

**Финансовая академия
при Правительстве Российской Федерации
Кафедра математики и финансовых приложений**

Солодовников А.С., Браилов А.В., Бабайцев В.А.

Теория вероятностей

Курс лекций

Москва, 2001.

УДК 33.51(075.8)
ББК 22.18
С60

Р е ц е н з е н т ы :

доктор экономических наук, профессор (ФА)

И.Н. Дрогобыцкий,

доктор физико-математических наук, профессор (ФА)

С.А. Семаков,

доктор физико-математических наук, профессор (МИИ)

А.В. Тищенко

Бабайцев В.А., Браилов А.В., Солодовников А.С.

Математика в экономике. Теория вероятностей: Курс лекций.

М.: Финансовая академия, 2002. 232 с.

В данной книге излагается математическая дисциплина – теория вероятностей, которая является теоретическим фундаментом математической статистики. После рассмотрения классических вероятностных схем вводится понятие вероятностного пространства на основе аксиоматики Колмогорова. Затем изучаются случайные величины и их числовые характеристики. Новым по сравнению с предыдущим изданием является понятие случайного вектора. Детально рассматривается многомерное нормальное распределение. Приводится доказательство центральной предельной теоремы.

ISBN 5-7942-0290-4

© Бабайцев В.А.

© Браилов А.В.

© Солодовников А.С.

© Финансовая академия

при Правительстве РФ, 2002

Предисловие

Теория вероятностей кажется не совсем обычной математической дисциплиной, так как имеет дело с особой категорией – со случайностью. Роль случая в нашей жизни, как известно, весьма значительна. Испытает ли ваше здоровье какое-либо потрясение в ближайшие несколько лет, попадет ли купленный вами автомобиль в аварию в течение ближайшего года, отдаст ли клиент коммерческого банка кредит в установленные сроки, – во всех подобных случаях сделать уверенный прогноз затруднительно. Это заставляет нас в каждом отдельном случае оценивать риск и прибегать к различным страховочным процедурам. Соответствующие расчеты основаны на теории вероятностей, являющейся в настоящее время краеугольным камнем всех естественных наук. Теперь можно видеть, насколько прав был французский ученый Лаплас, когда писал еще в 1812 году: "Замечательно, что наука, которая начала с рассмотрения азартных игр, обещает стать наиболее важным объектом человеческого знания... Ведь большей частью важнейшие жизненные вопросы являются на самом деле задачами из теории вероятностей".

В настоящем пособии излагаются основные понятия и формулы теории вероятностей.

Основные понятия и простейшие формулы

§ 1.1. Классическое определение вероятности

Рассмотрим опыт, состоящий в подбрасывании монеты. Ясно, что у этого опыта при каждом испытании может быть только два исхода: выпал герб или выпала цифра. Очевидно, что, во-первых, мы не можем предсказать, какая сторона монеты выпадет в очередном испытании, и, во-вторых, из соображений симметричности монеты эти два исхода можно считать равновероятными.

Следующий пример связан с подбрасыванием игральной кости – симметричного кубика с гранями, помеченными цифрами от 1 до 6. У данного опыта имеется шесть исходов: при каждом испытании может выпасть любая цифра в указанном диапазоне. Выпадение той или иной грани является случайным и равновероятным событием.

Легко усложнить любой из рассмотренных примеров. Допустим, что опыт состоит в том, что мы подбрасываем сразу три монеты. Тогда общее число исходов возрастет до восьми и мы можем все их перечислить. Обозначим для краткости выпадение герба буквой Г и цифры буквой Ц. Тогда все возможные исходы опыта можно записать в виде последовательностей из трех таких букв. Например, последовательность ГГЦ означает, что на первой и второй монетах выпал герб, а на третьей цифра. Таким образом, перечень всевозможных исходов имеет вид:

ГГГ, ГГЦ, ГЦГ, ГЦЦ, ЦГГ, ЦГЦ, ЦЦГ, ЦЦЦ.

Поскольку мы считаем, что все монеты симметричные, то можно считать все эти исходы равновероятными. Допустим, что мы хотим выяснить, насколько часто произойдет случайное событие A , состоящее в том, что цифра выпала ровно два раза. Глядя на этот перечень, легко увидеть, что благоприятных исходов три, а именно ГЦЦ, ЦГЦ, ЦЦГ. Разделив это число на общее число исходов, мы получим число – вероятность случайного события A

$$P(A) = \frac{3}{8}.$$

Рассмотрим математическую модель таких и подобных им примеров. Итак, *пространством элементарных событий* назовем конечное множество Ω , состоящее из *исходов* или *элементарных событий* $\omega_1, \dots, \omega_n$. *Случайным событием* назовем любое подмножество пространства элементарных событий $A \subset \Omega$, $A = \{\omega_{i_1}, \dots, \omega_{i_k}\}$ и будем говорить, что исходы $\omega_{i_1}, \dots, \omega_{i_k}$ *благоприятны* для события A . Будем говорить, что событие A *наступило*, если в опыте наблюдался один из благоприятных исходов. *Вероятностью* случайного события A называется отношение числа благоприятных исходов k к общему числу исходов n :

$$P(A) = \frac{k}{n}. \quad (1.1)$$

В этой модели отражены существенные (и идеализированные) черты рассмотренных нами опытов. А именно, все элементарные события равноправны, что характеризует симметричность исходов, и все они различны. Такое определение вероятности называется *классическим*.

В опыте с подбрасыванием монеты пространство элементарных событий состоит из двух исходов $\Omega = \{\Gamma, \Pi\}$, а вероятность выпадения герба $A = \{\Gamma\}$ равна $P(A) = \frac{1}{2}$. В опыте с подбрасыванием игральной кости пространство элементарных событий состоит из шести исходов $\Omega = \{1, 2, 3, 4, 5, 6\}$. Найдем вероятность случайного события A – выпадения четного числа, большего трех. В этом случае событие A состоит из двух исходов: $A = \{4, 6\}$, поэтому $P(A) = \frac{2}{6} = \frac{1}{3}$. В опыте с подбрасыванием трех монет пространство элементарных событий состоит из восьми исходов $\Omega = \{\Gamma\Gamma\Gamma, \Gamma\Gamma\Pi, \Gamma\Pi\Gamma, \Pi\Gamma\Gamma, \Pi\Gamma\Pi, \Pi\Pi\Gamma, \Pi\Pi\Pi\}$. Найдем вероятность случайного события A , состоящего в выпадении одинаковых сторон всех трех монет. Событие A состоит из двух благоприятных исходов: $A = \{\Gamma\Gamma\Gamma, \Pi\Pi\Pi\}$, так что $P(A) = \frac{2}{8} = \frac{1}{4}$.

В общем случае из определения вероятности случайного события (1.1) легко следует его свойство.

1. Вероятность случайного события заключена в пределах от 0 до 1.

$$0 \leq P(A) \leq 1. \quad (1.2)$$

Крайние значения вероятности: 0 и 1 тоже принимаются. Введем следующие определения.

Событие A , которое произойдет при любом испытании, называется *достоверным*. Это означает, что выполняется равенство $A = \Omega$. Например, в опыте с подбрасыванием игральной кости событие A , задаваемое условием «число выпавших очков положительно», будет достоверным.

2. Вероятность достоверного события равна 1.

Событие A , которое не может произойти ни при одном испытании, называется *невозможным*. Иными словами, выполняется равенство $A = \emptyset$. Например, событие A , задаваемое условием «при подбрасывании игральной кости выпало 7 очков», является невозможным.

3. Вероятность невозможного события равна 0.

Мы видим, что в рамках этой модели подсчет вероятности состоит в установлении того, как случайное событие A выражается через более простые события и, в конечном счете, через элементарные, а также подсчета чисел k и n .

§ 1.2. Алгебра случайных событий

При нахождении вероятностей приходится, естественно, учитывать связи между событиями. Формы таких связей весьма многообразны. Наиболее простые из них заключаются в том, что *одни события являются комбинациями других*. Далее мы ознакомимся с тремя основными видами комбинаций: *суммой событий, произведением событий, противоположным событием*.

1°. Сумма событий. Пусть с некоторым пространством элементарных событий связаны события A и B . Их *суммой* называется третье событие $A + B$, которое (по определению) состоит из исходов, благоприятных либо для A либо для B . Если мы условимся исход, благоприятный для события, обозначать знаком "1", а неблагоприятный – знаком "0", то полную характеристику события $A+B$ будет давать следующая таблица:

A	B	$A + B$
1	1	1
1	0	1
0	1	1
0	0	0

Аналогично определяется сумма трех событий, четырех и т.д. Вообще, сумма любого множества событий есть событие, состоящее из тех и только тех исходов, которые являются благоприятными хотя бы для одного из событий данного множества.

Приведем пример. Пусть пространство элементарных событий заключается в выборе наугад точки из области Ω , являющейся квадратом на плоскости (такой опыт осуществляет брошенный наугад бильярдный шар – после ряда отражений от бортов бильярдного стола шар останавливается в случайной точке). Если A означает попадание точки в верхнюю половину квадрата (рис. 1.1), а B – попадание в правую половину, то $A+B$ будет означать попадание в область, являющуюся объединением указанных половин.

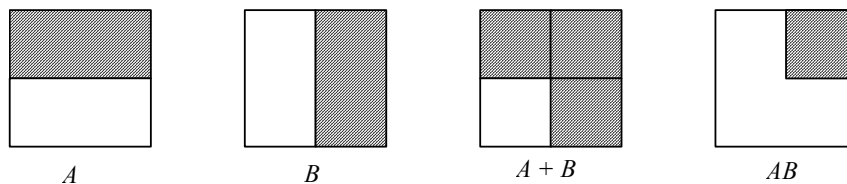


Рис. 1.1

Другой пример. Пусть в опыте с бросанием игральной кости событие A есть выпадение числа, кратного 2, а B – выпадение числа, кратного 3. Тогда $A + B$ будет выпадение хотя бы одного из чисел 2, 3, 4, 6.

2°. Произведение событий. Пусть A и B – два события. Их *произведением* называется третье событие AB , которое состоит из тех и только тех исходов, которые благоприятны одновременно для A и B .

Таблица, характеризующая событие AB , имеет вид:

A	B	AB
1	1	1
1	0	0

0	1	0
0	0	0

Аналогично определяется произведение любого множества событий. Это событие, состоящее из исходов, которые благоприятны для всех событий данного множества.

Если, например, A и B – события из приведенного выше примера с выбором точки внутри квадрата, то AB будет означать попадание точки в правую верхнюю четверть квадрата. В примере с бросанием игральной кости событие AB означает выпадение 6 очков. Если в том же примере в качестве A принять выпадение четного числа очков, а в качестве B – выпадение нечетного числа очков, то AB будет означать невозможное событие.

3°. Противоположное событие. Противоположное событие для события A обозначается $\bar{A}$. По определению в событие $\bar{A}$ входят те и только те исходы, которые не благоприятны для A .

Например, если A есть выпадение четного числа очков при бросании игральной кости, то $\bar{A}$ – это выпадение нечетного числа очков; если A – это попадание при выстреле, то $\bar{A}$ – промах; если A означает исправность всех элементов некоторой системы, то $\bar{A}$ – выход из строя хотя бы одного из элементов.

Таблица, характеризующая событие $\bar{A}$, выглядит так:

A	$\bar{A}$
1	0
0	1

Беря несколько событий $A, B, C, D, \dots$ и применяя к ним в любом порядке операции сложения и умножения, а также используя переход к противоположным событиям, можно строить различные комбинации, например: $AB + \bar{C}$, $(A\bar{C} + B)D$ и т.п. Читатель должен отчетливо понимать смысл подобных выражений, научиться быстро и безошибочно перечислять благоприятные или неблагоприятные исходы той или иной комбинации. Укажем, например, таблицу, характеризующую событие $AB + \bar{C}$:

A	B	C	$AB + \bar{C}$
1	1	1	1
1	1	0	1
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	0	1	0

0	0	0	1
---	---	---	---

Приведем еще один пример.

Покупаются три лотерейных билета; событие A_1 означает выигрыш по первому билету, A_2 – выигрыш по второму, A_3 – по третьему. Рассмотрим следующую комбинацию:

$$A_1 A_2 + A_2 A_3 + A_1 A_3. \quad (1.3)$$

Согласно определению операций сложения и умножения благоприятными исходами для события (1.3) являются любой из трех случаев: выигрывают 1-й и 2-й билеты, выигрывают 2-й и 3-й, выигрывают 1-й и 3-й. Другими словами, событие (1.3) означает выигрыш *не менее чем по двум билетам*.

Аналогичным образом, рассмотрев комбинацию

$$A_1 A_2 \bar{A}_3 + A_1 \bar{A}_2 A_3 + \bar{A}_1 A_2 A_3, \quad (1.4)$$

легко убедиться, что событие (1.4) означает выигрыш *ровно по двум билетам*.

4°. Следствия событий. Равные события.

По определению, *событие A влечет за собой событие B или событие B является следствием события A* (обозначение: $A \subset B$), если каждый исход, благоприятный для A , является благоприятным и для B .

Иными словами, все элементарные события, из которых состоит событие A , входят в событие B .

Например, пусть событие A состоит в том, что при бросании игральной кости выпало нечетное число, меньшее 5, а событие B – выпавшее число меньше 4. Легко видеть, что $A \subset B$.

Другой пример. Условие: если в семье муж старше жены (B) и жене больше 50 лет (C), то мужу больше 50 лет (A) – можно записать в виде импликации $BC \subset A$.

События A и B равны (обозначение: $A = B$) в случае, когда они являются следствиями друг друга.

Иными словами, события A и B равны, если всякий раз, когда наступает одно из них, наступает и другое, или они состоят из одних и тех же исходов.

Разумеется, равные события могут иметь отличающиеся по форме словесные описания. Например, события "не все студенты данного курса успешно сдали теорию вероятностей" и "по крайней мере один из студентов данного курса не сдал теорию вероятностей" равны, хотя и выражены различными оборотами речи.

5°. Разность событий.

Событие C является разностью событий B и A (обозначение: $C = B \setminus A$), если оно содержит все исходы, благоприятные для B , не являющиеся исходами, благоприятными для A .

Легко видеть, что выполняется равенство

$$B \setminus A = B \bar{A}.$$

И наоборот, противоположное событие можно выразить с помощью этой операции:

$$\bar{A} = \Omega \setminus A.$$

Надо иметь в виду, что так определенные операции сложения и вычитания событий все-таки отличаются от аналогичных действий с числами. В частности, событие $(A \setminus B) + (B \setminus A)$, вообще говоря, не равно невозможному событию. Рассмотрим пример.

Пример 1.1. Пусть в группе из 20 студентов имеются три подгруппы, состоящие из 10 студентов, которые знают английский и французский языки, 6 – знающих французский и немецкий, 4 – английский и немецкий. Рассмотрим событие A – студент знает английский язык, B – студент знает французский. Тогда $A \setminus B$ состоит из студентов, знающих английский и немецкий языки, а $B \setminus A$ из студентов, знающих французский и немецкий. Таким образом, событие $(A \setminus B) + (B \setminus A)$ состоит из 10 студентов, входящих во вторую и третью подгруппы.

6°. Некоторые тождества.

При рассмотрении операций над событиями часто приходится пользоваться двумя важными равенствами:

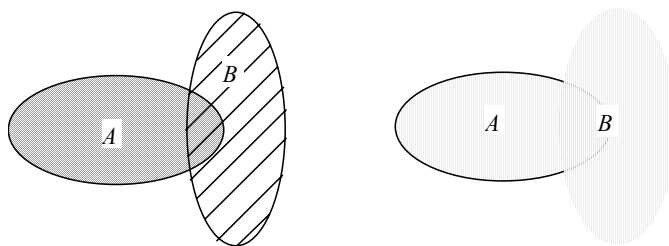
$$\overline{A_1 + A_2 + \dots + A_n} = \overline{A_1} \overline{A_2} \dots \overline{A_n}, \quad (1.5)$$

$$\overline{A_1 A_2 \dots A_n} = \overline{A_1} + \overline{A_2} + \dots + \overline{A_n}. \quad (1.6)$$

Проверим справедливость первого из них; второе проверяется аналогично.

Наступление события $A_1 + \dots + A_n$ означает, что наступает по меньшей мере одно из событий $A_1, A_2, \dots, A_n$. Наступление противоположного события $\overline{A_1 + \dots + A_n}$ означает, следовательно, что не наступает ни одно из событий $A_1, A_2, \dots, A_n$ или, по-другому, что наступают одновременно все события $\overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$; но это в точности означает наступление события $\overline{A_1} \overline{A_2} \dots \overline{A_n}$.

Для наглядного истолкования различных соотношений между событиями удобно пользоваться так называемыми *диаграммами Эйлера-Венна*. В этом случае каждое событие рассматривается как попадание случайно брошенной точки в некоторую область на плоскости; иначе говоря, каждое событие задается некоторой фигурой на плоскости. При таком истолковании событие $A + B$ будет не что иное, как попадание точки в область, являющуюся *объединением* фигур A и B (рис. 1.2), событие AB – попадание в область, являющуюся *пересечением* фигур A и B , а событие $\overline{A}$ – попадание в область, дополнительную к фигуре A . Позже мы увидим, что такой подход является универсальным: с определенной точки зрения (см. § 1.6, п. 2⁰) *каждое событие можно истолковать как некоторое множество, а операции $A + B$, AB и $\overline{A}$ над событиями – как операции объединения, пересечения и дополнения для множеств.*



§ 1.3. Элементы комбинаторики

В этом параграфе рассматриваются задачи комбинаторного характера. В каждой из них требуется подсчитать число различных вариантов, ответить на вопрос "сколько?" или "сколькими способами?". Например, интересно узнать, сколькими способами можно рассадить k людей в аудитории, где имеется n мест ($k \leq n$), каким количеством способов студент может набрать на сессии, состоящей из 4 экзаменов, сумму баллов не ниже 12, сколькими способами можно купить 10 акций трех предприятий и т.п.

Комбинаторика имеет весьма непосредственное отношение к теории вероятностей. Близость этих разделов обусловлена, прежде всего, классическим способом подсчета вероятностей. Формула

$$P(A) = \frac{k}{n},$$

где n – число всех исходов опыта, а k – число исходов, благоприятных для A , сводит вычисление $P(A)$ к нахождению двух чисел n и k ; последняя задача во многих случаях носит явно комбинаторный характер. Кроме теории вероятностей, комбинаторика используется в теории вычислительных машин, теории автоматов, в некоторых задачах экономики, биологии и т.д.

1°. Правило произведения. Будем рассматривать последовательности данной длины k :

$$(x_1, x_2, \dots, x_k),$$

состоящие из некоторых элементов $x_1, x_2, \dots, x_k$ (не обязательно различных). Условимся для краткости называть такие последовательности *строками*. Две строки $(x_1, x_2, \dots, x_k)$ и $(y_1, y_2, \dots, y_k)$ будем считать различными в том и только том случае, если хотя бы для одного номера i (из совокупности $1, 2, \dots, k$) элемент x_i отличен от y_i .

Правило произведения может быть сформулировано следующим образом.

Пусть элемент x_1 может быть выбран n_1 способами; при каждом выборе x_1 элемент x_2 может быть выбран n_2 способами; при каждом выборе пары x_1, x_2 элемент x_3 может быть выбран n_3 способами и т.д.; наконец, при каждом выборе $x_1, x_2, \dots, x_{k-1}$ элемент x_k может быть выбран n_k способами. Тогда число различных строк $(x_1, x_2, \dots, x_k)$ равно произведению $n_1 n_2 \dots n_k$.

Докажем это правило сначала для $k = 2$, т.е. для строк длины 2.

Обозначим через $a_1, a_2, \dots, a_n$ различные значения для x_1 . Среди строк (x_1, x_2) имеется ровно n_2 строк, начинающихся с a_1 (т.е. строк вида (a_1, x_2)), ровно n_2 строк, начинающихся с a_2 , и т.д. Следовательно, число всех строк (x_1, x_2) будет:

$$\underbrace{n_2 + n_2 + \dots + n_2}_{n_1 \text{ раз}} = n_1 n_2.$$

Пусть теперь $k = 3$. Любую строку

$$(x_1, x_2, x_3) \quad (1.7)$$

можно рассматривать как строку из двух объектов: строки (x_1, x_2) и элемента x_3 . Первый объект, по доказанному, может быть выбран $n_1 n_2$ способами; при любом из этих способов элемент x_3 , по условию, может быть выбран n_3 способами. Применяя опять-таки правило произведения для строк длины 2, получим, что число различных строк вида (1.7) будет

$$(n_1 n_2) \cdot n_3 = n_1 n_2 n_3.$$

Ясно, что такое же рассуждение можно применить к строкам длины 4, затем 5 и т.д.

Пример 1.2. Рассматриваются 5 различных языков. Сколько словарей нужно иметь для непосредственного перевода с любого языка на любой?

Решение. Любой словарь задается строкой (x_1, x_2) , где x_1 — язык, с которого делается перевод, а x_2 — язык, на который переводят. Объект x_1 может быть выбран 5 способами; при каждом выборе x_1 объект x_2 может быть выбран 4 способами. По правилу произведения находим, что число различных словарей будет $5 \cdot 4 = 20$.

Пример 1.3. Сколько можно составить пятизначных чисел так, чтобы любые две соседние цифры числа были различны?

Решение. Пятизначному числу с цифрами x_1, x_2, x_3, x_4, x_5 можно сопоставить строку $(x_1, x_2, x_3, x_4, x_5)$. При этом выбор цифры x_1 возможен 9 способами; если цифра x_1 выбрана, то для выбора x_2 имеется тоже 9 возможностей (x_2 может быть любой из цифр 0, 1, 2, ..., 9, отличной от x_1); после выбора x_1, x_2 для цифры x_3 имеется снова 9 возможностей и т.д. Применяя правило произведения, находим, что искомое количество чисел есть

$$9 \cdot 9 \cdot 9 \cdot 9 \cdot 9 = 9^5.$$

Пример 1.4. Сколько различных подмножеств имеет множество $A = \{a_1, a_2, \dots, a_n\}$, состоящее из n элементов?

Решение. Пусть X — подмножество в A . Сопоставим этому подмножеству строку $(x_1, x_2, \dots, x_n)$ длиной n — нечто вроде "шифра" подмножества X . А именно: положим x_1 равным 1 или 0, смотря по тому, входит или не входит элемент a_1 в подмножество X ; положим x_2 равным 1 или 0, смотря по тому, входит или не входит a_2 в X , и так далее. В результате каждому подмножеству X будет соответствовать строка длины n , состоящая из единиц и нулей. И обратно, любая строка длины n , состоящая из единиц и нулей, однозначно определяет некоторое подмножество X (например, в случае $n = 5$ строка $(0, 0, 0, 1, 1)$ определяет подмножество $X = \{a_4, a_5\}$). Но число различных строк по правилу произведения равно $2 \cdot 2 \cdot \dots \cdot 2 = 2^n$. Значит, число различных подмножеств множества A будет также 2^n .

2°. Формула включения-исключения. Для любого конечного множества A обозначим через $|A|$ число его элементов. Тогда для любых двух конечных множеств выполняется формула

$$|A \cup B| = |A| + |B| - |A \cap B|. \quad (1.8)$$

Очевидно, что если A и B не пересекаются, то $|A \cup B| = |A| + |B|$. Общий случай сводится к рассмотренному, поскольку $A \cup B = A \cup (B \setminus A)$ (см. рис. 1.3).

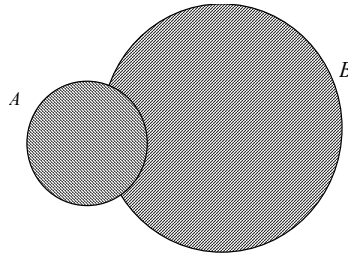


Рис. 1.3

Формула (1.8) обобщается на случай трех множеств, а именно

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|. \quad (1.9)$$

Графическая иллюстрация формулы (1.9) приведена ниже (рис. 1.4). Ее вывод мы предлагаем читателю.

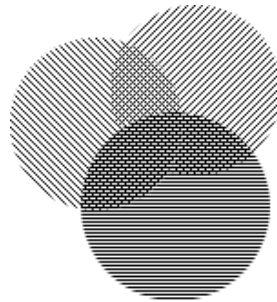


Рис. 1.4

Рассмотрим следующий пример.

Пример 1.5. В группе из 30 студентов 20 студентов (множество A) изучают английский язык, 15 студентов (B) — немецкий и 10 (C) — французский. При этом 8 студентов изучают одновременно английский и немецкий, 5 студентов — английский и французский, 4 — французский и немецкий. Сколько студентов изучают все три языка?

Решение. Имеем $|A| = 20$, $|B| = 15$, $|C| = 10$. По условию,

$$|A \cap B| = 8, \quad |A \cap C| = 5, \quad |B \cap C| = 4.$$

Применим формулу (1.9):

$$30 = 20 + 15 + 10 - 8 - 5 - 4 + |A \cap B \cap C|,$$

откуда видно, что искомое число равно 2.

Формула (1.9) по индукции легко обобщается на случай объединения любого числа множеств и в этом случае называется *формулой включения-исключения*.

3°. Размещения. Пусть X – множество, состоящее из n элементов. Любой упорядоченный набор k различных элементов множества X называется *размещением из n элементов по k* .

Для множества X , состоящего из трех элементов a, b, c , все размещения из двух элементов выглядят следующим образом:

$$(a, b), (a, c), (b, c), (b, a), (c, a), (c, b).$$

Число размещений обозначается A_n^k и вычисляется по формуле

$$A_n^k = n(n-1)\dots(n-k+1). \quad (1.10)$$

Для вывода этой формулы применим правило произведения. Действительно, для выбора первого элемента у нас имеется n возможностей, так как на первом месте может стоять любой элемент множества X . Фиксируя первый элемент, мы видим, что для выбора второго элемента у нас остается $n-1$ возможность, для выбора третьего при выбранных первых двух элементах $n-2$ возможности и т.д. Вид последнего множителя в формуле (1.10) обусловлен тем, что число множителей равно k .

Пример 1.6. Найти число способов распределения первых трех призовых мест для восьми участников финального забега.

Решение. Всякий такой способ является размещением из восьми участников по три. Поэтому число способов вычисляется по формуле (1.10):

$$A_8^3 = 8 \cdot 7 \cdot 6 = 336.$$

3°. Перестановки. Пусть X – множество, состоящее из n элементов. *Перестановкой* элементов множества X называется их расположение в каком-либо определенном порядке:

$$x_1, x_2, \dots, x_n.$$

Иными словами, перестановка является размещением из n элементов по n . Число различных перестановок обозначим P_n . Из формулы (1.10) следует, что справедлива формула

$$P_n = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n. \quad (1.11)$$

Заметим, что произведение $1 \cdot 2 \cdot \dots \cdot n$ обозначается $n!$ (читается " n факториал"). Итак, $P_n = n!$.

Например, 5 человек могут выстроиться в очередь (скажем, к кассе кинотеатра) $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ способами.

С помощью факториала формулу (1.10) можно переписать следующим образом

$$A_n^k = \frac{n!}{(n-k)!}. \quad (1.12)$$

Для доказательства достаточно умножить и разделить правую часть формулы (1.10) на $(n-k)!$.

Добавим к определению числа $n!$ равенство $0! = 1$, которое примем по определению.

4°. Сочетания. Число сочетаний. Пусть снова X – множество, состоящее из n элементов. Любое подмножество Y множества X , содержащее k элементов, называется *сочетанием k элементов из n* ; при этом, разумеется, $k \leq n$.

Число различных сочетаний k элементов из n обозначается C_n^k . Одной из важнейших формул комбинаторики является следующая формула для числа C_n^k :

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (1.13)$$

Ее можно преобразовать после очевидных сокращений следующим образом:

$$C_n^k = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}. \quad (1.14)$$

В частности,

$$C_n^0 = \frac{n!}{0!n!} = 1;$$

это вполне согласуется с тем, что в множестве X имеется только одно подмножество из 0 элементов – пустое подмножество.

Приведем доказательство формулы (1.13). Пусть Y – какое-либо подмножество множества X , содержащее k элементов. Составив всевозможные перестановки из этих элементов, получим все размещения элементов Y длиной k . Если указанную операцию проделать с *каждым* подмножеством Y , содержащим k элементов, то получим все размещения из n по k , число которых равно A_n^k . Получим формулу

$$C_n^k \cdot k! = A_n^k,$$

откуда следует формула (1.14) или (1.13) в зависимости от того, какую формулу для числа размещений подставить: (1.10) или (1.12).

Числа C_n^k обладают рядом замечательных свойств. Эти свойства, в конечном счете, выражают различные соотношения между подмножествами данного множества X . Их можно доказывать непосредственно, исходя из формулы (1.13), но более содержательными являются доказательства, опирающиеся на теоретико-множественные рассуждения.

1) Справедлива формула

$$C_n^k = C_n^{n-k}, \quad (1.15)$$

вытекающая из (1.13) очевидным образом. Смысл формулы (1.15) состоит в том, что имеется взаимнооднозначное соответствие между множеством всех k -членных подмножеств из X и множеством всех $(n-k)$ -членных подмножеств из X : чтобы установить это соответствие, достаточно каждому k -членному подмножеству Y сопоставить его дополнение во множестве X .

2) Справедлива формула

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n. \quad (1.16)$$

Поскольку сумма, стоящая в левой части, выражает собой число всех подмножеств множества X (C_n^0 есть число 0-членных подмножеств, C_n^1 – число 1-членных подмножеств и т.д.), то для доказательства формулы (1.15) достаточно сослаться на уже известный читателю факт (см. пример 4 из пункта 1°): число различных подмножеств n -членного множества X равно 2^n .

3) При любом k , $1 \leq k \leq n$, справедливо равенство

$$C_n^k = C_{n-1}^k + C_{n-1}^{k-1}. \quad (1.17)$$

Это равенство нетрудно получить с помощью формулы (1.13). В самом деле,

$$\begin{aligned} C_{n-1}^{k-1} + C_{n-1}^k &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-1-k)!} = \\ &= \frac{(n-1)!k}{k!(n-k)!} + \frac{(n-1)!(n-k)}{k!(n-k)!} = \frac{(n-1)!}{k!(n-k)!} (k + n - k) = \\ &= \frac{n!}{k!(n-k)!} = C_n^k. \end{aligned}$$

Вывод формулы (1.17), основанный на теоретико-множественных соображениях, мы предоставляем провести читателю. Укажем, что для этого следует выделить какой-то определенный элемент $a \in X$ и все k -членные подмножества разбить на две группы: подмножества, содержащие a , и подмножества, не содержащие a .

4) Рассмотрим так называемый *арифметический треугольник Паскаля*.

Равенство (1.17) позволяет вычислять значения C_n^k , если известны C_{n-1}^k и C_{n-1}^{k-1} . Иными словами, с помощью этого равенства можно последовательно вычислять C_n^k сначала при $n = 1$, затем при $n = 2$, $n = 3$ и т.д. Вычисления удобно записывать в виде треугольной таблицы:

				1						
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
1		5		10		10		5		1

в $(n + 1)$ строке которой по порядку стоят числа $C_n^0, C_n^1, \dots, C_n^n$. При этом крайние числа строки, т.е. C_n^0 и C_n^n , равны 1, а остальные числа находятся по формуле (1.17). Поскольку C_{n-1}^{k-1} и C_{n-1}^k располагаются в этой таблице строкой выше, чем число C_n^k , и находятся в этой строке слева и справа от него, то для получения числа C_n^k надо сложить находящиеся слева и справа от него числа предыдущей строки. Например, число 10 в шестой строке мы получаем, сложив числа 4 и 6 пятой строки.

Указанная таблица и есть как раз "арифметический треугольник Паскаля".

Пример 1.7. Пусть n и k – два целых числа, причем $n > 0$, $k \geq 0$. Сколько существует различных строк длиной n , состоящих из n букв a и b , с условием, что в каждой из этих строк буква a встречается k раз (и, следовательно, буква b – $n - k$ раз)?

Решение. Для примера приведем несколько строк с двумя буквами a и тремя b :

$$b a b b a, a a b b b, a b b a b.$$

Пусть $(x_1, x_2, \dots, x_n)$ – одна из строк указанного вида. Рассмотрим все номера i , такие, что $x_i = a$. Совокупность таких номеров является подмножеством множества $M = \{1, 2, \dots, n\}$, состоящим из k элементов. Обратно, если Y – любое подмножество множества M , состоящее из k элементов, то, положив $x_i = a$ для всех $i \in Y$ и $x_i = b$ для всех $i \in \bar{Y}$, получим строку $(x_1, x_2, \dots, x_n)$ требуемого вида. Значит, число указанных в задаче строк равно числу k -элементных подмножеств в n -элементном множестве M , т.е. равно числу C_n^k .

5°. Бином Ньютона. Из школьного курса читателю известны формулы:

$$(a + b)^2 = a^2 + 2ab + b^2, \quad (a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Обобщением этих формул является следующая формула, называемая обычно *формулой бинома Ньютона*:

$$(a + b)^n = C_n^0 a^0 b^n + C_n^1 a^1 b^{n-1} + C_n^2 a^2 b^{n-2} + \dots + C_n^{n-1} a^{n-1} b + C_n^n a^n b^0. \quad (1.18)$$

В этой формуле n может быть любым натуральным числом.

Вывод формулы (1.18) несложен. Прежде всего, запишем:

$$(a + b)^n = (a + b)(a + b) \dots (a + b), \quad (1.19)$$

где число перемножаемых скобок равно n . Из обычного правила умножения суммы на сумму вытекает, что выражение (1.19) равно сумме всевозможных произведений, которые можно составить следующим образом: любое слагаемое первой из сумм $a + b$ умножается на любое слагаемое второй суммы $a + b$, на любое слагаемое третьей суммы и т.д. Например, при $n = 3$ имеем:

$$(a + b)(a + b)(a + b) = aaa + aab + aba + abb + baa + bab + bba + bbb.$$

Из сказанного ясно, что слагаемым в выражении для $(a + b)^n$ соответствуют (взаимно однозначно) строки длиной n , составленные из букв a и b . Среди слагаемых будут встречаться подобные члены; очевидно, что таким членам соответствуют строки, содержащие одинаковое количество букв a . Но число строк, содержащих ровно k раз букву a , равно C_n^k (см. задачу в конце предыдущего пункта 4°). Значит, сумма всех членов, содержащих букву a множителем ровно k раз, равна $C_n^k a^k b^{n-k}$. Поскольку k может принимать значения $0, 1, 2, \dots, n-1, n$, то из нашего рассуждения следует формула (1.18).

Используя знак суммирования, формулу (1.18) можно записать короче:

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}.$$

Хотя формулу (1.18) называют именем Ньютона, в действительности она была открыта еще до Ньютона (например, ее знал Паскаль). Заслуга Ньютона состоит в том, что он нашел обобщение этой формулы на случай *нецелых* показателей.

Числа $C_n^0, C_n^1, \dots, C_n^n$, входящие в формулу (1.18), принято называть *биномиальными коэффициентами*. Из формулы (1.18) можно получить целый ряд свойств этих коэффициентов. Например, полагая $a = 1, b = 1$, получим: $2^n = C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n$, т.е. формулу (1.16). Если положить $a = 1, b = -1$, то будем иметь:

$$0 = C_n^0 - C_n^1 + C_n^2 - C_n^3 + \dots + (-1)^n C_n^n$$

или

$$C_n^0 + C_n^2 + C_n^4 + \dots = C_n^1 + C_n^3 + C_n^5 + \dots$$

§ 1.4. Некоторые примеры вычисления вероятностей

В § 1.1 мы рассмотрели классическое определение вероятности случайного события как отношение числа благоприятных исходов опыта к общему числу исходов – формулу (1.1)

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

В этом параграфе мы разберем ряд примеров непосредственного вычисления вероятности случайного события.

Пример 1.8. В урне находятся 10 шаров: 4 белых и 6 черных. Из урны наудачу извлекают один шар. Какова вероятность того, что он окажется черным (событие A)?

Решение. Представим себе, что шары снабжены номерами 1, 2, ..., 10, причем черные шары получили номера 1, 2, ..., 6.

Обозначим через A_i , где $i = 1, 2, \dots, 10$, исход опыта: извлечение шара с номером i . Интересующему нас событию A благоприятны исходы $A_1, A_2, \dots, A_6$. Значит, в данном случае $n = 10, k = 6$ и

$$P(A) = \frac{6}{10}.$$

Пример 1.9. Дважды бросается игральная кость. Какова вероятность того, что сумма очков при обоих бросаниях окажется больше 10 (событие A)?

Решение. Через A_{ij} обозначим исход опыта, состоящий в том, что при первом бросании выпало i очков, а при втором j . Тогда 36 событий

$$\begin{array}{cccc} A_{11} & A_{12} & \dots & A_{16} \\ A_{21} & A_{22} & \dots & A_{26} \\ \dots & \dots & \dots & \dots \\ A_{61} & A_{62} & \dots & A_{66} \end{array}$$

можно рассматривать как элементарные исходы опыта, заключающегося в двукратном бросании игральной кости. Действительно, при каждом осуществлении опыта наступает один и только один из этих исходов, а соображения "равноправия" (между гранями игральной кости, а также между первым

и вторым бросанием) позволяют считать указанные события равновероятными. Интересующему нас событию A благоприятны исходы A_{56}, A_{65}, A_{66} (остальные неблагоприятны). Отсюда имеем:

$$P(A) = \frac{k}{n} = \frac{3}{36} = \frac{1}{12}.$$

Пример 1.10. В лотерее разыгрывается 100 билетов. Выигрыши падают на 10 билетов. Некто покупает три билета. Какова вероятность того, что хотя бы один из них выиграет?

Решение. В данном случае опыт заключается в выборе наугад трех лотерейных билетов. Перенумеруем все возможные тройки билетов. В качестве номеров будут фигурировать числа $1, 2, \dots, n$, где $n = C_{100}^3 = \frac{100 \cdot 99 \cdot 98}{1 \cdot 2 \cdot 3} = 161700$.

Пусть A_i – исход опыта, заключающийся в покупке тройки с номером i . Тогда события $A_1, A_2, \dots, A_n$ можно рассматривать как все исходы данного опыта.

Интересующее нас событие A состоит в том, что хотя бы один из выбранных билетов оказался выигрышным. Благоприятными для A являются такие группы из трех билетов, которые содержат хотя бы один выигрышный билет, неблагоприятными – такие, в которых ни на один билет не падает выигрыш. Число неблагоприятных групп равно C_{90}^3 , следовательно, число благоприятных есть $k = C_{100}^3 - C_{90}^3$. Отсюда

$$\begin{aligned} P(A) &= \frac{k}{n} = 1 - \frac{C_{90}^3}{C_{100}^3} = 1 - \frac{90 \cdot 89 \cdot 88}{100 \cdot 99 \cdot 98} = \\ &= 1 - \left(1 - \frac{10}{100}\right) \left(1 - \frac{10}{99}\right) \left(1 - \frac{10}{98}\right). \end{aligned} \quad (1.20)$$

Полученное выражение приближенно равно:

$$1 - \left(1 - \frac{10}{100}\right) \left(1 - \frac{10}{99}\right) \left(1 - \frac{10}{98}\right) = 1 - (0,9)^3 = 0,271 \dots$$

Впрочем, выражение (1.20) нетрудно подсчитать точно. Такой подсчет дает $P(A) = 0,2735 \dots$

Рассмотрим в связи с последним примером еще один пример.

Пример 1.11. В условиях лотереи примера 1.10 выяснить, какое минимальное число билетов нужно купить, чтобы вероятность получения хотя бы одного выигрыша оказалась большей, чем 0,5.

Решение. Пусть покупаются m билетов. Обозначим вероятность выигрыша хотя бы по одному из них через p_m . Понятно, что с ростом m число p_m будет возрастать. Наша цель – найти наименьшее значение m , при котором это число больше 0,5.

Рассуждая, как в примере 1.10, получим:

$$p_m = 1 - \left(1 - \frac{10}{100}\right) \left(1 - \frac{10}{99}\right) \dots \left(1 - \frac{10}{100 - (m-1)}\right). \quad (1.21)$$

Следовательно, должно выполняться неравенство

$$p_m > 1 - \left(1 - \frac{10}{100}\right)^m = 1 - (0,9)^m.$$

Таким образом, для наших целей достаточно, чтобы выполнялось неравенство

$$1 - (0,9)^m > 0,5,$$

или $(0,9)^m > 0,5$. Логарифмируя по десятичному основанию и решая полученное неравенство, получим

$$m > \frac{\lg 0,5}{\lg 0,9} \approx 6,6.$$

Таким образом, искомое значение m равно 7. Непосредственное вычисление вероятности по формуле (1.21) дает следующие значения:

$$p_6 = 1 - \frac{C_{90}^6}{C_{100}^6} \approx 0,478, \quad p_7 = 1 - \frac{C_{90}^7}{C_{100}^7} \approx 0,533.$$

Многие задачи на подсчет вероятностей можно свести к так называемой *схеме случайного выбора*. Рассмотрим два основных варианта этой схемы: *выбор с возвращением* и *выбор без возвращения*.

1) *Выбор с возвращением*. Представим себе, что в некотором ящике собрано n различных предметов $a_1, a_2, \dots, a_n$. Из ящика наугад извлекается один из предметов, регистрируется, затем кладется обратно в ящик. Если осуществить k таких извлечений, то получим некоторую строку длиной k , составленную из элементов множества $X = \{a_1, a_2, \dots, a_n\}$. Она называется *выборкой с возвращением объема k* из множества X . Число различных выборок объема k согласно правилу произведения равно n^k .

Описанная процедура носит название *случайного выбора с возвращением*. Слово "случайный" в этом названии означает нечто большее, нежели просто тот факт, что состав выборки предсказать заранее невозможно. Мы условимся вкладывать в это слово следующий смысл: *все n^k выборок равновозможны*. Другими словами, опыт состоит из n^k исходов, и вероятность появления любой конкретной выборки равна $\frac{1}{n^k}$.

К схеме случайного выбора с возвращением можно свести большое число опытов. Например, бросание монеты можно представить как случайный выбор одного элемента из множества $X = \{\text{герб, цифра}\}$. Вместо двукратного бросания игральной кости можно рассматривать случайный выбор с возвращением двух элементов из множества $X = \{1, 2, 3, 4, 5, 6\}$. Выяснение дней рождения k случайных прохожих можно заменить случайным выбором с возвращением k элементов из множества $X = \{1, 2, 3, \dots, 365\}$ и т.д.

2) *Выбор без возвращения*. В этом случае выбранный предмет не кладется обратно в ящик и следующее извлечение производится из меньшего числа предметов. После k извлечений получаем строку длиной k без повторений. Число таких строк, как следует из правила произведения, будет равно числу размещений из n по k .

$$A_n^k = n(n-1)(n-2)\dots(n-k+1).$$

Случайный характер выбора понимается, как и выше, в том смысле, что опыт состоит из всех равновозможных выборок данной длины.

Пример 1.12. Пусть из совокупности n предметов извлекаются с возвращением k предметов. Найти вероятность того, что все предметы, составляющие выборку, окажутся различными (событие S).

Решение. В данном случае число всех элементарных исходов опыта равно n^k , а число исходов, благоприятных для события S , равно A_n^k . Отсюда искомая вероятность

$$P(S) = \frac{A_n^k}{n^k} = \frac{n(n-1)(n-2)\dots(n-(k-1))}{n^k}. \quad (1.22)$$

Остановимся на одном частном случае разобранного выше примера – так называемом *парадоксе дня рождения*.

Пример 1.13. На лекции присутствует k студентов. Какова вероятность того, что хотя бы у двух студентов дни рождения совпадают (событие T)?

Решение. Как уже отмечалось, выяснение дней рождения у k случайно собравшихся людей можно заменить выбором с возвращением k элементов из множества $X = \{1, 2, 3, \dots, 365\}$. Нам необходимо найти вероятность события T – совпадения дней рождения у каких-либо двух студентов. Событие, противоположное T , заключается в том, что все дни рождения различны – выше это событие было обозначено S . Формула (1.22) при $n = 365$ дает:

$$P(S) = \frac{365 \cdot 364 \dots (365 - k + 1)}{365^k},$$

откуда следует:

$$\begin{aligned} P(T) &= \frac{365^k - 365 \cdot 364 \dots (365 - k + 1)}{365^k} = \\ &= 1 - \frac{365 \cdot 364 \dots (365 - k + 1)}{365^k}. \end{aligned} \quad (1.23)$$

Найденное нами выражение для $P(T)$ зависит, естественно, от k – числа студентов на лекции. Подсчитав $P(T)$ для различных значений k , можно получить такую таблицу:

k	5	10	22	23	30	60
$P(T)$	0,027	0,117	0,476	0,507	0,706	0,994

(все знаки после запятой, начиная с четвертого, отброшены). Из таблицы видно, что если в аудитории находятся всего лишь 23 человека, то уже и тогда имеется более половины шансов на то, что, по крайней мере, у двух из них дни рождения совпадают!

Пример 1.14. Монету бросают 10 раз. Какова вероятность того, что герб при этом выпадет ровно 3 раза (и, следовательно, цифра выпадет 7 раз)?

Решение. Десятикратное бросание монеты можно рассматривать как составление строки длиной 10 (с повторениями) из элементов множества $X = \{\Gamma, \Pi\}$. Число всех строк такого рода равно 2^{10} . Строк, в которых элемент " Γ " встречается 3 раза, а " Π " входит 7 раз, будет C_{10}^3 . Отсюда искомая вероятность

$$p = \frac{C_{10}^3}{2^{10}} = \frac{10 \cdot 9 \cdot 8}{3! 2^{10}} = \frac{15}{128}.$$

Пример 1.15. Слово "каре́та", составленное из букв-кубиков, рассыпано на отдельные буквы, которые затем сложены в коробке. Из коробки наугад извлекают буквы одну за другой. Какова вероятность получить при таком извлечении слово "раке́та"?

Решение. Здесь нет схемы случайного выбора в прежнем понимании, так как буквы, сложенные в коробке, не все различны (три одинаковые буквы "а"). Представим себе, что одинаковые буквы (в данном случае a, a, a) индивидуализированы с помощью знаков 1, 2, 3 (превратились в a_1, a_2, a_3). Следовательно,

$$P(A) = \frac{3!}{6!} = \frac{1}{4 \cdot 5 \cdot 6} = \frac{1}{120}.$$

Пример 1.16. (задача о выборке). Партия готовых изделий содержит ровно N изделий – A стандартных и B бракованных ($A + B = N$). Из партии наудачу извлекают n изделий. Какова вероятность того, что в выборке будет a стандартных изделий и b бракованных (где $a + b = n$)?

Решение. Выбор n изделий из N возможен C_N^n равновероятными способами. Подсчитаем, в скольких случаях будет получаться выборка, содержащая a стандартных и b бракованных. Число различных групп, состоящих из a стандартных изделий, равно C_A^a . Число различных групп, состоящих из b бракованных изделий, равно C_B^b . По правилу произведения число различных выборок, содержащих a стандартных изделий и b бракованных, будет $C_A^a \cdot C_B^b$. Следовательно, вероятность получить выборку из a стандартных изделий и b бракованных равна $\frac{C_A^a C_B^b}{C_{A+B}^{a+b}}$.

§ 1.5. Теорема сложения вероятностей

В этом параграфе в рамках классического определения вероятности мы рассмотрим формулы, выражающие вероятность суммы событий через вероятности слагаемых. Справедлива следующая

Теорема 1.1. Для любых событий A и B выполняется формула

$$P(A + B) = P(A) + P(B) - P(AB). \quad (1.24)$$

Доказательство. Обозначим через $n = |\Omega|$ – общее число исходов, n_S – число исходов, благоприятных для события S . Тогда формулу (1.8) можно переписать следующим образом

$$n_{A+B} = n_A + n_B - n_{AB}. \quad (1.25)$$

Разделив почленно формулу (1.25) на n , получим формулу (1.24).

