

В.А. ФИГУРИН  
В.В. ОБОЛОНКИН

# ТЕОРИЯ ВЕРОЯТНОСТЕЙ

И

МАТЕМАТИЧЕСКАЯ  
СТАТИСТИКА

УЧЕБНОЕ ПОСОБИЕ



В.А.Фигурин, В.В.Оболонкин

# ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

*Допущено Министерством образования Республики Беларусь  
в качестве учебного пособия для студентов  
естественных специальностей высших учебных заведений*

1656

Минск  
ООО "Новое знание"  
2000

УДК 519.2(075.8)  
ББК 22.17я73  
Ф 49

*Рецензенты:*

каф. теории вероятностей и математической статистики  
Белорусского государственного университета; проф. *Р.М.Жевняк*

**Фигурин В.А., Оболонкин В.В.**

Ф 49 Теория вероятностей и математическая статистика: Учеб. пособие. — Мн.: ООО "Новое знание", 2000. — 208 с.

ISBN 985-6516-09-9.

В основу книги положен курс лекций, читаемый в Международном экологическом университете им. А.Д.Сахарова. Главной задачей, которую ставят перед собой авторы, является рассмотрение теории вероятностей как математического аппарата статистики. Это предопределило включение в книгу таких вопросов, как различные определения вероятности события, случайные величины и способы их описания, числовые характеристики случайных величин, основные законы распределения, системы случайных величин, закон больших чисел и центральные предельные теоремы, основные сведения из теории случайных процессов.

УДК 519.2(075.8)  
ББК 22.17я73

ISBN 985-6516-09-9

© Фигурин В.А., Оболонкин В.В., 2000  
© Оформление. ООО "Новое знание", 2000

## Введение

Для изучения каких-либо явлений производят наблюдения или опыты. Их результаты обычно регистрируют в виде значений некоторых наблюдаемых величин. Однако при повторении наблюдений всегда имеет место разброс результатов. Например, повторение измерений одной и той же величины одним и тем же прибором в совершенно, казалось бы, одинаковых условиях дает результаты, которые хоть немного, но все же отличаются друг от друга. Сколько бы раз ни повторялись измерения, невозможно заранее точно предсказать их результат. В этом смысле говорят, что результаты измерений — это случайная величина.

Появление таких случайных результатов связано либо со случайной природой самого исследуемого явления, либо с различными случайными воздействиями, которые неконтролируемо вносятся в процесс измерений. Но все же и в этом мире случайностей обнаруживаются определенные закономерности. Математический аппарат для изучения этих закономерностей дает теория вероятностей.

Таким образом, теория вероятностей представляет собой область науки, занимающуюся изучением закономерностей, присущих случайным величинам.

Первые работы, в которых появились основные понятия теории вероятностей, представляли собой попытки создания теории азартных игр (игры в кости и карты). Этот период относится к XVI—XVII векам. Становление теории вероятностей как математической науки связано с именем математика Я. Бернулли (1654—1705), доказавшим одну из теорем так называемого закона больших чисел.

Статистика возникла существенно раньше теории вероятностей. Еще в глубокой древности проводились переписи населения и велись земельные кадастры. Эти операции были связаны с наблюдениями и вычислениями. На протяжении веков статистика искала свой математический аппарат и нашла его в теории вероятностей. В результате возник такой раздел математики, как математическая статистика, в котором устанавливаются закономерности случайных явлений на основании обработки статистических данных — результатов наблюдений и измерений.

## Глава 1

# Основные понятия теории вероятностей

### 1.1. Случайные события

В теории вероятностей рассматриваются явления (опыты), которые при одном и том же комплексе начальных условий в зависимости от случайных обстоятельств заканчиваются различными исходами — *событиями*. При этом, говоря об одном и том же комплексе начальных условий, подразумевают то обстоятельство, что остаются без изменений основные существенные обстоятельства опыта.

В дальнейшем вместо того, чтобы говорить “комплекс начальных условий создан”, будем говорить кратко: “произведено испытание”. Таким образом, событие будем рассматривать как результат *испытания*. Термин “испытание” является общепринятым в теории вероятностей и заменяет собой термины *наблюдение, опыт, измерение*.

Событие, которое в результате проведенного испытания обязательно произойдет, называется *достоверным*. Например, если в урне имеются только белые шары, то при извлечении шара событие, состоящее в появлении белого шара, является достоверным. Достоверное событие будем обозначать буквой  $U$ .

Событие, которое в результате проведенного испытания не может наступить, называется *невозможным*. Если в урне имеются только белые шары, то при извлечении из урны шара событие, состоящее в том, что этот шар окажется черным, является невозможным событием. Невозможное событие будем обозначать буквой  $V$ .

*Случайное событие* — это событие, которое в результате проведенного испытания может либо произойти, либо не произойти. Случайные события будем обозначать буквами  $A, B, C, \dots$

Через  $\bar{A}$  будем обозначать событие, состоящее в том, что в результате проведенного испытания событие  $A$  не произойдет. События

$A$  и  $\bar{A}$  называются *противоположными* событиями. Очевидно, что свойство противоположности у них взаимное, то есть

$$\overline{\bar{A}} = A.$$

События называются *несовместными*, если появление одного из них исключает появление других в одном и том же испытании, то есть такие события не могут произойти одновременно. Если же появление одного из событий не исключает возможности появления других событий в данном испытании, то такие события называются *совместными*. Так, при извлечении из колоды одной игральной карты события появления валета и дамы являются несовместными, но события появления валета и пики являются совместными.

Несколько несовместных событий образуют *полную группу* (или *полную систему*) событий, если в результате испытания появится хотя бы одно из них. Другими словами, появление хотя бы одного из событий полной группы есть достоверное событие. Если стрелок произвел выстрел по мишени, то обязательно произойдет одно из двух событий: попадание или промах. Эти два несовместных события образуют простейшую полную группу событий.

События, составляющие полную группу событий, называются *равновозможными*, если есть основания считать, что в данном испытании ни одно из них не является более возможным, чем другое. Так, появление того или иного числа очков на брошенной игральной кости — равновозможные события. Действительно, предполагается, что игральная кость изготовлена из однородного материала, имеет форму правильного многогранника и наличие очков не оказывает влияния на выпадение любой грани.

При проведении испытания достоверное событие обязательно произойдет, а невозможное обязательно не произойдет. Среди событий, которые при испытании могут произойти, а могут и не произойти, на появление одних можно рассчитывать с большим основанием, на появление других — с меньшим основанием. Если, например, в урне белых шаров больше, чем черных (шары могут отличаться только цветом, и перед извлечением шара урну встряхивают, чтобы шары хорошо перемешались), то надеяться на появление белого шара при выпимании шаров из урны больше оснований, чем на появление черного шара.

Величина, определяющая, насколько значительны объективные основания рассчитывать на появление события, характеризуется как *вероятность события*. Необходимо подчеркнуть, что вероятность есть объективная величина, существующая независимо от испытания и обусловленная всем комплексом начальных условий, которые способствуют появлению события.

Объяснение, которое было дано понятию вероятности, не является математическим определением, так как оно не определяет этого понятия количественно. Существует несколько различных определений вероятности случайного события: классическое, статистическое, аксиоматическое и др. Каждому из них присущи недостатки, связанные либо с принятой концепцией, либо с возможными приложениями. Как ни странно, исчерпывающей математической формулировки понятия вероятности еще не существует.

## 1.2. Классическое определение вероятности

Классическое определение вероятности события основано на рассмотрении полной группы несовместных и равновозможных событий, которые называются случаями. Рассмотрим такую полную группу событий, состоящую из  $n$  случаев  $A_1, A_2, \dots, A_m, \dots, A_n$ . Пусть из общего числа случаев событию  $A$  благоприятствует  $m$  случаев.

Вероятность  $P$  события  $A$  в классическом определении равна отношению числа благоприятствующих ему, случаев к общему числу равновозможных случаев, то есть

$$P(A) = \frac{m}{n}, \quad m \leq n.$$

Согласно тому же определению вероятности события, получаем

$$P(A_1) = P(A_2) = \dots = P(A_n) := \frac{1}{n},$$

то есть равновозможные события имеют одинаковые вероятности.

Как следует из определения, вероятность события — всегда безразмерная величина.

Пусть в урне имеется пять белых и три черных шара. Тогда полная группа событий состоит из восьми равновозможных случаев,

а событие, состоящее в появлении белого шара, подразделяется на пять таких случаев. В связи с этим, вероятность появления белого шара

$$P(A) = \frac{5}{8}.$$

Аналогично вероятность появления черного шара

$$P(B) = \frac{3}{8},$$

так как здесь число благоприятствующих случаев равно трем.

Из классического определения вероятности события вытекают следующие ее свойства:

1. Вероятность достоверного события равна единице. Действительно, если событие  $U$  достоверно, то каждый исход испытания благоприятствует событию. В этом случае  $m = n$ , следовательно,

$$P(U) = \frac{m}{n} = \frac{n}{n} = 1.$$

2. Вероятность невозможного события равна нулю. В самом деле, если событие невозможно, то ни один из исходов испытания не благоприятствует событию. В этом случае  $m = 0$ , следовательно,

$$P(V) = \frac{m}{n} = \frac{0}{n} = 0.$$

3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей. Действительно, случайному событию благоприятствует лишь часть из общего числа исходов испытания. В этом случае  $0 < m < n$ ,  $0 < \frac{m}{n} < 1$ , следовательно,

$$0 < P(A) < 1.$$

Итак, вероятность любого события удовлетворяет неравенству

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

Классическое определение вероятности можно подвергнуть критике на том основании, что термин "равновозможные события" на самом деле означает "равновероятные события", то есть в проведенных рассуждениях содержится порочный круг.

### 1.3. Статистическое определение вероятности

Необходимый для классического определения вероятности анализ, основанный на рассмотрении полной группы равновозможных событий и выделении тех из них, которые благоприятны для рассматриваемого события, удается провести далеко не всегда. Например, событие, состоящее в том, что определенный атом радия распадается за время, не превосходящее  $t$ , не поддается исследованию в такой схеме. Поэтому определить вероятность этого события классическим методом нельзя.

Существует другой способ нахождения вероятности случайного события — оценка при помощи опыта. Допустим, что комплекс начальных условий, при котором может происходить рассматриваемое событие, многократно в точности повторяется сам или может быть многократно воспроизведен. Для краткости в таком случае условимся говорить, что производятся повторные испытания. Если при проведении  $N$  испытаний событие осуществилось  $m$  раз, то говорят, что *относительная частота*  $\nu$  события  $A$  равна

$$\nu(A) = \frac{m}{N}.$$

Из-за случайных причин относительная частота  $\nu$  (ее называют также *частотью*) при разном количестве испытаний и в разных сериях из  $N$  испытаний будет различной. Однако повседневная практика и специальные длительные эксперименты показывают, что с увеличением числа испытаний относительная частота обнаруживает свойство устойчивости. Это свойство состоит в том, что в различных испытаниях относительная частота изменяется мало (тем меньше, чем больше произведено испытаний), то есть колеблется около некоторого постоянного числа. Поэтому относительную частоту можно считать результатом измерения вероятности события:

$$P(A) \approx \nu(A) = \frac{m}{N}.$$

Например, в опытах по бросанию монеты относительная частота появления герба при 12 000 бросаний оказалась равной 0,5016, а в опыте с 24 000 бросаний — 0,5005.

В статистическом определении вероятностью события называется величина, около которой группируются относительные частоты этого события. Можно также сказать, что вероятностью события является предел, к которому стремится относительная частота при неограниченном увеличении числа испытаний:

$$P(A) = \lim_{N \rightarrow \infty} \frac{m}{N}.$$

Легко убедиться, что свойства вероятности, вытекающие из классического определения, сохраняются и при статистическом определении вероятности.

Если вероятность некоторого события близка к нулю, то в соответствии со сказанным следует, что при единичном испытании в подавляющем большинстве случаев такое событие не наступит. Естественно, возникает вопрос: насколько малой должна быть вероятность, чтобы можно было считать невозможным наступление некоторого события в единичном испытании? Ответ на него не однозначен и зависит от тех потерь, которые будут иметь место, если интересующее нас событие все-таки произойдет. Достаточно малую вероятность, при которой наступление события можно считать практически невозможным, называют *уровнем значимости*. Отметим, что на практике уровень значимости обычно принимают равным 0,01 (однопроцентный уровень значимости) или 0,05 (пятипроцентный уровень).

Событие  $A$  называется практически достоверным, если  $\bar{A}$  — практически невозможное событие.

## 1.4. Аксиоматическое построение теории вероятностей

Наиболее распространенная в настоящее время логическая схема построения основ теории вероятностей разработана А.Н. Колмогоровым в 1933 г. Основные черты этой схемы следующие.

