей. Какова вероятность выпадения: 1) хотя бы одной; 2) ровно одной; 3) ровно двух единиц? Найти точные значения и сравнить с формулой Пуассона.

**13.** Какова вероятность того, что среди 200 человек будет не менее четырех левшей, если левши в среднем составляют 1%?

**14.** Сколько изюма должна в среднем содержать булочка, чтобы вероятность иметь хотя бы одну изюмину в булке была не менее 0,99?

**15.** Вероятность попадания в цель при каждом выстреле равна 0,001. Найти вероятность попадания в цель двумя и более выстрелами при залпе в 5000 выстрелов

**16.** Найти приближенное выражение того, что число выпадений «1» при 12 000 бросании игральной кости заключено между 1900 и 2150.

**17.** В поселке $A$ 2500 жителей. Каждый из них примерно 6 раз в месяц ездит на поезде в город $B$, выбирая дни поездок по случайным мотивам независимо от остальных. Какой наименьшей вместительностью должен обладать поезд, чтобы он переполнялся в среднем не чаще одного раза в 100 дней? (Поезд идет раз в сутки.)

ГЛАВА 4

# СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

## § 1. Определения и примеры

В § 2 гл. 3 было дано определение случайной величины для дискретных вероятностных пространств. Пусть теперь $(\Omega, \mathfrak{F}, \mathsf{P})$ — произвольное вероятностное пространство. Случайной величиной $\xi$ назовем действительную функцию $\xi = \xi(\omega)$, $\omega \in \Omega$, такую, что при любом действительном $x$

$$\{\omega : \xi(\omega) < x\} \in \mathfrak{F}. \tag{1.1}$$

Если в $\mathfrak{F}$ включаются все подмножества $\Omega$, то (1.1), очевидно, выполняется. Событие в (1.1) более коротко будем иногда записывать в виде $\xi < x$. Так как $\mathfrak{F}$ является $\sigma$-алгеброй, то из (1.1) следует, что

$$\begin{gathered}(\xi \geqslant x) = \overline{(\xi < x)} \in \mathfrak{F}, \\ (x_1 \leqslant \xi < x_2) = (\xi < x_2) \setminus (\xi < x_1) \in \mathfrak{F}, \\ (\xi = x) = \bigcap_{n=1}^{\infty} \left( x \leqslant \xi < x + \frac{1}{n} \right) \in \mathfrak{F}.\end{gathered} \tag{1.2}$$

Таким образом, вероятности этих событий определены. Для вычисления вероятностей событий вида (1.2) не требуется полностью знать $\mathsf{P}(A)$, $A \in \mathfrak{F}$; достаточно при любом $x$ знать вероятность

$$F_\xi(x) = \mathsf{P}(\xi < x). \tag{1.3}$$

Функция $F_\xi(x)$ действительной переменной $x$, $-\infty < x < \infty$, называется функцией распределения случайной величины $\xi$. Так как

$$(\xi < x_2) = (x_1 \leqslant \xi < x_2) + (\xi < x_1),$$

то согласно аксиоме А4

$$\mathsf{P}(\xi < x_2) = \mathsf{P}(x_1 \leqslant \xi < x_2) + \mathsf{P}(\xi < x_1).$$

Отсюда и из (1.3) находим

$$\mathsf{P}(x_1 \leqslant \xi < x_2) = F_\xi(x_2) - F_\xi(x_1). \tag{1.4}$$

По формуле (1.5.1)

$$\mathsf{P}(\xi \geqslant x) = 1 - F_\xi(x).$$

Из (1.5.9), (1.4) и (1.2) получим

$$\mathsf{P}(\xi = x) = \lim_{n\to\infty}\left(F_\xi\left(x + \frac{1}{n}\right) - F_\xi(x)\right) = \\ = F_\xi(x+0) - F_\xi(x). \tag{1.5}$$

По функции распределения можно вычислить вероятности любых событий вида (1.2).

Иногда вместо $F_\xi(x)$ будем писать просто $F(x)$.

Пример 1. Два игрока по одному разу подбрасывают симметричную монету. Если выпал «герб», то первый игрок получает 1 рубль, а если выпала «решетка», то—отдает 1 рубль. Для описания данной игры естественно положить $\Omega = \{\Gamma, \text{Р}\}$ и $\mathsf{P}(\{\Gamma\}) = \mathsf{P}(\{\text{Р}\}) = \frac{1}{2}$. Случайная величина $\xi$, равная выигрышу первого игрока, определяется следующим образом:

$$\xi = \xi(\Gamma) = 1, \qquad \xi = \xi(\text{Р}) = -1.$$

Легко вычисляется функция распределения $F_\xi(x)$ величины $\xi$. Если $x \leqslant -1$, то множество $\{\xi < x\}$ является пустым и его вероятность равна 0. Если $-1 < x \leqslant 1$, то $\{\xi < x\} = \{\text{Р}\}$ и, следовательно, $F_\xi(x) = \mathsf{P}(\xi < x) = \frac{1}{2}$. При $x > 1$ имеем $\{\xi < x\} = \{\Gamma, \text{Р}\}$ и $\mathsf{P}(\xi < x) = 1$. Таким образом (рис. 4),

$$F_\xi(x) = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ \frac{1}{2}, & \text{если} \quad -1 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1. \end{cases}$$

Пример 2. Пусть в единичный квадрат $\Omega = \{(u, v): 0 \leqslant u \leqslant 1, 0 \leqslant v \leqslant 1\}$ наудачу брошена точка. Элементарными событиями $\omega$ являются точки квадрата $\Omega$; $\sigma$-алгебра $\mathfrak{F}$ здесь порождается квадрируемыми подмножествами квадрата. Вероятность события, являющегося квадрируемым подмножеством, равна его площади. Случайными величинами являются, например, функции $\xi=\xi(\omega)=\sqrt{u^2+v^2}$ — расстояние брошенной точки от начала координат, $\eta=\eta(u, v)=u$ — первая координата брошенной точки и т. д. Найдем

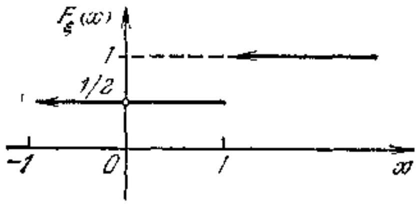


Рис. 4.

функцию распределения величины $\eta$. При $0<x<1$ имеем

$$\{(u, v):\ \xi(u, v)<x\}=\{(u, v):\ u<x\}.$$

Стороны этого прямоугольника равны 1 и $x$. Таким образом (рис. 5),

$$\mathbf{P}(\{(u, v):\ u<x\})=x$$

и

$$F_\eta(x)=\begin{cases} 0, & \text{если} \quad x\leqslant 0, \\ x, & \text{если} \quad 0<x\leqslant 1, \\ 1, & \text{если} \quad x>1. \end{cases}$$

Пример 3. Рассмотрим следующий опыт. Пусть один раз подбрасывается монета. Если выпал «герб», то на этом опыт заканчивается. Если выпала «решетка», то на отрезок $[0, 1]$ наудачу бросается точка. Для описания этого опыта введем следующее пространство элементарных событий:

$$\Omega=\{\Gamma;\ (\mathrm{P}, u)\},$$

где $0\leqslant u\leqslant 1$; $\sigma$-алгебра $\mathfrak{F}$ пусть порождается событиями

$$\{\Gamma\};\ \{(\mathrm{P}, u\}:\ a\leqslant u<b\},$$

где $0 \leqslant a \leqslant b \leqslant 1$. Положим

$$\mathsf{P}(\{\Gamma\}) = \frac{1}{2}, \quad \mathsf{P}(\{(\mathrm{P}, u):\ a \leqslant u < b\}) = \frac{b-a}{2}.$$

По вероятностям этих событий однозначно определяется вероятность на $\mathfrak{F}$. Рассмотрим случайную

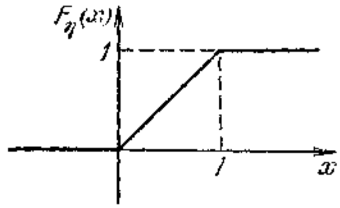


Рис. 5.

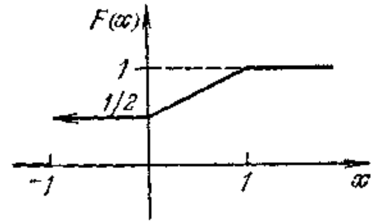


Рис. 6.

величину $\zeta$, определяемую равенствами

$$\zeta(\Gamma) = -1, \quad \zeta(\mathrm{P}, u) = u.$$

Так как

$$(\zeta < x) = \begin{cases} \varnothing, & \text{если} \quad x \leqslant -1, \\ \{\Gamma\}, & \text{если} \quad -1 < x \leqslant 0, \\ \{\Gamma\} + \{(\mathrm{P}, u):\ 0 \leqslant u < x\}, & \text{если} \quad 0 < x < 1, \end{cases}$$

то (рис. 6)

$$F_\zeta(x) = \begin{cases} 0, & \text{если} \quad x \leqslant -1, \\ \frac{1}{2}, & \text{если} \quad -1 < x \leqslant 0, \\ \frac{x+1}{2}, & \text{если} \quad 0 < x \leqslant 1, \\ 1, & \text{если} \quad x > 1. \end{cases}$$

## § 2. Свойства функций распределения

Пусть $F(x)$ — функция распределения некоторой случайной величины $\xi$. По определению функция $\xi(\omega)$ принимает только конечные значения. Формула (1.3) определяет $F(x)$ при любом действительном $x$, при $x = +\infty$ и при $x = -\infty$. Очевидно, что

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

Теорема 2.1. *Функция распределения $F(x)$ обладает следующими свойствами:*

1. *Если* $x_1 \leqslant x_2$, *то* $F(x_1) \leqslant F(x_2)$.
2. $\lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0$; $\lim\limits_{x\to+\infty} F(x) = F(+\infty) = 1$.
3. $\lim\limits_{x\to x_0-0} F(x) = F(x_0)$ (*непрерывность слева*).

Доказательство.

1. Так как $(\xi < x_1) \subset (\xi < x_2)$ при $x_1 \leqslant x_2$, то неравенство $F(x_1) \leqslant F(x_2)$ следует из (1.5.6).

2. Последовательность событий $(\xi < -n)$, $n = 1, 2, \ldots$, монотонно убывает, т. е.

$$(\xi < -1) \supset (\xi < -2) \supset \ldots \supset (\xi < -n) \supset (\xi < -n-1) \supset \ldots$$

и

$$\bigcap_{n=1}^{\infty} (\xi < -n) = \varnothing.$$

По формуле (1.5.9) получим

$$\lim_{n\to+\infty} F(-n) = \lim_{n\to+\infty} \mathsf{P}(\xi < -n) = \mathsf{P}(\varnothing) = 0.$$

Отсюда с учетом доказанной монотонности функции $F(x)$ получаем первое равенство во втором свойстве. Второе равенство доказывается аналогично с использованием монотонной последовательности $(\xi < n)$, $n = 1, 2, \ldots$

3. Пусть числовая последовательность $x_n$ возрастает и $\lim\limits_{n\to\infty} x_n = x_0$. Тогда

$$(\xi < x_n) \subset (\xi < x_{n+1}), \qquad \bigcup_{n=1}^{\infty} (\xi < x_n) = (\xi < x_0)$$

и по формуле (1.5.9)

$$\lim_{n\to\infty} \mathsf{P}(\xi < x_n) = \mathsf{P}(\xi < x_0).$$

Отсюда с учетом монотонности $F(x)$ получим третье свойство. Теорема доказана.

Иногда вместо (1.3) функцию $F(x)$ определяют по формуле $F(x) = \mathsf{P}(\xi \leqslant x)$. При таком определении функция $F(x)$ непрерывна справа: $\lim\limits_{x\to x_0+0} F(x) = F(x_0)$.

Любая функция $G(x)$, обладающая тремя свойствами, указанными в теореме 2.1, является функцией распределения некоторой случайной величины, т. е. можно построить вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ и определить на нем случайную величину $\xi$ такую, что $F_\xi(x)=G(x)$. Положим $\Omega=\{u\colon -\infty<u<\infty\}$. Обозначим $\mathfrak{F}_0$ алгебру, порожденную полуинтервалами $[u_1, u_2)$. На этих полуинтервалах зададим вероятность с помощью равенства

$$\mathsf{P}([u_1, u_2])=G(u_2)-G(u_1). \tag{2.1}$$

Эта формула однозначно определяет вероятность на алгебре $\mathfrak{F}_0$. По теореме о продолжении вероятности мы можем единственным образом продолжить заданную на $\mathfrak{F}_0$ вероятность на минимальную $\sigma$-алгебру $\mathfrak{F}$, порожденную $\mathfrak{F}_0$. Если мы теперь положим $\xi=\xi(u)=u$, то $F_\xi(x)=\mathsf{P}(\xi<x)=\mathsf{P}((-\infty, x))=G(x)-G(-\infty)=G(x)$. При изучении отдельной случайной величины можно заменить исходное вероятностное пространство $(\Omega, \mathfrak{F}, \mathsf{P})$ на вероятностное пространство $(\tilde{\Omega}, \tilde{\mathfrak{F}}, \tilde{\mathsf{P}})$, в котором $\tilde{\Omega}$ является числовой прямой, $\tilde{\mathfrak{F}}$ системой борелевских множеств, $\tilde{\mathsf{P}}$ определяется формулой (2.1) с $G(x)=F_\xi(x)$. Вероятность $\tilde{\mathsf{P}}$ называют распределением случайной величины.

Функция распределения однозначно определяет $\tilde{\mathsf{P}}$. Однако в ряде частных случаев распределения случайной величины можно задавать в более удобной форме. Отметим два типа распределений, которые часто встречаются в приложениях.

Случайная величина $\xi$ называется *величиной дискретного типа*, если существует конечное или счетное множество чисел $x_1, x_2, \ldots, x_n, \ldots$ (без предельных точек) таких, что

$$\mathsf{P}(\xi=x_n)=p_n>0, \quad n=1, 2, \ldots; \qquad \sum_{n=1}^{\infty} p_n=1.$$

Очевидно, для дискретной величины закон распределения полностью определяется указанием значений $x_n$, $n=1, 2, \ldots$, и вероятностей $p_n$, с которыми эти значения принимает случайная величина. Функция распределения дискретной величины является ступенчатой. В точке $x_n$, $n=1, 2, \ldots$, она имеет скачок, равный $p_n$.

Случайная величина $\xi$ называется *величиной абсолютно непрерывного типа*, если существует неотрицательная функция $p_\xi(x)$ такая, что при любом $x$

$$F_\xi(x) = \mathsf{P}(\xi < x) = \int_{-\infty}^{x} p_\xi(u)\,du.$$

Ниже мы будем всегда предполагать, что $p_\xi(x)$ непрерывна всюду, за исключением конечного числа точек. Функция $p_\xi(x)$ называется *плотностью распределения вероятностей*. Очевидно,

$$\mathsf{P}(a \leqslant \xi < b) = \int_a^b p_\xi(x)\,dx, \tag{2.2}$$

$$\mathsf{P}(\xi = a) = \lim_{n\to\infty} \mathsf{P}\left(a \leqslant \xi < a + \frac{1}{n}\right) = \lim_{n\to\infty} \int_a^{a+\frac{1}{n}} p_\xi(x)\,dx = 0.$$

Отсюда следует, что для абсолютно непрерывных величин

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \mathsf{P}(a < \xi \leqslant b) = \mathsf{P}(a \leqslant \xi < b) = \mathsf{P}(a < \xi < b).$$

Если $x$ является точкой непрерывности $p_\xi(x)$, то при $\Delta x \to 0$

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

Это равенство следует из (2.2). Плотность распределения обладает следующими свойствами:

1) $p_\xi(x) \geqslant 0$, $-\infty < x < \infty$;

2) $\int_{-\infty}^{\infty} p_\xi(x)\,dx = 1$;

3) $F'_\xi(x) = p_\xi(x)$ в точках непрерывности $p_\xi(x)$.

Плотность распределения полностью определяет распределение случайной величины. Функция распределения абсолютно непрерывной величины, очевидно, непрерывна.

Пример 3 из § 1 показывает, что существуют распределения, не принадлежащие ни одному из указанных типов.

Приведем ряд часто встречающихся абсолютно непрерывных распределений:

1. *Нормальное распределение*

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad -\infty < a < \infty; \quad \sigma > 0. \quad (2.3)$$

Случайная величина с плотностью распределения (2.3) называется нормально распределенной с параметрами $(a, \sigma)$.

2. *Показательное распределение*

$$p_\xi(x) = \begin{cases} \alpha e^{-\alpha x}, & \text{если} \quad x > 0, \\ 0, & \text{если} \quad x \leqslant 0. \end{cases} \quad (2.4)$$

3. *Равномерное распределение*

$$p_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b], \quad a < b. \end{cases} \quad (2.5)$$

Нетрудно показать, что существуют случайные величины, для которых функции (2.3)—(2.5) являются плотностями распределения. Функции (2.3)—(2.5) неотрицательны, и интеграл от них по всей числовой прямой равен единице.

В определении абсолютно непрерывного вероятностного пространства (см. п. 6.4 § 6 гл. I) положим $n = 1$, $\Omega = \{u: -\infty < u < \infty\}$, $\pi(u) = p_\xi(u)$. Случайная величина $\xi = \xi(u) = u$ имеет плотность распределения $p_\xi(x)$.

Примеры дискретных распределений:

1. *Биномиальное распределение*

$$\mathsf{P}(\xi = m) = C_n^m p^m (1-p)^{n-m}, \quad 0 < p < 1; \qquad m = 0, 1, 2, \ldots, n. \quad (2.6)$$

Здесь $n$ — натуральное число.

2. *Пуассоновское распределение*

$$\mathsf{P}(\xi=m)=\frac{\lambda^m}{m!}e^{-\lambda},\quad m=0,\ 1,\ 2,\ \ldots;\quad \lambda>0. \quad (2.7)$$

3. *Геометрическое распределение*

$$\mathsf{P}(\xi=m)=(1-p)^{m-1}p,\quad m=1,2,\ldots;\quad 0<p<1.$$

4. *Гипергеометрическое распределение*

$$\mathsf{P}(\xi=m)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n},\quad m=0,1,2,\ldots,\min(M,n).$$

Наиболее простым вероятностным пространством, на котором можно определить дискретную случайную величину с заданным распределением вероятностей, является дискретное вероятностное пространство (см. п. 6.2 § 6 гл. 1) с $\Omega$, состоящим из значений случайной величины. Например, случайную величину с пуассоновским распределением можно определить следующим образом. Пусть $\Omega=\{0, 1, 2, \ldots\}$. Положим $p_m=\frac{\lambda^m}{m!}e^{-\lambda}$. Эти числа удовлетворяют равенству (1.6.4):

$$\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}e^{-\lambda}=e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}=e^{-\lambda}\cdot e^{\lambda}=1.$$

За $\sigma$-алгебру событий $\mathscr{F}$ принимаем все подмножества множества $\Omega$. Вероятность зададим формулой (1.6.5). Случайная величина $\xi=\xi(m)=m$ имеет пуассоновское распределение. Аналогично можно определить величины с биномиальным, геометрическим и гипергеометрическим распределениями.

Случайные величины с указанными выше распределениями могут появляться в различных задачах. Например, в примере 1 (п. 6.1 § 6 гл. 1) естественно ввести случайную величину, равную числу белых шаров в выборке. Эта величина имеет гипергеометрическое распределение (см. (1.6.3)). В § 2 гл. 3 для схемы Бернулли была введена случайная величина $\mu_n$, равная числу успехов в первых $n$ испытаниях. Она имеет биномиальное распределение (см. (3.2.6)).

В вероятностном пространстве, соответствующем бесконечной последовательности независимых испытаний с двумя исходами в каждом (см. § 4 гл. 3), введем случайную величину $\tau$, равную числу испытаний до первого успеха включительно. Тогда событие $(\tau=k)$ определяется результатами первых $k$ испытаний: должен подряд $k-1$ раз произойти «неуспех» и на $k$-й раз «успех». В обозначениях § 4 гл. 3 имеем

$$(\tau=k)=B_k\,(\underbrace{0,\,0,\,\ldots,\,0,}_{k-1\text{ раз}}\,1)$$

и, следовательно, по формуле (3.4.1)

$$\mathsf{P}(\tau=k)=q^{k-1}p.$$

*Таким образом*, величина $\tau$ имеет геометрическое распределение. В рассмотренном примере вероятностное пространство не является дискретным, так как множество всех последовательностей из 0 и 1 — несчетно.

Приведем пример дискретной случайной величины, заданной на абсолютно непрерывном вероятностном пространстве (см. п. 6.4 § 6 гл. 1). Пусть $n=1$; $\Omega=\{u:\ -\infty<u<\infty\}$, $\pi(u)=1$, если $u\in[0,\ 1]$, $\pi(u)=0$ в остальных случаях. Положим $\xi=\xi(u)=1$, если $u\geqslant\frac{1}{2}$, и $\xi=\xi(u)=0$ если $u<\frac{1}{2}$. Очевидно, что величина $\xi$ является дискретной и

$$\mathsf{P}(\xi=1)=\mathsf{P}\left(\left\{u:\ u\geqslant\frac{1}{2}\right\}\right)=\int\limits_{1/2}^{\infty}\pi(u)\,du=\int\limits_{1/2}^{1}du=\frac{1}{2},$$

$$\mathsf{P}(\xi=0)=\mathsf{P}\left(\left\{u:\ u<\frac{1}{2}\right\}\right)=\int\limits_{-\infty}^{1/2}\pi(u)\,du=\int\limits_{0}^{1/2}du=\frac{1}{2}.$$

## § 3. Совместные распределения нескольких случайных величин

Пусть на вероятностном пространстве $(\Omega,\ \mathfrak{F},\ \mathsf{P})$ заданы случайные величины

$$\xi_1=\xi_1(\omega),\ \ \xi_2=\xi_2(\omega),\ \ldots,\ \ \xi_n=\xi_n(\omega),\qquad \omega\in\Omega.$$

Каждому $\omega$ эти величины ставят в соответствие $n$-мерный вектор. Функцию от переменных $x_1, x_2, \ldots, x_n$:

$$F_{\xi_1 \ldots \xi_n}(x_1, x_2, \ldots, x_n) = \mathsf{P}(\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n),$$

назовем *многомерной функцией распределения* случайного вектора $(\xi_1, \xi_2, \ldots, \xi_n)$.

Так же, как в одномерном случае, проверяется, что $F_{\xi_1 \ldots \xi_n}(x_1, x_2, \ldots, x_n)$ монотонна по каждому аргументу,

$$\lim_{\substack{x_1 \to +\infty \\ \cdots\cdots \\ x_n \to +\infty}} F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = 1, \qquad (3.1)$$

$$\lim_{x_k \to -\infty} F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = 0,$$

$$\lim_{x_n \to +\infty} F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = F_{\xi_1 \ldots \xi_{n-1}}(x_1, \ldots, x_{n-1}).$$

Аналогичное свойство выполняется при переходе к пределу по любому аргументу.

Случайный вектор $(\xi_1, \ldots, \xi_n)$ назовем *вектором дискретного типа*, если существует конечное или счетное множество точек $(x_{k1}, x_{k2}, \ldots, x_{kn})$, $k = 1, 2, \ldots$ (без предельных точек), таких, что

$$\mathsf{P}(\xi_1 = x_{k1}, \ \xi_2 = x_{k2}, \ \ldots, \ \xi_n = x_{kn}) = p_{x_{k1} x_{k2} \ldots x_{kn}} \qquad (3.2)$$

и

$$\sum_{(x_{k1}, \ldots, x_{kn})} p_{x_{k1}, \ldots, x_{kn}} = 1.$$

Случайный вектор дискретного типа можно определить на дискретном вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$, в котором $\Omega$ является множеством значений данного вектора.

Случайный вектор $(\xi_1, \ldots, \xi_n)$ назовем *вектором абсолютно непрерывного типа*, если существует функция $p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)$ такая, что при любых $x_1, x_2, \ldots, x_n$

$$F_{\xi_1 \xi_2 \ldots \xi_n}(x_1, x_2, \ldots, x_n) = \\ = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} p_{\xi_1 \xi_2 \ldots \xi_n}(u_1, \ldots, u_n)\, du_1 \ldots du_n.$$

Функция $p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)$ называется *плотностью распределения вероятностей* случайного вектора $(\xi_1, \ldots, \xi_n)$. Для любого квадрируемого множества $B$ в $n$-мерном пространстве имеем

$$\mathsf{P}((\xi_1, \ldots, \xi_n) \in B) = \iint\limits_B \ldots \int p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n. \tag{3.3}$$

Интеграл от плотности по всему $n$-мерному пространству равен 1. Случайный вектор с плотностью распределения $p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)$ можно задать следующим образом. В определении абсолютно непрерывного вероятностного пространства (см. п. 6.4 § 6 гл. I) положим $\pi(u_1, \ldots, u_n) = p_{\xi_1 \ldots \xi_n}(u_1, \ldots, u_n)$. Нетрудно проверить, что случайный вектор

$$\xi_1 = \xi_1(u_1, \ldots, u_n) = u_1, \ \ldots, \ \xi_n = \xi_n(u_1, \ldots, u_n) = u_n$$

имеет плотность распределения $p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)$.

В дальнейшем, когда мы будем говорить о случайных векторах дискретного или абсолютно непрерывного типа без указания вероятностного пространства, будет подразумеваться, что они определены на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$, в котором $\Omega$ состоит из множества значений рассматриваемого дискретного вектора или является евклидовом пространством в случае абсолютно непрерывного вектора.

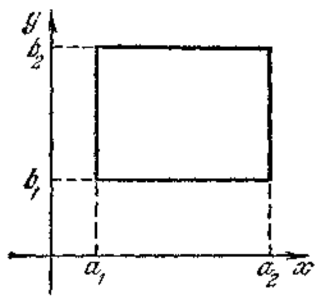

Рис. 7.

Ниже мы будем рассматривать случай $n = 2$. Результаты, полученные для $n = 2$, легко переносятся на общий случай. В § 1 для одномерной случайной величины была получена формула (1.4), по которой можно вероятность попадания значения случайной величины в полуинтервал выразить через функцию распределения. В двумерном случае найдем аналогичную формулу для вероятности попадания пары случайных величин в прямоугольник (рис. 7). Пусть

$$F(x, y) = F_{\xi\eta}(x, y) = \mathsf{P}(\xi < x, \eta < y).$$

Положим $A = (a_1 \leqslant \xi < a_2, b_1 \leqslant \eta < b_2)$. Тогда

$$(\xi < a_2, \eta < b_2) = A + (\xi < a_1, \eta < b_2) + (\xi < a_2, \eta < b_1). \tag{3.4}$$

Так как

$$(\xi < a_1, \eta < b_2) \cdot (\xi < a_2, \eta < b_1) = (\xi < a_1, \eta < b_1),$$

то по формуле (1.5.4)

$$\mathsf{P}((\xi < a_1, \eta < b_2) + (\xi < a_2, \eta < b_1)) =$$
$$= \mathsf{P}(\xi < a_1, \eta < b_2) + \mathsf{P}(\xi < a_2, \eta < b_1) - \mathsf{P}(\xi < a_1, \eta < b_1) =$$
$$= F(a_1, b_2) + F(a_2, b_1) - F(a_1, b_1). \tag{3.5}$$

Учитывая, что в (3.4) $A$ несовместно с другими слагаемыми, из (3.4) и (3.5) получим

$$\mathsf{P}(A) = \mathsf{P}(a_1 \leqslant \xi < a_2, b_1 \leqslant \eta < b_2) = F(a_2, b_2) +$$
$$+ F(b_1, a_1) - F(a_1, b_2) - F(a_2, b_1). \tag{3.6}$$

Если $(\xi, \eta)$ — абсолютно непрерывный вектор, то из (3.6) следует, что

$$\mathsf{P}(a_1 \leqslant \xi < a_2, b_1 \leqslant \eta < b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} p_{\xi, \eta}(x, y)\, dx\, dy. \tag{3.7}$$

Можно показать, что по вероятностям для прямоугольников однозначно определяется вероятность на квадрируемых множествах и из формулы (3.7) следует формула (3.3) при $n = 2$.

По двумерной функции распределения можно найти одномерные функции распределения. Так как $(\eta < +\infty) = \Omega$, то

$$(\xi < x) = (\xi < x, \eta < +\infty),$$

и, следовательно, с учетом свойства (3.1) получим

$$F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}(\xi < x, \eta < +\infty) = F_{\xi, \eta}(x, +\infty).$$

Таким образом,

$$F_\xi(x)=\lim_{y\to\infty}F_{\xi,\eta}(x,y)=F_{\xi,\eta}(x,\infty),$$
$$F_\eta(y)=\lim_{x\to\infty}F_{\xi,\eta}(x,y)=F_{\xi,\eta}(\infty,y). \tag{3.8}$$

Если $(\xi,\eta)$ — абсолютно непрерывный вектор, то

$$F_\xi(x)=\int\limits_{-\infty}^{x}\left(\int\limits_{-\infty}^{\infty}p_{\xi,\eta}(u,v)\,dv\right)du=\int\limits_{-\infty}^{x}p_\xi(u)\,du,$$

где функция

$$p_\xi(x)=\int\limits_{-\infty}^{\infty}p_{\xi,\eta}(x,v)\,dv \tag{3.9}$$

является плотностью распределения величины $\xi$. Аналогично находим

$$p_\eta(y)=\int\limits_{-\infty}^{\infty}p_{\xi,\eta}(u,y)\,du. \tag{3.10}$$

Пусть вектор $(\xi,\eta)$ дискретного типа:

$$\mathsf{P}(\xi=x_i,\ \eta=y_j)=p_{ij},\qquad i,j=1,2,\ldots, \tag{3.11}$$

и $\sum\limits_{i,j=1}^{\infty}p_{ij}=1$. Найдем одномерные распределения, соответствующие (3.11). Имеем

$$\mathsf{P}(\xi=x_i)=\sum_{j=1}^{\infty}\mathsf{P}(\xi=x_i,\ \eta=y_j)=\sum_{j=1}^{\infty}p_{ij}.$$

Пусть

$$p_{i\cdot}=\sum_{j=1}^{\infty}p_{ij},\qquad p_{\cdot j}=\sum_{i=1}^{\infty}p_{ij}. \tag{3.12}$$

В этих обозначениях получим

$$\mathsf{P}(\xi=x_i)=p_{i\cdot},\qquad \mathsf{P}(\eta=y_j)=p_{\cdot j}. \tag{3.13}$$

Мы получили формулы, позволяющие по совместному распределению двумерной величины находить одномерные распределения.

Рассмотрим два примера, показывающих, что по одномерным распределениям (3.13) без дополнительной информации нельзя восстановить совместное распределение случайных величин (3.11).

Пример 1. Пусть $\Omega=\{(1, 1), (-1, 1), (1, -1), (-1, -1)\}$ и все элементарные события равновероятны. Положим

$$\xi=\xi(i, j)=i, \quad \eta=\eta(i, j)=j, \qquad i, j=-1, 1. \tag{3.14}$$

Очевидно, при любой паре

$$\mathsf{P}(\xi=i, \eta=j)=\frac{1}{4} \tag{3.15}$$

и

$$\mathsf{P}(\xi=i)=\mathsf{P}(\xi=i, \eta=-1)+\mathsf{P}(\xi=i, \eta=1)=\frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$
$$i=-1, 1.$$

Аналогично получим

$$\mathsf{P}(\eta=j)=\frac{1}{2}, \qquad j=-1, 1.$$

Пример 2. Пусть $\Omega$, $\xi$ и $\eta$ такие же, как в примере 1. Положим

$$\mathsf{P}((1, 1))=\mathsf{P}((-1, -1))=\frac{1}{2}, \quad \mathsf{P}((-1, 1))=\mathsf{P}((1, -1))=0.$$

Тогда, например,

$$\mathsf{P}(\xi=-1, \eta=1)=0,$$

и, следовательно, совместное распределение $\xi$, $\eta$ не совпадает с (3.15). Одномерные распределения

$$\mathsf{P}(\xi=-1)=\mathsf{P}(\xi=-1, \eta=-1)=\frac{1}{2},$$
$$\mathsf{P}(\xi=1)=\mathsf{P}(\xi=1, \eta=1)=\frac{1}{2}$$

и

$$\mathsf{P}(\eta=-1)=\mathsf{P}(\eta=1)=\frac{1}{2}$$

совпадают с соответствующими распределениями предыдущего примера.

*Равномерным распределением* в плоской ограниченной области $G$ будем называть распределение с плотностью

$$p(x, y) = \begin{cases} \frac{1}{\text{пл. } G}, & \text{если} \quad (x, y) \in G, \\ 0, & \text{если} \quad (x, y) \notin G. \end{cases}$$

В следующем параграфе будет дано определение многомерного нормального распределения.

## § 4. Независимость случайных величин

Пусть $(\Omega, \mathfrak{F}, \mathsf{P})$ — произвольное вероятностное пространство. Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ называются *независимыми в совокупности* или просто *независимыми*, если при любых действительных $x_1, x_2, \ldots, x_n$

$$F_{\xi_1 \xi_2 \ldots \xi_n}(x_1, x_2, \ldots, x_n) = F_{\xi_1}(x_1) F_{\xi_2}(x_2) \cdot \ldots \cdot F_{\xi_n}(x_n). \tag{4.1}$$

Часто удобнее использовать следующее эквивалентное определение независимости: случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ *независимы*, если при любых борелевских множествах *) $B_1, B_2, \ldots, B_n$ имеет место равенство

$$\mathsf{P}(\xi_1 \in B_1, \xi_2 \in B_2, \ldots, \xi_n \in B_n) = \\ = \mathsf{P}(\xi_1 \in B_1) \mathsf{P}(\xi_2 \in B_2) \ldots \mathsf{P}(\xi_n \in B_n). \tag{4.2}$$

Так как бесконечные полуинтервалы являются борелевскими множествами, то (4.1) следует из (4.2), если положить $B_k = (-\infty, x_k)$, $k = 1, 2, \ldots, n$. Покажем, что при $n = 2$ из (4.1) следует (4.2) для любых полуинтервалов: $B_k = [a_k, b_k)$, $k = 1, 2$.

Теорема 4.1. *Если при любых $x, y$*

$$F_{\xi, \eta}(x, y) = F_{\xi}(x) F_{\eta}(y), \tag{4.3}$$

*то при любых $a_k < b_k$, $k = 1, 2$,*

$$\mathsf{P}(a_1 \leqslant \xi < b_1, a_2 \leqslant \eta < b_2) = \\ = \mathsf{P}(a_1 \leqslant \xi < b_1) \mathsf{P}(b_2 \leqslant \eta < b_2). \tag{4.4}$$

---

*) Читатель, пожелавший опустить определение борелевского множества, приведенное в гл. I, может и в дальнейшем выражения «борелевское множество $B$» заменять выражениями «множество $B$, для которого определена вероятность события $\xi \in B$».

Доказательство. Из формул (3.6) и (4.3) следует, что

$$\mathsf{P}(a_1 \leqslant \xi < b_1,\ a_2 \leqslant \eta < b_2) = \\ = F_\xi(a_2) F_\eta(b_2) + F_\xi(a_1) F_\eta(b_1) - \\ - F_\xi(a_1) F_\eta(b_2) - F_\xi(a_2) F_\eta(b_1) = \\ = (F_\eta(b_2) - F_\eta(a_2)) (F_\xi(b_1) - F_\xi(a_1)).$$

Отсюда получаем (4.4), так как

$$\mathsf{P}(a_1 \leqslant \xi < b_1) = F_\xi(b_1) - F_\xi(a_1),$$

$$\mathsf{P}(a_2 \leqslant \xi < b_2) = F_\eta(b_2) - F_\eta(a_2).$$

Теорема доказана.

Так как борелевские множества порождаются полуинтервалами, то из (4.4) следует (4.2) при $n=2$. Доказательство этого утверждения не будет приведено.

Условие независимости (4.4) удобно использовать для установления условий независимости в дискретном и абсолютно непрерывном случаях.

Теорема 4.2. *Пусть распределение величин $\xi$, $\eta$ задается формулой*

$$\mathsf{P}(\xi = x_i,\ \eta = y_j) = p_{ij} \geqslant 0, \qquad \sum_{i,\, j=1}^{\infty} p_{ij} = 1.$$

*Случайные величины $\xi$, $\eta$ независимы тогда и только тогда, когда при любых $i$, $j$*

$$p_{ij} = p_{i\cdot} \cdot p_{\cdot j}, \tag{4.5}$$

*где*

$$\mathsf{P}(\xi = x_i) = p_{i\cdot} = \sum_{j=1}^{\infty} p_{ij}, \qquad \mathsf{P}(\eta = y_j) = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij}.$$

Доказательство. Необходимость. Пусть для величин $\xi$, $\eta$ выполняется условие (4.4). Так как множество значений пары случайных величин $(x_i, y_j)$ по определению не имеет предельных точек, то каждую фиксированную точку $(x_i, y_j)$ можно заключить в прямоугольник $\{(x, y)\colon a_i < x < a_i',\ b_j < y < b_j'\}$ такой, что $x_i \in [a_i, a_i']$, $y_j \in [b_j, b_j']$, а другие значения случайных величин $\xi$, $\eta$ не входят в эти отрезки.

Тогда

$$\mathsf{P}(a_i < \xi < a_i', b_j < \eta < b_j') = \mathsf{P}(\xi = x_i, \eta = y_j) = p_{ij},$$
$$\mathsf{P}(a_i < \xi < a_i') = p_{i\cdot}, \quad \mathsf{P}(b_j < \eta < b_j') = p_{\cdot j}.$$

Отсюда и из (4.4) следует (4.5).

Достаточность. Пусть выполнено условие (4.5). Для произвольных полуинтервалов $[a_1, b_1)$, $[a_2, b_2)$ нужно доказать (4.4). В $[a_1, b_1)$ попадут некоторые числа из множества $\{x_1, x_2, \ldots, x_i, \ldots\}$. Обозначим

$$X = (i_1, i_2, \ldots, i_s) = \{i\colon\ x_i \in [a_1, b_1)\}.$$

В полуинтервал $[a_2, b_2)$ попадут числа из множества $\{y_1, y_2, \ldots, y_j, \ldots\}$. Положим

$$Y = (j_1, j_2, \ldots, j_t) = \{j\colon\ y_j \in [a_2, b_2)\}.$$

Используя (4.5), получим

$$\mathsf{P}(a_1 \leqslant \xi < b_1, a_2 \leqslant \eta < b_2) = \\ = \sum_{i \in X} \sum_{j \in Y} p_{ij} = \sum_{i \in X} \sum_{j \in Y} p_{i\cdot} p_{\cdot j}. \tag{4.6}$$

Так как

$$\sum_{i \in X} \sum_{j \in Y} p_{i\cdot} p_{\cdot j} = \Big(\sum_{i \in X} p_{i\cdot}\Big)\Big(\sum_{j \in Y} p_{\cdot j}\Big) \tag{4.7}$$

и

$$\begin{gathered} \mathsf{P}(a_1 \leqslant \xi < b_1) = \sum_{i \in X} p_{i\cdot}, \\ \mathsf{P}(a_2 \leqslant \eta < b_2) = \sum_{j \in Y} p_{\cdot j}, \end{gathered} \tag{4.8}$$

то из (4.6), (4.7), (4.8) следует (4.4). Теорема доказана.

В примере 1 из § 3 случайные величины $\xi$, $\eta$ независимы, так как равенства

$$\mathsf{P}(\xi = i, \eta = j) = \mathsf{P}(\xi = i)\,\mathsf{P}(\eta = j)$$

выполняются при любых $i$, $j$. Величины $\xi$, $\eta$ в примере 2 того же параграфа зависимы, так как

$$0 = \mathsf{P}(\xi = 1, \eta = -1) \neq \mathsf{P}(\xi = 1)\,\mathsf{P}(\eta = -1) = \frac{1}{4}.$$

Если для дискретных величин заданы одномерные распределения и еще известно, что величины независимы, то по формулам (4.5) можно найти совместное распределение.

Теорема 4.3. *Пусть $p_{\xi,\eta}(x, y)$ — плотность распределения случайных величин $\xi$, $\eta$. Случайные величины $\xi$, $\eta$ независимы тогда и только тогда, когда во всех точках непрерывности функций $p_{\xi,\eta}(x, y)$, $p_\xi(x)$, $p_\eta(y)$ имеем*

$$p_{\xi,\eta}(x, y) = p_\xi(x)\, p_\eta(y). \tag{4.9}$$

Доказательство. Необходимость. Пусть выполнено условие (4.4). Выражая левую и правую части (4.4) через плотности распределения вероятностей, при любом $B = \{(x, y)\colon a_1 \leqslant x < b_1,\ a_2 \leqslant y < b_2\}$ получим

$$\iint\limits_B p_{\xi,\eta}(x, y)\, dx\, dy = \int\limits_{a_1}^{b_1} p_\xi(x)\, dx \cdot \int\limits_{a_2}^{b_2} p_\eta(y)\, dy. \tag{4.10}$$

Так как

$$\int\limits_{a_1}^{b_1} p_\xi(x)\, dx \int\limits_{a_2}^{b_2} p_\eta(y)\, dy = \iint\limits_B p^\xi(x)\, p_\eta(y)\, dx\, dy,$$

то из (4.10) найдем

$$\iint\limits_B [p_{\xi,\eta}(x, y) - p_\xi(x)\, p_\eta(y)]\, dx\, dy = 0. \tag{4.11}$$

Пусть $(x, y)$, $x$, $y$ — точки непрерывности функций $p_{\xi,\eta}(x, y)$, $p_\xi(x)$, $p_\eta(y)$. Нужно показать, что в точке $(x, y)$ имеет место (4.9). Пусть это не так. Тогда в силу непрерывности подынтегральной функции в (4.11) найдется окрестность точки $(x, y)$, в которой функция отлична от нуля и сохраняет знак. Выбрав прямоугольник $B$, целиком лежащий внутри этой окрестности, получим противоречие с (4.11).