В качестве следствия получим наиболее часто используемый факт.

Правило сложения вероятностей. Если события A и B несовместны, то

$$P(A + B) = P(A) + P(B). \quad (1.26)$$

Д о к а з а т е л ь с т в о . Действительно, поскольку события A и B несовместны, то $P(AB) = 0$, поэтому из формулы (1.24) получим формулу (1.26).

Теорема сложения вероятностей (формулы (1.24), (1.26)) допускает обобщение на любое число слагаемых. Основой для вывода таких формул служит формула включения-исключения. Приведем формулировку теоремы для трех событий.

Теорема 1.2. Для любых трех событий A, B, C справедлива формула

$$P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC). \quad (1.27)$$

Из правила сложения вероятностей для двух событий вытекает одно простое, но важное следствие.

Следствие. Сумма вероятностей противоположных событий равна единице.

$$P(A) + P(\bar{A}) = 1. \quad (1.28)$$

Действительно, события A и $\bar{A}$ несовместны, так что можно применить формулу (1.26). При этом следует использовать то, что событие $A + \bar{A}$ – достоверное, поэтому его вероятность равна 1.

Очень часто используют вариант формулы (1.28), выражающий вероятность события A через вероятность противоположного события $\bar{A}$.

$$P(A) = 1 - P(\bar{A}). \quad (1.29)$$

Следующее неравенство связывает вероятности двух событий, одно из которых является следствием другого.

Следствие. Если событие B является следствием события A , то

$$P(A) \leq P(B). \quad (1.30)$$

Действительно, поскольку $A \subset B$, то $B = A + (B \setminus A)$ – сумма несовместных событий. Применяя правило сложения вероятностей, имеем

$$P(B) = P(A) + P(B \setminus A),$$

откуда следует неравенство (1.30).

Рассмотрим несколько примеров применения теоремы сложения вероятностей.

Пример 1.17. Первый пример основан на историческом факте. Шевалье де Мере обратился к известному математику и философу Паскалю с просьбой объяснить следующее эмпирическое наблюдение. Если бросить игральную кость четыре раза, то выгоднее держать пари, что по меньшей мере

один раз выпадет шестерка. Иными словами, надо показать, что вероятность указанного события больше $\frac{1}{2}$.

Решение. Обозначим через A событие, состоящее в выпадении хотя бы один раз шестерки при четырех бросаниях игральной кости. Рассмотрим противоположное событие $\bar{A}$ – шестерка не выпала ни разу. По правилу произведения $n = 6^4 = 1296$, $n_{\bar{A}} = 5^4 = 625$. Поэтому

$$P(\bar{A}) = \frac{n_{\bar{A}}}{n} = \frac{625}{1296}.$$

По формуле (1.29) имеем

$$P(A) = 1 - \frac{625}{1296} = \frac{671}{1296} > \frac{1}{2}.$$

Пример 1.18. Пять открыток случайным образом раскладывают по трем конвертам. Какова вероятность того, что в каждом конверте окажется хотя бы одна открытка?

Решение. Прежде всего определим, из каких исходов состоит размещение открыток по конвертам. Всякий вариант можно представить в виде строки $(x_1, x_2, x_3, x_4, x_5)$, где x_i – обозначает номер конверта, в который вложили i -ю открытку. Поэтому общее число исходов по правилу произведения равно $n = 3^5 = 243$.

Обозначим через A событие, вероятность которого мы находим, A_i , $i = 1, 2, 3$ – ни одну открытку не вложили в i -й конверт. Тогда

$$A = \overline{A_1 + A_2 + A_3},$$

так что

$$P(A) = 1 - P(A_1 + A_2 + A_3). \quad (1.31)$$

Для вычисления вероятности события $A_1 + A_2 + A_3$ применим формулу сложения вероятностей (1.27) для трех событий. При этом

$$P(A_1) = P(A_2) = P(A_3) = \frac{2^5}{3^5} = \frac{32}{243},$$

$$P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1}{243},$$

$$P(A_1 A_2 A_3) = 0,$$

поскольку $A_1 A_2 A_3$ – невозможное событие. Подставляя в формулу (1.27), получим

$$P(A_1 + A_2 + A_3) = 3 \frac{32}{243} - \frac{3}{243} = \frac{93}{243}.$$

Отсюда искомая вероятность

$$P(A) = 1 - \frac{93}{243} = \frac{150}{243}.$$

§ 1.6. Аксиомы теории вероятностей

1°. Предварительные соображения. До сих пор мы исходили из классического определения вероятности (1.1), которое основано на равноправности конечного числа исходов опыта, их симметричности. Оба этих предположения: симметричность и конечность могут быть не приложимы в некоторых ситуациях. Рассмотрим следующие примеры.

1. Статистика рождаемости показывает, что по многолетним наблюдениям частота рождения мальчиков несколько больше 0,51.

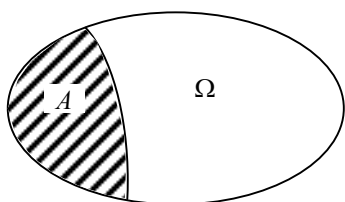
2. При бросании неправильной монеты опыт по-прежнему состоит из двух испытаний, однако вероятность выпадения герба или цифры может принимать любые значения, отличные от 0,5.

3. Пусть опыт состоит в подбрасывании симметричной монеты до тех пор, пока не появится герб. В этом случае мы сталкиваемся с ситуацией, когда число подбрасываний не ограничено. Пусть герб появился при k -й попытке. Нетрудно сообразить, что вероятность такого события (в опыте, в котором монета подбрасывается k раз) равна $\frac{1}{2^k}$. Сумма всех таких вероятностей равна 1, что следует из равенства

$$\frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^k} + \dots = 1.$$

Таким образом, число исходов опыта может быть бесконечным, в рассматриваемом случае – счетным.

4. Рассмотрим еще один пример. Пусть опыт заключается в бросании маленького шарика в область Ω на плоскости (см. рис.).



Будем считать исходом опыта попадание шарика в определенную точку области Ω . В данном примере множество элементарных событий – исходов опыта бесконечно. Идеализируя ситуацию, т.е. предполагая, что выбирается точка области Ω , можно считать, что множество исходов опыта совпадает с множеством всех точек области Ω . Событием A в этом случае является попадание точки в некоторое подмножество.

Такой подход к понятию события удобен еще и тем, что благодаря ему понятия суммы и произведения событий, а также противоположного события приобретают естественный теоретико-множественный смысл. А именно: сумма событий A и B превращается в *объединение соответствующих подмножеств*, произведение событий A и B – в *пересечение тех же подмножеств*, а противоположное событие $\bar{A}$ – в *дополнение к подмножеству A (в Ω)*.

Переходим к изложению современной аксиоматики теории вероятностей, которая принадлежит советскому математику А.Н. Колмогорову. Для большей ясности разобьем аксиомы на две группы: *аксиомы событий* и *аксиомы вероятностей*.

2°. Аксиомы событий. Мы исходим из того, что задано некоторое множество Ω . Элементы этого множества называются *элементарными событиями*. Что представляют собой элементарные со-

бытия, какова их конкретная природа, для нас это сейчас безразлично. Иногда мы будем называть элементарные события *элементарными исходами*.

Далее, мы предполагаем, что фиксирована некоторая (непустая) совокупность подмножеств множества Ω , называемых *событиями*. События в дальнейшем обозначаются заглавными латинскими буквами A, B, C и т.д., а вся совокупность событий – буквой S .

К совокупности S предъявим следующие два требования (аксиомы событий):

- I. Если множества $A_1, A_2, \dots$ (в конечном или счетном множестве) суть события, то их объединение тоже является событием.
- II. Если множество A является событием, то его дополнение (до Ω) есть тоже событие.

Множество Ω называют *пространством элементарных событий*. Из аксиом I и II легко следует, что Ω является событием: для доказательства достаточно взять любое событие A и рассмотреть объединение множества A и его дополнения.

Примем следующее соглашение. В тех случаях, когда элементарных событий конечное число (т.е. когда множество Ω состоит из конечного числа элементов), под событиями будем понимать все без исключения подмножества множества Ω (включая, разумеется, и пустое подмножество – оно является дополнением к Ω и потому обязано быть событием). В этом случае аксиомы I и II становятся, конечно, ненужными – они выполняются автоматически.

Условимся операцию объединения множеств обозначать знаком "+" (знаком сложения), операцию пересечения множеств – знаком "." (знаком умножения); наконец, если A – подмножество в Ω , то дополнение к A будем обозначать $\bar{A}$. Тогда аксиомы I и II запишутся следующим образом:

- I. $A_1 \in S, A_2 \in S, \dots \Rightarrow A_1 + A_2 + \dots \in S$;
- II. $A \in S \Rightarrow \bar{A} \in S$.

Укажем одно следствие из аксиом I и II.

Если множества $A_1, A_2, \dots$ суть события, то их пересечение – снова событие. Или, в символической записи:

$$A_1 \in S, A_2 \in S, \dots \Rightarrow A_1 A_2 \dots \in S.$$

Для доказательства воспользуемся соотношением

$$\overline{A_1 A_2 \dots} = \bar{A}_1 + \bar{A}_2 + \dots, \quad (1.32)$$

справедливым, как известно, для любых подмножеств $A_1, A_2, \dots$ фиксированного множества Ω . В силу аксиом II и I множество, записанное в правой части (1.32), есть событие; значит и $\overline{A_1 A_2 \dots}$ – событие. Снова применяя аксиому II, получим, что $A_1 A_2 \dots$ тоже событие.

Введем следующие определения.

Два события A и B , не имеющие (как два подмножества) общих элементов, называются *несовместными*. События A и $\bar{A}$ называются *противоположными*. Если $A \subset B$, то мы говорим, что событие B является *следствием* события A или что A *влечет* за собой B .

Мы завершаем данный пункт следующим определением.

Определение. Совокупность подмножеств S фиксированного множества Ω , удовлетворяющая аксиомам I и II, называется σ – **алгеброй множеств**.

3°. Аксиомы вероятностей. Теперь мы можем сформулировать аксиомы, задающие само понятие вероятности.

1. Каждому событию $A \in S$ поставлено в соответствие неотрицательное число $P(A)$, называемое **вероятностью события A** .

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

3. Если события $A_1, A_2, \dots$ попарно несовместны, то

$$P(A_1 + A_2 + \dots) = P(A_1) + P(A_2) + \dots$$

Заметим, что при бесконечном числе событий $A_1, A_2, \dots$ в правой части написанного равенства 3 стоит сумма ряда.

Очевидно, если множество Ω является конечным, то любая совокупность попарно не пересекающихся подмножеств состоит лишь из конечного числа подмножеств. Отсюда ясно, что для случая конечного Ω аксиома 3 равнозначна такому (в общем случае более слабому) требованию:

3'. $P(A + B) = P(A) + P(B)$, если A и B несовместны.

Чтобы подчеркнуть различие между аксиомами 3 и 3', часто называют аксиому 3' **аксиомой аддитивности**, а 3 – **аксиомой счетной аддитивности**.

Аксиомы 1 – 3 составляют основу всей теории вероятностей. Все теоремы этой теории, включая самые сложные, выводятся из них формально-логическим путем.

Укажем несколько примеров такого вывода.

Исходя из очевидного соотношения между подмножествами

$$A + \bar{A} = \Omega$$

и применяя аксиомы 2 и 3, получаем, что

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

Отсюда, принимая во внимание неотрицательность чисел $P(A)$ и $P(\bar{A})$, находим:

$$P(A) \leq 1$$

– факт, не содержащийся непосредственно ни в одной из аксиом 1 – 3.

Заметим, что доказательство теоремы сложения вероятностей (1.24) было основано на теоретико-множественных соображениях и потому может быть перенесено и на общий случай.

Действительно, рассмотрим очевидные соотношения между событиями (как между подмножествами множества Ω):

$$A = AB + A\bar{B}, \quad A + B = B + A\bar{B}$$

(рис. 1.5). Применяя к обоим равенствам аксиому сложения, получим два числовых равенства:

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

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

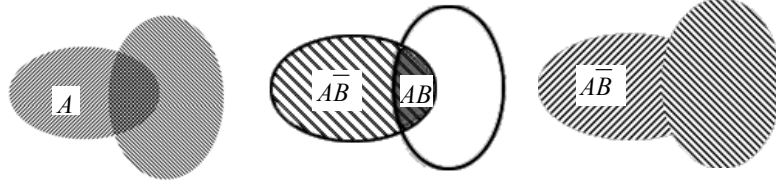


Рис. 1.5

Вычитая из последнего равенства предыдущее, приходим к формуле (1.24).

Докажем теперь две теоремы, устанавливающие как бы непрерывную зависимость $P(A)$ от события A .

Теорема 1.3. Если $A_1, A_2, \dots$ – последовательность событий таких, что каждое следующее является следствием предыдущего ($A_1 \subset A_2 \subset \dots A_n \subset A_{n+1} \dots$), то

$$\lim_{n \rightarrow \infty} P(A_n) = P(A_1 + A_2 + \dots). \quad (1.33)$$

Доказательство. Рассмотрим события:

$$B_1 = A_1, B_2 = \overline{A_1}A_2, B_3 = \overline{A_2}A_3, \dots$$

Нетрудно видеть, что события $B_i, i = 1, 2, 3, \dots$ попарно несовместны и для любого n выполняется равенство $A_n = B_1 + B_2 + \dots + B_n$. Отсюда, с учетом аксиомы 3, получаем

$$\lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} (P(A_1) + P(A_2) + \dots + P(A_n)) = \sum_{k=1}^{\infty} P(B_k) = P\left(\sum_{k=1}^{\infty} B_k\right).$$

Теперь осталось показать, что $\sum_{k=1}^{\infty} B_k = \sum_{k=1}^{\infty} A_k$.

Действительно, если наступает событие $\sum_{k=1}^{\infty} B_k$, то наступает какое-то из B_n . В этом случае наступают все события A_k с номерами $k \geq n$. Поэтому

$$\sum_{k=1}^{\infty} B_k \subset \sum_{k=1}^{\infty} A_k.$$

Если же наступает событие $\sum_{k=1}^{\infty} A_k$, то наступает хотя бы одно из $A_1, A_2, \dots$. Пусть A_n – наступившее событие с минимальным номером n . Поскольку n минимально, то событие A_{n-1} не наступает, следовательно, наступает событие $\overline{A_{n-1}}A_n = B_n$. Поэтому

$$\sum_{k=1}^{\infty} A_k \subset \sum_{k=1}^{\infty} B_k,$$

что завершает доказательство.

Теорема 1.4. Если $A_1, A_2, \dots$ – последовательность событий таких, что каждое предыдущее является следствием последующего ($A_1 \supset A_2 \supset \dots \supset A_n \supset A_{n+1} \supset \dots$), то

$$\lim_{n \rightarrow \infty} P(A_n) = P(A_1 A_2 \dots). \quad (1.34)$$

Д о к а з а т е л ь с т в о . Обозначим $B_i = \overline{A_i}$. Тогда последовательность событий B_i удовлетворяет условиям теоремы 1.3. Имеем

$$\lim_{n \rightarrow \infty} P(B_n) = P(B_1 + B_2 + \dots).$$

Поскольку $B_1 + B_2 + \dots = \overline{A_1} + \overline{A_2} + \dots = \overline{A_1 A_2 \dots}$, то

$$1 - \lim_{n \rightarrow \infty} P(A_n) = 1 - \lim_{n \rightarrow \infty} P(A_1 A_2 \dots),$$

что доказывает равенство (1.34).

4°. Вероятностные пространства. Предмет теории вероятностей. Аксиомы А.Н. Колмогорова дают весьма удобную математическую модель для исследования конкретных теоретико-вероятностных задач, или, по-другому, для описания опытов со случайными исходами. В этой схеме содержится вся информация о данном опыте, представляющая интерес с точки зрения теории вероятностей. Модель включает в себя три объекта:

- 1) множество Ω (называемое пространством элементарных событий или исходов опыта);
- 2) систему S подмножеств множества Ω (называемых событиями), удовлетворяющую аксиомам I, II пункта 2°, или, по-другому, σ – алгебру подмножеств;
- 3) функцию $P(A)$, определенную на алгебре множеств S и удовлетворяющую аксиомам 1, 2, 3 п. 3°.

Совокупность этих трех объектов условимся называть *вероятностной моделью данного опыта* или *вероятностным пространством*. С введением понятия вероятностного пространства появляется возможность определить предмет теории вероятностей в не вызывающих сомнений точных терминах, а именно: **теория вероятностей занимается изучением всевозможных вероятностных пространств**. Естественно, что эту теорию должны интересовать как теоремы, справедливые в *любом* вероятностном пространстве (таких теорем сравнительно немного, примером могут служить доказанные выше теоремы 1 и 2 или, скажем, формула (1.24)), так и теоремы, относящиеся к вероятностным пространствам тех или иных *специальных типов*. Изложенные выше построения свидетельствуют о том, что язык современной теории вероятностей есть язык теории множеств. Между тем прикладные задачи формулируются на другом, "практическом" языке (читатель может для сравнения еще раз прочитать примеры из § 1, где вообще нет упоминания о множествах). Поэтому для придания теоретико-вероятностным рассуждениям полной строгости желателен перевод с одного языка на другой. В дальнейшем при разборе конкретных задач мы, как правило, не будем делать такого перевода. Однако в тех случаях, когда отсутствие должной строгости может затруднить понимание вопроса, мы

полностью опишем вероятностное пространство, отвечающее данной задаче: укажем пространство Ω элементарных событий, систему S подмножеств, играющих роль событий, и определим во всем объеме функцию $P(A)$.

§ 1.7. Некоторые типы вероятностных пространств

В этом параграфе мы рассмотрим основные виды вероятностных пространств, изучаемые в данном учебнике. И начнем мы с классического типа.

Определение. Пусть Ω – конечное пространство элементарных событий, S – алгебра всех подмножеств Ω , называемых (случайными) **событиями**, **вероятностью** события $A \in S$ называется отношение $P(A) = \frac{|A|}{|\Omega|}$. Тройка (Ω, S, P) называется **классическим вероятностным пространством**.

По сути, до сих пор мы в основном занимались примерами и свойствами классических вероятностных пространств. Обобщением классических вероятностных пространств являются *конечные вероятностные пространства*. В этом случае исходы опыта $\omega_i \in \Omega$, вообще говоря, не являются равновероятными, и вероятность события считается как сумма вероятностей всех исходов, из которых оно состоит.

Например допустим, что в опыте с бросаниями игральной кости нас интересует только выпадение шестерки. Соответствующий опыт будет включать только два исхода: шестерка выпала или не выпала. Тогда алгебра множеств состоит всего из четырех событий: 1) невозможное событие $\emptyset$, $P(\emptyset) = 0$; 2) выпала шестерка A , $P(A) = \frac{1}{6}$; 3) шестерка не выпала $\bar{A}$, $P(\bar{A}) = \frac{5}{6}$; 4) достоверное событие $\Omega = A + \bar{A}$, $P(\Omega) = 1$.

Аналогичным образом можно определить *дискретные вероятностные пространства* в случае, когда число исходов опыта счетное.

Определение. Вероятностное пространство (Ω, S, P) называется **дискретным**, если множество элементарных событий Ω конечно или счетно, а S – σ –алгебра всех подмножеств Ω .

Читатель легко проверит, что и в этом случае все аксиомы вероятностного пространства выполняются.

В качестве примера рассмотрим опыт, состоящий в подбрасывании монеты до тех пор, пока она не выпадет одной и той же стороной. Запишем все исходы такого опыта:

$$\{\Gamma\Gamma, \Pi\Pi, \Pi\Gamma\Gamma, \Gamma\Pi\Pi, \Pi\Gamma\Pi\Pi, \Gamma\Pi\Gamma\Gamma, \dots\}.$$

Каждому исходу опыта, заканчивающемуся на k -м шаге, припишем вероятность $\frac{1}{2^k}$. Тогда

$P(\Gamma\Gamma) = P(\Pi\Pi) = \frac{1}{4}$, $P(\Pi\Gamma\Gamma) = P(\Gamma\Pi\Pi) = \frac{1}{8}$, Найдем вероятность того, что потребуется четное число бросаний.

$$P(\Gamma\Gamma + \Pi\Pi + \Pi\Gamma\Pi\Pi + \Gamma\Pi\Gamma\Gamma + \dots) = \frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \dots = \frac{\frac{1}{2}}{1 - \frac{1}{4}} = \frac{2}{3}.$$

Рассмотрим теперь пример не дискретного вероятностного пространства, а именно, положим $\Omega = [0, 1]$ – единичный отрезок числовой прямой. В качестве событий будут выступать открытые интервалы вида $(a, b) \subset [0, 1]$ и все множества, которые можно получить из них с помощью теоретико-множественных операций. Такие множества называются *борелевскими*. В качестве вероятности события $A = (a, b)$ возьмем длину: $P(A) = b - a$. Известен такой факт.

Теорема. На борелевских множествах A на отрезке $[0, 1]$ определена неотрицательная функция $m(A)$, обладающая свойством счетной аддитивности и совпадающая с длиной открытого интервала.

Таким образом, полагая для любого события $A \subset \Omega$

$$P(A) = m(A), \quad (1.35)$$

мы получаем вероятностное пространство, включающее следующие компоненты. Опыт Ω состоит в бросании точки в единичный отрезок, алгебра событий S совпадает с множеством борелевских множеств на отрезке $[0, 1]$, вероятность события A находится по формуле (1.35).

Так определенное вероятностное пространство обладает рядом свойств, отличных от свойств дискретных вероятностных пространств. Например, справедливо

Предложение. Вероятность события, состоящего из счетного множества точек, равна нулю.

Д о к а з а т е л ь с т в о . Докажем, что вероятность события, состоящего из одной точки x , равна нулю. Имеем равенство

$$\{x\} = \bigcap_{n=1}^{\infty} \left(x - \frac{1}{n}, x + \frac{1}{n} \right).$$

По теореме 4 § 1.6 $P(\{x\}) = \lim_{n \rightarrow \infty} P\left(x - \frac{1}{n}, x + \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \frac{2}{n} = 0$. Если событие A состоит из счетного

числа точек, то его можно представить в виде счетной суммы несовместных событий:

$$A = \{x_1\} + \{x_2\} + \dots$$

Воспользуемся аксиомой счетной аддитивности:

$$P(A) = P(\{x_1\}) + P(\{x_2\}) + \dots = 0 + 0 + \dots = 0.$$

Следствие 1. Вероятность любого из событий $A = [a, b]$, $A = (a, b]$, $A = [a, b)$ равна длине (a, b)

$$P(A) = b - a.$$

Например, событие $A = [a, b]$ можно представить в виде суммы несовместных событий $A = (a, b) + \{a\} + \{b\}$ и воспользоваться предложением.

Следствие 2. Вероятность события $A = Q$, состоящего из всех рациональных точек на отрезке $[0, 1]$, равна нулю. Вероятность события $A = I$, состоящего из всех иррациональных точек на отрезке $[0, 1]$, равна единице.

Для множества рациональных точек Q утверждение следует непосредственно из доказанного предложения. Имеем $I = \overline{Q}$, поэтому $P(I) = 1 - P(Q) = 1$.

Смысл следствия 2 состоит в том, что наудачу выбранная точка из единичного отрезка будет наверняка иррациональным числом, поскольку последних чисел гораздо больше, чем рациональных.

Борелевские множества могут быть устроены достаточно сложно и парадоксально. В качестве примера рассмотрим *канторово множество*. Оно получается индуктивным процессом удаления некоторых частей отрезка $[0, 1]$.

А именно, на первом шаге удалим центральную половину отрезка, т.е. интервал $\left(\frac{1}{4}, \frac{3}{4}\right)$. Остаются два отрезка: $\left[0, \frac{1}{4}\right]$ и $\left[\frac{3}{4}, 1\right]$.

На втором шаге из каждого из оставшихся отрезков удалим центральные половины, т.е. открытое множество $\left(\frac{1}{16}, \frac{3}{16}\right) \cup \left(\frac{13}{16}, \frac{15}{16}\right)$. Останутся четыре отрезка $\left[0, \frac{1}{16}\right]$, $\left[\frac{3}{16}, \frac{1}{4}\right]$, $\left[\frac{3}{4}, \frac{13}{16}\right]$ и $\left[\frac{15}{16}, 1\right]$.

Продолжая этот процесс, мы на k -м шаге удалим 2^{k-1} центральные половины, так что останется множество C_k , состоящий из 2^k отрезков, каждый длиной $\frac{1}{4^k}$. По определению, канторово множество

$C = \bigcap_{k=1}^{\infty} C_k$ состоит из точек, которые не были удалены ни на одном из бесконечного числа шагов.

Заметим, что $P(C_k) = 2^k \frac{1}{4^k} = \frac{1}{2^k}$, поэтому $P(C) = \lim_{n \rightarrow \infty} P(C_k) = 0$. С другой стороны, можно показать, что канторово множество содержит несчетное множество точек.

§ 1.8. Условная вероятность

При совместном рассмотрении двух случайных событий A и B часто возникает вопрос: насколько связаны эти события друг с другом, в какой мере наступление одного из них влияет на возможность наступления другого?

Простейшим примером связи между двумя событиями может служить причинная связь – когда наступление одного из событий ведет к обязательному осуществлению другого или же, наоборот, когда наступление одного события исключает шансы другого. Скажем, если событие A заключается в том, что выбранное наугад изделие данного предприятия не содержит брака, а событие B – в том, что изделие является первосортным, то ясно, что наступление B влечет за собой в обязательном порядке наступление A ; напротив, событие $\overline{A}$ исключает событие B .

Однако наряду с такими крайними случаями существует и много промежуточных, когда непосредственная причинная зависимость одного события от другого отсутствует, но *некоторая зависимость* все же имеется. Чтобы пояснить сказанное, приведем такой пример.

Бросается игральная кость. Событие A – выпадение четного числа очков, событие B – выпадение числа очков, большего чем 3. Очевидно, было бы неверно утверждать, что одно из этих событий влечет за собой другое или, наоборот, одно из них исключает другое. В то же время между событиями A и B имеется какая-то зависимость. В самом деле, из трех случаев, к которым сводится B (выпадение 4, 5 или 6 очков), событию A будут благоприятны два; поэтому, если считать наступившим событие B , то шансы события A будут $\frac{2}{3}$. В то же время при отсутствии предварительной информации об исходе бросания шансы события A оцениваются отношением $\frac{3}{6}$. Так как $\frac{2}{3} > \frac{3}{6}$, то следует признать, что наступление события B повышает шансы события A .

Для характеристики зависимости одних событий от других вводится понятие *условной вероятности*.

Определение. Пусть A и B – два случайных события по отношению к некоторому опыту Ω , причем $P(B) \neq 0$. Число $\frac{P(AB)}{P(B)}$ называется **вероятностью события A при условии, что наступило событие B** , или просто **условной вероятностью события A** .

Вероятность A при условии B обозначается $P_B(A) = P(A|B)$. Таким образом, по определению имеем следующее равенство:

$$P_B(A) = P(A|B) = \frac{P(AB)}{P(B)}. \quad (1.44)$$

Смысл числа $P_B(A)$ как вероятности события A при условии, что наступило событие B , можно подтвердить таким рассуждением. Пусть опыт имеет n равновероятных исходов (классическая схема), из которых n_A исходов благоприятны событию A , n_B – событию B и n_{AB} – событию AB . Если известно, что в результате опыта наступило событие B , то число возможных исходов сокращается до n_B , из них ровно n_{AB} исходов будут такими, при которых наступает и событие A . Таким образом, вероятность A при условии наступления B будет, согласно классическому способу подсчета вероятности, равна отношению $\frac{n_{AB}}{n_B}$ или

$$\frac{n_{AB}}{n} \cdot \frac{n_B}{n} = \frac{P(AB)}{P(B)},$$

что делает оправданным определение (1.44).

Из равенства (1.44), являющегося определением условной вероятности, следует:

$$P(AB) = P_B(A)P(B). \quad (1.45)$$

Это означает, что вероятность произведения двух событий равна вероятности одного из этих событий при условии другого, умноженной на вероятность самого условия.

Если $P(A) \neq 0$, то наряду с равенством (1.45) имеет место

$$P(BA) = P_A(B)P(A).$$

Сравнивая эти равенства, мы видим, что

$$P_B(A)P(B) = P_A(B)P(A). \quad (1.46)$$

Перепишем формулу (1.45) (с учетом (1.46)) в виде

$$P(A_1 A_2) = P(A_1)P_{A_1}(A_2). \quad (1.47)$$

Нетрудно видеть, что справедлива аналогичная формула для трех событий A_1, A_2, A_3 :

$$P(A_1 A_2 A_3) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3).$$

Действительно,

$$P(A_1 A_2 A_3) = P(A_1 A_2)P_{A_1 A_2}(A_3) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3).$$

Вообще, для n событий $A_1, A_2, \dots, A_n$ справедлива формула

$$P(A_1 A_2 \dots A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3) \dots P_{A_1 A_2 \dots A_{n-1}}(A_n). \quad (1.48)$$

Пример 1.19. Все грани игральной кости заклеены непрозрачной бумагой: грани 1, 2, 3 – красной, грани – 4, 5, 6 – черной. При бросании кости выпала черная грань. Какова вероятность того, что на этой грани стоит четное число?

Решение. Очевидно, мы должны найти условную вероятность $P_B(A)$, где событие A есть выпадение четного числа очков, а событие B – выпадение числа очков, большего 3. Имеем:

$$P_B(A) = \frac{P(AB)}{P(B)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}.$$

Для сравнения отметим, что безусловная вероятность события A (просто $P(A)$) равна $\frac{1}{2}$.

Пример 1.20. Слово "фото", составленное из букв-кубиков, рассыпано на отдельные буквы, которые затем сложены в коробке. Из коробки наугад извлекаются одна за другой все буквы. Какова вероятность того, что при этом снова появится слово "фото" (событие A)?

Решение. Введем обозначения для событий:

A_1 – первой извлечена буква "ф",

A_2 – второй извлечена буква "о",

A_3 – третьей извлечена буква "т",

A_4 – четвертой извлечена буква "о".

Очевидно, $A = A_1 A_2 A_3 A_4$.

По формуле (1.48) для $n = 4$ имеем

$$P(A) = P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3)P_{A_1A_2A_3}(A_4) = \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{12}.$$

Пример 1.21. В команде из 12 спортсменов 5 мастеров спорта. По жеребьевке из команды выбирают 3 спортсменов. Какова вероятность того, что все выбранные спортсмены являются мастерами спорта?

Решение. Укажем 2 способа решения, из которых первый способ состоит в непосредственном подсчете искомой вероятности по классической схеме, а второй – в применении формулы (1.48).

Первый способ. Представим себе урну, в которой 5 красных шаров и 7 белых. Красные шары соответствуют мастерам спорта, а белые – остальным спортсменам. Из этой урны наудачу извлекают 3 шара, и пусть событие A состоит в появлении 3 красных шаров. Тогда искомая вероятность равна

$$P(A) = \frac{C_5^3}{C_{12}^3} = \frac{1}{22}.$$

Второй способ. Из урны последовательно без возвращения извлекаются 3 шара. Введем обозначения: A_1 – первый шар красный, A_2 – второй шар красный, A_3 – третий шар красный. Тогда $A = A_1A_2A_3$ и по формуле (1.48) при $n = 3$ имеем

$$P(A) = P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3) = \frac{5}{12} \cdot \frac{4}{11} \cdot \frac{3}{10} = \frac{1}{22}.$$

§ 1.9. Независимые события и правило умножения вероятностей

1°. Независимость двух событий. Мы уже видели на примерах, что числа $P(A)$ и $P_B(A)$, вообще говоря, различны; другими словами, наступление события B может изменять вероятность события A . В связи с этим вводится следующее определение.

Определение. Мы говорим, что событие A не зависит от B , если выполняется равенство

$$P_B(A) = P(A) \quad (1.49)$$

Таким образом, A не зависит от B , если наступление B не оказывает влияния на вероятность A (или, говоря проще, наступление B не меняет шансов A).

Понятие независимости событий – одно из центральных в теории вероятностей.

Из равенства

$$P(AB) = P_B(A)P(B) \quad (1.50)$$

(см. (1.45)) мы можем вывести теперь такое заключение:

Если событие A не зависит от B , то справедливо равенство

$$P(AB) = P(A)P(B) \quad (1.51)$$

(вероятность произведения равна произведению вероятностей). Действительно, если в правой части равенства (1.50) заменить $P_B(A)$ множителем $P(A)$, то придем к (1.51).

Указанное выше предложение называют обычно *правилом умножения вероятностей*.

Справедливо и предложение, в известном смысле обратное: если выполняется равенство (1.51), причем $P(B) \neq 0$, то A не зависит от B . Действительно, из (1.51) следует $P(A) = \frac{P(AB)}{P(B)}$, а значит и $P(A) = P_B(A)$.

Согласно определению, данному в начале параграфа, говорить о независимости события A от B имеет смысл лишь при условии $P(B) \neq 0$. Однако в некоторых случаях такое ограничение представляется ненужным. Ввиду этого вводится более широкое понятие независимости, а именно: мы говорим, что событие A не зависит от B , если выполняется равенство (1.51). При этом уже безразлично, будет ли вероятность события B отлична от нуля, или же будет $P(B) = 0$. В случае когда вероятность B отлична от нуля, новое определение эквивалентно первоначальному.

В дальнейшем независимость A от B будет пониматься как выполнение равенства (1.51).

Заметим, что в случае когда $P(B) = 0$, равенство (1.51) выполняется автоматически. Действительно, выполняется равенство: $P(B) = P(BA) + P(B\bar{A})$. Положив $P(B) = 0$, получаем, что сумма двух неотрицательных чисел $P(BA)$ и $P(B\bar{A})$ равна нулю. Следовательно, каждое из них равно нулю в отдельности. Таким образом, если $P(B) = 0$, то $P(AB) = 0$, что влечет (1.51).

Так как $AB = BA$, то из равенства (1.51) следует также равенство $P(BA) = P(B)P(A)$. Это означает, что *если A не зависит от B , то и B не зависит от A* . Иначе говоря, отношение независимости является *симметричным*. Поэтому в дальнейшем мы можем говорить просто о *независимых событиях A и B* .

Пример 1.22. Из колоды игральных карт наугад выбирают одну карту. Пусть событие A заключается в том, что вынутая карта является "тузом", а событие B – в том, что карта красной масти ("бубновая" или "червовая"). Интуитивно ясно, что A не зависит от B (цена карты не зависит от масти). Проверим это подсчетом. Так как

$$P(A) = \frac{4}{36}, \quad P(B) = \frac{18}{36}, \quad P(AB) = \frac{2}{36},$$

то равенство (1.51) выполняется. Следовательно, события A и B независимы.

Нетрудно видеть, что *если события A и B независимы, то независимы также события $\bar{A}$ и B* . Действительно, из (1.51) следует:

$$P(B) - P(AB) = P(B) - P(A)P(B).$$

Если еще учесть, что справедливо равенство

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

то получим: $P(\bar{A}B) = P(B) - P(A)P(B)$.

Отсюда следует:

$$P(\bar{A}B) = P(B)(1 - P(A)) = P(\bar{A})P(B),$$

что означает независимость событий $\bar{A}$ и B .

Дважды применяя доказанное только что предложение, находим: *если независимы события A и B , то независимы и события $\bar{A}$ и $\bar{B}$.*

В практических вопросах для установления независимости одного события от другого редко прибегают к проверке равенства (1.51). Обычно при этом довольствуются интуитивными соображениями. Так, например, если бросают подряд две монеты, то ясно, что выпадение той или другой стороны на одной монете не оказывает никакого влияния на условия бросания другой, и, значит, следующие два события – выпадение герба на одной монете (событие A) и выпадение герба на другой (событие B), являются независимыми.

2°. Независимость n событий, $n > 2$. Дадим такое

Определение. События $A_1, A_2, \dots, A_n$ называются **независимыми**, если вероятность любого из них не меняется при наступлении какого угодно числа событий из остальных.

Другими словами, события $A_1, A_2, \dots, A_n$ независимы, если выполняется следующее условие: каково бы ни было подмножество $\{i_1, i_2, \dots, i_k\}$ множества $\{1, 2, \dots, n\}$, событие A_{i_1} независимо от произведения событий $A_{i_2}, A_{i_3}, \dots, A_{i_k}$.

Как мы уже знаем, независимость A_{i_1} от $A_{i_2} A_{i_3} \dots A_{i_k}$ записывается в виде соотношения

$$P(A_{i_1} A_{i_2} A_{i_3} \dots A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2} A_{i_3} \dots A_{i_k}). \quad (1.52)$$

Если такое же соотношение записать для подмножества $\{i_2, i_3, \dots, i_k\}$, то получим:

$$P(A_{i_2} A_{i_3} \dots A_{i_k}) = P(A_{i_2}) \cdot P(A_{i_3} \dots A_{i_k}). \quad (1.53)$$

Из (1.52) и (1.53), очевидно, следует:

$$P(A_{i_1} A_{i_2} A_{i_3} \dots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \cdot P(A_{i_3} \dots A_{i_k}).$$

Продолжив это рассуждение, придем, в конце концов, к следующему равенству:

$$P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k}). \quad (1.54)$$

Итак, если события $A_1, A_2, \dots, A_n$ независимы, то для любого подмножества $\{i_1, i_2, \dots, i_k\}$ множества $\{1, 2, \dots, n\}$ справедливо равенство (1.54) ("вероятность произведения равна произведению вероятностей").

Это предложение мы будем называть *правилом умножения вероятностей для n событий*.

Из сказанного следует, что выполнение равенств (1.54) есть *необходимое* условие для независимости событий $A_1, A_2, \dots, A_n$. Но это же условие является и *достаточным* (поскольку из равенств (1.54) очевидным образом следуют равенства (1.52)). Отсюда ясно, что определение независимости для n событий можно было бы сформулировать по-другому: события $A_1, A_2, \dots, A_n$ независимы, *если для них справедливы соотношения (1.54).*

Заметим, что при $n = 3$ соотношения (1.54) принимают следующий вид:

$$\left. \begin{aligned} P(A_1 A_2) &= P(A_1)P(A_2) \\ P(A_1 A_3) &= P(A_1)P(A_3) \\ P(A_2 A_3) &= P(A_2)P(A_3) \end{aligned} \right\}, \quad (1.55)$$

$$P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3). \quad (1.56)$$

Можно было бы показать, что из независимости событий $A_1, A_2, \dots, A_n$ вытекает, что любое из этих событий независимо от любой комбинации остальных. Мы рекомендуем читателю рассмотреть какой-либо частный случай этого утверждения: например, доказать, что из независимости событий A_1, A_2, A_3 вытекает независимость A_1 от $A_2 + A_3$. В этом случае задача будет сводиться к тому, чтобы из соотношений (1.55), (1.56) вывести равенство

$$P(A_1 \cdot (A_2 + A_3)) = P(A_1) \cdot P(A_2 + A_3).$$

Закljučая обсуждение вопроса о независимости событий, еще раз подчеркнем, что в большинстве случаев основанием для вывода о независимости служат интуитивные соображения. Обычный ход рассуждений таков: из конкретных условий рассматриваемого опыта делается заключение о независимости тех или иных событий и затем на этом основании пишется равенство (1.51) (или равенства (1.54), если число событий больше двух).

Приведем сначала пример из области техники.

Пример 1.23. Электрическая схема состоит из n последовательно соединенных блоков (рис. 1.6). Надежность (т.е. вероятность безотказной работы) каждого блока равна соответственно $p_1, p_2, \dots, p_n$. Считая выходы из строя различных блоков независимыми событиями, найти надежность всей схемы в целом.

Решение. Событие, заключающееся в исправной работе i -го блока, обозначим A_i ; исправность схемы в целом обозначим A . Так как блоки собраны последовательно, то A имеет место в том и только в том случае, когда имеют место все A_i . Поэтому

$$A = A_1 A_2 \dots A_n,$$

откуда в силу независимости событий $A_1, A_2, \dots, A_n$ следует:

$$P(A) = p_1 p_2 \dots p_n.$$

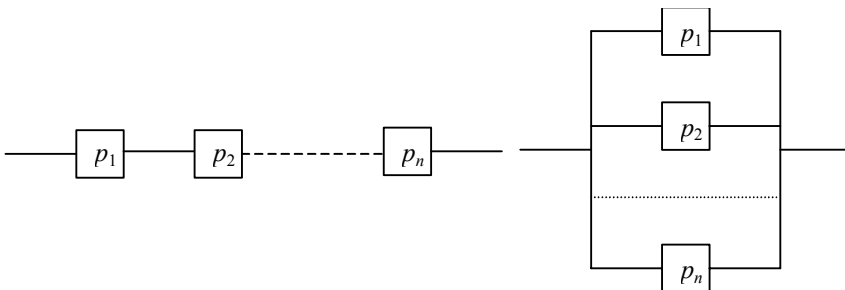


Рис. 1.6

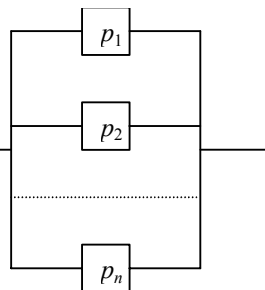


Рис. 1.7

Та же самая задача для схемы из параллельно соединенных блоков (рис. 1.7) приводит к другому ответу. В этом случае выход схемы из строя происходит лишь в том случае, когда выходят из строя все блоки. Это значит, что

$$\overline{A} = \overline{A_1} \overline{A_2} \dots \overline{A_n}$$

и, следовательно,

$$P(\overline{A}) = P(\overline{A_1})P(\overline{A_2}) \dots P(\overline{A_n}) = (1 - p_1)(1 - p_2) \dots (1 - p_n).$$

Таким образом, надежность всей схемы оказывается равной

$$P(A) = 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n).$$

Пример 1.24. Поставим в связи с последним примером такой вопрос. Пусть все блоки имеют одинаковую надежность p . Как велико должно быть число блоков n , чтобы надежность схемы, полученной путем параллельного соединения, превысила 0,99?

Решение. Чтобы получить ответ, необходимо решить неравенство $1 - (1 - p)^n > 0,99$ или $(1 - p)^n < 0,01$. Отсюда

$$n \lg(1 - p) < -2, \text{ т.е. } n > -\frac{2}{\lg(1 - p)}$$

(следует учесть, что $\lg(1 - p)$ есть число отрицательное). Например, при $p = 0,9$ получаем: $n > 2$.

Отвлекаясь на время от примеров, сделаем одно замечание, полезное при решении задач. На практике правило умножения применяют чаще всего *вместе* с правилом сложения. При этом событие A , вероятность которого требуется найти, стараются представить в виде суммы нескольких попарно несовместных слагаемых – *вариантов* события A : $A = A_1 + A_2 + \dots + A_n$, а каждый из вариантов, в свою очередь, – в виде произведения нескольких независимых событий. Тогда последовательное применение правил сложения и умножения позволяет в большинстве случаев найти ответ.

Пример 1.25. Имеются две урны. В первой находятся 1 белый шар, 3 черных и 4 красных, во второй – 3 белых, 2 черных и 3 красных. Из каждой урны наугад извлекают по одному шару, после чего сравнивают их цвета. Найти вероятность того, что цвета вытянутых шаров совпадают (событие A).

Решение. Обозначим событие, состоящее в извлечении из первой урны белого шара, через B_1 , черного – C_1 , красного – D_1 . Аналогичные события для второй урны обозначим B_2 , C_2 , D_2 . Событие A распадается на три варианта: B_1B_2 , C_1C_2 и D_1D_2 . Следовательно, $A = B_1B_2 + C_1C_2 + D_1D_2$. Применяя сначала правило сложения, а затем правило умножения, получим:

$$\begin{aligned} P(A) &= P(B_1)P(B_2) + P(C_1)P(C_2) + P(D_1)P(D_2) = \\ &= \frac{1}{8} \cdot \frac{3}{8} + \frac{3}{8} \cdot \frac{2}{8} + \frac{4}{8} \cdot \frac{3}{8} = \frac{21}{64}. \end{aligned}$$

Пример 1.26. Два игрока поочередно бросают монету. Выигрывает тот, у кого первым выпадает "герб". Какова вероятность выигрыша для игрока, бросающего монету первым?

Решение. Назовем игрока, делающего первый бросок, игроком I, а выигрыш им игры – событием A . Событие A распадается на варианты $\Gamma, \bar{\Gamma}\bar{\Gamma}\Gamma, \bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\Gamma, \dots$, где, например, $\bar{\Gamma}\bar{\Gamma}\Gamma$ означает, что при первом бросании выпал не герб, при втором также не герб, а при третьем – герб. Таким образом,

$$A = (\Gamma) + (\bar{\Gamma}\bar{\Gamma}\Gamma) + (\bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\Gamma) + \dots$$

Так как любые два варианта несовместны, то по правилу сложения имеем

$$P(A) = P(\Gamma) + P(\bar{\Gamma}\bar{\Gamma}\Gamma) + P(\bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\bar{\Gamma}\Gamma) + \dots = \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \dots = \frac{2}{3}.$$

Таким образом, игрок, бросающий первым, имеет вдвое больший шанс выиграть, чем второй игрок.

Пример 1.27. Брак в продукции завода вследствие дефекта A составляет 5%, причем среди забракованной по признаку A продукции 6% имеют дефект B ; в продукции же, свободной от дефекта A , дефект B составляет 2%. Найдите: а) вероятность дефекта B ; б) вероятность наличия дефекта (событие D).

Решение. Очевидно, имеем

$$B = AB + \bar{A}B$$

(дефект B возможен в двух случаях: при наличии A или при отсутствии A). Используя правила сложения и умножения, получим

$$\begin{aligned} P(B) &= P(AB) + P(\bar{A}B) = P(A)P_A(B) + P(\bar{A})P_{\bar{A}}(B) = \\ &= 0,05 \cdot 0,06 + 0,95 \cdot 0,02 = 0,022. \end{aligned}$$

Далее,

$$D = A + \bar{A}B$$

(дефектная деталь имеет либо дефект A , либо свободна от A , но имеет дефект B). Отсюда

$$\begin{aligned} P(D) &= P(A) + P(\bar{A}B) = P(A) + P(\bar{A})P_{\bar{A}}(B) = \\ &= 0,05 + 0,95 \cdot 0,02 = 0,069. \end{aligned}$$

§ 1.10. Формула полной вероятности и формула Байеса

Одним из эффективных методов подсчета вероятностей является *формула полной вероятности*, с помощью которой решается широкий круг задач.

Условимся говорить, что события $H_1, H_2, \dots, H_n$ образуют *полную группу*, если они попарно несовместны и при каждом осуществлении опыта обязательно наступает хотя бы одно из них, иными словами выполняются равенства $H_1 + H_2 + \dots + H_n = \Omega$, $H_i \cap H_j = \emptyset$, $i \neq j$.

Формула полной вероятности. Если события $H_1, H_2, \dots, H_n$ образуют полную группу, то для вероятности любого события A справедлива формула

$$P(A) = P_{H_1}(A)P(H_1) + P_{H_2}(A)P(H_2) + \dots + P_{H_n}(A)P(H_n). \quad (1.57)$$

События $H_1, H_2, \dots, H_n$, образующие полную группу, обычно называют *гипотезами* (по отношению к любому A). В такой интерпретации формула (1.57) читается так: *вероятность события A равна сумме произведений условных вероятностей этого события по каждой из гипотез на вероятности самих гипотез*.

Д о к а з а т е л ь с т в о . Событие A распадается на варианты: AH_1 (A наступает вместе с H_1), $AH_2, \dots, AH_n$. Иначе говоря, имеем

$$A = AH_1 + AH_2 + \dots + AH_n.$$

Так как $H_1, H_2, \dots, H_n$ попарно несовместны, то несовместны и события $AH_1, AH_2, \dots, AH_n$. Применяя правило сложения, находим: $P(A) = P(AH_1) + P(AH_2) + \dots + P(AH_n)$. Заменяв каждое слагаемое $P(AH_i)$ правой части произведением $P_{H_i}(A)P(H_i)$, получаем требуемое равенство (1.57).