При изучении какой-либо реальной задачи методами теории вероятностей прежде всего выделяется множество  $\Omega$ , называемое вероятностным пространством. Элементы  $\omega$  этого множества составляют всю совокупность возможных исходов наблюдения — элементарных событий. Всякое событие описывается множеством благоприятствующих ему элементарных событий и поэтому рассматривается как не-

которое множество элементарных событий. Например, если в опыте с игральной костью в качестве возможных исходов рассматривать случаи выпадения каждой из шести граней кубика, то соответствующим вероятностным пространством будет множество  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . Случайному событию  $A_1$  может соответствовать выпадение грани 2, случайному событию  $A_2$  — выпадение одной из граней 1, 2 или 3.

Поскольку каждое наблюдение должно иметь по крайней мере один исход, все вероятностное пространство  $\Omega$  соответствует достоверному событию, а пустое множество, скажем  $\emptyset$ , — невозможному событию. С некоторыми из событий  $A_i$  связываются определенные числа  $P(A_i)$ , называемые их вероятностями и удовлетворяющие следующим условиям (аксиомам):

- 1)  $0 \leq P(A_i) \leq 1$ ;

- 2)  $P(\Omega) = 1$ , то есть вероятность достоверного события равна единице;

- 3) вероятность события  $A$ , заключающегося в том, что наступит или событие  $A_1$ , или событие  $A_2, \dots$ , или событие  $A_n$ , где  $A_1, A_2, \dots, A_n$  — попарно несовместные события, составляет

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Это условие должно выполняться и для бесконечных последовательностей попарно несовместных событий. Вся теория вероятностей строится на этих трех аксиомах.

Подчеркнем, что исходные аксиомы постулируются и попытка доказать их лишена смысла. Единственным возможным критерием справедливости этих аксиом является степень, с которой теория, построенная на их основе, отражает реальность. Этот критерий, кстати, справедлив и в отношении любой другой естественнонаучной теории.

Приведем формальное определение вероятности для испытаний с бесконечным числом исходов. В подобных случаях пространство элементарных исходов может быть некоторой областью  $G$ , а под событием  $A$  можно понимать исходы, входящие в область  $g$  (рис. 1.1).

Пусть на область  $G$  наугад бросается "точка". Какова вероятность того, что она попадет в область  $g$ , являющуюся частью области  $G$ ?

Хотя каждое из множеств  $G$  и  $g$  содержит бесконечное множество точек, естественно принять, что "вместимость" множества  $G$  больше и притом во столько раз, во сколько площадь  $S_G$  области  $G$  превышает

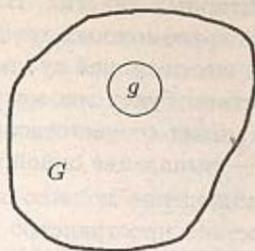


Рис. 1.1. К определению геометрической вероятности

площадь  $S_g$  области  $g$ . Приняв равновозможность всех возможных вариантов, естественно считать, что искомая вероятность равна

$$P(A) = \frac{S_g}{S_G}$$

В общем случае множества  $G$  и  $g$  могут иметь другую размерность, но приведенная формула сохраняет свой смысл с той лишь разницей, что множества в общем случае оцениваются своей мерой (длиной, площадью, объемом). Таким образом, в общем случае

$$P(A) = \frac{\text{мера } g}{\text{мера } G}$$

Вероятности, определяемые по такой схеме, часто называют геометрическими вероятностями.

*Рассмотрим пример.* На плоскости начерчены две концентрические окружности, радиусы которых 5 и 10 см соответственно. Требуется найти вероятность того, что точка, брошенная наугад в большой круг, попадет в кольцо, образованное концентрическими окружностями. Предполагается, что вероятность попадания этой точки в плоскую фигуру пропорциональна площади этой фигуры и не зависит от ее расположения относительно большого круга. Вычислим площадь кольца (фигуры  $g$ ):

$$S_g = \pi(10^2 - 5^2) = 75\pi.$$

Площадь большого круга (фигуры  $G$ ) составляет

$$S_G = \pi 10^2 = 100\pi.$$

Отсюда искомая вероятность равна

$$P = \frac{75\pi}{100\pi} = 0,75.$$

## 1.5. Сумма и произведение событий

*Суммой событий  $A$  и  $B$*  называется такое событие  $C$ , которое состоит в наступлении или события  $A$ , или события  $B$ , или в наступлении обоих событий  $A$  и  $B$ . Для обозначения суммы событий  $A$  и  $B$  применяется запись  $C = A + B$ .

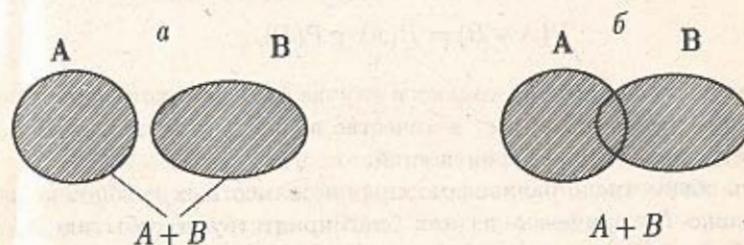


Рис. 1.2. Геометрическая интерпретация суммы двух событий:  $a$  — события  $A$  и  $B$  несовместны;  $б$  — события  $A$  и  $B$  совместны

Пусть событие  $A$  — попадание снаряда в область  $A$ , а событие  $B$  — попадание снаряда в область  $B$ . Тогда  $A + B$  — событие, состоящее в попадании снаряда в заштрихованную область (рис. 1.2).

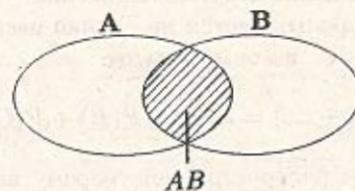


Рис. 1.3. Геометрическая интерпретация произведения двух событий

*Произведением событий  $A$  и  $B$*  называется такое событие  $C$ , которое состоит в осуществлении и события  $A$ , и события  $B$ . Для обозначения произведения событий применяется запись  $C = A \cdot B = AB$ .

Пусть  $A$  и  $B$  — события, рассмотренные выше. Тогда  $AB$  — это попадание снаряда в общую часть областей  $A$  и  $B$  (рис. 1.3). Аналогично определяются сумма и произведения большего числа событий.

## 1.6. Теорема сложения вероятностей

Пусть  $A$  и  $B$  представляют собой несовместные исходы одного и того же испытания. Для таких событий имеет место следующая теорема.

**Теорема сложения вероятностей:** вероятность суммы двух несовместных событий равна сумме вероятностей этих событий, то есть

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

Эта теорема доказывается только в случае классического определения вероятности и выступает в качестве аксиомы при аксиоматическом построении теории вероятностей.

Пусть общее число равновероятных несовместных исходов испытания равно  $N$ , причем  $n$  из них благоприятствуют событию  $A$ , а  $m$  — событию  $B$ , так что  $P(A) = n/N$  и  $P(B) = m/N$ . Так как события  $A$  и  $B$  совместно не осуществляются, то их сумме  $A+B$  благоприятствуют все  $n+m$  событий. Поэтому, согласно классическому определению вероятности, получим

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

Укажем на два обобщения теоремы сложения.

1. Эта теорема распространяется на случай нескольких событий. Если события  $A, B, C, \dots$  несовместны, то

$$P(A + B + C + \dots) = P(A) + P(B) + P(C) + \dots$$

2. Второе обобщение распространяет теорему на случай совместных событий. Если события  $A$  и  $B$  совместны, то

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

Для доказательства кроме событий  $A$  и  $B$  рассмотрим также события  $AB, \overline{AB}, \overline{A}\overline{B}, A\overline{B}$ . Первое из этих событий состоит в том, что

произошли оба события  $A$  и  $B$ , второе — в том, что событие  $A$  произошло, а событие  $B$  не произошло, и т.д. Очевидно, что отмеченные события составляют полную группу событий, так как они несовместны и при проведении испытания одно из них обязательно произойдет.

Теперь предположим, что удалось наблюдать полную систему  $N$  равновероятных событий, из которых событию  $AB$  благоприятствуют  $n_1$  событий, событию  $\overline{AB}$  —  $n_2$  событий, событию  $A\overline{B}$  —  $n_3$  событий и событию  $\overline{A}\overline{B}$  —  $n_4$  событий. Поскольку эти события составляют полную систему событий, то

$$n_1 + n_2 + n_3 + n_4 = N.$$

Соответствующие вероятности рассмотренных событий равны:

$$P(AB) = \frac{n_1}{N}, \quad P(\overline{AB}) = \frac{n_2}{N},$$

$$P(A\overline{B}) = \frac{n_3}{N}, \quad P(\overline{A}\overline{B}) = \frac{n_4}{N}.$$

Теперь заметим, что в данной полной группе равновероятных событий событию  $A$  благоприятны  $n_1 + n_3$  равновероятных событий (так как событие  $A$  произойдет, если произойдут события  $AB$  или  $A\overline{B}$ ), а событию  $B$  —  $n_1 + n_2$  равновероятных событий. Следовательно,

$$P(A) = \frac{n_1 + n_3}{N}, \quad P(B) = \frac{n_1 + n_2}{N}.$$

Далее рассмотрим событие  $A+B$ . Ему благоприятствуют  $n_1 + n_3$  равновероятных событий и, следовательно,

$$P(A+B) = \frac{n_1 + n_2 + n_3}{N}$$

или

$$P(A+B) = \frac{n_1 + n_3}{N} + \frac{n_1 + n_2}{N} - \frac{n_1}{N} = P(A) + P(B) - P(AB).$$

Следствия теоремы сложения:

1) сумма вероятностей  $n$  несовместных событий, составляющих полную группу, равна единице, то есть

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1;$$

2) сумма вероятностей противоположных событий равна единице, то есть

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

## 1.7. Теорема умножения вероятностей

Несколько событий могут находиться между собой в таких взаимоотношениях, что вероятность наступления какого-либо из событий зависит от того, произошло или не произошло другое событие. В связи с этим введем следующие определения.

Событие  $A$  называется статистически зависимым от событий  $B_1, B_2, \dots, B_k$ , если вероятность события  $A$  зависит от того, осуществились или не осуществились события  $B_1, B_2, \dots, B_k$ . Если же вероятность события  $A$  не связана с осуществлением событий  $B_1, B_2, \dots, B_k$ , то событие  $A$  называется статистически независимым от событий  $B_1, B_2, \dots, B_k$ .

Вероятность события  $A$ , вычисленная при условии, что произошли события  $B_1, B_2, \dots, B_k$ , называется *условной* и обозначается через  $P(A|B_1, B_2, \dots, B_k)$ . Если же при вычислении вероятности события  $A$  события  $B_1, B_2, \dots, B_k$  не принимаются во внимание, то вероятность  $P(A)$  называется *безусловной*.

Согласно определению, событие  $A$  статистически зависит от события  $B$ , если  $P(A|B) \neq P(A)$ . Если же  $P(A|B) = P(A)$ , то событие  $A$  не зависит от события  $B$ .

**Теорема умножения вероятностей:** вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что произошло первое:

$$P(AB) = P(A)P(B|A) = P(B)P(A|B). \quad (1.1)$$

Доказательство теоремы проведем для схемы случаев (классического определения вероятности).

Пусть общее число равновозможных исходов испытания равно  $N$ , причем событию  $A$  благоприятствуют  $k$  исходов, событию  $B$  —  $l$  исходов, событию  $AB$  —  $m$  исходов (рис. 1.4).

Вычислим условную вероятность  $P(B|A)$ . Если событие  $A$  произошло, то общее число равновозможных исходов равно  $k$ . Из них событию  $B$  благоприятствуют  $m$  исходов. Следовательно,

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

Но

$$\frac{m}{k} = \frac{m/N}{k/N},$$

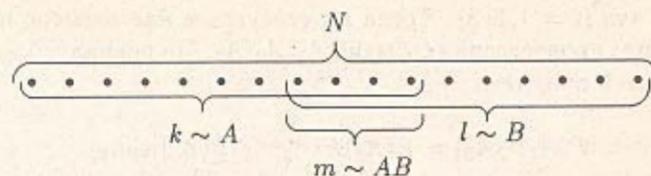


Рис. 1.4. К доказательству теоремы умножения вероятностей

то есть

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

что и доказывает первую часть формулы (1.1). Аналогичным образом можно показать, что

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

Теорему умножения вероятностей легко распространить на случай, когда событий больше двух. Например, для трех событий  $A, B$  и  $C$  формула умножения вероятностей имеет вид

$$P(ABC) = P((AB)C) = P(AB)P(C|AB) = P(A)P(B|A)P(C|AB).$$

Следствия теоремы умножения вероятностей:

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

2. Если  $A$  и  $B$  независимы, то формулу (1.1) можно записать в виде

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

то есть вероятность произведения двух независимых событий равна произведению вероятностей этих событий. Если события  $A_1, A_2, \dots, A_n$  независимы, то

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2) \dots P(A_n).$$

*Рассмотрим пример.* Пусть из полной колоды карт (52 карты) вынимают наугад три карты (без возврата). Вычислим вероятность того, что среди них не будет ни одного туза.