Достаточность. Пусть выполнено (4.9). Тогда при любом $B$ верно (4.11) и, следовательно, имеет место (4.10) и (4.4). Теорема доказана.

Условия независимости двух случайных величин дискретного и непрерывного типа ((4.5) и (4.10)) естественно переносятся и на случай нескольких величин. Так, например, случайные величины $\xi_1$, $\xi_2$, ..., $\xi_n$

абсолютно непрерывного типа независимы тогда и только тогда, когда

$$p_{\xi_1\xi_2\ldots\xi_n}(x_1, \ldots, x_n) = p_{\xi_1}(x_1) \ldots p_{\xi_n}(x_n) \quad (4.12)$$

в точках непрерывности рассматриваемых функций. Положим

$$p_{\xi_1\xi_2\ldots\xi_n}(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x-a_k)^2}{2\sigma_k^2}}. \quad (4.13)$$

Так как $p_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n) > 0$ и

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} p_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n =$$

$$= \prod_{k=1}^{n} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x-a_k)^2}{2\sigma_k}} dx_k = 1,$$

то функция $p_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n)$ является плотностью распределения. Если (4.13) проинтегрировать по всем переменным от $-\infty$ до $+\infty$, кроме $x_i$, то получится плотность распределения $\xi_i$:

$$p_{\xi_i}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-a_i)^2}{2\sigma_i^2}}.$$

Таким образом, случайные величины с плотностью распределения (4.13) независимы и каждая из них имеет нормальное распределение.

Многомерным нормальным распределением называется распределение с плотностью

$$p_{\xi_1\ldots\xi_n}(x_1, \ldots, x_n) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} e^{-\frac{1}{2} Q(x_1, \ldots, x_n)}, \quad (4.14)$$

где $Q = \sum_{i,j=1}^{n} a_{ij}(x_i - a_i)(x_j - a_j)$ — положительно определенная квадратичная форма, $|A|$ — определитель матрицы $A = \|a_{ij}\|$. Плотность распределения (4.13) является частным случаем (4.14). Случай $n=2$ рассмотрен в приложении 2.

Условие (4.5) обобщается на случай $n$ дискретных величин $\xi_1, \xi_2, \ldots, \xi_n$ следующим образом. Случайные величины с законом распределения (3.2) независимы тогда и только тогда, когда

$$\mathsf{P}(\xi_1 = x_{k-1}, \ldots, \xi_n = x_{kn}) = \\ = \mathsf{P}(\xi_1 = x_{k1}) \cdot \ldots \cdot \mathsf{P}(\xi_n = x_{kn}) \qquad (4.15)$$

при любых $x_{k1}, \ldots, x_{kn}$.

Рассмотрим пример, связанный со схемой Бернулли. В схеме Бернулли с $n$ испытаниями элементарными событиями являются последовательности вида (3.2.3). Пусть $\xi_k = 1$, $k = 1, 2, \ldots, n$, если в последовательности (3.2.3) на $k$-м месте был успех (У) и $\xi_k = 0$ в противном случае. Событие $(\xi_1 = \varepsilon_1, \xi_2 = \varepsilon_2, \ldots \ldots, \xi_n = \varepsilon_n)$, где $\varepsilon_k$ принимают значение 0 или 1, однозначно определяет элементарное событие и, следовательно,

$$\mathsf{P}(\xi_1 = \varepsilon_1, \xi_2 = \varepsilon_2, \ldots, \xi_n = \varepsilon_n) = \\ = p^{\varepsilon_1} q^{1-\varepsilon_1} p^{\varepsilon_2} q^{1-\varepsilon_2} \ldots p^{\varepsilon_n} q^{1-\varepsilon_n}, \qquad (4.16)$$

так как $p^{\varepsilon_k} q^{1-\varepsilon_k} = p$ при $\varepsilon_k = 1$, $p^{\varepsilon_k} q^{1-\varepsilon_k} = q$ при $\varepsilon_k = 0$. Заметим, что

$$\sum_{\varepsilon_l = 0}^{1} p^{\varepsilon_l} q^{1-\varepsilon_l} = p + q = 1.$$

Учитывая это замечание из (4.16), легко получить одномерные распределения. Например,

$$\mathsf{P}(\xi_1 = \varepsilon_1) = \\ = \sum_{\varepsilon_2=0}^{1} \sum_{\varepsilon_3=0}^{1} \ldots \sum_{\varepsilon_n=0}^{1} \mathsf{P}(\xi_1 = \varepsilon_1, \xi_2 = \varepsilon_2, \ldots, \xi_n = \varepsilon_n) = \\ = p^{\varepsilon_1} q^{1-\varepsilon_1} \sum_{\varepsilon_2=0}^{1} \ldots \sum_{\varepsilon_n=0}^{1} p^{\varepsilon_2} q^{1-\varepsilon_2} \ldots p^{\varepsilon_n} q^{1-\varepsilon_n} = p^{\varepsilon_1} q^{1-\varepsilon_1}.$$

Аналогично находим

$$\mathsf{P}(\xi_k = \varepsilon_k) = p^{\varepsilon_k} q^{1-\varepsilon_k}, \qquad k = 1, 2, \ldots, n. \qquad (4.17)$$

Используя (4.16), (4.17) для проверки (4.15), получим, что величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы. В § 2 гл. 3

была введена случайная величина $\mu_n$, равная числу успехов в $n$ испытаниях. Легко проверить, что

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n. \tag{4.18}$$

Это представление в дальнейшем будет часто использоваться.

## § 5. Функции от случайных величин

Пусть $(\Omega, \mathfrak{F}, \mathsf{P})$ — произвольное вероятностное пространство и $\xi = \xi(\omega)$, $\omega \in \Omega$, — некоторая случайная величина. Часто приходится рассматривать функции от случайных величин. Суперпозиция действительной функции $\xi$, заданной на $\Omega$, и функции $\varphi(x)$, заданной на действительной прямой, является функцией $\eta = \varphi[\xi(\omega)] = \eta(\omega)$, заданной на $\Omega$. Для дискретных вероятностных пространств функция $\eta$ является случайной величиной, так как никаких ограничений на функцию $\eta(\omega)$ не налагается. Для произвольных вероятностных пространств требуется, чтобы при любом $x$

$$(\eta < x) \in \mathfrak{F}. \tag{5.1}$$

При произвольной функции $\varphi$ условие (5.1) может оказаться не выполненным. Однако можно показать, что (5.1) имеет место, если при любом борелевском множестве $B$ множество $\{u: \varphi(u) \in B\}$ является борелевским множеством на числовой прямой (т. е. полный прообраз $\varphi^{-1}(B)$ множества $B$ при отображении, осуществляемом функцией $\varphi$, является борелевским множеством). Таким свойством обладают, например, непрерывные функции, функции с конечным числом точек разрыва. Таким образом, в практически важных случаях условие (5.1) оказывается выполненным.

Справедлива следующая теорема.

Теорема 5.1. *Если случайные величины $\xi_1$ и $\xi_2$ независимы, то независимы и случайные величины $\eta_1 = \varphi_1(\xi_1)$, $\eta_2 = \varphi_2(\xi_2)$.*

Доказательство. Пусть $B_1$ и $B_2$ — произвольные борелевские множества *). Тогда

$$\mathsf{P}(\eta_1 \in B_1,\ \eta_2 \in B_2) = \mathsf{P}(\varphi_1(\xi_1) \in B_1,\ \varphi_2(\xi_2) \in B_2) = \\ = \mathsf{P}(\xi_1 \in \varphi^{-1}(B_1),\ \xi_2 \in \varphi^{-1}(B_2)). \tag{5.2}$$

---

*) См. сноску на стр. 92.

Так как множества $\varphi_1^{-1}(B_1)$ и $\varphi_2^{-1}(B_2)$ борелевские, а случайные величины $\xi_1$ и $\xi_2$ независимы, то согласно определению независимости в форме (4.2) имеем

$$\mathsf{P}(\xi_1 \in \varphi_1^{-1}(B_1),\ \xi_2 \in \varphi_2^{-1}(B_2)) = \\ = \mathsf{P}(\xi_1 \in \varphi_1^{-1}(B_1))\,\mathsf{P}(\xi_2 \in \varphi_2^{-1}(B_2)). \qquad (5.3)$$

Очевидно, что

$$(\xi_i \in \varphi_i^{-1}(B_i)) = (\varphi_i(\xi_i) \in B_i) = (\eta_i \in B_i).$$

Отсюда и из равенств (5.2) и (5.3) получаем

$$\mathsf{P}(\eta_1 \in B_1,\ \eta_2 \in B_2) = \mathsf{P}(\eta_1 \in B_1)\,\mathsf{P}(\eta_2 \in B_2)$$

для любых борелевских множеств $B_1$, $B_2$. Теорема доказана.

Если на исходном вероятностном пространстве задано несколько случайных величин $(\xi_1, \xi_2, \ldots, \xi_n)$, то сложная функция $\eta = \varphi(\xi_1, \ldots, \xi_n)$ также является случайной величиной для достаточно «хороших» функций $\varphi(u_1, u_2, \ldots, u_n)$; например, достаточно потребовать, чтобы $\varphi(u_1, u_2, \ldots, u_n)$ была непрерывной. Можно доказать следующее утверждение, обобщающее теорему 5.1: если $\xi_1, \xi_2, \ldots, \xi_n, \eta_1, \eta_2, \ldots, \eta_m$ — независимые случайные величины, то случайные величины

$$\zeta_1 = \varphi_1(\xi_1, \ldots, \xi_n), \qquad \zeta_2 = \varphi_2(\eta_1, \ldots, \eta_m)$$

независимы.

*Закон распределения* новых случайных величин, являющихся функциями от старых, очевидно, определен, так как они заданы на том же вероятностном пространстве. Можно найти функцию распределения новой случайной величины $\eta = \varphi(\xi)$. Для этого достаточно знать только функцию распределения $\xi$. Действительно,

$$F_\eta(x) = \mathsf{P}(\varphi(\xi) < x) = \mathsf{P}(\xi \in \varphi^{-1}(-\infty, x)). \qquad (5.4)$$

Так как множество $\varphi^{-1}(-\infty, x)$ борелевское, то вероятность в правой части (5.4) может быть вычислена по $F_\xi(x)$. Аналогично находится функция распределения, когда $\eta = \varphi(\xi_1, \ldots, \xi_n)$.

Пример 1. Найдем закон распределения линейной функции от нормальной случайной величины. Пусть случайная величина $\xi$ имеет нормальное распределение с параметрами $(a, \sigma)$. Покажем, что случайная величина $\eta = A\xi + B$, $A \neq 0$, имеет нормальное распределение с параметрами $(Aa + B, |A|\sigma)$.

Плотность распределения $p_\xi(u)$ величины $\xi$ задается формулой (2.3). Пусть сначала $A > 0$. Тогда

$$F_\eta(x) = \mathsf{P}(A\xi + B < x) = \mathsf{P}\left(\xi < \frac{x-B}{A}\right) = \int_{-\infty}^{\frac{x-B}{A}} p_\xi(u)\, du.$$

Отсюда, так как $p_\xi(x)$ и $F_\eta(x)$ непрерывны при любом $x$, следует, что существует

$$p_\eta(x) = F'_\eta(x) = p_\xi\left(\frac{x-B}{A}\right) \cdot \left(\frac{x-B}{A}\right)'_x = \frac{1}{A} p_\xi\left(\frac{x-B}{A}\right). \quad (5.5)$$

Если $A < 0$, то

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

$$= 1 - \mathsf{P}\left(\xi \leqslant \frac{x-B}{A}\right) = 1 - \int_{-\infty}^{\frac{x-B}{A}} p_\xi(u)\, du$$

и

$$p_\eta(x) = -\frac{1}{A} p_\xi\left(\frac{x-B}{A}\right).$$

Объединяя эту формулу с (5.5) и используя (2.3), получим

$$p_\eta(x) = \frac{1}{|A|} p_\xi\left(\frac{x-B}{A}\right) =$$

$$= \frac{1}{|A|} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{x-B}{A} - a\right)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-a_1)^2}{2\sigma_1^2}},$$

где

$$a_1 = Aa + B, \quad \sigma_1 = |A|\sigma.$$

Теорема 5.2. *Если случайные величины $\xi_1$ и $\xi_2$ абсолютно непрерывны и независимы, то случайная вели-*

*чина $\xi_1+\xi_2$ тоже абсолютно непрерывна и ее плотность распределения определяется по формуле*

$$p_{\xi_1+\xi_2}(x)=\int_{-\infty}^{x} p_{\xi_1}(u)\,p_{\xi_2}(x-u)\,du. \qquad (5.6)$$

Доказательство. Найдем сначала

$$F_{\xi_1+\xi_2}(x)=\mathsf{P}(\xi_1+\xi_2<x)=\mathsf{P}((\xi_1,\xi_2)\in D_x),$$

где $D_x=\{(u,v):\ u+v<x\}$. Так как по условию теоремы величины $\xi_1$ и $\xi_2$ независимы, то их плотность совместного распределения согласно (4.9) равна

$$p_{\xi_1\xi_2}(u,v)=p_{\xi_1}(u)\,p_{\xi_2}(v),$$

и, следовательно, по формуле (3.3) найдем

$$F_{\xi_1+\xi_2}(x)=\iint_{D_x} p_{\xi_1}(u)\,p_{\xi_2}(v)\,du\,dv=$$
$$=\int_{-\infty}^{\infty} p_{\xi_1}(u)\left(\int_{-\infty}^{-u+x} p_{\xi_2}(v)\,dv\right)du.$$

Отсюда, полагая $v=z-u$, получим

$$F_{\xi_1+\xi_2}(x)=\int_{-\infty}^{\infty} p_{\xi_1}(u)\,du\int_{-\infty}^{x} p_{\xi_2}(z-u)\,dz=$$
$$=\int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} p_{\xi_1}(u)\,p_{\xi_2}(z-u)\,du\right)dz=\int_{-\infty}^{x} p_{\xi_1+\xi_2}(z)\,dz,$$

где

$$p_{\xi_1+\xi_2}(x)=\int_{-\infty}^{\infty} p_{\xi_1}(u)\,p_{\xi_2}(x-u)\,du.$$

Теорема доказана.

Отметим, что вместе с (5.6) верна формула

$$p_{\xi_1+\xi_2}(x)=\int_{-\infty}^{x} p_{\xi_1}(x-u)\,p_{\xi_2}(u)\,du.$$

Используя формулу (5.6), можно показать, что сумма независимых нормально распределенных случайных величин имеет нормальное распределение. В гл. 7 этот факт будет установлен при помощи характеристических функций.

Рассмотрим следующий пример.

Пример 2. Пусть случайные величины $\xi_1$ и $\xi_2$ независимы и имеют распределения соответственно равномерное (см. (2.5)) и показательное (см. (2.4)):

$$p_{\xi_1}(x)=\begin{cases}1, & \text{если } x\in[0,\ 1],\\ 0, & \text{если } x\notin[0,\ 1],\end{cases}$$

$$p_{\xi_2}(x)=\begin{cases}e^{-x}, & \text{если } x\geqslant 0,\\ 0, & \text{если } x<0.\end{cases}$$

Требуется найти $p_{\xi_1+\xi_2}(x)$.

По формуле (5.6) получим

$$p_{\xi_1+\xi_2}(x)=\int_{-\infty}^{\infty} p_{\xi_1}(u)\,p_{\xi_2}(x-u)\,du=\int_0^1 p_{\xi_2}(x-u)\,du.$$

Подынтегральная функция $p_{\xi_2}(x-u)>0$, если $0<u<x$, и $p_{\xi_2}(x-u)=0$ в остальных случаях. Таким образом, при $0<x\leqslant 1$

$$p_{\xi_1+\xi_2}(x)=\int_0^x e^{-(x-u)}\,du=1-e^{-x}$$

и при $x>1$

$$p_{\xi_1+\xi_2}(x)=\int_0^1 e^{-(x-u)}\,du=(e-1)\,e^{-x}.$$

Окончательно получим

$$p_{\xi_1+\xi_2}(x)=\begin{cases}0, & \text{если } x\leqslant 0,\\ 1-e^{-x}, & \text{если } 0<x\leqslant 1,\\ (e-1)\,e^{-x}, & \text{если } x>1.\end{cases}$$

### Задачи к главе 4

1. В задаче 20 гл. 1 доказать, что $\xi_1$, $\xi_2$ независимы.

2. Случайная величина $\xi$ распределена нормально с параметрами $a=0$, $\sigma=1$. а) Найти плотность распределения $p_{\xi^2}(x)$.

б) Найти плотность распределения величины $\eta=e^{\xi}$ (логарифмически — нормальное распределение).

3. Плотность распределения $\xi$ задана формулой

$$p_{\xi}(x)=\begin{cases}\dfrac{C}{x^2}, & \text{если } x\geqslant 1,\\ 0, & \text{если } x<1.\end{cases}$$

Найти постоянную $C$, плотность распределения $\eta=\dfrac{1}{\xi}$, $\mathsf{P}\left(\dfrac{1}{2}<\eta<\dfrac{3}{4}\right)$.

4. Случайная величина $\xi$ имеет показательное распределение. *Найти плотность распределения* $\eta=\dfrac{1}{1-\xi}$.

5. Функция распределения $F(x)$ величины $\xi$ строго монотонна и непрерывна. Найти закон распределения величины $\eta=F(\xi)$.

6. Случайные величины $\xi_1$, $\xi_2$ независимы, одинаково распределены и имеют показательное распределение. Найти плотность распределения их суммы.

7. Величины $\xi_1$ и $\xi_2$ независимы и равномерно распределены на отрезках $[0, 1]$, $[1, 2]$. Найти плотность распределения $\xi_1/\xi_2$.

8. Точка наудачу брошена на отрезок $[0, 1]$. Ее координата равна

$$0, \xi_1\xi_2\xi_3\ldots=\sum_{k=1}^{\infty}\frac{\xi_k}{10^k}.$$

Найти совместное распределение $\xi_1$, $\xi_2$. Будут ли эти величины *независимы*?

9. Совместное распределение $(\xi_1, \xi_2)$ задано формулами

$$\mathsf{P}(\xi_1=-1,\xi_2=-1)=\mathsf{P}(\xi_1=0,\xi_2=-1)=\mathsf{P}(\xi_1=1,\xi_2=-1)=\frac{1}{6},$$

$$\mathsf{P}(\xi_1=-1,\ \xi_2=1)=\frac{1}{4},\quad \mathsf{P}(\xi_1=0,\ \xi_2=1)=\mathsf{P}(\xi_1=1,\xi_2=1)=\frac{1}{8}.$$

Найти одномерные распределения $\xi_1$, $\xi_2$ и распределения величин $\eta_1=\xi_1+\xi_2$, $\eta_2=\xi_1\xi_2$.

10. В предыдущей задаче найти совместное распределение величин $\eta_1$, $\eta_2$.

11. Величины $\xi_1$ и $\xi_2$ независимы, $\mathsf{P}(\xi_1=0)=\mathsf{P}(\xi_1=1)=\dfrac{1}{2}$, $\xi_2$ равномерно распределена на отрезке $[0, 1]$. Найти распределение величины $\xi_1+\xi_2$.

12. Найти закон распределения суммы двух независимых случайных величин, каждая из которых имеет распределение Пуассона.

13. Пусть в полиномиальной схеме (см. § 2 гл. 3) $n$ произвольно, $N=3$. Закон распределения чисел исходов различных

типов $\xi_1$, $\xi_2$, $\xi_3$ определяется формулой (3.2.7). Найти $\mathbf{P}(\xi_1=m_1, \xi_2=m_2 \mid \xi_3=m_3)$, где $m_1+m_2+m_3=n$.

14. Обозначим $\tau$ число испытаний в схеме Бернулли до появления первого успеха включительно. Найти закон распределения $\tau$.

15. Величина $\tau^{(1)}$ равна числу испытаний в схеме Бернулли до 1-го успеха включительно, а $\tau^{(2)}$ — число испытаний, прошедших после 1-го успеха до 2-го успеха. Найти совместное распределение $\tau^{(1)}$, $\tau^{(2)}$. Доказать их независимость.

16. Найти закон распределения величины $\tau_m$, равной числу испытаний в схеме Бернулли до появления $m$-го успеха.

17. Величины $\xi_1$, $\xi_2$ независимы и одинаково распределены: $\mathbf{P}(\xi_1=k)=\mathbf{P}(\xi_2=k)=2^{-k}$, $k=1, 2, \ldots$ Найти распределение $\xi_1+\xi_2$ непосредственно по определению. Сравнить с задачей 16 при $m=2$, $p=q=1/2$.

18. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — независимые одинаково распределенные абсолютно непрерывные случайные величины. При каждом $\omega$ расположим числа $\xi_k(\omega)$, $k=1, 2, \ldots, n$, в порядке возрастания и перенумеруем: $\eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_n$. Таким образом, $\eta_1(\omega)$ является наименьшим из чисел $(\xi_1(\omega), \ldots, \xi_n(\omega))$; $\eta_n(\omega)$ — наибольшее из тех же чисел и т. д. Найти плотности распределений: 1) $\eta_1$; 2) $\eta_n$; 3) $\eta_m$, $1<m<n$; 4) $(\eta_m, \eta_k)$.

19. Пусть $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4\}$ и элементарные события равновероятны. Положим $\xi_1(\omega_1)=\xi_1(\omega_4)=\xi_2(\omega_2)=\xi_2(\omega_4)=\xi_3(\omega_3)=\xi_3(\omega_4)=1$. В остальных случаях $\xi_1$, $\xi_2$, $\xi_3$ равны 0. Доказать, что случайные величины $\xi_1$, $\xi_2$, $\xi_3$ попарно независимы, но не являются независимыми в совокупности.

20. Машина состоит из 10 000 деталей. Каждая деталь независимо от других оказывается неисправной с вероятностью $p_i$, причем для $n_1=1000$ деталей — $p_1=0{,}0003$; для $n_2$ деталей — $p_2=0{,}0005$ и для $n_3=7000$ деталей — $p_3=0{,}0001$. Машина не работает, если в ней неисправны хотя бы две детали. Найти вероятность того, что машина не будет работать.

ГЛАВА 5

# ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

## § 1. Математическое ожидание. Определения и примеры

Функция распределения $F_\zeta(x)$ полностью определяет распределение случайной величины. Однако в ряде задач достаточно использовать более простые характеристики случайной величины. Одной из таких характеристик является среднее значение случайной величины или математическое ожидание.

Рассмотрим сначала следующий пример. Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ и элементарным событиям $\omega_k$ приписаны вероятности $p_k \geqslant 0$, $k = 1, 2, \ldots, N$, $\sum\limits_{k=1}^{N} p_k = 1$. Положим $\zeta = \zeta(\omega_k) = x_k$, $k = 1, 2, \ldots, N$. Средним значением случайной величины $\zeta$ естественно назвать

$$p_1x_1 + p_2x_2 + \ldots + p_Nx_N. \tag{1.1}$$

Предположим, что имеется серия из $n$ независимых испытаний, каждое из которых заключается в том, что введенная выше случайная величина $\zeta$ принимает определенное значение (например, проводится $n$ измерений неизвестной величины $a$; однократное измерение за счет ошибки можно считать случайной величиной $\zeta$). Среднее арифметическое чисел, полученных в $n$ испытаниях, близко к (1.1). Это следует из близости значений частот к соответствующим вероятностям. Действительно, пусть получены следующие значения: $y_1, y_2, \ldots, y_n$, где $y_k \in (x_1, x_2, \ldots, x_N)$. Тогда

$$\frac{y_1+y_2+\ldots+y_n}{n} = \frac{n_1x_1+n_2x_2+\ldots+n_Nx_N}{n}. \tag{1.2}$$

Частоты $n_k/n$ значений $x_k$ должны быть близки к $p_k$ и, следовательно,

$$\frac{y_1+y_2+\ldots+y_n}{n} \approx p_1x_1+p_2x_2+\ldots+p_Nx_N. \tag{1.3}$$

Точная формулировка и доказательство свойства (1.3) будут даны в гл. 6.

Перейдем теперь к определению математического ожидания. Наиболее удобное и простое определение математического ожидания произвольной случайной величины вводится при помощи интеграла Лебега.

Математическим ожиданием случайной величины $\zeta$, заданной на вероятностном пространстве $(\Omega \mathfrak{F} \mathsf{P})$, называется число

$$\mathsf{M}\zeta = \int_{\Omega} \zeta(\omega)\, \mathsf{P}(d\omega),$$

если интеграл Лебега, стоящий в правой части равенства, существует (см. [2], стр. 74, 319—324). Используя интеграл Стилтьеса, можно $\mathsf{M}\zeta$ записать в виде

$$\mathsf{M}\zeta = \int_{-\infty}^{\infty} x\, dF_{\zeta}(x).$$

Чтобы не пользоваться теорией интеграла Лебега, ограничимся определениями математического ожидания для случайных величин, заданных на дискретном и абсолютно непрерывном вероятностных пространствах (см. § 6 гл. 1).

*Математическим ожиданием* $\mathsf{M}\zeta$ *случайной величины* $\zeta = \zeta(\omega_k)$, *заданной на дискретном вероятностном пространстве* $(\Omega, \mathfrak{F}, \mathsf{P})$ (§ 6 гл. 1), *называется число*

$$\mathsf{M}\zeta = \sum_{k=1}^{\infty} \zeta(\omega_k)\, p_k, \tag{1.4}$$

*если ряд абсолютно сходится.* Если этот ряд не сходится абсолютно, то говорят, что математическое ожидание не существует.

*Математическим ожиданием* $\mathsf{M}\zeta$ *случайной величины* $\zeta = \zeta(u_1, u_2, \ldots, u_n)$, *заданной на абсолютно непрерывном вероятностном пространстве* $(\Omega, \mathfrak{F}, \mathsf{P})$ (см. § 6 гл. 1), *называется число*

$$\mathsf{M}\zeta = \int \underset{\Omega}{\cdots} \int \zeta(u_1, \ldots, u_n)\, \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n, \tag{1.5}$$

*если интеграл абсолютно сходится.*

Если интеграл (1.5) не сходится абсолютно, то говорят, что математическое ожидание $\mathsf{M}\zeta$ не существует.

Пример 1. Найдем математическое ожидание выигрыша $\zeta$ первого игрока в примере 1 § 1 гл. 4. В этом примере

$$\Omega=\{\Gamma,\ P\},\quad \mathsf{P}(\{\Gamma\})=\mathsf{P}(\{P\})=\frac{1}{2},$$
$$\zeta(\Gamma)=1,\ \zeta(P)=-1.$$

По формуле (1.4)

$$\mathsf{M}\zeta=\zeta(\Gamma)\,\mathsf{P}(\{\Gamma\})+\zeta(P)\,\mathsf{P}(\{P\})=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot(-1)=0.$$

Пример 2. Пусть в дискретном вероятностном пространстве (§ 6 гл. 1) $\Omega=\{1,\ 2,\ \ldots,\ k,\ \ldots\}$, $p_k=\frac{1}{k(k+1)}$. Величины $p_k$ удовлетворяют (1.6.4):

$$\sum_{k=1}^{\infty}p_k=\sum_{k=1}^{\infty}\left(\frac{1}{k}-\frac{1}{k+1}\right)=1.$$

Положим $\xi_1=\xi_1(k)=(-1)^k$, $\xi_2=\xi_2(k)=(-1)^k k$. Так как ряд $\sum\limits_{k=1}^{\infty}\frac{(-1)^k}{k(k+1)}$ сходится абсолютно, то $\mathsf{M}\xi_1$ существует. По формуле (1.4)

$$\mathsf{M}\xi_1=\sum_{k=1}^{\infty}\xi_1(k)\,p_k=$$
$$=\sum_{k=1}^{\infty}\frac{(-1)^k}{k(k+1)}=\sum_{k=1}^{\infty}(-1)^k\left[\frac{1}{k}-\frac{1}{k+1}\right]=$$
$$=\sum_{k=1}^{\infty}\frac{(-1)^k}{k}+\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k+1}=1+2\sum_{k=1}^{\infty}\frac{(-1)^k}{k}.$$

Полагая в разложении

$$-\ln(1+x)=\sum_{k=1}^{\infty}(-1)^k\frac{x^k}{k}$$

$x=1$, получим $\sum\limits_{k=1}^{\infty}\frac{(-1)^k}{k}=-\ln 2$. Таким образом,

$\mathsf{M}\xi_1 = 1 - 2\ln 2$. *Однако* $\mathsf{M}\xi_2$ по определению не существует, так как ряд $\sum\limits_{k=1}^{\infty} |(-1)^k k| \cdot \frac{1}{k(k+1)}$ расходится.

Пример 3. Пусть в абсолютно непрерывном вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ (§ 6 гл. 1)

$$\Omega = \{(u, v): u^2 + v^2 \leqslant 9\}, \quad \pi(u, v) = \frac{1}{9\pi}, \quad (u, v) \in \Omega.$$

Рассмотрим случайные величины: $\xi_1 = \sqrt{u^2+v^2}$, $\xi_2 = k$, если $k - 1 \leqslant \sqrt{u^2+v^2} < k$, $k = 1, 2, 3$;

$$\xi_3 = \xi_3(u, v) = \begin{cases} \sqrt{u^2+v^2}, & \text{если } \sqrt{u^2+v^2} < 1, \\ 1, & \text{если } \sqrt{u^2+v^2} \geqslant 1. \end{cases}$$

По формуле (1.5) находим

$$\mathsf{M}\xi_k = \iint\limits_{\Omega} \xi_k(u, v) \frac{1}{9\pi} \, du \, dv.$$

При вычислении этого интеграла естественно перейти к полярным координатам: $u = \rho\cos\varphi$, $v = \rho\sin\varphi$. Тогда

$$\mathsf{M}\xi_k = \iint\limits_{\Omega} \xi_k(\rho\cos\varphi, \rho\sin\varphi) \frac{\rho}{9\pi} \, d\rho \, d\varphi.$$

Отсюда

$$\mathsf{M}\xi_1 = \iint\limits_{\Omega} \frac{\rho^2}{9\pi} \, d\rho \, d\varphi = \frac{2}{9}\int\limits_0^3 \rho^2 \, d\rho = \frac{2}{9} \cdot \frac{\rho^3}{3}\Big|_0^3 = 2,$$

$$\mathsf{M}\xi_2 = \frac{1}{9\pi}\int\limits_0^{2\pi} d\varphi \left( \int\limits_0^1 \rho \, d\rho + \int\limits_1^2 2\rho \, d\rho + \int\limits_2^3 3\rho \, d\rho \right) = \frac{22}{9},$$

$$\mathsf{M}\xi_3 = \frac{1}{9\pi}\int\limits_0^{2\pi} d\varphi \left( \int\limits_0^1 \rho^2 \, d\rho + \int\limits_1^3 \rho \, d\rho \right) = \frac{26}{27}.$$

Отметим, что в данном примере величина $\xi_1$ абсолютно непрерывна, $\xi_2$ — дискретна, а $\xi_3$ — смешанного типа. Действительно,

$$F_{\xi_1}(x) = \mathsf{P}(\xi_1 < x) = \iint\limits_{u^2+v^2<x^2} \frac{1}{9\pi} \, du \, dv = \frac{x^2}{9}, \quad x \in [0, 3],$$

и, следовательно,

$$p_{\xi_1}(x)=\frac{2}{9}x, \quad x\in[0,\ 3]; \qquad p_{\xi_1}(x)=0, \quad x\notin[0,\ 3].$$

Так как $(\xi_2=k)=\{(u,\ v):\ k-1\leqslant\sqrt{u^2+v^2}<k\}$, $k=1,\ 2,\ 3$, то

$$\mathsf{P}(\xi_2=1)=1/9, \quad \mathsf{P}(\xi_2=2)=1/3, \quad \mathsf{P}(\xi_2=3)=5/9.$$

Нетрудно проверить, что

$$F_{\xi_3}(x)=\begin{cases}0, & x<0,\\ x^2/9, & 0\leqslant x\leqslant 1,\\ 1, & x>1.\end{cases}$$

Если уже известна плотность распределения случайной величины или вероятности значений дискретной величины, то может оказаться, что математическое ожидание удобнее вычислять не по определению, а по формулам, которые мы получим в качестве частного случая следующих двух теорем.

Теорема 1.1. *Пусть* $(\xi_1, \xi_2, \ldots, \xi_n)$ — *дискретный случайный вектор, для которого*

$$\mathsf{P}(\xi_1=x_{k1},\ \xi_2=x_{k2},\ \ldots,\ \xi_n=x_{kn})=p_{x_{k1},\,x_{k2},\,\ldots,\,x_{kn}}\geqslant 0,$$

$$\sum_{x_{k1}\ldots x_{kn}} p_{x_{k1}x_{k2}\ldots x_{kn}}=1.$$

*Если ряд* $\sum\limits_{x_{k1}\ldots x_{kn}}|g(x_{k1},\ \ldots,\ x_{kn})|\,p_{x_{k1}\ldots x_{kn}}$ *сходится, то случайная величина* $\eta=g(\xi_1,\ \xi_2,\ \ldots,\ \xi_n)$ *имеет математическое ожидание*

$$\mathsf{M}\eta=\sum_{x_{k1}\ldots x_{kn}} g(x_{k1},\ \ldots,\ x_{kn})\,p_{x_{k1}\ldots x_{kn}}. \tag{1.6}$$

Теорема 1.2. *Пусть* $(\xi_1, \xi_2, \ldots, \xi_n)$ — *абсолютно непрерывный случайный вектор с плотностью распределения* $p_{\xi_1\ldots\xi_n}(x_1,\ x_2,\ \ldots,\ x_n)$. *Если интеграл*

$$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}|g(x_1,\ \ldots,\ x_n)|\,p_{\xi_1\ldots\xi_n}(x_1,\ \ldots,\ x_n)\,dx_1\ldots dx_n$$

*сходится, то математическое ожидание случайной величины* $\eta = g(\xi_1, \ldots, \xi_n)$ *существует и*

$$\mathsf{M}\eta = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n)\, p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n. \tag{1.7}$$

Отсюда, полагая $n=1$ и $g(x)=x$, получим формулы для вычисления математического ожидания случайной величины по плотности распределения или по вероятностям значений дискретной случайной величины.

Если $\sum_{k=1}^{\infty} \mathsf{P}(\xi = x_k) = 1$, то из формулы (1.6) получим

$$\mathsf{M}\xi = \sum_{k=1}^{\infty} x_k \mathsf{P}(\xi = x_k). \tag{1.8}$$

Если $p_\xi(x)$ — плотность распределения $\xi$, то из формулы (1.7) следует, что

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\, dx. \tag{1.9}$$

Доказательство теоремы 1.1. Проведем доказательство для случая, когда случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ заданы на дискретном вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ (см. § 6 гл. 1). По определению случайный вектор называется дискретным, если существует такое множество точек $(x_{k1}, x_{k2}, \ldots, x_{kn})$, $k = 1, 2, \ldots$, что

$$\sum_{k=1}^{\infty} \mathsf{P}(\xi_1 = x_{k1}, \ \ldots, \ \xi_n = x_{kn}) = 1. \tag{1.10}$$

Математическое ожидание случайной величины $\eta = \eta(\omega_m) = g(\xi_1(\omega_m), \xi_2(\omega_m), \ldots, \xi_n(\omega_m))$ по формуле (1.4) запишем в виде

$$\mathsf{M}\eta = \sum_{m=1}^{\infty} g(\xi_1(\omega_m), \xi_2(\omega_m), \ldots, \xi_n(\omega_m))\, p_m. \tag{1.11}$$

Преобразование (1.11) к виду (1.6) основано на «приведении подобных членов». Введем множества

$$A_k=\{m\colon\ \xi_1(\omega_m)=x_{k1},\ \ldots,\ \xi_n(\omega_m)=x_{kn}\}. \quad (1.12)$$

Очевидно, что любая пара множеств $A_k$ и $A_l$, $l\neq k$; не имеет общих элементов и

$$\bigcup_{k=1}^{\infty} A_k=\{1,\ 2,\ 3,\ \ldots,\ m,\ \ldots\}.$$

Если $\mathsf{M}\eta$ существует, то ряд (1.11) сходится абсолютно и его члены можно произвольно переставлять. Тогда (1.11) можно записать в виде

$$\mathsf{M}\eta=\sum_{k=1}^{\infty}\sum_{m\in A_k} g(\xi_1(\omega_m),\ \ldots,\ \xi_n(\omega_n))\,p_m.$$

Из этого равенства с учетом (1.12) получим

$$\mathsf{M}\eta=\sum_{k=1}^{\infty}\sum_{m\in A_k} g(x_{k1},\ \ldots,\ x_{kn})\,p_m=$$
$$=\sum_{k=1}^{\infty} g(x_{k1},\ \ldots,\ x_{kn})\sum_{m\in A_k} p_m.$$

Отсюда следует формула (1.6), так как

$$\sum_{m\in A_k} p_m=\mathsf{P}(\xi_1=x_{k1},\ \ldots,\ \xi_n=x_{kn})=p_{x_{k1}\ldots x_{kn}}.$$

Если абсолютно сходится ряд в (1.6), то, начав с (1.6), можно аналогичными рассуждениями получить формулу для $\mathsf{M}\eta$ в виде (1.11).

Доказательство (1.6) для случая, когда дискретный случайный вектор задан на абсолютно непрерывном вероятностном пространстве, а также доказательство теоремы 1.2 приведены не будут. Эти доказательства при условии использования только фактов из курса обычного анализа потребовали бы введения дополнительных лишних ограничений на рассматриваемые случайные величины. Приведенное для дискретных пространств доказательство дает достаточно хорошее представление о связи формул (1.4), (1.5) с (1.6), (1.7).

Воспользуемся формулами (1.8) и (1.9) для вычисления математических ожиданий случайных величин, распределения которых были приведены в § 2 гл. 4.

1. *Нормальное распределение.* Плотность распределения величины $\xi$ задана формулой (4.2.3). Заменяя в (1.9) $p_\xi(x)$ по формуле (4.2.3), получим

$$\mathsf{M}\xi = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx.$$

Заменой переменной интегрирования $x = \sigma y + a$ приведем $\mathsf{M}\xi$ к следующему виду:

$$\mathsf{M}\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma y + a)\, e^{-\frac{y^2}{2}}\, dy =$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\, dy + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy.$$

Так как функция $ye^{-\frac{y^2}{2}}$ нечетная, а $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ является плотностью нормального распределения с параметрами (0, 1), то в правой части последнего равенства первое слагаемое равно 0, а второе $a$. Таким образом, для нормально распределенной величины

$$\mathsf{M}\xi = a. \qquad (1.13)$$

2. *Показательное распределение.* По формулам (1.9) и (2.4) находим

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\, dx = \int_{0}^{\infty} \alpha x e^{-\alpha x}\, dx = \frac{1}{\alpha}.$$

3. *Равномерное распределение.* Подставляя (4.2.5) в (1.9), найдем

$$\mathsf{M}\xi = \int_{-\infty}^{\infty} x p_\xi(x)\, dx = \frac{1}{b-a} \int_{a}^{b} x\, dx.$$

Отсюда

$$\mathsf{M}\xi = \frac{a+b}{2}. \tag{1.14}$$

4. *Биномиальное распределение.* По формулам (1.8) и (3.2.6) получим

$$\mathsf{M}\xi = \sum_{m=0}^{n} mC_n^m p^m (1-p)^{n-m}.$$

Нетрудно проверить, что $mC_n^m = nC_{n-1}^{m-1}$. Тогда $\mathsf{M}\xi$ можно записать так:

$$\mathsf{M}\xi = np \sum_{m=1}^{n} C_{n-1}^{m-1} p^{m-1} (1-p)^{n-m} = np\,[p + (1-p)]^{n-1}.$$

Таким образом,

$$\mathsf{M}\xi = np. \tag{1.15}$$

5. *Пуассоновское распределение.* По формулам (1.8) и (4.2.7) находим

$$\mathsf{M}\xi = \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda} \cdot e^{\lambda}.$$

Следовательно,

$$\mathsf{M}\xi = \lambda. \tag{1.16}$$

## § 2. Свойства математического ожидания

Приведем основные свойства математического ожидания.

Теорема 2.1.

1°. *Если $C$ постоянная, то $\mathsf{M}C = C$.*

2°. *Если $C$ постоянная, то $\mathsf{M}(C\xi) = C\mathsf{M}\xi$.*

3°. *Для любых величин $\xi$*

$$|\mathsf{M}\xi| \leqslant \mathsf{M}|\xi|.$$

4°. *Для любых случайных величин $\xi_1$ и $\xi_2$*

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

*Если существуют какие-нибудь два из участвующих в равенстве математических ожиданий, то существует третье.*

5°. *Если случайные величины $\xi_1$ и $\xi_2$ независимы, то $M\xi_1\xi_2 = M\xi_1 \cdot M\xi_2$. Из существования любых двух математических ожиданий следует существование третьего.*

Доказательство. Свойства 1°, 2°, 3°, 4° следуют из соответствующих свойств интеграла или ряда. Докажем, например, свойство 4° для величин, определенных в абсолютно непрерывном вероятностном пространстве (см. § 6 гл. 1). Пусть $\xi_1$, $\xi_2$ заданы на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$. Тогда сумма $\xi_1 + \xi_2$ также является случайной величиной, определенной на том же вероятностном пространстве, и по формуле (1.5)

$$M(\xi_1+\xi_2) = \int\underset{\Omega}{\cdots}\int [\xi_1(u_1, \ldots, u_n) + \xi_2(u_1, \ldots, u_n)]\, \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n =$$

$$= \int\underset{\Omega}{\cdots}\int \xi_1(u_1, \ldots, u_n)\, \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n +$$

$$+ \int\underset{\Omega}{\cdots}\int \xi_2(u_1, \ldots, u_n)\, \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n = M\xi_1 + M\xi_2.$$

Здесь мы воспользовались тем, что интеграл от суммы двух функций равен сумме интегралов.