Пример 1.28. Имеются три урны. В первой находятся 5 белых и 3 черных шара, во второй – 4 белых и 4 черных, в третьей – 8 белых. Наугад выбирается одна из урн (это может означать, например, что сначала осуществляется выбор шара из вспомогательной урны, где находятся три шара с номерами 1, 2, 3 соответственно) и из нее наугад извлекается шар. Какова вероятность того, что он окажется черным (событие A)?

Р е ш е н и е . Шар может быть вытаскен из первой урны, либо из второй, либо из третьей; обозначим первое из этих событий H_1 , второе – H_2 , третье – H_3 . Так как имеются одинаковые шансы выбрать любую из урн, то $P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}$. Далее находим вероятности события A при каждом из условий H_1, H_2, H_3 :

$$P_{H_1}(A) = \frac{3}{8}, \quad P_{H_2}(A) = \frac{4}{8}, \quad P_{H_3}(A) = \frac{0}{8}.$$

Отсюда

$$\begin{aligned} P(A) &= P_{H_1}(A)P(H_1) + P_{H_2}(A)P(H_2) + P_{H_3}(A)P(H_3) = \\ &= \frac{1}{3} \cdot \frac{3}{8} + \frac{1}{3} \cdot \frac{4}{8} + \frac{1}{3} \cdot \frac{0}{8} = \frac{7}{24}. \end{aligned}$$

Пример 1.29. С первого станка-автомата на сборку поступают 40%, со второго и третьего – по 30% деталей. Среди деталей, выпущенных первым станком, 1% бракованных, вторым – 2%, третьим – 3%. Найти вероятность того, что поступившая на сборку деталь бракованная (событие A)?

Р е ш е н и е . Обозначим через H_i ($i = 1, 2, 3$) событие: деталь поступила с i -го станка. События H_1, H_2, H_3 являются гипотезами. По формуле полной вероятности

$$P(A) = P_{H_1}(A)P(H_1) + P_{H_2}(A)P(H_2) + P_{H_3}(A)P(H_3).$$

В данном случае

$$P(H_1) = 0,40, \quad P(H_2) = 0,30, \quad P(H_3) = 0,30,$$

$$P_{H_1}(A) = 0,01, \quad P_{H_2}(A) = 0,02, \quad P_{H_3}(A) = 0,03,$$

откуда следует

$$P(A) = 0,01 \cdot 0,4 + 0,02 \cdot 0,3 + 0,03 \cdot 0,3 = 0,019.$$

Таким образом, 1,9% всех деталей, поступающих на сборку, бракованные.

В тесной связи с формулой полной вероятности находится так называемая *формула Байеса*. Она относится к той же ситуации, что и формула полной вероятности (событие A может наступить только с одной из гипотез $H_1, H_2, \dots, H_n$). Формула Байеса решает следующую задачу.

Пусть произведен опыт и в результате него наступило событие A . Сам по себе этот факт еще не позволяет сказать, какое из событий $H_1, H_2, \dots, H_n$ имело место в проделанном опыте. Можно, однако, поставить такую задачу: найти вероятности

$$P_A(H_1), P_A(H_2), \dots, P_A(H_n)$$

каждой из гипотез в предположении, что наступило событие A . На этот вопрос и дает ответ формула Байеса. Понятно, что, вообще говоря, $P_A(H_i) \neq P(H_i)$. Так, в примере 1 вероятность гипотезы H_3

(шар извлечен из третьей урны) до того, как произведен опыт, равнялась $\frac{1}{3}$. Однако если опыт произведен и наступило событие A – вытасканный шар оказался черным, то это снижает шансы гипотезы H_3 до нуля. Послеопытная, "апостериорная", вероятность гипотезы H_3 будет в данном случае ниже, чем доопытная, "априорная".

Вывод формулы Байеса весьма прост. Мы имеем:

$$P(AH_i) = P_{H_i}(A)P(H_i),$$

а также

$$P(H_i A) = P_A(H_i)P(A).$$

Приравняв правые части, получим

$$P_A(H_i)P(A) = P_{H_i}(A)P(H_i),$$

откуда следует: $P_A(H_i) = \frac{P_{H_i}(A)P(H_i)}{P(A)}$, или, если воспользоваться формулой полной вероятности,

$$P_A(H_i) = \frac{P_{H_i}(A)P(H_i)}{P_{H_1}(A)P(H_1) + P_{H_2}(A)P(H_2) + \dots + P_{H_n}(A)P(H_n)}. \quad (1.58)$$

Это и есть требуемая формула. Запомнить ее нетрудно: в знаменателе стоит выражение для полной вероятности события A , в числителе – одно из слагаемых в этом выражении.

Рассмотрим два примера на применение формулы Байеса.

Пример 1.30. В некоторой отрасли 30% продукции производится фабрикой I, 25% продукции – фабрикой II, а остальная часть продукции – фабрикой III. На фабрике I в брак идет 1% всей производимой ею продукции, на фабрике II – 1,5%, на фабрике III – 2%.

Купленная покупателем единица продукции оказалась бракованной. Какова вероятность того, что она произведена фабрикой I?

Решение. Введем обозначения для событий: A – купленное изделие оказалось бракованным; H_1 – изделие произведено фабрикой I; H_2 – изделие произведено фабрикой II; H_3 – изделие произведено фабрикой III.

Имеем:

$$P(H_1) = 0,30; P(H_2) = 0,25; P(H_3) = 0,45;$$

$$P_{H_1}(A) = 0,01; P_{H_2}(A) = 0,015; P_{H_3}(A) = 0,02;$$

$$P(A) = 0,01 \cdot 0,30 + 0,015 \cdot 0,25 + 0,02 \cdot 0,45 = 0,015;$$

$$P_A(H_1) = \frac{0,01 \cdot 0,30}{0,015} = 0,20.$$

Таким образом, из всех бракованных изделий отрасли в среднем 20% выпускаются фабрикой I.

В данном примере апостериорная вероятность гипотезы H_1 оказалась ниже априорной (постарайтесь это объяснить).

Пример 1.31.¹ При обследовании больного имеется подозрение на одно из двух заболеваний H_1 и H_2 . Их вероятности в данных условиях: $P(H_1) = 0,6$, $P(H_2) = 0,4$. Для уточнения диагноза назначается анализ, результатом которого является положительная или отрицательная реакция. В случае болезни H_1 вероятность положительной реакции равна 0,9, отрицательной – 0,1; в случае H_2 положительная и отрицательная реакции равновероятны. Анализ произвели дважды, и оба раза реакция оказалась отрицательной (событие A). Требуется найти вероятность каждого заболевания после проделанных анализов.

Решение. В случае заболевания H_1 событие A происходит с вероятностью $0,1 \cdot 0,1 = 0,01$, а в случае заболевания H_2 – с вероятностью $0,5 \cdot 0,5 = 0,25$. Следовательно, по формуле Байеса имеем:

$$P_A(H_1) = \frac{P_{H_1}(A)P(H_1)}{P_{H_1}(A)P(H_1) + P_{H_2}(A)P(H_2)} = \frac{0,01 \cdot 0,6}{0,01 \cdot 0,6 + 0,25 \cdot 0,4} \approx 0,06,$$

$$P_A(H_2) = \frac{0,25 \cdot 0,4}{0,01 \cdot 0,06 + 0,25 \cdot 0,4} \approx 0,94.$$

Отсюда видно, что полученные результаты анализов дают веские основания предполагать болезнь H_2 .

¹ Сюжет этого примера заимствован из популярной книги Б.В. Гнеденко и А.Я. Хинчина «Элементарное введение в теорию вероятностей». М.: Наука, 1970.

§ 2.1. Схема Бернулли. Вероятности $P_n(k)$

В приложениях теории вероятностей часто встречается некоторый тип вероятностного пространства, называемый *схемой повторных независимых испытаний* или *схемой Бернулли*. Настоящая глава посвящена изучению этой схемы и связанных с ней теорем и задач.

Пример 2.1. Трижды бросается игральный кубик. Какова вероятность того, что при этом ровно два раза выпадет максимальное число (т.е. 6) очков?

Решение. В соответствии со сказанным обозначим через S опыт, состоящий в трехкратном бросании игральной кости. Через A обозначим выпадение 6 очков при том или ином бросании. Пространство элементарных событий для опыта S состоит из 2^3 строк

$$(A, A, A), (A, A, \bar{A}), (A, \bar{A}, A), (A, \bar{A}, \bar{A}), \\ (\bar{A}, A, A), (\bar{A}, A, \bar{A}), (\bar{A}, \bar{A}, A), (\bar{A}, \bar{A}, \bar{A}),$$

где, например, событие $(A, A, \bar{A})$ означает, что при первом и втором бросаниях наступило A (выпало 6 очков), а при третьем бросании — $\bar{A}$ (выпало менее 6 очков).

Наша цель — найти вероятность того, что при трех бросаниях дважды наступит A . Очевидно, это событие можно представить в виде $(A, A, \bar{A}) + (A, \bar{A}, A) + (\bar{A}, A, A)$, следовательно, его вероятность равна:

$$\left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} + \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} + \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} = 3 \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} = \frac{5}{72}.$$

Обобщая рассмотренный пример, будем исследовать случай, когда производится n одинаковых и независимых опытов, каждый из которых имеет только два исхода $\{A, \bar{A}\}$. Иными словами, некоторый опыт повторяется n раз, причем в каждом опыте некоторое событие A может появиться с вероятностью $P(A) = p$ или не появиться с вероятностью $P(\bar{A}) = q$, $q = 1 - p$ (в рассмотренном примере мы имели $n = 3$, $p = \frac{1}{6}$, $q = \frac{5}{6}$).

Пространство элементарных событий каждой серии испытаний содержит 2^n точек или последовательностей из символов A и $\bar{A}$. Такое вероятностное пространство и носит название *схемы Бернулли*. Задача же заключается в том, чтобы для данного k найти вероятность того, что при n -кратном повторении опыта событие A наступит k раз.

Для большей наглядности условимся каждое наступление события A рассматривать как *успех*, ненаступление A — как *неудачу*. Наша цель — найти вероятность того, что из n опытов ровно k окажутся успешными; обозначим это событие временно через B .

Событие B представляется в виде суммы ряда событий — вариантов события B . Чтобы фиксировать определенный вариант, нужно указать номера тех опытов, которые оканчиваются успехом. Например, один из возможных вариантов есть

$$\underbrace{(A, A, \dots, A)}_k, \underbrace{(\bar{A}, \bar{A}, \dots, \bar{A})}_{n-k}$$

(успех в 1, 2, ..., k -м опытах и неудача в остальных). Вообще, каждый вариант записывается в виде строки длиной n , в которой k компонент суть A , а остальные $n - k$ компонент — $\bar{A}$. Число всех вариантов равно, очевидно, C_n^k , а вероятность каждого варианта ввиду независимости опытов равна $p^k q^{n-k}$. Отсюда вероятность события B равна $C_n^k p^k q^{n-k}$.

Чтобы подчеркнуть зависимость полученного выражения от n и k , обозначим его $P_n(k)$. Итак,

$$P_n(k) = C_n^k p^k q^{n-k}. \quad (2.1)$$

Полученная формула носит название *формулы Бернулли*, а сами вероятности $P_n(k)$ называются *биномиальными вероятностями*. Такое название связано с тем, что числа $P_n(k)$ имеют непосредственное отношение к формуле бинома Ньютона. Полагая $1 - p = q$ и записав формулу бинома в виде

$$(q + p)^n = q^n + C_n^1 q^{n-1} p + C_n^2 q^{n-2} p^2 + \dots + C_n^{n-2} q^2 p^{n-2} + C_n^{n-1} q p^{n-1} + p^n,$$

можно увидеть, что выражение для $P_n(k)$ совпадает с $(k+1)$ -м членом указанного бинома. Итак,

$$(q + p)^n = P_n(0) + P_n(1) + \dots + P_n(n-1) + P_n(n).$$

Поскольку $q + p = 1$, то отсюда находим:

$$P_n(0) + P_n(1) + \dots + P_n(n-1) + P_n(n) = 1$$

(постарайтесь объяснить это равенство, исходя из теоретико-вероятностных соображений!).

Приведем несколько примеров на применение формулы Бернулли.

Пример 2.2. Монета бросается 10 раз. Какова вероятность того, что герб выпадает при этом ровно 3 раза?

Решение. В данном случае успехом считается выпадение герба, вероятность p этого события в каждом опыте равна $\frac{1}{2}$, так что $q = 1 - p = \frac{1}{2}$. Отсюда

$$P_{10}(3) = C_{10}^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} \cdot \frac{1}{2^{10}} = \frac{15}{128}.$$

Пример 2.3. Контрольная работа состоит из 5 вопросов. На каждый вопрос предлагается 4 варианта ответа, из которых только один правильный. Учащийся не готов к контрольной и поэтому выбирает ответы наугад. Какова вероятность того, что он правильно ответит на k вопросов ($k = 0, 1, 2, 3, 4, 5$)?

Решение. Рассматривая выбор ответа на тот или иной вопрос как отдельное испытание, а получение правильного ответа как некоторое событие A ("успех"), будем иметь: $n = 5$, $p = \frac{1}{4}$, $q = \frac{3}{4}$.

Наша задача – найти вероятности

$$P_5(0), P_5(1), P_5(2), P_5(3), P_5(4), P_5(5).$$

Имеем:

$$P_5(0) = \left(\frac{3}{4}\right)^5 = \frac{243}{1024} \approx 0,24,$$

$$P_5(1) = C_5^1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^4 = \frac{5 \cdot 81}{1024} \approx 0,4,$$

$$P_5(2) = C_5^2 \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = \frac{10 \cdot 27}{1024} \approx 0,26,$$

$$P_5(3) = C_5^3 \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 = \frac{10 \cdot 9}{1024} \approx 0,09,$$

$$P_5(4) = C_5^4 \left(\frac{1}{4}\right)^4 \left(\frac{3}{4}\right)^1 = \frac{5 \cdot 3}{1024} \approx 0,01,$$

$$P_5(5) = \left(\frac{1}{4}\right)^5 = \frac{1}{1024} \approx 0,001.$$

Наибольшей из полученных вероятностей оказалась $P_5(1)$.

Пример 2.4. В урне находятся 6 белых и 9 черных шаров. Из урны извлекают шар, фиксируют его цвет, после чего возвращают шар обратно в урну. Указанный опыт повторяют трижды. Какова вероятность того, что из трех вытащенных при этом шаров ровно два окажутся белыми?

Решение. Поскольку все три опыта производятся в неизменных условиях, мы имеем дело со схемой Бернулли. Вероятность успеха (извлечение белого шара) в каждом из опытов равна $p = \frac{6}{15} = \frac{2}{5}$, $q = \frac{3}{5}$. Следовательно, искомая вероятность

$$P_3(2) = C_3^2 \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right) = 3 \cdot \frac{4}{25} \cdot \frac{3}{5} = \frac{36}{125} \approx 0,288.$$

Если изменить условия предыдущего примера, потребовав, чтобы вытянутый шар не возвращался в урну, то придем к другому ответу. В этом случае уже нельзя будет говорить о неизменности условий всех трех опытов, ибо после каждого извлечения состав шаров в урне будет меняться; как следствие этого будет меняться и вероятность успеха.

Чтобы все же ответить на поставленный вопрос, можно воспользоваться классическим способом подсчета вероятностей. Так как из урны извлекаются три шара, то число возможных исходов будет C_{15}^3 . Из них благоприятными для интересующего нас события будут $C_6^2 \cdot C_9^1$ исходов. Следовательно, искомая вероятность будет:

$$\frac{C_6^2 \cdot 9}{C_{15}^3} = \frac{6 \cdot 5 \cdot 9 \cdot 3}{15 \cdot 14 \cdot 13} = \frac{27}{91} \approx 0,297.$$

Мы видим, что в случае безвозвратной выборки вероятность получилась несколько большей, чем в случае выборки с возвращением.

§ 2.2. Наиболее вероятное число успехов. Среднее число успехов

Если n фиксированно, то $P_n(k)$ превращается в некоторую функцию от аргумента k , принимающего значения $0, 1, 2, \dots, n$. Поставим вопрос, при каком значении аргумента эта функция достигает максимума, проще говоря, какое из чисел $P_n(0), P_n(1), P_n(2), \dots, P_n(n)$ является наибольшим? Итак, мы хотим выяснить, какое число успехов является наиболее вероятным при данном числе опытов n .

Из простых соображений можно предвидеть, что максимум достигается при значении k , близком к числу np . Действительно, поскольку вероятность успеха, или, что то же самое, вероятность события A в одном опыте равна p , то при n -кратном повторении опыта можно ожидать, что частота наступления события A будет близка к p ; следовательно, скорее всего число наступлений события A будет близко к np .

Подтвердим наше предположение расчетом.

Для этой цели рассмотрим два соседних числа $P_n(k)$ и $P_n(k+1)$.

Между ними имеет место одно из трех соотношений:

$$P_n(k) \leq P_n(k+1) \quad (2.2)$$

(меньше, равно или больше) или, что эквивалентно,

$$\frac{P_n(k)}{P_n(k+1)} \leq 1. \quad (2.3)$$

Подставляя вместо числителя и знаменателя их выражения по формулам

$$P_n(k) = C_n^k p^k q^{n-k}, \quad P_n(k+1) = C_n^{k+1} p^{k+1} q^{n-k-1}$$

и учитывая, что $C_n^{k+1} = C_n^k \cdot \frac{n-k}{k+1}$, получим вместо (2.3) соотношения

$$\frac{k+1}{n-k} \cdot \frac{q}{p} \leq 1$$

или

$$(k+1)q \leq (n-k)p.$$

Собирая все слагаемые с множителем k и учитывая, что $p+q=1$, получим эквивалентные (2.3) соотношения

$$k \leq np - q. \quad (2.4)$$

Обозначим число $np + p$ через α . Тогда (2.4) перепишется:

$$k \leq \alpha - 1.$$

Таким образом, для всех значений k меньших чем $\alpha - 1$ справедливо неравенство $P_n(k) < P_n(k+1)$, для $k = \alpha - 1$ (это возможно только в том случае, когда α – целое число) имеет место равенство $P_n(k) = P_n(k+1)$, наконец, при $k > \alpha - 1$ выполняется неравенство $P_n(k) > P_n(k+1)$. Тем самым при значениях k меньших $\alpha - 1$ функция $P_n(k)$ возрастает, а при значениях k больших $\alpha - 1$ убывает. Следовательно, если число α не является целым, то функция имеет единственный максимум; он достигается при ближайшем к α слева целом значении k , т.е. при таком целом k_0 , которое заключено между $\alpha - 1$ и α :

$$np - q < k_0 < np + p,$$

или

$$k_0 = [np + p]. \quad (2.5)$$

Если же α – целое число, то два равных между собой максимума достигаются при $k = \alpha - 1$ и $k = \alpha$. На рисунке 2.1 представлен график функции $P_n(k)$ для двух случаев: $n = 20, p = 0,7$ (тогда $\alpha = 13,7$) и $n = 20, p = 0,1$ ($\alpha = 1,1$). Точки, из которых состоит график, во втором случае для большей наглядности соединены отрезками.

Итак, ответ на вопрос, поставленный в начале параграфа, заключается в следующем. Если число $\alpha = np + p$ не является целым, то наиболее вероятное число успехов равно ближайшему к α слева целому числу. В случае когда α есть целое число, наиболее вероятное число успехов имеет два значения: $\alpha - 1$ и α .

Отметим, что в любом случае наиболее вероятное число успехов отличается от α меньше чем на единицу.

Пример 2.5. Пусть игральную кость бросают 20 раз. Каково наиболее вероятное число выпадений грани "1"?

Решение. В данном случае $n = 20, p = \frac{1}{6}, q = \frac{5}{6}$, откуда

$\alpha = np + p = 20 \cdot \frac{1}{6} + \frac{1}{6} = 3,5$. Поскольку α не является целым числом, то наибольшим среди чисел $P_{20}(0), P_{20}(1), \dots, P_{20}(20)$ будет $P_{20}(3)$. Следовательно, наиболее вероятное число выпадений грани "1" будет 3.

Найдем, чему равна вероятность такого числа выпадений. По формуле Бернулли имеем:

$$P_{20}(3) = C_{20}^3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{17} = 0,238\dots$$

При выяснении наиболее вероятного числа успехов в n опытах мы имели возможность убедиться в особой роли, которую играет в схеме Бернулли число np . Эта роль заключалась в том, что одно из двух ближайших к np целых чисел было наиболее вероятным числом успехов.

Оказывается, число np допускает и другую интерпретацию, притом значительно более важную. А именно: np можно рассматривать в определенном смысле как среднее число успехов в n опытах. Точная формулировка, а также строгое доказательство этого утверждения будут даны позднее.

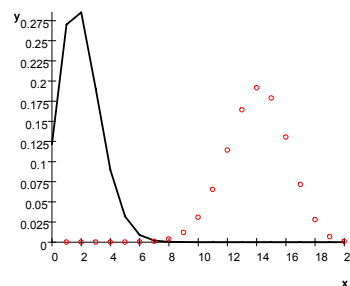
Пример 2.6. В условиях данного предприятия вероятность брака равна 0,05. Чему равно среднее число бракованных изделий на сотню?

Ответ. Искомое число $np = 100 \cdot 0,05 = 5$.

§ 2.3. Вероятности $P_n(k)$ при больших значениях n . Приближенные формулы Лапласа

1°. В приложениях часто возникает необходимость в вычислении вероятностей $P_n(k)$ для весьма больших значений n . Пусть, например, требуется решить такую задачу.

Пример 2.7. На некотором предприятии вероятность брака равна 0,02. Обследуются 500 изделий готовой продукции. Найти вероятность того, что среди них окажется ровно 10 бракованных (не-



трудно видеть, что при указанной выше вероятности брака число 10 есть наиболее вероятное число бракованных изделий из 500).

Решение. Рассматривая обследование каждого изделия как отдельный опыт, можно сказать, что производится 500 независимых опытов, причем в каждом из них событие A (изделие оказалось бракованным) наступает с вероятностью 0,02. По формуле Бернулли

$$P_{500}(10) = C_{500}^{10} (0,02)^{10} (0,98)^{490}.$$

Непосредственный подсчет этого выражения представляет известную сложность из-за громоздкости выражения.

Еще большую трудность нам пришлось бы испытать при попытке решить, например, такую задачу: в условиях последнего примера найти вероятность того, что число бракованных изделий среди 500 окажется в пределах, скажем, от 5 до 15. Поскольку событие $5 \leq k \leq 15$ (где k – число бракованных изделий) равно сумме событий $k=5, k=6, \dots, k=15$, то искомая вероятность равна $P_{500}(5) + P_{500}(6) + \dots + P_{500}(15)$. Подсчет такой суммы является достаточно трудным делом. Между тем задачи, подобные указанной, встречаются в приложениях весьма часто. Поэтому возникает необходимость в отыскании *приближенных формул* для вероятностей $P_n(k)$, а также для сумм вида

$$\sum_{k=k_1}^{k=k_2} P_n(k) \quad (2.6)$$

при больших значениях n . Мы укажем четыре формулы такого рода. Первые две из них мы приведем без доказательства – эти приближенные формулы принадлежат Лапласу и основаны на так называемых *предельных теоремах Лапласа*. Две другие приближенные формулы носят имя Пуассона и основаны на *предельной теореме Пуассона* (см. § 2.5). Название "предельная" в обоих случаях связано с тем, что упомянутые теоремы устанавливают поведение вероятностей $P_n(k)$ (или сумм вида (2.6)) при определенных условиях, в число которых обязательно входит условие $n \rightarrow \infty$.

2°. Приближенные формулы Лапласа. Первая из этих формул дает оценку для вероятности $P_n(k)$ при больших n . Эту формулу называют обычно "локальной".

Локальная приближенная формула Лапласа. При больших n справедливо приближенное равенство

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x), \quad (2.7)$$

где $q = 1 - p$, $x = \frac{k - np}{\sqrt{npq}}$, а $\varphi(x)$ обозначает следующую функцию:

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

Заметим, что для функции $\varphi(x)$, называемой *функцией Гаусса*, составлена таблица ее значений (см. в конце книги).

Свойства функции Гаусса.

1. $\varphi(x)$ – четная функция: $\varphi(-x) = \varphi(x)$.
2. $\varphi(x)$ принимает только положительные значения и имеет единственный максимум при

$$x = 0: \varphi(x) > 0, \max \varphi(0) = \frac{1}{\sqrt{2\pi}}.$$

3. $\varphi(x)$ обладает свойством **нормированности**: $\int_{-\infty}^{\infty} \varphi(x) dx = 1$.

Из этих свойств нуждается в пояснении только последнее. Оно следует из вычисления следующего несобственного интеграла, которое выполняется в курсе математического анализа:

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$$

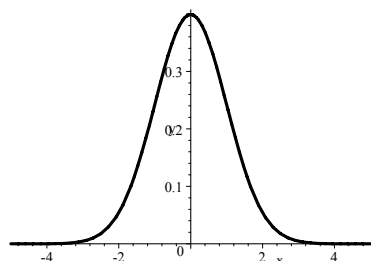


График функции Гаусса приведен ниже.

Обоснованием формулы (2.7) является, как уже отмечалось, одна из так называемых предельных формул Лапласа; рас-

считать ее здесь мы не будем. Таким образом, формулу (2.7) мы примем без доказательства.

Вторая приближенная формула Лапласа тесно связана с первой. Ее обычно называют "интегральной". Эта формула позволяет оценивать не отдельные вероятности $P_n(k)$, а суммы $\sum_{k=k_1}^{k=k_2} P_n(k)$.

Иначе говоря, она дает приближенное выражение для величины $P_n(k_1 \leq k \leq k_2)$ – вероятности того, что число наступлений события A в n опытах (число "успехов") окажется заключенным между заданными границами k_1 и k_2 .

Интегральная приближенная формула Лапласа. При больших n справедливо приближенное равенство

$$P_n(k_1 \leq k \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right), \quad (2.8)$$

где $\Phi(x)$ обозначает следующую функцию:

$$\Phi(x) = \int_0^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt. \quad (2.9)$$

Формулу (2.8) мы также примем без доказательства.

Остановимся на свойствах функции $\Phi(x)$, называемой *функцией Лапласа*.

Свойства функции Лапласа.

1. $\Phi(x)$ – нечетная функция: $\Phi(-x) = -\Phi(x)$.
2. При возрастании аргумента x от 0 до ∞ функция $\Phi(x)$ возрастает от нуля до 0,5.

Для доказательства свойства 1 запишем выражение для $\Phi(-x)$

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} e^{-\frac{t^2}{2}} dt$$

и выполним замену переменной $z = -t$, $dz = -dt$, при этом нижний предел интегрирования не изменится, а верхний станет равным x . Таким образом,

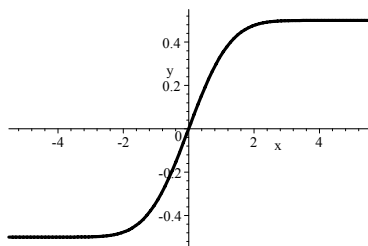
$$\Phi(-x) = -\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz = -\Phi(x).$$

Производная функции $\Phi(x)$ равна функции Гаусса $\varphi(x)$, которая принимает только положительные значения, поэтому функция Лапласа возрастает на всей области определения. Легко видеть, что

$$\Phi(0) = 0, \quad \Phi(\infty) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{2},$$

последнее свойство следует из четности и нормированности функции Гаусса.

График $\Phi(x)$ изображен ниже.



Функция $\Phi(x)$ табулирована; таблица ее значений приведена в конце книги. Из таблицы можно видеть, что уже при $x = 5$ значение $\Phi(x)$ отличается от $\frac{1}{2}$ меньше чем на $3 \cdot 10^{-8}$. По этой причине в таблице указаны значения $\Phi(x)$ лишь для x в пределах от 0 до 4.

По поводу "точности" формул (2.7) и (2.8) мы укажем здесь только следующее: точность существенно зависит от взаимоотношения величин n и p . Более определенно: точность улучшается с ростом произведения npq . Обычно формулами (2.7) и (2.8) пользуются, когда $npq \geq 10$. Отсюда, между прочим, видно, что чем ближе одно из чисел p или q к нулю (другое число в этом случае близко к единице), тем большим следует брать n . Поэтому в случае близости одной из величин p или q к нулю формулами (2.7) и (2.8) обычно не пользуются; для этого случая значительно более точными являются приближенные формулы Пуассона (см. следующий параграф).

Пример 2.8. Монету бросают 100 раз. Какова вероятность того, что при этом герб выпадет ровно 50 раз?

Решение. Имеем: $npq = 100 \cdot \frac{1}{2} \cdot \frac{1}{2} = 25 > 10$. Воспользовавшись приближенной формулой (2.7), получим:

$$P_{100}(50) \approx \frac{1}{\sqrt{25}} \varphi\left(\frac{50-50}{\sqrt{25}}\right) = \frac{1}{5} \varphi(0).$$

Из таблицы для функции $\varphi(x)$ найдем, что $\varphi(0) = 0,399\dots$. Отсюда $P_{100}(50) \approx 0,08$.

Итак, если выполнить опыт, состоящий в 100 бросаниях монеты, то вероятность того, что герб выпадет ровно 50 раз, равна 0,08.

Пример 2.9. Доведем до конца решение задачи из п. 1°, в которой требовалось найти $P_{500}(10)$, а также вероятность события $5 \leq k \leq 15$.

Решение. В данном случае $np = 500 \cdot 0,02 = 10$, $npq = 500 \cdot 0,02 \cdot 0,98 = 9,8$. Воспользовавшись приближенными формулами (2.7) и (2.8), получим:

$$P_{500}(10) \approx \frac{1}{\sqrt{9,8}} \varphi\left(\frac{10-10}{\sqrt{9,8}}\right) = \frac{1}{\sqrt{9,8}} \varphi(0)$$

и

$$P_{500}(5 \leq k \leq 15) \approx \Phi\left(\frac{15-10}{\sqrt{9,8}}\right) - \Phi\left(\frac{5-10}{\sqrt{9,8}}\right) \approx \Phi(1,597) - \Phi(-1,597).$$

Из таблиц для функций $\varphi(x)$ и $\Phi(x)$ находим: $\varphi(0) = 0,399\dots$ и $\Phi(\pm 1,597) = \pm 0,445\dots$. Отсюда

$$P_{500}(10) \approx 0,127, \quad P_{500}(5 \leq k \leq 15) \approx 0,89.$$

Отметим, что точные значения, полученные на компьютере по формуле Бернулли, равны

$$P_{500}(10) = 0,126, \quad P_{500}(5 \leq k \leq 15) = 0,887.$$

Мы видим, что точность формул (2.7) и (2.8) в данном случае вполне удовлетворительна.

Пример 2.10. У страховой компании имеется 12 тыс. клиентов. Каждый из них, страхуясь от несчастного случая, вносит 10 тыс. руб. Вероятность несчастного случая $p = 0,006$, а выплата пострадавшему составляет 1 млн руб. Какая прибыль обеспечивается страховой компанией с вероятностью 0,995?

Иначе говоря, на какую прибыль может рассчитывать страховая компания при уровне риска 0,005?

Решение. Суммарный взнос всех клиентов равен $12\,000 \cdot 10\,000 = 120$ млн руб. Прибыль Π компании зависит от числа k несчастных случаев и определяется равенством

$$\Pi = 120\,000 - 10\,k \text{ тыс. руб.}$$

Наша задача – найти такое число M , чтобы вероятность события $P(k > M)$ не превосходила 0,005; тогда с вероятностью 0,995 будет обеспечена прибыль $120\,000 - 10 \cdot M$ тыс. руб.

Неравенство $P(k > M) \leq 0,005$ равнозначно $P(k \leq M) \geq 0,995$. Учитывая, что $k \geq 0$, можем уточнить последнее неравенство:

$$P(0 \leq k \leq M) \geq 0,995.$$

Для оценки вероятности $P(0 \leq k \leq M)$ воспользуемся интегральной приближенной формулой Лапласа при $n = 12\,000$ и $p = 0,006$, $q = 0,994$ (в данном случае $npq > 10$). Имеем

$$P_{12000}(0 \leq k \leq M) \approx \Phi(x_2) - \Phi(x_1),$$

где $x_1 = \frac{0 - np}{\sqrt{npq}} = -\sqrt{\frac{np}{q}} = -\sqrt{\frac{72}{0,994}} = -8, \dots$, а $x_2 = \frac{M - 72}{\sqrt{72 \cdot 0,994}}$. Так как $x_1 < -8$, то $\Phi(x_1) = 0$. Итак,

необходимо найти M , при котором

$$\Phi\left(\frac{M - 72}{8,5}\right) + 0,5 \geq 0,995,$$

или $\Phi\left(\frac{M-72}{8,5}\right) \geq 0,495$. По таблице значений функции $\Phi(x)$ находим, что $\frac{M-72}{8,5}$ должно быть не меньше чем 2,58:

$$\frac{M-72}{8,5} \geq 2,58,$$

следовательно, $M \geq 72 + 22 = 94$. Итак, с вероятностью 0,995 компании гарантируется прибыль

$$\Pi = 120000 - 1000 \cdot 94 = 26000 \text{ тыс. руб.} = 26 \text{ млн. руб.}$$

3°. Оценка отклонения относительной частоты от вероятности.

Начнем снова с примера.

Пример 2.11. Игральную кость бросают 100 раз. Пусть A обозначает выпадение грани "1" и k – число наступлений события A при 100 бросаниях. Вообще говоря, дробь $\frac{k}{100}$ – относительная частота наступления события A – будет близка к $\frac{1}{6}$ (вероятности события A при одном бросании). Однако сколь тесной окажется эта близость, предугадать невозможно. Поставим следующую задачу: оценить вероятность события

$$\left| \frac{k}{100} - \frac{1}{6} \right| \leq 0,01,$$

т.е. вероятность того, что относительная частота наступления события A в 100 опытах отклонится от вероятности события A не более чем на 0,05.

Отложим на время решение этого примера и поставим задачу в общем виде. В условиях схемы Бернулли с заданными значениями n и p постараемся для данного $\varepsilon > 0$ оценить вероятность события

$$\left| \frac{k}{n} - p \right| \leq \varepsilon, \quad (2.10)$$

где, как и раньше, k – число успехов в n опытах. Неравенство (2.10) эквивалентно $|k - np| < \varepsilon n$, т.е.

$$-\varepsilon n \leq k - np \leq \varepsilon n;$$

или, что то же самое,

$$np - \varepsilon n \leq k \leq np + \varepsilon n.$$

Таким образом, речь идет о получении оценки для вероятности события

$$k_1 \leq k \leq k_2,$$

где $k_1 = np - \varepsilon n$, $k_2 = np + \varepsilon n$.

Применяя приближенную формулу (2.8), получим для данного случая

$$\begin{aligned} P\left(\left|\frac{k}{n} - p\right| \leq \varepsilon\right) &\approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right) = \\ &= \Phi\left(\frac{\varepsilon n}{\sqrt{npq}}\right) - \Phi\left(\frac{-\varepsilon n}{\sqrt{npq}}\right). \end{aligned}$$

С учетом нечетности функции Лапласа приходим к приближенному равенству

$$P\left(\left|\frac{k}{n} - p\right| \leq \varepsilon\right) \approx 2\Phi\left(\varepsilon \sqrt{\frac{n}{pq}}\right). \quad (2.11)$$

Приближенная формула (2.11) дает оценку вероятности того, что относительная частота наступления события A в n опытах отклонится от вероятности события A в одном опыте не больше чем на ε .

Возвращаясь к примеру 2.7 из начала этого параграфа, получим, в частности

$$P\left(\left|\frac{k}{100} - \frac{1}{2}\right| \leq 0,05\right) \approx 2\Phi\left(\frac{0,05 \cdot 100}{\sqrt{100 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right) = 2\Phi(1).$$

С помощью таблицы значений функции Лапласа найдем, что это число равно 0,68... . Иными словами вероятность того, что количество выпадений герба в 100 бросаниях будет колебаться от 45 до 55, равна 0,68.

Заметим, что при фиксированном ε и увеличении числа бросаний n указанная вероятность все время увеличивается. Данную тенденцию можно проследить по следующей таблице.

n	100	200	300	400	500	600	700	800	900	1000
p_n	0,68	0,84	0,92	0,95	0,97	0,985	0,99	0,995	0,997	0,998

Для $n = 1000$ вероятность того, что число выпадений герба будет колебаться в пределах от 450 до 550, равна 0,998. Это означает, что данное событие можно считать практически достоверным.

Заметим, что в равенстве (2.11) предел правой части при $n \rightarrow \infty$ равен 1, поскольку $\lim_{x \rightarrow \infty} \Phi(x) = 0,5$. Таким образом, с ростом n становится сколь угодно достоверным факт, что относи-

тельная частота случайного события $\frac{k}{n}$ отличается от вероятности p меньше чем на ε , где ε – произвольно малая величина.

На данном наблюдении основана *статистическая интерпретация вероятности*, которая представляет собой чисто практический подход к самому понятию вероятности.

§ 2.4. Статистическая интерпретация вероятности

Пусть A – случайное событие, связанное с некоторым опытом. Предположим, что опыт произведен N раз и при этом событие A наступило в N_A случаях. Составим отношение

$$w = \frac{N_A}{N}$$

– оно называется *относительной частотой* наступления события A в рассматриваемой серии опытов.

В соответствии с рассуждениями предыдущего параграфа примем следующее определение.

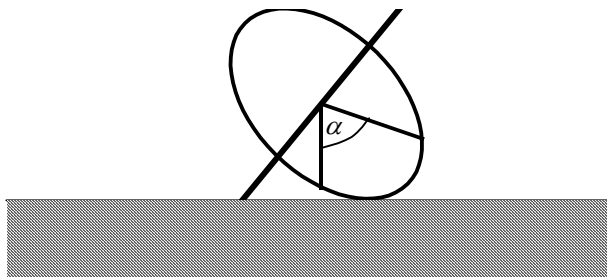
Вероятность случайного события – это связанное с данным событием постоянное число, вокруг которого колеблется относительная частота наступления этого события в длинных сериях опытов.

Заметим, что устойчивость относительной частоты представляет собой одну из простейших закономерностей, проявляющихся в сфере "случайного". Эта закономерность, в конечном счете, составляет основу всех приложений теории вероятностей к практике.

Приведенное выше определение часто называют "статистическим определением" вероятности. Оно не является, конечно, математическим в строгом смысле этого слова, так как невозможно экспериментально или с помощью мысленного опыта найти это число. Из предыдущего параграфа мы знаем, что частота случайного события в N испытаниях Бернулли колеблется около вероятности, и с ростом числа испытаний вероятность фиксированного отклонения стремится к единице. Но сам этот факт опирается на знание вероятности случайного события. Кроме того, в эксперименте надо будет остановиться на некотором конечном шаге, при этом получится только некоторое приближение неизвестной вероятности, которую мы так до конца и не узнаем.

Впрочем, сказанное еще не означает, что статистическое определение является чем-то лишним. Напротив, оно играет чрезвычайно важную роль, но не для самой теории, а для ее приложений. Всякий раз, когда мы найдем вероятности различных событий путем расчетов, практическое истолкование полученных вероятностей будет связано именно со статистическим определением. Другими словами, если найденная путем некоторого расчета (по формулам теории вероятностей) вероятность события A равна числу p , то реальная ценность этого результата состоит прежде всего в возможности такого предсказания: при большом числе опытов частота наступления события A будет близка к p .

Пример 2.12. Волчок приводится во вращение вокруг оси, после чего отпускается. Часть дискообразной поверхности волчка, допустим, некоторый сектор, содержащий угол α (см. рис.), – закрашена. Событие A – касание волчком пола (после остановки) в точке, принадлежащей закрашенной части. Все те же соображения "равноправия" (на этот раз между различными точками окружности) подсказывают нам, что число наступлений события A будет составлять примерно такую же долю от числа всех опытов, какую α составляет от 2π , иначе говоря, частота w будет колебаться около числа $\frac{\alpha}{2\pi}$. Следовательно, $P(A) = \frac{\alpha}{2\pi}$.



Заметим, что последний пример можно свести к схеме извлечения шара из урны. В самом деле, представим себе, что на окружности волчка каким-либо образом выделены промежутками в один градус 360 точек. Часть из них, допустим k точек, приходится на закрашенный сектор. Пренебрегая дугами, меньшими 1° , будем считать, что точка соприкосновения волчка с полом является одной из выделенных точек. Тогда рассматриваемый опыт можно уподобить извлечению одного шара из урны, в которой находятся 360 шаров, причем k из них являются черными. Событие A в такой интерпретации будет означать извлечение черного шара. Вероятность этого события равна $\frac{k}{360}$, т.е. $\frac{\alpha}{2\pi}$.

§ 2.5. Предельная теорема и приближенные формулы Пуассона

Как мы уже отмечали, точность приближенных формул Лапласа понижается по мере приближения одного из чисел, p или q , к нулю. Поэтому случай, когда p или q является малым числом, нуждается в особом исследовании. Будем считать для определенности, что мало p (в случае малого q нужно было бы просто изменить обозначения – вместо A рассмотреть событие $\bar{A}$, при этом числа p и q поменяются ролями). Итак, нас интересует оценка для вероятности того или иного числа успехов, когда сам по себе успех является *редким событием*. Примеры таких событий указать нетрудно: рождение близнецов, достижение столетнего возраста, получение крупного выигрыша в лотерее, авария на городском транспорте, "сбой" в автоматическом соединении, опечатка в книге и т.п.

Разумеется, многие из подобных событий вовсе не являются "успехами"; скорее следовало бы рассматривать их как несчастья. Тем не менее мы сохраним принятую ранее терминологию и каждое наступление интересующего нас события A будем по-прежнему называть "успехом".

Искомую оценку для $P_n(k)$ дает так называемая *предельная теорема Пуассона*. Разъясним ее содержание.

Выражение для $P_n(k)$: $P_n(k) = C_n^k p^k (1-p)^{n-k}$ представляет собой формально функцию трех переменных: n , p и k . Предположим, что k фиксированно, а n и p изменяются. Более точно: пусть $n \rightarrow \infty$, а $p \rightarrow 0$, притом так, что величина $\lambda = np$ остается постоянной ($\lambda = \text{const.}$).

Предельная теорема Пуассона. При указанных выше условиях справедливо равенство

$$\lim_{n \rightarrow \infty} P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (2.12)$$

Доказательство. Имеем:

$$P_n(k) = \frac{n(n-1)\dots(n-(k-1))}{k!} p^k (1-p)^{n-k},$$

и так как $p = \frac{\lambda}{n}$, то

$$P_n(k) = \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-(k-1)}{n} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

Первое из подчеркнутых в правой части выражений представляет собой произведение k множителей, стремящихся к 1; поэтому и все произведение стремится к 1. Третье подчеркнутое выражение

также стремится к 1. Что касается второго, то его можно записать в виде $\left[\left(1 - \frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}$. Замечая, что

выражение в квадратных скобках имеет пределом число $e = \lim_{\alpha \rightarrow 0} (1 + \alpha)^{\frac{1}{\alpha}}$, получим окончательно:

$$P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda} \xi, \text{ где } \xi \rightarrow 1. \text{ Отсюда тотчас следует равенство (2.12).}$$

Из предельной теоремы Пуассона вытекают следующие **приближенные формулы Пуассона**. При больших n и малых p справедливы приближенные равенства

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \quad (2.13)$$

$$\sum_{k=k_1}^{k=k_2} P_n(k) \approx \sum_{k=k_1}^{k=k_2} \frac{\lambda^k}{k!} e^{-\lambda}. \quad (2.14)$$

Исследование вопроса о точности формул (2.13) и (2.14) мы оставляем в стороне. Ограничимся лишь тем, что приведем без доказательства неравенство

$$\left| \sum_{k \in M} P_n(k) - \sum_{k \in M} \frac{\lambda^k}{k!} e^{-\lambda} \right| < np^2,$$

справедливое для любого множества M , принадлежащего $\{0, 1, 2, \dots\}$. В частности, если M состоит из одного числа k , имеем:

$$\left| P_n(k) - \frac{\lambda^k}{k!} e^{-\lambda} \right| < np^2.$$

Обратим внимание читателя на замечательную особенность приближенного равенства (2.13): чтобы найти вероятность того или иного числа успехов, вовсе не требуется знать ни n , ни p . Все определяется, в конечном счете, числом $\lambda = np$, которое является не чем иным (см. конец § 2.2), как **средним числом успехов**.

Для выражения $\frac{\lambda^k}{k!} e^{-\lambda}$, рассматриваемого как функция двух переменных k и λ , составлена таблица значений; она приведена в конце книги.

Пример 2.13. В тесто засыпают некоторое количество изюма, затем всю массу тщательно перемешивают, нарезают на равные доли и выпекают булочки с изюмом. Пусть N — число всех булок, а n — число всех изюмин. Требуется оценить вероятность того, что в случайно выбранной булке окажется ровно k изюмин.

Решение. Рассматривая бросание каждой изюмины в тесто как отдельный опыт, можно утверждать, что производится n опытов. Событие A — попадание изюмины в выбранную нами булку.

Его вероятность p в каждом опыте равна $\frac{1}{N}$ (условие, что тесто с изюмом тщательно перемешивается, означает для каждой изюмины одинаковую вероятность попадания в любую из булок, а число всех булок равно N). Так как обычно булок выпекается много, то можно считать p весьма малым числом. Что касается числа $\lambda = np$, то в данном случае оно есть $\frac{n}{N}$; это среднее число изюмин на одну

булку, его реальное значение практически легко найти. Наша цель заключается в оценке вероятности $P_n(k)$. Так как налицо все условия, при которых справедливо приближенное равенство (2.13), то можем записать ответ в виде

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

Например, для $\lambda = 8$ находим:

$$\begin{aligned} P_n(0) &\approx 0,000, & P_n(5) &\approx 0,092, & P_n(10) &\approx 0,099, \\ P_n(1) &\approx 0,003, & P_n(6) &\approx 0,122, & P_n(11) &\approx 0,072, \\ P_n(2) &\approx 0,011, & P_n(7) &\approx 0,139, & P_n(12) &\approx 0,048, \\ P_n(3) &\approx 0,029, & P_n(8) &\approx 0,139, & P_n(13) &\approx 0,030, \\ P_n(4) &\approx 0,057, & P_n(9) &\approx 0,124, & P_n(14) &\approx 0,017 \end{aligned}$$

и $P_n(k) < 0,001$ при $k > 14$.

Это означает, что из общего числа N булок приблизительно 0,3% содержат по одной изюмине, 1,1% содержат по две изюмины и т.д.

Как видно из решения данного примера, фактическое значение количества булок и числа изюмин не требуется для нахождения искомых вероятностей. Единственным числом, от которого зависит ответ, оказывается $\lambda = \frac{n}{N}$ – среднее число изюмин на булку. Этот факт находится в полном согласии с общим замечанием, сделанным ранее к формуле (2.13).

Разобранный пример, несмотря на его специфическое содержание, носит общий характер. Для иллюстрации приведем еще два подобных примера.

Пример 2.14. В книге из 500 страниц 20 опечаток. Какова вероятность того, что на случайно выбранной странице окажется ровно k опечаток?

(Проводя сравнение с примером 2.13, можно уподобить весь объем книги тесту, а отдельные страницы – булкам. Опечатки играют роль "изюмин").

Пример 2.15. В диспетчерский пункт поступают в среднем 3 заявки на такси в минуту. Найти вероятность того, что за определенные 2 минуты поступит ровно 5 заявок.

(Роль теста играет весь промежуток $[T_1, T_2]$ работы пункта, роль булок – последовательные 2-минутные промежутки. "Изюминами" являются поступающие заявки. Среднее число λ изюмин на 1 булку, как следует из условия, равно 5).