Обозначим через  $A_i$  событие, состоящее в том, что  $i$ -я вынутая карта не туз ( $i = 1, 2, 3$ ). Тогда интересующее нас событие представится в виде произведения событий  $A_1 \cdot A_2 \cdot A_3$ . По правилу умножения вероятностей получим

$$P(A_1 A_2 A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2).$$

Так как всех карт 52, а тузов из них 4, то при вынимании первой карты имеем для события  $A_1$  классическую модель с  $n = 52$ ,  $m = 48$ , поэтому  $P(A_1) = 48/52$ . При условии, что вынутая карта не туз, для события  $A_2|A_1$  снова имеем классическую модель с  $n = 51$ ,  $m = 47$ , поэтому  $P(A_2|A_1) = 47/51$ . Аналогично  $P(A_3|A_1 A_2) = 46/50$  и, следовательно,

$$P(A_1 A_2 A_3) = \frac{48}{52} \cdot \frac{47}{51} \cdot \frac{46}{50} = \frac{4324}{5525} = 0,783.$$

## 1.8. Формула полной вероятности

Во многих реальных ситуациях то или иное событие  $A$  может появиться лишь как случайное следствие одного из несовместных событий  $H_1, H_2, \dots, H_n$ , которые входят в некоторую полную группу событий (но могут ее и не составлять) и называются *гипотезами*. При этом термин "случайное следствие" означает, что каждая из гипотез  $H_1, H_2, \dots, H_n$  может повлечь за собой не только исход  $A$ , но и какие-то другие исходы. Предполагается, что вероятности гипотез  $P(H_1), P(H_2), \dots, P(H_n)$  и условные вероятности события  $A$  при каждой из гипотез, то есть вероятности  $P(A|H_i)$ , известны.

Ставится задача определить безусловную вероятность события  $A$ , то есть вероятность  $P(A)$ , при вычислении которой принимаются во внимание все случаи появления события  $A$ . Отметим, что событие  $A$  появляется тогда и только тогда, когда осуществляется одно из несовместных сложных событий  $AH_1, AH_2, \dots, AH_n$ . Заметим, что события  $AH_1, AH_2, \dots, AH_n$  несовместны так же, как и сами гипотезы  $H_1, H_2, \dots, H_n$ . Поэтому, согласно теореме сложения вероятностей, можно записать

$$P(A) = P(AH_1) + P(AH_2) + \dots + P(AH_n).$$

Для определения вероятностей событий  $AH_i$  можно применить теорему умножения:

$$P(AH_i) = P(H_i)P(A|H_i).$$

В результате получим формулу, которая носит название *формулы полной вероятности*:

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

*Рассмотрим пример.* Пусть однотипная продукция трех рабочих упакована в три одинаковых на вид ящика. Из одного (взятого произвольно) ящика наугад вынимается одна деталь. Чему равна вероятность  $P(A)$  того, что деталь окажется бракованной, если есть основания считать, что в первом ящике из 100 деталей негодных четыре, во втором из 120 негодных шесть, в третьем из 80 негодных восемь?

Рассмотрим следующие гипотезы:  $H_1$  — деталь взята из первого ящика,  $H_2$  — деталь взята из второго ящика,  $H_3$  — деталь взята из третьего ящика. В условии задачи сказано, что все гипотезы равновероятны, то есть

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

Из условия следует, что условные вероятности события  $A$  при каждой из гипотез составляют

$$P(A|H_1) = \frac{4}{100}; \quad P(A|H_2) = \frac{6}{120};$$

$$P(A|H_3) = \frac{8}{80} \quad !$$

Подставляя полученные значения  $P(H_i)$  и  $P(A|H_i)$  в формулу полной вероятности, находим

$$P(A) = \frac{1}{3} \cdot \frac{4}{100} + \frac{1}{3} \cdot \frac{6}{120} + \frac{1}{3} \cdot \frac{8}{80} = \frac{19}{300} = 0,063.$$

## 1.9. Формула Байеса

Ситуация, приводящая к формуле Байеса, может быть описана следующим образом.

Пусть интересующее нас событие  $A$  может появиться лишь как случайное следствие одной из несовместных гипотез  $H_1, H_2, \dots, H_n$  и условные вероятности события  $A$  при этих гипотезах  $P(A|H_1), P(A|H_2), \dots, P(A|H_n)$  известны. Пусть произведено испытание, и результатом его явилось событие  $A$ . Спрашивается: с какой из гипотез следует связывать появление события  $A$ ?

Отметим прежде всего, что поскольку изложенная ситуация является вероятностной, то ответ на поставленной вопрос будет иметь вероятностный характер. Для решения задачи необходимо вычислить условную вероятность каждой гипотезы при условии, что произошло событие  $A$ , и той из гипотез, которая будет иметь наибольшую вероятность  $P(H_i|A)$ , целесообразно отдать предпочтение.

Рассмотрим порядок вычисления условных вероятностей  $P(H_1|A), P(H_2|A), \dots, P(H_n|A)$ . На основании теоремы умножения для вероятности совместного появления события  $A$  и гипотезы  $H_i$  можно записать:

$$P(AH_i) = P(A)P(H_i|A) = P(H_i)P(A|H_i).$$

Это равенство дает

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)}.$$

Но по формуле полной вероятности имеем

$$\begin{aligned} P(A) &= P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_n)P(A|H_n) = \\ &= \sum_{k=1}^n P(H_k)P(A|H_k), \end{aligned}$$

следовательно,

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{k=1}^n P(H_k)P(A|H_k)}.$$

Это и есть формула Байеса, или формула для вероятности гипотезы после испытания.

*Рассмотрим пример.* Обнаружен факт сброса в водоем неочищенных стоков. Пусть известно, что потенциальными источниками загрязнения являются два предприятия, причем исходно вероятность того, что сброс произведен первым предприятием, оценивается в 90, а вторым — в 10%. Известно, что в 15% стока первого предприятия и в 92% второго ртуть превышает предельно допустимую концентрацию (ПДК). Определить, какому предприятию может принадлежать обнаруженный сброс, если взятая проба показывает превышение ПДК по ртути. Пусть  $A$  — событие, состоящее в том, что превышен ПДК по ртути;  $H_1$  и  $H_2$  — события (гипотезы), состоящие в том, что сброс произведен соответственно первым и вторым предприятиями. Тогда

$$P(A|H_1) = 0,15, \quad P(A|H_2) = 0,92;$$

$$P(H_1) = 0,9, \quad P(H_2) = 0,1.$$

Поэтому

$$P(H_1|A) = \frac{0,15 \cdot 0,9}{0,15 \cdot 0,9 + 0,92 \cdot 0,1} = 0,595;$$

$$P(H_2|A) = \frac{0,92 \cdot 0,1}{0,15 \cdot 0,9 + 0,92 \cdot 0,1} = 0,405.$$

Таким образом, предпочтение следует отдать гипотезе  $H_1$  (сбросы принадлежат первому предприятию).

## Случайные величины

## 2.1. Понятие случайной величины

Случайной называется величина, которая в результате испытания принимает то или иное значение, которое зависит от случайных обстоятельств данного испытания и заранее не может быть предсказано. Между случайной величиной и случайным событием существует тесная связь. Так, например, любому событию  $A$  можно поставить во взаимнооднозначное соответствие случайную величину, которая при появлении события  $A$  равна единице, а при появлении этого события — нулю. Однако во многих случаях предпочтительнее иметь дело со случайными величинами.

Примерами случайных величин могут служить число очков, появляющихся при бросании игральной кости; число выстрелов до первого попадания в цель; время безотказной работы какого-либо устройства. В первых двух примерах случайная величина является *дискретной*, а в последнем — *непрерывной*. Множество значений, которое в результате испытания может принимать случайная величина, в первом примере конечно, а в двух последних бесконечно. Бесконечное множество может быть счетным, то есть его элементы можно, как во втором примере, пронумеровать, и несчетным (третий пример).

## 2.2. Описание случайных величин

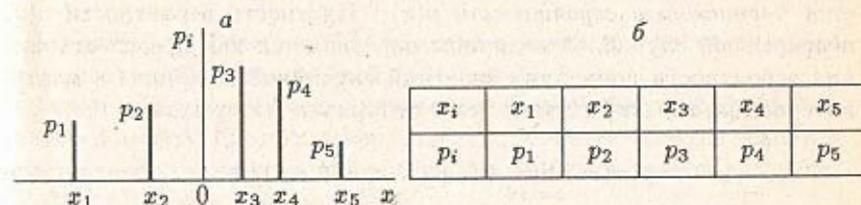
Дискретная случайная величина принимает конечное или счетное множество возможных значений. Для ее описания нужно прежде всего указать эти возможные значения. Однако характер случайной величины таким перечислением возможных значений не определяется. Необходимо еще знать, насколько часто будут осуществляться одни значения случайной величины и насколько редко — другие или, что то же самое, насколько вероятно появление тех или иных значений

случайной величины. Так, во втором примере для хорошего стрелка наиболее вероятное число выстрелов до первого попадания равно 1 или 2; для плохого эти цифры будут большими. Соотношение, в той или иной форме устанавливающее зависимость между возможными значениями случайной величины и их вероятностями, называется *законом распределения дискретной случайной величины*, или короче — *распределением*.

Случайные величины будем обозначать большими буквами латинского алфавита, а их возможные значения — малыми (например, случайная величина  $X$ , а ее значения —  $x_i$ ). Вероятность того, что в результате испытания случайная величина  $X$  примет значение  $x_i$ , будем обозначать через  $p_i$ :

$$p_i = P(X = x_i).$$

Тогда закон распределения дискретной случайной величины можно задать в виде графика, таблицы или аналитического выражения, как это показано на рис. 2.1:



$$p_i = f(x_i)$$

Рис. 2.1. Три способа задания закона распределения дискретной случайной величины: а — график, б — таблица и в — формула

Очевидно, что вероятности  $p_i$  должны удовлетворять так называемому условию нормировки:

$$p_1 + p_2 + \dots + p_n = \sum_{i=1}^n p_i = 1$$

Данное свойство вытекает из того, что если случайная величина принимает конечное число различных значений  $x_1, x_2, \dots, x_n$ , то эле-

независимые события  $X = x_1, X = x_2, \dots, X = x_n$  образуют полную группу попарно несовместных случайных событий, и поэтому сумма их вероятностей равна единице.

Если множество возможных значений случайной величины бесконечно, то второе условие заменяется следующим: бесконечный ряд  $p_i$  должен быть сходящимся, его сумма должна быть равна 1:

$$\sum_{i=1}^{\infty} p_i = 1.$$

Введенное понятие распределения утрачивает смысл для непрерывных случайных величин. Пусть, например, наугад разрезается нить некоторой длины  $L$ . Поскольку точек, в которых может быть сделан разрез, бесконечно много, то вероятность совпадения разреза с некоторой конкретной точкой оказывается исчезающе малой (равной нулю).

Закон распределения непрерывной случайной величины  $X$  задается *плотностью вероятности*  $p(x)$ . Плотность вероятности  $p(x)$  непрерывной случайной величины определяется как предел отношения вероятности попадания значений случайной величины в малый интервал  $[x, x + \Delta x]$  к длине этого интервала  $\Delta x$  при  $\Delta x \rightarrow 0$ :

$$p(x) = \lim_{\Delta x \rightarrow 0} P(x \leq X < x + \Delta x) / \Delta x.$$

Отсюда следует, что  $P(x \leq X < x + \Delta x) = p(x)\Delta x + o(\Delta x)$ , где  $o(\Delta x)$  — величина более высокого порядка малости по сравнению с  $\Delta x$ . Величина  $p(x)\Delta x$  называется элементом вероятности.

Плотность вероятности  $p(x)$  обладает следующими свойствами:

- 1) неотрицательностью (что связано с неотрицательностью вероятностей), то есть  $p(x) \geq 0$ ;
- 2) *нормировкой*, то есть

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

Это свойство отражает достоверность события ( $-\infty \leq X \leq \infty$ );

3) вероятность попадания случайной величины в полузамкнутый интервал  $[a, b)$  равна интегралу от плотности вероятности в этих пределах, то есть

$$P(a \leq X < b) = \int_a^b p(x) dx. \quad (2.1)$$

Поскольку вероятность — величина безразмерная, то, как видно из последнего выражения, плотность вероятности имеет размерность, обратную размерности случайной величины.

Пусть плотность вероятности случайной величины  $X$  имеет вид, приведенный на рис. 2.2. Найдем вероятность попадания случайной величины в интервал  $[a, b)$ .