Докажем свойство 5°. Пусть, например, $\xi_1$ и $\xi_2$ — абсолютно непрерывные величины и $p_{\xi_1\xi_2}(u, v)$ — их плотность распределения. Так как $\xi_1$ и $\xi_2$ независимы, то $p_{\xi_1\xi_2}(u, v) = p_{\xi_1}(u)\, p_{\xi_2}(v)$. По формуле (1.7) с $n = 2$ и $g(x_1, x_2) = x_1 x_2$ получим

$$M\xi_1\xi_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv p_{\xi_1}(u)\, p_{\xi_2}(v)\, du\, dv =$$

$$= \int_{-\infty}^{\infty} u p_{\xi_1}(u)\, du \cdot \int_{-\infty}^{\infty} v p_{\xi_2}(v)\, dv = M\xi_1 \cdot M\xi_2.$$

Из свойств 2° и 4° по индукции следует, что

$$\mathsf{M}(C_1\xi_1+\ldots+C_n\xi_n)=C_1\mathsf{M}\xi_1+\ldots+C_n\mathsf{M}\xi_n. \quad (2.1)$$

Приведенные выше свойства помогают при вычислении математических ожиданий. Рассмотрим два примера.

Пример 1. Пусть $\xi$ имеет нормальное распределение с параметрами $(a, \sigma)$. В § 1 было показано, что $\mathsf{M}\xi=a$. Найдем математическое ожидание величины $\eta=A\xi+B$. Используя свойства 1°, 2°, 4°, получим

$$\mathsf{M}\eta=A\mathsf{M}\xi+B=Aa+B. \quad (2.2)$$

В § 5 гл. 4 мы доказали, что величина $\eta$ имеет нормальное распределение с параметрами $(a_1\sigma_1)$, где $a_1=Aa+B$. Это значение естественно совпадает с (2.2).

Пример 2. Пусть $\mu_n$ — число успехов в $n$ испытаниях Бернулли с вероятностью $p$ успеха в отдельном испытании. Эта величина имеет биномиальное распределение. В § 1 было показано, что случайная величина с биномиальным распределением имеет математическое ожидание, равное $np$. Покажем, как этот результат можно получить при помощи представления $\mu_n$ в виде (4.4.18):

$$\mu_n=\xi_1+\xi_2+\ldots+\xi_n,$$

где $\xi_k$ — независимые одинаково распределенные случайные величины с $\mathsf{P}(\xi_k=1)=p$, $\mathsf{P}(\xi_k=0)=1-p$. По формуле (2.1)

$$\mathsf{M}\mu_n=\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_n.$$

Слагаемые легко вычисляются по формуле (1.8):

$$\mathsf{M}\xi_k=1\cdot p+0\cdot(1-p)=p, \qquad k=1, 2, \ldots, n.$$

Таким образом, $\mathsf{M}\mu_n=np$.

В качестве еще одного примера использования свойств математического ожидания докажем следующую теорему.

Теорема 2.2. *Если $A_1, A_2, \ldots, A_n$ — произвольные события, то*

$$\mathsf{P}(A_1+\ldots+A_n)=\sum_{k=1}^{n}\mathsf{P}(A_k)-$$
$$-\sum_{1\leqslant k_1<k_2\leqslant n}\mathsf{P}(A_{k_1}A_{k_2})+\ldots(-1)^{n+1}\mathsf{P}(A_1\ldots A_n)=$$
$$=\sum_{m=1}^{n}(-1)^{m+1}\sum_{1\leqslant k_1<k_2<\ldots<k_m\leqslant n}\mathsf{P}(A_{k_1}A_{k_2}\ldots A_{k_m}).$$

Доказательство. Из равенства

$$\overline{A_1+A_2+\ldots+A_n}=\overline{A}_1\overline{A}_2\ldots\overline{A}_n$$

следует, что

$$\mathsf{P}(A_1+A_2+\ldots+A_n)=1-\mathsf{P}(\overline{A}_1\overline{A}_2\ldots\overline{A}_n). \qquad (2.3)$$

Введем случайные величины $\eta_k$, $k=1, 2, \ldots, n$, положив

$$\eta_k=\eta_k(\omega)=\begin{cases}1, & \text{если} \quad \omega\in A_k,\\ 0, & \text{если} \quad \omega\in\overline{A}_k.\end{cases}$$

Очевидно, что случайная величина $\prod\limits_{k=1}^{n}(1-\eta_k)$ принимает два значения: значение 1, если все $\eta_k=0$, и значение 0 в остальных случаях. Событие $(\eta_1=\eta_2=\ldots=\eta_n=0)=\overline{A}_1\overline{A}_2\ldots\overline{A}_n$. Следовательно,

$$\mathsf{M}\prod_{k=1}^{n}(1-\eta_k)=\mathsf{P}(\overline{A}_1\overline{A}_2\ldots\overline{A}_n). \qquad (2.4)$$

Так как

$$\prod_{k=1}^{n}(1-\eta_k)=$$
$$=1-\sum_{k=1}^{n}\eta_k+\sum_{1\leqslant k_1<k_2\leqslant n}\eta_{k_1}\eta_{k_2}-\ldots(-1)^n\,\eta_1\eta_2\ldots\eta_n,$$

то

$$\mathsf{M}\prod_{k=1}^{n}(1-\eta_k)=$$
$$=1-\sum_{k=1}^{n}\mathsf{M}\eta_k+\sum_{1\leqslant k_1<k_2\leqslant n}\mathsf{M}\eta_{k_1}\eta_{k_2}-\ldots(-1)^n\,\mathsf{M}\eta_1\ldots\eta_n. \tag{2.5}$$

Случайная величина $\eta_{k_1}\eta_{k_2}\ldots\eta_{k_m}$ может быть определена следующим образом:

$$\eta_{k_1}\eta_{k_2}\ldots\eta_{k_m}=\begin{cases}1, & \text{если}\quad \omega\in A_{k_1}A_{k_2}\ldots A_{k_m},\\ 0, & \text{если}\quad \omega\in\overline{A_{k_1}A_{k_2}\ldots A_{k_m}}.\end{cases}$$

Следовательно,

$$\mathsf{M}\eta_{k_1}\eta_{k_2}\ldots\eta_{k_m}=\mathsf{P}(A_{k_1}A_{k_2}\ldots A_{k_m}). \tag{2.6}$$

Утверждение теоремы следует из формул (2.3), (2.4), (2.5), (2.6).

Приведенная в теореме формула обобщает формулу (1.5.4).

## § 3. Дисперсия

*Дисперсией* $\mathsf{D}\xi$ случайной величины $\xi$ называется число

$$\mathsf{D}\xi=\mathsf{M}(\xi-\mathsf{M}\xi)^2, \tag{3.1}$$

если математическое ожидание в правой части (3.1) существует. Дисперсия является мерой рассеяния значений случайной величины около ее математического ожидания. Величину $\sqrt{\mathsf{D}\xi}$ называют средним квадратическим отклонением.

Если воспользоваться свойствами математического ожидания, то правую часть (3.1) можно привести к следующему виду:

$$\mathsf{M}(\xi-\mathsf{M}\xi)^2=\mathsf{M}(\xi^2-2\xi\mathsf{M}\xi+(\mathsf{M}\xi)^2)=$$
$$=\mathsf{M}\xi^2-2\mathsf{M}\xi\cdot\mathsf{M}\xi+(\mathsf{M}\xi)^2.$$

Отсюда и из (3.1) следует, что

$$\mathsf{D}\xi = \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2. \tag{3.2}$$

Если случайная величина $\xi$ абсолютно непрерывна, то, полагая в формуле (1.7) $n=1$ и $g(x_1) = (x_1 - \mathsf{M}\xi)^2$, получим

$$\mathsf{D}\xi = \int_{-\infty}^{\infty} (x - \mathsf{M}\xi)^2 p_\xi(x)\,dx. \tag{3.3}$$

Для дискретной величины $\xi$ из (1.6) найдем

$$\mathsf{D}\xi = \sum_{k=1}^{\infty} (x_k - \mathsf{M}\xi)^2 \mathsf{P}(\xi = x_k), \tag{3.4}$$

где $\sum\limits_{k=1}^{\infty} \mathsf{P}(\xi = x_k) = 1$.

Приведем свойства дисперсии.

Теорема 3.1.

1°. *Для любой случайной величины $\xi$ имеем $\mathsf{D}\xi \geqslant 0$.*

2°. *Если $c$ постоянная, то $\mathsf{D}c = 0$.*

3°. *Если $c$ постоянная, то $\mathsf{D}(c\xi) = c^2\mathsf{D}\xi$.*

4°. *Если случайные величины $\xi_1$ и $\xi_2$ независимы, то*

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{D}\xi_1 + \mathsf{D}\xi_2. \tag{3.5}$$

Доказательство. Свойства 1°, 2°, 3° следуют непосредственно из определения и свойств математического ожидания. Докажем свойство 4°. По определению (3.1)

$$\begin{aligned}\mathsf{D}(\xi_1 + \xi_2) &= \mathsf{M}[(\xi_1 + \xi_2) - \mathsf{M}(\xi_1 + \xi_2)]^2 = \\ &= \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1) + (\xi_2 - \mathsf{M}\xi_2)]^2 = \\ &= \mathsf{M}[(\xi_1 - \mathsf{M}\xi_1)^2 + (\xi_2 - \mathsf{M}\xi_2)^2 + 2(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)].\end{aligned} \tag{3.6}$$

Отсюда, так как случайные величины $\xi_1 - \mathsf{M}\xi_1$, $\xi_2 - \mathsf{M}\xi_2$ независимы и

$$\begin{aligned}\mathsf{M}(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2) &= \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)\cdot\mathsf{M}(\xi_2 - \mathsf{M}\xi_2) = \\ &= (\mathsf{M}\xi_1 - \mathsf{M}\xi_1)(\mathsf{M}\xi_2 - \mathsf{M}\xi_2) = 0,\end{aligned}$$

следует, что

$$\mathsf{D}(\xi_1 + \xi_2) = \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^2 + \mathsf{M}(\xi_2 - \mathsf{M}\xi_2)^2 = \mathsf{D}\xi_1 + \mathsf{D}\xi_2.$$

Вычислим дисперсии некоторых случайных величин.

1. *Нормальное распределение.* По формулам (4.2.3) и (3.3) в этом случае имеем

$$\mathsf{D}\xi = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx.$$

Отсюда, полагая $x = \sigma y + a$, получим

$$\mathsf{D}\xi = \sigma^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}}\,dy = \sigma^2\frac{1}{\sqrt{2\pi}}\left(-\int_{-\infty}^{\infty} y\,d\left(e^{-\frac{y^2}{2}}\right)\right) =$$

$$= \sigma^2\left[-y\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\Bigg|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy\right].$$

Первое слагаемое в квадратных скобках равно 0, а второе равно 1. Таким образом,

$$\mathsf{D}\xi = \sigma^2.$$

В § 5 гл. 4 было показано, что линейная функция $\eta = A\xi + B$ от нормально распределенной случайной величины $\xi$ имеет нормальное распределение. В § 2 был указан способ вычисления $\mathsf{M}\eta$, не связанный с доказательством нормальности $\eta$ (см. формулу (2.2)). Найдем теперь $\mathsf{D}\eta$. Нетрудно показать, что случайная величина, являющаяся постоянной, независима с любой случайной величиной. Следовательно, при вычислении $\mathsf{D}\eta = \mathsf{D}(A\xi + B)$ можно воспользоваться формулой (3.5). Тогда

$$\mathsf{D}\eta = \mathsf{D}(A\xi + B) = \mathsf{D}(A\xi) + \mathsf{D}B = A^2\mathsf{D}\xi = A^2\sigma^2.$$

2. *Равномерное распределение.* По формулам (4.2.5), (3.3) и (1.13)

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

Отсюда

$$\mathsf{D}\xi = \frac{(b-a)^2}{12}, \tag{3.7}$$

3. *Пуассоновское распределение.* Найдем сначала $\mathsf{M}[\xi(\xi-1)]$. По формуле (1.6) с $n=1$ и $g(x)=x(x-1)$, воспользовавшись (4.2.7), получим

$$\mathsf{M}\xi(\xi-1)=\sum_{m=0}^{\infty} m(m-1)\frac{\lambda^m}{m!}e^{-\lambda}=\lambda^2\sum_{m=2}^{\infty}\frac{\lambda^{m-2}}{(m-2)!}e^{-\lambda}=\lambda^2.$$

Так как $\mathsf{M}\xi(\xi-1)=\mathsf{M}\xi^2-\mathsf{M}\xi$ и $\mathsf{M}\xi=\lambda$, то

$$\mathsf{M}\xi^2=\lambda^2+\lambda.$$

Подставляя это выражение в (3.2), получим

$$\mathsf{D}\xi=\lambda.$$

4. *Биномиальное распределение.* Так же, как в примере 2 из § 2, воспользуемся тем, что число успехов в схеме Бернулли имеет биномиальное распределение и представимо в виде суммы (4.4.18) с независимыми слагаемыми. Тогда

$$\mathsf{D}\mu_n=\mathsf{D}\xi_1+\mathsf{D}\xi_2+\ldots+\mathsf{D}\xi_n$$

и $\mathsf{D}\xi_k=\mathsf{M}\xi_k^2-(\mathsf{M}\xi_k)^2=p-p^2=p(1-p)$. Таким образом,

$$\mathsf{D}\mu_n=npq.$$

В формулировку теоремы Муавра—Лапласа (§ 3 гл. 3) входила линейная функция $(\mu_n-np)/\sqrt{npq}$. Теперь мы можем эту величину задать в таком виде: $(\mu_n-\mathsf{M}\mu_n)/\sqrt{\mathsf{D}\mu_n}$. При любой случайной величине $\xi$ для

$$\eta=\frac{\xi-\mathsf{M}\xi}{\sqrt{\mathsf{D}\xi}}$$

имеют место равенства

$$\mathsf{M}\eta=0,\qquad \mathsf{D}\eta=1.$$

При вычислении $\mathsf{M}\mu_n$ и $\mathsf{D}\mu_n$ мы воспользовались представлением (4.4.18). В ряде задач изучаемую случайную величину удается представить в виде суммы более простых зависимых величин. Пусть, например,

$$\eta_n=\xi_1+\xi_2+\ldots+\xi_n, \tag{3.8}$$

где случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ зависимы и каждая из них принимает значения 0 и 1. В этом случае можно воспользоваться формулой

$$M\eta_n = M\xi_1 + M\xi_2 + \ldots + M\xi_n.$$

Однако $D\eta_n$ уже не равна сумме дисперсий $D\xi_k$. Для вычисления $D\eta_n$ можно использовать формулу (3.2). Так как $\xi_k^2(\omega) = \xi_k(\omega)$, $\omega \in \Omega$ (действительно $\xi_k(\omega) = 1$ или $\xi_k(\omega) = 0$), то

$$\eta_n^2 = \sum_{k=1}^{n} \xi_k^2 + \sum_{k \neq l} \xi_k \xi_l = \sum_{k=1}^{n} \xi_k + \sum_{k \neq l} \xi_k \xi_l = \eta_n + \sum_{k \neq l} \xi_k \xi_l.$$

Таким образом,

$$D\eta_n = M\eta_n^2 - (M\eta_n)^2 = \sum_{k \neq l} M\xi_k \xi_l + M\eta_n - (M\eta_n)^2.$$

Суммой вида (3.8) является, например, случайная величина в задаче 7, стр. 128.

## § 4. Ковариация. Коэффициент корреляции

При доказательстве формулы (3.5) нам потребовалось вычислить $M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]$. Это число называется ковариацией случайных величин $\xi_1$, $\xi_2$ и обозначается $\operatorname{cov}(\xi_1, \xi_2)$. Таким образом,

$$\operatorname{cov}(\xi_1, \xi_2) = M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]. \qquad (4.1)$$

Отсюда, используя свойства математического ожидания, легко получить следующую формулу:

$$\operatorname{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1 \cdot M\xi_2. \qquad (4.2)$$

Очевидно, что

$$\operatorname{cov}(\xi, \xi) = D\xi, \quad \operatorname{cov}(\xi_1, \xi_2) = \operatorname{cov}(\xi_2, \xi_1).$$

Из (3.6) и (4.1) следует формула для дисперсии суммы двух произвольных (не обязательно зависимых) случайных величин:

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2 + 2\operatorname{cov}(\xi_1, \xi_2). \qquad (4.3)$$

Теорема 4.1. *Если для случайных величин* $\xi_1, \xi_2, \ldots, \xi_n$ *существуют* $\operatorname{cov}(\xi_i, \xi_j) = \sigma_{ij}$, $i, j = 1, \ldots, n$, *то при любых постоянных* $c_1, c_2, \ldots, c_n$ *имеем*

$$D(c_1\xi_1 + c_2\xi_2 + \ldots + c_n\xi_n) = \sum_{i, j=1}^{n} \sigma_{ij} c_i c_j. \qquad (4.4)$$

Доказательство. Положим

$$\eta_n = c_1\xi_1 + c_2\xi_2 + \ldots + c_n\xi_n.$$

Нетрудно проверить, что

$$\eta_n - \mathsf{M}\eta_n = \sum_{i=1}^{n} c_i(\xi_i - \mathsf{M}\xi_i)$$

и

$$(\eta_n - \mathsf{M}\eta_n)^2 = \sum_{i,\,j=1}^{n} c_i c_j (\xi_i - \mathsf{M}\xi_i)(\xi_j - \mathsf{M}\xi_j).$$

Вычисляя математическое ожидание от обеих частей последнего равенства, получим утверждение теоремы.

Правую часть (4.4) можно рассматривать как квадратичную форму от переменных $c_1, c_2, \ldots, c_n$. Так как при любых $c_1, c_2, \ldots, c_n$ дисперсия в левой части (4.4) неотрицательна, то квадратичная форма в правой части (4.4) неотрицательно определена. Квадратичная форма неотрицательно определена тогда и только тогда, когда неотрицательны все главные миноры матрицы, составленной из ее коэффициентов. Таким образом, из теоремы 4.1 получили следующее утверждение.

Определитель

$$\begin{vmatrix} \operatorname{cov}(\xi_1, \xi_1) & \operatorname{cov}(\xi_1, \xi_2) & \ldots & \operatorname{cov}(\xi_1, \xi_m) \\ \operatorname{cov}(\xi_2, \xi_1) & \operatorname{cov}(\xi_2, \xi_2) & \ldots & \operatorname{cov}(\xi_2, \xi_m) \\ \cdot & \cdot & \cdot & \cdot \\ \operatorname{cov}(\xi_m, \xi_1) & \operatorname{cov}(\xi_m, \xi_2) & \ldots & \operatorname{cov}(\xi_m, \xi_m) \end{vmatrix} \geqslant 0 \qquad (4.5)$$

для любых случайных величин $\xi_1, \xi_2, \ldots, \xi_m$; $m = 1, 2, \ldots$. При $m = 2$ неравенство (4.5) имеет вид

$$\begin{vmatrix} \mathsf{D}\xi_1 & \operatorname{cov}(\xi_1, \xi_2) \\ \operatorname{cov}(\xi_1, \xi_2) & \mathsf{D}\xi_2 \end{vmatrix} = \mathsf{D}\xi_1\,\mathsf{D}\xi_2 - \operatorname{cov}^2(\xi_1, \xi_2) \geqslant 0.$$

Отсюда

$$|\operatorname{cov}(\xi_1, \xi_2)| \leqslant \sqrt{\mathsf{D}\xi_1\,\mathsf{D}\xi_2}. \qquad (4.6)$$

В доказательстве формулы (3.5) было попутно получено, что для независимых случайных величин $\xi_1, \xi_2$ имеет место равенство

$$\operatorname{cov}(\xi_1, \xi_2) = 0. \qquad (4.7)$$

Таким образом, если $\operatorname{cov}(\xi_1, \xi_2) \neq 0$, то величины $\xi_1$ и $\xi_2$ зависимы. В качестве количественной характеристики степени зависимости случайных величин $\xi_1$ и $\xi_2$ используется коэффициент корреляции $\rho(\xi_1, \xi_2)$, определяемый следующим равенством:

$$\rho(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{\mathsf{D}\xi_1 \mathsf{D}\xi_2}}. \tag{4.8}$$

Свойства коэффициента корреляции:

1°. $|\rho(\xi_1, \xi_2)| \leqslant 1$.

2°. *Если $\xi_1$ и $\xi_2$ независимы, то* $\rho(\xi_1, \xi_2) = 0$.

3°. *Если $\xi_2 = A\xi_1 + B$, где $A$ и $B$ постоянные, то*

$$|\rho(\xi_1, \xi_2)| = 1.$$

Доказательство. Свойство 1° следует из (4.8) и (4.6); свойство 2° следует из (4.8) и (4.7). Докажем свойство 3°. Положим $\mathsf{M}\xi_1 = a$, $\mathsf{D}\xi_1 = \sigma^2$. Тогда

$$\mathsf{M}\xi_2 = Aa + B, \qquad \mathsf{D}\xi_2 = A^2\sigma^2,$$

$$\operatorname{cov}(\xi_1, \xi_2) = \mathsf{M}[(\xi_1 - a)(\xi_2 - \mathsf{M}\xi_2)] =$$
$$= \mathsf{M}[(\xi_1 - a)(A(\xi_1 - a))] = A\mathsf{D}\xi_1 = A\sigma^2$$

и, следовательно,

$$\rho(\xi_1, \xi_2) = \frac{A\sigma^2}{\sqrt{A^2\sigma^2 \cdot \sigma^2}} = \frac{A}{|A|}.$$

Таким образом, $|\rho(\xi_1, \xi_2)| = 1$.

Отметим, что равенство 0 коэффициента корреляции не является достаточным условием независимости случайных величин. Из равенства $\rho(\xi_1, \xi_2) = 0$ не следует независимость случайных величин (см. задачу 10 этой главы).

Наряду с рассмотренными выше числовыми характеристиками случайных величин часто используются моменты более высоких порядков. *Моментом порядка $k$* случайной величины $\xi$ называется число $\mathsf{M}\xi^k$. Число $\mathsf{M}(\xi - \mathsf{M}\xi)^k$ называется *центральным моментом порядка $k$*.

Пусть задан случайный вектор $(\xi_1, \xi_2, \ldots, \xi_n)$. Величины

$$\mathsf{M}\xi_1^{k_1}\xi_2^{k_2}\ldots\xi_n^{k_n}, \ \mathsf{M}(\xi_1 - \mathsf{M}\xi_1)^{k_1}\ldots(\xi_n - \mathsf{M}\xi_n)^{k_n}$$

называются соответственно *смешанным моментом порядка* $k=k_1+k_2+\ldots+k_n$ и *смешанным центральным моментом порядка* $k$. Вычислять моменты более высоких порядков можно по формулам (1.6), (1.7). Например, для абсолютно непрерывных величин

$$\mathsf{M}\xi^k=\int\limits_{-\infty}^{\infty}x^k p_\xi(x)\,dx,$$

$$\mathsf{M}(\xi_1-\mathsf{M}\xi_1)^{k_1}\ldots(\xi_n-\mathsf{M}\xi_n)^{k_n}=\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty}(x-\mathsf{M}\xi_1)^{k_1}\ldots$$
$$\ldots(x-\mathsf{M}\xi_n)^{k_n}p_{\xi_1\ldots\xi_n}(x_1,\ldots,x_n)\,dx_1\ldots dx_n$$

и т. д.

Отметим еще, что из существования момента $\mathsf{M}\xi^m$ следует существование моментов $\mathsf{M}\xi^k$, $k=1, 2, \ldots$ $\ldots, m-1$. Это утверждение следует из неравенств

$$|\xi(\omega)|^k\leqslant|\xi(\omega)|^m+1,\qquad \omega\in\Omega,$$
$$k=1, 2, \ldots, m-1.$$

## § 5. Условные распределения и условные математические ожидания

Пусть в пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$ определен случайный вектор $(\xi, \eta)$. Введем понятие условного распределения величины $\xi$ при условии, что задано значение $\eta$. Рассмотрим сначала дискретный случай. Пусть

$$\mathsf{P}(\xi=x_i,\ \eta=y_j)=p_{ij}>0,\qquad \sum_{i,j=1}^{\infty}p_{ij}=1,$$

$$\mathsf{P}(\xi=x_i)=\sum_{j=1}^{\infty}p_{ij}=p_{i\cdot}>0,\qquad \mathsf{P}(\eta=y_j)=\sum_{i=1}^{\infty}p_{ij}=p_{\cdot j}>0.$$

В § 1 гл. 2 было дано определение условной вероятности. По этому определению

$$\mathsf{P}(\xi=x_i\mid\eta=y_j)=\frac{\mathsf{P}(\xi=x_i,\ \eta=y_j)}{\mathsf{P}(\eta=y_j)}=\frac{p_{ij}}{p_{\cdot j}},\qquad(5.1)$$
$$i=1, 2, \ldots$$

Если фиксировать $y_j$ (или $j$), то вероятности (5.1) можно рассматривать как условное распределение величины $\xi$ при условии, что $\eta=y_j$. Условным математическим ожиданием случайной величины $\xi$ при условии, что $\eta=y_j$, называется число

$$\mathsf{M}(\xi \mid \eta=y_j)=\sum_{i=1}^{\infty} x_i \frac{p_{ij}}{p_{\cdot j}}. \tag{5.2}$$

Отметим, что приведенная здесь в качестве определения формула (5.2) аналогична формуле для вычисления безусловного математического ожидания (1.8). В случае дискретной величины $\eta$ можно дать более естественное определение $\mathsf{M}(\xi \mid \eta=y_j)$, аналогичное (1.4) и (1.5). Для этого нужно воспользоваться формулами (1.4) и (1.5) для вероятностного пространства $(\Omega, \mathfrak{F}, \mathsf{P}_{y_j})$; вероятность $\mathsf{P}_{y_j}$ определена для любого $A \in \mathfrak{F}$ следующей формулой:

$$\mathsf{P}_{y_j}(A)=\mathsf{P}(A \mid \eta=y_j).$$

Тогда в (1.4) и (1.5) нужно заменить $p_k$ и $\pi(x_1, x_2, \ldots, x_n)$ на

$$\frac{p_k}{\sum\limits_{\omega_l \in (\eta=y_j)} p_l}, \quad \frac{\pi(x_1, x_2, \ldots, x_n)}{\int\limits_{(\eta=y_j)} \cdots \int \pi(u_1, \ldots, u_n)\, du_1 \ldots du_n}.$$

Если левые части (5.1) и (5.2) рассматривать как функции от $y_j$, то можно считать *условное распределение* и *условное математическое ожидание случайными величинами*, определенными в исходном вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$. Тогда условное распределение и условное математическое ожидание $\xi$ при условии $\eta$ определяются соответственно формулами

$$\mathsf{P}(\xi=x_i \mid \eta)=\begin{cases} \mathsf{P}(\xi=x_i \mid \eta=y_j), & \text{если} \quad \omega \in (\eta=y_j), \\ j=1, 2, \ldots, \end{cases}$$

$$\mathsf{M}(\xi \mid \eta)=\begin{cases} \mathsf{M}(\xi \mid \eta=y_j), & \text{если} \quad \omega \in (\eta=y_j), \\ j=1, 2, \ldots, \end{cases}$$

где $\mathsf{P}(\xi=x_i \mid \eta=y_j)$, $\mathsf{M}(\xi \mid \eta=y_j)$ определены формулами (5.1) и (5.2). *Справедлива следующая* формула (*формула полного математического ожидания*):

$$\mathsf{M}\xi=\mathsf{M}[\mathsf{M}(\xi \mid \eta)]. \tag{5.3}$$

Здесь условное математическое ожидание в квадратных скобках рассматривается как случайная величина. Применяя формулу (1.8) к случайной величине $\mathsf{M}(\xi \mid \eta)$, можно (5.3) записать в следующей эквивалентной форме:

$$\mathsf{M}\xi=\sum_{j=1}^{\infty} \mathsf{P}(\eta=y_j)\, \mathsf{M}(\xi \mid \eta=y_j). \tag{5.4}$$

Докажем формулу (5.4). Воспользовавшись определением (5.2), получим

$$\sum_{j=1}^{\infty} \mathsf{P}(\eta=y_j)\,\mathsf{M}(\xi\mid\eta_j)=\sum_{j=1}^{\infty} \mathsf{P}(\eta=y_j)\sum_{i=1}^{\infty} x_i\,\frac{\mathsf{P}(\xi=x_i,\ \eta=y_j)}{\mathsf{P}(\eta=y_j)}=$$
$$=\sum_{i,\,j=1}^{\infty} x_i\mathsf{P}(\xi=x_i,\ \eta=y_j)=\sum_{i=1}^{\infty} x_i\left(\sum_{j=1}^{\infty}\mathsf{P}(\xi=x_i,\ \eta=y_j)\right).$$

Правая часть последнего равенства равна $\mathsf{M}\xi$, так как $\sum\limits_{j=1}^{\infty} \mathsf{P}(\xi=x_i,\ \eta=y_j)=\mathsf{P}(\xi=x_i)$. Покажем, что из формул (5.3) и (5.4) следует формула полной вероятности. Пусть

$$\xi=\begin{cases}1, & \text{если } \omega\in A,\\ 0, & \text{если } \omega\in\overline{A},\end{cases}\qquad \eta=\begin{cases}y_j, & \text{если } \omega\in B_j,\\ j=1,\,2,\,\ldots,\end{cases}$$

где $B_iB_j=\varnothing$, $i\neq j$, $\bigcup\limits_{j=1}^{\infty} B_j=\Omega$. Тогда

$$\mathsf{M}\xi=\mathsf{P}(A),\qquad \mathsf{P}(\eta=y_j)=\mathsf{P}(B_j),\qquad \mathsf{M}(\xi\mid\eta=y_j)=\mathsf{P}(A\mid B_j).$$

Подставляя эти выражения в (5.4), получим формулу полной вероятности со счетной системой событий $B_1$, $B_2$, ... :

$$\mathsf{P}(A)=\sum_{k=1}^{\infty}\mathsf{P}(B_k)\,\mathsf{P}(A\mid B_k). \tag{5.5}$$

Положим в (5.5)

$$A=(\xi\in C),\qquad B_k=(\eta=y_k),$$

где $C\subset(x_1,\,x_2,\,\ldots,\,x_n,\,\ldots)$, $(x_i,\,y_j)$ — значения $(\xi,\,\eta)$. Тогда при любом $C$

$$\mathsf{P}(\xi\in C)=\sum_{x_i\in C}\mathsf{P}(\eta=y_k)\,\mathsf{P}(\xi=x_i\mid\eta=y_k). \tag{5.6}$$

Свойства условного математического ожидания:

1°. $$\mathsf{M}[\varphi(\eta)\mid\eta]=\varphi(\eta). \tag{5.7}$$

2°. $$\mathsf{M}[\varphi(\eta)\,\xi\mid\eta]=\varphi(\eta)\,\mathsf{M}(\xi\mid\eta). \tag{5.8}$$

3°. $$\mathsf{M}(\xi_1+\xi_2\mid\eta)=\mathsf{M}(\xi_1\mid\eta)+\mathsf{M}(\xi_2\mid\eta). \tag{5.9}$$

4°. *Если $\xi$ и $\eta$ независимы, то*

$$\mathsf{M}(\xi\mid\eta)=\mathsf{M}\xi. \tag{5.10}$$

Равенства (5.7) — (5.10) верны при любом $\omega \in \Omega$. Равенства (5.7), (5.8), (5.10) следуют непосредственно из определения (5.2). Равенство (5.9) легко получить из определения, аналогичного (1.4), (1.5).

Перейдем к рассмотрению абсолютно непрерывного вектора $(\xi, \eta)$. Так как в этом случае $\mathsf{P}(\eta = y) = 0$ при любом $y$, то мы не сможем воспользоваться определением условной вероятности (2.1.1), как это мы сделали в дискретном случае. Назовем *условной плотностью распределения вероятностей* величины $\xi$ при условии, что $\eta = y$, следующую функцию:

$$p_\xi(x \mid \eta = y) = \frac{p_{\xi, \eta}(x, y)}{\int\limits_{-\infty}^{\infty} p_{\xi, \eta}(u, y)\, du}. \tag{5.11}$$

Условное математическое ожидание $\xi$ при условии, что $\eta = y$, определяется формулой

$$\mathsf{M}(\xi \mid \eta = y) = \int\limits_{-\infty}^{\infty} x p_\xi(x \mid \eta = y)\, dx. \tag{5.12}$$

Из (5.11) нетрудно получить, что при любых $a$ и $b$

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \int\limits_{-\infty}^{\infty} p_\eta(y) \left( \int\limits_a^b p_\xi(x \mid \eta = y)\, dx \right) dy. \tag{5.13}$$

Легко также проверяется формула, аналогичная (5.4):

$$\mathsf{M}\xi = \int\limits_{-\infty}^{\infty} p_\eta(y)\, \mathsf{M}(\xi \mid \eta = y)\, dy. \tag{5.14}$$

Если рассматривать (5.11) и (5.12) как случайные величины

$$p_\xi(x \mid \eta) = \begin{cases} p_\xi(x \mid \eta = y), & \text{если } \omega \in (\eta = y), \\ y \in (-\infty, \infty), \end{cases}$$

$$\mathsf{M}(\xi \mid \eta) = \begin{cases} \mathsf{M}(\xi \mid \eta = y), & \text{если } \omega \in (\eta = y), \\ y \in (-\infty, \infty), \end{cases}$$

то формулы (5.13) и (5.14) можно записать в виде

$$\mathsf{P}(a \leqslant \xi \leqslant b) = \mathsf{M}\left( \int\limits_a^b p_\xi(x \mid \eta)\, dx \right)$$

и в виде (5.3) соответственно. Для условного математического ожидания абсолютно непрерывных величин сохраняются свойства (5.7) — (5.10).

**Задачи к главе 5**

**1.** Найти математическое ожидание величины $\tau$, определенной в задаче 14 гл. 4.

**2.** Обозначим $\xi$ номер испытания, в котором появился нужный ключ (см. пример 3 из § 6 гл. 1). Найти $\mathsf{M}\xi$.

**3.** Решить задачу 2 в случае с возвращением ключей.

**4.** Найти $\mathsf{M}\tau_m$ величины $\tau_m$, определенной в задаче 16 гл. 4.

**5.** В задаче 8 гл. 2 обозначим $\tau$ время свободного пробега молекулы. Найти $\mathsf{M}\tau$, $\mathsf{D}\tau$.

**6.** Найти $\mathsf{M}(\xi_1+\xi_2)$ и $\mathsf{D}(\xi_1+\xi_2)$, где $\xi_1$, $\xi_2$ определены в задаче 8 гл. 4.

**7.** Пусть $\xi$ — число комбинаций НУ в $n+1$ испытаниях схемы Бернулли. Найти $\mathsf{M}\xi$, $\mathsf{D}\xi$.

**8.** Из 100 карточек с числами 00, 01, 02, ..., 98, 99 наудачу вынимается одна. Пусть $\eta_1$, $\eta_2$ соответственно сумма и произведение цифр на вынутой карточке. Найти $\mathsf{M}\eta_1$, $\mathsf{M}\eta_2$, $\mathsf{D}\eta_1$, $\mathsf{D}\eta_2$.

**9.** Для величин $\xi_1$, $\xi_2$, определенных в задаче 9 гл. 4, найти $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$, $\operatorname{cov}(\xi_1, \xi_2)$.

**10.** Совместное распределение величин $\xi_1$, $\xi_2$ определяется формулами $\mathsf{P}(\xi_1=0,\ \xi_2=1)=\mathsf{P}(\xi_1=0,\ \xi_2=-1)=\mathsf{P}(\xi_1=1,\ \xi_2=0)=\mathsf{P}(\xi_1=-1,\ \xi_2=0)=\frac{1}{4}$. Найти $\mathsf{M}\xi_1$, $\mathsf{M}\xi_2$, $\mathsf{D}\xi_1$, $\mathsf{D}\xi_2$, $\operatorname{cov}(\xi_1, \xi_2)$. Являются ли $\xi_1$, $\xi_2$ независимыми величинами?

**11.** Случайные величины $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ независимы; $\mathsf{D}\xi_i=\sigma^2$. Найти: а) *коэффициент корреляции величин* $\xi_1+\xi_2$, $\xi_3+\xi_4+\xi_5$, б) $\xi_1+\xi_2+\xi_3$, $\xi_3+\xi_4+\xi_5$.

**12.** Пусть $(\xi_1, \xi_2, \ldots, \xi_N)$ — дискретный случайный вектор с полиномиальным распределением (3.2.7). Найти $\mathsf{M}\xi_k$, $\operatorname{cov}(\xi_k, \xi_l)$, $k, l=1, 2, \ldots, N$.

**13.** В задаче 11 гл. 1 обозначим $\mu_0$ — число пустых ящиков. Найти $\mathsf{P}(\mu_0=0)$.

**14.** Для величины $\mu_0$, определенной в задаче 11 гл. 1 и задаче 13, найти $\mathsf{M}\mu_0$, $\mathsf{D}\mu_0$. Найти асимптотические формулы при $N\to\infty$, $\frac{n}{N}=\alpha=\mathrm{const}$.

**15.** По $n$ конвертам случайно разложено $n$ писем различным адресатам. Найти вероятность того, что хотя бы одно письмо попадет своему адресату. Найти предел этой вероятности при $n\to\infty$.

**16.** В $N$ ящиков *случайно и независимо* друг от друга бросают шары, пока не останется пустых ящиков. Найти $\mathsf{M}\nu$, где $\nu$ — число брошенных шаров.

**17.** Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы и $\mathsf{D}\xi_i=b^2$, $\mathsf{M}\xi_i=a$. Найти математическое ожидание величин

$$\eta=\frac{1}{n}\sum_{i=1}^{n}\xi_i, \qquad \theta=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\eta)^2.$$

**18.** Пусть $\xi$ принимает значения 0, 1, 2, ... Доказать, что

$$\mathsf{M}\xi = \sum_{m=1}^{\infty} \mathsf{P}(\xi \geqslant m).$$

**19*.** В задаче о случайном блуждании, рассмотренной в § 5 гл. 2, обозначим $\tau$ время до поглощения в точках 0 или $n$. Найти $\mathsf{M}_k = \mathsf{M}(\tau \mid \xi_0 = k)$, где $\xi_t$ — координата частицы в момент времени $t$. Указание: воспользоваться формулой (5.3) для составления уравнения в конечных разностях для $\mathsf{M}_k$.

**20.** Пусть $\mathsf{M}\xi_1 = a_1$, $\mathsf{M}\xi_2 = a_2$, $\mathsf{D}\xi_1 = \sigma_{11} > 0$, $\mathsf{D}\xi_2 = \sigma_{22} > 0$, $\operatorname{cov}(\xi_1, \xi_2) = \sigma_{12}$. Найти $\alpha$ и $\beta$, при которых выражение $\mathsf{M}(\xi_2 - \alpha\xi_1 - \beta)^2$ минимально.

ГЛАВА 6

# ЗАКОН БОЛЬШИХ ЧИСЕЛ

## § 1. Неравенство Чебышева

Числовые характеристики случайных величин, введенные в предыдущей главе, позволяют давать некоторые оценки распределений случайных величин.

Теорема 1.1. *Пусть* $\xi = \xi(\omega) \geqslant 0$ *при любом* $\omega \in \Omega$. *Если* $\mathsf{M}\xi$ *существует, то при любом* $\varepsilon > 0$

$$\mathsf{P}(\xi \geqslant \varepsilon) \leqslant \frac{\mathsf{M}\xi}{\varepsilon}. \qquad (1.1)$$

Доказательство. Проведем доказательство в случае, когда $\xi$ задана в абсолютно непрерывном вероятностном пространстве (см. § 6 гл. 1). По определению математического ожидания имеем

$$\mathsf{M}\xi = \int \underset{\Omega}{\dots} \int \xi(u_1, \dots, u_n)\, \pi(u_1, \dots, u_n)\, du_1\, du_2 \dots du_n.$$

Пусть

$$\Omega_\varepsilon = \{(u_1, \dots, u_n) \colon \xi(u_1, \dots, u_n) \geqslant \varepsilon\} \subseteq \Omega.$$

Введем случайную величину

$$\eta = \begin{cases} 0, & \text{если} \quad (u_1, \dots, u_n) \in \Omega \setminus \Omega_\varepsilon, \\ \varepsilon, & \text{если} \quad (u_1, \dots, u_n) \in \Omega_\varepsilon. \end{cases}$$

При любом $(u_1, \dots, u_n) \in \Omega$ имеем

$$\xi \geqslant \eta.$$

Умножим обе части этого неравенства на $\pi(u_1, u_2, \dots, u_n)$ и проинтегрируем по $\Omega$. Получим, что $\mathsf{M}\xi \geqslant \mathsf{M}\eta$. Отсюда следует утверждение теоремы, так как

$$\mathsf{M}\eta = \varepsilon \mathsf{P}(\Omega_\varepsilon) = \varepsilon \mathsf{P}(\xi \geqslant \varepsilon).$$

Доказанная теорема позволяет легко получить неравенство Чебышева.