§ 3.1. Описательный подход к понятию случайной величины

Почти в каждом из примеров, с которыми мы встречались в предыдущих главах, дело обстояло таким образом, что в результате опыта возникало некоторое *число*. Например, при бросании игральной кости выпадало то или иное число очков, при обследовании партии готовых изделий обнаруживалось то или иное число единиц брака. Следует сказать, что такое положение типично для теории вероятностей. Среди решаемых ею задач исключительно много таких, в которых исход опыта выражается некоторым числом X . При этом случайный характер исхода влечет за собой случайность числа X ; это означает, что при повторении опыта оно меняется непредвиденным образом. Приведем несколько примеров.

1. Бросается игральная кость; X – выпавшее число очков.
2. Покупается n лотерейных билетов; X – число выигрышей.
3. Из данной аудитории выбирается наугад один студент; X – его рост (скажем, в сантиметрах).
4. Проводится наблюдение над количеством осадков, выпадающих на данную местность в неделю; X – суммарный слой осадков (в сантиметрах).

Чтобы все примеры подобного рода уложить в единую схему, введем понятие *случайной величины*. Заметим, что с каждым из рассмотренных примеров можно связать некоторое вероятностное пространство, состоящее из пространства элементарных событий Ω , алгебры случайных событий S , функции вероятности, определенной на S .

Определение. *Случайной величиной, связанной с данным вероятностным пространством (Ω, S, P) , называется действительная функция X , определенная на пространстве элементарных событий Ω , такая, что для любого действительного числа x множество элементарных событий, для которых выполняется неравенство $X < x$, является событием.*

Более коротко это можно записать следующим образом:

$$(X < x) \in S. \quad (3.1)$$

Можно показать, что данное условие эквивалентно следующему.

Для любого открытого множества U на действительной прямой множество элементарных событий, для которых $X \in U$, является событием.

Иными словами, условие (3.1) эквивалентно следующему

$$(X \in U) \in S. \quad (3.2)$$

Определение. *Функцией распределения случайной величины X называется функция, определенная равенством*

$$F(x) = P(X < x). \quad (3.3)$$

Случайные величины мы будем обозначать в дальнейшем большими латинскими буквами: X , Y и т.д.

Данное выше определение тривиально выполняется для *любой функции*, определенной на пространстве элементарных событий *дискретного* вероятностного пространства. Более сложные случаи вероятностных пространств будут рассмотрены в дополнении. Если исходное вероятностное пространство дискретно, то случайная величина X называется *дискретной случайной величиной*. В этой главе мы будем рассматривать только такие величины.

Каждой случайной величине X соответствует некоторое множество чисел. Это – множество значений, которые может принимать величина X . Так, в первом из наших примеров множество значений состоит из шести чисел: 1, 2, 3, 4, 5, 6, во втором – из чисел 0, 1, 2, ..., n .

Различные случайные величины могут иметь одно и то же множество возможных значений. Чтобы проиллюстрировать это примером, представим себе, что имеются две игральные кости, причем одна сделана из однородного материала, а другая, скажем, склеена из двух кусков разной плотности. Обозначим через X число очков, выпадающих на первой кости, через Y – число очков на второй. Случайные величины X и Y имеют одно и то же множество возможных значений, а именно {1, 2, 3, 4, 5, 6}, однако ведут себя совершенно по-разному. Действительно, вероятности любого исхода для

первого опыта равны $\frac{1}{6}$; для второй кости вероятности событий $Y = 1, Y = 2, \dots, Y = 6$ будут совсем другими. Этот пример показывает, что знания одного лишь множества возможных значений недостаточно для полного описания случайной величины. Необходимо еще знать *вероятности значений*, которые принимает случайная величина. Любое правило, устанавливающее связь между возможными значениями случайной величины и их вероятностями, называется *законом распределения* случайной величины.

В практических приложениях случайных величин пространство Ω элементарных событий остается как бы "за кадром" – как правило, не очень просто дать точное описание совокупности факторов, составляющих реальное описание элементарного события ω . Например, если случайная величина X – это стоимость акции данного предприятия в тот или иной момент, то элементарное событие представляет собой совокупность огромного числа факторов: состояние рынка, положение данного предприятия, привлекательность его продукции на данный момент, конкурентоспособность и т.п. Тем не менее, несмотря на трудную обозримость множества таких факторов, мы считаем возможным рассматривать стоимость акции как случайную величину в указанном выше смысле, определяя вероятность события $X < x$ по ряду косвенных признаков (показаниям биржи).

§ 3.2. Свойства функции распределения

Напомним, что функцией распределения случайной величины X называется функция $F(x) = P(X < x)$.

Отметим прежде всего следующий факт: зная функцию $F(x)$, можно найти вероятность любого события вида $x_1 \leq X < x_2$. Действительно, воспользуемся очевидным соотношением между событиями:

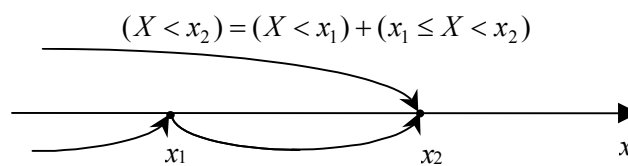


Рис. 3.1

(см. рис. 3.1: луч, расположенный левее точки x_2 , есть объединение луча, расположенного левее x_1 , и полуинтервала $[x_1, x_2)$). Если к этому соотношению применить правило сложения вероятностей, то получим:

$$F(x_2) = F(x_1) + P(x_1 \leq X < x_2).$$

Следовательно,

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1). \quad (3.4)$$

Формула (3.4) в дальнейших рассуждениях будет играть важную роль.

Установим теперь некоторые свойства функции распределения. В последующих рассуждениях будем ссылаться на теоремы 1.3 и 1.4 из § 6 главы 1, которые для краткости будем называть просто теоремой 3 и теоремой 4.

Функция $F(x)$ *неубывающая*, т.е. $F(x_2) \geq F(x_1)$, если $x_2 > x_1$. Это немедленно следует из формулы (3.4), если учесть, что величина, стоящая в левой части, неотрицательна (как вероятность некоторого события).

1. Справедливы равенства:

$$\lim_{x \rightarrow \infty} F(x) = 1, \quad \lim_{x \rightarrow -\infty} F(x) = 0. \quad (3.5)$$

Чтобы доказать первое из этих равенств, возьмем любую возрастающую последовательность чисел $x_1, x_2, \dots$, для которой $\lim_{n \rightarrow \infty} x_n = \infty$, и рассмотрим последовательность событий:

$$A_1 : X < x_1, \quad A_2 : X < x_2, \quad A_3 : X < x_3, \dots$$

Эта последовательность событий, из которых каждое последующее является следствием предыдущего. Сумма событий A_n есть достоверное событие $X < \infty$. Согласно теореме 3

$$\lim_{n \rightarrow \infty} P(A_n) = P(X < \infty) = 1.$$

Следовательно,

$$\lim_{n \rightarrow \infty} F(x_n) = 1,$$

что доказывает первое из равенств (3.5). Чтобы доказать второе, следует к событиям $A_1, A_2, \dots$ применить теорему 4.

Как известно, для неубывающей (вообще для монотонной) функции $F(x)$ при любом значении аргумента x существуют оба односторонних предела:

$$F(x-0) = \lim_{\alpha \rightarrow 0, \alpha > 0} F(x-\alpha),$$

$$F(x+0) = \lim_{\alpha \rightarrow 0, \alpha > 0} F(x+\alpha).$$

Оказывается, что в случае функции распределения первый из этих пределов совпадает со значением функции в самой точке x . А именно справедливо следующее свойство:

$$2. \quad F(x-0) = F(x) \text{ при любом } x. \quad (3.6)$$

Как говорят, в этом случае функция $F(x)$ непрерывна слева.

Для доказательства выберем какую-либо возрастающую последовательность чисел $x_1, x_2, \dots$, сходящуюся к некоторому значению x_0 . Если снова применить теорему 3 к последовательности событий

$$A_1 : X < x_1, A_2 : X < x_2, A_3 : X < x_3, \dots,$$

то получим $\lim_{n \rightarrow \infty} P(A_n) = P(X < x_0)$ или

$$\lim_{n \rightarrow \infty} F(x_n) = F(x_0). \quad (3.7)$$

Таким образом, каково бы ни было число x_0 , для любой возрастающей и сходящейся к x_0 последовательности $x_1, x_2, \dots$ справедливо равенство (3.7). Как известно из курса математического анализа, отсюда следует: $F(x-0) = F(x)$, т.е. (3.6).

Представляет интерес вопрос о *разрывах* функции $F(x)$. Точки разрыва соответствуют таким значениям x , для которых разность

$$F(x+0) - F(x-0) \quad (3.8)$$

отлична от нуля. Напомним, что в случае монотонной функции $F(x)$ величина (3.8) называется *скачком* функции в точке x . Мы докажем сейчас, что при любом x справедлива формула

$$F(x+0) - F(x-0) = P(X = x), \quad (3.9)$$

т.е. скачок функции распределения в точке x совпадает с вероятностью события $X = x$ (рис. 3.2).

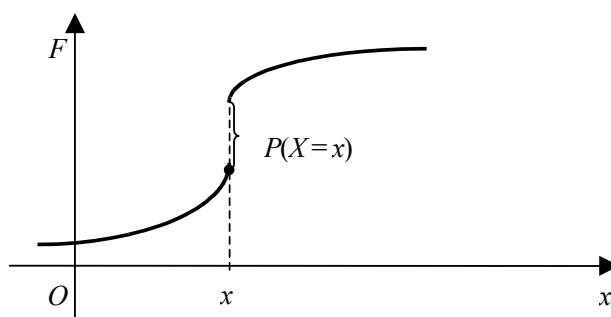


Рис. 3.2

Чтобы вывести формулу (3.9), рассмотрим какую-нибудь возрастающую последовательность $x_1, x_2, \dots$, сходящуюся к некоторой точке x_0 , и убывающую последовательность $y_1, y_2, \dots$, сходящуюся тоже к x_0 (рис. 3.3).

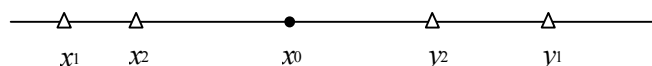


Рис. 3.3

Произведение событий

$$A_1 : X \in [x_1, y_1), A_2 : X \in [x_2, y_2), A_3 : X \in [x_3, y_3), \dots$$

есть, очевидно, $X = x_0$. Отсюда по теореме 4 имеем

$$\lim_{n \rightarrow \infty} P(A_n) = P(X = x_0)$$

или

$$\lim (F(y_n) - F(x_n)) = P(X = x_0), \quad (3.10)$$

но, как мы уже знаем, $\lim_{n \rightarrow \infty} F(x_n) = F(x_0)$. Поэтому из (3.10) следует

$$F(x+0) - F(x) = P(X = x),$$

что и требовалось получить.

Подводя итог, перечислим установленные выше свойства функции распределения $F(x)$:

1. $F(x)$ – неубывающая функция.
2. $F(-\infty) = 0, F(+\infty) = 1$.
3. $F(x)$ непрерывна слева в любой точке.

На рис. 3.4 изображен график одной из возможных функций $F(x)$.

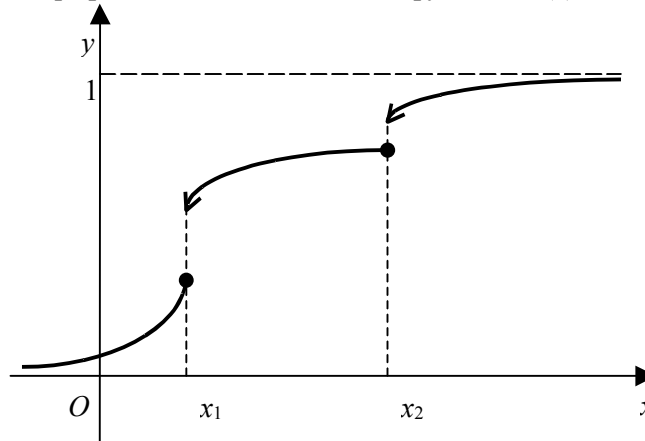


Рис. 3.4

Скачки функции равны соответственно $P(X = x_1), P(X = x_2)$. Для любой точки x , отличной от точек x_1, x_2 , т.е. точки x , в которой $F(x)$ непрерывна, имеем $P(X = x) = 0$.

§ 3.3. Дискретные случайные величины

Наиболее удобными для изучения являются так называемые *дискретные* случайные величины. Они характеризуются тем, что могут принимать лишь конечное или счетное множество значений.

Примеры дискретных величин:

1. X – число выигрышей на купленные n лотерейных билетов.
2. X – число вызовов, поступающих на телефонную станцию в течение часа.
3. X – число вкладов, внесенных в банк в течение дня.

И так далее. Впрочем, любую случайную величину X можно превратить в дискретную, если условиться принимаемые значения величины X округлять до ближайших (скажем, слева) целых значений. При таком подходе можно считать, что значения величины X – целые числа, а значит, X – дискретная случайная величина.

Для дискретной случайной величины X закон распределения может быть задан в виде таблицы:

X	x_1	x_2	...
P	p_1	p_2	...

(3.11)

где $x_1, x_2, \dots$ – возможные различные значения величины X , а $p_1, p_2, \dots$ – их вероятности. А именно p_1 есть вероятность x_1 (вероятность события $X = x_1$), p_2 есть вероятность значения x_2 и т.д. Числа $p_1, p_2, \dots$ связаны соотношением

$$p_1 + p_2 + \dots = 1,$$

которое вытекает из того факта, что события

$$X = x_1, X = x_2, \dots \quad (3.12)$$

попарно несовместны, а их сумма есть событие достоверное (при каждом осуществлении опыта величина X принимает одно и только одно из своих значений, т.е. наступает одно и только одно из событий (3.12)).

Выясним теперь, какой вид имеет функция распределения дискретной случайной величины X , характеризуемой таблицей (3.11).

Пусть x_0 – любое число. Среди чисел $x_1, x_2, \dots$ выделим те, которые меньше x_0 . Пусть ими будут $x_{i_1}, x_{i_2}, \dots$. Событие $X < x_0$ является суммой событий $X = x_{i_1}, X = x_{i_2}, \dots$, поэтому его вероятность равна $p_{i_1} + p_{i_2} + \dots$.

Итак,

$$F(x) = \sum_{\{i | x_i < x\}} p_i. \quad (3.13)$$

Формула (3.13) дает полную информацию о функции $F(x)$. На рисунке 3.5 изображен график этой функции для частного случая, когда X принимает только три значения: x_1, x_2, x_3 . Можно при этом считать $x_1 < x_2 < x_3$. График представляет собой ступенчатую ломаную со скачками в точках x_1, x_2, x_3 . Величины скачков равны соответственно p_1, p_2, p_3 . Левее x_1 график совпадает с осью Ox , правее x_3 – с прямой $y=1$. Аналогичная ступенчатая ломаная будет для любой дискретной случайной величины X .

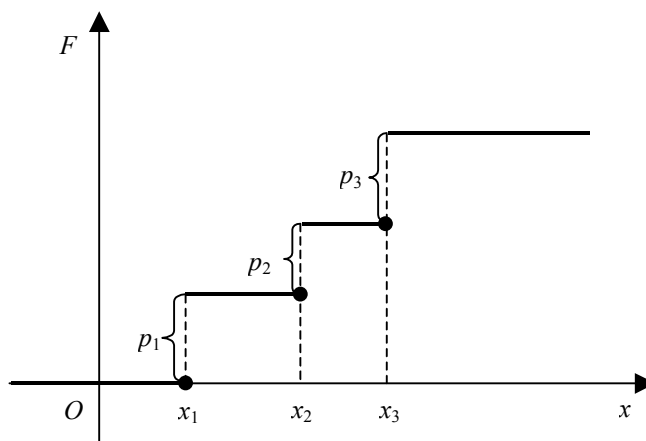


Рис. 3.5

Рассмотрим примеры дискретных случайных величин.

Пример 3.1. По мишени стреляют один раз с вероятностью попадания p . Случайная величина X – число попаданий. Очевидно, X может принимать только два значения: 1 и 0, причем их вероятности равны соответственно p и $q = 1 - p$. Действительно, при выстреле возможны два исхода: попадание (тогда $X = 1$) и промах ($X = 0$); вероятности этих событий суть p и q . В итоге получаем следующую таблицу

X	1	0
P	p	q

Пример 3.2. Дважды бросается игральная кость. Случайная величина X – сумма очков при обоих бросаниях. Возможные значения величины X суть числа 2, 3, ..., 12. Вероятности этих значений легко подсчитываются. Например, $P(X=10) = \frac{3}{36}$, так как из тридцати шести возможных исходов опыта событию $X=10$ благоприятны три. Найдя все вероятности, получим следующую таблицу:

X	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Пример 3.3. Монету бросают 5 раз. Случайная величина X – число выпадений герба. Возможные значения случайной величины X суть 0, 1, 2, 3, 4, 5. Их вероятности подсчитываются с помощью формулы Бернулли, например:

$$P(X=3) = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32}.$$

Производя все подсчеты, получим таблицу:

X	0	1	2	3	4	5
P	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

(3.14)

§ 3.4. Примеры типичных законов распределения дискретных величин

Пример 3.4. Геометрическое распределение. Производится ряд независимых опытов, в каждом из которых с одной и той же вероятностью наступает событие A . Опыты продолжаются до первого появления события A , после чего прекращаются. Рассматривается случайная величина X – число произведенных опытов. Составить для нее закон распределения.

Решение. Возможные значения величины X суть 1, 2, 3, Событие $X = n$ (n – любое натуральное) означает, что в первых $n - 1$ опытах событие A не наступает, а в n -м опыте наступает. Вероятность такого исхода равна:

$$\underbrace{q \cdot q \cdot \dots \cdot q}_{n-1} \cdot p = pq^{n-1},$$

где $q = 1 - p$. Следовательно, закон распределения величины X будет:

X	1	2	3	...	n	...
P	p	pq	pq^2	...	pq^{n-1}	...

Пример 3.5. Биномиальное распределение. Пусть производится определенное число n независимых опытов. В каждом из них с одной и той же вероятностью p может наступить некоторое событие A . Рассматриваемая случайная величина X – число наступлений событий A в n опытах. Соответствующая таблица имеет вид:

X	0	1	2	...	$n-1$	n
P	$P_n(0)$	$P_n(1)$	$P_n(2)$	...	$P_n(n-1)$	$P_n(n)$

(3.15)

где $P_n(k) = C_n^k p^k (1-p)^{n-k}$ ($k = 0, 1, 2, \dots, n$). Это непосредственно следует из формулы Бернулли.

Очевидно, таблица (3.14) есть частный случай таблицы (3.15). Этот частный случай соответствует значениям $n = 5$, $p = \frac{1}{2}$.

Закон распределения, характеризующийся таблицей (3.15), называют **биномиальным**. Такое название связано с уже известным читателю фактом: числа $P_n(0)$, $P_n(1)$, ..., $P_n(n)$ являются членами бинома $(q + p)^n$, где $q = 1 - p$.

Пример 3.6. Распределение Пуассона. Мы говорим, что случайная величина X распределена **по закону Пуассона**, если соответствующая таблица имеет вид:

X	0	1	2	...
P	p_0	p_1	p_2	...

(3.16)

где $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$ ($k = 0, 1, 2, \dots$). Здесь λ – фиксированное положительное число (разным значениям λ отвечают разные распределения Пуассона).

Легко проверить, что для написанной таблицы выполнено обязательное условие – сумма вероятностей всех возможных значений равна 1. Действительно,

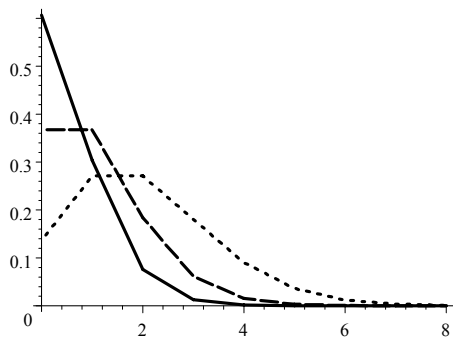


Рис. 3.6

На рисунке 3.6 показаны графики функции $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$ (как функции от k) для значений параметра $\lambda = 0,5$ (сплошная линия), 1 (пунктир) и 2 (штрих-пунктир). Каждый график представляет собой дискретный ряд точек; для большей наглядности точки соединены последовательно ломаной линией (так называемый *многоугольник распределения*).

Одна из причин, обуславливающих важную роль пуассоновского распределения для практики, заключается в его тесной связи с биномиальным распределением. Напомним (§ 2.5), что если в формуле Бернулли

$$P_n(k) = C_n^k p^k (1-p)^{n-k}$$

мы зафиксируем значение k и станем устремлять число опытов n к бесконечности, а вероятность p – к нулю, притом так, чтобы их произведение оставалось равным постоянному числу λ ($np = \lambda$), то будем иметь:

$$\lim P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (3.17)$$

Соотношение (3.17) показывает, что при описанном выше предельном переходе таблица (3.15) биномиального распределения переходит в таблицу (3.16) распределения Пуассона. Таким образом, распределение Пуассона является *предельным* для биномиального распределения при указанных выше условиях. Заметим, что с этим свойством распределения Пуассона – выражать биномиальное распределение при большом числе опытов и малой вероятности события – связано часто применяемое для него название: *закон редких явлений*.

§ 3.5. Системы дискретных случайных величин

До сих пор мы рассматривали случайные величины изолированно друг от друга, не касаясь вопроса об их взаимоотношениях. Однако в практических задачах часто встречаются ситуации, когда те или иные случайные величины приходится изучать *совместно*. В таких случаях говорят о *системе* нескольких случайных величин. Более точно: случайные величины образуют систему, если они определены на одном и том же пространстве элементарных событий Ω .

Систему двух случайных величин (X, Y) можно истолковывать как случайную точку на плоскости, систему трех случайных величин (X, Y, Z) – как случайную точку в трехмерном пространстве. Мы ограничимся в основном двумерным случаем.

Интуитивный подход к понятию системы двух случайных величин связан с представлением об опыте, результатом которого является пара чисел X, Y . Поскольку исход опыта мыслится как случайное событие, то предсказать заранее значения чисел X и Y невозможно (при повторении опыта они меняются непредвиденным образом). Приведем несколько примеров.

Пример 3.7. Дважды бросается игральная кость. Обозначим через X число очков при первом бросании, через Y – число очков во втором. Пара (X, Y) будет системой двух случайных величин.

Пример 3.8. Из некоторой аудитории наугад выбирается один студент; X – его рост (скажем, в сантиметрах), Y – вес (в килограммах).

Пример 3.9. В данном сельскохозяйственном районе выбирается произвольно участок посева пшеницы площадью 1 га; X – количество внесенных на этом участке удобрений, Y – урожай, полученный с участка.

Пример 3.10. Сравниваются письменные работы по математике и русскому языку; X – оценка за работу по математике, Y – за работу по русскому языку.

Список подобных примеров легко продолжить.

§ 3.6. Независимые дискретные случайные величины

1°. Общие замечания. Примеры. При рассмотрении системы двух случайных величин (X, Y) необходимо иметь в виду, что свойства системы не всегда исчерпываются свойствами самих величин X и Y . Иначе говоря, если мы знаем *все* о величине X и *все* о величине Y , то это еще не значит, что мы знаем *все* о системе (X, Y) . Дело в том, что между величинами X и Y может существовать зависимость, и без учета этой зависимости нельзя построить закон распределения системы (X, Y) .

Зависимость между случайными величинами в реальных условиях может быть различной. В некоторых случаях она оказывается столь сильной, что, зная, какое значение приняла величина X , можно в точности указать значение Y . Применяя традиционную терминологию, можно сказать, что в этих случаях зависимость между X и Y *функциональная* (впрочем, понятие функции от случайной величины еще нуждается в уточнениях, последние будут даны в § 3.7). С примерами такой зависимости мы постоянно встречаемся в природе и технике.

В то же время можно указать и примеры другого рода – когда зависимость между случайными величинами существует, но не носит строго выраженного функционального характера. Подобные примеры особенно характерны для таких областей науки и практики, как агротехника, биология, медицина, экономика и т. д., где развитие явлений, как правило, зависит от многих трудно поддающихся учету факторов. Известно, например, что обилие осадков в период созревания пшеницы приводит к повышению урожайности; однако это еще не означает, что связь между количеством X осадков и урожайностью Y (скажем, в расчете на 1 га) является функциональной; кроме осадков на урожайность оказывают влияние и другие факторы: тип почвы, количество внесенных удобрений, число солнечных дней и т. д. В подобных случаях, когда изменение одной величины влияет на другую лишь статистически, *в среднем*, принято говорить о *вероятностной связи* между величинами. Не приводя пока точных определений, рассмотрим несколько примеров. Они иллюстрируют разные степени зависимости между случайными величинами – от сильной, почти функциональной зависимости до практической независимости.

Пример 3.11. Пусть X – рост наугад выбранного взрослого человека (скажем, в сантиметрах), а Y – его вес (в килограммах). Зависимость между ростом и весом является весьма сильной, в первом приближении ее можно даже считать функциональной. Формула, приближенно выражающая эту зависимость, пишется обычно:

$$Y \text{ (кг)} = X \text{ (см)} - 100.$$

Пример 3.12. X – высота выбранного наугад дерева в лесу, Y – диаметр его основания. И здесь зависимость следует признать сильной, хотя и не в такой степени, как в предыдущем примере.

Пример 3.13. Из груды камней неправильной формы выбирают наугад один камень. Пусть X – его масса, а Y – наибольшая длина. Зависимость между X и Y носит сугубо вероятностный характер.

Пример 3.14. X – рост выбранного наугад взрослого человека, Y – его возраст. Наблюдения показывают, что эти величины практически независимы.

2°. Определение независимости случайных величин. Оставим пока в стороне вопрос о том, какими числами можно выразить степень зависимости между величинами X и Y . Ограничимся строгим определением *независимости* случайных величин.

Определение. Пусть задана система (X, Y) . Мы скажем, что величины X и Y *независимы*, если независимы события $X \in A$ и $Y \in B$, где A и B – любые два отрезка $[a_1, a_2]$ и $[b_1, b_2]$.

Иными словами выполняется равенство

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B). \quad (3.18)$$

В случае системы дискретного типа условие независимости X и Y можно представить в более обозримом виде:

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j), \quad (3.19)$$

где x_i – любое возможное значение величины X , а y_j – любое возможное значение величины Y .

Действительно, из (3.18) очевидным образом следует (3.19). Проверим, что и обратно, из (3.19) следует (3.18).

Пусть система (X, Y) характеризуется таблицей

$X \backslash Y$	y_1	y_2	$\dots$
x_1	p_{11}	p_{12}	$\dots$
x_2	p_{21}	p_{22}	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$

Положим $A = [a_1, a_2]$, $B = [b_1, b_2]$. Тогда

$$p_{ij} = P(X = x_i) \cdot P(Y = y_j) \quad (i, j = 1, 2, \dots)$$

(написанное равенство и есть как раз условие (3.19)). Отсюда

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{\{i, j | x_i \in A, y_j \in B\}} p_{ij} = \sum_{\{i, j | x_i \in A, y_j \in B\}} P(X = x_i) \cdot P(Y = y_j) = \\ &= \sum_{\{i | x_i \in A\}} P(X = x_i) \cdot \sum_{\{j | y_j \in B\}} P(Y = y_j) = P(X \in A) \cdot P(Y \in B), \end{aligned}$$

т.е. величины X и Y независимы.

§ 3.7. Функция от случайной величины. Действия над случайными величинами

Пусть X – случайная величина. Часто возникает необходимость в рассмотрении случайных величин Y вида:

$$Y = g(X), \quad (3.20)$$

где $g(x)$ – заданная числовая функция. Какой смысл вкладывается в запись (3.20), т. е. в понятие функции от случайной величины?

Предположим, что в результате опыта наступило событие

$$X = x$$

т. е. величина X приняла значение x . Тогда, по определению, мы считаем, что в данном опыте величина Y приняла значение $g(x)$. Ясно, что для дискретной случайной величины такое соглашение вполне определяет новую случайную величину Y . Что касается непрерывной случайной величины, то справедливо следующее утверждение.

Предложение 3.1. Если $g(x)$ непрерывная функция, то соотношение (3.20) определяет случайную величину Y .

Доказательство. Мы воспользуемся условием (3.2), эквивалентным определению случайной величины. Тем самым нам надо проверить, что для любого открытого множества U на числовой прямой множество элементарных событий, для которых

$$(g(X) \in U), \quad (3.21)$$

будет событием. Поскольку по условию $g(x)$ непрерывная функция и U – открытое множество, то множество $V = g^{-1}(U)$ также открытое. Условие (3.21) эквивалентно условию

$$(X \in V). \quad (3.22)$$

Но по определению (3.2) множество элементарных событий, определенного условием (3.22), является событием. Поэтому и условие (3.21) определяет событие, что и требовалось доказать.

Для любой функции (3.20) случайная величина

$$Y = g(X),$$

подобно X , имеет свой закон распределения. Каков этот закон? Ограничимся рассмотрением того случая, когда случайная величина X – дискретного типа. Пусть закон распределения X задан таблицей (3.11). По определению, закон распределения случайной величины Y задается таблицей (3.23), в кото-

рой первую строку (3.11) мы заменили на соответствующие значения функции $g(x)$, оставив без изменения вторую строку.

Y	$g(x_1)$	$g(x_2)$	$\dots$
P	p_1	p_2	$\dots$

(3.23)

Если среди значений Y имеются равные, то надо объединить соответствующие столбцы в один столбец, сложив соответствующие вероятности.

Пример 3.15. Пусть случайная величина X задана законом распределения:

X	-2	-1	0	1	2
P	0,2	0,2	0,2	0,2	0,2

Найти закон распределения случайной величины $Y = X^2$.

Решение. Для того чтобы найти закон распределения $Y = X^2$, возведем все значения в квадрат и получим следующую таблицу

Y	4	1	0	1	4
P	0,2	0,2	0,2	0,2	0,2

Далее объединим столбцы с равными значениями в один столбец и упорядочим значения Y по возрастанию, получим искомый закон распределения:

Y	0	1	4
P	0,2	0,4	0,4

Очень часто для случайных величин X и Y , образующих систему, приходится рассматривать их сумму и произведение. Поскольку закон распределения таких и подобных им операций над случайными величинами определяется аналогичным образом, будем считать, что мы рассматриваем случайную величину

$$Z = g(X, Y),$$

где $g(x, y)$ – некоторая числовая функция.

Итак, пусть система (X, Y) характеризуется таблицей

$X \backslash Y$	y_1	y_2	$\dots$
x_1	p_{11}	p_{12}	$\dots$
x_2	p_{21}	p_{22}	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$

(3.24)

смысл которой читателю известен. Величина

$$Z = g(X, Y)$$

также будет дискретной. Ее возможными значениями будут числа

$$z_{11} = g(x_1, y_1), z_{12} = g(x_1, y_2), \dots$$

Разберем два случая.

1. Все числа z_{ij} различны. Тогда событие $Z = z_{ij}$, т.е.

$$g(X, Y) = z_{ij},$$

наступает только тогда, когда одновременно наступают события $X = x_i$ и $Y = y_j$, следовательно, его вероятность будет равна

$$P(X = x_i, Y = y_j) = p_{ij}.$$

Таким образом, величина Z будет иметь закон распределения

z_{11}	z_{12}	$\dots$
p_{11}	p_{12}	$\dots$

(3.25)

2. Среди чисел z_{ij} имеются равные. Пусть, например, $z_{12} = z_{35} = a$ и никаких других z_{ij} , равных числу a , нет. Тогда событие $g(X, Y) = a$

будет суммой событий

$$(X = x_1, Y = y_2) \text{ и } (X = x_3, Y = y_5),$$

следовательно, его вероятность будет

$$p_{12} + p_{35}.$$

Подводя итог, можно сказать, что закон распределения величины $g(X, Y)$ будет выражаться таблицей (3.25), в которой столбцы с одинаковыми значениями z_{ij} следует объединить в один, сложив стоящие в них вероятности p_{ij} .

Пример 3.16. Пусть закон распределения системы случайных величин (X, Y) задается таблицей. Найти закон распределения их произведения.

$X \backslash Y$	1	2	3
-2	0,1	0,1	0,2
-1	0	0,2	0,1
0	0,2	0,1	0

Решение. Числа z_{ij} в данном случае будут

$$z_{11} = -2 \quad z_{12} = -4 \quad z_{13} = -6$$

$$z_{21} = -1 \quad z_{22} = -2 \quad z_{23} = -3$$

$$z_{31} = 0 \quad z_{32} = 0 \quad z_{33} = 0.$$

Поэтому "предварительный" закон распределения для $X \cdot Y$ будет

-2	-4	-6	-1	-2	-3	0	0	0
0,1	0,1	0,2	0	0,2	0,1	0,2	0,1	0

а окончательный

XY	-6	-4	-3	-2	0
P	0,2	0,1	0,1	0,3	0,3

Числовые характеристики дискретных случайных величин

Исчерпывающей характеристикой дискретной случайной величины является ее закон распределения. Но далеко не в каждой задаче нужно знать весь закон распределения. В ряде случаев можно обойтись одним или несколькими числами, отражающими наиболее важные особенности закона распределения. Такого рода числа называют *числовыми характеристиками* случайной величины (или соответствующего закона распределения). Их роль в теории вероятностей чрезвычайно велика; многие задачи удастся решить до конца, оставляя в стороне законы распределения и оперируя только числовыми характеристиками.

§ 4.1. Математическое ожидание случайной величины

Наиболее важное место среди числовых характеристик занимает так называемое *математическое ожидание* или *среднее значение* случайной величины.

Чтобы подойти естественным путем к понятию математического ожидания, будем рассуждать следующим образом. Пусть X – дискретная случайная величина, связанная с некоторым опытом. Предположим, что опыт осуществлен N раз и при этом величина X :

N_1 раз принимала значение x_1

N_2 раз принимала значение x_2

и т.д. Найдем среднее арифметическое всех значений, принятых величиной X в данной серии опытов. Оно запишется:

$$\frac{x_1 N_1 + x_2 N_2 + \dots}{N} = x_1 \frac{N_1}{N} + x_2 \frac{N_2}{N} + \dots$$

Дробь $\frac{N_1}{N}$ есть частота, с которой появлялось значение x_1 ; с увеличением числа опытов N эта

дробь будет приближаться к p_1 – вероятности события $X = x_1$. Аналогичным образом дробь $\frac{N_2}{N}$ будет приближаться к p_2 и т.д. В итоге получаем, что с увеличением N среднее арифметическое будет приближаться к числу

$$x_1 p_1 + x_2 p_2 + \dots$$

Дадим теперь такое

Определение. *Математическим ожиданием или средним значением дискретной случайной величины X с законом распределения*

x_1	x_2	...
p_1	p_2	...

(4.1)

называется число

$$M(X) = x_1 p_1 + x_2 p_2 + \dots \quad (4.2)$$

Таким образом, математическое ожидание дискретной случайной величины X равно сумме произведений возможных значений величины X на их вероятности.

Смысл числа $M(X)$ ясен из приведенного выше рассуждения. Он заключается в том, что около числа $M(X)$ колеблется среднее арифметическое значений, принимаемых величиной X в больших сериях опытов.

Из определения (4.2) непосредственно вытекает следующий факт.

Математическое ожидание постоянной величины равно ей самой.

Действительно, постоянную величину c можно рассматривать как дискретную случайную величину X , принимающую только одно значение c с вероятностью 1. Но тогда

$$M(X) = c \cdot 1 = c.$$

Рассмотрим ряд примеров на нахождение математического ожидания.

Пример 4.1. Найти математическое ожидание числа очков, выпадающих при бросании игральной кости.

Решение. Обозначим указанную случайную величину через X . Ее закон распределения имеет вид:

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Отсюда

$$M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3,5.$$

Пример 4.2. Взнос клиента в страховую компанию составляет 10 руб., а выплата компании при наличии страхового случая 1000 руб. Вероятность страхового случая 0,009. Какую прибыль в среднем имеет компания от одного клиента?

Решение. Пусть X – прибыль от одного клиента. Закон распределения для случайной величины X имеет вид

X	10	–990
P	0,991	0,009

Отсюда

$$M(X) = 9,91 - 8,91 = 1 \text{ руб.}$$

§ 4.2. Свойства математического ожидания

Ранее был отмечен такой факт: математическое ожидание постоянной величины равно ей самой. Добавим теперь к этому еще одно свойство:

$$M(cX) = cM(X) \quad (4.3)$$

– *постоянный множитель c можно вынести за знак математического ожидания.*

Для дискретной случайной величины равенство (4.3) очевидно: если X имеет закон распределения:

X	x_1	x_2	...
P	p_1	p_2	...

то cX будет иметь закон распределения:

cX	cx_1	cx_2	...
P	p_1	p_2	...

откуда следует:

$$M(cX) = cx_1p_1 + cx_2p_2 + \dots = c(x_1p_1 + x_2p_2 + \dots) = cM(X).$$

1°. Математическое ожидание суммы. Наиболее существенным из всех свойств математического ожидания является следующий факт, который часто называют *теоремой сложения* математических ожиданий.

Теорема 4.1. Математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий

$$M(X + Y) = M(X) + M(Y). \quad (4.4)$$

Доказательство. Пусть закон распределения системы (X, Y) выражается таблицей

$X \backslash Y$	y_1	y_2
x_1	p_{11}	p_{12}
x_2	p_{21}	p_{22}
$\dots$	$\dots$	$\dots$

где $x_1, x_2, \dots$ – возможные значения величины X , $y_1, y_2, \dots$ – возможные значения Y , а $p_{ij} = P(X = x_i, Y = y_j)$. Закон распределения величины $X + Y$ будет выражаться таблицей

$x_1 + y_1$	$x_1 + y_2$	$x_2 + y_1$	$x_2 + y_2$	$\dots$
p_{11}	p_{12}	p_{21}	p_{22}	$\dots$

где $p_{ij} = P(X = x_i, Y = y_j)$, причем в случае одинаковых сумм $x_i + y_j$ и $x_k + y_l$ соответствующие столбцы нужно объединить в один, сложив записанные в них вероятности. Отсюда ясно, что для $M(X + Y)$ будем иметь равенство

$$M(X + Y) = \sum_{i,j} (x_i + y_j) p_{ij}. \quad (4.5)$$

Перепишем формулу (4.5) следующим образом:

$$M(X + Y) = \sum_{i,j} x_i p_{ij} + \sum_{i,j} y_j p_{ij}.$$

Первую сумму правой части можно представить в виде

$$\sum_i \left(x_i \sum_j p_{ij} \right).$$

Выражение $\sum_j p_{ij}$, или подробнее $p_{i1} + p_{i2} + \dots$, есть вероятность того, что наступит какое-либо из событий

$$(X = x_i, Y = y_1), (X = x_i, Y = y_2), \dots,$$

т.е. в конечном счете, что X примет значение x_i безотносительно к тому, какое значение примет Y . Следовательно, это выражение равно $P(X = x_i)$. Отсюда

$$\sum_{i,j} x_i p_{ij} = \sum_i x_i P(X = x_i) = M(X).$$

Аналогично покажем, что

$$\sum_{i,j} y_j p_{ij} = \sum_j y_j P(Y = y_j) = M(Y).$$

В итоге имеем:

$$M(X + Y) = M(X) + M(Y),$$

что и требовалось получить.

Нетрудно видеть, что формула (4.4) обобщается на любое число слагаемых:

$$M(X_1 + X_2 + \dots + X_n) = M(X_1) + M(X_2) + \dots + M(X_n).$$

Заметим, что доказанные выше свойства математического ожидания

$$M(cX) = cM(X)$$

и

$$M(X + Y) = M(X) + M(Y)$$

называют обычно *свойствами линейности*. Оба свойства линейности можно записать в виде одной формулы:

$$M(aX + bY) = a M(X) + b M(Y).$$

2°. Математическое ожидание произведения. Доказанная в предыдущем пункте формула (4.4) наводит на мысль о возможности аналогичной формулы для случайной величины XY :

$$M(XY) = M(X) \cdot M(Y). \quad (4.6)$$

Однако формула (4.6) в общем случае неверна; чтобы в этом убедиться, достаточно взять в качестве X случайную величину с законом распределения:

-1	1
$\frac{1}{2}$	$\frac{1}{2}$

Тогда X^2 будет постоянной величиной 1 и равенство $M(X^2) = M(X) \cdot M(X)$ выполняться не будет ($1 \neq 0 \cdot 0$).

Докажем, что если случайные величины X и Y независимы, то формула (4.6) справедлива.

Теорема 4.2. Если случайные величины X и Y независимы, то математическое ожидание их произведения равно произведению их математических ожиданий, т. е.

$$M(XY) = M(X) \cdot M(Y).$$

Эту важную теорему часто называют *теоремой умножения математических ожиданий*.

Д о к а з а т е л ь с т в о . Возможные значения X обозначим, как и ранее, $x_1, x_2, \dots$, возможные значения Y — $y_1, y_2, \dots$.

Применив такие же рассуждения, как при выводе формулы (4.4), получим равенство

$$M(XY) = \sum_{i,j} x_i y_j p_{ij},$$

где p_{ij} есть вероятность события $X = x_i, Y = y_j$.

Ввиду независимости величин X и Y имеем:

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j). \quad (4.7)$$

Обозначив

$$P(X = x_i) = r_i, P(Y = y_j) = s_j,$$

перепишем равенство (4.7) в виде

$$p_{ij} = r_i s_j.$$

Итак,

$$M(XY) = \sum_{i,j} x_i y_j r_i s_j = \sum_{i,j} x_i r_i y_j s_j.$$

Преобразуя полученное равенство, будем иметь:

$$M(XY) = \left(\sum_i x_i r_i \right) \left(\sum_j y_j s_j \right) = M(X)M(Y),$$

что и требовалось получить.

§ 4.3. Дисперсия случайной величины

Различные случайные величины могут иметь одно и то же математическое ожидание. Рассмотрим, например, величины X и Y , законы распределения которых заданы таблицами вида:

X	-0,01	0,02
P	$\frac{2}{3}$	$\frac{1}{3}$

Y	-100	100
P	$\frac{1}{2}$	$\frac{1}{2}$

Математические ожидания их одинаковы (равны нулю):

$$M(X) = -0,01 \cdot \frac{2}{3} + 0,02 \cdot \frac{1}{3} = 0,$$

$$M(Y) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} = 0.$$

Однако характер распределения величин X и Y существенно различен. В то время как величина X может принимать лишь значения, мало отличающиеся от ее математического ожидания, значения величины Y значительно удалены от $M(Y)$.

Аналогичных примеров можно привести много. В двух различных географических местностях могут оказаться одинаковые средние уровни осадков, в двух учреждениях с различным соотношением низкооплачиваемых и высокооплачиваемых служащих может оказаться одна и та же средняя заработная плата и т. д.

Укажем еще один пример. Пусть имеются два различных прибора для измерения одной и той же физической величины. Практически результат измерения никогда не совпадает точно с измеряемой величиной: каждое измерение сопровождается случайной ошибкой. Обозначим ошибку при измерении первым прибором через X , вторым – через Y . Очевидно, X и Y – случайные величины. Предполагая оба прибора исправными, будем иметь $M(X) = 0$, $M(Y) = 0$ (написанные равенства означают, что измерения свободны от систематической ошибки). Однако сам по себе этот факт не означает одинаковой точности обоих приборов. Вполне может быть, что для одного из приборов ошибка (по абсолютной величине) принимает в среднем большие значения, чем для другого; проще говоря, один из приборов может оказаться более "разболтанным", давать большее рассеивание результатов измерения, чем другой.

1°. Определение дисперсии и среднего квадратичного отклонения. Рассмотренные в начале параграфа примеры убеждают нас в необходимости вести еще одну числовую характеристику – для измерения степени *рассеивания*, *разброса* значений, принимаемых случайной величиной X , вокруг ее математического ожидания.

Обсудим вопрос о возможных подходах к оценке рассеивания.

На первый взгляд может показаться естественным в качестве характеристики рассеивания взять математическое ожидание разности $X - m$, где m – математическое ожидание величины X . Заметим, что случайную величину $X - m$ называют *отклонением* (имеется в виду отклонение величины X от ее математического ожидания) и обозначают $\overset{\circ}{X}$:

$$\overset{\circ}{X} = X - m.$$

Оказывается, однако, что для характеристики рассеивания число $M(\overset{\circ}{X})$ ничего не дает – оно всегда равно нулю. Проверка этого факта тривиальна:

$$M\left(\overset{\circ}{X}\right) = M(X - m) = M(X) - M(m) = m - m = 0.$$

В действительности степень рассеивания, т. е. степень *удаленности X от m* , должна определяться не самим отклонением, а его *абсолютной величиной* $|X - m|$. Поэтому более логично в качестве меры рассеивания принять не $M(X - m)$, а $M(|X - m|)$. Такая мера рассеивания называется *средним линейным отклонением*, однако пользоваться ею не очень удобно; читатель по собственному опыту знает, сколь трудно оперировать с формулами, содержащими абсолютные величины. Имеется и другое возражение против числа $M(|X - m|)$; определенная таким образом мера рассеивания не обладает хорошими свойствами. В этом отношении она сильно уступает той мере рассеивания, которая вводится ниже и носит название *дисперсии*.

Определение. *Дисперсией* случайной величины X называется число

$$D(X) = M[(X - m)^2]. \quad (4.8)$$

Другими словами, *дисперсия есть математическое ожидание квадрата отклонения*.

Из определения (4.8) легко вытекают следующие свойства дисперсии.

1. *Дисперсия принимает только неотрицательные значения*

$$D(X) \geq 0.$$

Дисперсия дискретной величины X равна нулю тогда и только тогда, когда $X = c$ – константа.

Действительно, запишем развернутую формулу дисперсии дискретной случайной величины X , имеющей закон распределения, заданный таблицей

X	x_1	x_2	$\dots$
P	p_1	p_2	$\dots$

(4.9)

$$D(X) = (x_1 - m)^2 p_1 + (x_2 - m)^2 p_2 + \dots \quad (4.10)$$

Правая часть формулы (4.10) есть сумма неотрицательных слагаемых, причем можно считать, что все вероятности p_i положительные числа. Поэтому $D(X) = 0$ тогда и только тогда, когда для всех i выполняются равенства $x_i = m$. А это и означает, что случайная величина X принимает постоянное значение.

2. При умножении случайной величины X на постоянное число c ее дисперсия умножается на c^2

$$D(cX) = c^2 D(X). \quad (4.11)$$

Действительно,

$$\begin{aligned} D(cX) &= M\left[(cX - M(cX))^2\right] = M\left[c^2(X - M(X))^2\right] = \\ &= c^2 M\left[(X - M(X))^2\right] = c^2 D(X). \end{aligned}$$

3. Если к случайной величине X прибавить константу a , то дисперсия не изменится

$$D(X+a) = D(X). \quad (4.12)$$

Очевидно, что

$$X + a - M(X + a) = X - M(X),$$

т.е. отклонение случайной величины не изменится, следовательно, не изменится и ее дисперсия.

По свойству 1 дисперсия принимает только неотрицательные значения, поэтому определено число

$$\sigma(X) = \sqrt{D(X)}, \quad (4.13)$$

которое носит название **среднего квадратического отклонения** величины X . Если дисперсия характеризует средний размер квадрата отклонения, то число $\sigma(X)$ можно рассматривать как некоторую среднюю характеристику самого отклонения, точнее, величины $|X - m|$.