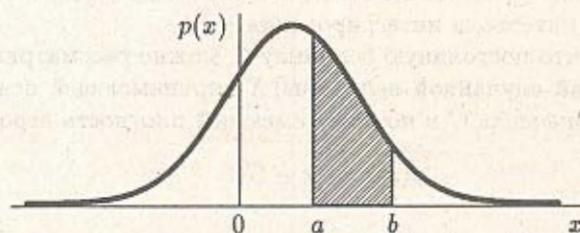


Рис. 2.2. К объяснению геометрического смысла вероятности попадания случайной величины в интервал  $[a, b)$

Как видно из формулы 2.1, искомая вероятность равна площади заштрихованной криволинейной трапеции.

Дискретные случайные величины, как и непрерывные, можно описывать с помощью плотности вероятности. Поэтому часто не делается существенного различия в использовании терминов “закон распределения” и “плотность вероятности”.

Используя определение  $\delta$ -функции

$$\delta(x - x_0) = \begin{cases} \infty, & \text{если } x = x_0, \\ 0, & \text{если } x \neq x_0, \end{cases}$$

$$\int_{x_0-\epsilon}^{x_0+\epsilon} \delta(x - x_0) dx = 1 \text{ при любом } \epsilon > 0$$

и ее свойство

$$\int_a^b f(x) \delta(x - x_0) dx = f(x_0),$$

закон распределения дискретной случайной величины  $p_i = P(X = x_i)$  можно записать в виде плотности вероятности

$$p(x) = \sum_{i=1}^n p_i \delta(x - x_i).$$

При этом для плотности вероятности дискретной случайной величины будут выполняться первые два из приведенных выше свойств плотности вероятности. Однако последнее свойство следует конкретизировать, так как в граничной точке может быть сосредоточена отличная от нуля вероятность, если в этой точке аргумент  $\delta$ -функции равен нулю. Поэтому применительно к дискретной случайной величине в интеграле (2.1) левый конец  $a$  нужно включить, а правый  $b$  исключить из интервала интегрирования.

Очевидно, что постоянную величину  $C$  можно рассматривать как частный случай случайной величины  $X$ , принимающей всякий раз одно и то же значение  $C$  и поэтому имеющей плотность вероятности

$$p(x) = \delta(x - C).$$

### 2.3. Функция распределения

Универсальной характеристикой, одинаково пригодной как для дискретных, так и непрерывных случайных величин, является функция распределения случайной величины. Функцией распределения случайной величины  $X$  называется функция  $F(x)$ , равная вероятности того, что случайная величина  $X$  примет значение меньшее, чем  $x$  (для всех  $x$  на числовой оси):

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

Для дискретной случайной величины  $X$  функция распределения равна сумме вероятностей тех ее значений  $x_k$ , которые меньше  $x$ :

$$F(x) = \sum_{x_k < x} p_k.$$

Рассмотрим построение функции распределения для дискретной случайной величины.

Пусть вероятности отдельных значений случайной величины  $X$  заданы графиком (рис. 2.3).

Пока аргумент функции  $F(x)$  остается меньшим или равным  $x_1$ , функция распределения  $F(x)$ , очевидно, равна нулю, так как нет ни одного значения  $x$ , которое было бы меньше  $x_1$ . В точке  $(x_1+0)$  функция  $F(x)$  скачком принимает значение  $p_1$  и остается постоянной в интервале от  $x_1$  до  $x_2$ . В точке  $(x_2+0)$  функция  $F(x)$  скачком возрастает

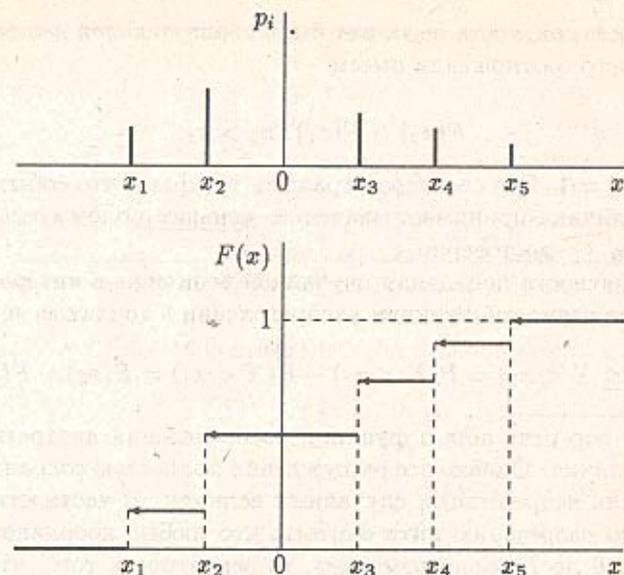


Рис. 2.3. К построению функции распределения дискретной случайной величины

возрастает до величины  $(p_1+p_2)$ , так как событие “принять случайной величине значение, меньшее  $(x_2+0)$ ” распадается на два несовместных события: “принять значение  $x_1$ ” с вероятностью  $p_1$  и “принять значение  $x_2$ ” с вероятностью  $p_2$ . Аналогичным образом производится построение функции распределения для остальных  $x$ . Теперь становится понятным и другое, иногда встречающееся в литературе, название функции распределения — накопленная (или кумулятивная) вероятность.

Функция распределения обладает следующими свойствами:

1.  $F(-\infty) = 0$ . Это свойство отражает тот факт, что нет значений случайной величины, которые были бы меньше, чем отрицательная бесконечность.

2.  $F(x)$  — неубывающая функция. Действительно, пусть  $x_2 > x_1$ . Тогда

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

или

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

Так как вероятность не может быть отрицательной величиной, то из последнего соотношения имеем:

$$F(x_2) \geq F(x_1), \quad x_2 > x_1.$$

3.  $F(\infty) = 1$ . Это свойство отражает тот факт, что событие "случайная величина принимает значение, меньшее положительной бесконечности" — достоверно.

4. Вероятность попадания случайной величины в интервал от  $x_1$  до  $x_2$  равна разности функции распределения в точках  $x_1$  и  $x_2$ :

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

До сих пор речь шла о функции распределения дискретных случайных величин. Однако все рассуждения полностью сохраняют свой смысл и для непрерывных случайных величин. В частности, если в примере по разрезанию нити считать, что любые координаты в интервале от 0 до  $L$  равновозможны, то вероятность того, что разрез придется на участок от нуля до  $x$ , будет равна  $x/L$  и, следовательно,  $F(x) = x/L$ . В общем случае функция распределения  $F(x)$  будет представлять собой некоторую функцию, удовлетворяющую приведенным выше свойствам.

Функция распределения непрерывной случайной величины — функция непрерывная (рис. 2.4).

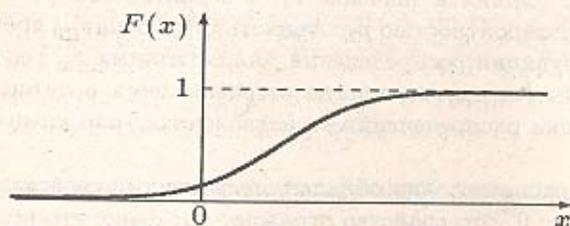


Рис. 2.4. Пример функции распределения непрерывной случайной величины

На практике часто встречаются случайные величины *смешанного типа*, которые могут и непрерывно заполнять некоторый промежуток, и принимать отдельные дискретные значения. Функции распре-

деления таких случайных величин имеют точки разрыва первого рода (рис. 2.5).

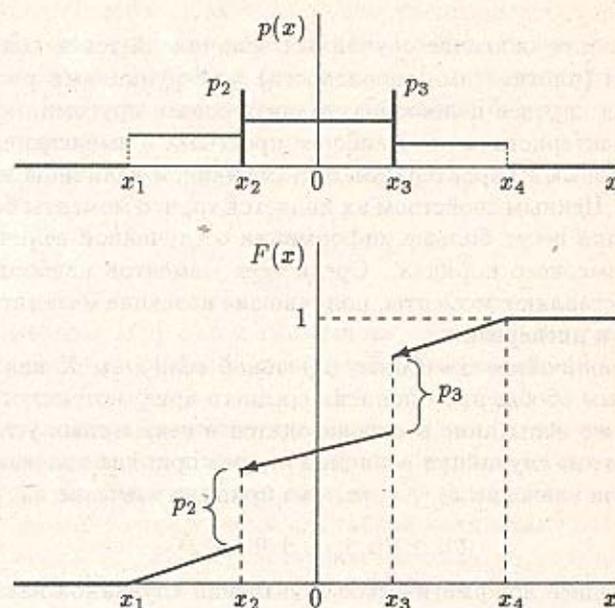


Рис. 2.5. Пример плотности вероятности и функции распределения случайной величины смешанного типа

Последнее из вышеперечисленных свойств функции распределения остается в силе и тогда, когда случайная величина является дискретной или смешанной, но при этом следует помнить, что интервал полузамкнутый, то есть левая точка включается в интервал, а правая нет.

Для непрерывной случайной величины  $X$  в соответствии с определением

$$F(x) = \int_{-\infty}^x p(x) dx.$$

Отсюда следует, что производная от функции распределения равна плотности вероятности:

$$F'(x) = \frac{dF(x)}{dx} = p(x).$$

## 2.4. Числовые характеристики случайных величин

Хотя полное описание случайных величин дается законами распределения (плотностями вероятности) или функциями распределения, в ряде случаев целесообразно оперировать другими, более простыми характеристиками. Наиболее простыми в вычислении и важными числовыми характеристиками случайной величины являются моменты. Ценным свойством их является то, что моменты более низкого порядка несут больше информации о случайной величине, чем моменты высокого порядка. Среди всех моментов наибольший интерес представляют моменты, получившие название математического ожидания и дисперсии.

Математическое ожидание случайной величины  $X$  является вероятностным обобщением понятия среднего арифметического. Пусть одно и то же испытание воспроизводится в неизменных условиях  $N$  раз и при этом случайная величина  $m_1$  раз приняла значение  $x_1$ ,  $m_2$  раз приняла значение  $x_2$ , ...,  $m_n$  раз приняла значение  $x_n$ , где

$$m_1 + m_2 + \dots + m_n = N.$$

Найдем среднее арифметическое  $\bar{x}$  значений случайной величины  $X$  в произведенных  $N$  испытаниях:

$$\bar{x} = \frac{m_1 x_1 + m_2 x_2 + \dots + m_n x_n}{N} = x_1 \frac{m_1}{N} + x_2 \frac{m_2}{N} + \dots + x_n \frac{m_n}{N}.$$

При большом числе испытаний относительные частоты  $\frac{m_1}{N}$ ,  $\frac{m_2}{N}$ , ...,  $\frac{m_n}{N}$  событий, состоящих соответственно в том, что случайная величина принимает значения  $x_1$ ,  $x_2$ , ...,  $x_n$ , группируются около вероятностей  $p_1$ ,  $p_2$ , ...,  $p_n$  этих значений. Следовательно, среднее арифметическое значение случайной величины в достаточно длинной серии испытаний будет группироваться вокруг величины

$$m_x = x_1 p_1 + x_2 p_2 + \dots + x_n p_n,$$

которая называется *математическим ожиданием*, или *средним значением*, случайной величины.

Итак, математическое ожидание дискретной случайной величины — это ее значение, около которого в опытах достаточной длины группируются средние арифметические ее наблюдаемых значений.

При этом слово “группируются” следует понимать в том же смысле, в котором оно употребляется в отношении частот событий, которые группируются при больших  $N$  вокруг соответствующих вероятностей. Это значит, что сколько-нибудь существенные отклонения среднего арифметического случайной величины в  $N$  испытаниях от математического ожидания при достаточно больших  $N$  будут происходить достаточно редко.

Предыдущую формулу можно записать в виде

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

Далее символом  $M[\cdot]$  будем обозначать операцию математического ожидания над случайной величиной, указанной в квадратных скобках.

Если сравнить выражение для математического ожидания с формулой для определения центра тяжести системы материальных точек, становится понятным еще одно название математического ожидания — *центр распределения* случайной величины. Если на оси  $x$  в точках  $x_1$ ,  $x_2$ , ...,  $x_n$  сосредоточены массы  $p_1$ ,  $p_2$ , ...,  $p_n$ , то центр тяжести такой системы находится по формуле

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

которая совпадает с формулой для математического ожидания в силу условия

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

Отметим, что математическое ожидание дискретной случайной величины может не совпадать ни с одним из ее возможных значений, то есть может находиться между этими значениями.

Если множество возможных значений дискретной случайной величины  $X$  бесконечно, то математическое ожидание определяется суммой бесконечного ряда

$$M[X] = \sum_{i=1}^{\infty} x_i p_i$$

при условии, что этот ряд абсолютно сходится. В противном случае говорят, что математическое ожидание не существует.