Теорема 1.2 (неравенство Чебышева). *Если случайная величина* $\xi$ *имеет дисперсию, то при любом* $\varepsilon > 0$

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) \leqslant \frac{\mathsf{D}\xi}{\varepsilon^2}.$$

Доказательство. Случайная величина $\eta = (\xi - \mathsf{M}\xi)^2 \geqslant 0$ при всех $\omega \in \Omega$ и $\mathsf{M}\eta = \mathsf{M}(\xi - \mathsf{M}\xi)^2 = \mathsf{D}\xi$ конечно. Следовательно, можно воспользоваться неравенством (1.1). Таким образом,

$$\mathsf{P}(|\xi - \mathsf{M}\xi| \geqslant \varepsilon) = \mathsf{P}(\eta \geqslant \varepsilon^2) \leqslant \frac{\mathsf{M}\eta}{\varepsilon^2} = \frac{\mathsf{D}\xi}{\varepsilon^2}.$$

Неравенство Чебышева 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. Пусть еще $\mathsf{D}\delta_k = b^2$. За значение неизвестной величины $a$ принимают обычно среднее арифметическое результатов измерений. Тогда ошибка в определении числа $a$ будет равна

$$\eta_n = \frac{\delta_1 + \delta_2 + \ldots + \delta_n}{n}$$

и

$$\mathsf{D}\eta_n = \frac{1}{n^2}(\mathsf{D}\delta_1 + \ldots + \mathsf{D}\delta_n) = \frac{b^2}{n}, \qquad \mathsf{M}\eta_n = 0.$$

Предположим, что нам нужно, чтобы ошибка $\eta_n$ не превосходила $\Delta$ с достаточно большой вероятностью. Например,

$$\mathsf{P}(|\eta_n| < \Delta) > 0{,}99.$$

Это неравенство можно записать в эквивалентном виде

$$\mathsf{P}(|\eta_n| \geqslant \Delta) \leqslant 0{,}01. \tag{1.2}$$

По неравенству Чебышева имеем

$$\mathsf{P}(|\eta_n| \geqslant \Delta) \leqslant \frac{\mathsf{D}\eta_n}{\Delta^2} = \frac{b^2}{n\Delta^2},$$

Следовательно, (1.2) будет выполнено, если

$$\frac{b^2}{n\Delta^2} \leqslant 0{,}01 \text{ или } n > 100\frac{b^2}{\Delta^2}.$$

Таким образом, мы получили оценку числа измерений, необходимого для получения заданной точности. В рассматриваемой задаче оценка для $n$ является завышенной. Ее можно улучшить, если воспользоваться тем, что $\eta_n$ является суммой независимых случайных величин. Будет показано (см. теорему 2.1 гл. 8), что при больших $n$ величина $\eta_n$ имеет распределение, близкое к нормальному. Однако если о случайной величине ничего не известно, кроме математического ожидания и дисперсии, то оценку, которую дает неравенство Чебышева, улучшить нельзя. Укажем распределение случайной величины, для которой неравенство Чебышева при данном $b^2 = \mathsf{D}\xi$ и $\varepsilon > b$ обращается в равенство. Пусть

$$\mathsf{P}(\xi = 0) = 1 - \frac{b^2}{\varepsilon^2}, \quad \mathsf{P}(\xi = -\varepsilon) = \mathsf{P}(\xi = \varepsilon) = \frac{b^2}{2\varepsilon^2}.$$

Тогда

$$\mathsf{M}\xi = 0, \ \mathsf{D}\xi = \mathsf{M}\xi^2 = (-\varepsilon)^2\frac{b^2}{2\varepsilon^2} + \varepsilon^2\frac{b^2}{2\varepsilon^2} = b^2$$

и

$$\mathsf{P}(|\xi| \geqslant \varepsilon) = \frac{b^2}{\varepsilon^2}.$$

Нетрудно получить еще ряд полезных неравенств типа неравенства Чебышева. Пусть $f(x)$ — неубывающая неотрицательная функция. Если существует $\mathsf{M}f(\xi)$, то при любом $\varepsilon > 0$

$$\mathsf{P}(\xi \geqslant \varepsilon) \leqslant \frac{\mathsf{M}f(\xi)}{f(\varepsilon)}. \qquad (1.3)$$

Доказательство проводится так же, как доказательство теоремы 1.1. Из (1.3) нетрудно получить еще два неравенства.

Если $\mathsf{M}e^{\lambda\xi}$ конечно, то

$$\mathsf{P}(\xi \geqslant x) \leqslant e^{-\lambda x}\mathsf{M}e^{\lambda\xi}, \quad \lambda > 0. \qquad (1.4)$$

Если $\mathsf{M}|\xi|^m$ конечно, то

$$\mathsf{P}(|\xi| \geqslant x) \leqslant x^{-m}\mathsf{M}|\xi|^m, \quad x > 0; \quad m > 0. \qquad (1.5)$$

## § 2. Закон больших чисел

Во введении был отмечен экспериментальный факт, состоящий в том, что в некоторой длинной серии опытов частота появления события $A$ сближается с определенным числом, которое можно рассматривать как вероятность события $A$. В математической модели серии опытов этот факт будет доказан. Сначала докажем более общую теорему.

Теорема 2.1. *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *попарно независимы и*

$$\lim_{n\to\infty} \frac{1}{n^2}\sum_{k=1}^{n} \mathsf{D}\xi_k = 0, \tag{2.1}$$

*то для любого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} - \frac{\mathsf{M}\xi_1+\mathsf{M}\xi_2+\ldots+\mathsf{M}\xi_n}{n}\right| < \varepsilon\right) = 1.$$

Доказательство. Положим

$$\eta_n = \frac{\xi_1+\xi_2+\ldots+\xi_n}{n}.$$

Утверждение теоремы равносильно тому, что при любом $\varepsilon > 0$

$$\lim_{n\to\infty} \mathsf{P}(|\eta_n - \mathsf{M}\eta_n| \geqslant \varepsilon) = 0. \tag{2.2}$$

Так как случайные величины $\xi_1, \ldots, \xi_n$ попарно независимы, то

$$\mathsf{D}\eta_n = \frac{1}{n^2}\sum_{k=1}^{n} \mathsf{D}\xi_k. \tag{2.3}$$

По неравенству Чебышева

$$\mathsf{P}(|\eta_n - \mathsf{M}\eta_n| \geqslant \varepsilon) \leqslant \frac{\mathsf{D}\eta_n}{\varepsilon^2}.$$

Отсюда, воспользовавшись (2.3) и (2.1), получим (2.2). Теорема доказана.

Случайные величины $\xi_1, \ldots, \xi_n, \ldots$ называют *некоррелированными*, если $\operatorname{cov}(\xi_i, \xi_j) = 0$ при любых $i, j, i \neq j$.

В условии теоремы 2.1 можно вместо попарной независимости величин $\xi_1, \ldots, \xi_n, \ldots$ потребовать, чтобы они были некоррелированными, так как для некоррелированных величин сохранится формула (2.3). Если для величин $\xi_1, \ldots, \xi_n, \ldots$ выполнено утверждение теоремы 2.1, то говорят, что к ним применим закон больших чисел.

Отметим некоторые частные случаи этой теоремы.

Теорема 2.2 (теорема Чебышева). *Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ попарно независимы и*

$$\mathsf{D}\xi_k \leqslant C, \quad k = 1, 2, \ldots, \tag{2.4}$$

*где $C$ — некоторая постоянная, то при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\xi_1 + \ldots + \xi_n}{n} - \frac{\mathsf{M}\xi_1 + \ldots + \mathsf{M}\xi_n}{n}\right| < \varepsilon\right) = 1.$$

Теорема 2.2 следует из теоремы 2.1, так как из условия (2.4) следует (2.1).

Теорема 2.3. *Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ одинаково распределены, попарно независимы и имеют конечные дисперсии, то при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\xi_1 + \ldots + \xi_n}{n} - a\right| < \varepsilon\right) = 1,$$

*где $a = \mathsf{M}\xi_k$.*

Теорема 2.3 следует из теоремы 2.2. Действительно, $\mathsf{D}\xi_k$, $k = 1, 2, \ldots$, существуют и равны между собой. Следовательно, выполнено условие (2.4). В гл. 8 теорема 2.2 при помощи характеристических функций будет доказана без предположения о существовании дисперсии.

Теорема 2.4 (теорема Бернулли). *Пусть $\mu_n$ — число успехов в $n$ испытаниях Бернулли и $p$ — вероятность успеха в отдельном испытании. Тогда при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) = 1.$$

Доказательство. Воспользуемся для $\mu_n$ представлением (4.4.18):

$$\mu_n = \xi_1 + \xi_2 + \ldots + \xi_n,$$

где $\xi_k$, $k=1, 2, \ldots, n$, независимы и $\mathsf{P}(\xi_k=1)=p$, $\mathsf{P}(\xi_k=0)=1-p=q$. Дисперсии величин $\xi_k$, очевидно, существуют и $\mathsf{M}\xi_k=p$.

Таким образом, доказываемая теорема сразу следует из теоремы 2.3.

Схема Бернулли является математической моделью серии опытов, повторяющихся в неизменных условиях. В каждом опыте может произойти событие $A$, которое мы назвали успехом. Согласно теореме Бернулли частота $\mu_n/n$ наступления события $A$ сближается с вероятностью $p$. Этот же факт установлен экспериментально.

Математической моделью последовательности из $n$ измерений неизвестной величины $a$ является случайный вектор $(\xi_1, \xi_2, \ldots, \xi_n)$, где $\xi_k$, $k=1, 2, \ldots, n$, независимы, одинаково распределены,

$$\mathsf{M}\xi_k=a, \qquad \mathsf{D}\xi_k=b^2.$$

По теореме 2.3 среднее арифметическое

$$\frac{\xi_1+\ldots+\xi_n}{n}$$

при больших $n$ мало отличается от измеряемой величины $a$ с вероятностью, близкой к 1.

### Задачи к главе 6

1. 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 $\Delta$, если

$$\mathsf{P}\left(\left|\frac{x_1+x_2+\ldots+x_{10}}{10}-a\right|<\Delta\right)=0{,}99.$$

2. Пусть случайная величина $\xi$ имеет нормальное распределение с параметрами $(a, \sigma)$. Оценить по неравенству Чебышева

$$\mathsf{P}(|\xi-a|>2\sigma).$$

Сравнить с точным значением этой вероятности.

3. Доказать неравенства (1.3), (1.4), (1.5).

**4.** Применим ли закон больших чисел к последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_k, \ldots$, если

$$\mathsf{P}(\xi_k = \sqrt{k}) = \mathsf{P}(\xi_k = -\sqrt{k}) = \frac{1}{2\sqrt{k}}, \quad \mathsf{P}(\xi_k = 0) = 1 - \frac{1}{\sqrt{k}}?$$

**5.** Применим ли закон больших чисел к последовательности независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_k, \ldots$, если они нормально распределены с $\mathsf{M}\xi_k = 0$, $\mathsf{D}\xi_k = ck^\alpha$, $c > 0$, $\alpha > 0$ — некоторые постоянные?

**6.** Доказать, что к последовательности $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ применим закон больших чисел, если $|\operatorname{cov}(\xi_k, \xi_l)| \leqslant C$ для всех $k, l = 1, 2, \ldots$ и $\operatorname{cov}(\xi_k, \xi_l) \to 0$ при $|k - l| \to \infty$.

ГЛАВА 7

# ПРОИЗВОДЯЩИЕ И ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

## § 1. Производящие функции. Определение и свойства

Среди дискретных случайных величин особенно часто используются величины, принимающие только значения $k=0, 1, 2, \ldots$,

$$\sum_{k=0}^{\infty} \mathsf{P}(\xi=k)=1.$$

Такие случайные величины называют *целочисленными*. Производящие функции являются удобным аппаратом исследования распределений целочисленных случайных величин. Производящей функцией $A(x)$ последовательности действительных чисел $a_0, a_1, a_2, \ldots$ называется степенной ряд

$$A(x)=a_0+a_1x+a_2x^2+\ldots, \tag{1.1}$$

если он сходится при $|x|<R$, где $R>0$, $x$ — действительная переменная. Производящую функцию

$$\varphi(x)=\sum_{k=0}^{\infty} p_k x^k \tag{1.2}$$

последовательности чисел $p_k$, $k=0, 1, 2, \ldots$, такой, что $p_k \geqslant 0$ и

$$\sum_{k=0}^{\infty} p_k=1, \tag{1.3}$$

будем называть *вероятностной производящей функцией*. Если $p_k=\mathsf{P}(\xi=k)$, $k=0, 1, 2, \ldots$, то мы иногда будем вместо $\varphi(x)$ писать $\varphi_\xi(x)$. По теореме 5.1.1

$$\varphi_\xi(x)=\sum_{k=0}^{\infty} x^k \mathsf{P}(\xi=k)=\mathsf{M}x^\xi. \tag{1.4}$$

Теорема 1.1. *Пусть* $\varphi(x)$ — *вероятностная производящая функция, определенная формулой* (1.2). *Тогда*

1°. $\varphi(x)$ *определена в каждой точке отрезка* $[-1, 1]$.

2°. $\varphi(1)=1$.

3°. *Соответствие, устанавливаемое формулой* (1.2), *между множеством производящих функций* $\varphi(x)$ *и множеством распределений* $\{p_k\}$ *является взаимно однозначным.*

Доказательство. Первое утверждение теоремы следует из того, что степенной ряд в (1.2) мажорируется при $|x| \leqslant 1$ сходящимся рядом в левой части (1.3). Равенство $\varphi(1)=1$ совпадает с (1.3). Третье утверждение теоремы является следствием единственности разложения функции в ряд Тейлора. Теорема доказана.

Используя формулы для коэффициентов ряда Тейлора, можно явно указать распределение $\{p_k\}$, соответствующее вероятностной производящей функции $\varphi(x)$. Имеем

$$p_k = \frac{1}{k!}\varphi^{(k)}(0), \qquad k=0,\ 1,\ 2,\ \ldots$$

Пример 1. Пусть случайная величина $\xi$ имеет биномиальное распределение

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

По формуле (1.4)

$$\varphi_\xi(x) = \sum_{k=0}^{n} C_n^k p^k q^{n-m} x^k = (px+q)^n.$$

Таким образом,

$$\varphi_\xi(x)=(px+q)^n. \tag{1.5}$$

По производящей функции легко определить моменты случайной величины. Особенно просто находятся математические ожидания

$$\mathsf{M}\xi(\xi-1)\ldots(\xi-k+1). \tag{1.6}$$

Математическое ожидание (1.6) будем называть $k$-м факториальным моментом. Зная факториальные моменты 1-го и 2-го порядков, можно найти дисперсию по следующей формуле:

$$\mathsf{D}\xi = \mathsf{M}\xi(\xi-1) + \mathsf{M}\xi - (\mathsf{M}\xi)^2. \qquad (1.7)$$

Теорема 1.2. *Если конечен $k$-й факториальный момент, то существует левосторонняя производная $\varphi_\xi^{(k)}(1)$ и*

$$\mathsf{M}\xi(\xi-1)\ldots(\xi-k+1) = \varphi_\xi^{(k)}(1); \qquad (1.8)$$

*в частности,*

$$\mathsf{M}\xi = \varphi_\xi'(1). \qquad (1.9)$$

Доказательство. При любом $|x|<1$ функцию $\varphi(x)$ можно дифференцировать сколько угодно раз. Таким образом, при любом $k$ определена $k$-я производная

$$\varphi^{(k)}(x) = \sum_{m=0}^{\infty} m(m-1)\ldots(m-k+1)x^{m-k}\mathsf{P}(\xi=m). \qquad (1.10)$$

По условию теоремы конечен $k$-й факториальный момент

$$\mathsf{M}\xi(\xi-1)\ldots(\xi-k+1) = \sum_{m=0}^{\infty} m(m-1)\ldots$$
$$\ldots(m-k+1)\mathsf{P}(\xi=m),$$

являющийся суммой ряда (1.10) в точке $x=1$. Следовательно, по теореме Абеля $\varphi^{(k)}(x)$ непрерывна в точке $x=1$ и

$$\lim_{x\to 1-0}\varphi^{(k)}(x) = \sum_{m=0}^{\infty} m(m-1)\ldots(m-k+1)\mathsf{P}(\xi=m) =$$
$$=\varphi^{(k)}(1).$$

Теорема доказана.

Пример 2. Найдем $\mathsf{M}\xi$ и $\mathsf{D}\xi$ случайной величины $\xi$, распределенной по биномиальному закону. Дифференцируя (1.5) два раза по $x$, получим

$$\varphi_\xi'(x) = np(px+q)^{n-1},$$
$$\varphi_\xi''(x) = n(n-1)p^2(px+q)^{n-2}.$$

Отсюда, воспользовавшись (1.8), при $x=1$ получим

$$\mathsf{M}\xi=\varphi_\xi'(1)=np(p+q)^{n-1}=np,$$

$$\mathsf{M}\xi(\xi-1)=\varphi_\xi''(1)=n(n-1)p^2(p+q)^{n-2}=n(n-1)p^2.$$

По формуле (1.7)

$$\mathsf{D}\xi=n(n-1)p^2+np-n^2p^2=np(1-p).$$

Легко вычисляется $k$-й факториальный момент биномиального распределения. Так как

$$\varphi_\xi^{(k)}(x)=n(n-1)\ldots(n-k+1)p^k(px+q)^{n-k},$$

то

$$\mathsf{M}\xi(\xi-1)\ldots(\xi-k+1)=n(n-1)\ldots(n-k+1)p^k. \tag{1.11}$$

Применение производящих функций к изучению сумм независимых целочисленных случайных величин основано на следующей теореме.

Теорема 1.3. *Если случайные величины* $\xi_1, \xi_2, \ldots$ $\ldots, \xi_n$ *независимы, то*

$$\varphi_{\xi_1+\ldots+\xi_n}(x)=\varphi_{\xi_1}(x)\cdot\ldots\cdot\varphi_{\xi_n}(x).$$

Доказательство. Воспользуемся представлением производящей функции в виде математического ожидания (1.4). Тогда

$$\varphi_{\xi_1+\ldots+\xi_n}(x)=\mathsf{M}x^{\xi_1+\ldots+\xi_n}=\mathsf{M}(x^{\xi_1}\cdot\ldots\cdot x^{\xi_n}). \tag{1.12}$$

Случайные величины $x^{\xi_1}$, $x^{\xi_2}$, $\ldots$, $x^{\xi_n}$ независимы как функции от независимых случайных величин. Следовательно,

$$\mathsf{M}(x^{\xi_1}\cdot\ldots\cdot x^{\xi_n})=\mathsf{M}x^{\xi_1}\cdot\ldots\cdot\mathsf{M}x^{\xi_n}.$$

Отсюда и из (1.12) следует утверждение теоремы, так как $\mathsf{M}x^{\xi_k}=\varphi_{\xi_k}(x)$.

Пример 3. Найдем производящую функцию биномиального распределения при помощи теоремы 1.3. Число успехов $\mu_n$ в схеме Бернулли имеет биномиальное распределение. Представим $\mu_n$ в виде суммы независимых слагаемых (см. (4.4.18)):

$$\mu_n=\xi_1+\xi_2+\ldots+\xi_n,$$

где $\mathsf{P}(\xi_k=1)=1-\mathsf{P}(\xi_k=0)=p$, $k=1, 2, \ldots, n$. По теореме 1.3

$$\varphi_{\mu_n}(x)=\varphi_{\xi_1}(x)\varphi_{\xi_2}(x)\ldots\varphi_{\xi_n}(x).$$

Производящие функции слагаемых получим по формуле (1.4):

$$\varphi_{\xi_k}(x)=xp+x^0q=px+q, \qquad k=1, \ldots, n.$$

Таким образом,

$$\varphi_{\mu_n}(x)=(px+q)^n.$$

Найдем производящую функцию случайного числа случайных слагаемых.

Теорема 1.4. *Пусть целочисленные величины* $\nu$, $\xi_1, \xi_2, \ldots, \xi_n$ *независимы при любом* $n=1, 2, 3, \ldots$; $\xi_k$, $k=1, 2, \ldots$, *одинаково распределены. Положим*

$$\zeta_\nu=\xi_1+\xi_2+\ldots+\xi_\nu, \qquad \zeta_0=0.$$

*Тогда*

$$\varphi_{\zeta_\nu}(x)=\varphi_\nu[\varphi_{\xi_1}(x)]. \tag{1.13}$$

Доказательство. Вероятность события $(\zeta_\nu=m)$ представим в виде

$$\mathsf{P}(\zeta_\nu=m)=\sum_{k=0}^{\infty}\mathsf{P}(\nu=k,\ \zeta_\nu=m).$$

Так как $(\nu=k,\ \zeta_\nu=m)=(\nu=k,\ \zeta_k=m)$ и события $(\nu=k)$, $(\zeta_k=m)$ независимы, то

$$\mathsf{P}(\nu=k,\ \zeta_\nu=m)=\mathsf{P}(\nu=k,\ \zeta_k=m)=$$
$$=\mathsf{P}(\nu=k, \xi_1+\ldots+\xi_k=m)=\mathsf{P}(\nu=k)\,\mathsf{P}(\xi_1+\ldots+\xi_k=m).$$

Тогда

$$\mathsf{P}(\zeta_\nu=m)=\sum_{k=0}^{\infty}\mathsf{P}(\nu=k)\mathsf{P}(\xi_1+\xi_2+\ldots+\xi_k=m).$$

Умножив обе части этого равенства на $x^m$ и просуммировав по $m=0, 1, 2, \ldots$, получим

$$\varphi_{\zeta_\nu}(x)=\sum_{k=0}^{\infty}\mathsf{P}(\nu=k)\left(\sum_{m=0}^{\infty}x^m\mathsf{P}(\xi_1+\ldots+\xi_k=m)\right).$$

Ряд в круглых скобках является производящей функцией распределения суммы $\xi_1 + \ldots + \xi_k$. Так как слагаемые $\xi_l$, $l = 1, \ldots, k$, независимы и одинаково распределены, то

$$\sum_{m=0}^{\infty} x^m \mathsf{P}(\xi_1 + \ldots + \xi_k = m) = \varphi_{\xi_1 + \ldots + \xi_k}(x) = [\varphi_{\xi_1}(x)]^k.$$

Следовательно,

$$\varphi_{\zeta_\nu}(x) = \sum_{k=0}^{\infty} \mathsf{P}(\nu = k)[\varphi_{\xi_1}(x)]^k = \varphi_\nu[\varphi_{\xi_1}(x)].$$

Теорема доказана.

Если воспользоваться условными математическими ожиданиями и формулой полного математического ожидания (5.5.3), то можно дать более короткое доказательство теоремы 1.4. Действительно, по формуле (5.5.3)

$$\varphi_{\zeta_\nu}(x) = \mathsf{M}x^{\zeta_\nu} = \mathsf{M}(\mathsf{M}(x^{\zeta_\nu} \mid \nu)).$$

Так как

$$\mathsf{M}(x^{\zeta_\nu} \mid \nu) = \mathsf{M}(x^{\xi_1} \ldots x^{\xi_\nu} \mid \nu) = [\varphi_{\xi_1}(x)]^\nu,$$

то

$$\varphi_{\zeta_\nu}(x) = \mathsf{M}[(\varphi_{\xi_1}(x))^\nu] = \varphi_\nu[\varphi_{\xi_1}(x)].$$

В намеченном доказательстве основная трудность заключается в проверке равенства

$$\mathsf{M}(x^{\xi_1 + \ldots + \xi_\nu} \mid \nu) = (\mathsf{M}x^{\xi_1})^\nu$$

или эквивалентного ему равенства

$$\mathsf{M}(x^{\xi_1 + \ldots + \xi_\nu} \mid \nu = k) = \mathsf{M}x^{\xi_1 + \ldots + \xi_k}.$$

Интуитивно оно не вызывает сомнений.

Установим еще свойство непрерывности соответствия множества производящих функций множеству распределений.

Теорема 1.5. *Пусть при любом фиксированном $n$, $n = 1, 2, \ldots$, последовательность $\{p_k(n)\}$ является* *распределением вероятностей, т. е.*

$$p_k(n) \geqslant 0, \qquad \sum_{k=0}^{\infty} p_k(n) = 1. \tag{1.14}$$

*Для того чтобы при любом фиксированном* $k$

$$\lim_{n \to \infty} p_k(n) = p_k \tag{1.15}$$

*и*

$$\sum_{k=0}^{\infty} p_k = 1, \tag{1.16}$$

*необходимо и достаточно, чтобы при любом* $x \in [0, 1)$

$$\lim_{n \to \infty} \varphi_n(x) = \varphi(x), \tag{1.17}$$

*где*

$$\varphi_n(x) = \sum_{k=0}^{\infty} p_k(n) x^k, \qquad \varphi(x) = \sum_{k=0}^{\infty} p_k x^k, \qquad \varphi(1) = 1.$$

Доказательство. Пусть выполнено (1.15) и (1.16). Представим разность $\varphi_n(x) - \varphi(x)$ в виде

$$\varphi_n(x) - \varphi(x) = \\ = \sum_{k=0}^{N} (p_k(n) - p_k) x^k + \sum_{k=N+1}^{\infty} p_k(n) x^k - \sum_{k=N+1}^{\infty} p_k x^k,$$

где $N$ — некоторое целое число. Пусть $x \in [0, 1)$ фиксирован. Докажем (1.17). Так как $0 \leqslant p_k(n) \leqslant 1$, $0 \leqslant p_k \leqslant 1$, то для любого фиксированного $\varepsilon > 0$ можно подобрать $N$ так, чтобы при любом $n$

$$\sum_{k=N+1}^{\infty} p_k(n) x^k \leqslant \sum_{k=N+1}^{\infty} x^k < \frac{\varepsilon}{3}, \qquad \sum_{k=N+1}^{\infty} p_k x^k < \frac{\varepsilon}{3}.$$

Тогда при данном $N$

$$|\varphi_n(x) - \varphi(x)| \leqslant \sum_{k=0}^{N} |p_k(n) - p_k| x^k + \frac{2}{3}\varepsilon.$$

Сумма в правой части этого неравенства может быть сделана меньше $\varepsilon/3$ при достаточно больших $n$, так как содержит конечное число стремящихся к нулю слагаемых. Равенство $\varphi(1)=1$ следует из (1.16).

Пусть теперь выполнено (1.17). Докажем (1.15) от противного. Предположим, что (1.15) не выполнено. Тогда найдутся две подпоследовательности $n'_m$ и $n''_m$ такие, что

$$\lim_{n'_m\to\infty} p_k(n'_m)=\tilde{p}_k, \quad \lim_{n''_m\to\infty} p_k(n''_m)=\tilde{\tilde{p}}_k, \tag{1.18}$$

причем $\{\tilde{p}_k\}$ и $\{\tilde{\tilde{p}}_k\}$ не совпадают. Тогда по доказанной части теоремы из (1.18) следует, что

$$\lim_{n'_m\to\infty} \varphi_{n'_m}(x)=\tilde{\varphi}(x)=\sum_{k=0}^{\infty}\tilde{p}_k x^k,$$

$$\lim_{n''_m\to\infty} \varphi_{n''_m}(x)=\tilde{\tilde{\varphi}}(x)=\sum_{k=0}^{\infty}\tilde{\tilde{p}}_k x^k$$

и $\tilde{\varphi}(x)\neq\tilde{\tilde{\varphi}}(x)$. Это невозможно, так как предел (1.17) существует. Теорема доказана.

Пример 4. Пусть $\mu_n$ — число успехов в $n$ испытаниях Бернулли и $p_n$ — вероятность успеха в одном испытании. Будем предполагать, что существует $\lim\limits_{n\to\infty} np_n=\lambda$, $0<\lambda<\infty$. Воспользуемся теоремой 1.5 для вычисления $\lim\limits_{n\to\infty}\mathbf{P}(\mu_n=m)$. Положим $\lambda_n=np_n$. По формуле (1.5)

$$\varphi_{\mu_n}(x)=\left(\frac{\lambda_n}{n}x+1-\frac{\lambda_n}{n}\right)^n=\left[1+\frac{\lambda_n}{n}(x-1)\right]^n.$$

Следовательно,

$$\lim_{n\to\infty}\varphi_{\mu_n}(x)=e^{\lambda(x-1)}=\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}e^{-\lambda}\cdot x^m.$$

Отсюда по теореме 1.5

$$\lim_{n\to\infty}\mathbf{P}(\mu_n=m)=\frac{\lambda^m}{m!}e^{-\lambda}.$$

Таким образом, получили новое доказательство теоремы Пуассона.

## § 2. Характеристические функции. Определения и свойства

Производящие функции определены для целочисленных случайных величин. Для исследования распределений произвольных случайных величин вводятся характеристические функции. Комплекснозначной случайной величиной будем называть функцию $\xi_1(\omega)+i\xi_2(\omega)$, где $\omega\in\Omega$, $(\xi_1, \xi_2)$ — случайный вектор. По определению положим

$$\mathsf{M}(\xi_1+i\xi_2)=\mathsf{M}\xi_1+i\,\mathsf{M}\xi_2. \tag{2.1}$$

Для математического ожидания от комплекснозначной случайной величины легко проверяются свойства 1°—4°, приведенные в § 2 гл. 5. В дальнейшем мы ими будем пользоваться без дополнительных оговорок. Понятие независимости для комплекснозначных случайных величин не будет введено и свойство 5° использоваться не будет.

*Характеристической функцией* действительной случайной величины $\xi$ называется

$$f_\xi(t)=\mathsf{M}e^{it\xi}, \tag{2.2}$$

где $t$ — действительное число, $-\infty<t<\infty$.

Если случайная величина $\xi$ дискретна, то по теореме 5.1.1

$$\mathsf{M}(\cos t\xi)=\sum_{k=1}^{\infty}\cos(tx_k)\,\mathsf{P}(\xi=x_k),$$

$$\mathsf{M}(\sin t\xi)=\sum_{k=1}^{\infty}\sin(tx_k)\,\mathsf{P}(\xi=x_k).$$

Отсюда и из (2.1) получим

$$f_\xi(t)=\sum_{k=1}^{\infty}e^{itx_k}\,\mathsf{P}(\xi=x_k). \tag{2.3}$$

Используя теорему 5.1.2, для характеристической функции абсолютно непрерывной величины $\xi$ будем иметь

$$f_\xi(t)=\int_{-\infty}^{\infty}e^{itx}\,p_\xi(x)\,dx. \tag{2.4}$$

Если случайная величина $\xi$ определена на дискретном или абсолютно непрерывном вероятностном пространстве, то по формулам (1.4) и (1.5) соответственно получим

$$f_\xi(t)=\sum_{k=1}^{\infty} e^{it\xi(\omega_k)}p_k, \qquad \mathsf{P}(\{\omega_k\})=p_k \tag{2.5}$$

и

$$f_\xi(t)=\int\limits_{\Omega}\ldots\int e^{it\xi(u_1,\ldots,u_n)}\,\pi(u_1,\ldots,u_n)\,du_1\ldots du_n. \tag{2.6}$$

Теорема 2.1. *Пусть $f_\xi(t)$ — характеристическая функция случайной величины $\xi$. Тогда*

1°. *$f_\xi(t)$ определена при любом $t\in(-\infty,\infty)$.*

2°. *$f_\xi(0)=1$, $|f_\xi(t)|\leqslant 1$.*

3°. *Если $\eta=a\xi+b$, где $a$ и $b$ постоянные, то $f_\eta(t)=$ $=e^{itb}f_\xi(ab)$.*

4°. *Соответствие, устанавливаемое формулой* (2.2), *между множеством характеристических функций $f_\xi(t)$ и множеством функций распределения, является взаимно однозначным.*

Докажем 1°—3°. Так как при любом действительном $t$

$$|e^{itx}|\leqslant 1, \qquad -\infty<x<\infty,$$

то доказательство первых двух утверждений легко следует из формул (2.3)—(2.6). Третье утверждение получим из следующих равенств:

$$f_\eta(t)=\mathsf{M}e^{it(a\xi+b)}=e^{itb}\mathsf{M}e^{it\xi a}=e^{itb}f_\xi(at).$$

Для дискретных и абсолютно непрерывных величин по функции распределения определяются соответственно вероятности значений и плотность распределения. Тогда по формулам (2.3) и (2.4) однозначно определяется $f_\xi(t)$. Обратное утверждение следует из формулы обращения

$$F_\xi(x_2)-F_\xi(x_1)=\frac{1}{2\pi}\lim_{N\to\infty}\int\limits_{-N}^{N}\frac{e^{-itx_1}-e^{-itx_2}}{it}f_\xi(t)\,dt. \tag{2.7}$$

Доказательство этой формулы приводится в более полных курсах теории вероятностей (см., например, А. А. Боровков [2]).

Найдем характеристические функции некоторых распределений.

Пример 1. Пусть $\xi$ — целочисленная случайная величина с производящей функцией $\varphi_\xi(x)$. Очевидно (см. (1.2) и (2.3)), что

$$f_\xi(t)=\varphi_\xi(e^{it}).$$

Пример 2. Если $\mathbf{P}(\xi=a)=1$, то по формуле (2.3)

$$f_\xi(t)=e^{ita}.$$

Пример 3. Пусть $\xi$ нормально распределена с параметрами (0, 1). Тогда

$$p_\xi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \qquad f_\xi(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{itx-\frac{x^2}{2}}\,dx.$$

Так как при любом $t$ в силу нечетности подынтегральной функции

$$\int\limits_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\sin tx\,dx=0,$$

то $f_\xi(t)$ имеет действительные значения и

$$f_\xi(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\cos tx\,dx.$$

При формальном дифференцировании этого равенства по $t$ справа получается интеграл, сходящийся равномерно по $t\in(-\infty,\infty)$. Следовательно,

$$f'_\xi(t)=-\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}xe^{-\frac{x^2}{2}}\sin tx\,dx.$$

Отсюда интегрированием по частям получим

$$f'_\xi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin tx\,de^{-\frac{x^2}{2}}=$$

$$=\frac{1}{\sqrt{2\pi}}\left[\left(\sin tx\,e^{-\frac{x^2}{2}}\right)\Big|_{-\infty}^{+\infty}-t\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\cos tx\,dx\right]=$$
$$=-tf_\xi(t).$$

Таким образом, функция $f_\xi(t)$ удовлетворяет уравнению

$$\frac{df_\xi(t)}{dt}=-tf_\xi(t)$$

и начальному условию $f_\xi(0)=1$. Отсюда

$$f_\xi(t)=e^{-\frac{t^2}{2}}.$$

Пример 4. Пусть случайная величина $\xi$ распределена нормально с параметрами $(a,\sigma)$. Найдем $f_\xi(t)$. Положим $\eta=\frac{\xi-a}{\sigma}$. Случайная величина $\eta$ имеет нормальное распределение с параметрами $(0, 1)$, и, следовательно, $f_\eta(t)=e^{-\frac{t^2}{2}}$. Тогда по утверждению 3° теоремы 2.1, получим

$$f_\xi(t)=f_{\sigma\eta+a}(t)=e^{ita}f_\eta(\sigma t)=e^{ita-\frac{\sigma^2t^2}{2}}.$$

Таким образом,

$$f_\xi(t)=e^{ita-\frac{\sigma^2t^2}{2}}. \tag{2.8}$$

Зная характеристическую функцию, можно легко найти моменты случайной величины.

Теорема 2.2. *Если существует $k$-й момент $\mathsf{M}|\xi|^k<\infty$, $k\geqslant 1$, то существует непрерывная $k$-я производная $f_\xi(t)$ и*

$$f_\xi^{(k)}(0)=i^k\mathsf{M}\xi^k.$$

Доказательство. Докажем теорему, например, для абсолютно непрерывных величин. Если существует

$k$-й момент, то существуют все моменты меньшего порядка. Так как

$$\left|\int_{-\infty}^{\infty} ix e^{itx} p_\xi(x)\,dx\right| \leqslant \int_{-\infty}^{\infty} |x|\, p_\xi(x)\,dx = \mathsf{M}|\xi| < \infty,$$

то интеграл в левой части неравенства сходится равномерно по $t$. Следовательно, можно дифференцировать под знаком интеграла

$$f'_\xi(t) = i\int_{-\infty}^{\infty} xe^{itx} p_\xi(x)\,dx, \qquad f'_\xi(0) = i\,\mathsf{M}\xi.$$

Пусть теперь существует производная порядка $l$, $l<k$ и

$$f_\xi^{(l)}(t) = i^l \int_{-\infty}^{\infty} x^l e^{itx} p_\xi(x)\,dx.$$

Отсюда

$$f_\xi^{(l+1)}(t) = i^{l+1} \int_{-\infty}^{\infty} x^{l+1} e^{itx} p_\xi(x)\,dx,$$

так как интеграл в правой части последнего равенства сходится равномерно по $t$. Таким образом,

$$f_\xi^{(l+1)}(0) = i^{l+1}\,\mathsf{M}\xi^{l+1}.$$

Теорема доказана.

В следующей главе нам потребуются разложения характеристической функции по степеням $t$ в окрестности точки $t=0$. Если существует $\mathsf{M}\xi$, то

$$f_\xi(t) = 1 + it\,\mathsf{M}\xi + t\varepsilon(t), \tag{2.9}$$

где $\varepsilon(t) \to 0$ при $t \to 0$. Для доказательства (2.9) представим $f_\xi(t)$ в виде $f_\xi(t) = u(t) + iv(t)$. Из существования $\mathsf{M}\xi$ следует существование $f'_\xi(t)$, а также $u'(t)$ и $v'(t)$. По формуле Тейлора с остаточным членом в форме Пеано

$$u(t) = u(0) + tu'(0) + t\varepsilon_1(t),$$
$$v(t) = v(0) + tv'(0) + t\varepsilon_2(t),$$

где $\varepsilon_1(t), \varepsilon_2(t) \to 0$ при $t \to 0$. Складывая два последних равенства, получим (2.9), так как $f'_\xi(0) = it\,\mathsf{M}\xi$. Аналогично проверяется следующее утверждение. Если $\mathsf{M}\xi^2$ существует, то

$$f_\xi(t) = 1 + it\,\mathsf{M}\xi - \frac{t^2}{2}\mathsf{M}\xi^2 + t^2\varepsilon(t), \qquad (2.10)$$

где $\varepsilon(t) \to 0$ при $t \to 0$.

В предположении существования $\mathsf{M}\xi^3$ дадим другую оценку остаточного члена в (2.10).

Из равенства

$$e^{ix} - 1 = \int_0^x ie^{iu}\,du$$

следует, что $|e^{ix} - 1| \leqslant \int_0^{|x|} du = |x|$, так как $|ie^{iu}| \leqslant 1$.

Так как

$$e^{ix} - 1 - ix = i\int_0^x (e^{iu} - 1)\,du,$$

то воспользовавшись оценкой $|e^{ix} - 1| \leqslant |x|$, получим, что

$$|e^{ix} - 1 - ix| \leqslant \int_0^{|x|} u\,du = \frac{|x|^2}{2}.$$

Применив эту оценку к равенству

$$e^{ix} - 1 - ix + \frac{x^2}{2} = i\int_0^x (e^{iy} - 1 - iy)\,dy,$$

получим

$$\left|e^{ix} - 1 - ix + \frac{x^2}{2}\right| \leqslant \frac{|x|^3}{6}.$$

Отсюда

$$e^{itx} = 1 + itx - \frac{t^2x^2}{2} + R_1(t, x),$$

$$|R_1(t, x)| \leqslant \frac{1}{6}|tx|^3.$$

Заменим в этом равенстве $x$ на $\xi$ и вычислим математическое ожидание от обеих частей. Получим следующее разложение:

$$\begin{gathered} f_\xi(t)=1+it\,\mathsf{M}\xi-\frac{t^2}{2}\mathsf{M}\xi^2+R(t), \\ |R(t)\leqslant\frac{|t|^3}{6}\mathsf{M}|\xi|^3. \end{gathered} \tag{2.11}$$

Для характеристических функций имеет место теорема, аналогичная теореме 1.3.

Теорема 2.3. *Если случайные величины* $\xi_1, \xi_2, \ldots, \xi_n$ *независимы, то*

$$f_{\xi_1+\xi_2+\ldots+\xi_n}(t)=f_{\xi_1}(t)\,f_{\xi_2}(t)\cdot\ \ldots\ \cdot f_{\xi_n}(t).$$

Доказательство. Достаточно доказать теорему для $n=2$, так как общее утверждение можно будет получить по индукции. По определению характеристической функции

$$\begin{aligned} f_{\xi_1+\xi_2}(t)&=\mathsf{M}e^{it(\xi_1+\xi_2)}=\mathsf{M}e^{it\xi_1}\cdot e^{it\xi_2}= \\ &=\mathsf{M}(\cos t\xi_1+i\sin t\xi_1)(\cos t\xi_2+i\sin t\xi_2). \end{aligned} \tag{2.12}$$

Дальше нужно сделать следующее: 1) перемножить выражения, стоящие в круглых скобках, и перейти к сумме математических ожиданий; 2) математические ожидания от произведений заменить на произведение математических ожиданий (это возможно, так как функции от независимых случайных величин являются независимыми случайными величинами); 3) полученное выражение вновь разложить на множители. В результате этих преобразований знак математического ожидания в правой части (2.12) появится перед каждым слагаемым в круглых скобках. Таким образом,

$$\begin{aligned} f_{\xi_1+\xi_2}(t)&= \\ &=(\mathsf{M}\cos t\xi_1+i\mathsf{M}\sin t\xi_1)(\mathsf{M}\cos t\xi_2+i\mathsf{M}\sin t\xi_2)= \\ &=\mathsf{M}e^{it\xi_1}\cdot\mathsf{M}e^{it\xi_2}=f_{\xi_1}(t)\cdot f_{\xi_2}(t). \end{aligned}$$

Теорема доказана.