В заключение укажем весьма полезную общую формулу, вытекающую из определения дисперсии:

$$D(X) = M(X^2) - M^2(X). \quad (4.14)$$

Доказательство этой формулы основано на известных читателю свойствах математического ожидания. Мы имеем:

$$\begin{aligned} D(X) &= M[(X - m)^2] = M(X^2 - 2mX + m^2) = \\ &= M(X^2) - 2mM(X) + M(m^2) = \\ &= M(X^2) - 2m^2 + m^2 = M(X^2) - m^2. \end{aligned}$$

2°. Вычисление дисперсии. Наряду с формулой (4.10) часто используют и формулу (4.14). В развернутом виде для дискретной случайной величины X с законом распределения (4.9) она выглядит следующим образом:

$$D(X) = \sum_i x_i^2 p_i - m^2. \quad (4.15)$$

Пример 4.3. Пусть X – число очков, выпадающих при бросании игральной кости. Распределение величины X описывается таблицей:

[illegible]

По формуле (4.15), учитывая, что в данном случае $m = 3,5$, находим:

$$D(X) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} - 3,5^2 = \frac{35}{12}.$$

3°. Нормированные случайные величины.

Определение. Случайная величина Y называется **нормированной**, если ее математическое ожидание равно нулю, а дисперсия – единице:

$$M(Y) = 0, D(Y) = 1.$$

От любой случайной величины X можно перейти к нормированной случайной величине Y с помощью линейного преобразования:

$$Y = \frac{X - m}{\sigma},$$

где m – математическое ожидание, а σ – среднее квадратичное отклонение величины X . Нормированность Y проверяется весьма просто:

$$M(Y) = \frac{1}{\sigma} M(X - m) = 0,$$

$$D(Y) = \frac{1}{\sigma^2} D(X - m) = \frac{1}{\sigma^2} D(X) = 1.$$

§ 4.4. Дисперсия суммы случайных величин. Ковариация

Пусть X и Y – две случайные величины. Положим,

$$Z = X + Y.$$

По теореме сложения математических ожиданий будем иметь:

$$M(Z) = M(X) + M(Y).$$

Вычитая это равенство из предыдущего, получим:

$$\overset{\circ}{Z} = \overset{\circ}{X} + \overset{\circ}{Y},$$

где $\overset{\circ}{X}$ обозначает, как и раньше, отклонение величины X . Отсюда

$$\overset{\circ}{Z}^2 = \overset{\circ}{X}^2 + \overset{\circ}{Y}^2 + 2\overset{\circ}{X}\overset{\circ}{Y}.$$

Найдем теперь дисперсию величины $X + Y$. Имеем

$$D(X + Y) = D(Z) = M\left(\overset{\circ}{Z}^2\right) = M\left(\overset{\circ}{X}^2\right) + M\left(\overset{\circ}{Y}^2\right) + 2M\left(\overset{\circ}{X}\overset{\circ}{Y}\right),$$

откуда

$$D(X + Y) = D(X) + D(Y) + 2M\left(\overset{\circ}{X}\overset{\circ}{Y}\right). \quad (4.16)$$

Число $M\left(\overset{\circ}{X}\overset{\circ}{Y}\right)$ имеет особое значение для характеристики системы случайных величин (X, Y) .

Его называют **ковариацией** или **корреляционным моментом** случайных величин X и Y и обозначают $\text{Cov}(X, Y)$. Таким образом, по определению,

$$\text{Cov}(X, Y) = M\left(\overset{\circ}{X}\overset{\circ}{Y}\right).$$

Формула (4.16) принимает теперь следующий вид:

$$D(X + Y) = D(X) + D(Y) + 2\text{Cov}(X, Y) \quad (4.17)$$

– дисперсия суммы равна сумме дисперсий плюс удвоенная ковариация.

Ковариация или корреляционный момент, как свидетельствует его название (от латинского слова *correlatio* – соответствие, взаимосвязь), играет определенную роль при оценке зависимости X и Y . Основное свойство ковариации выражается следующим предложением.

Если величины X и Y независимы, то их ковариация равна нулю.

Действительно, пусть X и Y независимы. Тогда, очевидно, величины $\overset{\circ}{X}$ и $\overset{\circ}{Y}$ будут тоже независимы. Отсюда вытекает, что математическое ожидание произведения $\overset{\circ}{X}\overset{\circ}{Y}$ будет:

$$M\left(\overset{\circ}{X}\overset{\circ}{Y}\right) = M\left(\overset{\circ}{X}\right)M\left(\overset{\circ}{Y}\right) = 0 \cdot 0 = 0.$$

Из доказанного предложения следует: если $\text{Cov}(X, Y) \neq 0$, то величины X и Y не могут быть независимыми. Таким образом, неравенство нулю ковариации определенно свидетельствует о наличии связи между величинами X и Y .

Из формулы (4.17) и доказанного выше свойства ковариации получаем важное следствие, носящее название *теоремы сложения дисперсий для независимых случайных величин*.

Теорема 4.3. *Если величины X и Y независимы, то дисперсия их суммы равна сумме их дисперсий:*

$$D(X + Y) = D(X) + D(Y).$$

§ 4.5. Числовые характеристики типичных законов распределения

Теперь мы готовы к тому, чтобы найти математическое ожидание и дисперсию основных законов распределения: биномиального, пуассоновского и геометрического. Но прежде мы рассмотрим числовые характеристики простейшего закона распределения – *индикатора случайного события A* .

Пример 4.4. Пусть закон распределения случайной величины X имеет вид

X	0	1
P	q	p

Найти математическое ожидание и дисперсию X .

Решение. Очевидно, что $M(X) = p$. Случайная величина X^2 имеет тот же закон распределения, поэтому

$$D(X) = M(X^2) - M^2(X) = p - p^2 = p(1 - p) = pq.$$

Таким образом, для индикатора случайного события

$$M(X) = p, D(X) = pq. \quad (4.18)$$

1°. Биномиальный закон.

Пусть производится n независимых испытаний, в каждом из которых может появиться событие A с вероятностью p , так что вероятность противоположного события $\bar{A}$ равна $q = 1 - p$. Рассмотрим случайную величину X – число появлений события A в n опытах. Мы знаем, что величина X распределена по биномиальному закону, т.е. для любого целого числа k , $0 \leq k \leq n$, справедлива формула

$$P_n(k) = C_n^k p^k q^{n-k}.$$

Непосредственное нахождение числовых характеристик в данном случае, используя закон распределения, является делом возможным, но затруднительным. Применим теорему сложения математических ожиданий и дисперсий для независимых величин. Для этого представим случайную величину X в виде суммы индикаторов события A для каждого испытания:

$$X = X_1 + X_2 + \dots + X_n.$$

Согласно примеру 4.4 математическое ожидание и дисперсия каждого слагаемого находится по формуле (4.18):

$$M(X_i) = p, D(X_i) = pq, 1 \leq i \leq n.$$

По теореме сложений математических ожиданий $M(X) = np$. Поскольку случайные величины X_i независимы, то дисперсии тоже складываются, так что $D(X) = npq$.

Окончательно имеем

$$M(X) = np, D(X) = npq. \quad (4.19)$$

2°. Закон Пуассона.

Закон Пуассона задается таблицей:

X	0	1	2	3	...
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!} e^{-\lambda}$	$\frac{\lambda^3}{3!} e^{-\lambda}$	...

Отсюда имеем:

$$M(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Таким образом, параметр λ , характеризующий данное пуассоновское распределение, есть не что иное как математическое ожидание величины X . Это легко понять, если вспомнить, что формулы Пуассона получились как предельный вариант формул Бернулли, когда $n \rightarrow \infty$, причем $\lambda = np$. Поскольку для биномиального закона математическое ожидание равно np , то неудивительно, что для пуассоновского закона $M(X) = \lambda$. Более того, мы можем предположить, что дисперсия X тоже будет равна λ , поскольку для биномиального закона $D(X) = npq$ и $q \rightarrow 1$ при $n \rightarrow \infty$. Действительно, непосредственный подсчет дисперсии подтверждает это предположение, однако мы не приводим его здесь из-за сложности выкладок. Ниже мы выведем эти формулы более простым способом. Таким образом, для закона Пуассона

$$M(X) = \lambda, D(X) = \lambda. \quad (4.20)$$

3°. Геометрический закон.

Напомним, что геометрический закон распределения связан с последовательностью испытаний Бернулли, которые проводятся до первого появления события A . Как обычно, вероятность появления события A в одном испытании равна p , противоположного события — $q = 1 - p$. Закон распределения случайной величины X — числа испытаний был найден ранее и имеет вид

X	1	2	3	...	n	...
P	p	pq	pq^2	...	pq^{n-1}	...

Запишем выражение для математического ожидания

$$\begin{aligned} M(X) &= 1 \cdot p + 2 \cdot pq + 3 \cdot pq^2 + \dots + n \cdot pq^{n-1} + \dots = \\ &= p(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots) \end{aligned}$$

Ряд, записанный в скобках, получается почленным дифференцированием геометрической прогрессии

$$q + q^2 + q^3 + \dots + q^n + \dots$$

Следовательно,

$$M(X) = p \left(\frac{1}{1-q} - 1 \right)' = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}.$$

Формула для дисперсии $D(X) = \frac{q}{p^2}$ будет получена ниже. Таким образом, для геометрического закона распределения

$$M(X) = \frac{1}{p}, D(X) = \frac{q}{p^2}. \quad (4.21)$$

Пример 4.5. Проводится 160 независимых испытаний, в каждом из которых подбрасывается 5 монет. Пусть X — число испытаний, в которых выпало 3 герба. Найти математическое ожидание и дисперсию случайной величины X .

Решение. Очевидно, что величина X распределена по биномиальному закону, а вероятность того, что в данном испытании выпадет 3 герба, можно вычислить по формуле Бернулли

$$p = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32} = \frac{5}{16}, \quad q = 1 - p = \frac{11}{16}.$$

Согласно формуле (4.19)

$$M(X) = np = 160 \cdot \frac{5}{16} = 50, \quad D(X) = npq = 160 \cdot \frac{5}{16} \cdot \frac{11}{16} = \frac{275}{8}.$$

Пример 4.6. В спортивной лотерее каждую неделю на 100 билетов разыгрывается 5 палаток и 5 рюкзаков. Турист решил каждую неделю покупать по одному билету до тех пор, пока он не выиграет палатку и рюкзак. Найти среднее время (в неделях) реализации этого намерения.

Решение. Обозначим через X число розыгрышей лотереи до реализации желания туриста. $X = X_1 + X_2$, где X_1 – число розыгрышей до первого выигрыша, X_2 – число розыгрышей до очередного желаемого выигрыша (рюкзака, если первый выигрыш был палаткой, или наоборот). Случайные величины X_1, X_2 распределены по геометрическому закону, причем для X_1 параметр p равен $\frac{10}{100} = \frac{1}{10}$, для X_2 – $\frac{5}{100} = \frac{1}{20}$. По теореме сложения вероятностей и формуле (4.21) имеем

$$M(X) = M(X_1) + M(X_2) = 10 + 20 = 30.$$

Пример 4.7. Игральная кость подбрасывается до тех пор, пока не выпадут все грани. Найти математическое ожидание и дисперсию числа бросаний.

Решение. Случайная величина X – число бросаний до выпадения всех граней представима в виде суммы шести независимых случайных величин

$$X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6.$$

Здесь X_i – число бросаний после выпадения $i - 1$ различных граней до выпадения i разных граней. Каждая случайная величина X_i распределена по геометрическому закону, причем параметр $p = \frac{7-i}{6}$. Применяя теорему сложения для математических ожиданий и дисперсий, а также формулы (4.21), имеем

$$M(X) = 6 \left(\frac{1}{6} + \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + 1 \right) = 14,7,$$

$$D(X) = \frac{\frac{1}{6}}{\left(\frac{5}{6}\right)^2} + \frac{\frac{2}{6}}{\left(\frac{4}{6}\right)^2} + \frac{\frac{3}{6}}{\left(\frac{3}{6}\right)^2} + \frac{\frac{4}{6}}{\left(\frac{2}{6}\right)^2} + \frac{\frac{5}{6}}{\left(\frac{1}{6}\right)^2} = 38,99.$$

Замечание. Последний пример допускает обобщение на случай, когда в каждом испытании разыгрывается N равновероятных возможностей и успех наступает в случае полного набора всех этих случаев. Например, речь может идти о составлении полного набора картинок и т.п. Формула для математического ожидания будет выглядеть следующим образом

$$M(X) = N \left(1 + \frac{1}{2} + \dots + \frac{1}{N} \right).$$

Для больших значений N выражение в скобках можно заменить на $\ln N$ и получить приближенную формулу $M(X) \approx N \ln N$.

§ 4.6. Производящие функции

Мы видели, что дискретные случайные величины, рассмотренные в предыдущих параграфах, принимали только *целые* значения $X = 0, 1, 2, \dots$. Нахождение числовых характеристик таких величин упрощается, если рассмотреть *производящие функции*.

Определение. *Производящей функцией* дискретной целочисленной случайной величины X с законом

$P(X = k) = p_k, k = 0, 1, 2, \dots$ называется функция, заданная степенным рядом

$$M(s^X) = P(s) = p_0 + p_1s + p_2s^2 + \dots \quad (4.22)$$

Поскольку все коэффициенты этого ряда не превосходят 1, то сравнение с геометрической прогрессией показывает, что этот ряд сходится, по крайней мере, для значений $|s| < 1$. Из свойства закона распределения видно, что $P(1) = 1$, так что область сходимости ряда содержит точку $s = 1$.

Теорема 4.4. *Производящая функция суммы независимых случайных величин X и Y равна произведению производящих функций слагаемых*

$$P_{X+Y}(s) = P_X(s) \cdot P_Y(s). \quad (4.23)$$

Доказательство. Имеем по определению

$$P_{X+Y}(s) = M(s^{X+Y}) = M(s^X s^Y) = M(s^X) M(s^Y) = P_X(s) P_Y(s).$$

Пример 4.8. Найти производящую функцию для биномиального закона.

Решение. Если вспомнить, что формула Бернулли получается из разложения бинома, то легко сообразить, что

$$P(s) = (q + ps)^n. \quad (4.24)$$

Можно также вспомнить, что случайная величина X , распределенная по биномиальному закону, представляется в виде суммы n независимых слагаемых – индикаторов появления события A . Очевидно, что для одного слагаемого производящая функция равна $q + ps$. Теперь осталось применить (4.23).

Пример 4.9. Найти производящую функцию для геометрического закона распределения.

Решение. Имеем $P(X = k) = pq^{k-1}, k = 1, 2, \dots$, поэтому

$$\begin{aligned} P(s) &= ps + pqs^2 + pq^2s^3 + \dots = \frac{p}{q} (qs + q^2s^2 + q^3s^3 + \dots) = \\ &= \frac{p}{q} \left(\frac{1}{1-qs} - 1 \right) = \frac{ps}{1-qs}. \end{aligned}$$

Данное равенство справедливо для всех значений s , удовлетворяющих неравенству $|qs| < 1$, для которых мы применили формулу суммы бесконечно убывающей геометрической прогрессии. Таким образом,

$$P(s) = \frac{ps}{1-qs}, |s| < \frac{1}{q}. \quad (4.25)$$

Пример 4.10. Найти производящую функцию для распределения Пуассона.

Решение. Для пуассоновского закона с параметром $\lambda > 0$ имеем $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \dots$. Поэтому

$$P(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)},$$

причем все ряды сходятся для любых значений аргумента s . Окончательно

$$P(s) = e^{\lambda(s-1)}, -\infty < s < \infty. \quad (4.26)$$

В качестве следствия получим теорему.

Теорема 4.5. *Сумма независимых случайных величин, распределенных по закону Пуассона, распределена по тому же закону.*

Доказательство. Пусть X и Y – независимые случайные величины, распределенные по закону Пуассона с параметрами λ и μ . Тогда их производящие функции находятся по формуле (4.26):

$$P_X(s) = e^{\lambda(s-1)}, \quad P_Y(s) = e^{\mu(s-1)},$$

а производящая функция суммы $X + Y$ находится согласно теореме 4.4

$$P_{X+Y}(s) = P_X(s)P_Y(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)}.$$

Отсюда видно, что сумма будет распределена по закону Пуассона с параметром $\lambda + \mu$, что и требовалось доказать.

Зная производящую функцию дискретной случайной величины X , нетрудно найти ее математическое ожидание и дисперсию.

Теорема 4.6. Для дискретной случайной величины X с производящей функцией $P(s)$ выполняются следующие соотношения:

$$M(X) = P'(1), \quad (4.27)$$

$$D(X) = P''(1) + P'(1) - P'^2(1). \quad (4.28)$$

Доказательство. Дифференцируя почленно ряд (4.22) два раза, имеем

$$P'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1},$$

$$P''(s) = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2} = \sum_{k=2}^{\infty} k^2 p_k s^{k-2} - \sum_{k=2}^{\infty} k p_k s^{k-2}.$$

Подставляя $s = 1$, получим

$$P'(1) = \sum_{k=1}^{\infty} k p_k = M(X),$$

$$P''(1) = \sum_{k=1}^{\infty} k^2 p_k - \sum_{k=1}^{\infty} k p_k = D(X) + M^2(X) - M(X),$$

откуда легко получить формулы (4.25), (4.26).

Комбинируя полученные выражения для производящих функций биномиального, геометрического и пуассоновского законов (4.23), (4.24), (4.25) с формулами (4.26), (4.27), теперь нетрудно найти основные числовые характеристики этих законов.

1. Биномиальный закон.

Из выражения (4.23) для производящей функции получим

$$P(s) = np(q + ps)^{n-1}, \quad P''(s) = n(n-1)p^2(q + ps)^{n-2}.$$

Подставляя $s = 1$ и учитывая, что $p + q = 1$, имеем

$$P'(1) = np, \quad P''(1) = n(n-1)p^2.$$

Используя формулы (4.26), (4.27), получим

$$M(X) = np, \quad n(n-1)p^2 = D(X) + n^2p^2 - np,$$

откуда $D(X) = npq$.

2. Геометрический закон.

Дифференцируя два раза производящую функцию, имеем

$$P'(s) = \frac{p}{(1-qs)^2}, \quad P''(s) = \frac{2pq}{(1-qs)^3}.$$

Отсюда

$$M(X) = P'(1) = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}, \quad P''(1) = \frac{2pq}{(1-q)^3} = \frac{2q}{p^2} = D(X) + \frac{1}{p^2} - \frac{1}{p} = D(X) + \frac{q}{p^2},$$

откуда $D(X) = \frac{q}{p^2}$, что и требовалось.

3. Закон Пуассона.

Имеем

$$P'(s) = \lambda e^{\lambda(s-1)}, \quad P''(s) = \lambda^2 e^{\lambda(s-1)},$$

поэтому $P'(1) = \lambda$, $P''(1) = \lambda^2$. Подставляя найденные значения в формулы (4.26), (4.27), получим $M(X) = \lambda$, $D(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Непрерывные случайные величины и типичные законы распределения

§ 5.1. Непрерывные и абсолютно непрерывные случайные величины

Определение. Случайная величина X называется **непрерывной**, если ее функция распределения $F(x)$ непрерывна в любой точке x .

Принимая во внимание формулу

$$F(x+0) - F(x) = P(x = x_0),$$

доказанную ранее (см. § 3.2), можно дать другое определение непрерывной случайной величины: случайная величина *непрерывна*, если вероятность каждого ее отдельного значения равна нулю.

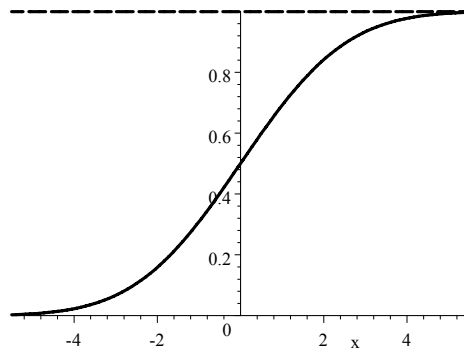


Рис. 5.1

Таким образом, если в случае дискретной случайной величины функция распределения $F(x)$ набирает рост только за счет отдельных «скачков», то в случае непрерывной случайной величины функция $F(x)$ вообще не имеет «скачков», т.е. является непрерывной функцией (рис. 5.1).

Поскольку вероятность отдельных значений непрерывной случайной величины X равна нулю, то вероятность попадания X в какой-либо промежуток от a до b не зависит от того, включаются или нет концы промежутка:

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X < b) = P(a \leq X < b) = \\ &= P(a < X \leq b) = F(b) - F(a). \end{aligned} \quad (5.1)$$

Определение. Случайная величина X называется **абсолютно непрерывной**, если найдется неотрицательная функция $f(x)$, называемая *плотностью распределения*, такая, что для $a < b$ вероятность попадания X в промежуток $[a, b]$ получается путем интегрирования данной функции

$$P(a \leq X \leq b) = \int_a^b f(x) dx. \quad (5.2)$$

Полагая $a = -\infty$, $b = x$, получим выражение функции распределения через функцию плотности

$$F(x) = \int_{-\infty}^x f(t) dt. \quad (5.3)$$

Установим некоторые свойства плотности.

1. Функция плотности принимает только неотрицательные значения

$$f(x) \geq 0. \quad (5.4)$$

Это свойство выполняется по определению.

2. Функция плотности обладает свойством **нормированности**

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (5.5)$$

Равенство (5.5) получается из формулы (5.2), если положить $a = -\infty$, $b = \infty$.

3. Во всех точках, где функция плотности непрерывна, выполняется равенство

$$f(x) = F'(x). \quad (5.6)$$

Непосредственно вытекает из формулы (5.3). Из соотношений (5.3), (5.6) понятно, почему иногда функцию плотности называют *дифференциальной функцией* распределения, а функцию распределения – *интегральной функцией*.

Итак, плотность вероятности всегда ≥ 0 и удовлетворяет условию нормированности. Следовательно, график $f(x)$ представляет собой кривую, лежащую в верхней полуплоскости, причем площадь, ограниченная этой кривой и осью Ox , равна 1 (рис. 5.2).

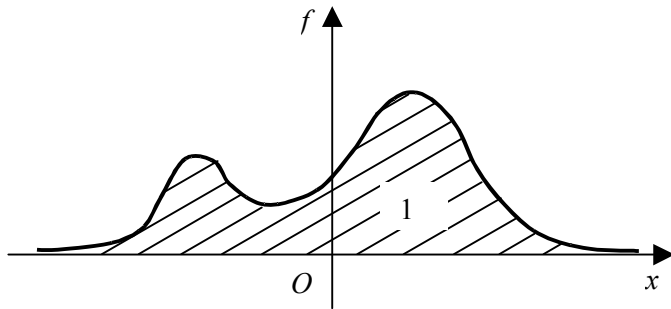


Рис. 5.2

Поясним смысл самого названия «плотность вероятности».

Согласно равенству (5.2), справедливо равенство

$$P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} f(x) dx, \quad (5.7)$$

По теореме о среднем интеграл, стоящий в правой части, равен $f(\alpha)(x_2 - x_1)$, где α — некоторая точка из интервала (x_1, x_2) . Отсюда

$$\frac{P(x_1 \leq X < x_2)}{x_2 - x_1} = f(\alpha). \quad (5.8)$$

Представим себе, что интервал (x_1, x_2) стягивается к некоторой точке x_0 , причем в этой точке функция $f(x)$ непрерывна. Тогда $f(\alpha)$ будет стремиться к числу $f(x_0)$, и мы получим:

$$\lim_{x_1, x_2 \rightarrow x_0} \frac{P(x_1 \leq x < x_2)}{x_2 - x_1} = f(x_0).$$

Отношение, стоящее под знаком предела, есть своего рода "вероятность на единицу длины" интервала (x_1, x_2) . Предел этого отношения естественно рассматривать как *плотность вероятности в самой точке* x_0 . Мы видим, таким образом, что во всякой точке x_0 , где $f(x)$ непрерывна, число $f(x)$ совпадает с естественно понимаемой плотностью вероятности в точке x_0 . Отсюда и название для функции $f(x)$ — "плотность вероятности".

В заключение дадим оценку для вероятности попадания абсолютно непрерывной величины X в малый промежуток $[x_0, x_0 + \Delta x]$. Согласно (5.8), можем записать

$$P(x_0 \leq X < x_0 + \Delta x) = f(\xi)\Delta x,$$

где ξ — некоторая точка из $[x_0, x_0 + \Delta x]$. Поэтому, если в точке x_0 функция $f(x)$ непрерывна, имеем приближенное равенство

$$P(x_0 \leq X \leq x_0 + \Delta x) \approx f(x_0)\Delta x.$$

Заметим, что выражение $f(x_0)\Delta x$ часто называют "*элементом вероятности*".

В приложениях часто встречается случай, когда значения непрерывной случайной величины X заполняют не всю числовую ось, а только некоторый отрезок $[a, b]$. Дадим следующее

Определение. Абсолютно непрерывная случайная величина X называется **сосредоточенной на отрезке** $[a, b]$, если функция плотности равна нулю вне этого промежутка.

Формулы (5.5) и (5.3) в этом случае примут следующий вид

$$\int_a^b f(x)dx = 1. \quad (5.9)$$

$$F(x) = \int_a^x f(x)dx, \quad a \leq x \leq b. \quad (5.10)$$

В заключение этого параграфа коснемся вопроса о взаимоотношениях между двумя классами случайных величин – дискретными и непрерывными случайными величинами.

Самое существенное для понимания этих взаимоотношений заключается в том, что любую непрерывную случайную величину можно получить предельным переходом из дискретных случайных величин.

Наглядное представление об этом дает рисунок 5.3, на котором ступенчатая ломаная изображает график функции распределения $F(x)$ для дискретной случайной величины X . Представим себе, что к множеству возможных значений величины X добавляются все новые точки, так что число ступенек на кривой $F(x)$ будет становиться все больше, а сами ступеньки – все мельче. Тогда ступенчатая ломаная будет все более приближаться к плавной кривой, а дискретная случайная величина – к непрерывной.

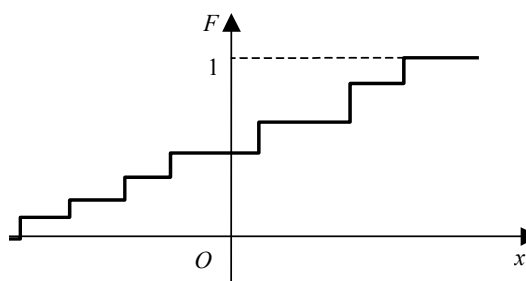


Рис. 5.3

§ 5.2. Типичные законы распределения

В этом параграфе мы ознакомимся с важными примерами непрерывных случайных величин. Соответствующие им законы распределения носят названия: *закон равномерного распределения на отрезке, показательный и нормальный законы.*

1°. Закон равномерного распределения на отрезке. Мы скажем, что случайная величина X , сосредоточенная на отрезке $[a, b]$, **равномерно распределена** на этом отрезке, если ее функция плотности равна константе:

$$f(x) = c(\text{const}), \quad a \leq x \leq b. \quad (5.11)$$

Значение постоянной c определяется из условия:

$$\int_{-\infty}^{\infty} f(x)dx = 1, \quad (5.12)$$

которому удовлетворяет любая плотность вероятности. В данном случае это условие принимает вид:

$$c(b-a) = 1,$$

откуда следует:

$$c = \frac{1}{b-a}.$$

График $f(x)$ изображен на рисунке 5.4.

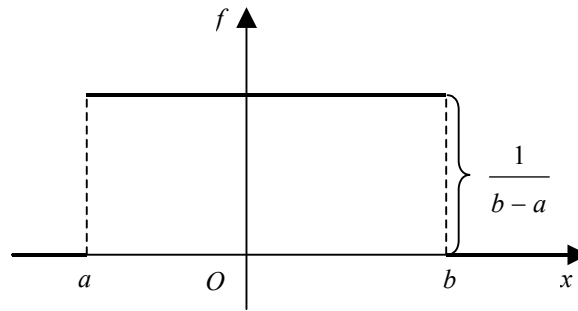


Рис. 5.4

Если какой-либо отрезок $[\alpha, \beta]$ целиком содержится в $[a, b]$, то вероятность попадания в него случайной величины X равна:

$$\int_{\alpha}^{\beta} f(x) dx = \int_{\alpha}^{\beta} c dx = cl = \frac{l}{b-a}, \quad (5.13)$$

где l – длина отрезка $[\alpha, \beta]$. Таким образом, вероятность попадания в любую часть отрезка $[a, b]$ пропорциональна длине этой части. В то же время вероятность попадания в любой отрезок, не перекрывающийся с $[a, b]$, очевидно, равна нулю.

Заметим, что формулу (5.12) можно записать в виде

$$P(\alpha \leq X \leq \beta) = \frac{\text{длина } [\alpha, \beta]}{\text{длина } [a, b]}, \quad (5.14)$$

делающем ее особенно "геометрической".

Охарактеризуем в общих чертах круг задач, приводящих к равномерным распределениям.

Можно сказать, что с равномерным распределением мы сталкиваемся всякий раз, когда по условиям опыта величина X принимает значения в конечном промежутке $[a, b]$, причем все значения из $[a, b]$ возможны в одинаковой степени, ни одно из них не имеет в этом смысле преимуществ перед другими. Вот несколько примеров такого рода:

1. X – время ожидания на стоянке автобуса (величина X распределена равномерно на отрезке $[0, l]$, где l – интервал движения между автобусами).
2. X – ошибка при взвешивании случайно выбранного предмета, получающаяся от округления результата взвешивания до ближайшего целого числа (величина X имеет равномерное распределение на отрезке $\left[-\frac{1}{2}, \frac{1}{2}\right]$, где за единицу принята цена деления шкалы).
3. X – угол между фиксированным радиусом приведенного во вращение волчка (рис. 5.5) и радиусом, идущим в точку касания волчком пола после остановки (величина X распределена равномерно на отрезке $[0, 2\pi]$).

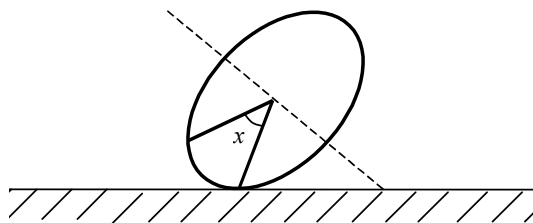


Рис. 5.5

Пример 5.1. Найти функцию распределения случайной величины X , распределенной равномерно на отрезке $[a, b]$.

Решение. Связь между функцией распределения и плотностью вероятности дается формулой (5.9):

$$F(x) = \int_a^x f(t) dt.$$

Подставляя сюда функцию $f(t)$, получим:

для $x < a$ $F(x) = 0,$

для $a \leq x < b$ $F(x) = \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a},$

для $x > b$ $F(x) = 1.$

График $F(x)$ показан на рисунке 5.6.

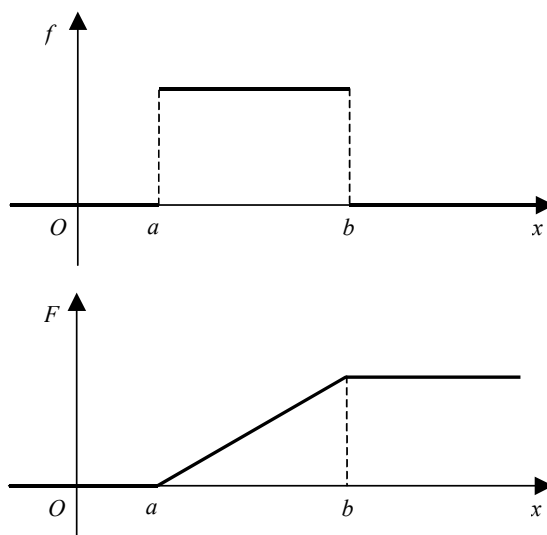


Рис. 5.6

2°. Показательный закон. Говорят, что случайная величина X , принимающая только неотрицательные значения, распределена по *показательному закону*, если для некоторого положительного параметра $\lambda > 0$ функция плотности имеет вид:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0. \quad (5.15)$$

Графики функций плотности для значений параметра $\lambda = 1, 2, 0,5$ приведен на рисунке 5.7.

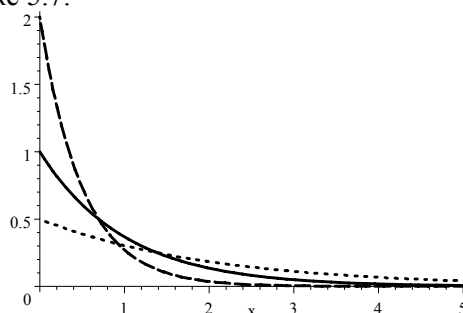


Рис. 5.7

Пример 5.2. Найти функцию распределения случайной величины, распределенной по показательному закону.

Решение. Функцию распределения найдем по формуле

$$F(x) = \int_0^x f(t) dt.$$

Подставляя выражение для функции плотности, получим

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x}.$$

Таким образом, для неотрицательных x выражение для функции распределения имеет вид:

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0. \quad (5.16)$$

Графики функций распределения для значений параметра $\lambda = 1, 2, 0,5$ приведены на рис. 5.8.

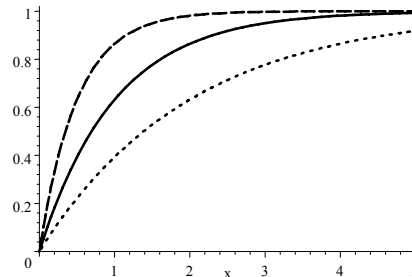


Рис. 5.8

Типичным примером непрерывной случайной величины, распределенной по показательному закону, может служить время между появлениями двух последовательных событий простейшего потока случайных событий. Например, такими потоками будут поток пассажиров, прибывающих на остановку, поток метеоритов в космическом пространстве, поток клиентов, приходящих в банк. Наиболее простой пример случайной величины, распределенной по показательному закону, – длительность службы электролампы.

3°. Закон нормального распределения на прямой. Закон нормального распределения на прямой, или, иначе, **нормальный закон**, – наиболее часто встречающийся на практике закон распределения. О причинах, обуславливающих его широкую распространенность в природе, технике, на производстве и т. д., мы расскажем в главе 11, пока же просто дадим описание этого закона.

Определение. Мы говорим, что непрерывная случайная величина X подчиняется **нормальному закону распределения**, если она имеет плотность вероятности следующего специального вида:

$$f(x) = A e^{-\lambda(x-a)^2},$$

где A , λ и a – постоянные, причем $A > 0$, $\lambda > 0$.

Соотношение (5.4), которому должна удовлетворять функция $f(x)$, показывает, что между параметрами A , λ и a должна существовать некоторая связь:

$$A \int_{-\infty}^{\infty} e^{-\lambda(x-a)^2} dx = 1.$$

Если в последнем интеграле произвести замену $\sqrt{\lambda}(x-a) = t$, то получим соотношение

$$\frac{A}{\sqrt{\lambda}} \int_{-\infty}^{\infty} e^{-t^2} dt = 1.$$

Поскольку интеграл от функции e^{-t^2} по всей оси равен $\sqrt{\pi}$, то имеем отсюда:

$$A \sqrt{\frac{\pi}{\lambda}} = 1. \quad (5.17)$$

Положим

$$\lambda = \frac{1}{2\sigma^2} \quad (\sigma > 0)$$

(т.е. введем вместо λ другой параметр σ), тогда из (5.17) будет следовать:

$$A = \frac{1}{\sigma\sqrt{2\pi}}.$$

Окончательно имеем:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \quad (5.18)$$

Это стандартная запись функции плотности нормального закона распределения. В нее входят два параметра: a и σ . Их теоретико-вероятностный смысл будет указан позже. Часто для краткой записи нормального закона распределения применяют обозначение $N(a, \sigma)$. *Стандартным нормальным законом* $N(0, 1)$ называют случай, когда $a = 0$, $\sigma = 1$.

На рисунке 5.9 показаны графики функции $f(x)$ для различных значений параметра σ . Каждый из них имеет вид холма, склоны которого полого спускаются к оси Ox . Максимальная высота холма равна $\frac{1}{\sigma\sqrt{2\pi}}$; она достигается при $x = a$. Если увеличивать значение параметра σ , то вершина холма

будет опускаться, но зато поднимутся склоны (ведь общая площадь между графиком и осью Ox должна оставаться равной 1). Что касается параметра a , то его значение не влияет на форму графика; с изменением a график только смещается в направлении оси Ox .

Методами дифференциального исчисления нетрудно проверить, что график $f(x)$ имеет две точки перегиба: одна отвечает значению $x = a + \sigma$, другая – значению $x = a - \sigma$ (рис. 5.10).

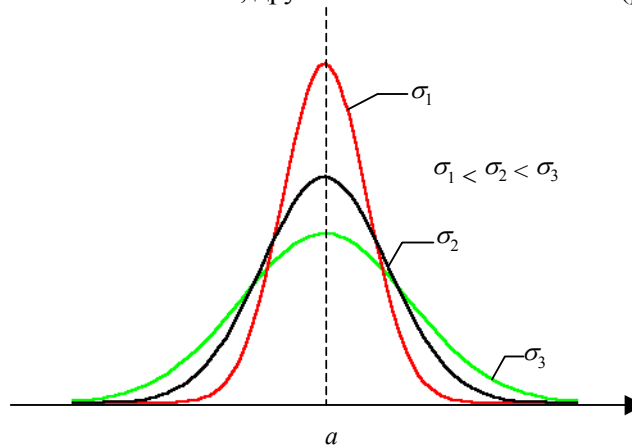


Рис. 5.9

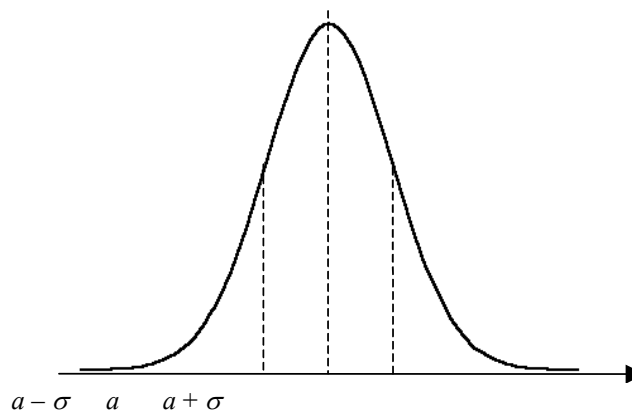


Рис. 5.10

Пример 5.3. Найти функцию распределения нормальной случайной величины X .

Решение. Пусть случайная величина X распределена по нормальному закону с плотностью (5.17). Для нахождения функции распределения воспользуемся общей формулой

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt.$$

Если в последнем интеграле произвести замену $\frac{t-a}{\sigma} = z$, то получим

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

$$= 0,5 + \Phi\left(\frac{x-a}{\sigma}\right),$$

где $\Phi(x)$ есть функция Лапласа, введенная в § 2.3:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

Таким образом, справедлива формула

$$F(x) = 0,5 + \Phi\left(\frac{x-a}{\sigma}\right), \quad (5.19)$$

которая и решает поставленную задачу.

Интересно отметить, что функция Лапласа является универсальным средством для решения всевозможных задач, связанных с нормальным законом распределения. В частности, из формулы (5.18) нетрудно получить *вероятность попадания нормальной случайной величины $N(a, \sigma)$ в данный промежуток $[\alpha, \beta]$* :

$$P(\alpha \leq X \leq \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \quad (5.20)$$

Формула (5.19) получается непосредственно из формулы (5.18) согласно общему правилу:

$$P(\alpha \leq X \leq \beta) = F(\beta) - F(\alpha).$$

В частности, вероятность попадания в отрезок $[a-3\sigma, a+3\sigma]$ равна:

$$\Phi(3) - \Phi(-3) = 2\Phi(3) = 0,997\dots;$$

как мы видим, эта вероятность отличается от 1 на весьма малую величину. Отсюда следует, что событие $a-3\sigma \leq X \leq a+3\sigma$ является практически достоверным, т.е. что практически возможные значения величины X расположены на отрезке $[a-3\sigma, a+3\sigma]$. Этот факт носит название "правило трех сигм".

§ 5.3. Числовые характеристики непрерывных случайных величин

1°. ε – приближение к случайной величине. Математическое ожидание случайной величины общего вида.

Прежде чем дать определение математического ожидания в общем случае, введем понятие ε – *приближение к случайной величине X* .

Пусть ε – положительное число. Рассмотрим на числовой оси бесконечную в обе стороны последовательность точек

$$\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots, \quad (5.21)$$

где $x_i = i\varepsilon$ (рис. 5.11). Условимся называть такую последовательность ε – *сетью*.

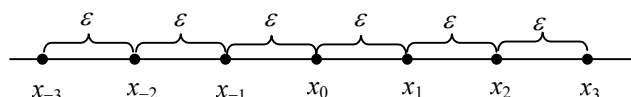


Рис. 5.11

Определим новую случайную величину X_ε , которая принимает дискретный ряд значений (5.20), причем вероятность любого из x_i есть

$$p_i = P(x_i \leq X < x_{i+1}) \quad (i = 0, \pm 1, \pm 2, \dots).$$

Случайную величину X_ε будем называть ε – *приближением* к случайной величине X .

Заметим, что условие $\sum p_i = 1$, обязательное для дискретного распределения, выполняется, поскольку

$$\sum_{i=-\infty}^{i=\infty} p_i = P(-\infty < X < \infty) = 1.$$

Если предположить, что величина X связана с некоторым опытом, то взаимоотношения между X и X_ε можно истолковать следующим образом. Пусть произведен опыт, и в результате него величина X приняла значение, принадлежащее некоторому промежутку $x_i \leq X < x_{i+1}$. Тогда величина X_ε , по определению, принимает в том же опыте значение x_i . При таком соглашении события $X_\varepsilon = x_i$ и $x_i \leq X < x_{i+1}$ будут равнозначны, и, следовательно, их вероятности будут совпадать; это как раз и соответствует данному выше определению величины X_ε . Итак, величина X_ε получается путем округления значений, принимаемых величиной X , до ближайших слева точек ε – сети.

Отсюда, между прочим, видно, что переход от случайной величины X к дискретной случайной величине X_ε имеет вполне реальный смысл. Например, такой переход совершается всякий раз, когда значения величины X считываются с измерительного прибора, причем показания прибора округляются до ближайшего слева деления шкалы.

Наши рассуждения показывают, что по мере приближения ε к нулю различие между X_ε и X становится все менее существенным. Естественно поэтому принять такое определение.

Определение. *Математическим ожиданием* случайной величины X называется число

$$M(X) = \lim_{\varepsilon \rightarrow 0} M(X_\varepsilon). \quad (5.22)$$

Если указанный справа предел не существует, то математическое ожидание величины X также считается несуществующим.

Итак,

$$M(X) = \lim_{\varepsilon \rightarrow 0} \sum_{i=-\infty}^{i=\infty} x_i P(x_i \leq X < x_{i+1}), \quad (5.23)$$

где $x_i = i\varepsilon$ ($i = \dots, -2, -1, 0, 1, 2, \dots$).

Пусть $F(x)$ – функция распределения величины X . Принимая во внимание формулу

$$P(x_k \leq X < x_{k+1}) = F(x_{k+1}) - F(x_k),$$

можно записать определение (5.22) в следующем виде:

$$M(X) = \lim_{\varepsilon \rightarrow 0} \sum_{k=-\infty}^{\infty} x_k (F(x_{k+1}) - F(x_k)). \quad (5.24)$$

Напомним еще раз, что x_k обозначает число $k\varepsilon$ ($k \in Z$).

2°. Математическое ожидание случайной величины, имеющей плотность вероятности.

Справедлива следующая

Теорема 5.1. Пусть X – абсолютно непрерывная случайная величина с непрерывной функцией плотности $f(x)$. Тогда

$$M(X) = \int_{-\infty}^{\infty} xf(x)dx, \quad (5.25)$$

если только указанный интеграл сходится.

Доказывать формулу (5.25) мы не будем, а ограничимся лишь идеей доказательства. Заменяя в формуле (5.24) каждую разность $F(x_{k+1}) - F(x_k)$ по теореме Лагранжа:

$$F(x_{k+1}) - F(x_k) = F'(x_k^*) \Delta x_k = f(x_k^*) \Delta x_k,$$

где $x_k^* \in [x_k, x_{k+1}]$, получим

$$M(X) = \lim_{\varepsilon \rightarrow 0} \sum_{k=-\infty}^{k=\infty} x_k f(x_k^*) \Delta x_k.$$

Стоящая справа сумма напоминает интегральную сумму для интеграла (5.25). Можно показать, что при $\varepsilon \rightarrow 0$ (т.е. когда все $\Delta x_k \rightarrow 0$) эта сумма сходится к интегралу (5.25).

Формула (5.25) допускает важное обобщение.

Теорема 5.2. Пусть X – абсолютно непрерывная случайная величина с непрерывной функцией плотности $f(x)$, и пусть $g(x)$ – некоторая непрерывная функция. Тогда

$$M[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx. \quad (5.26)$$

Очевидно, теорема 5.1 есть частный случай теоремы 5.2, когда $g(x) = x$.

В случае, когда абсолютно непрерывная случайная величина X сосредоточена на промежутке $[a, b]$, для формул (5.25), (5.26) справедливы их аналоги:

$$M(X) = \int_a^b x f(x) dx, \quad (5.27)$$

$$M[g(X)] = \int_a^b g(x) f(x) dx. \quad (5.28)$$

3°. Дисперсия абсолютно непрерывной случайной величины.

Дисперсия абсолютно непрерывной случайной величины X с функцией плотности $f(x)$ и математическим ожиданием $m = M(X)$ определяется таким же равенством, как и для дискретной величины

$$D(X) = M(X - m)^2. \quad (5.29)$$

Из равенства (5.26) следует, что справедлива следующая формула

$$D(X) = \int_{-\infty}^{\infty} (x - m)^2 f(x) dx. \quad (5.30)$$

Поскольку формула (5.29) может быть записана в следующем виде

$$D(X) = M(X^2) - m^2, \quad (5.31)$$

то формулу (5.30) можно представить таким образом

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - m^2. \quad (5.32)$$

В случае когда абсолютно непрерывная случайная величина X сосредоточена на промежутке $[a, b]$, формулы (5.30), (5.32) примут вид

$$D(X) = \int_a^b (x - m)^2 f(x) dx, \quad (5.33)$$

$$D(X) = \int_a^b x^2 f(x) dx - m^2. \quad (5.34)$$

Дисперсия непрерывной случайной величины определяет степень рассеивания значений, принимаемых случайной величиной, вокруг ее математического ожидания.

Среднее квадратичное отклонение, или стандартное отклонение, непрерывной случайной величины X определяется так же, как и для дискретной случайной величины:

$$\sigma(X) = \sqrt{D(X)}. \quad (5.35)$$

4°. Свойства математического ожидания и дисперсии.

Для математического ожидания и дисперсии непрерывной случайной величины X сохраняются свойства числовых характеристик дискретной случайной величины. Напомним эти свойства.

1. Математическое ожидание суммы равно сумме математических ожиданий.

$$M(X + Y) = M(X) + M(Y). \quad (5.36)$$

2. Постоянный множитель выносится из под знака математического ожидания.

$$M(cX) = cM(X), \quad c = \text{const}. \quad (5.37)$$

3. Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий.

$$M(XY) = M(X)M(Y), \quad X \text{ и } Y \text{ независимы.} \quad (5.38)$$

4. Постоянный множитель выносится из под знака дисперсии в квадрате.

$$D(cX) = c^2 D(X), \quad c = \text{const}. \quad (5.39)$$

5. Если к случайной величине прибавить константу, то дисперсия не изменится.

$$D(c + X) = D(X), \quad c = \text{const}. \quad (5.40)$$

6. Дисперсия суммы независимых случайных величин равна сумме дисперсий этих величин.

$$D(X + Y) = D(X) + D(Y), \quad X \text{ и } Y \text{ независимы.} \quad (5.41)$$

По поводу доказательства свойств 1-6 сделаем одно общее замечание. Для дискретных случайных величин эти свойства были проверены. Математическое ожидание непрерывной случайной величины было определено с помощью предельного перехода ε – приближения дискретной случайной величины. Поскольку предельный переход не нарушает сформулированных свойств, то они сохраняются и для непрерывных случайных величин.

5°. Числовые характеристики типичных законов распределения.

Найдем математическое ожидание и дисперсию для основных законов распределения: равномерного, показательного и нормального.