Если случайная величина является непрерывной, то можно провести аналогичные рассуждения, рассматривая непрерывную случайную величину как предельный случай дискретной. В результате получим формулу

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

которая принимается за определение математического ожидания непрерывной случайной величины. Если в этой формуле несобственный интеграл не является абсолютно сходящимся, то говорят, что математическое ожидание не существует.

Последней формуле можно дать следующую механическую интерпретацию. Рассмотрим стержень (вообще говоря, бесконечной длины), масса которого меняется по длине пропорционально  $p(x)$ . Тогда  $M[X]$  дает координаты центра тяжести этого стержня.

Математическое ожидание имеет следующие свойства:

1. Математическое ожидание неслучайной (постоянной) величины равно этой величине. Действительно, плотность вероятности неслучайной величины  $C$  равна  $p(x) = \delta(x - C)$ . Отсюда

$$M[X] = \int_{-\infty}^{\infty} C\delta(x - C) dx = C \int_{-\infty}^{\infty} \delta(x - C) dx = C.$$

2. Математическое ожидание имеет размерность случайной величины.

3. Математическое ожидание случайной величины  $X$ , имеющей симметричную плотность вероятности относительно прямой  $x = a$ , то есть  $p(a - x) = p(a + x)$ , равно  $M[X] = a$  (рис. 2.6). В частности, если плотность вероятности  $p(x)$  — четная функция, то есть  $p(x) = p(-x)$ , то  $M[X] = 0$ .

4. Неслучайную величину можно выносить за знак математического ожидания:

$$M[CX] = CM[X].$$

5. Математическое ожидание суммы (разности) случайных величин равно сумме (разности) их математических ожиданий.

Доказательство последнего утверждения проведем для суммы (разности) двух случайных величин  $X$  и  $Y$ , предположив для упрощения,

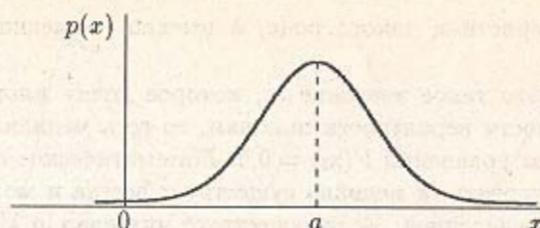


Рис. 2.6. Математическое ожидание случайной величины с симметричной плотностью вероятности

что каждая из случайных величин имеет лишь два возможных значения:  $x_1, x_2$  с вероятностями  $p_1, p_2$  и  $y_1, y_2$  с вероятностями  $q_1, q_2$ . При этом возможными значениями величины  $X \pm Y$  будут  $x_1 \pm y_1, x_1 \pm y_2, x_2 \pm y_1, x_2 \pm y_2$ . Обозначим вероятности этих значений соответственно через  $p_{11}, p_{12}, p_{21}, p_{22}$ .

Из определения математического ожидания имеем

$$M[X \pm Y] = (x_1 \pm y_1)p_{11} + (x_1 \pm y_2)p_{12} + (x_2 \pm y_1)p_{21} + (x_2 \pm y_2)p_{22}$$

или

$$M[X \pm Y] = x_1(p_{11} + p_{12}) + x_2(p_{21} + p_{22}) \pm y_1(p_{11} + p_{21}) \pm y_2(p_{12} + p_{22}).$$

Покажем, что  $p_{11} + p_{12} = p_1$ . Событие, состоящее в том, что  $X$  примет значение  $x_1$  (вероятность этого события равна  $p_1$ ), влечет за собой событие, которое состоит в том, что  $X \pm Y$  примет значение  $x_1 \pm y_1$  или  $x_1 \pm y_2$  (вероятность этого события по теореме сложения вероятностей равна  $p_{11} + p_{12}$ ). Отсюда следует, что  $p_{11} + p_{12} = p_1$ . Аналогично  $p_{21} + p_{22} = p_2, p_{11} + p_{21} = q_1, p_{12} + p_{22} = q_2$ . С учетом этих равенств получим

$$M[X \pm Y] = (x_1p_1 + x_2p_2) \pm (y_1q_1 + y_2q_2) = M[X] \pm M[Y].$$

Аналогичным образом это свойство доказывается для суммы (разности) двух случайных величин не только с двумя, но и с большим числом возможных значений, а также для суммы (разности) нескольких случайных величин.

Кроме математического ожидания, которое, как отмечалось, характеризует расположение "центра" закона распределения, вводят

еще две характеристики такого рода, а именно — медиану  $x_l$  и моду  $x_m$ .

*Медиана* — это такое значение  $x_l$ , которое делит площадь под графиком плотности вероятности пополам, то есть медиана является любым корнем уравнения  $F(x_l) = 0,5$ . Математическое ожидание может не существовать, а медиана существует всегда и может быть неоднозначно определенной. Если существует интервал  $[a, b]$ , на котором  $F(x) = 0,5$ , то любая точка этого интервала может быть медианой (рис. 2.7).

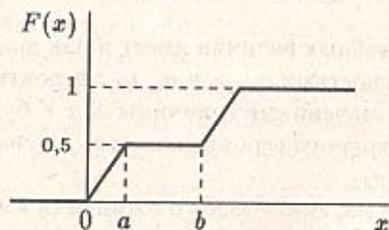


Рис. 2.7. Любое значение случайной величины в интервале  $[a, b]$  может быть принято в качестве медианы

В приложениях иногда рассматривают *квантили*. Квантиль  $x_p$  порядка  $p$  есть корень уравнения  $F(x_p) = p$ , где  $p$  — некоторое данное число,  $0 < p < 1$ . Медиана является квантилем порядка  $1/2$ , то есть  $x_l = x_{1/2}$ .

*Модой* непрерывной случайной величины называется значение  $x_m$ , соответствующее максимуму плотности вероятности. Если плотность вероятности имеет один максимум, то она называется *унимодальной*. При наличии двух или более максимумов плотности вероятности называется *бимодальной* или *мультимодальной*. Если для дискретной случайной величины возможные ее значения расположены в порядке возрастания, то точка  $x_\nu$  называется модой, если  $p_\nu > p_{\nu-1}$  и  $p_\nu > p_{\nu+1}$ .

Соотношения между математическим ожиданием, медианой и модой для некоторых плотностей вероятностей показаны на рис. 2.8. Математическое ожидание «чувствительно» к «хвостам» закона распределения, медиана менее чувствительна к ним, а на моду крайние значения вообще не влияют. Очевидно, что для унимодального симметричного закона распределения все три характеристики совпадают.

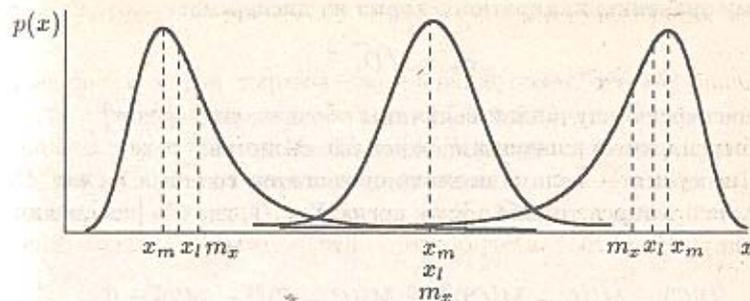


Рис. 2.8. Соотношения между математическим ожиданием, модой и медианой для некоторых распределений

Математическое ожидание случайной величины характеризует положение центра распределения. Второй отличительной особенностью случайной величины является степень разброса значений этой величины относительно центра распределения. Наиболее употребительной оценкой такого разброса является дисперсия.

*Дисперсия* случайной величины — числовая характеристика, равная математическому ожиданию квадрата отклонения случайной величины от ее математического ожидания:

$$D_x = M[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 p(x) dx = D[X].$$

Здесь  $D[\cdot]$  обозначает операцию нахождения дисперсии величины, указанной в скобках.

Часто используют понятие *центрированной случайной величины*  $\overset{\circ}{X} = X - m_x$ . Тогда дисперсия случайной величины — математическое ожидание квадрата соответствующей ей центрированной случайной величины:

$$D_x = M[\overset{\circ}{X}^2] = D[X].$$

Поскольку дисперсия имеет размерность квадрата случайной величины, то для характеристики меры рассеяния значений случайной величины относительно математического ожидания пользуются средним квадратическим отклонением  $\sigma_x$ , имеющим размерность случайной величины. *Среднее квадратическое отклонение* равно положи-

тельному значению квадратного корня из дисперсии:

$$\sigma_x = \sqrt{D_x}.$$

Часто дисперсию случайной величины обозначают через  $\sigma_x^2$ .

Дисперсия имеет следующие основные свойства:

1. Дисперсия — величина неотрицательная, то есть  $D_x \geq 0$ . При этом  $D_x = 0$  тогда и только тогда, когда  $X = C$ , где  $C$  — постоянная величина:

$$D[C] = M[(C - M[C])^2] = M[(C - C)^2] = M[0] = 0.$$

2. Постоянный множитель можно выносить за знак дисперсии, предварительно возведя его в квадрат. Действительно,

$$\begin{aligned} D[kX] &= M[(kX - M[kX])^2] = M[(kX - kM[X])^2] = \\ &= M[(k(X - M[X]))^2] = k^2 M[(X - M[X])^2] = k^2 D[X]. \end{aligned}$$

3. Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадратом ее математического ожидания, то есть

$$D[X] = M[X^2] - (M[X])^2.$$

Покажем справедливость этого выражения:

$$\begin{aligned} D[X] &= M[(X - M[X])^2] = M[X^2 - 2XM[X] + (M[X])^2] = \\ &= M[X^2] - 2M[XM[X]] + M[(M[X])^2] = \\ &= M[X^2] - 2(M[X])^2 + (M[X])^2 = M[X^2] - (M[X])^2. \end{aligned}$$

По этой формуле часто вычисляют дисперсию на практике.

Математическое ожидание и дисперсия являются частными случаями соответственно начальных моментов и центральных моментов распределения случайной величины. Начальный момент порядка  $\nu$  определяется выражением:

$$M[X^\nu] = m_\nu = \int_{-\infty}^{\infty} x^\nu p(x) dx.$$

При  $\nu = 0$  имеем  $m_0 = 1$ , то есть этот момент равен площади под кривой плотности вероятности. При  $\nu = 1$  имеем  $m_1 = m_x = M[X]$ , при  $\nu = 2$  имеем  $m_2 = M[X^2]$ .

Центральный момент  $\nu$ -го порядка дается выражением:

$$M[(X - m_x)^\nu] = \mu_\nu = \int_{-\infty}^{\infty} (x - m_x)^\nu p(x) dx.$$

При  $\nu = 2$  этот момент является дисперсией.

Для дискретных случайных величин

$$m_\nu = \sum_{i=1}^n x_i^\nu p_i;$$

$$\mu_\nu = \sum_{i=1}^n (x_i - m_x)^\nu p_i.$$

Центральный момент третьего порядка  $\mu_3$  может служить критерием для оценки асимметрии закона распределения относительно оси, параллельной оси ординат и проходящей через математическое ожидание случайной величины. В качестве показателя асимметрии вводят коэффициент асимметрии

$$A = \frac{\mu_3}{\sqrt{D_x^3}}.$$

Для симметричных законов распределения  $\mu_3 = 0$  и, следовательно,  $A = 0$ . Чем больше коэффициент  $A$  отличается от нуля, тем существеннее асимметрия закона распределения.

Центральный момент четвертого порядка  $\mu_4$  используется при вычислении так называемого коэффициента эксцесса, характеризующего островершинность или плосковершинность распределения:

$$E = \frac{\mu_4}{D_x^2} - 3.$$

## 2.5. Характеристические функции

Вместо закона распределения или функции распределения для описания случайной величины можно использовать так называемую характеристическую функцию. Пользуясь этой функцией, можно упростить расчет числовых характеристик случайных величин, а также, как будет показано позднее, упростить задачу нахождения плотности вероятности суммы нескольких независимых случайных величин.

Характеристические функции по отношению к законам распределения играют в теории вероятностей ту же роль, какую играют спектральные плотности по отношению к временным функциям в теории сигналов.

Характеристическая функция случайной величины  $X$  определяется как математическое ожидание случайной величины  $e^{j\omega X}$ , то есть

$$\theta(j\omega) = M[e^{j\omega X}] = \int_{-\infty}^{\infty} e^{j\omega x} p(x) dx,$$

где  $\omega$  — вещественная переменная,  $j = \sqrt{-1}$  — мнимая единица.

Легко увидеть, что характеристическая функция является преобразованием Фурье от закона распределения. Отметим, что обычно в прямом преобразовании Фурье используют экспоненту в отрицательной степени.