Пример 5. Пусть $\xi_1$ и $\xi_2$ независимы и нормально распределены с параметрами $(a_1, \sigma_1)$, $(a_2, \sigma_2)$ соответственно. Найдем распределение суммы $\xi_1+\xi_2$. По формуле (2.8)

$$f_{\xi_1}(t)=e^{ia_1t-\frac{\sigma_1^2t^2}{2}},\qquad f_{\xi_2}(t)=e^{ia_2t-\frac{\sigma_2^2t^2}{2}}.$$

Отсюда по теореме 2.3 получим

$$f_{\xi_1+\xi_2}(t)=f_{\xi_1}(t)\,f_{\xi_2}(t)=e^{iat-\frac{\sigma^2t^2}{2}},$$

где $a=a_1+a_2$, $\sigma^2=\sigma_1^2+\sigma_2^2$. Так как полученная характеристическая функция является характеристической функцией нормального распределения, то по теореме 2.1 (4°) сумма $\xi_1+\xi_2$ распределена нормально.

Пусть задана последовательность функций распределения $F_n(x)$, $n=1, 2, \ldots$ Последовательность $\{F_n(x)\}$ *слабо сходится* к функции распределения $F(x)$, если

$$\lim_{n\to\infty} F_n(x)=F(x)$$

при любом $x$, в котором $F(x)$ непрерывна.

Теорема 2.4. *Если последовательность функций распределения $F_n(x)$, $n=1, 2, \ldots$, слабо сходится к непрерывной функции распределения $F(x)$, то $F_n(x)\to F(x)$ при $n\to\infty$ равномерно по $x\in(-\infty, +\infty)$.*

Доказательство. Из монотонности и ограниченности функции $F(x)$ следует, что для любого $\varepsilon>0$ можно выбрать конечное число точек $x_1<x_2<\ldots<x_N$ так, чтобы на каждом из следующих множеств

$$(-\infty=x_0,\ x_1),\ [x_1,\ x_2),\ \ldots,\ [x_N,\ x_{N+1}=+\infty)$$

приращение функции $F(x)$ не превосходило $\varepsilon$. Пусть $x\in[x_k,\ x_{k+1})$. Так как

$$F_n(x)-F(x)=\\=(F_n(x)-F_n(x_k))+(F_n(x_k)-F(x_k))+(F(x_k)-F(x))$$

и

$$|F_n(x)-F_n(x_k)|\leqslant F_n(x_{k+1})-F_n(x_k)\leqslant\\ \leqslant|F_n(x_{k+1})-F(x_{k+1})|+|F(x_k)-F_n(x_k)|+\\+F(x_{k+1})-F(x_k),$$

$$|F(x_k)-F(x)|\leqslant F(x_{k+1})-F(x_k)$$

в силу монотонности $F_n(x)$ и $F(x)$, то

$$|F_n(x)-F(x)| \leqslant |F_n(x_{k+1})-F(x_{k+1})|+ \\ +2|F_n(x_k)-F(x_k)|+2(F(x_{k+1})-F(x_k)). \quad (2.13)$$

По условию теоремы $F_n(x_k) \to F(x_k)$ при $n \to \infty$. Следовательно, из (2.13) с учетом выбора $[x_k, x_{k+1})$ получим, что

$$|F_n(x)-F(x)| < 5\varepsilon$$

при $n > n_0(k)$. Полагая $n_0 = \max\limits_{0 \leqslant k \leqslant N} n_0(k)$, получим $|F_n(x)-F(x)| < 5\varepsilon$ при $n > n_0$ и любых $x$. Теорема доказана.

Приведем формулировку теоремы о непрерывности соответствия множества характеристических функций множеству функций распределения. Пусть $f_n(t)$, $n = 1, 2, \ldots,$ — последовательность характеристических функций и $F_n(x)$, $n=1, 2, \ldots,$ — последовательность соответствующих функций распределения.

Теорема 2.5. *Если $f_n(t) \to f(t)$ при $n \to \infty$ для любого $t$ и $f(t)$ непрерывна при $t=0$, то*

1) *$f(t)$ — характеристическая функция, соответствующая некоторой функции распределения $F(x)$;*

2) *$F_n(x)$ слабо сходится к $F(x)$ при $n \to \infty$. Обратно, если $F_n(x)$ слабо сходится к функции распределения $F(x)$, то $f_n(t) \to f(t)$, где $f(t)$ — характеристическая функция, соответствующая функции распределения $F(x)$.*

В следующей главе эта теорема будет использована для получения предельных распределений для сумм независимых случайных величин.

Приведем формулировку еще одной теоремы, которая часто используется при исследовании предельных распределений. Пусть

$$F_n(x) = \mathsf{P}(\xi_n < x), \quad n=1, 2, \ldots,$$

— последовательность функций распределения. Обозначим $m_n^{(k)} = \mathsf{M}\,\xi_n^k$.

Теорема 2.6. *Если при всех $k$, $k=0, 1, 2, \ldots,$ $\lim\limits_{n\to\infty} m_n^{(k)} = m^{(k)} < \infty$, то существует функция распре-*

*деления* $F(x) = \mathsf{P}(\xi < x)$ *такая, что* $m^{(k)} = \mathsf{M}\xi^k$. *Если этому условию удовлетворяет единственная функция* $F(x)$, *то* $F_n(x)$ *при* $n \to \infty$ *слабо сходится к* $F(x)$.

**Задачи к главе 7**

1. Пусть $\xi$ — неотрицательная целочисленная величина с производящей функцией $\varphi(x)$. Найти производящую функцию распределения величины $2\xi + 1$. Найти $\sum\limits_{n=0}^{\infty} x^n \mathsf{P}(\xi \geqslant n)$, $\sum\limits_{n=0}^{\infty} x^n \mathsf{P}(\xi \leqslant n)$, $\sum\limits_{n=0}^{\infty} x^n \mathsf{P}(\xi = 2n)$.

2. Найти производящие функции величин $\tau$, $\tau_m$, определенных в задачах 14, 16 гл. 4. Найти $\mathsf{M}\tau$, $\mathsf{D}\tau$, $\mathsf{M}\tau_m$, $\mathsf{D}\tau_m$.

3. Найти производящую функцию величины $\nu$, определенной в задаче 16 гл. 5.

4. Найти производящую функцию случайной величины, распределенной по закону Пуассона. Доказать, что сумма независимых пуассоновских величин имеет пуассоновское распределение.

5. Случайные величины $\nu$, $\xi_1, \xi_2, \ldots, \xi_n$ независимы при любом $n = 1, 2, \ldots$; $\mathsf{P}(\xi_k = 1) = 1 - \mathsf{P}(\xi_k = 0) = p$; $\nu$ имеет распределение Пуассона. Найти производящую функцию величины $\eta = \xi_1 + \ldots + \xi_\nu$, $\eta = 0$, если $\nu = 0$.

6. Вычислить характеристические функции следующих законов распределения: а) биномиального; б) Пуассона; в) показательного; г) равномерного на отрезке $[-1, 1]$.

7. Найти законы распределения, соответствующие характеристическим функциям: $\cos t$, $\cos^2 t$, $\sum\limits_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \cos kt$.

8. Величины $\xi_1$, $\xi_2$ независимы и одинаково распределены, их характеристическая функция равна $f(t)$. Найти характеристическую функцию величины $\xi_1 - \xi_2$.

9. Величины $\xi_1$, $\xi_2$, $\xi_3$ независимы и нормально распределены с параметрами $(1, 1)$, $(0, 2)$, $(-1, 1)$. Найти: а) $\mathsf{P}(\xi_1 + \xi_2 + \xi_3 < 0)$; б) $\mathsf{P}(|2\xi_1 - \xi_2 + \xi_3| < 3)$.

10. Доказать теорему Пуассона при помощи теоремы 2.6.

ГЛАВА 8

# ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

## § 1. Закон больших чисел

В этом параграфе будет доказана теорема 2.3 гл. 6 без предположения конечности дисперсии. Случайные величины бесконечной последовательности $\xi_1, \xi_2, \ldots$ $\ldots, \xi_n, \ldots$ называются независимыми, если при любом $n$ независимы величины $\xi_1, \xi_2, \ldots, \xi_n$.

Теорема 1.1. *Пусть случайные величины $\xi_1, \xi_2, \ldots$ $\ldots, \xi_n, \ldots$ независимы, одинаково распределены и имеют математическое ожидание $\mathsf{M}\xi_k = a$, $k = 1, 2, \ldots$ Тогда при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathsf{P}\left(\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} - a\right| < \varepsilon\right) = 1.$$

Доказательство. Характеристические функции $f_{\xi_k}(t)$, $k = 1, 2, \ldots,$ одинаковы. Поэтому можно положить $f_{\xi_k}(t) = f(t)$. Из существования $\mathsf{M}\xi_k$ следует, что верно разложение (7.2.9):

$$f(t) = 1 + ita + t\varepsilon(t), \tag{1.1}$$

где $\varepsilon(t) \to 0$ при $t \to 0$. Положим

$$\eta_n = \frac{\xi_1+\xi_2+\ldots+\xi_n}{n}.$$

Так как случайные величины независимы, то по теореме 2.3 гл. 7

$$f_{\xi_1+\ldots+\xi_n}(t) = [f(t)]^n.$$

Отсюда по теореме 2.1 (3°) гл. 7 находим

$$f_{\eta_n}(t) = \left[f\left(\frac{t}{n}\right)\right]^n.$$

Заменяя в этом равенстве $f(t/n)$ по формуле (1.1), при любом фиксированном $t$ и $n \to \infty$ получим

$$f_{\eta_n}(t) = \left[1 + i\frac{t}{n}a + \frac{t}{n}\varepsilon\left(\frac{t}{n}\right)\right]^n \to e^{ita}, \tag{1.2}$$

так как $\varepsilon(t/n) \to 0$ при $n \to \infty$. Таким образом, при любом $t$ последовательность $f_{\eta_n}(t)$ сходится к функции $e^{ita}$, являющейся характеристической функцией постоянной величины $a$. Функция распределения $F(x)$ постоянной $a$ равна

$$F(x) = \begin{cases} 1, & \text{если } x > a, \\ 0, & \text{если } x \leqslant a. \end{cases} \tag{1.3}$$

По теореме 2.5 гл. 7 при любом $x \neq a$

$$\lim_{n\to\infty} F_{\eta_n}(x) = F(x), \tag{1.4}$$

так как $x = a$ — единственная точка разрыва функции $F(x)$. Пусть задано $\varepsilon > 0$. Тогда

$$\mathsf{P}(|\eta_n - a| < \varepsilon) = \mathsf{P}(a - \varepsilon < \eta_n < a + \varepsilon) \geqslant$$
$$\geqslant \mathsf{P}\left(a - \frac{\varepsilon}{2} \leqslant \eta_n < a + \varepsilon\right) = F_{\eta_n}(a+\varepsilon) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right).$$

Точки $x = a + \varepsilon$ и $x = a - \varepsilon/2$ являются точками непрерывности функции (1.3). Следовательно, при $n \to \infty$

$$F_{\eta_n}(a+\varepsilon) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right) \to F(a+\varepsilon) - F\left(a - \frac{\varepsilon}{2}\right) = 1.$$

Отсюда и из неравенств

$$1 \geqslant \mathsf{P}\Big(|\eta_n - a| < \varepsilon\Big) \geqslant F_{\eta_n}\Big(a + \varepsilon\Big) - F_{\eta_n}\left(a - \frac{\varepsilon}{2}\right)$$

следует утверждение теоремы.

## § 2. Центральная предельная теорема

В § 3 гл. 3 была доказана теорема Муавра — Лапласа, согласно которой число успехов $\mu_n$ в $n$ испытаниях схемы Бернулли при больших $n$ имеет распределение, близкое к нормальному. Если воспользоваться тем, что $\mu_n$ представляется в виде суммы независимых слагаемых (4.4.18), то теорему Муавра — Лапласа можно сформулировать в следующем виде.

*Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ независимы, $\mathsf{P}(\xi_n=1)=1-\mathsf{P}(\xi_n=0)=p$, то при $n\to\infty$, $0<p<1$*

$$\mathsf{P}\left(\frac{\xi_1+\xi_2+\ldots+\xi_n-n\mathsf{M}\xi_1}{\sqrt{n\mathsf{D}\xi_1}}<x\right)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du \quad (2.1)$$

*равномерно по $x\in(-\infty, \infty)$.*

Утверждение (2.1) сохраняется при достаточно общих предположениях о законе распределения слагаемых $\xi_k$.

Докажем следующую теорему.

Теорема 2.1. *Если случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ 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 $a=\mathsf{M}\xi_n$, $\sigma^2=\mathsf{D}\xi_n$.*

Доказательство. Положим

$$\eta_n=\frac{\xi_1+\xi_2+\ldots+\xi_n-na}{\sigma\sqrt{n}}=\sum_{k=1}^{n}\frac{\xi_k-a}{\sigma\sqrt{n}}.$$

Разложим характеристическую функцию величин $\xi_k-a$ по формуле (7.2.10):

$$f_{\xi_k-a}(t)=1+it\mathsf{M}(\xi_k-a)-\frac{t^2}{2}\mathsf{M}(\xi_k-a)^2+t^2\varepsilon(t),$$

где $\varepsilon(t)\to 0$ при $t\to 0$. Так как $\mathsf{M}(\xi_k-a)=0$ и $\mathsf{M}(\xi_k-a)^2=\sigma^2$, то

$$f_{\xi_k-a}(t)=1-\frac{t^2\sigma^2}{2}+t^2\varepsilon(t). \quad (2.2)$$

Отметим, что в этом равенстве функция $\varepsilon(t)$ не зависит от $k$. По теореме 2.3 гл. 7

$$f_{\sum\limits_{k=1}^{n}(\xi_k-a)}(t)=\left(1-\frac{t^2\sigma^2}{2}+t^2\varepsilon(t)\right)^n.$$

Отсюда

$$f_{\eta_n}(t)=f_{\sum\limits_{k=1}^{n}(\xi_k-a)}\left(\frac{t}{\sigma\sqrt{n}}\right)=\left(1-\frac{t^2}{2}\cdot\frac{1}{n}+\frac{t^2}{n\sigma^2}\varepsilon\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n$$

и, следовательно, для любого фиксированного $t$ при $n\to\infty$

$$f_{\eta_n}(t)\to e^{-\frac{t^2}{2}}.$$

Предельная характеристическая функция является характеристической функцией нормального распределения с параметрами (0, 1). Отсюда по теореме 2.5 гл. 7 следует слабая сходимость $F_{\eta_n}(x)$. Из слабой сходимости по теореме 2.4 гл. 7 следует равномерная сходимость, так как предельная функция распределения непрерывна. Теорема 2.1 доказана.

Условия сходимости функций распределения разнораспределенных слагаемых к нормальному закону содержатся в теореме Ляпунова. Приведем без доказательства ее упрощенную формулировку.

Теорема 2.2. *Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые случайные величины, имеющие конечный третий абсолютный момент. Положим*

$$a_n=\mathsf{M}\xi_n,\qquad b_n^2=\mathsf{D}\xi_n,\qquad c_n^3=\mathsf{M}\,|\xi_n-a_n|^3,$$

$$A_n=\sum_{k=1}^{n}a_k,\qquad B_n^2=\sum_{k=1}^{n}b_k^2,\qquad C_n^3=\sum_{k=1}^{n}c_k^3,$$

*Если при $n\to\infty$*

$$\frac{C_n}{B_n}\to 0, \tag{2.3}$$

*то при $n\to\infty$*

$$\mathsf{P}\left(\frac{\xi_1+\ldots+\xi_n-A_n}{B_n}<x\right)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}\,du.$$

При ссылках на теоремы о сходимости сумм распределений к нормальному закону удобно использовать понятие асимптотической нормальности. Если функции распределения последовательности случайных величин

$(\eta_n - A_n)/\sqrt{B_n}$ сходятся при $n \to \infty$ к функции $\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\left(-\frac{u^2}{2}\right) du$, то говорят, что *случайная величина* $\eta_n$ *при* $n \to \infty$ *асимптотически нормальна с параметрами* $(A_n, \sqrt{B_n})$.

В гл. 3 теорема Муавра — Лапласа, являющаяся частным случаем теорем 2.1 и 2.2, была применена к оценке отклонения частоты от вероятности. В качестве примера применения теорем 2.1 и 2.2 оценим число испытаний в методе Монте-Карло, необходимое для вычисления кратного интеграла с заданной точностью. Пусть функция $f(\mathrm{x}) = f(x_1, x_2, \ldots, x_s)$ определена на $s$-мерном единичном кубе $V$. Требуется вычислить интеграл

$$a = \int \underset{V}{\ldots} \int f(x)\, dx.$$

Пусть известна постоянная $C$ такая, что $|f(\mathrm{x})| \leqslant C$, $\mathrm{x} \in V$. Обозначим $\xi = (\xi_1, \ldots, \xi_n)$ случайный вектор, равномерно распределенный на $V$. Тогда $p_\xi(x_1, \ldots, x_s) = 1$, если $x \in V$, и $p_\xi(x_1, \ldots, x_s) = 0$ в противном случае. Математическое ожидание случайной величины $\eta = f(\xi)$ найдем по формуле (5.1.7):

$$\mathsf{M}\eta = \int \underset{\Omega}{\ldots} \int f(x_1, \ldots, x_s)\, p_\xi(x_1, \ldots, x_s)\, dx_1 \ldots dx_s =$$
$$= \int \underset{V}{\ldots} \int f(\mathrm{x})\, d\mathrm{x} = a.$$

Таким образом, $\mathsf{M}\eta$ совпадает со значением вычисляемого интеграла. Так как $|f(\mathrm{x})| \leqslant C$, то

$$\sigma^2 = \mathsf{D}\eta = \int \underset{V}{\ldots} \int (f(\mathrm{x}) - a)^2\, d\mathrm{x} \leqslant 4C^2.$$

Пусть теперь случайные векторы $\xi_k = (\xi_{k1}, \ldots, \xi_{ks})$, $k = 1, \ldots, n$, независимы *) и распределены равномерно

*) Случайные векторы $\xi_k$, $k = 1, \ldots, n$, независимы, если для любых $s$-мерных прямоугольников $B_k$, $k = 1, \ldots, n$,

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathsf{P}(\xi_1 \in B_1) \ldots \mathsf{P}(\xi_n \in B_n).$$

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

$$\zeta_n = \frac{\eta_1 + \ldots + \eta_n}{n}$$

при больших $n$ близка к постоянной $a = M\eta_k$. Предположим, что нужно вычислить $a$ с точностью $\Delta$. Оценим вероятность

$$\mathsf{P}(|\zeta_n - a| < \Delta) = \mathsf{P}\left(\left|\frac{\zeta_n - na}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{\sigma}\right).$$

Так как $\sigma < 2C$, то

$$\mathsf{P}(|\zeta_n - a| < \Delta) \geqslant \mathsf{P}\left(\left|\frac{\zeta_n - na}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{2C}\right)$$

и при больших $n$

$$\mathsf{P}\left(\left|\frac{\zeta_n - na}{\sigma/\sqrt{n}}\right| < \frac{\Delta\sqrt{n}}{2C}\right) \approx 2\Phi_0\left(\frac{\Delta\sqrt{n}}{2C}\right).$$

По заданной малой вероятности нежелательного события $|\zeta_n - a| \geqslant \Delta$ можно так же, как в § 3 гл. 3, найти $n$.

Применение метода Монте-Карло к вычислению интегралов, а также сравнение его с другими методами даются в книге С. М. Ермакова [5], гл. 4.

### Задачи к главе 8

1. Складывается $10^4$ 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 2.1 и 2.2.

3. Пусть случайная величина $\xi_\lambda$ распределена по закону Пуассона с параметром $\lambda$. Найти

$$\lim_{\lambda \to \infty} \mathsf{P}\left(\frac{\xi_\lambda - \lambda}{\sqrt{\lambda}} < x\right).$$

4. Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы;

$$\mathsf{P}(\xi_k = 1) = 1 - \mathsf{P}(\xi_k = 0) = p_k(n), \quad k = 1, 2, \ldots$$

Найти $\lim\limits_{n \to \infty} \mathsf{P}(\xi_1 + \ldots + \xi_n = m)$, если $p_1(n) + \ldots + p_n(n) \longrightarrow \lambda$ при $n \longrightarrow \infty$.

ГЛАВА 9

# ЦЕПИ МАРКОВА

## § 1. Определение. Основные свойства

Определение цепи Маркова как частного случая общей схемы испытаний было дано в гл. 3. Дадим прямое определение. *Последовательностью $T$ испытаний, являющихся цепью Маркова с $N$ состояниями, будем называть вероятностное пространство* $(\Omega, \mathfrak{F}, \mathsf{P})$, *в котором*

$$\Omega = \{(l_0\, l_1 \ldots l_T)\}, \quad l_t = 1, 2, \ldots, N, \quad t = 0, 1, \ldots, T, \tag{1.1}$$

*и вероятности* $p(l_0 l_1 \ldots l_T)$, *приписываемые элементарным событиям* $\omega = (l_0 l_1 \ldots l_T)$, *задаются формулой*

$$p(l_0 l_1 \ldots l_T) = q_{l_0} p_{l_0 l_1} p_{l_1 l_2} \ldots p_{l_{T-1} l_T}, \tag{1.2}$$

*где*

$$q_k \geqslant 0, \quad k = 1, \ldots, N, \quad \sum_{k=1}^{N} q_k = 1$$

*и элементы матрицы*

$$P = \begin{pmatrix} p_{11} & p_{12} & \ldots & p_{1N} \\ p_{21} & p_{22} & \ldots & p_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ p_{N1} & p_{N2} & \ldots & p_{NN} \end{pmatrix} \tag{1.3}$$

*удовлетворяют условию*

$$p_{ij} \geqslant 0, \quad i, j = 1, \ldots, N; \quad \sum_{j=1}^{N} p_{ij} = 1, \quad i = 1, \ldots, N. \tag{1.4}$$

Алгебра событий $\mathfrak{F}$ состоит из всех подмножеств $\Omega$; вероятность определяется формулой (1.6.6):

$$\mathsf{P}(A) = \sum_{(l_0 l_1 \ldots l_T) \in A} \mathsf{P}(l_0 l_1 \ldots l_T).$$

Приведенное здесь определение цепи Маркова является более узким по сравнению с определением § 1 гл. 3. Здесь числа $p_{ij}$ (1.3) ($i, j = 1, 2, \ldots, N$), являющиеся вероятностями перехода из состояния $i$ в состояние $j$, не зависят от номера испытания, как в гл. 3. Такие цепи Маркова называют однородными. Утверждение теоремы 1.2 верно только для однородных цепей Маркова.

Условие (3.1.6) было проверено в гл. 3 для произвольной последовательности испытаний. Матрицы вида (1.3), для которых выполнено (1.4), называются стохастическими. Цепи Маркова часто используются для описания различных систем, которые могут находиться в одном из $N$ состояний и в дискретные моменты времени меняют состояние. В этом случае номер испытания $t$ естественно интерпретировать как время, а символ $l_t$ — как состояние системы в момент времени $t$. С цепью Маркова естественно связать последовательность случайных величин $\xi_t = \xi_t(l_0 l_1 \ldots l_T) = l_t$, $t = 0, 1, \ldots, T$, являющихся номером состояния, в котором находится система в момент времени $t$.

Из (1.2) следует, что

$$\mathsf{P}(\xi_0 = l_0, \xi_1 = l_1, \ldots, \xi_T = l_T) = q_{l_0} p_{l_0 l_1} p_{l_1 l_2} \ldots p_{l_{T-1} l_T}.$$

Так как

$$(\xi_0 = L_0, \xi_1 = L_1, \ldots, \xi_t = L_t) = \\ = \{\omega\colon l_0 = L_0, \ldots, l_t = L_t\} = \\ = \{L_0 L_1 \ldots L_t l_{t+1} \ldots l_T\} = A_{L_0 \ldots L_t},$$

то

$$\mathsf{P}(\xi_0 = L_0, \xi_1 = L_1, \ldots, \xi_t = L_t) = \\ = \sum_{\omega \in A_{L_0 \ldots L_t}} q_{L_0} p_{L_0 L_1} \ldots p_{L_{t-1} L_t} p_{L_t l_{t+1}} \ldots p_{l_{T-1} l_T} = \\ = q_{L_0} p_{L_0 L_1} \ldots p_{L_{t-1} L_t} \sum_{l_{t+1}, \ldots, l_T = 1}^{N} p_{L_t l_{t+1}} p_{l_{t+1} l_{t+2}} \ldots p_{l_{T-1} l_T}.$$

Из свойства (1.4) следует, что сумма в последнем равенстве равна 1. Таким образом,

$$P(\xi_0 = l_0, \xi_1 = l_1, \ldots, \xi_t = l_t) = q_{l_0} p_{l_0 l_1} \ldots p_{l_{t-1} l_t} \quad (1.5)$$

при любых $l_0, l_1, \ldots, l_t = 1, 2, \ldots, N$ и любом $t = 0, 1, 2, \ldots, T$.

Докажем основное свойство цепи Маркова, определенной формулами (1.1)—(1.4).

Теорема 1.1. *Если $\xi_t$ — номер состояния цепи Маркова в момент $t$, то*

$$\mathsf{P}(\xi_{t+1}=j \mid \xi_0=i_0,\ \xi_1=i_1,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i)= \\ =\mathsf{P}(\xi_{t+1}=j \mid \xi_t=i) \qquad (1.6)$$

*и, кроме того,*

$$\mathsf{P}(\xi_{t+1}=j \mid \xi_t=i)=p_{ij}$$

*при любых* $t=0, 1, 2, \ldots, T-1$ *и любых* $i_0, i_1, \ldots, i_{t-1}, i, j=1, 2, \ldots, N$.

Доказательство. Используя определение условной вероятности и (1.5), получим

$$\mathsf{P}(\xi_{t+1}=j \mid \xi_0=i_0,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i)= \\ =\frac{\mathsf{P}(\xi_0=i_0,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i,\ \xi_{t+1}=j)}{\mathsf{P}(\xi_0=i_0,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i)}= \\ =\frac{q_{i_0}p_{i_0i_1}\ldots p_{i_{t-1}i}p_{ij}}{q_{i_0}p_{i_0i_1}\ldots p_{i_{t-1}i}}=p_{ij}. \qquad (1.7)$$

Правую часть (1.6) запишем в виде

$$\mathsf{P}(\xi_{t+1}=j \mid \xi_t=i)=\frac{\mathsf{P}(\xi_t=i,\ \xi_{t+1}=j)}{\mathsf{P}(\xi_t=i)}= \\ =\frac{\sum \mathsf{P}(\xi_0=i_0,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i,\ \xi_{t+1}=j)}{\sum \mathsf{P}(\xi_0=i_0,\ \ldots,\ \xi_{t-1}=i_{t-1},\ \xi_t=i)}= \\ =\frac{\sum q_{i_0}p_{i_0i_1}\ldots p_{i_{t-1}i}p_{ij}}{\sum q_{i_0}p_{i_0i_1}\ldots p_{i_{t-1}i}}=p_{ij}, \qquad (1.8)$$

где суммирование проводится по каждому из индексов $i_0, i_1, \ldots, i_{t-1}$ от 1 до $N$. Утверждение теоремы следует из (1.7) и (1.8). Свойство (1.6) можно интерпретировать следующим образом: состояние системы в момент $t+1$ при известном состоянии в момент $t$ не зависит от поведения системы в более далеком прошлом (в моменты $s<t$).

Так же, как (1.6), можно доказать следующее равенство:

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_u \in B_u,\ u=0, 1, \ldots, s-1,\ \xi_s=i)= \\ =\mathsf{P}(\xi_{t+s}=j \mid \xi_s=i), \qquad (1.9)$$

при любых $i, j, s, t, B_u \subset \{1, 2, \ldots, N\}$. Если дискретные случайные величины $\xi_0, \xi_1, \ldots, \xi_T$, принимающие значения $1, 2, \ldots, N$, удовлетворяют (1.6) или (1.9), то говорят, что величины $\xi_t$ являются цепью Маркова.

Теорема 1.2. *При любом* $s$

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_s=i)=\mathsf{P}(\xi_t=j \mid \xi_0=i).$$

Доказательство. По формуле (1.5)

$$\mathsf{P}(\xi_s=i,\ \xi_{t+s}=j)=$$

$$=\sum_{i_0 \ldots i_{s-1} i_{s+1} \ldots i_{t+s-1}=1}^{N} q_{i_0} p_{i_0 i_1} \ldots p_{i_{s-1} i} p_{i i_{s+1}} \ldots p_{i_{t+s-1} j}=$$

$$=\sum_{i_0 i_1 \ldots i_{s-1}=1}^{N} q_{i_0} p_{i_0 i_1} \ldots p_{i_{s-1} i} \times$$

$$\times \sum_{i_{s+1} \ldots i_{t+s-1}=1}^{N} p_{i i_{s+1}} \ldots p_{i_{t+s-1} j}. \qquad (1.10)$$

Первый множитель в правой части этого равенства равен $\mathsf{P}(\xi_s=i)$. Если во втором множителе переменные индексы $i_{s+1}, \ldots, i_{t+s-1}$, по которым проводится суммирование, обозначить $l_1, \ldots, l_{t-1}$, то (1.10) можно переписать так:

$$\mathsf{P}(\xi_s=i,\ \xi_{t+s}=j)=\mathsf{P}(\xi_s=i)\,\frac{\sum\limits_{l_1 \ldots l_{t-1}=1}^{N} q_i p_{i l_1} \ldots p_{l_{t-1} j}}{q_i}. \qquad (1.11)$$

Второй множитель в правой части (1.11) равен $\mathsf{P}(\xi_t=j \mid \xi_0=i)$. Таким образом, поделив обе части (1.11) на $\mathsf{P}(\xi_s=i)$, получим утверждение теоремы.

Так как условная вероятность $\mathsf{P}(\xi_{t+s}=j \mid \xi_s=i)$ не зависит от $s$, то мы можем положить

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_s=i)=P_{ij}(t). \qquad (1.12)$$

Функции $P_{ij}(t)$ в правой части (1.12) называют вероятностями перехода.

Рассмотрим несколько примеров.

Пример 1. Блуждание частицы по целым точкам отрезка $[0, n]$, описанное в § 5 гл. 2, является цепью Маркова, в которой

$$q_l = 0, \quad l \neq k; \quad q_k = 1;$$

$$\begin{aligned} &p_{l,\,l+1} = p, \quad p_{l,\,l-1} = 1-p = q, \quad l = 1, 2, \ldots, n-1; \\ &p_{0,\,0} = p_{n,\,n} = 1; \; p_{0,\,l} = 0, \quad l = 1, 2, \ldots, n; \\ &p_{n,\,l} = 0, \quad l = 0, 1, \ldots, n-1. \end{aligned} \tag{1.13}$$

Из свойств цепей Маркова сразу следует (2.5.1). Действительно, по теореме 1.2

$$\mathsf{P}(\xi_{t+1} = n \mid \xi_1 = k+1) = \mathsf{P}(\xi_t = n \mid \xi_0 = k+1),$$
$$\mathsf{P}(\xi_{t+1} = n \mid \xi_1 = k-1) = \mathsf{P}(\xi_t = n \mid \xi_0 = k-1).$$

Пример 2. Пусть

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad q_1 = 1, \quad q_2 = q_3 = 0.$$

Так как переходы из 2-го и 3-го состояний в 1-е не происходят, то

$$\mathsf{P}(\xi_t = 1) = \mathsf{P}(\xi_0 = 1,\ \xi_1 = 1,\ \ldots,\ \xi_t = 1) =$$
$$= \mathsf{P}(\xi_0 = 1)\,\mathsf{P}(\xi_1 = 1 \mid \xi_0 = 1)\ldots\mathsf{P}(\xi_t = 1 \mid \xi_{t-1} = 1) = \left(\frac{1}{3}\right)^t$$

и вероятность $\mathsf{P}(\xi_t = 1)$, рассматриваемая как функция от $t$, имеет

$$\lim_{t \to \infty} \mathsf{P}(\xi_t = 1) = 0. \tag{1.14}$$

Цепь Маркова была определена как конечная последовательность испытаний. Можно ввести беконечную последовательность испытаний, связанных в цепь Маркова. Распространение определения с конечной последовательности испытаний на бесконечную проводится так же, как для схемы Бернулли в § 4 гл. 3. Вероятности событий, которые являются объединением конечного числа событий вида $(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s)$, вычисляются по формулам для цепи Маркова с конечным числом испытаний. Если в примере 2 определить цепь Маркова с бесконечным числом испытаний, то (1.14) является вероятностью остаться в состоянии 1.

Состояние $i$ цепи Маркова называется *несущественным*, если существуют состояние $j$ и число $t_0$ такие, что $P_{ij}(t_0)>0$, $P_{ji}(t)=0$ при любом $t$. В противном случае состояние называется существенным.

В примере 1 существенными состояниями являются 0 и $n$, а остальные — несущественные. В примере 2 состояние 1 — несущественнное, а 2 и 3 — существенные. Можно доказать, что система, описываемая цепью Маркова, уходит из несущественных состояний с вероятностью 1.

Пример 3. Положим

$$P=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad q_1=1,\ q_2=0.$$

Если в момент $t=0$ система находилась в первом состоянии, то в нечетные моменты времени система будет находиться в состоянии 2, а в четные — в состоянии 1. Таким образом,

$$P_{12}(t)=\frac{1-(-1)^t}{2}, \quad P_{11}(t)=\frac{1+(-1)^t}{2}. \tag{1.15}$$

Пример 4. Рассмотрим $T+1$ испытаний Бернулли; вероятность успеха в испытании с номером $t$, $t=0, 1, 2, \ldots, T$, равна $p$, $0<p<1$. Пусть случайная величина $\xi_t=1$, если общее число успехов в испытаниях с номерами $s=0, 1, 2, \ldots, t$ было нечетным и $\xi_t=2$ в противном случае.

Интуитивно равенство

$$\mathsf{P}(\xi_{t+1}=j \mid \xi_0=i_0,\ \xi_1=i_1, \ldots, \xi_{t-1}=i_{t-1},\ \xi_t=i)=\\ =\mathsf{P}(\xi_{t+1}=j \mid \xi_t=i), \tag{1.16}$$
$$i, j=1, 2,$$

очевидно: если в момент $t$ четность числа успехов известна, то четность в следующий момент определяется только результатом очередного испытания, а испытания по условию независимы. Равенство (1.16) легко доказать, если воспользоваться тем, что событие

$$(\xi_0=i_0,\ \xi_1=i_1,\ \ldots,\ \xi_t=i,\ \xi_{t+1}=j) \qquad (1.17)$$

однозначно определяет исходы в испытаниях Бернулли с номерами $0, 1, \ldots, t+1$, а событие $(\xi_t=i,\ \xi_{t+1}=j)$ является конечной суммой событий вида (1.17). Из (1.16) следует, что $\xi_t$ является цепью Маркова. Непосредственно из определения $\xi_t$ находим

$$p_{11}=\mathsf{P}(\xi_1=1 \mid \xi_0=1)=p_{22}=\mathsf{P}(\xi_1=2 \mid \xi_0=2)=1-p=q,$$

так как состояние не меняется, если успех не появился. Аналогично получим

$$p_{12}=p_{21}=p, \qquad q_1=p, \qquad q_2=q.$$

Нетрудно найти вероятности перехода $P_{ij}(t)$. Например,

$$P_{21}(t)=\mathsf{P}(\xi_t=1 \mid \xi_0=2)=\sum_{k=0}^{\left[\frac{t-1}{2}\right]} C_t^{2k-1}p^{2k+1}q^{t-2k+1}.$$

## § 2. Уравнение для вероятностей перехода

Вычисление вероятностей событий

$$(\xi_{t_1}=l_1,\ \xi_{t_2}=l_2,\ \ldots,\ \xi_{t_s}=l_s) \qquad (2.1)$$

в виде сумм вероятностей соответствующих элементарных событий затруднительно. Довольно просто вероятность события (2.1) выражается через вероятности перехода $P_{ij}(t)$. По формуле полной вероятности

$$\mathsf{P}(\xi_{t_1}=l_1,\ \ldots,\ \xi_{t_s}=l_s)=$$
$$=\sum_{k=1}^{N}\mathsf{P}(\xi_0=k)\,\mathsf{P}(\xi_{t_1}=l_1,\ \ldots,\ \xi_{t_s}=l_s \mid \xi_0=k). \qquad (2.2)$$

Используя (1.9) и теорему 1.2, получим

$$\begin{aligned}
&\mathsf{P}(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s \mid \xi_0=k)= \\
&\quad=\mathsf{P}(\xi_{t_1}=l_1 \mid \xi_0=k)\,\mathsf{P}(\xi_{t_2}=l_2 \mid \xi_0=k,\ \xi_{t_1}=l_1)\times\ldots \\
&\quad\ldots\times\mathsf{P}(\xi_{t_s}=l_s \mid \xi_0=k,\ \xi_{t_1}=l_1, \ldots, \xi_{t_{s-1}}=l_{s-1})= \\
&\qquad=\mathsf{P}(\xi_{t_1}=l_1 \mid \xi_0=k)\,\mathsf{P}(\xi_{t_2}=l_2 \mid \xi_{t_1}=l_1)\ldots \\
&\qquad\quad\ldots\mathsf{P}(\xi_{t_s}=l_s \mid \xi_{t_{s-1}}=l_{s-1})= \\
&\qquad\quad=P_{kl_1}(t_1)\,P_{l_1l_2}(t_2-t_1)\ldots P_{l_{s-1}l_s}(t_s-t_{s-1}).
\end{aligned}$$

Отсюда и из (2.2) найдем

$$\mathsf{P}(\xi_{t_1}=l_1, \ldots, \xi_{t_s}=l_s)= \\
=\sum_{k=1}^{N} q_k P_{kl_1}(t_1)\,P_{l_1l_2}(t_2-t_1)\ldots P_{l_{s-1}l_s}(t_s-t_{s-1}). \qquad (2.3)$$

Для вычислений по формуле (2.3) нужно уметь находить $P_{ij}(t)$.

Теорема 2.1. *При любых $s, t$*

$$P_{ij}(t+s)=\sum_{k=1}^{N} P_{ik}(s)\,P_{kj}(t), \qquad i, j=1, 2, \ldots, N. \qquad (2.4)$$

Доказательство. Вычислим вероятность $\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i)$ по формуле полной вероятности (2.3.1), положив $B_k=(\xi_s=k)$:

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i)= \\
=\sum_{k=1}^{N} \mathsf{P}(\xi_s=k \mid \xi_0=i)\,\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i,\ \xi_s=k). \qquad (2.5)$$

Из равенства (1.9) и теоремы 1.2 следует

$$\mathsf{P}(\xi_{t+s}=j \mid \xi_0=i,\ \xi_s=k)= \\
=\mathsf{P}(\xi_{t+s}=j \mid \xi_s=k)=\mathsf{P}(\xi_t=j \mid \xi_0=k).$$

Отсюда и из равенств (2.5) и (1.12) получаем утверждение теоремы. Определим матрицу $P(t)=\|P_{ij}(t)\|$. В матричной записи (2.4) имеет вид

$$P(t+s)=P(s)\,P(t). \qquad (2.6)$$

Так как $P_{ij}(1)=p_{ij}$, то $P(1)=P$, где $P$ — матрица (1.3). Из (2.6) следует

$$P(t)=(P(1))^t=P^t. \qquad (2.7)$$

Результаты, полученные в теории матриц, позволяют по формуле (2.7) вычислять $P_{ij}(t)$ и исследовать их поведение при $t \to \infty$. Некоторые теоремы о неотрицательных матрицах приведены в работе Б. А. Севастьянова [15]; подробное исследование стохастических матриц и цепей Маркова содержится в книге В. И. Романовского [14].

## § 3. Стационарное распределение. Теорема о предельных вероятностях

Вектор $q=(q_1, q_2, \ldots, q_N)$, входящий в определение цепи Маркова, задает распределение в нулевой момент времени

$$q_k = \mathsf{P}(\xi_0 = k), \qquad k = 1, 2, \ldots, N.$$

Распределение в произвольный момент времени $t$ можно найти, воспользовавшись формулой полной вероятности

$$\mathsf{P}(\xi_t = j) = \sum_{k=1}^{N} q_k P_{kj}(t). \tag{3.1}$$

Может оказаться, что $\mathsf{P}(\xi_t = j)$ не зависит от времени. Назовем *стационарным распределением* вектор $q^* = (q_1^*, \ldots, q_N^*)$, удовлетворяющий условиям

$$q_k^* \geqslant 0, \qquad k = 1, \ldots, N;$$

$$\sum_{k=1}^{N} q_k^* = 1, \quad \sum_{k=1}^{N} q_k^* p_{kj} = q_j^*, \quad j = 1, 2, \ldots, N, \tag{3.2}$$

где $p_{ij}$ — элементы матрицы (1.3).