Пример 5.4. Найти числовые характеристики случайной величины X , равномерно распределенной на отрезке $[a; b]$.

Решение. Для нахождения математического ожидания воспользуемся формулой (5.25):

$$M(X) = \int_a^b xf(x)dx = \int_a^b cdx,$$

где c – постоянная плотность вероятности на отрезке $[a; b]$, равная $\frac{1}{b-a}$.

Отсюда

$$M(X) = c \frac{x^2}{2} \Big|_a^b = c \frac{b^2 - a^2}{2} = \frac{b+a}{2}.$$

Мы получили таким образом, что числу $M(X)$ соответствует середина отрезка $[a; b]$. Ввиду равномерности распределения X на отрезке $[a; b]$ такой ответ можно было бы предвидеть заранее.

Для нахождения дисперсии воспользуемся формулой (5.34):

$$\begin{aligned} D(X) &= \frac{1}{b-a} \int_a^b x^2 dx - \left(\frac{a+b}{2} \right)^2 = \frac{1}{b-a} \frac{b^3 - a^3}{3} - \left(\frac{a+b}{2} \right)^2 = \\ &= \frac{b^2 + ab + a^2}{3} - \left(\frac{a+b}{2} \right)^2 = \frac{(b-a)^2}{12}. \end{aligned}$$

Таким образом, для равномерного закона найдены все числовые характеристики:

$$M(X) = \frac{a+b}{2}, \quad D(X) = \frac{(b-a)^2}{12}, \quad \sigma(X) = \frac{b-a}{2\sqrt{3}}. \quad (5.42)$$

Пример 5.5. Найти числовые характеристики случайной величины X , распределенной по показательному закону с плотностью $f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$, $x \geq 0$.

Решение. Для нахождения математического ожидания воспользуемся формулой (5.25):

$$M(X) = \int_0^{\infty} x f(x) dx = \int_0^{\infty} \lambda x e^{-\lambda x} dx. \quad (5.43)$$

Найдем интеграл (5.43) методом интегрирования по частям, полагая $u = x$, $dv = e^{-\lambda x} d(\lambda x)$, так что $du = dx$, $v = -e^{-\lambda x}$. Получим

$$M(X) = -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = -\frac{e^{-\lambda x}}{\lambda} \Big|_0^{\infty} = \frac{1}{\lambda}.$$

Таким образом, $M(X) = \frac{1}{\lambda}$. Чтобы найти дисперсию, воспользуемся формулой (5.32):

$$D(X) = \int_0^{\infty} x^2 f(x) dx - [M(X)]^2 = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2}. \quad (5.44)$$

Для нахождения интеграла (5.44) надо дважды применить формулу интегрирования по частям. Вычисления, полностью аналогичные соответствующим действиям для математического ожидания, предоставляется выполнить читателю. В результате получится $D(X) = \frac{1}{\lambda^2}$. Таким образом, мы нашли все искомые величины:

$$M(X) = \frac{1}{\lambda}, D(X) = \frac{1}{\lambda^2}, \sigma(X) = \frac{1}{\lambda}. \quad (5.45)$$

Пример 5.6. Найти математическое ожидание случайной величины X , распределенной по нормальному закону $N(a, \sigma)$ с плотностью

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

Решение. Имеем:

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

В последнем интеграле произведем замену переменной: $\frac{x-a}{\sigma} = t$.

Получим:

$$M(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t + a) e^{-\frac{t^2}{2}} dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dt + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt.$$

Первый из двух интегралов правой части равен нулю ввиду нечетности подынтегральной функции, второй же равен a по свойству нормированности функции плотности стандартного нормального закона $N(0, 1)$.

В итоге получаем:

$$M(X) = a. \quad (5.46)$$

Этим выяснен теоретико-вероятностный смысл параметра a , входящего в выражение для нормального закона: *параметр a совпадает с математическим ожиданием величины X .*

Равенство (5.46) можно было бы получить и более простым путем, используя одно соображение общего характера, а именно: если плотность $f(x)$ четна относительно точки $x = a$ (т.е.

$f(a-t) = f(a+t)$ для всех x), то $M(X) = a$. Действительно, если в интеграле $\int_{-\infty}^{\infty} xf(x)dx$ произвести за-

мену $x = a + t$, то получим интеграл $\int_{-\infty}^{\infty} (a+t)f(a+t)dt$ или

$$a \int_{-\infty}^{\infty} f(a+t)dt + \int_{-\infty}^{\infty} tf(a+t)dt, \quad (5.47)$$

но первый из двух интегралов (5.46) равен 1, а второй равен нулю (ввиду нечетности функции $f(a+t)$ относительно $t=0$). Следовательно, $M(X) = a$.

Поскольку в данном случае $M(X)$ совпадает с a , то для нахождения дисперсии удобнее воспользоваться формулой (5.30), которая в данном случае принимает следующий вид:

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Произведя замену $\frac{x-a}{\sigma} = t$, получим:

$$D(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt.$$

Интеграл, стоящий в правой части, равен¹ $\sqrt{2\pi}$. Следовательно,

$$D(X) = \sigma^2.$$

Тем самым мы выяснили смысл параметра σ , входящего в выражение для функции плотности нормального закона: *σ есть среднее квадратичное отклонение величины X .*

¹ Применяя интегрирование по частям, получим :

$$\int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt = - \int_{-\infty}^{\infty} t d \left(e^{-\frac{t^2}{2}} \right) = -te^{-\frac{t^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}.$$

Начальные и центральные моменты случайных величин

§ 6.1. Определение и основные свойства

Мы уже видели в предыдущей главе, какое большое значение имеют такие числовые характеристики распределения, как математическое ожидание и дисперсия. Обобщением этих понятий являются *начальные и центральные моменты*.

Определение. *Начальным моментом* порядка k ($k \in \mathbf{N}$) случайной величины X называют математическое ожидание k -й степени X :

$$\nu_k(X) = M(X^k). \quad (6.1)$$

Из определения ясно, что начальный момент первого порядка совпадает с математическим ожиданием.

$$\nu_1(X) = M(X). \quad (6.2)$$

Для дискретных случайных величин с законом распределения $P(X = x_i) = p_i$ ($i \in \mathbf{N}$) формула для начального момента порядка k выглядит следующим образом:

$$\nu_k(X) = \sum_{k=1}^{\infty} x^k p_k, \quad (6.3)$$

для абсолютно непрерывных случайных величин с плотностью $f(x)$:

$$\nu_k(X) = \int_{-\infty}^{\infty} x^k f(x) dx, \quad (6.4)$$

если ряд (6.3) и интеграл (6.4) сходятся абсолютно.

Определение. *Центральным моментом* порядка k ($k \in \mathbf{N}$) случайной величины X называют математическое ожидание k -й степени отклонения $\overset{\circ}{X} = X - m$, где m – математическое ожидание X :

$$\mu_k(X) = M\left(\overset{\circ}{X}^k\right). \quad (6.5)$$

Мы уже видели, что первый центральный момент равен нулю, а второй центральный момент совпадает с дисперсией случайной величины X :

$$\mu_1(X) = 0, \mu_2(X) = D(X). \quad (6.6)$$

Для дискретных случайных величин формула для центрального момента порядка k выглядит следующим образом:

$$\mu_k(X) = \sum_{k=1}^{\infty} (x - m)^k p_k, \quad (6.7)$$

для непрерывных случайных величин

$$\mu_k(X) = \int_{-\infty}^{\infty} (x - m)^k f(x) dx. \quad (6.8)$$

Разлагая бином $\overset{\circ}{X}^k = (X - m)^k$ и пользуясь свойствами математического ожидания, нетрудно получить формулы, выражающие центральные моменты через начальные. Например,

$$(X - m)^2 = X^2 - 2mX + m^2,$$

откуда

$$\mu_2 = M[(X - m)^2] = M(X^2 - 2mX + m^2) =$$

$$= M(X^2) - 2mM(X) + m^2 = \nu_2 - \nu_1^2,$$

так что

$$\mu_2 = \nu_2 - \nu_1^2. \quad (6.9)$$

Полученная формула полностью идентична с уже известной формулой для дисперсии. Аналогично получаются следующие формулы, вывод которых предоставляется читателю:

$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3, \quad (6.10)$$

$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4. \quad (6.11)$$

Отметим следующие свойства третьего и четвертого центральных моментов.

Свойство. Если величины X и Y независимы, то

$$\mu_3(X + Y) = \mu_3(X) + \mu_3(Y). \quad (6.12)$$

Доказательство. Используя формулу куба суммы и свойство математического ожидания, имеем

$$\begin{aligned} \mu_3(X + Y) &= M[(X + Y - m_X - m_Y)^3] = M[(X - m_X) + (Y - m_Y)]^3 = \\ &= M[(X - m_X)^3] + 3M[(X - m_X)^2(Y - m_Y)] + \\ &+ 3M[(X - m_X)(Y - m_Y)^2] + M[(Y - m_Y)^3]. \end{aligned}$$

По теореме умножения математического ожидания для независимых величин получим

$$\mu_3(X + Y) = \mu_3(X) + 3\mu_2(X)\mu_1(Y) + 3\mu_1(X)\mu_2(Y) + \mu_3(Y).$$

Поскольку центральные моменты первого порядка равны нулю, то формула (6.12) отсюда следует непосредственно.

Доказанная формула непосредственно обобщается на произвольное число независимых слагаемых:

$$\mu_3(X_1 + X_2 + \dots + X_n) = \mu_3(X_1) + \mu_3(X_2) + \dots + \mu_3(X_n). \quad (6.13)$$

Для моментов четвертого порядка аналогичное свойство выглядит более сложным образом.

Свойство. Для независимых случайных величин X и Y выполняется формула

$$\mu_4(X + Y) = \mu_4(X) + \mu_4(Y) + 6D(X)D(Y). \quad (6.14)$$

Доказательство полностью аналогично предыдущему и оставляется читателю. Для произвольного числа независимых слагаемых формула (6.14) приобретает следующий вид:

$$\mu_4\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mu_4(X_i) + 6 \sum_{\substack{i,j=1 \\ i < j}}^n D(X_i)D(X_j). \quad (6.15)$$

§ 6.2. Вычисление моментов. Производящие функции

Рассмотрим ряд примеров нахождения моментов высших порядков для типичных законов распределения.

Пример 6.1. Найти все центральные моменты до четвертого порядка включительно для случайной величины X , распределенной по биномиальному закону с параметрами $n, p, q = 1 - p$.

Решение. Что касается нахождения центральных моментов первого и второго порядка, то ответ нам известен:

$$\mu_1(X) = 0, \mu_2(X) = D(X) = npq.$$

Чтобы найти μ_3 , воспользуемся формулой (6.13) и представлением случайной величины X в виде суммы индикаторов появления события A в i -м испытании:

$$X = X_1 + X_2 + \dots + X_n.$$

Для каждого слагаемого X_i третий центральный момент находится непосредственно

$$\mu_3(X_i) = (0-p)^3 q + (1-p)^3 p = -p^3 q + pq^3 = pq(p^2 - q^2) = pq(1-2p).$$

Поэтому

$$\mu_3(X) = npq(1-2p). \quad (6.16)$$

Для нахождения центрального момента четвертого порядка воспользуемся формулой (6.15). Вначале найдем моменты для каждого слагаемого:

$$\begin{aligned} \mu_4(X_i) &= p^4 q + q^4 p = pq(p^3 + q^3) = pq(p+q)(p^2 - pq + q^2) = \\ &= pq((p+q)^2 - 3pq) = pq(1-3pq). \end{aligned}$$

Отсюда

$$\mu_4(X) = npq(1-3pq) + 6 \frac{n(n-1)}{2} pq,$$

откуда окончательно получаем

$$\mu_4(X) = npq[1 + 3(n-2)pq]. \quad (6.17)$$

Для нахождения всех начальных моментов непрерывной случайной величины проще использовать производящую функцию моментов.

Определение. Производящей функцией моментов случайной величины X называют математическое ожидание случайной величины e^{sX} , где s – действительный параметр:

$$\psi_X(s) = M(e^{sX}). \quad (6.18)$$

Производящая функция моментов позволяет находить все начальные моменты случайной величины X . Справедлива следующая теорема.

Теорема 6.1. Если случайная величина X имеет начальный момент порядка n , то производящая функция $\psi_X(s)$ n раз дифференцируема по s , и для всех $k \leq n$ выполняется соотношение

$$\nu_k = \psi_X^{(k)}(0). \quad (6.19)$$

Если оставить в стороне вопрос о дифференцируемости данной функции, то соотношение (6.20) получается формально дифференцированием соотношения (6.19) k раз по s

$$\psi_X^{(k)}(s) = M(X^k e^{sX}).$$

Подставляя в последнее соотношение $s = 0$, получим

$$\psi_X^{(k)}(0) = M(X^k) = \nu_k,$$

откуда непосредственно вытекает равенство (6.19).

Сравнивая разложение производящей функции моментов в ряд Маклорена с равенствами (6.19), получим ее разложение в ряд с использованием начальных моментов:

$$\psi_X(s) = 1 + \nu_1 s + \frac{\nu_2}{2!} s^2 + \frac{\nu_3}{3!} s^3 + \dots \quad (6.20)$$

Пример 6.2. Найти производящую функцию моментов для случайной величины X , распределенной по нормальному закону $N(a, \sigma)$.

Решение. Имеем по определению

$$\psi_X(t) = \int_{-\infty}^{\infty} f(x) e^{sx} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Для вычисления этого интеграла рассмотрим отдельно показатель экспоненты, стоящей под знаком интеграла. Имеем

$$sx - \frac{(x-a)^2}{2\sigma^2} = -\frac{x^2 - 2(a + \sigma^2 s)x + a^2}{2\sigma^2}.$$

Выделим в числителе последней дроби полный квадрат по x . После небольшого числа тождественных преобразований получим:

$$-\frac{(x-a-\sigma^2 s)^2}{2\sigma^2} + as + \frac{\sigma^2 s^2}{2}.$$

Таким образом, выражение для характеристической функции приобретет вид:

$$\psi_X(s) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-a-\sigma^2 s)^2}{2\sigma^2}} dx \cdot e^{as + \frac{\sigma^2 s^2}{2}}.$$

Поскольку первый множитель равен 1 как интеграл от плотности нормальной случайной величины $N(a + \sigma^2 s, \sigma)$, то окончательное выражение для характеристической функции нормального закона имеет вид:

$$\psi_X(s) = e^{as + \frac{\sigma^2 s^2}{2}}. \quad (6.21)$$

Особенно простой вид имеет характеристическая функция стандартного нормального закона $N(0, 1)$:

$$\psi_X(s) = e^{\frac{s^2}{2}}. \quad (6.22)$$

Теперь нам будет нетрудно найти все начальные моменты стандартной нормальной случайной величины. Вместо того чтобы дифференцировать функцию (6.22), можно разложить ее в ряд и сравнить с рядом (6.20):

$$\psi_X(s) = 1 + \frac{s^2}{2} + \frac{s^4}{2^2 \cdot 2!} + \dots = 1 + \frac{\nu_2}{2!} t^2 + \frac{\nu_4}{4!} t^4 + \dots,$$

откуда имеем выражения для начальных моментов:

$$\nu_2 = 1, \nu_4 = \frac{4!}{2^2 \cdot 2!} = 3, \dots, \nu_{2k} = \frac{(2k)!}{2^k \cdot k!}. \quad (6.23)$$

§ 6.3. Моменты высших порядков и числовые характеристики закона распределения

Математическое ожидание и дисперсия определяют наиболее важные черты закона распределения: его положение и степень рассеянности. Для более подробной его характеристики используют моменты высшего порядка и такие числовые характеристики, как *асимметрию* и *эксцесс*.

Определение. *Асимметрией* распределения называют отношение третьего центрального момента к кубу стандартного отклонения:

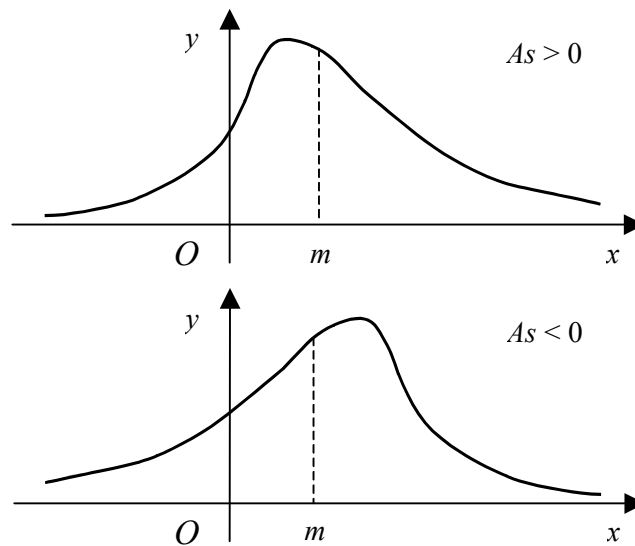
$$As = \frac{\mu_3}{\sigma^3}. \quad (6.24)$$

Замечание. Асимметрия случайной величины X совпадает с третьим начальным (центральным) моментом соответствующей нормированной случайной величины.

Действительно, по определению

$$\begin{aligned} As(X) &= \frac{M[(X-m)^3]}{\sigma^3} = M\left[\left(\frac{X-m}{\sigma}\right)^3\right] = \\ &= \nu_3\left(\frac{X-m}{\sigma}\right) = \mu_3\left(\frac{X-m}{\sigma}\right). \end{aligned} \quad (6.25)$$

Поскольку все центральные моменты нечетных порядков равны нулю для распределений, симметричных относительно математического ожидания, то асимметрия в этом случае равна нулю. Далее изображены два типичных случая непрерывных распределений. В первом случае $As > 0$ и кривая распределения скошена влево относительно вертикальной прямой $x = m$, где m – математическое ожидание. Во втором асимметрия отрицательна и кривая распределения скошена вправо.



Из формулы (6.16) легко получить формулу для асимметрии биномиального закона:

$$As = \frac{1-2p}{\sqrt{npq}}. \quad (6.26)$$

В частности, мы видим, что в этом случае асимметрия положительна, если $p < 1/2$, и отрицательна, если $p > 1/2$. Например, при $n = 100$, $p = 0,2$, $q = 0,8$, имеем $As = 0,15$; при $n = 100$, $p = 0,8$, $q = 0,2$, $As = -0,15$.

В качестве примера вычисления асимметрии непрерывной случайной величины найдем ее для показательного закона.

Воспользуемся формулой (6.24). Поскольку для показательного закона $M(X) = \sigma(X) = \frac{1}{\lambda}$, то имеем

$$As(X) = \int_0^{\infty} \left(\frac{x - \frac{1}{\lambda}}{\frac{1}{\lambda}} \right)^3 \lambda e^{-\lambda x} dx = \int_0^{\infty} (\lambda x - 1)^3 \lambda e^{-\lambda x} dx.$$

Выполним замену переменной $t = \lambda x - 1$. Интеграл примет следующий вид:

$$As(X) = \int_{-1}^{\infty} t^3 e^{-t-1} dt = e^{-1} \int_{-1}^{\infty} t^3 e^{-t} dt. \quad (6.27)$$

Первообразную для последнего интеграла проще всего найти методом неопределенных коэффициентов:

$$\int t^3 e^{-t} dt = (At^3 + Bt^2 + Ct + D)e^{-t}.$$

Дифференцируя обе части и приравнявая коэффициенты при одинаковых степенях t , получим

$$\int t^3 e^{-t} dt = -(t^3 + 3t^2 + 6t + 6)e^{-t}.$$

Подставляя найденный интеграл в (6.26), получим окончательно

$$As(X) = 2. \quad (6.28)$$

Определение. Эксцессом распределения называется величина

$$Ex = \frac{\mu_4}{\sigma^4} - 3. \quad (6.29)$$

Поскольку для стандартного нормального распределения $N(0, 1)$ мы нашли, что $\mu_4 = 3$ (см. (6.22)), то для нормального распределения эксцесс равен нулю. В частности, вычисляя эксцесс неизвестного распределения, мы можем судить о близости его к нормальному по этой числовой характеристике.

Для биномиального закона

$$Ex = \frac{1 - 6pq}{npq}. \quad (6.30)$$

Действительно, воспользуемся формулой (6.18). Имеем

$$Ex = \frac{npq[1 + 3(n-2)pq]}{(npq)^2} - 3 = \frac{1 + 3npq - 6pq}{npq} - 3 = \frac{1 - 6pq}{npq},$$

что и требовалось. В частности, при $n = 100$, $p = 0,2$, $q = 0,8$ имеем $Ex = 0,0025$. Это еще раз нам подтверждает то обстоятельство, что биномиальный закон при больших n хорошо согласуется с нормальным законом.

Наконец, для показательного закона вычисление эксцесса может быть выполнено способом, аналогичным тому, который был использован для вычисления асимметрии, и оставляется читателю.

Завершая рассмотрение асимметрии и эксцесса, сформулируем ряд их общих свойств.

Свойство 1. Асимметрия и эксцесс инвариантны относительно линейной замены случайной величины:

$$As(aX + b) = As(X), \quad (6.31)$$

$$Ex(aX + b) = Ex(X). \quad (6.32)$$

Данное свойство следует из того, что для любого натурального числа k выполняется равенство

$$\mu_k(aX + b) = M\{[aX + b - M(aX + b)]^k\} = a^k \mu_k(X).$$

Таким образом, асимметрия и эксцесс не меняются при сдвигах и растяжениях и их можно использовать в качестве характеристик формы распределения.

Свойство 2. Для независимых случайных величин $X_1, \dots, X_n$ имеем

$$As(X_1 + \dots + X_n) = a_1 As(X_1) + \dots + a_n As(X_n), \quad (6.33)$$

$$Ex(X_1 + \dots + X_n) = b_1 Ex(X_1) + \dots + b_n Ex(X_n), \quad (6.34)$$

где

$$a_i = \frac{\sigma^3(X_i)}{(\sigma^2(X_1) + \dots + \sigma^2(X_n))^{\frac{3}{2}}}, \quad b_i = \frac{\sigma^4(X_i)}{(\sigma^2(X_1) + \dots + \sigma^2(X_n))^2},$$

$i = 1, \dots, n$.

В частности, когда все $X_1, \dots, X_n$ имеют одинаковую дисперсию, то

$$As(X_1 + \dots + X_n) = n^{-\frac{3}{2}} (As(X_1) + \dots + As(X_n)), \quad (6.35)$$

$$Ex(X_1 + \dots + X_n) = n^{-2} (Ex(X_1) + \dots + Ex(X_n)). \quad (6.36)$$

Свойство 2 вытекает из доказанных ранее свойств центральных моментов μ_3 и μ_4 (см. формулы (6.13) и (6.15)).

Заметим, что в случае одинаково распределенных независимых случайных величин $X_1, \dots, X_n$ асимметрия и эксцесс их суммы стремятся к нулю, когда $n \rightarrow \infty$:

$$As(X_1 + \dots + X_n) = n^{-\frac{1}{2}} As(X_1) \rightarrow 0,$$

$$Ex(X_1 + \dots + X_n) = n^{-1} Ex(X_1) \rightarrow 0.$$

Случайные векторы (многомерные случайные величины)

В практических применениях теории вероятности очень часто приходится сталкиваться с задачами, когда результат опыта описывается не одной случайной величиной, а двумя или более, которые образуют *систему*. При этом такая система в общем случае не представляет собой механический набор нескольких случайных величин, она образует некоторый объект, в характеристике которого имеет значение и то, как эти случайные величины взаимодействуют между собой. В § 3.5, 3.6 мы уже затрагивали этот вопрос для дискретных случайных величин, теперь мы хотим рассмотреть его в самом общем случае.

Для простоты изложения будем систематически рассматривать двумерный случай.

§ 7.1. Случайные векторы

Определение. Упорядоченная пара случайных величин (X, Y) , определенных на одном и том же пространстве элементарных событий Ω , называется **системой случайных величин, двумерным случайным вектором** или **двумерной случайной величиной**.

Систему случайных величин можно рассматривать как случайную точку на координатной плоскости, либо как *случайный вектор* (см. рис. 7.1, 7.2).

Для обычной случайной величины X случайным событием являлось любое множество элементарных событий, удовлетворяющих условию $X \in B$, где B – борелевское множество на прямой.

Подмножество $B \subseteq \mathbf{R}^2$ числовой плоскости называется *борелевским*, если оно может быть получено из открытых или замкнутых множеств в $\mathbf{R}^2$ с помощью конечного или счетного числа теоретико-множественных операций. Борелевскими множествами являются: точки, прямые, открытые и замкнутые многоугольники, полуплоскости, круги и т.д. Так же как и в одномерном случае, ситуация упрощается с помощью рассмотрения функции распределения.

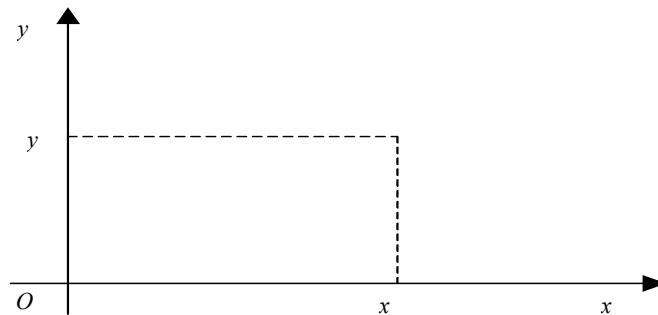


Рис. 7.1

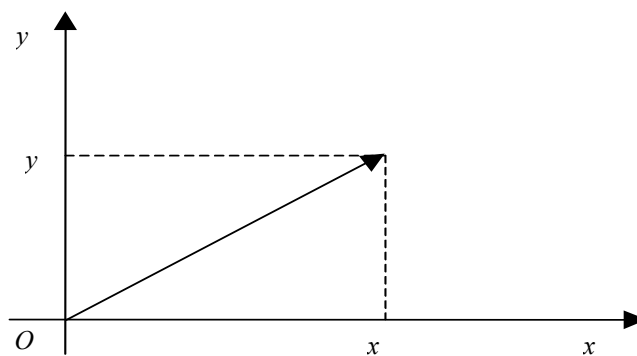


Рис. 7.2

Определение. *Функцией распределения* двумерной случайной величины (X, Y) называется *вероятность совместного выполнения двух событий $X < x$ и $Y < y$*

$$F_{X,Y}(x, y) = P(X < x, Y < y). \quad (7.1)$$

Геометрической интерпретацией (7.1) может служить рис. 7.3, на котором значением функции распределения может служить вероятность попадания случайной величины (X, Y) в бесконечный квадрант $Q(x, y)$ с вершиной в точке (x, y) , лежащий левее и ниже ее.

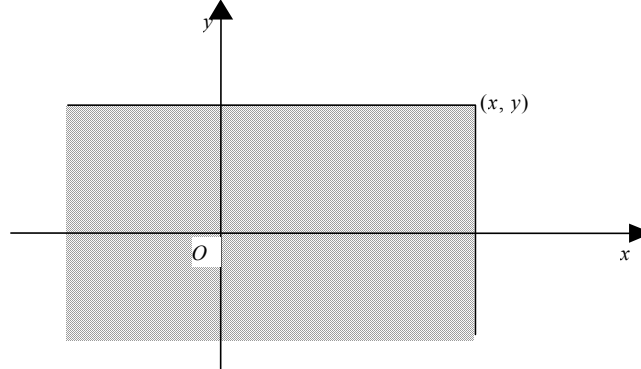


Рис. 7.3

Случайные векторы (X_1, Y_1) и (X_2, Y_2) *одинаково распределены*, когда их функции распределения совпадают. Для одинаково распределенных случайных векторов (X_1, Y_1) и (X_2, Y_2) вероятность попадания точек (X_1, Y_1) и (X_2, Y_2) в какое-либо борелевское множество $B \subseteq \mathbf{R}^2$ одна и та же:

$$P\{(X_1, Y_1) \in B\} = P\{(X_2, Y_2) \in B\}.$$

Свойства функции распределения.

1. $F_{X,Y}(x, y)$ не убывает по обоим аргументам, т.е.

$$\text{при } x_1 < x_2 \quad F(x_1, y) \leq F(x_2, y),$$

$$\text{при } y_1 < y_2 \quad F(x, y_1) \leq F(x, y_2).$$

2. $F_{X,Y}(x, y)$ непрерывна слева по обоим аргументам, т.е.

$$\lim_{x \rightarrow x_0 + 0} F(x, y) = F(x_0, y),$$

$$\lim_{y \rightarrow y_0 + 0} F(x, y) = F(x, y_0).$$

3. $P(x_1 \leq X < x_2, y_1 \leq Y < y_2) =$

$$= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1).$$

4. $\lim_{x \rightarrow -\infty} F_{X,Y}(x, y) = \lim_{y \rightarrow -\infty} F_{X,Y}(x, y) = 0.$

5. $\lim_{x \rightarrow +\infty} F_{X,Y}(x, y) = F_Y(y), \quad \lim_{y \rightarrow +\infty} F_{X,Y}(x, y) = F_X(x).$

6. $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F_{X,Y}(x, y) = 1.$

Первое свойство очевидно, поскольку при

$$x_1 < x_2 \quad Q(x_1, y) \subset Q(x_2, y).$$

Доказательство свойства 2, по сути дела, копирует доказательство непрерывности слева функции распределения случайной величины и оставляется читателю.

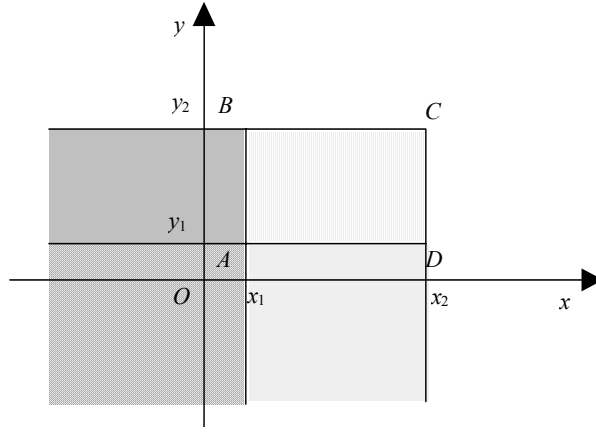


Рис. 7.4

Для доказательства свойства 3 обозначим через R прямоугольник с вершинами $A(x_1, y_1)$, $B(x_1, y_2)$, $C(x_2, y_2)$, $D(x_2, y_1)$ (см. рис. 7.4). Тогда выполняются два равенства

$$Q(x_2, y_2) = R \cup (Q(x_2, y_2) \setminus R), \quad (7.2)$$

$$Q(x_2, y_2) \setminus R = Q(x_1, y_2) \cup Q(x_2, y_1). \quad (7.3)$$

Применяя теорему сложения вероятностей (заметим, что $Q(x_1, y_2) \cap Q(x_2, y_1) = Q(x_1, y_1)$) к (7.2) и (7.3), получим

$$F(x_2, y_2) = P(x_1 \leq X < x_2, y_1 \leq Y < y_2) + F(x_1, y_2) + F(x_2, y_1) - F(x_1, y_1),$$

откуда свойство 3 следует непосредственно.

Для доказательства свойства 4 рассмотрим произвольную убывающую последовательность $x_n \rightarrow -\infty$. Тогда очевидно, что $\bigcap_{n=1}^{\infty} Q_{x_n, y} = \emptyset$ и мы можем воспользоваться свойством вероятности

$$\lim_{n \rightarrow \infty} P(Q_{x_n, y}) = P\left(\bigcap_{n=1}^{\infty} Q_{x_n, y}\right) = 0.$$

Доказательство свойств 5 и 6 аналогично доказательству свойства 4 и оставляется читателю. Особое значение играют *абсолютно непрерывные случайные векторы*.

§ 7.2. Абсолютно непрерывные случайные векторы

Определение. Случайный вектор (X, Y) называется **абсолютно непрерывным**, если найдется неотрицательная функция $f_{X,Y}(x, y)$, называемая **плотностью распределения**, такая, что для любого множества $G \subset \mathbf{R}^2$, для которого существует интеграл $\iint_G f_{X,Y}(x, y) dx dy$, вероятность попадания точки (X, Y) в G находится по формуле

$$P\{(X, Y) \in G\} = \iint_G f_{X,Y}(x, y) dx dy. \quad (7.4)$$

Особенно часто используется случай, когда область G – прямоугольник: $\{a \leq x \leq b, c \leq y \leq d\}$. В этом случае формула (7.4) принимает вид

$$P((X, Y) \in G) = \int_c^d \int_a^b f_{X,Y}(x, y) dx dy. \quad (7.5)$$

Из свойств интеграла следует, что если (X, Y) – абсолютно непрерывный случайный вектор, то вероятность попадания точки (X, Y) в какую-либо линию (прямую, гиперболу, ...) равна 0.

Функция распределения $F_{X,Y}(x, y)$ абсолютно непрерывного случайного вектора (X, Y) является непрерывной и может быть представлена согласно (7.5) в виде несобственного интеграла

$$F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(s,t) ds dt. \quad (7.6)$$

Плотность распределения обладает следующими свойствами:

1. $f_{X,Y}(x,y) \geq 0, (x,y) \in \mathbf{R}^2$ (неотрицательность).
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1$ (условие нормировки).
3. $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$ в точке непрерывности $f_{X,Y}(x,y)$.

Первое свойство выполняется по определению. Второе свойство следует из свойства 6 функции распределения. Свойство 3 получается, если найти смешанную производную от обеих частей равенства (7.6).

Компоненты X, Y абсолютно непрерывного случайного вектора (X, Y) являются также абсолютно непрерывными. Плотности распределения $f_X(x), f_Y(y)$ случайных величин X и Y выражаются через плотность совместного распределения:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx. \quad (7.7)$$

Формулы (7.7) непосредственно вытекают из свойства 5 функции распределения.

Случайный вектор (X, Y) называется *сосредоточенным на множестве* $G \subset \mathbf{R}^2$, если $P\{(X,Y) \in G\} = 1$. В этом случае все пределы интегрирования необходимо связать с видом области G .

§ 7.3. Независимость компонент случайного вектора

Определение. Говорят, что для случайного вектора (X, Y) его компоненты X и Y **независимые случайные величины**, если для любых борелевских множеств A и B на прямой

$$P((X \in A)(Y \in B)) = P(X \in A)P(Y \in B). \quad (7.8)$$

В случае абсолютно непрерывного случайного вектора (X, Y) условие (7.8) можно выразить в терминах функций распределения и плотности.

Теорема 7.1. Для абсолютно непрерывного случайного вектора (X, Y) компоненты X и Y независимы тогда и только тогда, когда выполняется одно из условий

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \quad (7.9)$$

$$f_{X,Y}(x,y) = f_X(x)f_Y(y). \quad (7.10)$$

Доказательство. Положим $A = (X < x), B = (Y < y)$. Тогда в левой части (7.8) мы получим $F_{X,Y}(x,y)$, а в правой части произведение $F_X(x)F_Y(y)$, что доказывает (7.9). Вычисляя смешанную производную по x, y от обеих частей (7.9), получим равенство (7.10) согласно свойству 3 функции плотности.

Наоборот, если выполнено (7.10), то равенство (7.8) будет выполнено ввиду соответствующего свойства двойного интеграла и определения функции плотности. Теорема доказана полностью.

Таким образом, компоненты X, Y абсолютно непрерывного случайного вектора (X, Y) являются независимыми случайными величинами в том и только в том случае, если произведение их плотностей совпадает с плотностью совместного распределения.

§ 7.4. Числовые характеристики случайного вектора

Теорема 7.2. Для математического ожидания функции $\varphi(x,y)$ от компонент случайного вектора (X, Y) справедлива формула

$$M[\varphi(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x,y) f_{X,Y}(x,y) dx dy. \quad (7.11)$$

Мы не приводим доказательство этой теоремы, которая является аналогом теоремы 5.2 для одномерного случая.

Мы видели, что в одномерном случае основные числовые характеристики случайной величины выражались через начальные и центральные моменты. Дадим аналогичное определение для случайного вектора.

Определение. *Начальным моментом порядка (k, l) называется математическое ожидание функции $x^k y^l$:*

$$\nu_{k,l} = M(X^k Y^l). \quad (7.12)$$

Из формулы (7.11) следует

$$\nu_{k,l} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^l f_{X,Y}(x, y) dx dy. \quad (7.13)$$

Определение. *Центральным моментом порядка (k, l) называется математическое ожидание функции $(x - m_X)^k (y - m_Y)^l$, где $m_X = M(X)$, $m_Y = M(Y)$:*

$$\mu_{k,l} = M\left(\overset{\circ}{X}^k \overset{\circ}{Y}^l\right). \quad (7.14)$$

Из формулы (7.11) следует

$$\mu_{k,l} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^k (y - m_Y)^l f_{X,Y}(x, y) dx dy. \quad (7.15)$$

Числа k и l характеризуют порядок момента по отношению к каждой компоненте случайного вектора. Число $r = k + l$ называют *суммарным порядком* момента. Соответственно суммарному порядку моменты можно разделить на моменты первого, второго и т.д. порядка. Мы рассмотрим более подробно моменты первого и второго порядка.

Первые начальные моменты – это нам уже знакомые математические ожидания случайных величин X и Y .

Действительно,

$$\begin{aligned} \nu_{1,0} &= M(X^1 Y^0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy = \\ &= \int_{-\infty}^{\infty} x dx \int_{-\infty}^{\infty} f(x, y) dy = \int_{-\infty}^{\infty} x f_X(x) dx = M(X). \end{aligned}$$

Аналогично

$$\nu_{0,1} = M(Y).$$

Точка с координатами $(M(X), M(Y))$ характеризует центр системы случайных величин, вокруг которого происходит рассеивание возможных значений.

Кроме первых моментов широко применяют вторые центральные моменты, которые бывают трех типов. Два из них дают знакомые нам дисперсии компонент X и Y :

$$\mu_{2,0} = M\left(\overset{\circ}{X}^2\right) = D(X), \quad \mu_{0,2} = M\left(\overset{\circ}{Y}^2\right) = D(Y),$$

которые характеризуют рассеивание возможных значений случайных величин X и Y вдоль осей x и y .

Особую роль в определении взаимодействия компонент играет *второй смешанный центральный момент*

$$\mu_{1,1} = M\left(\overset{\circ}{X} \overset{\circ}{Y}\right). \quad (7.16)$$

Мы уже рассматривали эту характеристику дискретных систем случайных величин, которую называли ковариацией. Она имеет важное значение и для непрерывных случайных векторов.

Определение. *Ковариацией или корреляционным моментом* случайного вектора (X, Y) называют величину

$$\begin{aligned}\text{Cov}(X, Y) &= M((X - m_X)(Y - m_Y)) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y)f(x, y)dx dy.\end{aligned}\quad (7.17)$$

Ковариация обладает следующими свойствами:

1. $\text{Cov}(X, Y) = M(XY) - M(X)M(Y)$.
2. $\text{Cov}(X, X) = D(X)$.
3. $D(X+Y) = D(X) + D(Y) + 2\text{Cov}(X, Y)$.
4. Если X и Y независимы, то $\text{Cov}(X, Y) = 0$.
5. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
6. $\text{Cov}(aX, Y) = \text{Cov}(X, aY) = a\text{Cov}(X, Y)$.
7. $\text{Cov}(X+Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$.
8. $\text{Cov}(X, Y+Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$.

Проверка этих свойств выполняется так же, как и в дискретном случае.

Если $\text{Cov}(X, Y) = 0$, то случайные величины X и Y называются *некоррелированными*. Таким образом, согласно свойству 4 из независимости X и Y следует их некоррелированность. Обратное утверждение неверно.

Ковариация $\text{Cov}(X, Y)$ может использоваться как характеристика взаимосвязи X и Y . Например, положительный знак $\text{Cov}(X, Y) > 0$ свидетельствует о том, что в колебательной динамике случайных величин X и Y преобладают отклонения от средних значений в одном направлении. Для подобного сравнения случайных величин, однако, больше подходит безразмерная характеристика – *коэффициент корреляции*, определяемый формулой

$$\rho_{XY} = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}. \quad (7.18)$$

Свойства коэффициента корреляции:

1. $\rho_{XY} = \rho_{YX}$.
2. $|\rho_{XY}| \leq 1$.
3. Условие $|\rho_{XY}| = 1$ равнозначно существованию констант $\alpha, \beta \in \mathbf{R}$, таких, что равенство $Y = \alpha + \beta X$ выполняется с вероятностью 1.

Заметим, что свойство 3 ковариации с использованием коэффициента корреляции можно переписать в следующем виде

$$\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y) + 2\rho(X, Y)\sigma(X)\sigma(Y). \quad (7.19)$$

Свойство 1 очевидно. Для доказательства свойств 2 и 3 рассмотрим случайную величину

$$Z = \sigma_Y X - \rho \sigma_X Y \quad (7.20)$$

и вычислим ее дисперсию

$$\sigma^2(Z) = \sigma_Y^2 \sigma_X^2 + \rho^2 \sigma_X^2 \sigma_Y^2 - 2\rho \sigma_Y \sigma_X \rho \sigma_Y \sigma_X = 2\sigma_X^2 \sigma_Y^2 (1 - \rho^2). \quad (7.21)$$

Поскольку левая часть (7.21) неотрицательна, то и правая, что доказывает свойство 2.

Если $|\rho| = 1$, то дисперсия Z равна 0, так что с вероятностью 1 эта случайная величина равна константе α . Таким образом, из (7.20) имеем

$$\sigma_Y X - \rho \sigma_X Y = \alpha,$$

что и доказывает линейную зависимость между случайными величинами X и Y .

Пусть равенство $Y = \alpha + \beta X$ выполняется с вероятностью 1. По свойству дисперсии имеем

$$\sigma^2(Y) = \beta^2 \sigma^2(X),$$

так что

$$\sigma(Y) = |\beta| \sigma(X). \quad (7.22)$$

Вычислим ковариацию X и Y

$$\text{Cov}(X, Y) = \text{Cov}(X, \alpha + \beta X) = \text{Cov}(X, \beta X) = \beta \sigma^2(X). \quad (7.23)$$

Подставляя (7.19), (7.20) в формулу для коэффициента корреляции (7.18), получим

$$\rho(X, Y) = \frac{\beta \sigma^2(X)}{\sigma(X) |\beta| \sigma(X)} = \pm 1,$$

что и требовалось доказать.

§ 7.5. Системы n случайных величин

До сих пор мы рассматривали числовые характеристики системы двух случайных величин. В приложениях часто рассматриваются системы n случайных величин.

Пусть случайная величина Y является линейной комбинацией случайных величин X_1 и X_2

$$Y = a_1 X_1 + a_2 X_2.$$

Найдем дисперсию Y

$$\begin{aligned} D(Y) &= D(a_1 X_1) + D(a_2 X_2) + 2\text{Cov}(a_1 X_1, a_2 X_2) = \\ &= a_1^2 D(X_1) + a_2^2 D(X_2) + 2a_1 a_2 \text{Cov}(X_1, X_2). \end{aligned} \quad (7.24)$$

Формулу (7.24) можно переписать и в таком виде

$$D(Y) = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + 2a_1 a_2 \rho \sigma_1 \sigma_2, \quad (7.25)$$

где ρ – коэффициент корреляции X_1 и X_2 . Нашей задачей является обобщение формул (7.24), (7.25) на случай n слагаемых. А именно, если случайная величина Y является линейной комбинацией n случайных величин $X_1, \dots, X_n$, то справедливы формулы

$$D(Y) = \sum_{i,j=1}^n \text{Cov}(X_i, X_j) a_i a_j, \quad (7.26)$$

$$D(Y) = \sum_{i,j=1}^n \rho(X_i, X_j) \sigma_i \sigma_j a_i a_j. \quad (7.27)$$

Действительно, формула (7.24) непосредственно вытекает из (7.26) для случая $n = 2$, поскольку $\text{Cov}(X_1, X_2) = \text{Cov}(X_2, X_1)$ и $\text{Cov}(X_i, X_i) = D(X_i)$, $i = 1, 2$. Формула (7.27) является следствием (7.26), если учесть, что

$$\text{Cov}(X_i, X_j) = \rho(X_i, X_j) \sigma_i \sigma_j, \quad i, j = 1, \dots, n.$$

Этим формулам можно придать более компактный вид, если воспользоваться матричными обозначениями. Для этого рассмотрим следующие определения.

Определение. Для набора случайных величин $X_1, X_2, \dots, X_n$ **ковариационной матрицей** $C = (c_{ij})$ и **корреляционной матрицей** $R = (\rho_{ij})$ называют квадратные матрицы порядка n , составленные из всех парных ковариаций $c_{ij} = \text{Cov}(X_i, X_j)$ и всех коэффициентов корреляции $\rho_{ij} = \rho(X_i, X_j)$, $i, j = 1, \dots, n$.

Свойства.

1. Ковариационная и корреляционные матрицы являются симметричными матрицами.
2. Ковариационная и корреляционные матрицы неотрицательно определены.
3. Определители этих матриц неотрицательны. Более того, $\det R \leq 1$.

Свойство 1 следует непосредственно из аналогичного свойства для ковариаций. Свойство 2 следует из формул (7.26), (7.27), так как левая часть принимает только неотрицательные значения. Первая часть свойства 3 отсюда следует непосредственно. Неравенство $\det R \leq 1$ является обобщением неравенства $\rho \leq 1$ и оставляется без доказательства.

Пусть C – ковариационная матрица случайных величин $X_1, X_2, \dots, X_n$ и $A = (a_1, a_2, \dots, a_n)^T$ – произвольный ненулевой вектор констант. Тогда для случайной величины $Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$ выполняется соотношение

$$D(Y) = A^T C A \geq 0, \quad (7.28)$$

которое служит матричной формой записи формулы (7.26). Заметим, что условие $D(Y) = 0$ равносильно равенству $CA = 0$, означающему вырожденность матрицы C .

§ 7.6. Равномерное распределение

Равномерное распределение в плоской области G . Начнем с разбора следующей задачи. На отрезке длиной 1 выбирают наугад и независимо друг от друга две точки. Какова вероятность того, что расстояние между ними окажется не больше l (где $l < 1$)?

Считая, что данный отрезок есть отрезок $[0, 1]$ числовой оси, обозначим через x координату первой точки и через y – координату второй. Исходом опыта является пара чисел X, Y , удовлетворяющих условиям:

$$0 \leq X \leq 1, 0 \leq Y \leq 1.$$

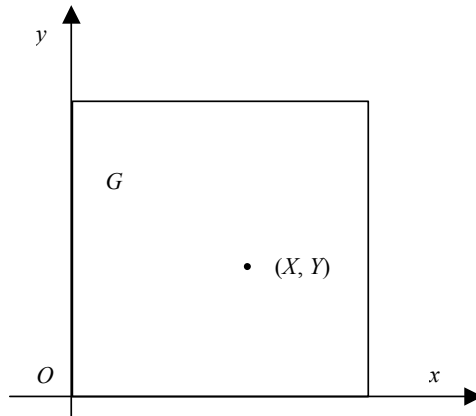


Рис. 7.5

Иначе говоря, исход опыта – случайная точка квадрата G , изображенного на рис. 7.5.

Рассмотрим отдельно случайную величину X . Ее закон распределения в условии задачи не оговорен; он будет зависеть от того, какой смысл мы придадим слову "наугад" в формулировке задачи. Наиболее естественное толкование этого слова заключается в том, что все значения величины X на отрезке $[0, 1]$ *равноправны*, или, говоря точнее, величина X равномерно распределена на отрезке $[0, 1]$. То же самое относится, конечно, и к Y . Итак, будем считать, что плотности вероятности для X и Y равны соответственно:

$$f_1(x) = \begin{cases} 1, & \text{для } 0 \leq x \leq 1, \\ 0, & \text{для остальных } x, \end{cases}$$

$$f_2(y) = \begin{cases} 1, & \text{для } 0 \leq y \leq 1, \\ 0, & \text{для остальных } y. \end{cases}$$

Ввиду независимости X и Y плотность распределения для системы (X, Y) будет:

$$f(x, y) = f_1(x)f_2(y),$$

т.е.

$$f(x, y) = \begin{cases} 1, & \text{для } 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{для остальных } x, y. \end{cases}$$

Для того, чтобы расстояние между точками не превышало l , нужно, чтобы выполнялось неравенство

$$|x - y| \leq l. \quad (7.29)$$

На рис. 7.6 штриховкой отмечена область внутри квадрата G , которая отвечает неравенству (7.29). Обозначим эту область через H . Неравенство (7.29) эквивалентно условию $(X, Y) \in H$, следовательно, вероятность его осуществления

$$P((X, Y) \in H) = \iint_H f(x, y) dx dy = \iint_H dx dy = S_H,$$

где S_H – площадь области H , равная $1 - (1 - l)^2$.