В случае дискретной случайной величины  $X$

$$\theta(j\omega) = \sum_{i=1}^n e^{j\omega x_i} p_i.$$

Плотность вероятности может быть найдена как обратное преобразование Фурье от характеристической функции:

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta(j\omega) e^{-j\omega x} d\omega.$$

Как уже отмечалось, при помощи характеристической функции можно достаточно просто находить числовые характеристики случайных величин, а именно моменты, если они существуют.

Дифференцируя  $\theta(j\omega)$  по  $\omega$ , получаем:

$$\frac{d^{\nu} \theta(j\omega)}{d\omega^{\nu}} = j^{\nu} \int_{-\infty}^{\infty} x^{\nu} e^{j\omega x} p(x) dx.$$

Положив  $\omega = 0$ , найдем простую связь между значениями производных характеристической функции при  $\omega = 0$  и начальными моментами:

$$\left. \frac{d^{\nu} \theta(j\omega)}{d\omega^{\nu}} \right|_{\omega=0} = \frac{d^{\nu} \theta(0)}{d\omega^{\nu}} = j^{\nu} \int_{-\infty}^{\infty} x^{\nu} e^{j0x} p(x) dx = j^{\nu} \int_{-\infty}^{\infty} x^{\nu} p(x) dx = j^{\nu} m_{\nu}.$$

Отсюда

$$m_{\nu} = j^{-\nu} \left. \frac{d^{\nu} \theta(j\omega)}{d\omega^{\nu}} \right|_{\omega=0}.$$

Таким образом, если интересующие нас моменты существуют, то эта формула дает простой способ вычисления их путем дифференцирования характеристической функции.

Для нахождения центральных моментов пользуются натуральным логарифмом характеристической функции. Эту функцию называют *логарифмической характеристической функцией*.

## 2.6. Преобразование законов распределения и моментов

Пусть известен закон распределения  $p_i = P(X = x_i)$  дискретной случайной величины  $X$  или плотность вероятности  $p_x(x)$  непрерывной случайной величины и нужно найти закон распределения  $p_i = P(Y = y_i)$  или плотность вероятности  $p_y(y)$  случайной величины  $Y$ , связанной со случайной величиной  $X$  детерминированной функциональной зависимостью  $Y = g(X)$ .

Рассмотрим сначала преобразование закона распределения дискретной случайной величины. Пусть дискретная случайная величина  $X$  принимает значения  $x_i$  с вероятностями  $p_i = P(X = x_i)$ ,  $i = 1, 2, \dots$ . Возможными значениями функции  $Y = g(X)$  будут  $y_i = g(x_i)$ , и остается только определить вероятности этих значений.

Если функция  $g(x)$  однозначная, то есть все значения  $y_i = g(x_i)$  различны, то случайная величина  $Y$  примет значение  $y_i$  тогда и только тогда, когда величина  $X$  примет значение  $x_i$ . Следовательно,

$$P(Y = y_i) = P(X = x_i) = p_i, \quad i = 1, 2, \dots,$$

так что таблицы распределения величин  $X$  и  $Y$  отличаются только строкой возможных значений (рис. 2.9).

Рассмотрим в качестве примера линейное преобразование

$$Y = aX + b.$$

Тогда получим закон распределения  $p_i = P(Y = y_i)$ , который представлен на рис. 2.10.

$X$	$x_1$	$x_2$	...
$p_i(x)$	$p_1$	$p_2$	...

$Y$	$g(x_1)$	$g(x_2)$	...
$p_i(y)$	$p_1$	$p_2$	...

Рис. 2.9. Преобразование закона распределения дискретной случайной величины при однозначной функциональной зависимости

$Y$	$ax_1 + b$	$ax_2 + b$	$ax_3 + b$	...
$p_i(y)$	$p_1$	$p_2$	$p_3$	...

Рис. 2.10. Закон распределения линейно преобразованной дискретной случайной величины

В этом примере все значения  $ax_i + b$  при различных  $x_i$  различны.

Если функция  $g(x)$  неоднозначная, то есть существует несколько значений величины  $X$ , для которых  $g(x_i)$  принимает одно и то же значение, то в соответствии с правилом сложения вероятностей для получения вероятности события  $Y = y_i$  нужно сложить вероятности соответствующих значений  $x_k$ :

$$P(Y = y_i) = \sum p_k,$$

где сумма берется по всем индексам  $k$ , для которых значения  $g(x_k)$  совпадают со значениями  $g(x_i) = y_i$ .

*Рассмотрим пример.* Пусть случайная величина  $X$  принимает три значения:  $-1, 0, 1$  с вероятностями соответственно  $p_1, p_2, p_3$  (рис. 2.11). Пусть  $Y = X^2$ . Тогда получим закон распределения случайной величины  $Y$  в виде таблицы на рис. 2.11.

$X$	$-1$	$0$	$1$
$p_i(x)$	$p_1$	$p_2$	$p_3$

$Y$	$0$	$1$
$p_j(y)$	$p_2$	$p_1 + p_3$

Рис. 2.11. Закон распределения квадратично преобразованной дискретной случайной величины  $X$

Таким образом, случайная величина  $Y = X^2$  принимает только два значения:  $0$  с вероятностью  $p_2$  и  $1$  вероятностью  $p_1 + p_3$ .

Теперь рассмотрим случай, когда случайная величина непрерывна. Пусть известна плотность вероятности  $p_x(x)$  случайной величины

$X$ , и нужно найти плотность вероятности  $p_y(y)$  случайной величины  $Y = g(X)$ , где  $g(X)$  — однозначная дифференцируемая функция (рис. 2.12). Тогда существует однозначная обратная функция  $X = h(Y)$ .

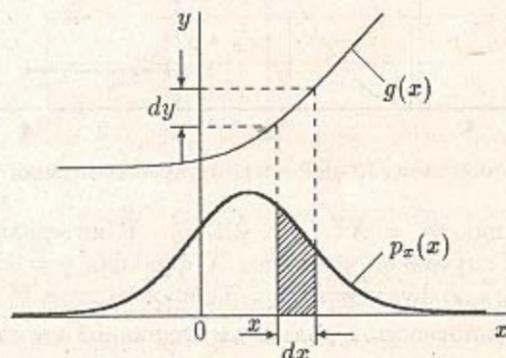


Рис. 2.12. Взаимно однозначное преобразование

Так как случайные величины связаны однозначной детерминированной зависимостью и полученное значение случайной величины  $X$  заключено в интервале  $[x, x + dx]$ , то величина  $Y$  будет находиться в интервале  $[y, y + dy]$ , где  $y = g(x)$ ;  $dy = g'(x) dx$  (рис. 2.12). Отсюда следует, что вероятности этих двух событий равны, то есть выполняется свойство инвариантности дифференциала вероятности:

$$p_x(x) dx = p_y(y) dy.$$

Отсюда

$$p_y(y) = p_x(x) \left| \frac{dx}{dy} \right| = p_x(h(y)) |h'(y)|.$$

Поскольку при  $dx > 0$  дифференциал  $dy > 0$ , когда функция  $g(x)$  возрастающая, и  $dy < 0$ , когда функция  $g(x)$  убывающая, а плотности вероятности не могут быть отрицательными, то в последней формуле следует брать модуль производной.

*Рассмотрим пример.* Пусть случайная величина имеет плотность вероятности

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

Вид этой плотности вероятности приведен на рис. 2.13,а.

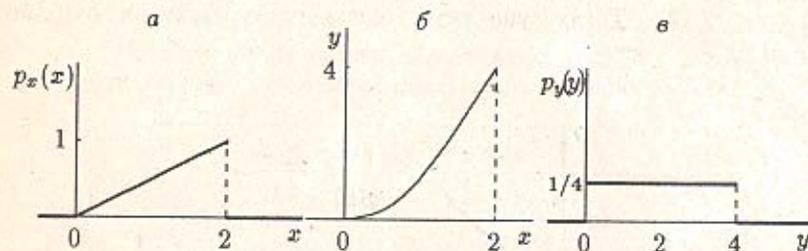


Рис. 2.13. Пример взаимно однозначного преобразования

Рассмотрим функцию  $Y = X^2$  (рис. 2.13,б). В интервале  $[0, 2]$  возможных значений случайной величины  $X$  функция  $y = x^2$  однозначна и имеет обратную функцию  $x = \sqrt{y}$  с производной  $\frac{dx}{dy} = \frac{1}{2\sqrt{y}}$ , ( $0 < y \leq 4$ ). Поэтому плотность вероятности случайной величины  $Y$  равна

$$p_y(y) = \begin{cases} \frac{1}{2}\sqrt{y} \frac{1}{2\sqrt{y}} = \frac{1}{4}, & \text{при } 0 \leq y \leq 4, \\ 0, & \text{при } y < 0, y > 4, \end{cases}$$

то есть случайная величина  $Y$  распределена равномерно в интервале  $[0, 4]$  (рис. 2.13,в).

Рассмотрим теперь случай, когда функция  $Y = g(X)$  неоднозначная, то есть нескольким значениям  $x$  соответствует одно значение  $y$ . Пусть имеются две ветви обратной функции —  $h_1(y)$  и  $h_2(y)$  (рис. 2.14).

В рассматриваемом случае выполнение неравенства

$$y \leq Y \leq y + dy$$

обеспечивается двумя несовместными возможностями

$$x_1 \leq X < x_1 + dx_1 \quad \text{или} \quad x_2 + dx_2 \leq X < x_2.$$

Поэтому вероятность выполнения неравенства  $y \leq Y \leq y + dy$  должна равняться сумме вероятностей выполнения каждого из двух последних неравенств:

$$p_y(y) dy = p_x(x_1) dx_1 + p_x(x_2) |dx_2|.$$

Выразив в правой части  $x$  через  $y$ , окончательно получим

$$p_y(y) = p_x(h_1(y)) |h_1'(y)| + p_x(h_2(y)) |h_2'(y)|.$$

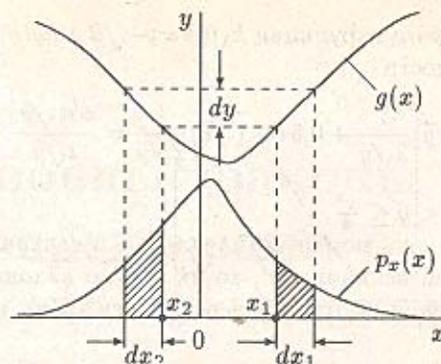


Рис. 2.14. Двухзначное нелинейное преобразование

Если имеется большее число ветвей обратной функции, то следует брать сумму по всем ветвям.

Отметим, что при неоднозначном преобразовании  $Y = g(X)$  нарушается однозначное соответствие между плотностями вероятности случайных величин  $X$  и  $Y$ , то есть по найденной плотности вероятности  $p_y(y)$  нельзя однозначно восстановить плотность вероятности  $p_x(x)$ .

Рассмотрим пример. Пусть требуется найти плотность вероятности случайной величины  $Y = X^2$ , где  $X$  — непрерывная случайная величина с плотностью вероятности  $p_x(x) = 0,5 \cos x$  на промежутке  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  (рис. 2.15,а).

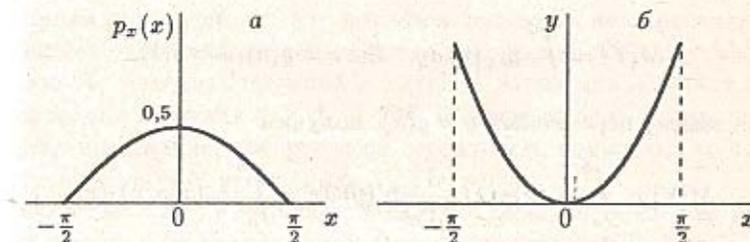


Рис. 2.15. Пример двухзначного нелинейного преобразования

Функция  $y = x^2$  монотонно убывает на интервале  $[-\frac{\pi}{2}, 0]$  и возрастает на интервале  $[0, \frac{\pi}{2}]$  (рис. 2.15,б). Обратными на интервале

$[-\frac{\pi}{2}, \frac{\pi}{2}]$  к функции  $y = x^2$  будут функции  $h_1(y) = -\sqrt{y}$  и  $h_2(y) = \sqrt{y}$ . Отсюда плотность вероятности

$$p_y(y) = 0,5 \cos(-\sqrt{y}) \frac{1}{2\sqrt{y}} + 0,5 \cos(\sqrt{y}) \frac{1}{2\sqrt{y}} = \frac{\cos \sqrt{y}}{2\sqrt{y}}$$

с областью определения  $0 < y \leq \frac{\pi^2}{4}$ .

Если требуется найти лишь моменты (начальные и центральные) преобразованной случайной величины  $Y$ , то их можно находить без предварительного вычисления плотности вероятности  $p_y(y)$ , пользуясь формулами

$$M[Y^n] = M[g^n(X)] = \int_{-\infty}^{\infty} g^n(x) p_x(x) dx,$$

$$M[(Y - m_y)^n] = M[(g(X) - m_y)^n] = \int_{-\infty}^{\infty} (g(x) - m_y)^n p_x(x) dx,$$

где  $m_y = M[Y]$ .