Если в цепи Маркова $q = q^*$, то при любом $t$

$$\mathsf{P}(\xi_t = k) = q_k^*, \qquad k = 1, 2, \ldots, N.$$

Это утверждение следует по индукции из (3.1) и (3.2).

Приведем без доказательства формулировку теоремы о предельных вероятностях для одного важного класса цепей Маркова.

Теорема 3.1. *Если при некотором $t_0 > 0$ все элементы матрицы $P^{t_0}$ положительны, то для любых $i, j = 1, 2, \ldots, N$ при $t \to \infty$*

$$P_{ij}(t) = q_j^* + \varepsilon_{ij}(t), \tag{3.3}$$

*где* $q^* = (q_1^*, \ldots, q_N^*)$ — *стационарное распределение с* $q_k^* > 0$, $k = 1, \ldots, N$, $\varepsilon_{ij}(t) = O(h^t)$, $h$ — *некоторая постоянная, удовлетворяющая неравенствам* $0 < h < 1$.

Доказательство этой теоремы, опирающееся на факты из теории матриц, содержится в книге В. И. Романовского [14]; прямым методом доказывается теорема из § 19 книги Б. В. Гнеденко [4]. В приложении 1 приведено доказательство теоремы 3.1.

Так как $P(t_0) = P^{t_0}$, то по условию теоремы из любого состояния можно попасть в любое за время $t_0$ с положительной вероятностью. Условия теоремы исключают цепи, являющиеся в некотором смысле периодическими (см. пример 3 из § 1).

Если выполнены условия теоремы 3.1, то вероятность того, что система находится в некотором состоянии $j$, через большой промежуток времени не зависит от начального распределения.

Действительно, из (3.3) и (3.1) следует, что при любом начальном распределении $q_i = \mathsf{P}(\xi_0 = i)$, $i = 1, \ldots, N$,

$$\lim_{t \to \infty} \mathsf{P}(\xi_t = j) = q_j^*, \qquad j = 1, 2, \ldots, N.$$

Рассмотрим примеры цепей Маркова, для которых условия теоремы 3.1 не выполнены. Нетрудно проверить, что такими примерами являются примеры 1, 2, 3 из § 1. В примере 1 при $t \to \infty$ вероятности перехода имеют пределы, но эти пределы зависят от начального состояния. Например, при $p = q = \frac{1}{2}$ (см. § 5 гл. 2)

$$\lim_{t \to \infty} P_{kl}(t) = 0, \qquad 0 < l < n, \qquad k = 0, 1, \ldots n,$$

$$\lim_{t \to \infty} P_{k0}(t) = 1 - \frac{k}{n}, \quad \lim_{t \to \infty} P_{kn}(t) = \frac{k}{n},$$

$$k = 0, 1, \ldots, n.$$

В примере 3 пределы вероятностей (1.15) при $t \to \infty$, очевидно, не существуют.

Найдем стационарное распределение в примере 2. Нужно найти вектор $q^*=(q_1^*, q_2^*, q_3^*)$, удовлетворяющий условиям (3.2):

$$q_k^* \geqslant 0, \quad k=1, 2, 3; \qquad \begin{cases} q_1^*\frac{1}{3} = q_1^*, \\ q_1^*\frac{1}{3}+q_2^*\frac{1}{2}+q_3^*\frac{1}{2}=q_2^*, \\ q_1^*\frac{1}{3}+q_2^*\frac{1}{2}+q_3^*\frac{1}{2}=q_3^*. \end{cases}$$
$$q_1^*+q_2^*+q_3^*=1;$$

Отсюда $q_1^*=0$, $q_2^*=q_3^*=1/2$. Таким образом, стационарное распределение существует, но не все координаты вектора $q^*$ положительны.

В примере 4 условия теоремы 3.1 выполнены уже при $t_0=1$. Стационарное распределение $(q_1^*, q_2^*)$ удовлетворяет уравнениям

$$q_1^*+q_2^*=1, \qquad q_1^*p_{11}+q_2^*p_{21}=q_1^*, \qquad q_1^*p_{12}+q_2^*p_{22}=q_2^*,$$

где $p_{11}=p_{22}=p$, $p_{12}=p_{21}=q$. Отсюда $q_1^*=q_2^*=1/2$ и по теореме 3.1 $\lim\limits_{t\to\infty} P_{ij}(t)=1/2$, $i$, $j=1, 2$.

Во всех разобранных примерах оказалось, что система однородных уравнений

$$(p_{jj}-1)\,q_j^*+\sum_{k\neq j} q_k^*p_{kj}=0, \qquad j=1, 2, \ldots, N, \tag{3.4}$$

относительно неизвестных $q_1^*, \ldots, q_N^*$ имеет ненулевое решение. Нетрудно проверить, что определитель матрицы коэффициентов системы (3.4) равен нулю (если все столбцы определителя сложить с первым, то в силу свойства (1.4) получим столбец из нулей).

В § 2 гл. 3 для полиномиальной схемы были введены случайные величины, равные числу исходов данного типа. Введем аналогичные величины для цепей Маркова. Пусть $\tau_k(t)$, $k=1, \ldots, N$,— число попаданий системы в состояние $k$ за время $t$. Тогда $\tau_k(t)/t$ — частота попаданий системы в состояние $k$. Используя формулы (3.3), можно доказать, что частота $\tau_k(t)/t$ при $t\to\infty$ сближается с $q_k^*$. Для этого нужно получить асимптотические формулы для $\mathsf{M}\tau_k(t)$ и $\mathsf{D}\tau_k(t)$ и воспользоваться неравенством Чебышева. Приведем вывод формулы для $M(\tau_j(t)\mid\xi_0=i)=m_{ij}(t)$. Представим $\tau_j(t)$ в виде

$$\tau_j(t)=\eta_0+\eta_1+\ldots+\eta_t, \tag{3.5}$$

где $\eta_s=1$, если $\xi_s=j$, и $\eta_s=0$ в противном случае. Так как

$$M(\eta_s\mid\xi_0=i)=P_{ij}(s),$$

то, воспользовавшись свойствами математического ожидания и формулой (3.3), получим

$$m_{ij}(t)=\sum_{s=0}^{t}P_{ij}(s)=q_j^*(t)+\sum_{s=0}^{t}\varepsilon_{ij}(s).$$

Второе слагаемое в правой части этого равенства в силу теоремы 3.1 является частной суммой сходящегося ряда. Положив

$$\alpha_{ij}=\sum_{s=0}^{\infty}\varepsilon_{ij}(s),$$

получим

$$m_{ij}(t)=q_j^*(t)+\alpha_{ij}+O(h^t),\qquad 0<h<1, \tag{3.6}$$

так как

$$\sum_{s=0}^{t}\varepsilon_{ij}(s)=\alpha_{ij}-\sum_{s=t+1}^{\infty}\varepsilon_{ij}(s),\qquad |\varepsilon_{ij}(s)|<K\cdot h^s.$$

Из формулы (3.6), в частности, следует, что

$$M\left(\frac{\tau_j(t)}{t}\,\middle|\,\xi_0=i\right)\to q_j^*,\qquad j=1,2,\ldots,N,$$

при $t\to\infty$. Аналогично можно получить формулу для $M[\tau_j(t)(\tau_j(t)-1)\mid\xi_0=i]$, которая используется для вычисления дисперсии.

### Задачи к главе 9

1. Игральная кость все время перекладывается случайным образом с одной грани равновероятно на любую из четырех соседних граней независимо от предыдущего. К какому пределу стремится при $t\to\infty$ вероятность того, что в момент времени $t$

кость лежит на грани «6», если в момент $t=0$ она находилась в этом же положении ($t=0, 1, 2, \ldots$)?

**2.** Пусть $\xi_t$, $t=1, 2, \ldots, T$,— независимые одинаково распределенные случайные величины, $\mathsf{P}(\xi_t=1)=p$, $\mathsf{P}(\xi_t=-1)=1-p$. Являются ли цепью Маркова величины:

а) $\eta_t=\xi_t\xi_{t+1}$;

б) $\eta_t=\xi_1\xi_2\ldots\xi_t$;

в) $\eta_t=\varphi(\xi_t, \xi_{t+1})$,

где $\varphi(-1, -1)=1$, $\varphi(-1, 1)=2$, $\varphi(1, -1)=3$, $\varphi(1, 1)=4$? Для цепей Маркова найти матрицы вероятностей перехода за один шаг.

**3.** Пусть $\xi_t$ — номер состояния в цепи Маркова в момент времени $t$, $\mathsf{P}(\xi_0=1)=1$ и матрица вероятностей перехода за единицу времени равна

$$\begin{pmatrix} 3/7 & 3/7 & 1/7 \\ 1/11 & 2/11 & 8/11 \\ 1/11 & 4/11 & 6/11 \end{pmatrix}; \qquad \eta_t=\begin{cases} 1, & \text{если } \xi_t=1, \\ 2, & \text{если } \xi_t\neq 1. \end{cases}$$

Показать, что последовательность $\eta_t$ является цепью Маркова. Найти соответствующую ей матрицу вероятностей перехода.

**4.** По двум урнам разложено $N$ черных и $N$ белых шаров так, что каждая урна содержит $N$ шаров. Число черных шаров в первой урне определяет состояние системы. В каждый момент времени выбирают случайно по одному шару из урны и выбранные шары меняют местами. Найти вероятности перехода и показать, что стационарные вероятности $q_k^*=C_N^kC_N^{N-k}/C_{2N}^N$, $k=1,\ldots,N$.

**5*.** Пусть в цепи Маркова с состояниями 1 и 2 матрица вероятностей перехода имеет вид $P=\|p_{ij}\|$, где $p_{11}=1-\alpha$, $p_{12}=\alpha$, $p_{21}=\beta$, $p_{22}=1-\beta$. Найти вероятности перехода за время $t$ и стационарные вероятности.

**6*.** В *цепи Маркова, определенной* в задаче 5, обозначим $\tau(t)$ число попаданий в состояние «1» и $\xi(t)$ номер состояния в момент $t$. Доказать, что $\mathsf{M}(\tau(t)\,|\,\xi(0)=i)$ и $\mathsf{D}(\tau(t)\,|\,\xi(0)=i)$ можно при $t\to\infty$ представить в виде

$$\mathsf{M}(\tau(t)\,|\,\xi(0)=i)=At+B+o(1),$$

$$\mathsf{D}(\tau(t)\,|\,\xi(0)=i)=Ct+o(t).$$

Определить постоянные $A$, $B$, $C$. Указание: воспользоваться представлением (9.3.5), решением задачи 5 и формулой (9.2.3).

**7*.** Доказать, что для любого $\varepsilon>0$

$$\lim_{t\to\infty}\mathsf{P}\left(\left|\frac{\tau(t)}{t}-A\right|<\varepsilon\right)=1,$$

где $\tau(t)$ и $A$ определены в задаче 6.

ГЛАВА 10

# ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

## § 1. Задачи математической статистики. Понятие выборки

Пусть требуется измерить некоторую величину $a$. Результаты измерений $x_1(\omega)$, $x_2(\omega)$, ..., $x_n(\omega)$ естественно рассматривать как значения случайных величин, $x_1$, $x_2$, ..., $x_n$, полученные в данном опыте с исходом $\omega$. Если измерительный прибор не дает систематической ошибки, то можно предположить, что $Mx_k = a$. Таким образом, по результатам наблюдений $x_1$, $x_2$, ..., $x_n$, нужно определить неизвестный параметр $a$. Это типичная *задача оценки неизвестных параметров*. Общая ошибка измерения часто складывается из большого числа ошибок, каждая из которых невелика. В такой ситуации на основании центральной предельной теоремы становится правдоподобным следующее предположение (гипотеза): случайные величины $x_k$ имеют нормальное распределение. Таким образом, мы пришли к *задаче статистической проверки гипотезы о законе распределения*.

В § 4 гл. 1 рассматривались две различные модели (модель 1 и модель 2) подбрасывания двух монет. Вероятность выпадения монет одинаковыми сторонами в модели 1 оказалась равной 2/3, а в модели 2 — равной 1/2. По результатам подбрасываний двух монет требуется выбрать одну из двух моделей (или гипотез).

В перечисленных выше задачах оценка неизвестного параметра или выбор гипотезы проводится по результатам наблюдений или измерений. Математической моделью $n$ независимых измерений является совокупность из $n$ независимых случайных величин.

Назовем *случайной выборкой объема $n$* (или просто *выборкой*) случайный вектор $x_1$, $x_2$, ..., $x_n$, где $x_k$,

$k=1, \ldots, n$,— независимые случайные величины с одинаковой функцией распределения

$$\mathsf{P}(x_k < x) = F(x), \qquad k=1, 2, \ldots, n.$$

В задачах, связанных с измерением, часто оказывается, что $F(x)$ является нормальным распределением. В задаче с подбрасыванием двух монет имеем

$$\mathsf{P}(x_k=1)=p_l, \quad \mathsf{P}(x_k=0)=1-p_l, \quad l=1, 2;$$
$$k=1, 2, \ldots, n,$$

где $x_k=1$, если в $k$-м испытании монеты выпали одинаковыми сторонами; $p_1=2/3$, $p_2=1/2$.

Иногда требуется рассматривать выборки с зависимыми наблюдениями. При оценке доли брака в партии изделий используется схема выбора без возвращения (см. § 6 гл. 1). В этом случае наблюдения оказываются зависимыми.

## § 2. Оценка неизвестных параметров распределения по выборке

**2.1. Точечные оценки.** Предположим, что функция распределения, соответствующая выборке $x_1, x_2, \ldots, x_n$, зависит от неизвестного параметра $\theta$:

$$\mathsf{P}(x_k < x) = F(x; \theta).$$

*Оценкой* $\theta_n^*$ *параметра* $\theta$ называется произвольная функция $\theta_n^* = \theta_n^*(x_1, x_2, \ldots, x_n)$. Таким образом, $\theta_n^*$ является случайной величиной. Естественно потребовать, чтобы значения оценки в большинстве опытов были близки к значению оцениваемого параметра. Будем называть оценку $\theta_n^*$ *несмещенной оценкой параметра* $\theta$, если при любом $n$

$$\mathsf{M}\theta_n^* = \theta. \tag{2.1}$$

*Оценка* $\theta_n^*$ *называется состоятельной*, если для любого $\varepsilon > 0$ при $n \to \infty$

$$\mathsf{P}(|\theta_n^* - \theta_n| < \varepsilon) \to 1. \tag{2.2}$$

Условиям (2.1) и (2.2) может удовлетворять несколько разных оценок одного и того же параметра. Тогда из них естественно выбрать ту, у которой дисперсия наименьшая. Для широкого класса распределений можно указать точную нижнюю грань дисперсий оценок одного и того же параметра (см. [10], гл. 32, стр. 521, (32.3.3a)).

Рассмотрим два примера.

Пример 1. Пусть $\theta = a = \mathsf{M}x_k$. Положим

$$\theta_n^* = \theta_n^*(x_1, x_2, \ldots, x_n) = \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n},$$

$$\tilde{\theta}_n^* = \tilde{\theta}_n^*(x_1, x_2, \ldots, x_n) = x_1.$$

Используя свойства математического ожидания, легко проверить, что обе оценки являются несмещенными:

$$\mathsf{M}\theta_n^* = \frac{1}{n}(\mathsf{M}x_1 + \mathsf{M}x_2 + \ldots + \mathsf{M}x_n) = \frac{na}{n} = a,$$

$$\mathsf{M}\tilde{\theta}_n^* = \mathsf{M}x_1 = a.$$

Оценка $\tilde{\theta}_n^*$, очевидно, не является состоятельной. Найдем $\mathsf{D}\theta_n^*$. Так как $x_1, \ldots, x_n$ независимы, то

$$\mathsf{D}\theta_n^* = \frac{1}{n^2}(\mathsf{D}x_1 + \ldots + \mathsf{D}x_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n},$$

где $\sigma^2 = \mathsf{D}x_k$, $k = 1, 2, \ldots, n$. По неравенству Чебышева

$$\mathsf{P}(|\bar{x} - a| > \varepsilon) \leqslant \frac{\sigma^2}{\varepsilon^2 n}.$$

Отсюда следует, что $\theta_n^* = \bar{x}$ является состоятельной оценкой параметра $\theta = a = \mathsf{M}x_k$.

Пример 2. Найдем оценку функции распределения $F(x) = \mathsf{P}(x_k < x)$, $k = 1, 2, \ldots, n$, в точке $x$. Пусть $\mu_n(x)$ — число элементов выборки, меньших $x$. Для каждого элемента выборки $x_k$ может произойти одно из двух событий: $(x_k < x)$, $(x_k \geqslant x)$. Вероятности этих событий, очевидно, равны

$$p = \mathsf{P}(x_k < x) = F(x), \qquad q = \mathsf{P}(x_k \geqslant x) = 1 - F(x).$$

Таким образом, всевозможные произведения, составленные из событий указанных двух типов, образуют схему Бернулли, в которой $\mu_n(x)$ — число наступлений события $(x_k < x)$. Покажем, что функция

$$\hat{F}_n(x) = \frac{\mu_n(x)}{n} \tag{2.3}$$

является несмещенной и состоятельной оценкой $F(x)$. В схеме Бернулли $\mathsf{M}\mu_n = np$.

Следовательно,

$$\mathsf{M}\hat{F}_n(x) = \frac{1}{n}\mathsf{M}\mu_n(x) = \frac{1}{n}\cdot np = F(x).$$

Несмещенность доказана. Состоятельность оценки $\hat{F}_n(x)$ следует из теоремы 2.4 гл. 6.

**2.2. Метод наибольшего правдоподобия для нахождения оценок параметров. Метод моментов.** *Если функция распределения $F(x, \theta) = \mathsf{P}(x_k < x)$, $k = 1, \ldots, n$, имеет плотность $p(x, \theta)$, то функцией правдоподобия называется*

$$L = L(x_1, \ldots, x_n, \theta) = p(x_1, \theta)\, p(x_2, \theta) \ldots p(x_n, \theta). \tag{2.4}$$

*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*

$$L = L(x_1, x_2, \ldots, x_n, \theta) = p_{x_1}(\theta)\, p_{x_2}(\theta) \ldots p_{x_n}(\theta). \tag{2.5}$$

При фиксированных $x_1, \ldots, x_n$ функцию $L$ будем рассматривать как функцию параметра $\theta$. По методу наибольшего правдоподобия за оценку параметра $\theta$ принимается значение аргумента $\theta$, при котором $L$ имеет максимальное значение. Метод наибольшего правдоподобия для широкого класса распределений приводит к состоятельным оценкам, распределенным асимптотически нормально и имеющим наименьшую дисперсию по сравнению с другими оценками того же параметра.

Оценки наибольшего правдоподобия удовлетворяют уравнению

$$\frac{\partial \ln L}{\partial \theta} = 0,$$

если $\theta$ — одномерный параметр. Если $\theta = (\theta_1, \ldots, \theta_s)$, то оценки параметров $\theta_1, \ldots, \theta_s$ удовлетворяют системе уравнений

$$\frac{\partial \ln L}{\partial \theta_k} = 0, \quad k = 1, 2, \ldots, s.$$

Пример 3. Пусть величины $x_k$, $k = 1, 2, \ldots, n$, имеют нормальное распределение. Неизвестными параметрами являются $a = \mathsf{M}x_k$, $b = \sigma^2 = \mathsf{D}x_k$. Найдем их оценки наибольшего правдоподобия. По формуле (2.3)

$$L = L(x_1, \ldots, x_n, a, b) = \left(\frac{1}{\sqrt{2\pi b}}\right)^n \exp\left(-\sum_{k=1}^{n} \frac{(x_k - a)^2}{2b}\right)$$

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 $a^*$ и $b^*$ получим систему

$$\begin{cases} \dfrac{\partial \ln L}{\partial a} = \dfrac{1}{b^*}\displaystyle\sum_{k=1}^{n}(x_k - a^*) = 0, \\ \dfrac{\partial \ln L}{\partial b} = -\dfrac{n}{2b^*} + \dfrac{1}{2(b^*)^2}\displaystyle\sum_{k=1}^{n}(x_k - a^*)^2 = 0. \end{cases}$$

Из первого уравнения $a^* = \frac{1}{n}\sum_{k=1}^{n} x_k = \bar{x}$. Подставляя это значение во второе уравнение, найдем

$$b^* = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

Оценка $a^*$ рассматривалась в примере 1. Можно показать (см. задача 17 гл. 5), что $\mathsf{M}b^* = \frac{n-1}{n}b$. Оценка $\frac{n}{n-1}b^*$ является несмещенной и состоятельной (см. задачу 2 этой главы).

Пример 4. Найдем оценку наибольшего правдоподобия для вероятности $p$ успеха в схеме Бернулли. По формуле (4.4.16) $L=L(x_1, \ldots, x_n, p)=\prod_{k=1}^{n} p^{x_k}(1-p)^{1-x_k}$, $x_k=0, 1$, ($x_k=1$, 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

$$\frac{\partial \ln L}{\partial p}=\sum_{k=1}^{n}\left(\frac{x_k}{p^*}-\frac{1-x_k}{1-p^*}\right)=\frac{\bar{x}n}{p^*}-\frac{n}{1-p^*}-\frac{n\bar{x}}{1-p^*}=0$$

и $p^*=\bar{x}$. Так как для числа успехов $\mu_n$ имеем равенство $\mu_n=x_1+x_2+\ldots+x_n$, то $p^*=\frac{\mu_n}{n}$. Оценка $p^*$ является несмещенной. Состоятельность $p^*$ следует из теоремы 2.4 гл. 6. Из теоремы Муавра—Лапласа следует асимптотическая нормальность $p^*$.

Другим методом, который иногда используется для оценки параметров, является метод моментов. Пусть $x_1, x_2, \ldots, x_n$ — случайная выборка. Случайные величины

$$m_r=\frac{1}{n}\sum_{k=1}^{n}(x_k-\bar{x})^r$$

называют *выборочными* или *эмпирическими моментами*. В примере 1 было показано, что $\bar{x}$ является состоятельной оценкой $\mathsf{M}x_k$. Можно показать, что $m_r$ является состоятельной оценкой $\delta_r=\mathsf{M}(x_k-\mathsf{M}x_k)^r$. Если функция распределения $\mathsf{P}(x_k<x)$ зависит от $s$ неизвестных параметров $\theta_1, \theta_2, \ldots, \theta_s$, то $\delta_r=\delta_r(\theta_1, \ldots, \theta_s)$. Неизвестные параметры $\theta_i$ можно выразить в виде функций от теоретических моментов $\delta_r$. Если в этих функциях заменить $\delta_r$ на $m_r$, то получим оценки параметров $\theta_1, \theta_2, \ldots, \theta_s$.

**2.3. Интервальные оценки.** Предположим, что нам удалось найти две функции $\underline{\theta}=\underline{\theta}(x_1, \ldots, x_n)$ и $\bar{\theta}=\bar{\theta}(x_1, \ldots, x_n)$ такие, что $\underline{\theta}<\bar{\theta}$ при всех $x_1, \ldots$

..., $x_n$ и при любых значениях $\theta$

$$\mathsf{P}(\underline{\theta}(x_1, \ldots, x_n) < \theta < \overline{\theta}(x_1, \ldots, x_n)) = 1 - 2\alpha,$$

т. е. вероятность того, что случайный интервал $(\underline{\theta}, \overline{\theta})$ накроет неизвестный параметр $\theta$, не зависит от параметра $\theta$. В этом случае *интервал* $(\underline{\theta}, \overline{\theta})$ *называют доверительным интервалом для неизвестного параметра* $\theta$, *соответствующим доверительной вероятности* $1-2\alpha$. В ряде случаев функции $\underline{\theta}$ и $\overline{\theta}$, обладающие указанными свойствами, можно найти.

Пусть имеется выборка $x_1, \ldots, x_n$, величины $x_k$ нормально распределены с параметрами $(a, \sigma)$, параметр $\sigma$ известен. Найдем доверительный интервал для $a$. Случайная величина $\bar{x} = (x_1 + \ldots + x_n)/n$ имеет нормальное распределение и

$$\mathsf{M}\bar{x} = a, \qquad \mathsf{D}\bar{x} = \frac{\sigma^2}{n}.$$

Тогда величина

$$\frac{\bar{x} - a}{\sigma/\sqrt{n}}$$

распределена нормально с параметрами $(0, 1)$, и, следовательно, ее распределение не зависит от $a$. Определяя $u_\alpha$ как решение уравнения

$$\frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-\frac{x^2}{2}}\, dx = \alpha, \tag{2.6}$$

получим

$$\mathsf{P}\left(\left|\frac{\bar{x} - a}{\sigma/\sqrt{n}}\right| < u_\alpha\right) = 1 - 2\alpha$$

или

$$\mathsf{P}\left(\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - 2\alpha. \tag{2.7}$$

Таким образом, мы нашли доверительный интервал $\left(\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}},\ \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right)$ для параметра $a$.

Более естественной является ситуация, когда оба параметра $a$ и $\sigma$ неизвестны. В этом случае для $a$ тоже можно найти доверительный интервал. Дадим сначала определение распределений $\chi^2$ и Стьюдента.

Пусть $\xi$, $\xi_1$, $\xi_2$, ..., $\xi_n$ — независимые нормально распределенные случайные величины с параметрами (0, 1). Распределение случайной величины

$$\chi_n^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$$

называется распределением $\chi^2$ (*хи* — квадрат) с $n$ степенями свободы. Распределение величины $\tau_n = \xi \Big/ \sqrt{\chi_n^2 \cdot \frac{1}{n}}$ называется распределением Стьюдента с $n$ степенями свободы.

В книге Н. В. Смирнова, И. В. Дунина-Барковского [16] показано, что

$$\mathsf{P}\left(\bar{x} - t_{\alpha,\, n-1}\frac{s}{\sqrt{n-1}} < a < \bar{x} + t_{\alpha,\, n-1}\frac{s}{\sqrt{n-1}}\right) = 1 - \alpha, \tag{2.8}$$

где $t_{\alpha,\, n-1}$ определяется условием

$$\mathsf{P}(|\tau_{n-1}| < t_{\alpha,\, n-1}) = 1 - \alpha,$$

$\tau_{n-1}$ имеет распределение Стьюдента с $n-1$ степенью свободы,

$$s^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

Сравнивая (2.8) с (2.7), можно заметить, что замена в (2.7) параметра $\sigma^2$ его оценкой $s^2$ приводит к замене нормального распределения распределением Стьюдента.

Доверительные интервалы для параметра $a$ были приведены в случае, когда $x_k$ имеют нормальное распределение. При больших $n$ можно найти для параметра $a = \mathsf{M}x_k$ приближенные оценки интервального типа без предположения нормальности $x_k$. Пусть $\sigma^2 = \mathsf{D}x_k$. Тогда по центральной предельной теореме распределение величины

$$\eta_n = \frac{x_1 + x_2 + \ldots + x_n - na}{\sigma\sqrt{n}} = \frac{\bar{x} - a}{\sigma/\sqrt{n}}$$

близко к нормальному с параметрами (0, 1), и, следовательно,

$$\mathsf{P}\left(\left|\frac{\bar{x}-a}{\sigma/\sqrt{n}}\right| < u_\alpha\right) \approx 1-2\alpha$$

или

$$\mathsf{P}\left(\bar{x}-u_\alpha\frac{\sigma}{\sqrt{n}} < a < \bar{x}+u_\alpha\frac{\sigma}{\sqrt{n}}\right) \approx 1-2\alpha. \quad (2.9)$$

На практике в последнем приближенном равенстве неизвестный параметр $\sigma^2$ заменяют его оценкой $s^2$. Естественно это приводит к уменьшению точности (2.9).

## § 3. Статистическая проверка гипотез

**3.1. Проверка гипотез о законе распределения.** Рассмотрим следующую задачу. По выборке $x_1, x_2, \ldots, x_n$ нужно решить, является ли заданная функция $F(x)$ функцией распределения случайной величины $x_k$, $k=1, \ldots, n$. Ограничимся случаем, когда $F(x)$ непрерывна. Разобьем числовую ось на следующие интервалы:

$$(-\infty=z_0,\ z_1),\ (z_1,\ z_2),\ \ldots,\ (z_r,\ z_{r+1}=+\infty), \quad (3.1)$$

где $z_1 < z_2 < \ldots < z_{r-1} < z_r$. Если величины $x_k$ имеют своей функцией распределения $F(x)$, то можно найти следующие вероятности:

$$p_1 = \mathsf{P}(x_k \in (z_0,\ z_1)) = F(z_1),$$
$$p_l = \mathsf{P}(x_k \in (z_l,\ z_{l+1})) = F(z_{l+1}) - F(z_l),$$
$$l = 1, 2, \ldots, r-1,$$
$$p_r = \mathsf{P}(x_k \in (z_r,\ z_{r+1})) = 1 - F(z_r).$$

Со случайными величинами $x_1, x_2, \ldots, x_n$ естественно связана полиномиальная схема с $n$ испытаниями, в которой результатом $k$-го испытания является попадание значения $x_k$ в какой-либо интервал (3.1). Обозначим $m_l = m_l(x_1, x_2, \ldots, x_n)$ число значений среди $x_1(\omega), x_2(\omega), \ldots, x_n(\omega)$, попавших в $(z_l, z_{l+1})$. Если наше предположение о законе распределения $x_k$ верно, то $m_l$ должны быть близки к

$$\mathsf{M}m_l = np_l, \qquad l = 1, 2, \ldots, r.$$

Общее отклонение всех $m_i$ можно измерять, например, суммой

$$\sum_{i=1}^{r} |m_i - np_i|.$$

Обычно в качестве меры отклонения используют

$$\eta_{n,r} = \sum_{i=1}^{r} \frac{(m_i - np_i)^2}{np_i}. \qquad (3.2)$$

Оказывается (см. [16], гл. 7, § 3), что для любого $x$ при $n \to \infty$

$$\mathsf{P}(\eta_{n,r} < x) \to \mathsf{P}(\chi^2_{r-1} < x), \qquad (3.3)$$

где случайная величина $\chi^2_{r-1}$ имеет распределение $\chi^2$ с *$r-1$ степенями свободы*; $r$, $p_1 > 0$, ..., $p_r > 0$ постоянны. Используя (3.3), можно выбрать такое «критическое» значение $C$, что при нашей гипотезе вероятность события

$$\eta_{n,r} > C \qquad (3.4)$$

очень мала и его можно считать практически невозможным. Если же событие (3.4) фактически наблюдалось, то говорят, что выборка обнаруживает значимое отклонение от проверяемой гипотезы. Это указывает на несовместимость нашей гипотезы с наблюденными значениями $x_1(\omega), \ldots, x_n(\omega)$. Таким образом, правило проверки или *статистический критерий состоит в том, что гипотеза отвергается, если произошло событие* (3.4) *и гипотеза не противоречит наблюдениям, если произошло противоположное событие.*

Вероятность события (3.4) $\alpha = \mathsf{P}(\eta_{n,r} > C)$ в случае, когда гипотеза верна, является вероятностью отвергнуть правильную гипотезу. Если задать $\alpha$, то число $C$ можно, воспользовавшись (3.3), приближенно определить из уравнения $\alpha = \mathsf{P}(\chi^2_{r-1} > C)$. В конце книги приведены таблицы чисел $\chi^2_{m,\alpha}$, для которых $\mathsf{P}(\chi^2_m > \chi^2_{m,\alpha}) = \alpha$. Выбор $\alpha$ зависит от рассматриваемой практической задачи. Часто допускают значения $\alpha = 0{,}01$, $\alpha = 0{,}05$. Если используется критерий с

$\alpha=0{,}01$, то примерно в $1\%$ случаев будет отвергаться правильная гипотеза.

**3.2. Выбор из двух гипотез.** Пусть в случайной выборке $x_1, \ldots, x_n$ величины $x_k$ нормально распределены с параметрами $(a, \sigma)$. Параметр $\sigma$ известен, а относительно $a$ имеется два предположения, или две *гипотезы* $H_0$ *и* $H_1$, согласно которым $a=a_0$ и $a=a_1$ $(a_0<a_1)$ соответственно. Одну из гипотез называют основной (например, $H_0$) или нулевой, а другую — конкурирующей. В § 2 было показано, что несмещенной и состоятельной оценкой математического ожидания $a$ является $\bar{x}=(x_1+\ldots+x_n)/n$. Таким образом, в рассматриваемой задаче естественно выбрать ту гипотезу, к параметру $a$ которой ближе $\bar{x}$. При любой из гипотез ($H_0$ или $H_1$) $\bar{x}$ имеет нормальное распределение с $\mathsf{D}\bar{x}=\frac{\sigma^2}{n}$; $\mathsf{M}\bar{x}=a$ зависит от гипотезы. Следовательно, распределения $\bar{x}$ при $H_0$ и $H_1$ различны. Вероятности событий, вычисляемые при гипотезах $H_0$ и $H_1$, будем обозначать соответственно $\mathsf{P}_0$ и $\mathsf{P}_1$.

Выберем некоторое число $C$, $a_0<C<a_1$. Критерий можно сформулировать следующим образом: если $\bar{x}>C$, то принимается $H_1$, а если $\bar{x}<C$, то принимается $H_0$.

Если верна гипотеза $H_0$ и произошло событие $\bar{x}>C$, то принимается $H_1$. Вероятность отвергнуть $H_0$, когда она верна, равна

$$\alpha=P_0(\bar{x}>C).$$

Если верна гипотеза $H_1$ и произошло событие $\bar{x}<C$, то принимается $H_0$. Вероятность принять $H_0$, когда верна $H_1$, равна

$$\beta=\mathsf{P}_1(\bar{x}<C).$$

*Вероятности $\alpha$ и $\beta$ называют вероятностями ошибок 1-го и 2-го рода соответственно. Вероятность принять гипотезу $H_1$, когда она верна, называют мощностью критерия.* Очевидно, мощность равна $1-\beta$. Среди критериев с фиксированной вероятностью $\alpha$ естественно выбирать критерий с большей мощностью.

Пусть $\alpha$ задано. Тогда

$$\alpha = P_0(\bar{x} > C) = P_0\left(\frac{\bar{x}-a_0}{\sigma/\sqrt{n}} > (C-a_0)\frac{\sqrt{n}}{\sigma}\right).$$

Так как $\sqrt{n}(\bar{x}-a_0)/\sigma$ имеет нормальное распределение с параметрами (0, 1), то $(C-a_0)\frac{\sqrt{n}}{\sigma} = u_\alpha$, где $u_\alpha$ определено уравнением (2.6). Отсюда

$$C = a_0 + u_\alpha \frac{\sigma}{\sqrt{n}}. \tag{3.5}$$

Тогда ошибка второго рода равна

$$\beta = P_1(\bar{x} < C) = P_1\left(\frac{\bar{x}-a_1}{\sigma/\sqrt{n}} < (C-a_1)\frac{\sqrt{n}}{\sigma}\right).$$

Отсюда

$$(C-a_1)\frac{\sqrt{n}}{\sigma} = u_{1-\beta} = -u_\beta.$$

Подставляя в это равенство значение $C$ из (3.5), получим

$$u_\alpha + u_\beta = \frac{a_1-a_0}{\sigma}\sqrt{n}. \tag{3.6}$$

Для заданных $\alpha$ и $\beta$ равенство (3.6) используется для нахождения $n$. Если $n$ менять нельзя, то (3.6) используется для выбора $\alpha$ и $\beta$. При заданных $\alpha$ и $n$ величина $\beta$ определяется однозначно. Величины $\alpha$ и $\beta$ выбираются в зависимости от степени нежелательности связанных с ними ошибочных решений. Иногда «степень нежелательности» можно выразить довольно точно. Пусть, например, при некоторой упрощенной проверке бракованное изделие может быть пропущено с вероятностью $\alpha$ и хорошее изделие принято за бракованное с вероятностью $\beta$. Если бракованное изделие продано, то за его гарантийный ремонт нужно платить $P$ рублей. Если хорошее изделие забраковано, то теряется его стоимость $Q$ рублей. Пусть в проверяемой партии из $N$ изделий примерно $M$ бракованных. Тогда средние потери при контроле этой партии равны

$$(M\alpha)\cdot P + (N-M)\beta Q. \tag{3.7}$$

*Для выбора $\alpha$ и $\beta$ нужно найти минимум* (3.7) *при наличии связи* (3.6).

Перейдем теперь к задаче выбора модели опыта, состоящего в подбрасывании двух монет. Обозначим $\mu_n$ число выпадений двух монет одинаковой стороной в $n$ опытах. Вместо моделей 1 и 2 мы будем теперь говорить о гипотезах $H_1$ и $H_0$. Рассмотрим величину $\mu_n/n$. Так как $\mu_n/n$ имеет при гипотезе $H_l$, $l=0,1$, биномиальное распределение, то

$$\mathsf{M}\frac{\mu_n}{n}=a_l,\qquad \mathsf{D}\frac{\mu_n}{n}=\frac{\sigma_l}{n},$$

где

$$a_1=\frac{2}{3},\quad \sigma_1=\frac{2}{3}\cdot\frac{1}{3};\quad a_0=\frac{1}{2},\quad \sigma_0^2=\frac{1}{4}. \tag{3.8}$$

Если $\mu_n/n>C$, то будем принимать $H_1$, а в противном случае $H_0$. Если биномиальное распределение заменить аппроксимирующим его нормальным, то для вычисления $C$ сохранится формула (3.5) с $\sigma=\sigma_0$, а вместо (3.6) получим

$$\sigma_0u_\alpha+\sigma_1u_\beta=(a_1-a_0)\sqrt{n}. \tag{3.9}$$

При $\alpha=\beta=0{,}05$ имеем $u_\alpha=u_\beta=1{,}65$. Подставляя эти значения и (3.8) в (3.9), получим

$$n\geqslant\frac{(\sigma_0u_\alpha+\sigma_lu_\beta)^2}{(a_1-a_0)^2}\approx\left(\frac{1{,}65\left(\frac{1}{2}+\frac{\sqrt{2}}{3}\right)}{\frac{2}{3}-\frac{1}{3}}\right)^2\approx 79{,}2,$$

и, следовательно, можно положить $n=80$. По формуле (3.5)

$$C=0{,}5+1{,}65\frac{1}{2\sqrt{80}}\approx 0{,}61.$$

Таким образом, гипотеза $H_1$ принимается, если $\mu_n/n>0{,}61$, в противном случае принимается $H_0$.

**Задачи к главе 10**

1. Используя таблицу случайных чисел, получить реализацию случайной выборки $x_1, x_2, \ldots, x_n$, где $x_k$ равномерно распределены на отрезке $[0, 1]$. Вычислить

$$a^* = \bar{x} = \frac{x_1 + \ldots + x_n}{n}, \qquad (\sigma^2)^* = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2.$$

Сравнить с $\mathsf{M}x_k$ и $\mathsf{D}x_k$. Значения $x_k$ взять с точностью до $10^{-3}$, $n = 100$.

2*. Пусть $x_1, x_2, \ldots, x_n$ — произвольная случайная выборка с $\mathsf{M}x_k = a$, $\mathsf{D}x_k = \sigma^2$, $\mathsf{M}(x_k - a)^4 < \infty$. Показать (см. задачу 14 к гл. 5), что

$$(\sigma^2)^* = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2$$

является несмещенной оценкой $\sigma^2$. Доказать, что эта оценка является состоятельной. Указание: воспользовавшись равенством

$$(\sigma^2)^* = \frac{1}{n-1}\sum_{k=1}^{n}(y_k - \bar{y})^2,$$

где $y_k = x_k - a$, $\bar{y} = (y_1 + \ldots + y_n)/n$, доказать, что

$$\mathsf{D}\left((\sigma^2)^*\right) = \frac{\delta}{n} + O\left(\frac{1}{n^2}\right),$$

$\delta > 0$ — некоторая постоянная.

3. Пусть $x_1, x_2, \ldots, x_n$ — случайная выборка, в которой $x_k$ равномерно распределена на отрезке $[a, b]$. Являются ли оценки

$$a^* = \min\{x_1, \ldots, x_n\}, \qquad b^* = \max\{x_1, \ldots, x_n\}$$

несмещенными оценками $a$ и $b$ соответственно? Указание: воспользоваться решением задачи 18 гл. 4.

4. Пусть $x_1, x_2, \ldots, x_n$ — случайная выборка, в которой плотность распределения $x_k$ равна $e^{\alpha - x}$ при $x \geqslant \alpha$. Является ли оценка $\alpha^* = -\frac{1}{n} + \min(x_1, x_2, \ldots, x_n)$ несмещенной и состоятельной оценкой $\alpha$?

5. Используя метод наибольшего правдоподобия, найти по выборке $x_1, x_2, \ldots, x_n$, где $\mathsf{P}(x_k = m) = \frac{\lambda^m}{m!}e^{-\lambda}$, оценку параметра $\lambda$. Будет ли эта оценка несмещенной и состоятельной?

6. Переменные $x$ и $y$ связаны линейной зависимостью $y = \alpha + \beta x$. Для определения неизвестных параметров $\alpha$ и $\beta$ при заданных значениях $x_i$ измерены значения

$$y_i = \alpha + \beta x_i + \delta_i, \qquad i = 1, \ldots, n,$$

где ошибки измерения $\delta_i$ независимы, нормально распределены, $M\delta_i = 0$, $D\delta_i = \sigma^2$. Функция правдоподобия

$$L(y_1, \ldots, y_n; x_1, \ldots, x_n; \alpha, \beta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2}Q},$$

где $Q = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2$. Минимизируя $Q$, найти оценки наибольшего правдоподобия параметров $\alpha$, $\beta$, $\sigma^2$ в предположении, что $\sum_{i=1}^{n} x_i = 0$. (Если условие $\sum x_i = 0$ не выполнено, то можно перейти к новым переменным $x_i' = x_i - \left(\sum_{i=1}^{n} x_i\right)\Big/ n$.)