Отсюда искомая вероятность будет:

$$1 - (1 - l)^2. \quad (7.30)$$

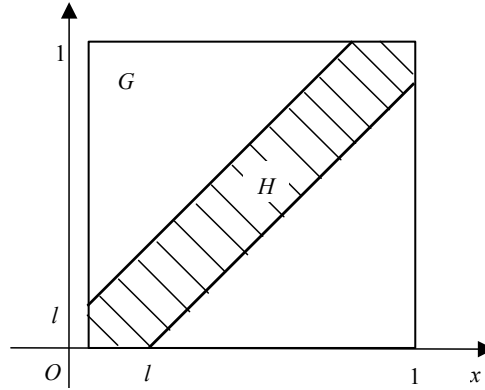


Рис. 7.6

Задачу о выборе двух точек на отрезке можно сформулировать как задачу о "встрече двух лиц". Представим себе, что два человека договорились встретиться в определенном месте между двенадцатью и часом дня. При этом было условлено, что пришедший на место свидания будет ждать другого только в течение 20 минут. Какова вероятность того, что встреча состоится, если каждый из договорившихся приходит, когда ему вздумается (но между двенадцатью и часом дня), не согласуя свои действия с другим? Нетрудно понять, что математическое содержание этой задачи в точности такое же, как в задаче о выборе двух точек на отрезке $[0,1]$; при этом задаваемый вопрос состоит в том, с какой вероятностью расстояние между точками окажется не больше чем $\frac{1}{3}$. Действительно, если обозначить через X время прихода одного из данных лиц и через Y – время прихода второго, то условие встречи запишется:

$$|X - Y| \leq \frac{1}{3},$$

после чего аналогия с задачей о выборе точек становится очевидной. Полагая в выражении (7.30) $l = \frac{1}{3}$, найдем, что искомая вероятность встречи будет $1 - \left(1 - \frac{1}{3}\right)^2 = \frac{5}{9}$.

Распределение, с которым мы столкнулись в только что разобранный задаче, относится к числу равномерных распределений. В двумерном случае равномерное распределение задается с помощью области G конечной площади; при этом плотность вероятности постоянна в G , а за пределами G равна нулю:

$$f(x, y) = \begin{cases} c & \text{для } (x, y) \in G, \\ 0 & \text{в остальных точках.} \end{cases}$$

Значение постоянной c можно определить из условия:

$$\iint_G f(x, y) dx dy = 1,$$

которое дает $cS_G = 1$ или $c = \frac{1}{S_G}$, где S_G обозначает площадь области G .

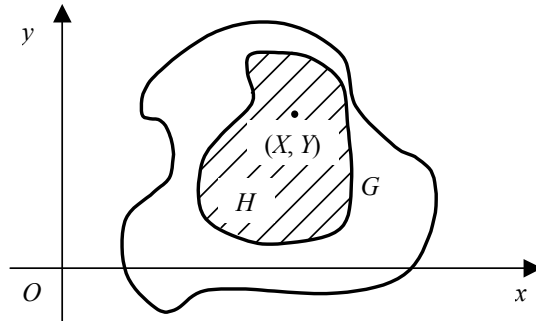


Рис. 7.7

Если H – какая-либо часть области G (см. рис. 7.7), то вероятность попадания случайной точки (X, Y) в область H равна интегралу по этой области от постоянной функции $\frac{1}{S_G}$, следовательно,

$$P((X, Y) \in H) = \frac{S_H}{S_G}. \quad (7.31)$$

Заметим, что равномерное распределение можно рассматривать не только в двумерном случае, но и в случае пространства любого числа измерений. Равномерное распределение на отрезке прямой было введено в § 1.5; для этого случая мы имели в упомянутом параграфе формулу (1.25), которую теперь запишем следующим образом:

$$P(X \in H) = \frac{\text{длина } H}{\text{длина } G}, \quad (7.32)$$

где G есть отрезок числовой оси, представляющий собой множество всех возможных положений точки X , а H – любой отрезок, содержащийся в G . Аналогичным образом для равномерного распределения в пространственной области G имеем:

$$P((X, Y, Z) \in H) = \frac{\text{объем } H}{\text{объем } G}. \quad (7.33)$$

§ 7.7. Условные распределения и условные математические ожидания

Мы знаем, что если дан случайный вектор (X, Y) и известна функция плотности $f(x, y)$, то мы можем определить законы распределения каждой компоненты. Пусть, наоборот, нам известны законы распределения каждой компоненты, заданные с помощью функции распределения или плотности. Сможем ли мы найти совместный закон распределения случайного вектора (X, Y) ? Ответ будет утвердительным только в том случае, когда случайные величины X и Y независимы. В общем случае нам нужно знать зависимость между величинами, входящими в систему. Эта зависимость может быть определена из знания так называемых *условных законов распределения*, т.е. закона распределения одной случайной величины при условии, что другая случайная величина приняла определенное значение.

Если случайные величины X и Y дискретные, то их условные законы распределения могут быть определены, используя теорему умножения вероятностей:

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}, \quad (7.34)$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}. \quad (7.35)$$

Если известен закон распределения $P(X = x_i, Y = y_j) = p_{ij}$, то эти формулы можно записать в следующем виде:

$$P(X = x_i | Y = y_j) = \frac{p_{ij}}{\sum_i p_{ij}}, \quad (7.36)$$

$$P(Y = y_j | X = x_i) = \frac{p_{ij}}{\sum_j p_{ij}}. \quad (7.37)$$

Нетрудно проверить, что любая из формул 7.34–37 определяет закон распределения, т.е. сумма вероятностей равна 1.

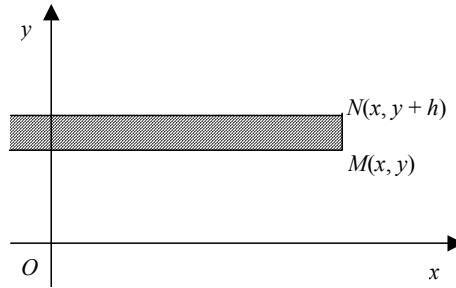


Рис. 7.8

Для непрерывных случайных величин вероятность отдельного значения равна 0, поэтому мы не можем сразу применить указанные формулы. Поэтому рассмотрим точку $M(x, y)$ на плоскости и примыкающую к ней горизонтальную полосу ширины h (см. рис. 7.8). Вероятность попадания системы (X, Y) в эту полосу можно найти, используя теорему умножения вероятностей

$$P(X < x, y \leq Y < y + h) = P(y \leq Y < y + h)P(X < x | y \leq Y < y + h). \quad (7.38)$$

Если устремить h к нулю, то второй множитель даст *условное распределение* X при условии, что Y принимает значение y .

Определение. Функция

$$F(x | y) = \lim_{h \rightarrow 0} \frac{P(X < x, y \leq Y < y + h)}{P(y \leq Y < y + h)}. \quad (7.39)$$

называется **условным распределением** X при условии, что Y принимает значение y .

Аналогично определяется условное распределение $F(y | x)$.

Для абсолютно непрерывных случайных величин условные плотности определяют аналогичным образом

$$f(x | y) = \frac{f(x, y)}{f_Y(y)}, \quad f(y | x) = \frac{f(x, y)}{f_X(x)}. \quad (7.40)$$

Заметим, что полученные функции можно рассматривать как функции плотности некоторой случайной величины, иными словами, они обладают свойствами неотрицательности и нормированности. Действительно, из (7.40) сразу следует, что $f(x | y) \geq 0$. С другой стороны

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

Если случайные величины X и Y независимы, то условные плотности равны безусловным

$$f(x | y) = f_X(x), \quad f(y | x) = f_Y(y).$$

Пример 7.1. Пусть двумерная случайная величина имеет функцию плотности $f(x, y) = x + y$, сосредоточенную в квадрате $0 \leq x \leq 1, 0 \leq y \leq 1$. Найти условные плотности.

Решение. Найдём плотности компонент

$$f_X(x) = \int_0^1 (x + y) dy = x + \frac{1}{2}, \quad f_Y(y) = \int_0^1 (x + y) dx = \frac{1}{2} + y.$$

Отсюда

$$f(x | y) = \frac{x + y}{y + 0,5}, \quad f(y | x) = \frac{x + y}{x + 0,5}.$$

Поскольку

$$f(x, y) = x + y \neq f_X(x)f_Y(y) = (x + 0,5)(y + 0,5),$$

то величины X и Y зависимы.

Зная как находятся условные законы распределения, можно поставить задачу о нахождении *условных математических ожиданий*.

Определение. *Условным математическим ожиданием* случайной величины Y при условии, что случайная величина X приняла значение x , называется величина

$$y(x) = M(Y | X = x) = \int_{-\infty}^{\infty} yf(y | x)dy. \quad (7.41)$$

Аналогично определяется *условное математическое ожидание* X при условии, что $Y = y$

$$x(y) = M(X | Y = y) = \int_{-\infty}^{\infty} xf(x | y)dx. \quad (7.42)$$

Заметим, что, вообще говоря, условное математическое ожидание, определенное формулой (7.41), не является числом, а выражается в виде некоторой функции, зависящей от x . Более того, каждую из функций (7.41), (7.42) можно рассматривать соответственно как функцию от случайной величины X и Y . Поэтому можно найти их математическое ожидание

$$\begin{aligned} M(M(Y | X)) &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} yf(y | x)dy \right\} f_X(x)dx = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} yf(x, y)dy \right\} dx = \\ &= \int_{-\infty}^{\infty} y \left\{ \int_{-\infty}^{\infty} f(x, y)dx \right\} dy = \int_{-\infty}^{\infty} yf_Y(y)dy = M(Y). \end{aligned}$$

Сформулируем полученное соотношение в виде теоремы.

Теорема 7.3. *Выполняются следующие соотношения*

$$M(M(Y | X)) = M(Y), \quad (7.43)$$

$$M(M(X | Y)) = M(X). \quad (7.44)$$

Если не прибегать к излишней строгости, соотношения (7.43) и (7.44) можно выразить словами: «Математическое ожидание от условного математического ожидания дает математическое ожидание исходной величины».

Пример 7.2. В условиях примера 7.1 найти условное математическое ожидание случайной величины Y .

Решение. Подставляя найденное выражение для условной плотности Y в (7.41), получим

$$M(Y | X) = \int_0^1 y \frac{x+y}{x+0,5} dy = \frac{1}{x+0,5} \left(\frac{xy^2}{2} + \frac{y^3}{3} \right) \Big|_0^1 = \frac{\frac{1}{2}x + \frac{1}{3}}{x + \frac{1}{2}} = \frac{3x+2}{3(2x+1)}.$$

В данном случае условное математическое ожидание Y представляет собой функцию, определенную на отрезке $[0, 1]$, которая медленно убывает от значения $M(Y | X = 0) = 2/3$ до значения $M(Y | X = 1) = 5/9$.

Проверим выполнение (7.43) в данном частном случае. Имеем

$$M(M(Y | X)) = \int_0^1 \frac{3x+2}{3(2x+1)} \left(x + \frac{1}{2} \right) dx = \frac{1}{6} \int_0^1 (3x+2) dx = \frac{1}{6} \left(\frac{3}{2} + 2 \right) = \frac{7}{12}.$$

С другой стороны

$$M(Y) = \int_0^1 y \left(y + \frac{1}{2} \right) dy = \frac{1}{3} + \frac{1}{4} = \frac{7}{12},$$

что совпадает с предыдущим результатом.

§ 7.8. Функциональные преобразования случайных векторов

Мы уже рассматривали функции от случайных величин, в частности, мы определили, как связаны функции плотности случайных величин X и $\varphi(X)$, где $\varphi(x)$ – дифференцируемая функция, имеющая обратную.

Пусть $X = (X_1, X_2)$ – случайный вектор, сосредоточенный в области $G \subseteq \mathbf{R}^2 = (x_1, x_2)$ и

$$\begin{cases} y_1 = g_1(x_1, x_2), \\ y_2 = g_2(x_1, x_2), \end{cases} \quad (7.45)$$

где g_1, g_2 – дифференцируемые функции, отображающие область G взаимно однозначно на область $H \subseteq \mathbf{R}^2 = (y_1, y_2)$, причем существует обратное отображение области H на область G , осуществляемое с помощью отображения

$$\begin{cases} x_1 = h_1(y_1, y_2), \\ x_2 = h_2(y_1, y_2). \end{cases} \quad (7.46)$$

Поскольку функции g_1, g_2 – непрерывны, то формулы

$$\begin{cases} Y_1 = g_1(X_1, X_2), \\ Y_2 = g_2(X_1, X_2) \end{cases}$$

определяют случайный вектор $Y = (Y_1, Y_2)$. Наша задача – зная совместную функцию плотности $f_{X_1, X_2}(x_1, x_2)$ для X , найти функцию плотности $f_{Y_1, Y_2}(y_1, y_2)$ для Y .

Для этого заметим, что для любого борелевского множества $B \subseteq G$ и его образа при отображении (7.45) $C \subseteq H$ выполняется равенство

$$P(X \in B) = P(Y \in C).$$

Это равенство можно переписать с помощью интегралов

$$\iint_B f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = \iint_C f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2. \quad (7.47)$$

Используя формулы замены (7.46) в двойном интеграле, формулу (7.47) представим в следующем виде

$$\iint_C f_{X_1, X_2}(h_1(y_1, y_2), h_2(y_1, y_2)) |J| dy_1 dy_2 = \iint_C f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2,$$

где J – якобиан замены (7.46)

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix}. \quad (7.48)$$

Окончательно получим

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(h_1(y_1, y_2), h_2(y_1, y_2)) |J|. \quad (7.49)$$

Формулы (7.49), (7.48) дают нам искомое выражение совместной функции плотности случайного вектора Y через функцию плотности вектора X .

Особое значение в дальнейшем будет играть *линейная замена переменных*. В этом случае формулы (7.45) приобретут вид

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2, \\ y_2 = a_{21}x_1 + a_{22}x_2 \end{cases}$$

с невырожденной матрицей A коэффициентов, так что обратная замена выглядит так

$$\begin{cases} x_1 = b_{11}y_1 + b_{12}y_2, \\ x_2 = b_{21}y_1 + b_{22}y_2 \end{cases}$$

с матрицей коэффициентов B , обратной матрице A

$$B = A^{-1}.$$

Формула (7.49) примет вид

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(b_{11}y_1 + b_{12}y_2, b_{21}y_1 + b_{22}y_2) |J|, \quad (7.50)$$

где якобиан J равен определителю матрицы B

$$J = \det(B) = \det(A)^{-1}. \quad (7.51)$$

Функции от случайных величин

В приложениях часто приходится рассматривать случайные величины, которые являются некоторыми функциями случайных величин. Например, необходимо рассмотреть квадрат случайной величины или произведение двух случайных величин. При этом часто необходимо выяснить закон распределения получающейся случайной величины. В данной главе мы рассматриваем задачи, связанные с решением такого типа проблем.

§ 8.1. Функции от случайной величины

Понятие функции случайной величины уже рассматривалось в § 3.7. Там же мы установили, что для непрерывной функции $\varphi(x)$ $Y = \varphi(X)$ будет снова случайной величиной.

В данном параграфе мы будем рассматривать задачу нахождения функции плотности $Y = \varphi(X)$, когда случайная величина X является абсолютно непрерывной случайной величиной с функцией плотности $f(x)$, а данная функция $\varphi(x)$ непрерывно дифференцируема.

Пример 8.1. Пусть X – абсолютно непрерывная случайная величина с функцией плотности $f(x)$. Найти функцию плотности $g(x)$ случайной величины

$$Y = aX + b. \quad (8.1)$$

Решение. Пусть $F(x)$ – функция распределения случайной величины X , $G(x)$ – функция распределения Y . Выразим $G(x)$ через $F(x)$. По определению

$$G(x) = P(Y < x). \quad (8.2)$$

Подставляя выражение (8.1) в (8.2), получим

$$G(x) = P(aX + b < x). \quad (8.3)$$

Возможны два случая. Если $a > 0$, то решая неравенство $aX + b < x$ относительно X , получим эквивалентное неравенство $X < \frac{x-b}{a}$, так что

$$G(x) = F\left(\frac{x-b}{a}\right), \quad a > 0. \quad (8.4)$$

Дифференцируя обе части по x , находим искомую зависимость:

$$g(x) = \frac{1}{a} f\left(\frac{x-b}{a}\right), \quad a > 0. \quad (8.5)$$

Если $a < 0$, то неравенство $aX + b < x$ эквивалентно неравенству $X > \frac{x-b}{a}$, так что формула (8.4) приобретает вид

$$G(x) = 1 - F\left(\frac{x-b}{a}\right), \quad a < 0, \quad (8.6)$$

откуда формула для плотности:

$$g(x) = -\frac{1}{a} f\left(\frac{x-b}{a}\right), \quad a < 0. \quad (8.7)$$

Формулы (8.5), (8.6) можно записать единообразно с использованием абсолютной величины. Это легко запомнить, если сообразить, что так как левая часть принимает только положительные значения, то и числовой множитель в правой части должен быть положительным. Итак, заключительная формула для искомой плотности имеет вид:

$$g(x) = \frac{1}{|a|} f\left(\frac{x-b}{a}\right). \quad (8.8)$$

В качестве следствия получим следующее утверждение.

Теорема 8.1. Если $X \sim N(m, \sigma)$ – случайная величина, распределенная по нормальному закону, то $Y = aX + b$ также распределена по нормальному закону, причем

$$Y \sim N(am + b, |a|\sigma). \quad (8.9)$$

Доказательство. Запишем функцию плотности случайной величины X :

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

Тогда согласно (8.8) имеем

$$g(x) = \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(\frac{x-b}{a} - m\right)^2}{2\sigma^2}\right) = \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-am-b)^2}{2a^2\sigma^2}\right),$$

откуда нетрудно увидеть, что $g(x)$ – функция плотности нормальной случайной величины с параметрами $M(Y) = am + b$ и $\sigma(Y) = |a|\sigma$, что и требовалось.

Пример 8.2. Пусть $X \sim N(0, 1)$ – стандартная нормальная случайная величина и $Y = X^2$. Найти плотность $g(x)$ случайной величины Y .

Решение. Найдем сначала функцию распределения Y . Имеем

$$G(x) = P(Y < x) = P(X^2 < x) = P(-\sqrt{x} < X < \sqrt{x}) = 2\Phi(\sqrt{x}),$$

где $\Phi(x)$ – функция Лапласа. Чтобы найти искомую функцию плотности, необходимо продифференцировать обе части равенства. Учитывая, что производная функции Лапласа есть функция Гаусса, получим

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

так что искомая формула для плотности имеет вид

$$g(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, x > 0. \quad (8.10)$$

Рассмотренные примеры служат двум целям: во-первых, дать подход к нахождению функции плотности в типичных случаях и, во-вторых, использовать полученные формулы в дальнейших приложениях.

§ 8.2. Закон распределения суммы двух случайных величин

Пусть дана система двух случайных величин (X, Y) с плотностью распределения $f(x, y)$. В этом параграфе мы будем решать важную для приложений задачу о нахождении функции плотности случайной величины

$$Z = X + Y.$$

Заметим, что для любой непрерывной функции $\varphi(x, y)$ и любых случайных величин X и Y случайной величиной будет и $\varphi(X, Y)$. Это устанавливается такими же рассуждениями, как и для функции от одной переменной.

Теорема 8.2. Если система случайных величин (X, Y) распределена с плотностью $f(x, y)$, то случайная величина $Z = X + Y$ имеет плотность $h(z)$, которая определяется формулой

$$h(z) = \int_{-\infty}^{\infty} f(x, z-x) dx = \int_{-\infty}^{\infty} f(z-y, y) dy. \quad (8.11)$$

Доказательство. Найдем функцию распределения случайной величины Z . Для этого рассмотрим неравенство $Z < z$ или $X + Y < z$. На координатной плоскости точки, удовлетворяющие этому неравенству, образуют область D , лежащую ниже прямой $x + y = z$ (см. рис. 8.1).

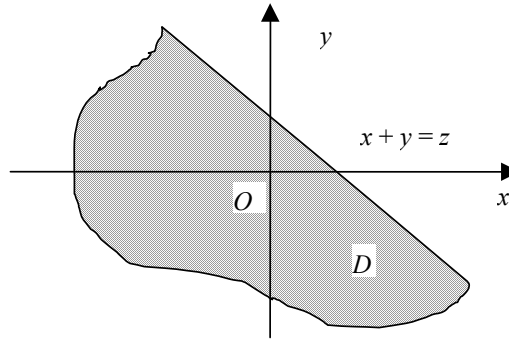


Рис. 8.1

Тогда функцию распределения для случайной величины Z можно выразить с помощью двойного интеграла, который можно свести к повторному:

$$H(z) = P(Z < z) = P(X + Y < z) = \iint_D f(x, y) dx dy = \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) dx dy = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{z-x} f(x, y) dy \right\} dx.$$

Дифференцируя по z , получим выражение для функции плотности:

$$h(z) = \int_{-\infty}^{\infty} f(x, z-x) dx. \quad (8.12)$$

Поскольку выражение для плотности симметрично относительно переменных x и y , то выражение (8.12) можно переписать в эквивалентном виде:

$$h(z) = \int_{-\infty}^{\infty} f(z-y, y) dy. \quad (8.13)$$

Тем самым мы получили искомое выражение (8.11) и доказательство теоремы закончено полностью.

Особого внимания заслуживает случай, когда случайные величины X и Y *независимы*. Мы знаем, что в этом случае

$$f(x, y) = f_1(x)f_2(y),$$

где $f_1(x)$ – функция плотности для X , $f_2(y)$ – функция плотности для Y . Поэтому формула (8.11) принимает вид

$$h(z) = \int_{-\infty}^{\infty} f_1(x)f_2(z-x)dx = \int_{-\infty}^{\infty} f_1(z-y)f_2(y)dy. \quad (8.14)$$

Отвлекаясь от существа решаемой задачи, можно заметить, что полученная нами формула (8.14) определяет весьма интересную операцию над функциями: каждой паре функций $f_1(x)$ и $f_2(x)$ ставится в соответствие новая функция $h(z)$, определяемая формулой (8.14). Эта операция называется *свертыванием*, а функция $h(z)$ – *сверткой функций* $f_1(x)$ и $f_2(x)$. Часто формулу (8.14) короче записывают так

$$h = f_1 * f_2.$$

Пример 8.3. Даны две независимые случайные величины X и Y , распределенные равномерно соответственно на отрезках $[0, m]$ и $[0, n]$ ($m \leq n$). Найти функцию плотности их сумм $Z = X + Y$.

Решение. Поскольку X и Y сосредоточены на отрезках $[0, m]$ и $[0, n]$, то случайная величина Z сосредоточена на отрезке $[0, m+n]$ и формула (8.14) принимает вид

$$h(z) = \int_{-\infty}^{\infty} f_1(x)f_2(z-x)dx.$$

При $0 \leq z \leq m$ пределы интегрирования сводятся к промежутку $[0, z]$, поскольку при отрицательном x обе плотности равны нулю, а при $x > z$ обращается в нуль второй множитель $f_2(z-x)$. Имеем

$$h(z) = \int_0^z \frac{1}{mn} dx = \frac{z}{mn}.$$

При $m \leq z \leq n$ интегрирование следует проводить на промежутке $[0, m]$, так как только в этом случай обе плотности не равны нулю. Поэтому

$$h(z) = \int_0^m \frac{1}{mn} dx = \frac{m}{mn} = \frac{1}{n}.$$

И наконец, при $n \leq z \leq m+n$ промежуток интегрирования имеет вид $[z-n, m]$, так что

$$h(z) = \int_{z-n}^m \frac{1}{mn} dx = \frac{m+n-z}{mn}.$$

Итак,

$$h(z) = \begin{cases} \frac{z}{mn}, & 0 \leq z \leq m, \\ \frac{1}{n}, & m \leq z \leq n, \\ \frac{m+n-z}{mn}, & n \leq z \leq m+n \end{cases}. \quad (8.15)$$

На рис. 8.2 изображен график соответствующей функции плотности, который называется *законом распределения Симпсона*.

Если $m = n$, т.е. когда случайные величины X и Y распределены одинаково, график из трапециевидного приобретает треугольный вид.

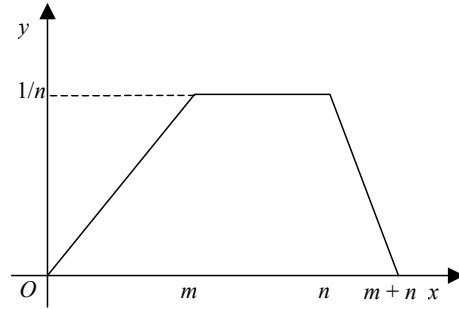


Рис. 8.2

Следующий пример связан с нахождением закона распределения величины χ^2

$$Y = X_1^2 + \dots + X_n^2, \quad (8.16)$$

где все $X_1, \dots, X_n$ – стандартные нормальные случайные величины. Мы уже нашли функцию плотности одного слагаемого (см. пример 8.2). Формула плотности (8.10) является частным случаем семейства так называемых *гамма плотностей*, а именно

$$f_{\alpha, \nu}(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\alpha x}, \quad x > 0, \alpha > 0, \nu > 0. \quad (8.17)$$

Формула (8.10) $g(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, x > 0$ является частным случаем (8.17), а именно

$$g(x) = f_{\frac{1}{2}, \frac{1}{2}}(x),$$

поскольку $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Заметим, что свойство нормированности плотности (8.17) следует из определения Г-функции. Рассмотрим свойства гамма плотности.

Свойство 1. $f_{\alpha,\mu} * f_{\alpha,\nu} = f_{\alpha,\mu+\nu}$.

Свойство 2. $M(X) = \frac{\nu}{\alpha}$.

Свойство 3. $D(X) = \frac{\nu}{\alpha^2}$.

Свойство 1 выводится с помощью формулы (8.14) и замены переменной $x = tv$. Докажем свойство 2. Имеем

$$M(X) = \int_0^{\infty} x f_{\alpha,\nu}(x) dx = \frac{1}{\Gamma(\nu)} \alpha^{\nu} \int_0^{\infty} x^{\nu} e^{-\alpha x} dx.$$

К последнему интегралу применим формулу интегрирования по частям. Положим $u = x^{\nu}$, $dv = e^{-\alpha x} dx$, $du = \nu x^{\nu-1} dx$, $v = -\frac{1}{\alpha} e^{-\alpha x}$. Тогда получим

$$M(X) = -\frac{\alpha^{\nu}}{\alpha \Gamma(\nu)} \left(x^{\nu} e^{-\alpha x} \Big|_0^{\infty} - \int_0^{\infty} \nu x^{\nu-1} e^{-\alpha x} dx \right).$$

Первое выражение в скобках равно нулю, поскольку $\alpha > 0$. Поэтому

$$M(X) = \frac{\nu}{\alpha} \int_0^{\infty} f_{\alpha,\mu}(x) dx.$$

Интеграл от функции плотности равен единице по свойству нормированности, и мы получаем искомую формулу.

Свойство 3 доказывается совершенно аналогично, только интегрировать по частям придется дважды. Вывод предлагаем провести читателю.

Пример 8.4. Найти закон распределения случайной величины χ^2 , определенной с помощью формулы (8.16).

Решение. Случайная величина χ^2 есть сумма независимых случайных величин, каждая из которых согласно примеру 8.2 подчиняется гамма распределению с параметрами $\alpha = \nu = \frac{1}{2}$. По свойству 1 сумма n таких величин также подчиняется гамма распределению с параметрами $\alpha = \frac{1}{2}$, $\nu = \frac{n}{2}$. По свойству 2 $M(\chi^2) = n$ и по свойству 3 $D(\chi^2) = 2n$. Подытожим полученные результаты в следующей теореме.

Теорема 8.3. Функция плотности случайной величины χ^2 дается формулой

$$\chi^2(x) = f_{\frac{1}{2}, \frac{n}{2}}(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x > 0. \quad (8.18)$$

Математическое ожидание и дисперсия равны соответственно

$$M(\chi^2) = n, \quad (8.19)$$

$$D(\chi^2) = 2n. \quad (8.20)$$

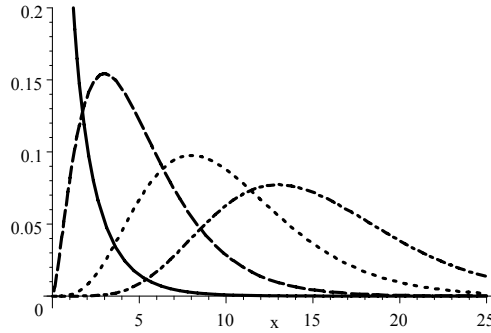


Рис. 8.3

Графики плотности χ^2 для числа степеней свободы, равных 1, 5, 10 и 15, приведены на рис. 8.3.

Из рисунка видно, что графики плотности *не являются* симметричными относительно математического ожидания, однако с ростом числа степеней свободы они становятся все более симметричными.

Наряду с суммой случайных величин в приложениях часто приходится иметь дело с частным непрерывных случайных величин. Для нахождения функции плотности частного сделаем сначала предварительное замечание.

Пусть F – функция распределения, зависящая от некоторого параметра t , а g – некоторая функция плотности. Рассмотрим функцию

$$H(x) = \int_{-\infty}^{\infty} F(x, t) g(t) dt. \quad (8.21)$$

Отметим следующие свойства $H(x)$.

Свойство 1. $H(x)$ – монотонно возрастающая функция.

Действительно, производная $H(x)$ дается выражением

$$H'(x) = \int_{-\infty}^{\infty} F'_x(x, t) g(t) dt = \int_{-\infty}^{\infty} f(x, t) g(t) dt,$$

где $f(x, t)$ – производная монотонно возрастающей функции и, следовательно, принимает положительные значения. Поэтому подынтегральная функция принимает положительные значения и $H'(x) > 0$.

Свойство 2. Значения функции $H(x)$ принадлежат отрезку $[0, 1]$.

Очевидно, что $H(x) \geq 0$, поскольку подынтегральная функция в (8.21) неотрицательна. С другой стороны, $F(x, t) \leq 1$ как функция распределения, так что интеграл (8.21) можно оценить следующим образом:

$$H(x) = \int_{-\infty}^{\infty} F(x, t) g(t) dt \leq \int_{-\infty}^{\infty} 1 \cdot g(t) dt = 1$$

по свойству функции плотности.

Свойство 3. Если функция $F(x, t)$ имеет непрерывную плотность $f(x, t)$, то функция $H(x)$ имеет плотность, равную

$$h(x) = \int_{-\infty}^{\infty} f(x, t) g(t) dt. \quad (8.22)$$

Для проверки этого свойства достаточно доказать свойство нормированности. Иными словами, надо проверить, что

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

Имеем

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, t) g(t) dx dt = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(x, t) dx \right\} g(t) dt = \int_{-\infty}^{\infty} g(t) dt = 1.$$

Доказанные свойства позволяют рассматривать функцию $h(x)$ как функцию плотности.

Теперь мы можем найти функцию плотности частного случайных величин.

Теорема 8.4. Пусть даны непрерывные случайные величины X и Y , причем Y сосредоточена на положительной полуоси, функции плотности которых соответственно равны $f(x)$ и $g(x)$. Тогда функцию плотности случайной величины $Z = X/Y$ можно найти по формуле

$$h(x) = \int_{-\infty}^{\infty} f(xt)tg(t)dt. \quad (8.23)$$

Доказательство. Сначала заменим случайную величину Y на число y . Имеем $F_Z(x) = P(X/y < x) = P(X < xy) = F_X(xy)$. Дифференцируя обе части по x , получим

$$h(x) = f(xy)y.$$

Рассматривая параметр y как случайную величину Y с плотностью $g(y)$, мы можем применить формулу (8.22). Таким образом, чтобы получить искомую функцию плотности, следует проинтегрировать $h(x)$ по всей оси, заменив переменную y на t , что и дает (8.23).

Следующая теорема использует формулу (8.23), чтобы найти функцию плотности, которая будет использована для описания *распределения Фишера*.

Теорема 8.5. Если случайные величины X и Y имеют плотности $f_{\frac{1}{2}, n}(x)$ и $f_{\frac{1}{2}, m}(x)$, то случайная величина $Z = X/Y$ имеет плотность

$$h(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)x^{\frac{n}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)(1+x)^{\frac{m+n}{2}}}, x > 0. \quad (8.24)$$

Доказательство. Подставляя в формулу (8.23) выражения для плотностей (8.17), получим

$$\begin{aligned} h(x) &= \int_0^{\infty} \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{n}{2}} (xt)^{\frac{n}{2}-1} e^{-\frac{xt}{2}} t \frac{1}{\Gamma\left(\frac{m}{2}\right)} \left(\frac{1}{2}\right)^{\frac{m}{2}} t^{\frac{m}{2}-1} e^{-\frac{t}{2}} dt = \\ &= \frac{x^{\frac{n}{2}-1}}{2^{\frac{m+n}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} t^{\frac{m+n}{2}-1} e^{-\frac{t(1+x)}{2}} dt. \end{aligned}$$

В последнем интеграле произведем замену $s = \frac{t(1+x)}{2}$, $t = \frac{2s}{1+x}$, $dt = \frac{2ds}{1+x}$. Тогда он приобретет вид

$$h(x) = \frac{x^{\frac{n}{2}-1}}{2^{\frac{m+n}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right) (1+x)^{\frac{m+n}{2}-1}} \frac{2}{1+x} \int_0^{\infty} s^{\frac{m+n}{2}-1} e^{-s} ds.$$

После простых преобразований выражения, стоящего перед интегралом, который равен $\Gamma\left(\frac{m+n}{2}\right)$, получим искомую формулу для плотности.

Определение. Распределением Фишера называется распределение случайной величины $F = \frac{m}{n} \frac{X}{Y}$, где X и Y распределены по закону χ^2 с числом степеней свободы соответственно n и m .

Применяя теорему 8.5, получим выражение для плотности величины, распределенной по закону Фишера.

Теорема 8.6. *Случайная величина F имеет плотность*

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right) m^{\frac{m}{2}} n^{\frac{n}{2}}}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \frac{x^{\frac{n}{2}-1}}{(m+nx)^{\frac{m+n}{2}}}, x > 0.$$

Графики функций плотности для трех случаев: а) $n = 8, m = 6$ (сплошная линия); б) $n = 18, m = 16$ (штриховая линия); в) $n = 28, m = 14$ (пунктирная линия) – приведены на рис. 8.4.

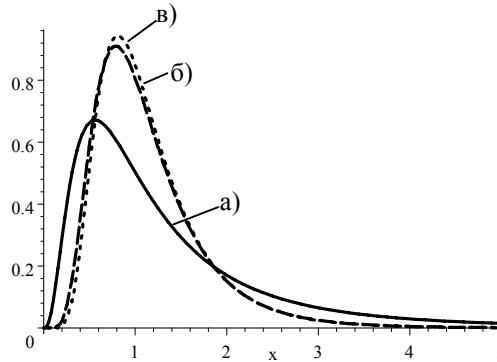


Рис. 8.4

Наряду с распределением χ^2 и Фишера – Снедекора в математической статистике используется *распределение Стьюдента*.

Определение. *Случайная величина подчинена закону распределения Стьюдента с n степенями свободы, если она имеет вид*

$$t_n = \frac{Z}{\sqrt{\chi^2(n)/n}}, \quad (8.25)$$

где $Z \sim N(0, 1)$ – стандартная нормальная величина.

Плотность распределения Стьюдента с n степенями свободы дается формулой

$$S_n(t) = B_n \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}, \quad (8.26)$$

где нормирующий коэффициент

$$B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)} \Gamma\left(\frac{n-1}{2}\right)}. \quad (8.27)$$

В частности,

$$B_4 = \frac{\Gamma(2)}{\sqrt{3\pi} \cdot \Gamma(3/2)} = \frac{1!}{\sqrt{3\pi} \cdot \frac{1}{2} \Gamma\left(\frac{1}{2}\right)} = \frac{2}{\pi\sqrt{3}} \text{ и}$$

$$S_4(t) = \frac{2}{\pi\sqrt{3}} \left(1 + \frac{t^2}{3}\right)^{-2} = \frac{6\sqrt{3}}{\pi(t^2 + 3)^2}.$$

Мы не будем выводить формулу (8.26), хотя все методы, необходимые для вывода, были изложены выше.

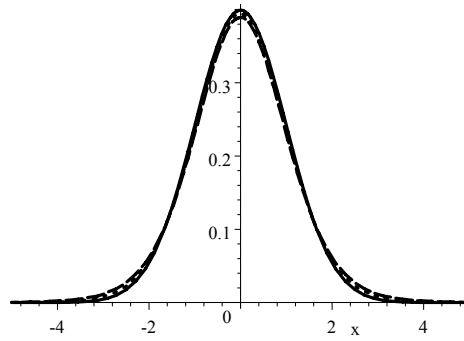


Рис. 8.5

При достаточно большом числе степеней свободы n распределение Стьюдента приближается к стандартному нормальному распределению. В подтверждение этого на рис. 8.5 приведены графики функций $S_n(t)$ для значений n , равных 10 (пунктир) и 20 (точки), и кривая Гаусса (сплошная). На рисунке видно, что даже при достаточно крупном масштабе по вертикали три кривые почти сливаются. Тем не менее для функции распределения Стьюдента обычно приводится таблица наиболее часто используемых значений.

Нормальное распределение на плоскости

§ 9.1. Определение и основные свойства

Определение. Случайный вектор $X = (X_1, X_2)$ называют *распределенным по нормальному закону* (с центром в начале координат), если его функция плотности имеет вид

$$f(x_1, x_2) = \frac{1}{\gamma} e^{-\frac{q(x_1, x_2)}{2}}, \quad (9.1)$$

где γ – некоторая постоянная, $q(x_1, x_2)$ – положительно определенная квадратичная форма.

Заметим, что данное определение является прямым обобщением определения нормально распределенной случайной величины в одномерном случае с плотностью

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}. \quad (9.2)$$

Если центр распределения находится в точке (m_1, m_2) , то плотность определяется из (9.1) как $f(x_1 - m_1, x_2 - m_2)$.

Вид константы, так же как и вероятностное истолкование коэффициентов квадратичной формы q , мы найдем, если воспользуемся теорией приведения квадратичных форм к главным осям.

Здесь нам лучше перейти к матричным обозначениям. Обозначим через $\vec{x} = (x_1, x_2)$ – вектор-строку координат в $\mathbf{R}^2$, Q – матрицу квадратичной формы, так что квадратичная форма q запишется так:

$$q(\vec{x}) = \vec{x} Q \vec{x}^T. \quad (9.3)$$

Напомним, что в курсе линейной алгебры доказывается теорема о приведении квадратичной формы (9.3) к главным осям. А именно, найдется ортогональное преобразование O такое, что квадратичная форма в новых координатах $\vec{y} = (y_1, y_2)$

$$\vec{y} = \vec{x} O \quad (9.4)$$

приводится к сумме квадратов. При этом матрица квадратичной формы в новых координатах

$$Q' = O^T Q O \quad (9.5)$$

примет диагональный вид.

Применим изложенную теорию к функции плотности (9.1). Для этого воспользуемся формулами (7.46), (7.47) из главы 7. Поскольку определитель ортогональной матрицы равен ± 1 , то модуль якобиана линейного преобразования равен 1, так что функция плотности в новых координатах получается подстановкой

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(b_{11}y_1 + b_{12}y_2, b_{21}y_1 + b_{22}y_2).$$

Поэтому левую часть можно записать в виде, аналогичном (9.2):

$$f_{Y_1, Y_2}(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{y_1^2}{2\sigma_1^2} - \frac{y_2^2}{2\sigma_2^2}}. \quad (9.6)$$

Поскольку функция совместной плотности в (9.6) представляется в виде произведения функций плотности каждого аргумента, то случайные величины Y_1, Y_2 сами нормальны и независимы. Таким образом, справедлива

Теорема 9.1. Существует ортогональное преобразование с матрицей O такое, что компоненты случайного вектора $Y = XO$ – независимые нормально распределенные случайные величины.

Обозначим через $C(Y)$ матрицу ковариаций случайного вектора $Y = (Y_1, Y_2)$. Легко видеть, что $C(Y)$ имеет диагональный вид

$$C(Y) = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}. \quad (9.7)$$

Поскольку

$$C(Y) = O^T C(X) O, \quad (9.8)$$

а по таким же формулам преобразуется квадратичная форма q , то первый вывод: *матрицей квадратичной формы q в (9.1) является матрица, обратная к матрице ковариаций случайной величины X .*

Теперь исследуем вид константы γ в (9.1). Мы уже отмечали, что при линейном преобразовании O константа в формуле совместной плотности не меняется. Для случайного вектора Y в (9.6) выполняется соотношение

$$\gamma^2(Y) = 2\pi \det C(Y).$$

Из формулы (9.8) следует соотношение для определителей

$$\det C(Y) = (\det O)^2 \det C(X) = \det C(X),$$

откуда следует, что

$$\gamma^2(X) = \gamma^2(Y) = 2\pi \det C(Y) = 2\pi \det C(X). \quad (9.9)$$

Получаем второй вывод: константа γ имеет вид, заданный формулой (9.9). Ввиду важности полученных утверждений сформулируем теорему.

Теорема 9.2. *Функция плотности нормального вектора $X = (X_1, X_2)$ с матрицей ковариаций C имеет вид*

$$f_X(x_1, x_2) = \frac{1}{2\pi\sqrt{\det C}} e^{-\frac{\bar{x}^T C^{-1} \bar{x}}{2}}. \quad (9.10)$$

Эту формулу можно записать, не используя матричных обозначений. Имеем формулу для ковариационной матрицы

$$C = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad (9.11)$$

где ρ – коэффициент корреляции случайных величин X_1 и X_2 . Отсюда

$$\det C = (1 - \rho^2) \sigma_1^2 \sigma_2^2, \quad (9.12)$$

$$C^{-1} = \frac{1}{(1 - \rho^2) \sigma_1^2 \sigma_2^2} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}. \quad (9.13)$$

Подставляя (9.12), (9.13) в формулу (9.10), получим явное выражение для функции плотности нормального вектора:

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left(\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right). \quad (9.14)$$

Заменяя x_i на $x_i - m_i$, мы получим общую формулу для нормальной плотности с центром в (m_1, m_2)

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{q}{2(1-\rho^2)}\right),$$

$$q = \left(\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2}\right). \quad (9.15)$$

Заметим, что если $\rho = 0$, т.е. случайные величины X_1 и X_2 *некоррелированы*, (9.14) превращается в формулу (9.6) для *независимых величин*. Таким образом, справедлива

Теорема 9.3. Для нормально распределенного случайного вектора $X = (X_1, X_2)$ понятия независимости и некоррелированности компонент X_1 и X_2 эквивалентны.

Отметим также важный случай $|\rho| = 1$, когда формулу (9.14) использовать нельзя, т.е. двумерная плотность не существует. Из анализа коэффициента корреляции мы знаем, что условие $|\rho| = 1$ эквивалентно линейной зависимости компонент X_1 и X_2 . В этом случае, хотя обе компоненты могут быть распределены нормально, однако распределение пары (X_1, X_2) сосредоточено на прямой и, следовательно, вырожденно.

§ 9.2. Условные нормальные распределения

Теорема 9.4. Для нормального вектора $X = (X_1, X_2)$ условная плотность одной из компонент при фиксированном значении другой является нормальной.

Доказательство. Переходя к отклонениям $\overset{\circ}{X}_i = X_i - m_i$, $m_i = M(X_i)$, если это необходимо, будем считать вектор X центрированным. Будем считать для определенности, что фиксировали значение второй компоненты $X_2 = x_2$. Выделим полный квадрат по переменной x_1 , все остальное отнесем к константе.

Напишем явное выражение для полученной плотности. Чтобы найти константу, воспользуемся выражением (7.36) для нахождения условной плотности. Имеем

$$\lambda = \frac{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}{\sigma_2\sqrt{2\pi}} = \sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}.$$

Заметим, что выделить полный квадрат из квадратичной формы

$$ax_1^2 - 2bx_1x_2 + cx_2^2$$

проще всего по формуле $a\left(x_1 - \frac{b}{a}x_2\right)^2 + \dots$. В соответствии с формулой (9.14) положим $a = \frac{1}{\sigma_1^2}$,

$b = \frac{\rho}{\sigma_1\sigma_2}$ и получим $\frac{1}{\sigma_1^2}\left(x_1 - \rho\frac{\sigma_1}{\sigma_2}x_2\right)^2 + \dots$. Таким образом, явное выражение для условной плотности выглядит следующим образом

$$f(x_1 | x_2) = \frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2\sigma_1^2(1-\rho^2)}\left(x_1 - \rho\frac{\sigma_1}{\sigma_2}x_2\right)^2\right), \quad (9.16)$$

что и заканчивает доказательство теоремы.

В общем случае произведем замену x_i на $x_i - m_i$ и получим

$$f(x_1 | x_2) = \frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2\sigma_1^2(1-\rho^2)}\left(x_1 - m_1 - \rho\frac{\sigma_1}{\sigma_2}(x_2 - m_2)\right)^2\right). \quad (9.17)$$

Аналогично выглядит формула, если поменять местами переменные x_1 и x_2 :

$$f(x_2 | x_1) = \frac{1}{\sigma_2\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(x_2 - m_2 - \rho\frac{\sigma_2}{\sigma_1}(x_1 - m_1)\right)^2\right). \quad (9.18)$$

Из формул (9.17), (9.18) легко получаются выражения для условных математических ожиданий и дисперсий.

$$M(X_1 | X_2) = m_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - m_2), \quad (9.19)$$

$$D(X_1 | X_2) = \sigma_1^2(1-\rho^2), \quad (9.20)$$

$$M(X_2 | X_1) = m_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - m_1), \quad (9.21)$$

$$D(X_2 | X_1) = \sigma_2^2(1-\rho^2). \quad (9.22)$$

Обратим внимание читателя, что условные математические ожидания представляют из себя *линейные функции*.

§ 9.3. Геометрическая интерпретация двумерного нормального закона

Рассмотрим поверхность распределения, представляющую график функции (9.16). Она имеет вид холма с вершиной в точке (m_1, m_2) (см. рис. 9.1).

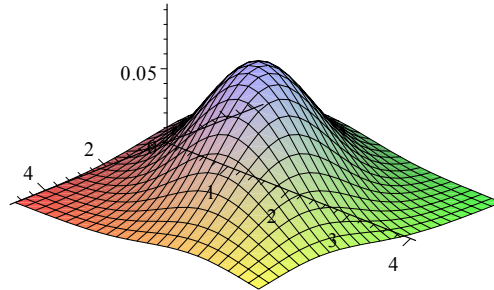


Рис. 9.1

В сечениях этой поверхности вертикальными плоскостями получим кривые, подобные кривым нормальной плотности в одномерном случае. В сечениях горизонтальными плоскостями будут получаться эллипсы. Запишем уравнение такого эллипса, который называется *эллипсом рассеивания*

$$\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} = c^2, \quad (9.23)$$

где c – некоторая константа. Мы уже видели, что в главных осях $O\xi_1\xi_2$ уравнение (9.23) принимает канонический вид

$$\frac{\xi_1^2}{\sigma_{\xi_1}^2} + \frac{\xi_2^2}{\sigma_{\xi_2}^2} = c^2. \quad (9.24)$$

При этом к уравнению (9.24) можно перейти с помощью переноса начала координат в точку (m_1, m_2) и поворота осей координат на угол α (см. рис. 9.2), который находится из соотношения

$$\operatorname{tg} 2\alpha = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}. \quad (9.25)$$

При $\sigma_1 = \sigma_2$, т.е. в случае, когда эллипс рассеивания превращается в окружность, угол $\alpha = \frac{\pi}{4}$.