Покажем справедливость первой формулы для математического ожидания случайной величины  $Y$ . Пусть  $X$  принимает значения только на интервале  $[a, b]$ , а  $g(X)$  — определенная на этом интервале непрерывно дифференцируемая функция, у которой  $g'(x) > 0$  для всех  $x$ , принадлежащих  $[a, b]$ . Плотность вероятности случайной величины  $Y$  определяется, как известно, по формуле

$$p_y(y) = p_x(h(y)) |h'(y)|.$$

Поэтому

$$M[Y] = \int_c^d y p_y(y) dy, \quad \text{где } c = g(a), d = g(b).$$

Сделав замену переменных  $y = g(x)$ , получим

$$M[Y] = \int_a^b p_x(x) g(x) \frac{1}{h'(y)} h'(y) dx = \int_a^b p_x(x) g(x) dx.$$

Таким образом,

$$M[Y] = M[g(X)] = \int_a^b g(x) p_x(x) dx.$$

## Глава 3

# Основные законы распределения случайных величин

## 3.1. Биномиальное распределение

Биномиальное распределение, или распределение Бернулли, описывает дискретные события следующего типа. Пусть некоторое испытание повторяется  $N$  раз и результаты, полученные в каждом испытании, не зависят друг от друга. Пусть в каждом испытании может произойти или не произойти событие  $A$ , вероятность осуществления которого в отдельном испытании одна и та же и равна  $p$ . Например, положим в урну  $N$  одинаковых шаров, пометив предварительно  $M$  шаров меткой  $A$ , так что вероятность вынуть шар с этой меткой равна  $M/N = p$ . Затем, вынув из урны наугад один шар, запишем, есть на нем метка или нет. Вернем шар в урну, тщательно перемешаем шары и повторим этот процесс до получения  $n$  записей наличия метки  $A$ . Такую последовательность испытаний называют последовательностью независимых испытаний по схеме Бернулли.

Будем рассматривать эту последовательность независимых испытаний как одно сложное испытание, с которым связана случайная величина  $X$  — число появлений события  $A$ . Величина  $X$  может принимать только значения  $0, 1, 2, \dots, N$ . Найдем закон ее распределения, то есть выражение, которое дает вероятность появления события  $A$  точно  $n$  раз при  $N$  испытаниях  $P(X = n)$ .

Событие  $X = N$  означает появление события  $A$  во всех испытаниях, поэтому по правилу умножения вероятностей  $P(X = N) = p^N$ . Аналогично  $P(X = 0) = q^N$ , где  $q = (1 - p)$  — вероятность наступления события, противоположного событию  $A$ .

Событие  $X = n$  ( $n = 1, 2, \dots, N - 1$ ) означает, что за  $N$  испытаний событие  $A$  наступит точно  $n$  раз и, значит,  $N - n$  раз наступит

противоположное событие  $\bar{A}$ . Вероятность того, что событие  $A$  наступит в первых  $n$  испытаниях и не наступит в остальных, может быть подсчитана по правилу умножения вероятностей:

$$P(\underbrace{A \cdot A \cdot \dots \cdot A}_n \cdot \underbrace{\bar{A} \cdot \bar{A} \cdot \dots \cdot \bar{A}}_{(N-n)}) = p^n q^{N-n}.$$

Эта вероятность не зависит от того, в каких именно испытаниях наступит событие  $A$ . Поэтому по правилу сложения вероятностей искомая вероятность  $P(X = n)$  равна вероятности  $p^n q^{N-n}$ , умноженной на число способов, которыми можно выбрать  $n$  натуральных чисел из  $N$ , то есть на число сочетаний  $C_N^n$  из  $N$  элементов по  $n$ :

$$P(X = n) = C_N^n p^n q^{N-n}, \quad (3.1)$$

$$\text{где } C_N^n = \frac{N!}{(N-n)!n!} = \frac{N(N-1)(N-2)\dots(N-n+1)}{n!}.$$

Данная формула остается справедливой и для крайних значений  $n = 0$  и  $n = N$ , поскольку  $0! = 1$  и  $C_N^0 = C_N^N = 1$ .

Распределение вероятностей, определяемое формулой (3.1), называется *биномиальным распределением* (или *распределением Бернулли*). Название "биномиальное" связано с тем, что вероятности (3.1) совпадают с соответствующими членами разложения бинома  $(p+q)^N$  по степеням  $p$ :

$$(p+q)^N = \sum_{n=0}^N C_N^n p^n q^{N-n} = q^N + Npq^{N-1} + \dots + C_N^n p^n q^{N-n} + \dots + p^N.$$

Отсюда сразу видно, что сумма всех вероятностей (3.1) равна единице, так как  $p+q = 1$ , то есть выполняется условие нормировки биномиального распределения.

Найдем математическое ожидание и дисперсию числа появлений события  $A$  в  $N$  испытаниях:

$$\begin{aligned} M[X] &= \sum_{n=0}^N nP(X = n) = \sum_{n=0}^N \frac{nN!}{(N-n)!n!} p^n q^{N-n} = \\ &= Np \sum_{n=1}^N \frac{(N-1)!}{(N-n)!(n-1)!} p^{n-1} q^{N-n} = Np \sum_{s=0}^{N-1} \frac{(N-1)!}{(N-s-1)!s!} p^s q^{N-s-1}, \end{aligned}$$

где  $s = n - 1$ .

Пусть  $N - 1 = k$ . Тогда

$$M[X] = Np \sum_{s=0}^k C_k^s p^s q^{k-s} = Np.$$

Для вычисления дисперсии найдем сначала  $M[X^2]$ :

$$\begin{aligned} M[X^2] &= \sum_{n=0}^N n^2 P(X = n) = \sum_{n=0}^N (n^2 - n + n) P(X = n) = \\ &= \sum_{n=0}^N n(n-1) P(X = n) + \sum_{n=0}^N n P(X = n). \end{aligned}$$

Второе слагаемое равно  $Np$ , а первое обращается в ноль для  $n = 0$  и  $n = 1$ . Поэтому

$$\begin{aligned} M[X^2] &= \sum_{n=2}^N n(n-1) \frac{N!}{(N-n)!n!} p^n q^{N-n} + Np = \\ &= \sum_{n=2}^N \frac{N!}{(N-n)!(n-2)!} p^n q^{N-n} + Np. \end{aligned}$$

Упростим первое слагаемое, вынеся за сумму  $N(N-1)p^2$ :

$$\sum_{n=2}^N \frac{N!}{(N-n)!(n-2)!} p^n q^{N-n} = N(N-1)p^2 \sum_{n=2}^N \frac{(N-2)!}{(N-n)!(n-2)!} p^{n-2} q^{N-n}.$$

Полагая здесь  $N - 2 = k$ , получим

$$N(N-1)p^2 \sum_{n=2}^{k+2} \frac{k!}{(k-(n-2))!(n-2)!} p^{n-2} q^{k-(n-2)}.$$

Примем далее  $n - 2 = m$ . Тогда получим

$$N(N-1)p^2 \sum_{m=0}^k \frac{k!}{(k-m)!m!} p^m q^{k-m} = N(N-1)p^2,$$

поскольку, как уже указывалось, сумма равна 1.

Итак,

$$M[X^2] = N(N-1)p^2 + Np.$$

Вычислим дисперсию:

$$\begin{aligned} D[X] &= M[X^2] - (M[X])^2 = N(N-1)p^2 + Np - N^2p^2 = \\ &= N^2p^2 - Np^2 + Np - N^2p^2 = Np(1-p) = Npq. \end{aligned}$$

Отсюда среднее квадратическое отклонение

$$\sigma_x = \sqrt{Npq}.$$

Изобразим биномиальный закон распределения графически для числа испытаний  $N = 8$  и вероятностей  $p$ , равных 0,1, 0,3 и 0,5 (рис. 3.1). Для этих случаев  $M[X]$  и  $\sigma_x$  соответственно равны:

$$M[X] = 0,8, \sigma_x = 0,85; M[X] = 2,4, \sigma_x = 1,30; M[X] = 4,0, \sigma_x = 1,41.$$

Приведенные на рис. 3.1 значения вероятностей  $P(X = n)$  являются округленными.

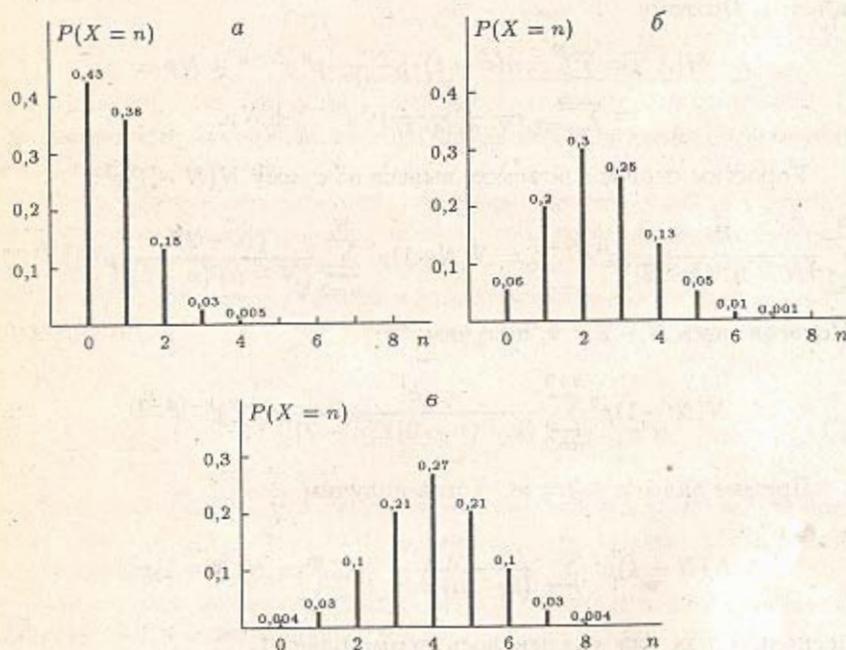


Рис. 3.1. Биномиальный закон распределения для  $N = 8$ :  
а —  $p = 0,1$ ; б —  $p = 0,3$ ; в —  $p = 0,5$

Биномиальный закон распределения весьма часто приходится применять в условиях, когда число независимых испытаний велико. Вычисление вероятностей по формуле (3.1) при этом усложняется. По-

этому представляет интерес асимптотическое приближение для биномиального закона, справедливое при больших  $N$ . Здесь могут встретиться два случая:

1) когда  $N \rightarrow \infty$ , математическое ожидание  $Np$  тоже неограниченно возрастает (случай постоянного  $p$ ); при этом биномиальное распределение сходится к нормальному закону, который будет рассмотрен позднее;

2) при  $N \rightarrow \infty$  произведение  $Np = \lambda$ , то есть математическое ожидание рассматриваемой случайной величины остается конечным (это означает, что вероятность события  $p = \lambda/N$  стремится к нулю). В данном случае распределение сходится к распределению Пуассона.

### 3.2. Распределение Пуассона

Рассмотрим второй случай асимптотического приближения биномиального распределения, когда  $N \rightarrow \infty$ , а  $Np = \lambda$  имеет конечное значение. Если в формулу для биномиального распределения подставить  $p = \lambda/N$ , то можно записать:

$$P(X = n) = \frac{N(N-1)(N-2)\dots(N-(n-1))}{n!} \left(\frac{\lambda}{N}\right)^n \left(1 - \frac{\lambda}{N}\right)^{N-n}.$$

Найдем предел правой части при  $N \rightarrow \infty$ :

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{N(N-1)(N-2)\dots(N-(n-1))}{n!} \left(\frac{\lambda}{N}\right)^n \left(1 - \frac{\lambda}{N}\right)^{N-n} &= \\ = \frac{\lambda^n}{n!} \lim_{N \rightarrow \infty} \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \dots \left(1 - \frac{n-1}{N}\right) \left(1 - \frac{\lambda}{N}\right)^{N-n} &= \\ = \frac{\lambda^n}{n!} \lim_{N \rightarrow \infty} \left(1 - \frac{\lambda}{N}\right)^N \lim_{N \rightarrow \infty} \left(1 - \frac{\lambda}{N}\right)^{-n} &= \end{aligned}$$

Воспользуемся хорошо известными из теории пределов формулами:

$$\lim_{N \rightarrow \infty} \left(1 - \frac{\lambda}{N}\right)^N = e^{-\lambda}; \quad \lim_{N \rightarrow \infty} \left(1 - \frac{\lambda}{N}\right)^{-n} = 1.$$

Тогда окончательно получим распределение вероятностей, называемое *распределением Пуассона*:

$$P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad n = 0, 1, 2, \dots$$

Величина  $\lambda = Np$  в биномиальном распределении имеет смысл математического ожидания. В распределении Пуассона  $\lambda$  — тоже математическое ожидание. В этом нетрудно убедиться непосредственным вычислением. Действительно,

$$\begin{aligned} M[X] &= \sum_{n=0}^{\infty} nP(X=n) = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda} = \\ &= \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda e^{-\lambda} \sum_{\mu=0}^{\infty} \frac{\lambda^{\mu}}{\mu!} = \lambda e^{-\lambda} e^{\lambda} = \lambda. \end{aligned}$$

Покажем, что дисперсия распределения Пуассона также равна  $\lambda$ . Для этого вычислим сначала второй начальный момент:

$$\begin{aligned} M[X^2] &= \sum_{n=0}^{\infty} n^2 \frac{\lambda^n}{n!} e^{-\lambda} = \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} = \\ &= \lambda \left( \sum_{n=1}^{\infty} (n-1) \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} + \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} \right). \end{aligned}$$

Но

$$\sum_{n=1}^{\infty} (n-1) \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} = \sum_{\mu=0}^{\infty} \mu \frac{\lambda^{\mu}}{\mu!} e^{-\lambda} = \lambda,$$

$$\sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} = 1,$$

поэтому  $M[X^2] = \lambda(\lambda + 1)$ . Отсюда

$$D[X] = M[X^2] - (M[X])^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda.$$

Вид распределения Пуассона при различных значениях параметра  $\lambda$  приведен на рис. 3.2.