ГЛАВА 11

# ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

## § 1. Понятие о случайных процессах

Случайные процессы являются моделями многих реальных процессов. Два примера таких реальных процессов (работа телефонной станции и броуновское движение) обсуждались в § 2 гл. 1.

*Случайным процессом* называется семейство случайных величин $\xi_t = \xi_t(\omega)$, заданных на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$. Параметр $t$ обычно интерпретируется как время. Если $t \in (0, 1, 2, \ldots)$, то говорят, что $\xi_t$ — процесс с дискретным временем; если же $t \in [0, T]$ то $\xi_t$ — процесс с непрерывным временем.

Действительную функцию $\xi_t(\omega_0)$ при фиксированном $\omega_0$ называют *реализацией* или *траекторией случайного процесса*. Если фиксировать $t = t_0$, то $\xi_{t_0}(\omega)$ является обычной случайной величиной. Функции распределения

$$F_t(x) = \mathsf{P}(\xi_t < x)$$

задают распределение значений процесса в каждый момент времени. Зная только $F_t(x)$ при всех $t$, мы еще ничего не можем сказать о зависимости значений процесса в какие-либо фиксированные моменты времени. Таким образом, естественно задать всевозможные совместные распределения

$$F_{t_1 t_2 \ldots t_n}(x_1, x_2, \ldots, x_n) = \mathsf{P}(\xi_{t_1} < x_1, \ldots, \xi_{t_n} < x_n), \tag{1.1}$$

где $0 \leqslant t_1 < t_2 < \ldots < t_n$, $n = 1, 2, \ldots$ В гл. 4 мы показали, что можно построить вероятностное пространство и определить на нем случайную величину с заданной функцией распределения. Аналогичная задача построения подходящего вероятностного пространства для семейства случайные величин $\xi_t$ с заданными распределениями (1.1) является значительно более сложной задачей и выходит за рамки настоящей книги (см. А. Н. Колмогоров [8]). Частный случай этой задачи был разобран в § 4 гл. 3.

В следующих параграфах мы опишем несколько типов случайных процессов и вычислим для них вероятности отдельных событий. Тип процесса будет определяться свойствами, которыми должны обладать величины $\xi_t$. Построения подходящих вероятностных пространств, на которых можно определить $\xi_t$ с заданными свойствами, приведены не будут.

## § 2. Пуассоновский процесс

Случайный процесс $\xi_t$ с непрерывным временем называется *процессом с независимыми приращениями*, если при любых $0 \leqslant t_1 < t_2 < \ldots < t_n$, $t_1, \ldots, t_n \in [0, \infty)$, $n = 1, 2, \ldots$, случайные величины

$$\xi_{t_1}, \ \xi_{t_2} - \xi_{t_1}, \ \ldots, \ \xi_{t_n} - \xi_{t_{n-1}}$$

независимы. *Пуассоновским процессом называется процесс $\xi_t$, удовлетворяющий следующим условиям:*

1°. *$\xi_t$ — процесс с независимыми приращениями.*

2°. *При любых $t_1 < t_2$, $s$ приращения $\xi_{t_2} - \xi_{t_1}$, $\xi_{t_2+s} - \xi_{t_1+s}$ одинаково распределены (однородность по времени).*

3°. $\xi_0(\omega) = 0$, $\omega \in \Omega$.

4°. *При* $h \to 0$

$$\mathsf{P}(\xi_h = 0) = 1 - \lambda h + o(h), \qquad \mathsf{P}(\xi_h \geqslant 2) = o(h),$$
$$\mathsf{P}(\xi_h = 1) = \lambda h + o(h), \qquad 0 < \lambda < \infty.$$

Пуассоновский процесс $\xi_t$ можно задать на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathsf{P})$, где множество $\Omega$ совпадает с множеством ступенчатых функций, у которых имеются только единичные скачки, а моменты времени, соответствующие скачкам, не имеют точек накопления.

Теорема 2.1. *Если $\xi_t$ — пуассоновский процесс, то*

$$\mathsf{P}(\xi_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \ldots$$

Доказательство. Пусть $\varphi_t(x)=\mathsf{M}x^{\xi_t}$ — производящая функция $\xi_t$. По свойству 1° величины $\xi_{t+h}-\xi_t$, $\xi_t$ независимы. Следовательно,

$$\varphi_{t+h}(x)=\mathsf{M}x^{\xi_{t+h}}=\mathsf{M}x^{(\xi_{t+h}-\xi_t)+\xi_t}=\varphi_t(x)\,\mathsf{M}x^{\xi_{t+h}-\xi_t}.$$

Так как по свойству 2° величины $\xi_{t+h}-\xi_t$ и $\xi_h-\xi_0=\xi_h$ одинаково распределены, то

$$\mathsf{M}x^{\xi_{t+h}-\xi_t}=\mathsf{M}x^{\xi_h}$$

и при $h\to 0$ согласно 4°

$$\mathsf{M}x^{\xi_h}=1-\lambda h+x\lambda h+o(h),\qquad |x|\leqslant 1.$$

Таким образом,

$$\varphi_{t+h}(x)=\varphi_t(x)\,(1-\lambda h+x\lambda h+o(h)).$$

Отсюда

$$\frac{\varphi_{t+h}(x)-\varphi_t(x)}{h}=-\lambda(x-1)\,\varphi_t(x)+o(1).$$

Переходя к пределу при фиксированном $x$ и $h\to 0$, получим

$$\frac{d\varphi_t(x)}{dt}=-\lambda(x-1)\,\varphi_t(x).$$

Решение этого уравнения с начальным условием $\varphi_0(x)=1$ определяется формулой

$$\varphi_t(x)=e^{\lambda t(x-1)}=\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}e^{-\lambda t}x^k.$$

Полученная производящая функция является производящей функцией распределения Пуассона. Теорема доказана.

Используя доказанную теорему, можно найти совместные распределения $\xi_{t_1},\ldots,\xi_{t_n}$ при любых $t_1<\ldots<t_n$. Очевидно, что

$$(\xi_{t_1}=k_1,\ \xi_{t_2}=k_2,\ \ldots,\ \xi_{t_n}=k_n)=$$
$$=(\xi_{t_1}=k_1,\ \xi_{t_2}-\xi_{t_1}=k_2-k_1,\ \ldots,\ \xi_{t_n}-\xi_{t_{n-1}}=k_n-k_{n-1}), \tag{2.1}$$

где $k_1 \leqslant k_2 \leqslant \ldots \leqslant k_n$. Отсюда, учитывая независимость и однородность приращений, получим

$$\mathsf{P}(\xi_{t_1}=k_1, \ldots, \xi_{t_n}=k_n) = \frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \cdot \frac{(\lambda(t_2-t_1))^{k_2-k_1}}{(k_2-k_1)!} e^{-\lambda(t_2-t_1)} \cdot \ldots \cdot \frac{(\lambda(t_n-t_{n-1}))^{k_n-k_{n-1}}}{(k_n-k_{n-1})} e^{-\lambda(t_n-t_{n-1})}. \tag{2.2}$$

Два свойства пуассоновского процесса сформулированы в задачах 4 и 5.

## § 3. Винеровский процесс

В § 2 был рассмотрен процесс, в котором изменения происходят скачками. Здесь мы определим процесс, имеющий непрерывные траектории. *Винеровским процессом называется случайный процесс $\xi_t$, удовлетворяющий условиям:*

1°. *$\xi_t$ — процесс с независимыми приращениями.*

2°. *При любых $t_1 < t_2$, $s$ приращения $\xi_{t_2}-\xi_{t_1}$; $\xi_{t_2+s}-\xi_{t_1+s}$ одинаково распределены.*

3°. $\xi_0(\omega)=0$, $\omega \in \Omega$.

4°. *При $h \to 0$*

$$\mathsf{M}\xi_h = ah + o(h), \qquad \mathsf{M}|\xi_h|^3 = o(h).$$

$$\mathsf{M}\xi_h^2 = bh + o(h), \qquad -\infty < a < \infty, \quad 0 < b < \infty.$$

Теорема 3.1. *Если $\xi_t$ — винеровский процесс, то*

$$\mathsf{P}(\xi_t < x) = \frac{1}{\sqrt{2\pi bt}} \int_{-\infty}^{x} e^{-\frac{(u-at)^2}{2bt}}\, du.$$

Доказательство. Положим $f_t(s) = \mathsf{M}e^{is\xi_t}$. Здесь аргумент характеристической функции обозначен буквой $s$. Пусть $s$ фиксировано. Так же, как в доказательстве теоремы 2.1, найдем

$$f_{t+h}(s) = f_t(s)\,\mathsf{M}e^{is(\xi_{t+h}-\xi_t)} = f_t(s)\,\mathsf{M}e^{is\xi_h}.$$

По формуле (7.2.11)

$$\mathsf{M}e^{is\xi_h} = 1 + is\mathsf{M}\xi_h - \frac{s^2}{2}\mathsf{M}\xi_h^2 + O(|s|^3\mathsf{M}|\xi_h|^3).$$

Отсюда и из свойства 4° следует

$$\mathsf{M}e^{is\xi_h} = 1 + is(ah) - \frac{s^2bh}{2} + o(h).$$

Таким образом,

$$f_{t+h}(s) = \left(1 + is(ah) - \frac{s^2bh}{2} + o(h)\right) f_t(s)$$

и

$$\frac{f_{t+h}(s) - f_t(s)}{h} = \left(isa - \frac{s^2b}{2}\right) f_t(s) + o(1).$$

Переходя в этом равенстве к пределу при $h \to 0$, получим

$$\frac{df_t(s)}{dt} = \left(isa - \frac{s^2b}{2}\right) f_t(s).$$

Отсюда, учитывая начальное условие $f_0(s) = 1$, найдем

$$f_t(s) = e^{is(at) - \frac{(bt)s^2}{2}}.$$

Полученная характеристическая функция величины $\xi_t$ соответствует нормальному распределению с параметрами $(at, \sqrt{bt})$. Теорема доказана.

Найдем совместное распределение $\xi_{t_1}, \ldots, \xi_{t_n}$. Совместное распределение приращений $\eta_k = \xi_{t_{k+1}} - \xi_{t_k}$, $k = 1, 2, \ldots, n$, $(\xi_{t_0} = 0)$ легко выписывается. Тогда, используя равенства

$$\xi_{t_1} = \eta_1, \quad \xi_{t_2} = \eta_1 + \eta_2, \ \ldots, \quad \xi_{t_n} = \eta_1 + \eta_2 + \ldots + \eta_n,$$

получим

$$\mathsf{P}(\xi_{t_1} < x_1, \ \xi_{t_2} < x_2, \ \ldots, \ \xi_{t_n} < x_n) =$$
$$= \mathsf{P}(\eta_1 < x_1, \ \eta_1 + \eta_2 < x_2, \ \ldots, \ \eta_1 + \ldots + \eta_n < x_n) =$$
$$= \int\limits_{a} \cdots \int \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi b(t_k - t_{k-1})}} e^{-\frac{[u_k - a(t_k - t_{k-1})]^2}{2b(t_k - t_{k-1})}} du_1 \ldots du_n,$$

где

$$G=\{(u_1, \ldots, u_n)\colon\ u_1<x_1,\ u_1+u_2<x_2, \ldots \\ \ldots, u_1+u_2+\ldots+u_n<x_n\}.$$

## § 4. Ветвящийся процесс

Ветвящийся процесс — это процесс размножения и превращения частиц, в котором частицы развиваются независимо друг от друга. Дадим определение ветвящегося процесса с дискретным временем. Пусть

$$\xi_k(t), \quad k=1, 2, \ldots, n, \ldots; \quad t=0, 1, 2, \ldots,$$

— независимые одинаково распределенные целочисленные случайные величины с общей производящей функцией

$$\varphi(x)=\sum_{m=0}^{\infty} x^m p_m.$$

Величина $\xi_k(t)$ будет интерпретироваться как число непосредственных потомков частицы с номером $k$, существовавшей в момент $t$. Случайный процесс $\mu_t$, определяемый равенствами

$$\mu_0=1, \quad \mu_{t+1}=\begin{cases} \xi_1(t)+\xi_2(t)+\ldots+\xi_{\mu_t}(t), & \text{если } \mu_t>0, \\ 0, & \text{если } \mu_t=0, \end{cases}$$

называется *ветвящимся*.

При исследовании ветвящихся процессов часто используются производящие функции. Положим

$$\Phi_t(x)=\mathsf{M}x^{\mu_t}.$$

Отметим, что

$$\Phi_0(x)=x, \qquad \Phi_1(x)=\varphi(x).$$

Так как $\mu_{t+1}$ является суммой случайного числа независимых слагаемых, то, положив в теореме 1.4 гл. 7,

$$\varphi_{\zeta_\nu}(x)=\Phi_{t+1}(x), \qquad \varphi_\nu(x)=\Phi_t(x), \qquad \varphi_{\xi_1}(x)=\varphi(x),$$

получим

$$\Phi_{t+1}(x)=\Phi_t(\varphi(x)). \tag{4.1}$$

Это функциональное уравнение позволяет найти производящую функцию $\Phi_t(x)$ при любом $t$. Полагая $t=1, 2, \ldots$ в (4.1), получим

$$\Phi_2(x)=\varphi(\varphi(x)), \qquad \Phi_3(x)=\varphi(\varphi(\varphi(x))), \ldots$$

Очевидно, что $\Phi_t(x)$ является $t$-й итерацией функции $\varphi(x)$: $\Phi_t(x)=\varphi(\varphi \ldots \varphi(x) \ldots)$.

Найдем $A_t=M\mu_t$. Продифференцируем (4.1) по $x$:

$$\frac{d\Phi_{t+1}(x)}{dx}=\frac{d\Phi_t}{d\varphi}\frac{d\varphi}{dx}.$$

Отсюда, полагая $x=1$ и $a=A_1=\varphi'(1)$, получим $A_{t+1}=A_t a$ и, следовательно, $A_t=a^t$. Поведение $A_t$ при $t\to\infty$ качественно различается в следующих трех случаях: $a<1$, $a=1$, $a>1$. При $a<1$ среднее число частиц $A_t$ стремится к 0 при $t\to\infty$, $A_t=1$ при $a=1$ и $A_t\to\infty$ в случае $a>1$. Ветвящиеся процессы с $a<1$, $a=1$ и $a>1$ называют соответственно *докритическими*, *критическими* (если $\varphi(x)\neq x$) и *надкритическими*. Асимптотические свойства ветвящихся процессов в этих трех случаях существенно различны.

Исследуем условия вырождения процесса. Обозначим $C$ событие, состоящее в том, что $\mu_t=0$ при некотором $t$. Очевидно, что

$$(\mu_t=0)\subset(\mu_{t+1}=0), \qquad t=1, 2, \ldots \tag{4.2}$$

Событие $C$ можно представить в виде $C=\bigcup\limits_{t=1}^{\infty}(\mu_t=0)$. Согласно (1.5.9)

$$\lambda=P(C)=\lim_{t\to\infty}\mathsf{P}(\mu_t=0).$$

Если $\lambda=1$, то процесс называется *вырождающимся*. Докритические процессы ($a<1$) являются вырождающимися, так как при $t\to\infty$

$$1-\mathsf{P}(\mu_t=0)=\mathsf{P}(\mu_t>0)=\mathsf{P}\left(\mu_t>\frac{1}{2}\right)\leqslant 2M\mu_t=2a^t\to 0.$$

Покажем, что вероятность $\lambda$ удовлетворяет уравнению

$$\varphi(x)=x. \tag{4.3}$$

Так как $\Phi_t(x)$ является $t$-й итерацией $\varphi(x)$, то наряду с (4.1) имеет место равенство

$$\Phi_{t+1}(x)=\varphi(\Phi_t(x)). \tag{4.4}$$

Очевидно, что $\Phi_t(0)=\mathsf{P}(\mu_t=0)$. Полагая $x=0$ в (4.4), получим

$$\mathsf{P}(\mu_{t+1}=0)=\varphi(\mathsf{P}(\mu_t=0)). \tag{4.5}$$

Отсюда при $t\to\infty$ найдем

$$\lambda=\varphi(\lambda). \tag{4.6}$$

Таким образом, $\lambda$ является корнем уравнения (4.3). В тех случаях, когда (4.3) имеет единственный корень

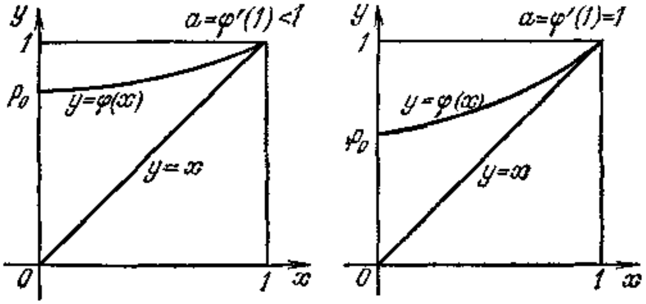


Рис. 8.

на отрезке $[0, 1]$, этот корень совпадает с $\lambda$. Пусть $\varphi(x)\neq x$. При $x\in[0, 1]$ имеем $\varphi'(x)>0$, $\varphi''(x)>0$. Следовательно, $\varphi(x)$ на отрезке $[0, 1]$ возрастает и обращена выпуклостью книзу. Параметр $a=\varphi'(1)$ определяет угловой коэффициент касательной к графику $y=\varphi(x)$ в точке $x=1$. Следовательно, при $a=1$ график $y=\varphi(x)$ касается $y=x$ в точке $x=1$; при $a<1$ касательная к $y=\varphi(x)$ проходит выше $y=x$ (рис. 8). Таким образом, при $a\leqslant 1$ уравнение (4.3) имеет на отрезке $[0, 1]$ единственный корень $x=1$, и, следовательно, ветвящийся процесс вырождается, если $a\leqslant 1$.

При $a>1$ касательная к графику $y=\varphi(x)$ проходит ниже $y=x$ (рис. 9) и уравнение (4.3) имеет на отрезке $[0,\ 1]$ два корня $x_1,\ x_2$ $(0 \leqslant x_1 < x_2 = 1)$. Покажем, что $\lambda = x_1$. Для этого достаточно показать, что $\mathsf{P}(\mu_t=0) \leqslant x_1$ при любом $t$. Проведем доказательство по индукции. При $t=1$

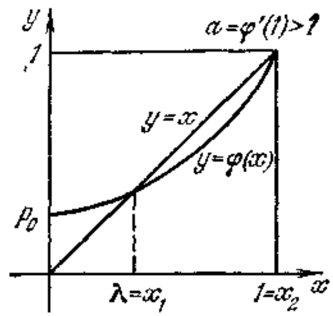


Рис. 9.

$$\mathsf{P}(\mu_1=0)=p_0 \leqslant p_0+p_1x_1+p_2x_1^2+\ldots=\varphi(x_1)=x_1,$$

так как $x_1$ — корень уравнения (4.3). Предположим, что $\mathsf{P}(\mu_t=0) \leqslant x_1$. Тогда, воспользовавшись (4.5), получим

$$\mathsf{P}(\mu_{t+1}=0)=\varphi(\mathsf{P}(\mu_t=0)) \leqslant \leqq \varphi(x_1)=x_1.$$

Таким образом, при любом $t$

$$\mathsf{P}(\mu_t=0) \leqslant x_1.$$

Переходя к пределу в этом неравенстве, получим $\lambda \leqslant x_1$. Отсюда следует, что $\lambda = x_1$, так как $\lambda$ и $x_1$ — корни уравнения (4.3) и $x_1$ — наименьший корень.

Таким образом, в случае надкритических процессов $\lambda = P(C) < 1$ и, следовательно, надкритические процессы не являются вырождающимися.

Для вырождающихся процессов можно ввести случайную величину $\tau$ — время до вырождения процесса. Функция распределения $\tau$ есть

$$\mathsf{P}(\tau < t) = \mathsf{P}(\mu_t = 0).$$

Найдем асимптотическую формулу при $t \to \infty$ для

$$Q(t) = \mathsf{P}(\tau \geqslant t) = 1 - \mathsf{P}(\mu_t = 0).$$

Уравнение (4.5) для $Q(t)$ запишется в виде

$$Q(t+1) = 1 - \varphi(1 - Q(t)). \qquad (4.7)$$

Рассмотрим сначала докритический случай. Для докритических процессов $a = \varphi'(1) < 1$. Предположим еще, что конечен второй факториальный момент

$$b = \mathsf{M}\mu_1(\mu_1 - 1) = \varphi''(1). \qquad (4.8)$$

Используя формулу Тейлора, запишем (4.7) в виде

$$Q(t+1)=1-\varphi(1)+\varphi'(1)\,Q(t)-\\ -\frac{Q^2(t)}{2}\varphi''(1-\theta Q(t)),\quad 0<\theta<1.$$

Отсюда

$$Q(t+1)=aQ(t)-\frac{b_t}{2}Q^2(t), \tag{4.9}$$

где

$$b_t=\varphi''(1-\theta Q(t)),\quad 0<\theta<1, \tag{4.10}$$

и

$$\frac{Q(s+1)}{Q(s)}=a\left(1-\frac{b_s}{2a}Q(s)\right),\quad s=1,2,\ldots.$$

Перемножив эти равенства от $s=t_0$ до $s=t-1$ $(0<t_0<t-1)$, получим

$$Q(t)=K_t a^t,$$

где

$$K_t=Q(t_0)\prod_{s=t_0}^{t-1}\left(1-\frac{b_s}{2a}Q(s)\right). \tag{4.11}$$

Так как

$$Q(s)=\mathsf{P}\left(\mu_s\geqslant\frac{1}{2}\right)\leqslant 2\mathsf{M}\mu_s=2a^s,\quad 0<a<1, \tag{4.12}$$

и $b_s\to b$ при $s\to\infty$, то при достаточно большом $t_0$ сомножители в (4.11) положительны. Из неравенства (4.12) следует сходимость ряда

$$\sum_{s=t_0}^{\infty}\ln\left(1-\frac{b_s}{2a}Q(s)\right)=\ln K\quad 0<K<\infty,$$

и оценка

$$\sum_{s=t}^{\infty}\ln\left(1-\frac{b_s}{2a}Q(s)\right)=O(a^t),\quad t\to\infty.$$

Отсюда при $t\to\infty$

$$K_t=K(1+O(a^t)).$$

Таким образом, *если* $0<a<1$ *и конечен момент* (4.8), *то при* $t\to\infty$

$$Q(t)=Ka^t(1+O(a^t)).$$

Рассмотрим критический процесс. В этом случае $a=1$, $b>0$ и уравнение (4.9) запишется в виде

$$Q(t+1)=Q(t)-\frac{b_t}{2}Q^2(t), \tag{4.13}$$

где $b_t \to b$ при $t\to\infty$. Отсюда

$$\frac{Q(t+1)}{Q(t)}=1-\frac{b_t}{2}Q(t)\to 1, \quad t\to\infty. \tag{4.14}$$

Равенство (4.13) запишем в виде

$$Q(t+1)=Q(t)-\frac{b}{2}Q(t)Q(t+1)+\varepsilon(t), \tag{4.15}$$

где

$$\varepsilon(t)=\frac{b}{2}Q(t)Q(t+1)-\frac{b_t}{2}Q^2(t).$$

Используя (4.14), при $t\to\infty$ получим

$$\frac{\varepsilon(t)}{Q(t)Q(t+1)}=\frac{b}{2}-\frac{b_t}{2}\frac{Q(t)}{Q(t+1)}\to 0. \tag{4.16}$$

Отсюда и из (4.15) вытекает, что

$$\frac{1}{Q(s+1)}=\frac{1}{Q(s)}+\frac{b}{2}+\delta(s),$$

где $\delta(s)\to 0$ при $s\to\infty$.

Суммируя последнее равенство от $s=1$ до $s=t-1$, при $t\to\infty$ получим

$$\frac{1}{Q(t)}=1+\frac{bt}{2}+\sum_{s=1}^{t-1}\delta(s)=$$

$$=\frac{bt}{2}\left(1+\frac{2}{bt}+\frac{2}{bt}\sum_{s=1}^{t-1}\delta(s)\right)=\frac{bt}{2}(1+o(1)).$$

Таким образом, *для критических ветвящихся процессов с конечным моментом* (4.8) *при* $t\to\infty$ *имеем*

$$Q(t)=\frac{2}{bt}(1+o(1)).$$

Мы рассматривали ветвящийся процесс, начинающийся с одной частицы. Общее число частиц $\mu_t$ в процессе, начавшемся с $n$ частиц, можно представить в виде

$$\mu_t=\mu_t^{(1)}+\mu_t^{(2)}+\ldots+\mu_t^{(n)}, \quad \mu_0=n,$$

где случайные величины $\mu_t^{(k)}$, $k=1, \ldots, n$, независимы и имеют такое же распределение, как рассмотренная выше величина $\mu_t$ с $\mu_0=1$. Величину $\mu_t^{(k)}$ можно интерпретировать как число потомков $k$-й начальной частицы в момент $t$. Нетрудно найти $\mathsf{M}\mu_t$:

$$\mathsf{M}\mu_t=\mathsf{M}\mu_t^{(1)}+\ldots+\mathsf{M}\mu_t^{(n)}=n\mathsf{M}\mu_t^{(1)}=na^t,$$

где

$$a=\mathsf{M}\mu_1^{(1)}=\varphi'(1).$$

Событие $(\mu(t)>0)$ можно представить в виде

$$(\mu_t>0)=\bigcup_{k=1}^{n}(\mu_t^{(k)}>0)=\overline{\bigcap_{k=1}^{n}(\mu_t^{(k)}=0)}.$$

Отсюда, учитывая независимость величин $\mu_t^{(k)}$, $k=1, \ldots, n$, находим

$$\mathsf{P}(\mu_t>0)=1-\prod_{k=1}^{n}\mathsf{P}(\mu_t^{(k)}=0)=1-(1-Q(t))^n,$$

где

$$Q(t)=\mathsf{P}(\mu_t^{(1)}>0).$$

Таким образом, зная поведение $\mu_t$ с $\mu_0=1$, нетрудно исследовать $\mu_t$ с $\mu_0=n$.

Мы рассмотрели некоторые свойства наиболее простого ветвящегося процесса. Систематическое изложение теории ветвящихся процессов дается в книге Б. А. Севастьянова [15].

### Задачи к главе 11

1. Найти $\mathsf{M}\xi_t\xi_{t+s}$, где $\xi_t$ — пуассоновский процесс.

Указание: воспользоваться независимостью и однородностью приращений.

2. Обозначим $\tau_k$, $k=1, 2, \ldots$, время между $(k-1)$-м и $k$-м скачками пуассоновского процесса $\xi_t$. Найти функцию распределения $\tau$ (времени до первого скачка).

Указание: $(\tau_1>t)=(\xi_t=0)$.

3. Доказать, что

$$\mathsf{P}(\tau_1>t+s\mid\tau_1>s)=\mathsf{P}(\tau_1>t),$$

где $\tau_1$ определена в задаче 2.

4*. Найти совместное распределение величин $\tau_1$, $\tau_2$, определенных в задаче 2. Доказать независимость $\tau_1$, $\tau_2$.

Указание: $(\tau_1 > t_1, \quad \tau_1+\tau_2 > t_2)=(\xi_{t_1}=0, \quad \xi_{t_2}=0)+$ $+(\xi_{t_1}=0, \ \xi_{t_2}=1)$.

5*. Пусть $\nu, \eta_1, \eta_2, \ldots$ — независимые случайные величины; $\mathsf{P}(\nu=k)=\frac{(\lambda T)^k}{k!}e^{-\lambda T}$; $\eta_l$ равномерно распределена на отрезке $[0, T]$. Обозначим $\xi_t$ число величин $\eta_l$, удовлетворяющих неравенству $\eta_l < t$, $l=1, 2, \ldots, \nu$, если $\nu > 0$ и $\xi_t=0$ при $\nu=0$. Найти вероятность

$$\mathsf{P}(\xi_{t_1}=k_1, \ \xi_{t_2}-\xi_{t_1}=k_2-k_1, \ \xi_T-\xi_{t_2}=k_3-k_2).$$

Сравнить с (11.2.2).

6. Найти $\mathsf{M}\xi_t\xi_{t+s}$, где $\xi_t$ — винеровский процесс, $a=0$. (См. указание к задаче 1).

7. *Пусть $\mu_t$ — ветвящийся процесс с* $\varphi(x)=\mathsf{M}x^{\mu_1}=px^2+1-p$. Найти

$$\lambda=\lim_{t\to\infty}\mathsf{P}(\mu_t=0).$$

8. В задаче 7 при $p=1/2$ найти $B_t=\mathsf{M}\mu_t(\mu_t-1)$ и $\mathsf{D}\mu_t$.

Указание: используя (11.4.4), получить уравнение для $B_t$.

9. Случайный процесс $\xi_t$ с непрерывным временем называется цепью Маркова, если множество значений $\xi_t$ конечно и выполняется условие вида (9.1.9). Пусть в цепи Маркова с двумя состояниями (1 и 2) при $h\to 0$

$$\mathsf{P}(\xi_h=2\mid\xi_0=1)=P_{12}(h)=\alpha h+o(h),$$

$$\mathsf{P}(\xi_h=1\mid\xi_0=2)=P_{21}(h)=\beta h+o(h).$$

Найти $P_{ij}(t)=\mathsf{P}(\xi_t=j\mid\xi_0=i)$, $i, j=1, 2$.

Указание: составить для $P_{ij}(t)$ систему дифференциальных уравнений.

10*. Пусть $\tau_k(t)$, $k=1, 2$, — суммарная длительность пребывания в состоянии $k$ за время $t$ в цепи Маркова, определенной в задаче 9. Составить дифференциальное уравнение для

$$f_k(t, s)=\mathsf{M}(e^{is(\tau_1(t)-\tau_2(t))}\mid\xi_0=k).$$

Указание: воспользоваться формулой полного математического ожидания.

11*. Движение точки по прямой управляется цепью Маркова, определенной в задаче 9. Если в цепи Маркова состояние 1, то точка движется вправо со скоростью $v_1$, а если состояние 2 — то влево со скоростью $v_2$. Пусть $\eta(t)$ — координата точки в момент $t$. Найти

$$M_k(t, x)=\mathsf{M}(\eta(t)\mid\eta(0)=x, \quad \xi(0)=k),$$
$$k=1, 2.$$

Указание: составить дифференциальное уравнение для $M_k(t, x)$; использовать равенство $M_k(t, x)=x+M_k(t, 0)$.

12*. Решить задачу 11, используя следующее равенство: $\eta(t) = v_1\tau_1(t) - v_2\tau_2(t) + x$, где $\tau_1(t)$, $\tau_2(t)$ определены в задаче 10.

13*. Движение точки по прямой управляется цепью Маркова, определенной в задаче 9. Если в цепи Маркова состояние 1, то точка движется вправо со скоростью $v$, а в состоянии 2 точка движется влево с постоянным ускорением $a$. Пусть $\eta(t)$ — координата точки в момент $t$. Обладает ли процесс $\eta(t)$ свойством вида (9.1.9)? Составить интегральное уравнение для

$$M_k(t, x) = \mathsf{M}(\eta(t) \mid \eta(0) = x, \ \xi(0) = k), \quad k = 1, 2.$$

Найти $M_k(t, x)$.

Указание: $M_k(t, x) = x + M_k(t, 0)$; использовать преобразование Лапласа.

## 1. Доказательство теоремы о предельных вероятностях в цепи Маркова

Приведем доказательство теоремы 3.1 гл. 9 без использования теории матриц. Докажем сначала две леммы.

*Положим*

$$v_l(t) = \min_i P_{il}(t), \quad V_l(t) = \max_j P_{jl}(t).$$

Лемма 1. *При любых $l$ существуют пределы*

$$\lim_{t\to\infty} v_l(t) = q_l, \quad \lim_{t\to\infty} V_l(t) = Q_l$$

и

$$q_l \leqslant Q_l.$$

Доказательство. Используя уравнение (9.2.4) с $s=1$, получим

$$v_j(t+1) = \min_i \sum_{k=1}^{N} p_{ik} P_{kj}(t) \geqslant \min_i \sum_{k=1}^{N} p_{ik} v_j(t) = v_j(t),$$

$$V_j(t+1) = \max_i \sum_{k=1}^{N} p_{ik} P_{kj}(t) \leqslant \max_i \sum_{k=1}^{N} p_{ik} V_j(t) = V_j(t).$$

Таким образом, последовательности $v_j(t)$ и $V_j(t)$ монотонны и ограничены. Отсюда следует утверждение леммы 1.

**Лемма 2.** *Если выполнены условия теоремы* 9.3.1, *то существуют постоянные* $C$, $\delta$ ($C>0$, $0<\delta<1$) *такие, что*

$$V_j(nt_0) - v_j(nt_0) \leqslant C\delta^n, \quad n = 1, 2, \ldots$$

Доказательство. *Для любых $i$, $j$*

$$0 = \sum_{k=1}^{N} P_{ik}(t_0) - \sum_{k=1}^{N} P_{jk}(t_0) = \sum_k{}^{+} \Delta_k(i,j) + \sum_k{}^{-} \Delta_k(i,j),$$

где $\Delta_k(i,j) = P_{ik}(t_0) - P_{jk}(t_0)$, $\sum^{+}$ означает суммирование по всем $k$, при которых $\Delta_k(i,j)$ *положительны, а* $\sum^{-}$ — суммирование по остальным $k$. Отсюда

$$\sum_k{}^{+} \Delta_k(i,j) = -\sum_k{}^{-} \Delta_k(i,j). \tag{1}$$

Так как в условиях теоремы 9.3.1 вероятности перехода $P_{ij}(t_0) > 0$ при всех $i$, $j$, то при любых $i$, $j$

$$\sum_k^+ \Delta_k(i, j) = \sum_k^+ (P_{ik}(t_0) - P_{jk}(t_0)) < \sum_{k=1}^N P_{ik}(t_0) = 1 \tag{2}$$

и в силу конечности числа состояний

$$\delta = \max_{i, j} \sum_k^+ (P_{ik}(t_0) - P_{jk}(t_0)) < 1. \tag{3}$$

Оценим теперь разность $V_j(nt_0) - v_j(nt_0)$. Используя уравнение (9.2.4) с $s = t_0$, $t = nt_0$, получим

$$V_l((n+1)\,t_0) - v_l((n+1)\,t_0) = \max_{i, j} [P_{il}((n+1)\,t_0) - P_{jl}((n+1)\,t_0)] =$$

$$= \max_{i, j} \sum_{k=1}^N [P_{ik}(t_0) - P_{jk}(t_0)]\, P_{kl}(nt_0) =$$

$$= \max_{i, j} \sum_{k=1}^N \Delta_k(i, j)\, P_{kl}(nt_0) \leqslant$$

$$\leqslant \max_{i, j} \left\{ \sum_k^+ \Delta_k(i, j)\, V_l(nt_0) + \sum_k^- \Delta_k(i, j)\, v_l(nt_0) \right\}.$$

Отсюда, используя (1), (2) и (3), найдем

$$V_l((n+1)\,t_0) - v_l((n+1)\,t_0) \leqslant$$

$$\leqslant \max_{i, j} \left\{ (V_l(nt_0) - v_l(nt_0)) \sum_k^+ \Delta_k(i, j) \right\} =$$

$$= \delta\,(V_l(nt_0) - v_l(nt_0)).$$

Отсюда и из неравенства (3) следует утверждение леммы.

Перейдем к доказательству теоремы. Так как последовательности $v_l(t)$, $V_l(t)$ монотонны, то

$$0 < v_l(t) \leqslant q_l \leqslant Q_l \leqslant V_l(t). \tag{4}$$

Отсюда и из леммы 2 находим

$$0 \leqslant Q_l - q_l \leqslant U_l(nt_0) - v_l(nt_0) \leqslant C\delta^n.$$

Следовательно, при $n \to \infty$ получим $Q_l = q_l = q_l^*$ и

$$|P_{ij}(t) - q_j^*| \leqslant V_j(t) - v_j(t) \leqslant$$

$$\leqslant V_j\left(\left[\frac{t}{t_0}\right] t_0\right) - v_j\left(\left[\frac{t}{t_0}\right] t_0\right) \leqslant C\delta^{\frac{t}{t_0} - 1}. \tag{5}$$

Положительность $q_l^*$ следует из неравенств (4). Переходя к пределу при $t \to \infty$ и $s = 1$ в уравнении (9.2.4), получим, что $q_l^*$ удовлетворяют уравнению (9.3.2). Таким образом, теорема доказана.

## § 2. Двумерное нормальное распределение

В § 4 гл. 4 было дано определение $n$-мерного нормального распределения (см. (4.4.14)). Рассмотрим более детально случай $n=2$. В формуле (4.4.14) положим $n=2$,

$$a_{11}=\frac{1}{(1-\rho^2)\sigma_1^2}, \qquad a_{22}=\frac{1}{(1-\rho^2)\sigma_2^2},$$

$$a_{12}=a_{21}=-\frac{\rho}{(1-\rho^2)\sigma_1\sigma_2}, \qquad |\rho|<1.$$

Тогда

$$|A|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=\frac{1}{(1-\rho^2)\sigma_1^2\sigma_2^2}.$$

Формула (4.4.14) в случае двумерного нормального распределения запишется в виде

$$p_{\xi_1\xi_2}(x_1,\ x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2}Q(x_1,\ x_2)}, \tag{1}$$

где

$$Q(x_1,x_2)=\frac{1}{1-\rho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2}+\frac{(x_2-a_2)^2}{\sigma_2^2}-2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2}\right]. \tag{2}$$

Найдем одномерные плотности распределения. По формуле (4.3.9)

$$p_{\xi_1}(x_1)=\int_{-\infty}^{\infty}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2}Q(x_1,\ x_2)}dx_2.$$

Отсюда, воспользовавшись заменой $v=\frac{x_2-a_2}{\sigma_2}$ и формулой

$$Q(x_1,\ x_2)=\frac{1}{1-\rho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2}+v^2-2\rho v\frac{x_1-a_1}{\sigma_1}\right]=$$

$$=\frac{1}{1-\rho^2}\left(v-\rho\frac{x_1-a_1}{\sigma_1}\right)^2+\left(\frac{x_1-a_1}{\sigma_1}\right)^2,$$

получим

$$p_{\xi_1}(x_1)=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2}}\int_{-\infty}^{\infty}e^{-\frac{1}{2(1-\rho^2)}\left(v-\rho\frac{x_1-a_1}{\sigma_1}\right)^2}dv.$$

Полагая $z=\left(v-\rho\frac{x_1-a_1}{\sigma_1}\right)\Big/\sqrt{1-\rho^2}$, найдем

$$p_{\xi_1}(x_1)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2}}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz$$

и, следовательно,

$$p_{\xi_1}(x_1)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2}}.$$

Из этой формулы получим заменой $a_1$, $\sigma_1$ на $a_2$, $\sigma_2$ формулу плотности распределения $\xi_2$. Таким образом, одномерные распределения величин $\xi_1$, $\xi_2$ являются нормальными с параметрами $(a_1, \sigma_1)$, $(a_2, \sigma_2)$ соответственно. Следовательно,

$$\mathsf{M}\xi_1=a_1,\quad \mathsf{D}\xi_1=\sigma_1^2,\quad \mathsf{M}\xi_2=a_2,\quad \mathsf{D}\xi_2=\sigma_2^2.$$

Полагая в формуле (5.1.7)

$$n=2,\quad g(x_1, x_2)=(x_1-a_1)(x_2-a_2),$$

получим

$$\operatorname{cov}(\xi_1, \xi_2)=\mathsf{M}(\xi_1-a_1)(\xi_2-a_2)=$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-a_1)(x_2-a_2)e^{-\frac{1}{2}Q(x_1, x_2)}dx_1\,dx_2.$$

Отсюда заменой переменных

$$y_1=\frac{x_1-a_1}{\sigma_1\sqrt{1-\rho^2}},\quad y_2=\frac{x_2-a_2}{\sigma_2\sqrt{1-\rho^2}}$$

получим

$$\operatorname{cov}(\xi_1, \xi_2)=\frac{(1-\rho^2)\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y_1y_2e^{-\frac{1}{2}(y_1^2-2\rho y_1y_2+y_2^2)}dy_1\,dy_2=$$

$$=\frac{(1-\rho^2)\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{y_2e^{-\frac{y_2^2}{2}(1-\rho^2)}\left(\int_{-\infty}^{\infty}\frac{y_1}{\sqrt{2\pi}}e^{-\frac{(y_1-\rho y_2)^2}{2}}dy_1\right)\right\}dy_2.$$

Интеграл в круглых скобках можно рассматривать как математическое ожидание нормально распределенной случайной величины с параметрами $(\rho y_2, 1)$. Следовательно, этот интеграл равен $\rho y_2$ и

$$\operatorname{cov}(\xi_1, \xi_2)=(1-\rho^2)\sigma_1\sigma_2\int_{-\infty}^{\infty}\frac{y_2^2}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}(1-\rho^2)}dy_2.$$

Последний интеграл является дисперсией нормально распределенной случайной величины с параметрами $(0,\ 1/\sqrt{1-\rho^2})$. Следовательно, этот интеграл равен $1/(1-\rho^2)$ и $\operatorname{cov}(\xi_1,\ \xi_2)=\rho\sigma_1\sigma_2$. Отсюда для коэффициента корреляции получим

$$\rho_{\xi_1,\ \xi_2}=\frac{\operatorname{cov}(\xi_1,\ \xi_2)}{\sqrt{D\xi_1 D\xi_2}}=\rho.$$

*Таким образом, все параметры, входящие в (1), (2), являются моментами случайных величин* $\xi_1,\ \xi_2$.