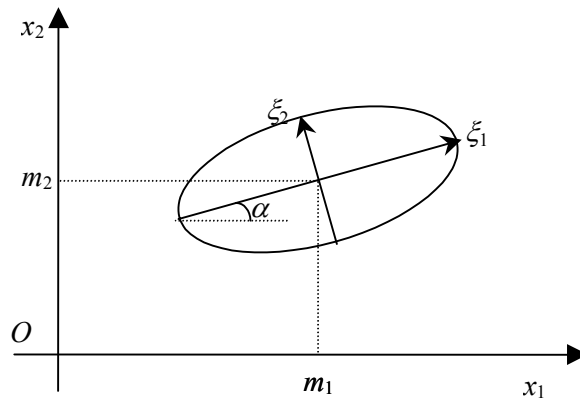


Рис. 9.2

§ 9.4. Многомерный нормальный закон

Мы вели изложение нормального закона на плоскости таким образом, чтобы обобщение на n -мерный случай получалось непосредственно. Поэтому мы приведем соответствующие формулировки без особых комментариев.

Определение. *Случайный вектор $X = (X_1, \dots, X_n)$ называют распределенным по **нормальному закону** (с центром в начале координат), если его функция плотности имеет вид*

$$f(x_1, \dots, x_n) = \frac{1}{\gamma} e^{-\frac{q(x_1, \dots, x_n)}{2}}, \quad (9.26)$$

где γ – некоторая постоянная, $q(x_1, \dots, x_n)$ – положительно определенная квадратичная форма.

Если центр распределения находится в точке $(m_1, \dots, m_n)$, то плотность определяется из (9.26) как $f(x_1 - m_1, \dots, x_n - m_n)$.

Так же, как и в двумерном случае, справедлива

Теорема 9.5. *Функция плотности нормального вектора $X = (X_1, \dots, X_n)$ с матрицей ковариаций C имеет вид*

$$f_X(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det C}} e^{-\frac{\bar{x}^T C^{-1} \bar{x}}{2}}. \quad (9.27)$$

Условием существования плотности (9.27) является невырожденность матрицы ковариаций C , что эквивалентно условию

$$\det C \neq 0.$$

Предмет теории вероятностей, как мы знаем, составляют закономерности, свойственные массовым случайным событиям. Простейшая из них – устойчивость частоты – лежит в основе всех приложений теории вероятностей к практике. Если попытаться в немногих словах отразить общий смысл подобных закономерностей, то придем к такому заключению. Пусть производится большая серия однотипных опытов. Исход каждого отдельного опыта является случайным, неопределенным. Однако, несмотря на это, средний результат всей серии опытов утрачивает случайный характер, становится закономерным. Под законом больших чисел в теории вероятностей понимается ряд теорем, в каждой из которых устанавливается факт приближения средних характеристик большого числа опытов к некоторым определенным постоянным.

В основе доказательства этих теорем лежит важное неравенство, установленное в 1845 г. П.Л. Чебышевым.

§ 10.1. Неравенство Чебышева

Пусть X – случайная величина с математическим ожиданием m . Выберем какое-либо положительное число ε и рассмотрим событие

$$|X - m| \geq \varepsilon. \quad (10.1)$$

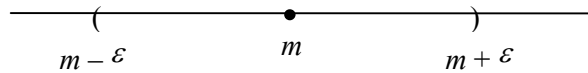


Рис. 10.1

Геометрический смысл этого события заключается в том, что значение случайной величины X попадает в область на числовой оси, получающуюся удалением из всей оси интервала от $m - \varepsilon$ до $m + \varepsilon$ (см. рис. 10.1). С возрастанием ε эта область сужается, следовательно, вероятность попадания в нее (т.е. вероятность события (10.1)) становится все меньше. Неравенство Чебышева замечательно тем, что устанавливает для этой вероятности весьма простую оценку.

Неравенство Чебышева. Пусть имеется случайная величина X с математическим ожиданием m и дисперсией D . Каково бы ни было положительное число ε , вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на ε , ограничена сверху числом $\frac{D}{\varepsilon^2}$:

$$P(|X - m| \geq \varepsilon) \leq \frac{D}{\varepsilon^2}. \quad (10.2)$$

Доказательство. Неравенство (10.2) является следствием другого неравенства, также принадлежащего Чебышеву: если случайная величина X может принимать только неотрицательные значения (т.е. $P(X < 0) = 0$), то вероятность того, что принятое ею значение окажется не меньше единицы, не превосходит числа m – математического ожидания X :

$$P(X \geq 1) \leq m. \quad (10.3)$$

Докажем сначала неравенство (10.3).

Пусть X – дискретная случайная величина, принимающая неотрицательные значения $x_1, x_2, \dots$ с вероятностями $p_1, p_2, \dots$ соответственно. Имеем:

$$P(X \geq 1) = \sum_{x_i \geq 1} p_i.$$

Если каждое слагаемое p_i суммы, стоящей справа, умножить на соответствующее значение x_i , то правая часть не уменьшится, поскольку $x_i \geq 1$. Мы придем тогда к неравенству

$$P(X \geq 1) \leq \sum_{x_i \geq 1} x_i p_i.$$

Выражение, стоящее справа, опять-таки не уменьшится, если распространить суммирование на все возможные значения X : действительно, при этом добавятся неотрицательные слагаемые $x_i p_i$, отвечающие таким номерам i , для которых $x_i < 1$. Итак,

$$P(X \geq 1) \leq \sum_i x_i p_i.$$

Но последняя сумма по определению совпадает с $M(X)$. Тем самым неравенство (10.3) доказано для случая дискретной величины ξ .

Пусть теперь X – какая угодно случайная величина, удовлетворяющая, однако, условию $P(X < 0) = 0$ и имеющая математическое ожидание m . Взяв какое-либо $\varepsilon > 0$, рассмотрим случайную величину $X_\varepsilon - \varepsilon$ – приближение к X (см. § 5.3, п. 1°). По доказанному $P(X_\varepsilon \geq 1) \leq M(X_\varepsilon)$. Полученное неравенство имеет место для любого $\varepsilon > 0$, в частности для $\varepsilon = \frac{1}{n}$, где n – какое угодно натуральное число. Поскольку в этом случае $P(X_\varepsilon \geq 1) = P(X \geq 1)$ (объясните!), то при любом натуральном n имеем: $P(X \geq 1) \leq M(X_\varepsilon)$. При $n \rightarrow \infty$ правая часть этого неравенства стремится к пределу, равному $M(X)$. Отсюда и следует (10.3).

Мы можем теперь вернуться к рассмотрению неравенства (10.2). Фигурирующее в нем событие $|X - m| \geq \varepsilon$ равносильно

$$\frac{(X - m)^2}{\varepsilon^2} \geq 1.$$

Случайная величина $\frac{(X - m)^2}{\varepsilon^2}$ принимает лишь неотрицательные значения, и к ней можно применить неравенство (10.3). Получим:

$$P(|X - m| \geq \varepsilon) = P\left(\frac{(X - m)^2}{\varepsilon^2} \geq 1\right) \leq M\left[\frac{(X - m)^2}{\varepsilon^2}\right] = \frac{1}{\varepsilon^2} M[(X - m)^2] = \frac{D}{\varepsilon^2},$$

или в конечном итоге

$$P(|X - m| \geq \varepsilon) \leq \frac{D}{\varepsilon^2}.$$

Неравенство Чебышева (10.2) можно записать в эквивалентной форме, если перейти от события $|X - m| \geq \varepsilon$ к противоположному событию $|X - m| < \varepsilon$. Тогда получим:

$$P(|X - m| < \varepsilon) > 1 - \frac{D}{\varepsilon^2}. \quad (10.4)$$

В качестве приложения оценим вероятность события

$$|X - m| < 3\sigma,$$

где σ – среднее квадратичное отклонение величины X .

Полагая в (10.4) $\varepsilon = 3\sigma$, получим в правой части неравенства число $1 - \frac{D}{9\sigma^2}$, т.е. $\frac{8}{9}$. Таким образом, в общем случае справедливо утверждение.

Правило 3σ. Для любой случайной величины X выполняется неравенство

$$P(|X - m| < 3\sigma) \geq \frac{8}{9}. \quad (10.5)$$

В действительности для подавляющего большинства встречающихся на практике случайных величин эта вероятность значительно ближе к единице, чем $\frac{8}{9}$. Например, для нормального распределения, как мы видели, она равна 0,997....

Обычно, если закон распределения случайной величины неизвестен, но указаны параметры m и σ , принимают, что диапазон практически возможных значений случайной величины есть интервал $(m - 3\sigma, m + 3\sigma)$ ("правило трех сигм").

§ 10.2. Различные формы закона больших чисел

1°. Теорема Чебышева. Неравенство Чебышева позволяет доказать ряд важных теорем, объединенных одним общим названием "закон больших чисел". Основная из этих теорем принадлежит самому П.Л. Чебышеву.

Теорема 10.1. (теорема Чебышева). Пусть имеется бесконечная последовательность $X_1, X_2, \dots$ независимых случайных величин с одним и тем же математическим ожиданием m и дисперсиями, ограниченными одной и той же постоянной:

$$M(X_1) = M(X_2) = \dots = m, \\ D(X_1) < C, D(X_2) < C, \dots$$

Тогда, каково бы ни было положительное число ε , вероятность события

$$\left| \frac{X_1 + X_2 + \dots + X_n}{n} - m \right| < \varepsilon$$

стремится к единице при $n \rightarrow \infty$.

Доказательство. Положим,

$$s_n = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

В силу свойств математического ожидания имеем:

$$M(X_n) = \frac{1}{n} \sum_{i=1}^n M(X_i) = \frac{nm}{n} = m.$$

Далее, так как величины $X_1, X_2, \dots, X_n$ независимы, то

$$D(s_n) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) \leq \frac{nc}{n^2} = \frac{c}{n}.$$

Сопоставив полученное неравенство с неравенством Чебышева:

$$P(|s_n - M(s_n)| < \varepsilon) > 1 - \frac{D(s_n)}{\varepsilon^2},$$

будем иметь:

$$P(|s_n - m| < \varepsilon) > 1 - \frac{c}{n\varepsilon^2}. \quad (10.6)$$

Это показывает, что с ростом n вероятность события $|s_n - m| < \varepsilon$ стремится к 1.

Смысл теоремы Чебышева можно пояснить следующим примером. Пусть требуется измерить некоторую физическую величину m . В силу неизбежных ошибок результат измерения будет случайной величиной. Обозначим эту величину X ; ее математическое ожидание будет совпадать с измеряемой величиной m , а дисперсия равна некоторой величине D (характеризующей точность измерительного прибора). Произведем n независимых измерений и обозначим:

X_1 – результат первого измерения;

X_2 – результат второго измерения

и т.д. Совокупность величин $X_1, X_2, \dots, X_n$ представляет собой систему n независимых случайных величин, каждая из которых имеет тот же закон распределения, что и сама величина X . Среднее арифметическое s_n этих величин тоже является, конечно, случайной величиной. Однако с увеличением n эта величина почти перестает быть случайной, она все более приближается к постоянной m . Точная количественная формулировка этой близости состоит в том, что событие $|s_n - m| < \varepsilon$ становится как угодно достоверным при достаточно большом n .

Тем самым оправдывается рекомендуемый в практической деятельности способ получения более точных результатов измерений: одна и та же величина измеряется многократно, и в качестве ее значения берется среднее арифметическое полученных результатов измерений.

Близость к $M(X)$ среднего арифметического опытных значений величины X уже подчеркивалась нами при самом введении понятия математического ожидания. Однако соответствующее рассуждение относилось только к дискретным случайным величинам; кроме этого, само высказывание о близости мотивировалось соображениями эмпирического характера. В противоположность этому теорема Чебышева дает точную характеристику близости среднего арифметического к $M(X)$, и притом для любой случайной величины.

2°. Теорема Бернулли. Из теоремы Чебышева в качестве следствия можно получить другую важную теорему, которая впервые была доказана Я. Бернулли и опубликована в 1713 году.

Теорема 10.2. (теорема Бернулли). Пусть производится n независимых опытов, в каждом из которых с вероятностью p может наступить некоторое событие A . Рассмотрим случайную величину k – число наступлений события A в n опытах. Каково бы ни было положительное число ε , вероятность события

$$\left| \frac{k}{n} - p \right| < \varepsilon$$

стремится к единице при $n \rightarrow \infty$.

Иначе говоря, как бы ни было мало ε , с увеличением числа опытов становится сколь угодно достоверным тот факт, что частота наступления события A отличается от вероятности этого события меньше чем на ε .

Чтобы вывести теорему Бернулли из теоремы Чебышева, достаточно заметить, что

$$k = X_1 + X_2 + \dots + X_n,$$

где X_i есть индикатор наступления события A в i -м опыте ($i=1,2,\dots,n$). Случайные величины $X_1, X_2, \dots, X_n$ имеют один и тот же закон распределения с математическим ожиданием $M(X_i) = p$ и дисперсией $D(X_i) = pq$.

Таким образом, все условия теоремы Чебышева выполняются, и для среднего арифметического величин $X_1, X_2, \dots, X_n$, т.е. для $\frac{k}{n}$, справедливо предельное соотношение

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 1.$$

Отметим попутно следующий факт. Поскольку

$$M\left(\frac{k}{n}\right) = \frac{1}{n} M(k) = \frac{1}{n} np = p,$$

$$D\left(\frac{k}{n}\right) = \frac{1}{n^2} npq = \frac{pq}{n},$$

то неравенство Чебышева применительно к случайной величине $\frac{k}{n}$ дает:

$$\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) > 1 - \frac{pq}{n\varepsilon^2}. \quad (10.7)$$

Мы получаем, таким образом, оценку, хотя и весьма грубую, для вероятности отклонения частоты события A в серии из n опытов от вероятности события A в одном опыте.

3°. Заключительные замечания. Возвращаясь к вопросу о практическом значении изложенных выше теорем, заметим следующее. В каждой области человеческой деятельности, как правило, существует определенный "уровень значимости" вероятностей, т.е. такое граничное число α , что события, имеющие вероятность меньшую α , считаются практически невозможными. Если в качестве примера взять, скажем, производство изделий бытового назначения, то указанный уровень будет находиться где-то в пределах 0,01 – 0,001. Действительно, известие о том, что вероятность брака в изделии имеет порядок 0,001, не заставит покупателя тревожиться за свою покупку; вероятность 0,001 мала настолько, что можно рассматривать бракованную лампу, часы, велосипед, пылесос и т.п. как практически невозможное явление. Однако та же самая вероятность 0,001 уже не устроила бы нас,

если бы речь шла, скажем, об изготовлении частей авиационного двигателя; в этом случае уровень значимости должен быть много меньше чем 0,001.

Если подойти с этой точки зрения, например, к неравенству (10.7), то можно сделать следующее заключение. Пусть дано число $\varepsilon > 0$. Выберем n столь большим, чтобы величина $\frac{pq}{n\varepsilon^2}$ была меньше того уровня значимости α , о котором шла речь выше. При таком выборе n событие

$$\left| \frac{k}{n} - p \right| < \varepsilon$$

будет практически достоверным. Иначе говоря, мы можем быть практически уверены, что при столь большом числе опытов частота наступления события A будет с точностью ε совпадать с его вероятностью.

Обратим внимание читателя еще раз на один аспект теоремы Бернулли. Он связан с сопоставлением двух подходов к понятию вероятности: статистического, основанного на понятии относительной частоты и не являющегося строго математическим, и чисто математического подхода, базирующегося на рассмотрении аксиом (во времена Я. Бернулли такой подход был неизвестен). Представим себе, что на множестве всех событий, связанных с *данным* опытом, определена функция $P(A)$, удовлетворяющая всем аксиомам теории вероятностей (аксиомам 1–3 § 1.6, п. 3°). *Вслед за аксиомами будут, конечно, выполняться и их следствия, в частности теорема Бернулли.* Возникает вопрос, соответствуют ли числа $P(A)$ реальным относительным частотам событий. Оказывается, что близость числа $P(A)$ к относительной частоте вовсе не обязательно проверять для всякого события A ; достаточно убедиться в такой близости лишь для некоторых A . А именно, рассмотрим те события A , для которых $P(A)$ близко к 1. Тогда можно утверждать следующее: если для таких «формально достоверных» событий числа $P(A)$ будут близки к относительным частотам (это означает, что такие события будут практически достоверными), то и для любых событий числа $P(A)$ будут близки к относительным частотам. Действительно, из теоремы Бернулли будет тогда следовать, что для всякого события A близость между $P(A)$ и относительной частотой этого события (при большом числе опытов) является событием практически достоверным.

Центральная предельная теорема

До сих пор мы говорили об устойчивости средних характеристик большого числа опытов, точнее, об устойчивости сумм вида

$$s_n = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Однако не следует все же забывать, что величина s_n — случайная, а значит, она имеет некоторый закон распределения. Оказывается — и этот замечательный факт составляет содержание группы теорем, объединяемых общим названием "*центральная предельная теорема*", — что при весьма общих условиях закон распределения s_n близок к нормальному закону.

Существенный вклад в формулировку и доказательство центральной предельной теоремы внес русский ученый А.А. Ляпунов. Он же разработал метод *характеристических функций*, широко используемый в теории вероятностей.

§ 11.1. Характеристические функции

Определение. Пусть X — абсолютно непрерывная случайная величина с плотностью $f(x)$. *Характеристической функцией* случайной величины X называется функция, определенная для действительных значений t формулой

$$\varphi_X(t) = M(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx, \quad (11.1)$$

где i — мнимая единица.

Поскольку по теореме Эйлера

$$e^{itx} = \cos(tx) + i \sin(tx),$$

то характеристическую функцию можно записать в следующем виде

$$\varphi_X(t) = u(t) + iv(t),$$

где

$$u(t) = \int_{-\infty}^{\infty} \cos(tx) f(x) dx, \quad v(t) = \int_{-\infty}^{\infty} \sin(tx) f(x) dx. \quad (11.2)$$

Замечание. В курсе математического анализа формула (11.1) определяет преобразование Фурье функции $f(x)$. Там же доказано, что если $\varphi_X(t)$ выражается через функцию $f(x)$ с помощью преобразования Фурье, то и $f(x)$ находится из $\varphi_X(t)$ с помощью обратного преобразования Фурье

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt. \quad (11.3)$$

В частности, функция плотности однозначно восстанавливается по своей характеристической функции.

Свойства характеристических функций.

1. Пусть случайная величина Y пропорциональна случайной величине X :

$$Y = aX.$$

Тогда

$$\varphi_{aX}(t) = \varphi_X(at). \quad (11.4)$$

Действительно,

$$\varphi_{aX}(t) = M(e^{itaX}) = M(e^{i(at)X}) = \varphi_X(at).$$

2. Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых.

Пусть $X_1, X_2, \dots, X_n$ – независимые случайные величины с характеристическими функциями

$$\varphi_{X_1}(t), \varphi_{X_2}(t), \dots, \varphi_{X_n}(t).$$

Имеем

$$\varphi_{X_1+X_2+\dots+X_n}(t) = M\left(e^{it(X_1+X_2+\dots+X_n)}\right) = M\left(e^{itX_1} \cdot e^{itX_2} \cdot \dots \cdot e^{itX_n}\right).$$

Поскольку случайные величины X_i независимы, то все множители e^{itX_k} , $k = 1, \dots, n$ также будут независимы. Используя свойство математического ожидания, получим

$$\begin{aligned} \varphi_{X_1+X_2+\dots+X_n}(t) &= M\left(e^{itX_1}\right) \cdot M\left(e^{itX_2}\right) \cdot \dots \cdot M\left(e^{itX_n}\right) = \\ &= \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \dots \cdot \varphi_{X_n}(t), \end{aligned}$$

что и требовалось доказать.

3. *Характеристическая функция выражается через производящую функцию моментов*

$$\varphi_X(t) = \psi_X(it). \quad (11.5)$$

По определению,

$$\psi_X(t) = M\left(e^{tX}\right),$$

откуда (11.5) следует немедленно.

Пример 11.1. Найти характеристическую функцию стандартной нормальной случайной величины $X \sim N(0, 1)$.

Решение. Мы уже находили производящую функцию моментов X

$$\psi_X(t) = e^{\frac{t^2}{2}}.$$

Заменяя t на it и учитывая, что $i^2 = -1$, получим

$$\varphi_X(t) = e^{-\frac{t^2}{2}}. \quad (11.6)$$

Из свойства 2 следует, что характеристическая функция свертки двух функций равна произведению характеристических функций сомножителей.

Пример 11.2. Найти характеристическую функцию суммы $Z = X + Y$, где $X \sim N(0, \sigma_1)$, $Y \sim N(0, \sigma_2)$ – нормальные величины.

Решение. Сначала найдем характеристическую функцию для одного слагаемого. Мы знаем, что $X = \sigma U$, где $U \sim N(0, 1)$ – стандартная нормальная случайная величина. На основании свойства 1 и предыдущего примера имеем

$$\varphi_X(t) = \varphi_U(\sigma t) = e^{-\frac{(\sigma t)^2}{2}} = e^{-\frac{\sigma^2 t^2}{2}}.$$

Взяв произведение характеристических функций слагаемых, получим

$$\varphi_Z(t) = \varphi_X(t) \varphi_Y(t) = e^{-\frac{\sigma_1^2 t^2}{2}} \cdot e^{-\frac{\sigma_2^2 t^2}{2}} = e^{-\frac{(\sigma_1^2 + \sigma_2^2) t^2}{2}}. \quad (11.7)$$

Мы видим, что (11.7) дает характеристическую функцию нормальной случайной величины $Z \sim N\left(0, \sqrt{\sigma_1^2 + \sigma_2^2}\right)$, т.е. *сумма нормальных случайных величин распределена по нормальному закону.*

§ 11.2. Центральная предельная теорема для одинаково распределенных слагаемых

Так как величина s_n лишь постоянным множителем отличается от суммы:

$$X_1 + X_2 + \dots + X_n, \quad (11.8)$$

то в общих чертах содержание центральной предельной теоремы может быть высказано следующим образом.

Распределение суммы большого числа независимых случайных величин при весьма общих условиях близко к нормальному распределению.

Этим и определяется особая роль нормального закона, поскольку с суммами большого числа случайных слагаемых приходится иметь дело весьма часто как в самой теории вероятностей, так и в ее многочисленных приложениях. Чуть позже мы поясним этот факт примерами из разных областей.

Чтобы подготовить точную формулировку центральной предельной теоремы, поставим два вопроса.

1. Какой точный смысл содержится в утверждении о том, что закон распределения суммы (11.8) близок к нормальному закону?
2. При каких условиях имеет место эта близость?

Чтобы ответить на эти вопросы, рассмотрим не просто большое число, а *бесконечную последовательность* случайных величин:

$$X_1, X_2, \dots, X_n, \dots$$

Составим "частичные" суммы:

$$S_n = X_1 + X_2 + \dots + X_n \quad (n=1, 2, \dots). \quad (11.9)$$

От каждой из случайных величин S_n перейдем к "нормированной" случайной величине (см. §4.3, п. 3°):

$$S'_n = \frac{S_n - M(S_n)}{\sqrt{D(S_n)}}, \quad (11.10)$$

математическое ожидание которой равно 0, а дисперсия 1. Ответ на первый из поставленных вопросов теперь можно сформулировать в виде равенства

$$\lim_{n \rightarrow \infty} P(a \leq S'_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad (11.11)$$

(где a и b – любые, $a < b$), означающего, что закон распределения случайной величины S'_n с ростом n приближается к нормальному закону распределения с математическим ожиданием 1 и дисперсией 0.

Если обозначить через $F_n(x)$ функцию распределения для каждой из случайных величин S'_n , $F_0(x)$ – функцию распределения стандартного нормального закона, то условие (11.10) можно выразить по другому, а именно, для любого x

$$F_n(x) \xrightarrow{n \rightarrow \infty} F_0(x). \quad (11.12)$$

В курсе математического анализа устанавливается, что условие (11.12) можно проверять для последовательности соответствующих характеристических функций:

$$\varphi_n(x) \xrightarrow{n \rightarrow \infty} \varphi_0(x). \quad (11.13)$$

Разумеется, из того факта, что величина S'_n имеет приближенно нормальное распределение, следует, что и величина S_n распределена приближенно нормально. Ведь связь между S'_n и S_n линейная:

$$S'_n = \alpha_n S_n + \beta_n \quad \left(\text{где } \alpha_n = \frac{1}{\sqrt{D(S_n)}}, \beta_n = -\frac{M(S_n)}{\sqrt{D(S_n)}} \right),$$

а, как мы знаем, нормальный закон распределения сохраняется при линейном преобразовании случайной величины.

Все эти предварительные замечания позволяют доказать нам следующую теорему.

Теорема 11.1. *Если $X_1, X_2, \dots, X_n, \dots$ – последовательность независимых случайных величин, имеющих один и тот же закон распределения с математическим ожиданием m и дисперсией σ^2 , то при неограниченном увеличении n закон распределения нормализованной суммы (11.10) неограниченно приближается к нормальному.*

Д о к а з а т е л ь с т в о . Рассуждения, предшествовавшие формулировке теоремы, показывают, что нам достаточно рассмотреть случай нормированных слагаемых X_k , так что можно положить $m = 0$, $\sigma = 1$. Поскольку в этом случае $\sigma(S_n) = \sqrt{n}$, то случайная величина

$$Y_n = \frac{S_n}{\sqrt{n}} \quad (11.14)$$

будет нормированной: $M(Y_n) = 0$, $\sigma(Y_n) = 1$.

Пусть

$$\varphi(t) = \varphi_{X_k}(t), \varphi_n(t) = \varphi_{Y_n}(t)$$

– характеристические функции слагаемых X_k и случайной величины Y_n соответственно. Для доказательства теоремы достаточно проверить, что

$$\lim_{n \rightarrow \infty} \varphi_n(t) = e^{-\frac{t^2}{2}}. \quad (11.15)$$

Используя свойства 1, 2 характеристических функций, получим

$$\varphi_n(t) = \left(\frac{t}{\sqrt{n}} \right)^n. \quad (11.16)$$

Разложим функцию $\varphi(t)$ в ряд Маклорена до второго порядка включительно. Для этого воспользуемся разложением в ряд производящей функции моментов (7.18) и соотношением (11.5):

$$\psi(t) = 1 + \nu_1 t + \frac{\nu_2}{2} t^2 + \alpha(t) t^2, \quad (11.17)$$

где $\alpha(t) \rightarrow 0$ когда $t \rightarrow 0$. Поскольку $m = 0$, $\sigma = 1$, а $\nu_1 = m$, $\nu_2 = m^2 + \sigma^2$, то разложение (11.17) упрощается до следующего

$$\psi(t) = 1 + \frac{t^2}{2} + \alpha(t) t^2. \quad (11.18)$$

Заменяя t на it , получим разложение для характеристической функции

$$\varphi(t) = 1 - \frac{t^2}{2} + \alpha(t) t^2. \quad (11.19)$$

Подставим (11.19) в выражение (11.16) и получим

$$\varphi_n(t) = \left(1 - \frac{t^2}{n} + \alpha\left(\frac{t}{\sqrt{n}}\right) \frac{t^2}{n} \right)^n. \quad (11.20)$$

Перейдем к пределу при $n \rightarrow \infty$. Из второго замечательного предела следует, что

$$\lim_{n \rightarrow \infty} \left(1 + \frac{\beta(n)}{n} \right)^n = e^{\lim_{n \rightarrow \infty} \beta(n)}.$$

Поскольку $\alpha\left(\frac{t}{\sqrt{n}}\right) \rightarrow 0$, когда $n \rightarrow \infty$, то имеем окончательно

$$\lim_{n \rightarrow \infty} \varphi_n(t) = e^{-\frac{t^2}{2}},$$

что и требовалось доказать.

В качестве графической иллюстрации центральной предельной теоремы рассмотрим случай, когда все X_i распределены равномерно на отрезке $[0, 1]$. Математическое ожидание и стандартное отклонение случайной величины S_n легко находятся

$$M(S_n) = \frac{n}{2}, \sigma(S_n) = \frac{\sqrt{n}}{2\sqrt{3}}. \quad (11.21)$$

Построим на одном графике функции плотности S_n и нормальной случайной величины с такими же параметрами. Ниже приведены соответствующие графики (см. рис. 11.1-3) для случаев $n = 2, 5, 10$. В последнем случае обе кривые практически сливаются.

$$n = 2$$

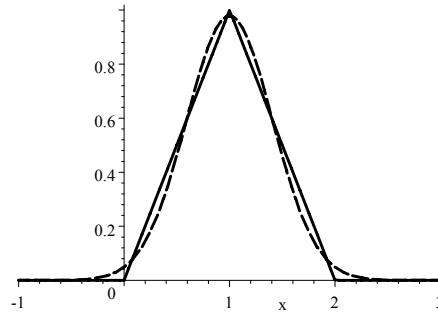


Рис. 11.1

$$n = 5$$

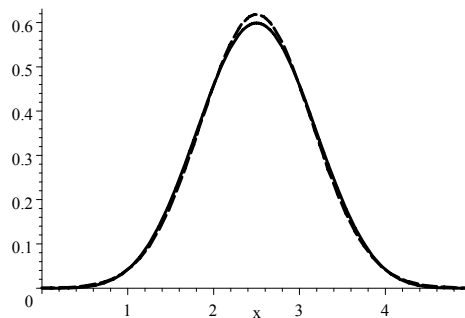


Рис. 11.2

$$n = 10$$

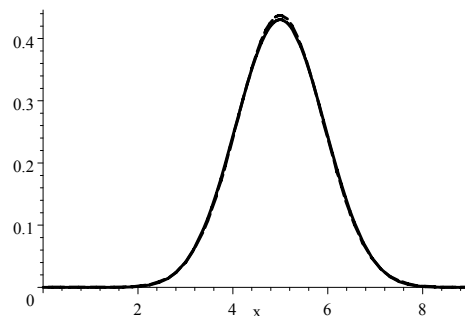


Рис. 11.3

§ 11.3. Центральная предельная теорема в общем случае

По поводу условий, которые следует наложить на величины $X_1, X_2, \dots, X_n$ в общем случае, можно высказать следующие соображения. Вычитая из равенства (11.13) равенство

$$M(S_n) = M(X_1) + M(X_2) + \dots + M(X_n),$$

получим:

$$\overset{\circ}{S}_n = \overset{\circ}{X}_1 + \overset{\circ}{X}_2 + \dots + \overset{\circ}{X}_n,$$

где $\overset{\circ}{X}$, как обычно, означает отклонение случайной величины X от ее математического ожидания. Общий смысл условий, накладываемых на величины $X_1, X_2, \dots$ состоит в том, что *отдельные отклонения $\overset{\circ}{X}_k$ ($k = 1, 2, \dots, n$) должны быть равномерно малы по сравнению с суммарным отклонением $\overset{\circ}{S}_n$.*

Точную формулировку условий, при которых справедливо предельное соотношение (11.10), дал в 1901 году выдающийся русский математик А. М. Ляпунов. Оно заключается в следующем.

Пусть для каждой из величин X_k ($k = 1, 2, \dots$) существуют числа

$$d_k = M(\overset{\circ}{X}_k^2), a_k = M(|\overset{\circ}{X}_k|^3)$$

(заметим, что d_k есть дисперсия случайной величины X_k , а a_k – так называемый "абсолютный центральный момент третьего порядка"). Если при $n \rightarrow \infty$ предел

$$\lim \frac{\sum_{k=1}^n a_k}{\left(\sum_{k=1}^n d_k\right)^{\frac{3}{2}}} = 0,$$

то будем говорить, что последовательность $X_1, X_2, \dots$ удовлетворяет *условию Ляпунова*.

Простейший частный случай, когда выполняется условие Ляпунова, это случай, когда все величины $X_1, X_2, \dots$ имеют *один и тот же* закон распределения. Тогда, разумеется, $d_1 = d_2 = \dots = d$ и $a_1 = a_2 = \dots = a$, следовательно,

$$\frac{\sum_{k=1}^n a_k}{\left(\sum_{k=1}^n d_k\right)^{\frac{3}{2}}} = \frac{na}{(nd)^{\frac{3}{2}}} = \frac{a}{d^{\frac{3}{2}}} \cdot \frac{1}{\sqrt{n}} \rightarrow 0, \text{ когда } n \rightarrow \infty.$$

На этом примере можно хорошо видеть, в чем находит свое проявление равномерная малость слагаемых $\overset{\circ}{X}_k$: величина $\sum_{k=1}^n a_k$ имеет порядок n , а величина $\left(\sum_{k=1}^n d_k\right)^{\frac{3}{2}}$ – порядок $n^{\frac{3}{2}}$, тем самым отношение первой величины ко второй стремится к 0.

Теперь мы в состоянии сформулировать центральную предельную теорему в форме А.М. Ляпунова.

Теорема 11.2. *Если последовательность $X_1, X_2, \dots$ независимых случайных величин удовлетворяет условию Ляпунова, то справедливо предельное соотношение (11.10).*

Иначе говоря, в этом случае закон распределения нормированной суммы X_n сходится к нормальному закону с параметрами:

$$m = 0, \sigma = 1.$$

Доказательство теоремы выходит за рамки нашего курса.

§ 11.4. Применения центральной предельной теоремы

1°. Обоснование роли нормального закона. Допустим, что производится измерение какой-либо физической величины. На результат измерения влияет огромное количество случайных факторов, таких как колебание атмосферных условий, сотрясение измерительного прибора, усталость наблюдателя и т.п. Каждый из этих факторов, взятый в отдельности, порождает ничтожную ошибку X_k в измерении данной величины. Результирующая ошибка S_n будет, следовательно, суммой огромного числа малых случайных величин X_k ; и хотя закон распределения каждой из этих величин нам неизвестен, тем не менее можно уверенно заключить, что вся сумма S_n будет иметь закон распределения, близкий к нормальному.

В полном соответствии со сказанным выше при математической обработке результатов измерений исходят из следующего постулата: случайная ошибка измерения подчиняется нормальному закону распределения. Из двух параметров этого закона один, а именно математическое ожидание, равен нулю. Второй параметр – среднее квадратичное отклонение – характеризует в известном смысле точность измерений.

Другой важный пример, иллюстрирующий роль нормального распределения в приложениях теории вероятностей, дает массовое производство, существующее во многих отраслях современной промышленности. В процессе массового производства изготавливаются большие партии однотипных изделий. Все наиболее существенные характеристики выпускаемых изделий должны, естественно, соответствовать определенному стандарту. Однако в действительности наблюдаются отклонения от стандарта, которые порождаются причинами случайного характера (следует учесть, что выпуск изделия связан, как правило, с большим числом операций, некоторые из них не могут быть выполнены

абсолютно точно). Каждая из этих причин сама по себе порождает лишь ничтожную ошибку X_k , но, складываясь, такие ошибки могут давать вполне ощутимые отклонения от стандарта. И здесь, так же как в случае ошибок измерений, имеются все основания считать, что суммарное отклонение от стандарта следует нормальному распределению.

Подобных примеров можно привести очень много из самых различных областей науки и практики. Они объясняют, почему нормальный закон так часто возникает в задачах прикладного характера.

2°. Связь с приближенной формулой Лапласа. Пусть производится n независимых опытов, в каждом из которых с одной и той же вероятностью p наступает событие A . Рассмотрим случайную величину S_n – число наступлений события A в n опытах. Очевидно,

$$S_n = X_1 + X_2 + \dots + X_n,$$

где X_k обозначает число наступлений события A в k -м опыте, $k = 1, 2, \dots, n$. Случайные величины X_k имеют один и тот же закон распределения, так что условия теоремы 11.1 здесь налицо. Но тогда должна быть справедлива формула (11.13), которая в данном случае принимает вид:

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad (11.22)$$

(напомним, что $M(S_n) = np$, а $D(S_n) = npq$, где $q = 1 - p$). Это равенство носит название интегральной предельной теоремы Лапласа. Покажем, что из него следует интегральная приближенная формула Лапласа (формула (2.8)).

Событие

$$a \leq \frac{S_n - np}{\sqrt{npq}} \leq b$$

равнозначно условию

$$np + a\sqrt{npq} \leq v \leq np + b\sqrt{npq}.$$

Положим

$$k_1 = np + a\sqrt{npq}, \quad k_2 = np + b\sqrt{npq},$$

так что

$$a_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad b = \frac{k_2 - np}{\sqrt{npq}}.$$

Теперь левую часть формулы (11.22) можно записать в виде:

$$\lim_{n \rightarrow \infty} P(k_1 \leq v \leq k_2), \quad (11.23)$$

правую же часть, учитывая соотношение

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx = \Phi(b) - \Phi(a)$$

(где $\Phi(x)$ – функция Лапласа, см. (2.9)), можно представить как

$$\Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right). \quad (11.24)$$

Приравнявая выражение, стоящее под знаком предела в (11.23), к выражению (11.24), получаем приближенное равенство

$$P(k_1 \leq v \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right),$$

которое есть не что иное, как интегральная приближенная формула Лапласа.

Рассмотрим ряд примеров задач на нормальный закон распределения.

Пример 11.3. Завод изготавливает шарики для подшипников. Каждый шарик должен иметь один и тот же диаметр d . Однако в силу ряда причин, неизбежных в условиях массового производства, фактический диаметр несколько отличается от d . Обозначим через X разность между фактическим диаметром и числом d . По соображениям, изложенным в п. 1° предыдущего параграфа, можно принять, что величина X подчиняется нормальному закону распределения с математическим ожиданием 0 и некоторым средним квадратичным отклонением σ (характеризующим среднюю точность изготовления шариков).

Каждый шарик, сойдя с конвейера, проходит контроль. Последний состоит в том, что шарик пропускается через отверстия диаметром $d - \varepsilon$ и $d + \varepsilon$ (см. рис. 11.4). Все шарики, которые свободно проходят через большее отверстие, но застревают в меньшем, поступают в готовую продукцию; остальные шарики бракуются. Найти вероятность того, что случайно выбранный с конвейера шарик будет забракован.

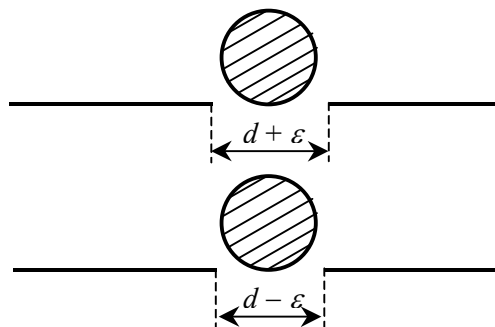


Рис. 11.4

Решение. Условием успешного прохождения шарика через контроль является выполнение неравенств

$$-\varepsilon \leq X \leq \varepsilon.$$

Имеем (см. формулу (5.20)):

$$P(-\varepsilon \leq X \leq \varepsilon) = \Phi\left(\frac{\varepsilon - 0}{\sigma}\right) - \Phi\left(\frac{-\varepsilon - 0}{\sigma}\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

Поэтому вероятность того, что шарик окажется бракованным, равна $1 - 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$.

Пример 11.4. Для определения точности измерительного прибора произведено сравнение его показаний с показаниями контрольного (высокоточного) прибора. Это сравнение показало, что 75% всех ошибок данного прибора не превосходят по абсолютной величине 2 мкм. Считая, что ошибка измерения подчиняется нормальному закону с математическим ожиданием 0, найти среднее квадратичное отклонение σ .

Решение. Обозначим ошибку при измерении на данном приборе через X . По условию X есть случайная величина, подчиненная нормальному закону распределения с плотностью

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

В произведенной серии измерений событие $-2 \leq X \leq 2$ имело частоту 0,75. Считая, что число проделанных измерений достаточно велико, и заменяя частоту вероятностью, запишем:

$$P(-2 \leq X \leq 2) = 0,75.$$

Отсюда $\Phi\left(\frac{2}{\sigma}\right) - \Phi\left(\frac{-2}{\sigma}\right) = 0,75$ или $2\Phi\left(\frac{2}{\sigma}\right) = 0,75$. Решая уравнение $2\Phi(x) = 0,75$ (с помощью

таблицы значений функции $\Phi(x)$), находим: $\frac{2}{\sigma} = 1,15$, откуда $\sigma = 1,74$.

Литература

- Бабайцев В.А., Браилов А.В., Солодовников А.С. Математика в экономике: Теория вероятностей: Курс лекций. Ч. II. М.: Финансовая академия, 1999.
- Боровков А.А. Теория вероятностей. М.: Эдиториал УРСС, 1999.
- Вентцель Е.С. Теория вероятностей. М.: Высшая школа, 1998.
- Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1998.
- Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. М.: Инфра-М, 1997.
- Солодовников А.С. Математика в экономике: Теория вероятностей: Курс лекций. Ч. I. М.: Финансовая академия, 1998.
- Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, 1984. Т. 1, 2.
- Фигурин В.А., Оболонкин В.В. Теория вероятностей и математическая статистика. Минск: Новое знание, 2000.

Таблица значений функций

$$\text{Гаусса } \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ и Лапласа } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
0,00	0,3989	0,0000	0,40	0,3683	0,1554	0,80	0,2897	0,2881
01	3989	0040	41	3683	1591	81	2874	2910
02	3989	0080	42	3653	1628	82	2850	2939
03	3988	0120	43	3637	1664	83	2827	2967
04	3986	0160	44	3621	1700	84	2803	2995
05	3984	0199	45	3605	1736	85	2780	3023
06	3982	0239	46	3589	1772	86	2756	3051
07	3980	0279	47	3572	1808	87	2732	3078
08	3977	0319	48	3555	1844	88	2709	3106
09	3973	0359	49	3538	1879	89	2685	3133
0,10	0,3970	0,0398	0,50	0,3521	0,1915	0,90	0,2661	0,3159
11	3965	0438	51	3503	1950	91	2637	3186
12	3961	0478	52	3485	1985	92	2613	3212
13	3956	0517	53	3467	2019	93	2589	3238
14	3951	0557	54	3448	2054	94	2565	3264
15	3945	0596	55	3429	2088	95	2541	3289
16	3939	0636	56	3410	2123	96	2516	3315
17	3932	0675	57	3391	2157	97	2492	3340
18	3925	0714	58	3372	2190	98	2468	3365
19	3918	0753	59	3352	2224	99	2444	3389
0,20	0,3910	0,0793	0,60	0,3332	0,2257	1,00	0,2420	0,3413
21	3902	0832	61	3312	2291	01	2396	3438
22	3894	0871	62	3292	2324	02	2371	3461
23	3885	0910	63	3271	2357	03	2347	3484
24	3876	0948	64	3251	2389	04	2323	3508
25	3867	0987	65	3230	2422	05	2299	3531
26	3857	1026	66	3209	2454	06	2275	3554
27	3847	1064	67	3187	2486	07	2251	3577
28	3836	1103	68	3166	2517	08	2227	3599
29	3825	1141	69	3144	2549	09	2203	3621

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
0,30	0,3814	0,1179	0,70	0,3123	0,2580	1,10	0,2179	0,3643
31	3802	1217	71	3101	2611	11	2155	3665
32	3790	1255	72	3079	2642	12	2131	3686
33	3778	1293	73	3056	2673	13	2107	3708
34	3765	1331	74	3034	2703	14	2083	3729
35	3752	1368	75	3011	2734	15	2059	3749
36	3739	1406	76	2989	2764	16	2036	3770
37	3726	1443	77	2966	2794	17	2012	3790
38	3712	1480	78	2943	2823	18	1989	3810
39	3697	1517	79	2920	2852	19	1965	3830
1,20	0,1942	0,3849	1,70	0,0940	0,4554	2,40	0,0224	0,4918
21	1919	3869	71	0925	4564	42	0213	4922
22	1895	3888	72	0909	4573	44	0203	4927
23	1872	3907	73	0893	4582	46	0194	4931
24	1849	3925	74	1878	4591	48	0184	4934
25	1826	3944	75	0863	4599	50	0175	4938
26	1804	3962	76	0848	4608	52	0167	4941
27	1781	3980	77	0833	4616	54	0158	4945
28	1758	3997	78	0818	4625	56	0151	4948
29	1736	4015	79	0804	4633	58	0143	4951
1,30	0,1714	0,4032	1,80	0,0790	0,4641	2,60	0,0136	0,4953
31	1691	4049	81	0775	4649	62	0129	4956
32	1669	4066	82	0761	4656	64	0122	4959
33	1647	4082	83	0748	4664	66	0116	4961
34	1626	4099	84	0734	4671	68	0110	4963
35	1604	4115	85	0721	4678	70	0104	4965
36	1582	4131	86	0707	4686	72	0099	4967
37	1561	4147	87	0694	4693	74	0093	4969
38	1539	4162	88	0681	4699	76	0088	4971
39	1518	4177	89	0669	4706	78	0084	4973
1,40	0,1497	0,4192	1,90	0,0656	0,4713	2,80	0,0079	0,4974
41	1476	4207	91	0644	4719	82	0075	4976
42	1456	4222	92	0632	4726	84	0071	4977
43	1435	4236	93	0620	4732	86	0067	4979
44	1415	4251	94	0608	4738	88	0063	4980
45	1394	4265	95	0596	4744	90	0060	4981
46	1374	4279	96	0584	4750	92	0056	4982
47	1354	4292	97	0573	4756	94	0053	4984
48	1334	4306	98	0562	4761	96	0050	4985
49	1315	4319	99	0551	4767	98	0047	4986

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
1,50	0,1295	0,4332	2,00	0,0540	0,4772	3,00	0,00443	0,49865
51	1276	4345	02	0519	4783			
52	1257	4357	04	0498	4793	3,10	00327	49903
53	1238	4370	06	0478	4803	3,20	00238	49931
54	1219	4382	08	0459	4812			
55	1200	4394	10	0440	4821	3,30	00172	49952
56	1182	4406	12	0422	4830	3,40	00123	49966
57	1163	4418	14	0404	4838			
58	1145	4429	16	0387	4846	3,50	00087	49977
59	1127	4441	18	0371	4854			
1,60	0,1109	0,4452	2,20	0,0355	0,4861	3,60	00061	49984
61	1092	4463	22	0339	4868	3,70	00042	49989
62	1074	4474	24	0325	4875	3,80	00029	49993
63	1057	4484	26	0310	4881			
64	1040	4495	28	0297	4887	3,90	00020	49995
65	1023	4505	30	0283	4893	4,00	0001338	499968
66	1006	4515	32	0270	4898			
67	0989	4525	34	0258	4904	4,50	000016	499997
68	0973	4535	36	0246	4909	5,00	0000015	49999997
69	0957	4545	38	0235	4913			

Таблица значений функции $\frac{\lambda^k}{k!} e^{-\lambda}$

[illegible]

Учебное издание

Владимир Алексеевич Бабайцев

Андрей Владимирович Браилов

Александр Самуилович Солодовников

**Математика в экономике.
Теория вероятностей**

Курс лекций

Художественный редактор	В.А. Селин
Техническое редактирование и компьютерная верстка	Л.Б. Галкиной
Корректор	Т.Н. Кузнецова

Лицензия ИД № 01322 от 24.03.2000.

Подписано в печать 24.04.2002.

Формат 60х84/16. Гарнитура Таймс

Усл.п.л. 13,49. Усл.-кр.отт. 13,72. Уч.-изд.л. 6,53.

Тираж 1800 экз.

**Финансовая академия
при Правительстве Российской Федерации**
125468, Москва, Ленинградский проспект, 49

Отпечатано в типографии "Нефтяник"
Москва, Софийская наб., 26