В § 4 гл. 5 было показано, что для независимых случайных величин $\rho_{\xi_1,\ \xi_2}=0$. Для случайных величин, совместное распределение которых нормально, верно и обратное утверждение. Действительно, если $\rho=0$, то плотность распределения (1) представляется в виде произведения одномерных плотностей.

Приведем без доказательства следующее утверждение: *если* $(\xi_1,\ \xi_2)$ *распределены нормально, то совместное распределение случайных величин*

$$\eta_1=c_{11}\xi_1+c_{12}\xi_2,$$
$$\eta_2=c_{21}\xi_1+c_{22}\xi_2,$$

*где* $C=\|c_{ij}\|$ — *невырожденная матрица, также нормально.*

Используя свойства математических ожиданий, нетрудно найти параметры нормального распределения величин $(\eta_1,\ \eta_2)$. Действительно,

$$M\eta_1=c_{11}a_1+c_{12}a_2,\quad M\eta_2=c_{21}a_1+c_{22}a_2,$$
$$D\eta_1=c_{11}^2\sigma_1^2+c_{12}^2\sigma_2^2+2c_{11}c_{12}\rho\sigma_1\sigma_2,$$
$$D\eta_2=c_{21}^2\sigma_1^2+c_{22}^2\sigma_2^2+2c_{21}c_{22}\rho\sigma_1\sigma_2,$$
$$\operatorname{cov}(\eta_1,\ \eta_2)=c_{11}c_{21}\sigma_1^2+c_{22}c_{12}\sigma_2^2+(c_{11}c_{22}+c_{21}c_{12})\,\rho\sigma_1\sigma_2.$$

Если случайные величины $\xi_1,\ \xi_2$ независимы, $\sigma_1=\sigma_2$ и матрица $C$ ортогональна, то величины $\eta_1,\ \eta_2$ независимы, так как $\rho=0$,

$$c_{11}^2+c_{12}^2=c_{21}^2+c_{22}^2=1,\qquad c_{11}c_{21}+c_{22}c_{12}=0$$

и, следовательно, $\operatorname{cov}(\eta_1,\ \eta_2)=0$.

# СЛУЧАЙНЫЕ ЧИСЛА

Приведенные в таблице цифры можно рассматривать как реализации независимых и одинаково распределенных случайных величин, принимающих значения 0, 1, ..., 9 с одной и той же вероятностью, равной 0,1.

Таблица 1

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 09 | 73 | 25 | 33 | 76 | 52 | 01 | 35 | 86 | 34 | 67 | 35 | 48 | 76 |
| 37 | 54 | 20 | 48 | 05 | 64 | 89 | 47 | 42 | 96 | 24 | 80 | 52 | 40 | 37 |
| 08 | 42 | 26 | 89 | 53 | 19 | 64 | 50 | 93 | 03 | 23 | 20 | 90 | 25 | 60 |
| 99 | 01 | 90 | 25 | 29 | 09 | 37 | 67 | 07 | 15 | 38 | 31 | 13 | 11 | 65 |
| 12 | 80 | 79 | 99 | 70 | 80 | 15 | 73 | 61 | 47 | 64 | 03 | 23 | 66 | 53 |
| 80 | 95 | 90 | 91 | 17 | 39 | 29 | 27 | 49 | 45 | 66 | 06 | 57 | 47 | 17 |
| 20 | 63 | 61 | 04 | 02 | 00 | 82 | 29 | 16 | 65 | 31 | 06 | 01 | 08 | 05 |
| 15 | 95 | 33 | 47 | 64 | 35 | 08 | 03 | 36 | 06 | 85 | 26 | 97 | 76 | 02 |
| 88 | 67 | 67 | 43 | 97 | 04 | 43 | 62 | 76 | 59 | 63 | 57 | 33 | 21 | 35 |
| 98 | 95 | 11 | 68 | 77 | 12 | 17 | 17 | 68 | 33 | 73 | 79 | 64 | 57 | 53 |
| 34 | 07 | 27 | 68 | 50 | 36 | 69 | 73 | 61 | 70 | 65 | 81 | 33 | 98 | 85 |
| 45 | 57 | 18 | 24 | 06 | 35 | 30 | 34 | 26 | 14 | 86 | 79 | 90 | 74 | 39 |
| 02 | 05 | 16 | 56 | 92 | 68 | 66 | 57 | 48 | 18 | 73 | 05 | 38 | 52 | 47 |
| 05 | 32 | 54 | 70 | 48 | 90 | 55 | 35 | 75 | 48 | 28 | 46 | 82 | 87 | 09 |
| 03 | 52 | 96 | 47 | 78 | 35 | 80 | 83 | 42 | 82 | 60 | 93 | 52 | 03 | 44 |
| 11 | 19 | 92 | 91 | 70 | 98 | 52 | 01 | 77 | 67 | 14 | 90 | 56 | 86 | 07 |
| 23 | 40 | 30 | 97 | 32 | 11 | 80 | 50 | 54 | 31 | 39 | 80 | 82 | 77 | 32 |
| 18 | 62 | 38 | 85 | 79 | 83 | 45 | 29 | 96 | 34 | 06 | 28 | 89 | 80 | 83 |
| 83 | 49 | 12 | 56 | 24 | 88 | 68 | 54 | 02 | 00 | 86 | 50 | 75 | 84 | 01 |
| 35 | 27 | 38 | 84 | 35 | 99 | 59 | 46 | 73 | 48 | 87 | 51 | 76 | 49 | 69 |
| 22 | 10 | 94 | 05 | 58 | 60 | 97 | 09 | 34 | 33 | 50 | 50 | 07 | 39 | 98 |
| 50 | 72 | 56 | 82 | 48 | 29 | 40 | 52 | 42 | 01 | 52 | 77 | 56 | 78 | 51 |
| 13 | 74 | 67 | 00 | 78 | 18 | 47 | 54 | 06 | 10 | 68 | 71 | 17 | 78 | 17 |
| 36 | 76 | 66 | 79 | 51 | 90 | 36 | 47 | 64 | 93 | 29 | 60 | 91 | 10 | 62 |
| 91 | 82 | 60 | 89 | 28 | 93 | 78 | 56 | 13 | 68 | 23 | 47 | 83 | 41 | 13 |
| 65 | 48 | 11 | 76 | 74 | 17 | 46 | 85 | 09 | 50 | 58 | 04 | 77 | 69 | 74 |
| 80 | 12 | 43 | 56 | 35 | 17 | 72 | 70 | 80 | 15 | 45 | 31 | 82 | 23 | 74 |
| 74 | 35 | 09 | 98 | 17 | 77 | 40 | 27 | 72 | 14 | 43 | 23 | 60 | 02 | 10 |
| 69 | 91 | 62 | 68 | 03 | 66 | 25 | 22 | 91 | 48 | 36 | 93 | 68 | 72 | 03 |
| 09 | 89 | 32 | 05 | 05 | 14 | 22 | 56 | 85 | 14 | 46 | 42 | 75 | 67 | 88 |
| 73 | 03 | 95 | 71 | 86 | 40 | 21 | 81 | 65 | 44 | 91 | 49 | 91 | 45 | 23 |
| 21 | 11 | 57 | 82 | 53 | 14 | 38 | 55 | 37 | 63 | 80 | 33 | 69 | 45 | 98 |
| 45 | 52 | 16 | 42 | 37 | 96 | 28 | 60 | 26 | 55 | 44 | 10 | 48 | 19 | 49 |
| 76 | 62 | 11 | 39 | 90 | 94 | 40 | 05 | 64 | 18 | 12 | 55 | 07 | 37 | 42 |
| 96 | 29 | 77 | 88 | 22 | 54 | 38 | 21 | 45 | 98 | 63 | 60 | 64 | 93 | 29 |
| 68 | 47 | 92 | 76 | 86 | 46 | 16 | 28 | 35 | 54 | 94 | 75 | 08 | 99 | 23 |
| 26 | 94 | 03 | 68 | 58 | 70 | 29 | 73 | 41 | 35 | 53 | 14 | 03 | 33 | 40 |
| 85 | 15 | 74 | 79 | 54 | 32 | 97 | 92 | 65 | 75 | 57 | 60 | 04 | 08 | 81 |
| 11 | 10 | 00 | 20 | 40 | 12 | 86 | 07 | 46 | 97 | 96 | 64 | 48 | 94 | 39 |
| 16 | 50 | 53 | 44 | 84 | 40 | 21 | 95 | 25 | 63 | 43 | 65 | 17 | 70 | 82 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 37 | 08 | 92 | 00 | 48 | 61 | 19 | 69 | 04 | 46 | 26 | 45 | 74 | 77 | 74 |
| 42 | 05 | 08 | 23 | 41 | 15 | 47 | 44 | 52 | 66 | 95 | 27 | 07 | 99 | 53 |
| 22 | 22 | 20 | 64 | 13 | 94 | 55 | 72 | 85 | 73 | 67 | 89 | 75 | 43 | 87 |
| 28 | 70 | 72 | 58 | 15 | 42 | 48 | 11 | 62 | 13 | 97 | 34 | 40 | 87 | 21 |
| 07 | 20 | 73 | 17 | 90 | 23 | 52 | 37 | 83 | 17 | 73 | 20 | 88 | 98 | 37 |
| 51 | 92 | 43 | 37 | 29 | 65 | 39 | 45 | 95 | 93 | 42 | 58 | 26 | 05 | 27 |
| 59 | 36 | 78 | 38 | 48 | 82 | 39 | 61 | 01 | 18 | 33 | 21 | 15 | 94 | 66 |
| 54 | 62 | 24 | 44 | 31 | 91 | 19 | 04 | 25 | 92 | 92 | 92 | 74 | 59 | 73 |
| 16 | 86 | 84 | 87 | 67 | 03 | 07 | 11 | 20 | 59 | 25 | 70 | 14 | 66 | 70 |
| 68 | 93 | 59 | 14 | 16 | 26 | 25 | 22 | 96 | 63 | 05 | 52 | 28 | 25 | 62 |
| 04 | 49 | 35 | 24 | 94 | 75 | 24 | 63 | 38 | 24 | 45 | 86 | 25 | 10 | 25 |
| 00 | 54 | 99 | 76 | 54 | 64 | 05 | 18 | 81 | 59 | 96 | 11 | 96 | 38 | 96 |
| 35 | 96 | 31 | 53 | 07 | 26 | 89 | 80 | 93 | 54 | 33 | 35 | 13 | 54 | 62 |
| 59 | 80 | 80 | 83 | 91 | 45 | 42 | 72 | 68 | 42 | 83 | 60 | 94 | 97 | 00 |
| 46 | 05 | 88 | 52 | 36 | 01 | 89 | 09 | 22 | 86 | 77 | 28 | 14 | 40 | 77 |
| 61 | 96 | 27 | 93 | 35 | 65 | 33 | 71 | 24 | 72 | 32 | 17 | 90 | 05 | 97 |
| 54 | 69 | 28 | 23 | 91 | 23 | 28 | 72 | 95 | 29 | 69 | 23 | 46 | 14 | 06 |
| 77 | 97 | 45 | 00 | 24 | 90 | 10 | 33 | 93 | 33 | 19 | 56 | 54 | 14 | 30 |
| 13 | 02 | 12 | 48 | 92 | 78 | 56 | 52 | 01 | 06 | 45 | 15 | 51 | 49 | 38 |
| 93 | 91 | 08 | 36 | 47 | 70 | 61 | 74 | 29 | 41 | 94 | 86 | 43 | 19 | 94 |
| 87 | 37 | 92 | 52 | 41 | 05 | 56 | 70 | 70 | 07 | 86 | 74 | 31 | 71 | 57 |
| 20 | 11 | 74 | 52 | 04 | 15 | 95 | 66 | 00 | 00 | 18 | 74 | 39 | 24 | 23 |
| 01 | 75 | 87 | 53 | 79 | 40 | 41 | 92 | 15 | 85 | 66 | 67 | 43 | 68 | 06 |
| 19 | 47 | 60 | 72 | 46 | 43 | 66 | 79 | 45 | 43 | 59 | 04 | 79 | 00 | 33 |
| 36 | 16 | 81 | 08 | 51 | 34 | 88 | 88 | 15 | 53 | 01 | 54 | 03 | 54 | 56 |
| 85 | 39 | 41 | 18 | 38 | 98 | 08 | 62 | 48 | 26 | 45 | 24 | 02 | 84 | 04 |
| 97 | 11 | 89 | 63 | 38 | 33 | 18 | 51 | 62 | 32 | 41 | 94 | 15 | 09 | 49 |
| 84 | 96 | 28 | 52 | 07 | 80 | 95 | 10 | 04 | 06 | 96 | 38 | 27 | 07 | 74 |
| 20 | 82 | 66 | 95 | 41 | 79 | 75 | 24 | 91 | 40 | 71 | 96 | 12 | 82 | 96 |
| 05 | 01 | 45 | 11 | 76 | 18 | 63 | 33 | 25 | 37 | 98 | 14 | 50 | 65 | 71 |
| 44 | 99 | 90 | 88 | 96 | 39 | 09 | 47 | 34 | 07 | 35 | 44 | 13 | 18 | 80 |
| 89 | 43 | 54 | 85 | 81 | 88 | 69 | 54 | 19 | 94 | 37 | 54 | 87 | 30 | 43 |
| 20 | 15 | 12 | 33 | 87 | 25 | 01 | 62 | 52 | 98 | 94 | 62 | 46 | 11 | 71 |
| 69 | 86 | 10 | 25 | 91 | 74 | 85 | 22 | 05 | 39 | 00 | 38 | 75 | 95 | 79 |
| 31 | 01 | 02 | 46 | 74 | 05 | 45 | 56 | 14 | 27 | 77 | 93 | 89 | 19 | 36 |
| 74 | 02 | 94 | 39 | 02 | 77 | 55 | 73 | 22 | 70 | 97 | 79 | 01 | 71 | 19 |
| 54 | 17 | 84 | 56 | 11 | 80 | 99 | 33 | 71 | 43 | 05 | 33 | 51 | 29 | 69 |
| 11 | 66 | 44 | 98 | 83 | 52 | 07 | 98 | 48 | 27 | 59 | 38 | 17 | 15 | 39 |
| 48 | 32 | 47 | 79 | 28 | 31 | 24 | 96 | 47 | 10 | 02 | 29 | 53 | 68 | 70 |
| 69 | 07 | 49 | 41 | 38 | 87 | 63 | 79 | 19 | 76 | 35 | 58 | 40 | 44 | 01 |
| 52 | 52 | 75 | 80 | 21 | 80 | 81 | 45 | 17 | 48 | 09 | 18 | 82 | 00 | 97 |
| 56 | 12 | 71 | 92 | 55 | 36 | 04 | 09 | 03 | 24 | 90 | 04 | 58 | 54 | 97 |
| 09 | 97 | 33 | 34 | 40 | 88 | 46 | 12 | 33 | 56 | 73 | 18 | 95 | 02 | 07 |
| 32 | 30 | 75 | 75 | 46 | 15 | 02 | 00 | 99 | 94 | 75 | 76 | 87 | 64 | 90 |
| 10 | 51 | 82 | 16 | 15 | 01 | 84 | 87 | 69 | 38 | 54 | 01 | 64 | 40 | 56 |



| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 32 | 82 | 53 | 95 | 27 | 04 | 22 | 08 | 63 | 04 | 88 | 38 | 98 | 73 | 74 |
| 51 | 98 | 15 | 06 | 54 | 94 | 93 | 88 | 19 | 97 | 91 | 87 | 07 | 61 | 50 |
| 47 | 67 | 72 | 62 | 69 | 62 | 29 | 06 | 44 | 64 | 27 | 12 | 46 | 70 | 18 |
| 20 | 97 | 18 | 17 | 49 | 90 | 42 | 91 | 22 | 72 | 95 | 37 | 50 | 58 | 71 |
| 66 | 28 | 13 | 10 | 03 | 00 | 68 | 22 | 73 | 98 | 20 | 71 | 45 | 32 | 95 |
| 64 | 27 | 85 | 80 | 44 | 08 | 35 | 86 | 99 | 10 | 78 | 54 | 24 | 27 | 85 |
| 68 | 47 | 66 | 46 | 59 | 28 | 30 | 60 | 22 | 64 | 81 | 33 | 31 | 05 | 91 |
| 41 | 36 | 18 | 27 | 60 | 53 | 84 | 08 | 62 | 33 | 81 | 59 | 41 | 36 | 28 |
| 93 | 82 | 34 | 31 | 78 | 91 | 75 | 75 | 37 | 41 | 61 | 61 | 36 | 22 | 69 |
| 07 | 70 | 61 | 78 | 13 | 89 | 41 | 59 | 26 | 94 | 00 | 39 | 75 | 83 | 91 |
| 13 | 66 | 15 | 88 | 73 | 04 | 61 | 89 | 75 | 53 | 31 | 22 | 30 | 84 | 20 |
| 40 | 51 | 00 | 78 | 93 | 92 | 60 | 46 | 04 | 75 | 94 | 11 | 90 | 18 | 40 |
| 51 | 21 | 59 | 02 | 90 | 28 | 46 | 66 | 87 | 95 | 77 | 76 | 22 | 07 | 91 |
| 50 | 26 | 39 | 02 | 12 | 55 | 78 | 17 | 65 | 14 | 83 | 48 | 34 | 70 | 55 |
| 12 | 60 | 71 | 76 | 46 | 48 | 94 | 97 | 23 | 06 | 94 | 54 | 13 | 74 | 08 |

## НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Значения функции $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$

Таблица 2

| $x$ | 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 $u_\alpha$**

**Функция $u_\alpha$ определяется равенством** $\alpha = \frac{1}{\sqrt{2\pi}} \int_{u_\alpha}^{\infty} e^{-\frac{x^2}{2}} dx$

Таблица 3

| $\alpha$ | 0,001 | 0,005 | 0,010 | 0,015 | 0,020 | 0,025 | 0,030 | 0,035 | 0,040 | 0,045 | 0,050 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $u_\alpha$ | 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 4

| $k$ \ $\lambda$ | 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 |

| $k$ \ $\lambda$ | 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 |

| $k$ \ $\lambda$ | 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 |

# РАСПРЕДЕЛЕНИЕ СТЬЮДЕНТА

### Значения функции $t_{\alpha, n}$

Функция $t_{\alpha, n}$ определяется равенством $\mathsf{P}(|\tau_n| < t_{\alpha, n}) = 1 - \alpha$, где случайная величина $\tau_n$ имеет распределение Стьюдента с $n$ степенями свободы. Плотность распределения $\tau_n$ равна

$$p_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \cdot \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

Таблица 5

| $n$ \ $\alpha$ | 0,10 | 0,05 | 0,02 | 0,01 |
|---|---|---|---|---|
| 5 | 2,015 | 2,571 | 3,365 | 4,032 |
| 6 | 1,943 | 2,447 | 3,143 | 3,707 |
| 7 | 1,895 | 2,365 | 2,998 | 3,499 |
| 8 | 1,860 | 2,306 | 2,896 | 3,355 |
| 9 | 1,833 | 2,262 | 2,821 | 3,250 |
| 10 | 1,812 | 2,228 | 2,764 | 3,169 |
| 12 | 1,782 | 2,179 | 2,681 | 3,055 |
| 14 | 1,761 | 2,145 | 2,624 | 2,977 |
| 16 | 1,746 | 2,120 | 2,583 | 2,921 |
| 18 | 1,734 | 2,101 | 2,552 | 2,878 |
| 20 | 1,725 | 2,086 | 2,528 | 2,845 |
| 22 | 1,717 | 2,074 | 2,508 | 2,819 |
| 30 | 1,697 | 2,042 | 2,457 | 2,750 |
| $\infty$ | 1,645 | 1,960 | 2,326 | 2,576 |

# $\chi^2$-РАСПРЕДЕЛЕНИЕ

## Значения функции $\chi^2_{m,\,\alpha}$

Функция $\chi^2_{m,\,\alpha}$ определяется равенством $P\left(\chi^2_m > \chi^2_{m,\,\alpha}\right) = \alpha$, где случайная величина $\chi^2_m$ имеет $\chi^2$-распределение с $m$ степенями свободы. Плотность распределения $\chi^2_m$ равна

$$p_{\chi^2_m}(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0.$$

Таблица 6

| $m$ \ $\alpha$ | 0,10 | 0,05 | 0,02 | 0,01 | 0,005 |
|---|---|---|---|---|---|
| 1 | 2,7 | 3,8 | 5,4 | 6,6 | 7,9 |
| 2 | 4,6 | 6,0 | 7,8 | 9,2 | 11,6 |
| 3 | 6,3 | 7,8 | 9,8 | 11,3 | 12,8 |
| 4 | 7,8 | 9,5 | 11,7 | 13,3 | 14,9 |
| 5 | 9,2 | 11,1 | 13,4 | 15,1 | 16,3 |
| 6 | 10,6 | 12,6 | 15,0 | 16,8 | 18,6 |
| 7 | 12,0 | 14,1 | 16,6 | 18,5 | 20,3 |
| 8 | 13,4 | 15,5 | 18,2 | 20,1 | 21,9 |
| 9 | 14,7 | 16,9 | 19,7 | 21,7 | 23,6 |
| 10 | 16,0 | 18,3 | 21,2 | 23,2 | 25,2 |
| 11 | 17,3 | 19,7 | 22,6 | 24,7 | 26,8 |
| 12 | 18,5 | 21,0 | 24,1 | 26,2 | 28,3 |
| 13 | 19,8 | 22,4 | 25,5 | 27,7 | 29,8 |
| 14 | 21,1 | 23,7 | 26,9 | 29,1 | 31 |
| 15 | 22,3 | 25,0 | 28,3 | 30,6 | 32,5 |
| 16 | 23,5 | 26,3 | 29,6 | 32,0 | 34 |
| 17 | 24,8 | 27,6 | 31,0 | 33,4 | 35,5 |
| 18 | 26,0 | 28,9 | 32,3 | 34,8 | 37 |
| 19 | 27,2 | 30,1 | 33,7 | 36,2 | 38,5 |
| 20 | 28,4 | 31,4 | 35,0 | 37,6 | 40 |
| 21 | 29,6 | 32,7 | 36,3 | 38,9 | 41,5 |
| 22 | 30,8 | 33,9 | 37,7 | 40,3 | 42,5 |
| 23 | 32,0 | 35,2 | 39,0 | 41,6 | 44,0 |
| 24 | 33,2 | 36,4 | 40,3 | 43,0 | 45,5 |
| 25 | 34,4 | 37,7 | 41,6 | 44,3 | 47 |

# ОТВЕТЫ К ЗАДАЧАМ

## Глава 1

2. 1) $A$; 2) $AB$; 3) $B+AC$. 3. 1), 2), 4) — неверны; 3) — верно. 4. 1) $A\bar{B}\bar{C}$; 2) $AB\bar{C}$; 3) $ABC$; 4) $A+B+C$, 5) $A\bar{B}\bar{C}+\bar{A}B\bar{C}+\bar{A}\bar{B}C$; 6) $\bar{A}\bar{B}\bar{C}$; 7) $(A+B+C)-ABC$. 5. 5/36, 2/3, 13/36 7. 15/16, 2/3, 1. 8. $1/n$. 9. 0,72; 0,27; 0,01. 10. $1-\left(\frac{5}{6}\right)^4=0{,}5177\ldots$, $1-\left(\frac{35}{36}\right)^{24}=0{,}4914\ldots$ 11. $\frac{C_n^2 n!}{n^n}$. 12. $\frac{12!}{12^{12}}$. 13. $\frac{2^{2r}C_n^{2r}}{C_{2n}^{2r}}$, $\frac{n2^{2r-2}C_{n-1}^{2r-2}}{C_{2n}^{2r}}$. 14. $C_{90}^{10}/C_{100}^{10}$. 15. $\frac{10C_{43}^3}{C_{49}^6}$. 16. $\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\longrightarrow C_n^m\alpha^m(1-\alpha)^{n-m}$. 17. $\frac{(a-2r)^2}{a^2}$, $1-\frac{4r^2}{a^2}$. 18. $1-\frac{(T-l)^2}{T^2}$. 19. $F'(x)=1$, $x\in[0,1]$. 20. $F'(x)=x\in[0,1]$; $F'(x)=2-x$, $x\in[1,2]$. 21. $F'(x)=\frac{x^2}{2}$, $x\in[0,1]$; $F'(x)=3x-x^2-\frac{3}{2}$, $x\in[1,2]$, $F'(x)=\frac{x^2}{2}-3x+\frac{9}{2}$, $x\in[2,3]$. 22. 1/12. 23. 1/4. 24. $\mathsf{P}(\eta_1<x)=x(1-\ln x)$, $x\in(0,1)$; $\mathsf{P}(\eta_2<x)=1-(1-x)^2$, $x\in(0,1)$, $\mathsf{P}(\eta_3<x)=x^2$, $x\in(0,1)$.

## Глава 2

1. 0,5. 2. $1-\frac{10\cdot 5^9}{6^{10}-5^{10}}$. 4. 9/35. 5. 31/36. 6. 20/21. 7. 9/16. 8. $e^{-\lambda t}$. 9. $\pi_0=\frac{\beta}{\alpha_1+\alpha_2+\beta}$, $\pi_1=\frac{\alpha_1}{\alpha_1+\beta}$, $\pi_2=\frac{\alpha_2}{\alpha_1+\beta}\pi_0$. 10. $\pi_0=1-\theta$; $\pi_k=(1-\theta)\theta^k$, $k\geqslant 1$.

## Глава 3

1. $\mathsf{P}(A_k)=\frac{b}{b+r}$, $\mathsf{P}(A_1\mid A_2)=\frac{b+c}{b+r+c}$. 2. 3/5; 2/5. 4. 8/32; 7/32. 5. $C_{10}^4(5/72)^4(1-5/72)^6$. 6. $(0{,}9)^n\leqslant 0{,}1$, $n\geqslant 22$. 7. 96.

8. 0,0588. 9. $\frac{5^2\cdot 7\cdot 9}{4}(1/\pi)^4(1-2/\pi)^6$. 10. $2^{-2n}C_{2n}^n$.
11. $C_n^m C_{n-1}^{l-1} p^{m+l} q^{2n-m-l}$. 12. 1) 0,6651...; 0,6321...; 2) 0,40187...; 0,3679...; 3) 0,2009...; 0,1839... 13. 0,14288. 14. 5. 15. 0,95957. 16. 0,99. 17. 547.

## Глава 4

2. a) $\frac{1}{\sqrt{2\pi x}}e^{-x/2}$, $x>0$. 3. $C=1$; $p_\eta(x)=1$, $x\in[0,1]$; 1/4.
4. $p_\eta(x)=\frac{\alpha}{x^2}e^{-\alpha\left(1-\frac{1}{x}\right)}$, $x\in(-\infty,0)\cup(1,+\infty)$. 5. $p_\eta(x)=1$, $x\in[0,1]$. 6. $\alpha^2xe^{-\alpha x}$, $x>0$. 7. $p(x)=3/2$, $x\in\left(0,\frac{1}{2}\right)$; $p(x)=\frac{1-x^2}{2x^2}$, $x\in\left(\frac{1}{2},1\right)$. 8. $\mathsf{P}(\xi_1=i,\ \xi_2=j)=\frac{1}{100}$, $i,j=0,1,\ldots,9$.
11. $p_{\xi_1+\xi_2}(x)=\frac{1}{2}$, $x\in[0,2]$. 13. $C_{n-m_2}^{m_1}\left(\frac{p_1}{p_1+p_2}\right)^{m_1}\times\left(\frac{p_2}{p_1+p_2}\right)^{n-m_2-m_1}$. 14. $\mathsf{P}(\tau=m)=q^{m-1}p$, $m\geqslant 1$. 15. $\mathsf{P}(\tau^{(1)}=m_1,\ \tau^{(2)}=m_2)=q^{m_1+m_2-2}p^2$. 16. $\mathsf{P}(\tau^{(m)}=k)=C_{k-1}^{m-1}p^mq^{k-m}$.
17. $\mathsf{P}(\xi_1+\xi_2=k)=\frac{k-1}{2^k}$. 18. 1) $n(1-F_{\xi_1}(x))^{n-1}p_{\xi_1}(x)$;
2) $n[F_{\xi_1}(x)]^{n-1}p_{\xi_1}(x)$; 3) $\frac{n!}{(m-1)!\,(n-m)!}[F_{\xi_1}(x)]^{m-1}[1-F_{\xi_1}(x)]^{n-m}$; 4) $\frac{n!}{(m-1)!\,(k-m)!\,(n-k)!}[F_{\xi_1}(x_1)]^{m-1}[F_{\xi_1}(x_2)-F_{\xi_1}(x_1)]^{k-m}\cdot[1-F_{\xi_1}(x_2)]^{n-k}p_{\xi_1}(x_1)\,p_{\xi_1}(x_2)$.
20. $\lambda=n_1p_1+n_2p_2+n_3p_3=2$, $1-e^{-\lambda}(1+\lambda)=0{,}59399\ldots$

## Глава 5

1. $1/p$. 2. $n+1$. 3. $\mathsf{P}(\xi=k)=\left(1-\frac{1}{n}\right)^{k-1}\frac{1}{n}$; $\mathsf{M}\xi=n$.
4. $m/p$. 5. $\mathsf{M}\tau=\frac{1}{\lambda}$, $\mathsf{D}\tau=\frac{1}{\lambda^2}$. 6. 9; 16,5. 7. $\mathsf{M}\xi=npq$, $\mathsf{D}\xi=npq(p^3+q^3)+2p^2q^2$. 8. $\mathsf{M}\eta_1=9$, $\mathsf{M}\eta_2=20{,}25$, $\mathsf{D}\eta_1=16{,}5$, $\mathsf{D}\eta_2=402{,}1875$. 9. $\mathsf{M}\xi_1=-\frac{1}{8}$, $\mathsf{M}\xi_2=0$, $\mathsf{D}\xi_1=\frac{133}{192}$, $\mathsf{D}\xi_2=1$, $\operatorname{cov}(\xi_1,\xi_2)=-\frac{1}{8}$. 10. $\mathsf{M}\xi_1=\mathsf{M}\xi_2=\operatorname{cov}(\xi_1,\xi_2)=0$, $\mathsf{D}\xi_1=\mathsf{D}\xi_2=0{,}5$; величины зависимы. 11. а) 0; б) 1/3. 12. $\mathsf{M}\xi_k=np_k$;

$\operatorname{cov}(\xi_k, \xi_l) = np_k(\delta_{kl} - p_l)$. **13.** $\mathsf{P}(\mu_0 = 0) = \sum_{l=0}^{N} (-1)^l C_N^l \left(1 - \frac{l}{N}\right)^n$.
$\Big[$У к а з а н и е: $(\mu_0 > 0) = \bigcup_{l=1}^{N} A_l$, где $A_l$ — событие, состоящее в том, что ящик с номером $l$ пустой.$\Big]$ **14.** $\mathsf{M}\mu_0 \sim Ne^{-\alpha}$, $\mathsf{D}\mu_0 \sim$ $\sim Ne^{-\alpha}\left[1 - (1+\alpha)e^{-\alpha}\right]$ при $N \longrightarrow \infty$. **15.** $\sum_{k=1}^{n} (-1)^{k+1} \frac{1}{k!} \longrightarrow$ $\longrightarrow 1 - e^{-1}$ при $n \longrightarrow \infty$. **16.** $\mathsf{M}\nu = N \sum_{k=1}^{N} \frac{1}{k}$. **17.** $\mathsf{M}\eta = a$, $\mathsf{M}\theta = b^2$.
**19.** $M_k = \frac{k}{q-p} - \frac{n}{q-p} \cdot \frac{1-(q/p)^k}{1-(q/p)^n}$, $p \neq q$; $M_k = k(n-k)$, $p = q = \frac{1}{2}$. **20.** $\alpha = \frac{\sigma_{12}}{\sigma_{11}}$, $\beta = -a_1 \frac{\sigma_{12}}{\sigma_{11}} + a_2$.

### Глава 6

**1.** 0,08. **2.** 0,25; 0,0455. **4.** Да. **5.** При $\alpha < 1$ применим; при $\alpha \geqslant 1$ неприменим.

### Глава 7

**1.** $x\varphi(x^2)$, $\frac{1 - x\varphi(x)}{1-x}$, $\frac{\varphi(x)}{1-x}$, $\frac{1}{2}\left[\varphi(\sqrt{x}) + \varphi(-\sqrt{x})\right]$.
**2.** $\frac{px}{1-qx}$; $\mathsf{M}\tau = 1/p$, $\mathsf{D}\tau = q/p^2$; $\left(\frac{px}{1-qx}\right)^m$; $\mathsf{M}\tau_m = m/p$. $\mathsf{D}\tau_m =$ $= mq/p^2$. **3.** $\prod_{k=1}^{N} \left[\left(1 - \frac{k-1}{N}\right) \cdot \frac{x}{1 - \frac{k-1}{N}x}\right]$. [У к а з а н и е: $\nu_N = \delta_1 + \delta_2 + \ldots + \delta_N$, где $\delta_1 = 1$, $\delta_l$ — число размещаемых шаров с момента заполнения $l-1$ ящиков до заполнения $l$ ящиков.]
**4.** $e^{-\lambda(x-1)}$. **5.** $e^{\lambda p(x-1)}$. **6.** а) $(pe^{it} + q)^n$; б) $e^{\lambda(e^{it}-1)}$; в) $\frac{1}{1 - it/\alpha}$;
г) $\frac{\sin t}{t}$. **7.** $\mathsf{P}(\xi = 1) = \mathsf{P}(\xi = -1) = \frac{1}{2}$, $\mathsf{P}(\xi = 2) = \mathsf{P}(\xi = -2) = \frac{1}{4}$,
$\mathsf{P}(\xi = 0) = \frac{1}{2}$, $\mathsf{P}(\xi = k) = \mathsf{P}(\xi = -k) = 2^{-k-1}$, $k = 1, 2, \ldots$ **8.** $|f(t)|^2$.
**9.** а) 0,5; б) 0,6568.

### Глава 8

1. $0{,}75 \cdot 10^{-m+2}$. 3. Нормальное распределение с параметрами $(0,\ 1)$. 4. $\frac{\lambda^m}{m!} e^{-\lambda}$.

### Глава 9

1. $1/6$. 2. а) Нет, если $p \neq q$ (например, $\mathsf{P}(\eta_3 = 1/\eta_2 = -1) = 1/2$, $\mathsf{P}(\eta_3 = 1 \mid \eta_1 = 1,\ \eta_2 = -1) = 2pq$. б) Да, $\mathsf{P}(\eta_{t+1} = 1 \mid \eta_t = 1) = p$, $\mathsf{P}(\eta_{t+1} = 1 \mid \eta_t = -1) = q$. в) Да, $\begin{pmatrix} q & p & 0 & 0 \\ 0 & 0 & q & p \\ q & p & 0 & 0 \\ 0 & 0 & q & p \end{pmatrix}$. 3. $\begin{pmatrix} 3/7 & 4/7 \\ 1/11 & 10/11 \end{pmatrix}$.

5. $$\begin{pmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{pmatrix} = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{(1-\alpha-\beta)^t}{\alpha+\beta} \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}.$$

### Глава 10

3. $\mathsf{M}a^* = a + \frac{b-a}{n+1}$, $\mathsf{M}b^* = b - \frac{b-a}{n+1}$. 4. Да; $\mathsf{M}\alpha^* = 0$, $\mathsf{D}\alpha^* = 1/n^2$. 5. $\lambda^* = (x_1 + x_2 + \ldots + x_n)/n$. 6. $\alpha^* = (y_1 + \ldots + y_n)/n$, $\beta^* = (x_1 y_1 + \ldots + x_n y_n)/(x_1^2 + \ldots + x_n^2)$, $(\sigma^2)^* = \frac{1}{n}\sum_{i=1}^{n} (y_i - \alpha^* - \beta^* x_i)^2$.

### Глава 11

1. $\lambda t + \lambda^2 t(t+s)$. 2. $\mathsf{P}(\tau_1 < t) = 1 - e^{-\lambda t}$, $t > 0$. 4. $\mathsf{P}(\tau_1 < t_1,\ \tau_2 < t_2) = (1 - e^{-\lambda t_1})(1 - e^{-\lambda t_2})$, $t_1 > 0$, $t_2 > 0$. 5. $\frac{(\lambda t_1)^{k_1}}{k_1!} e^{-\lambda t_1} \times \frac{[\lambda (t_2 - t_1)]^{k_2 - k_1}}{(k_2 - k_1)!} e^{-\lambda (t_2 - t_1)} \frac{[\lambda (T - t_2)]^{k_3 - k_2}}{(k_3 - k_2)!} e^{-\lambda (T - t_2)}$. 6. $bt$.
7. $\lambda = (1-p)/p$. 8. $\mathsf{D}\mu_t = B_t = t$, $\mathsf{M}\mu_t = 1$.

9. $$\begin{pmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{pmatrix} = \frac{1}{\alpha+\beta} \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + \frac{e^{-(\alpha+\beta)t}}{\alpha+\beta} \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}.$$

10. $$\begin{cases} \dfrac{\partial f_1}{\partial t} = -(\alpha - is) f_1 + \alpha f_2, \\[2mm] \dfrac{\partial f_2}{\partial t} = \beta f_1 - (\beta + is) f_2, \quad f_1(0,\ s) = f_2(0,\ s) = 1. \end{cases}$$

11.
$$\begin{cases} \dfrac{\partial M_1}{\partial t} - v_1 \dfrac{\partial M_1}{\partial t} = -\alpha M_1 + \alpha M_2, \\ \dfrac{\partial M_2}{\partial t} + v_2 \dfrac{\partial M_2}{\partial x} = \beta M_1 - \beta M_2, \end{cases}$$

$$\begin{cases} \dfrac{dm_1}{dt} = -\alpha m_1 + \alpha m_2 + v_1, \\ \dfrac{dm_2}{dt} = \beta m_1 - \beta m_2 - v_2, \end{cases} \quad \text{где} \quad \begin{matrix} m_k(t) = M_k(t, 0), \\ m_1(0) = m_2(0) = 0. \end{matrix}$$

13.
$$\begin{cases} M_1(t, x) = e^{-\alpha t}(x + vt) + \displaystyle\int_0^t \alpha e^{-\alpha u} M_2(t-u, x+vu)\, du, \\ M_2(t, x) = e^{-\beta t}\left(x - \dfrac{at^2}{2}\right) + \displaystyle\int_0^t \beta e^{-\beta u} M_1\left(t-u, x - \dfrac{au^2}{2}\right) du, \end{cases}$$

$$\begin{cases} m_1(t) = \dfrac{v}{\alpha}(1 - e^{-\alpha t}) + \displaystyle\int_0^t \beta e^{-\beta u} m_1(t-u)\, du, \\ m_2(t) = -\dfrac{2}{\beta^2}[1 - (1 + \beta t) e^{-\beta t}] + \displaystyle\int_0^t \beta e^{-\beta u} m_1(t-u)\, du, \end{cases}$$

где $m_k(t) = M_k(t, 0)$, $M_k(t, x) = x + m_k(t)$, $m_k(0) = 0$.

## ЛИТЕРАТУРА

[1] Большев Л. Н., Смирнов Н. В., Таблицы математической статистики, «Наука», М., 1965.
[2] Боровков А. А., Теория вероятностей, «Наука», М., 1976.
[3] Гельфонд А. О., Исчисление конечных разностей (изд. 3-е), «Наука», М., 1967.
[4] Гнеденко Б. В., Курс теории вероятностей (изд. 5-е), «Наука», М., 1969.
[5] Ермаков С. М., Метод Монте-Карло и смежные вопросы, «Наука», М., 1971.
[6] Кендалл М., Моран П., Геометрические вероятности (перев. с англ.), «Наука», М., 1972.
[7] Коваленко И. Н., Филиппова А. А., Теория вероятностей и математическая статистика, «Высшая школа», М., 1973.
[8] Колмогоров А. Н. *), Основные понятия теории вероятностей, «Наука», М., 1974.
[9] Колчин В. Ф., Севастьянов Б. А., Чистяков В. П., Случайные размещения, «Наука», М., 1976.
[10] Крамер Г., Математические методы статистики (перев. с англ., изд. 2-е), «Мир», М., 1975.
[11] Мешалкин Л. Д., Сборник задач по теории вероятностей, Изд-во МГУ, 1963.
[12] Прохоров Ю. В., Розанов Ю. А., Теория вероятностей (Основные понятия, предельные теоремы, случайные процессы) (изд. 2-е), «Наука», М., 1973.
[13] Розанов Ю. А., Случайные процессы, «Наука», М., 1971.
[14] Романовский В. И., Дискретные цепи Маркова, Гостехиздат, М., 1949.
[15] Севастьянов Б. А., Ветвящиеся процессы, «Наука», М., 1971.
[16] Смирнов Н. В., Дунин-Барковский И. В., Курс теории вероятностей и математической статистики для технических приложений (изд. 3-е), «Наука», М., 1969.
[17] Феллер В., Введение в теорию вероятностей и ее приложения, т. I (перев. с англ.), «Мир», М., 1967.
[18] Ченцов Н. Н., Чистяков В. П., Теория вероятностей (методические указания), Изд-во МГУ, 1961.

*) Книга А. Н. Колмогорова «Основные понятия теории вероятностей» была впервые издана в 1933 г. на немецком языке.

## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