Отметим, что при малых  $\lambda$  наблюдается асимметрия закона распределения. С ростом  $\lambda$  имеется тенденция к симметрии.

В ряде задач распределение Пуассона выступает не как асимптотическое, а как совершенно точное. Например, часто приходится иметь дело с распределением событий во времени (появление импульсов или электронов и т.д.). В этой связи рассмотрим еще один вывод распределения Пуассона, который позволит более четко установить условия его возникновения и пределы применимости.

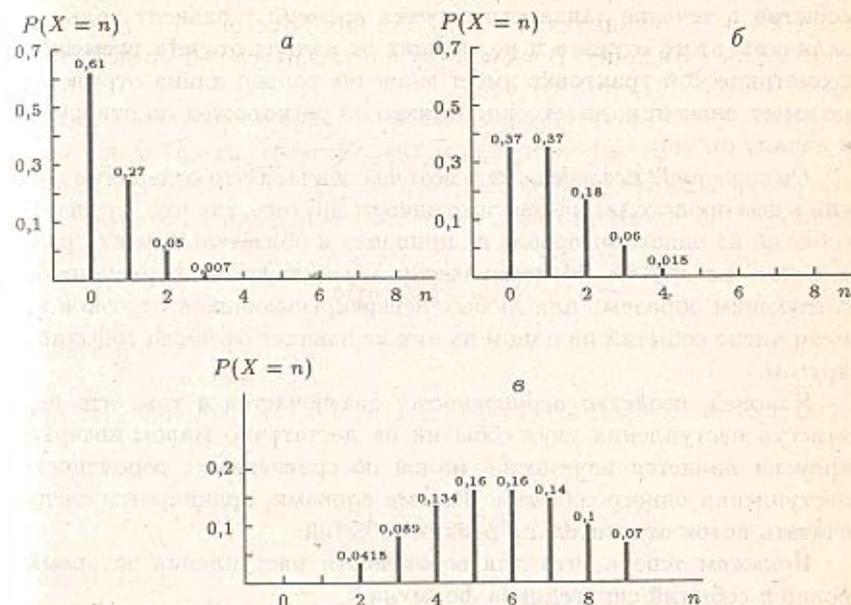


Рис. 3.2. Распределение Пуассона: а —  $\lambda = 0,5$ ; б —  $\lambda = 1,0$ ; в —  $\lambda = 6,0$

Введем сначала некоторые понятия. Потоком событий называется последовательность событий, наступающих одно за другим в случайные моменты времени. Примером могут служить поток вызовов на телефонной станции, поток импульсов, число отказов при работе некоторой системы, последовательность распада частиц при радиоактивном распаде. Геометрически поток событий можно изобразить в виде точек на оси времени (рис. 3.3). Это так называемый *точечный процесс (поток)*.

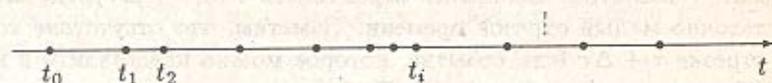


Рис. 3.3. Точечный процесс (поток)

Рассмотрим поток событий, обладающий следующими тремя свойствами: стационарностью, отсутствием последствия и ординарностью.

В стационарном потоке вероятность наступления некоторого числа событий в течение заданного отрезка времени  $\tau$  зависит только от величины этого отрезка и не зависит от начала отсчета времени. В геометрической трактовке имеет значение только длина отрезка  $\tau$  и не имеет значения, далеко или близко он расположен по отношению к началу отсчета.

Отсутствие последствия в потоке означает, что отдельные события в нем происходят независимо одно от другого, так что "сгущения" событий на одном интервале не приводят к обязательным их "разрежениям" на другом. Математически это требование формулируется следующим образом: для любых неперекрывающихся отрезков времени число событий на одном из них не зависит от числа событий на другом.

Наконец, свойство *ординарности* заключается в том, что вероятность наступления двух событий на достаточно малом интервале времени является исчезающе малой по сравнению с вероятностью наступления одного события. Иными словами, ординарным следует считать поток относительно редких событий.

Покажем теперь, что для вероятности наступления за время  $\tau$  ровно  $n$  событий справедлива формула

$$P(X = n, \tau) = \frac{\lambda^n}{n!} e^{-\lambda},$$

где  $\lambda$  — среднее число событий в потоке, приходящееся на интервал длительностью  $\tau$ .

Если обозначить через  $\nu$  среднее число событий в потоке за единицу времени, то эту формулу можно записать в виде

$$P(X = n, \tau) = \frac{(\nu\tau)^n}{n!} e^{-\nu\tau}. \quad (3.2)$$

Докажем ее справедливость при  $n = 0$ .

Предположим, что вероятность  $P(0, \tau)$  отсутствия событий на интервале  $\tau$  известна. Вычислим вероятность  $P(0, \tau + \Delta\tau)$ , где  $\Delta\tau$  — достаточно малый отрезок времени. Заметим, что отсутствие точек на отрезке  $\tau + \Delta\tau$  есть событие, которое можно представить в виде произведения двух событий  $A$  и  $B$ . Одно из них есть отсутствие точек на интервале  $\tau$ , второе — отсутствие их на интервале  $\Delta\tau$ . В силу независимости этих событий (отсутствие последствия в потоке) имеем

$$P(0, \tau + \Delta\tau) = P(0, \tau)P(0, \Delta\tau).$$

Но

$$P(0, \Delta\tau) = 1 - P(1, \Delta\tau) - P(2, \Delta\tau) - \dots \approx 1 - P(1, \Delta\tau),$$

так как в силу ординарности потока вероятностью наступления за время  $\Delta\tau$  двух или более событий можно пренебречь.

Остается найти  $P(1, \Delta\tau)$ . С этой целью запишем выражение для математического ожидания числа точек на интервале  $\Delta\tau$ . С одной стороны,

$$M[n, \Delta\tau] = \nu\Delta\tau,$$

а с другой —

$$M[n, \Delta\tau] = 0 \cdot P(0, \Delta\tau) + 1 \cdot P(1, \Delta\tau) + 2 \cdot P(2, \Delta\tau) + \dots \approx P(1, \Delta\tau).$$

Отсюда

$$P(1, \Delta\tau) = \nu\Delta\tau.$$

Следовательно,

$$P(0, \tau + \Delta\tau) = P(0, \tau)(1 - \nu\Delta\tau),$$

откуда

$$\frac{P(0, \tau + \Delta\tau) - P(0, \tau)}{\Delta\tau} = -\nu P(0, \tau).$$

При  $\Delta\tau \rightarrow 0$  из этого соотношения имеем

$$\frac{dP(0, \tau)}{d\tau} = -\nu P(0, \tau).$$

Это дифференциальное уравнение первого порядка. Его решение относительно  $P(0, \tau)$  при очевидном начальном условии  $P(0, 0) = 1$  имеет следующий вид:

$$P(0, \tau) = e^{-\nu\tau},$$

что совпадает с (3.2) при  $n = 0$ .

Аналогичным образом могут быть получены формулы и для других значений  $n$ .

### 3.3. Экспоненциальное распределение

В качестве модели, описывающей это распределение, рассмотрим интервалы времени  $T$  между двумя последовательными событиями в простейшем потоке, описанном ранее.

Для нахождения плотности вероятности времени  $T$  подсчитаем вероятность того, что  $t < T < t + \Delta t$  ( $t > 0$ ), то есть вероятность того, что за промежуток времени  $t$  от предыдущего события следующее событие не произошло, а за добавочный промежуток времени  $\Delta t$  оно произошло ровно один раз. Тогда, учитывая свойства потока, получим

$$P(t < T < t + \Delta t) = P(0, t)P(1, \Delta t) \approx e^{-\nu t} \nu \Delta t.$$

Отсюда плотность вероятности

$$p(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T < t + \Delta t)}{\Delta t} = \nu e^{-\nu t} \text{ при } t > 0.$$

При  $t < 0$  очевидно, что  $p(t) = 0$ .

Таким образом, плотность вероятности временного интервала между двумя соседними событиями описывается выражением

$$p(t) = \begin{cases} \nu e^{-\nu t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Распределение такого вида при  $\nu > 0$  носит название *экспоненциального распределения*. Часто его называют также *односторонним показательным распределением*. Вид экспоненциального распределения показан на рис. 3.4.

Найдем функцию распределения экспоненциального закона:

$$F(t) = \int_0^t p(\tau) d\tau = \nu \int_0^t e^{-\nu \tau} d\tau = 1 - e^{-\nu t}.$$

Итак,

$$F(t) = \begin{cases} 1 - e^{-\nu t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Математическое ожидание случайной величины, распределенной по экспоненциальному закону

$$M[T] = \int_0^{\infty} t p(t) dt = \int_0^{\infty} t \nu e^{-\nu t} dt = \frac{1}{\nu}.$$

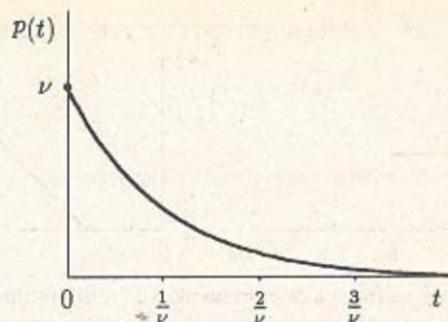


Рис. 3.4. Экспоненциальное распределение

Можно показать, что дисперсия  $T$

$$D[T] = \int_0^{\infty} (t - \frac{1}{\nu})^2 \nu e^{-\nu t} dt = \frac{1}{\nu^2}.$$

### 3.4. Равномерное распределение

Непрерывная случайная величина  $X$ , которая принимает значения только на интервале от  $a$  до  $b$  с постоянной плотностью вероятности, называется *распределенной по равномерному закону*. Из определения следует, что плотность вероятности задается в виде

$$p(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x > b. \end{cases}$$

Найдем функцию распределения данной случайной величины:

$$F(x) = \begin{cases} 0, & x < a, \\ \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

Графики функций  $p(x)$  и  $F(x)$  приведены на рис. 3.5.

Найдем математическое ожидание и дисперсию равномерно распределенной случайной величины:

$$M[X] = \int_{-\infty}^{\infty} x p(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}.$$

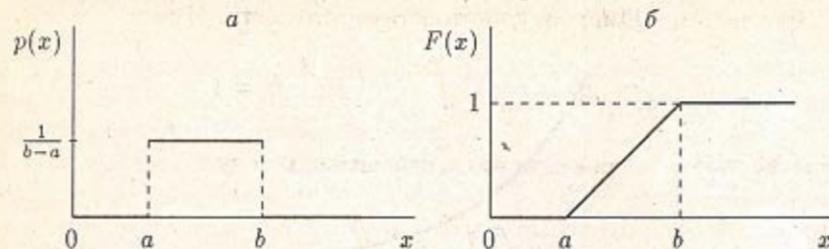


Рис. 3.5. Графики плотности вероятности (а) и функции распределения (б) для равномерно распределенной величины

$$D[X] = \int_{-\infty}^{\infty} (x - m_x)^2 p(x) dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}.$$

По равномерному закону может быть распределена и дискретная случайная величина. В этом случае закон распределения называется *дискретным равномерным* распределением и записывается в виде

$$p_i = P(X = x_i) = \frac{1}{k}, \quad i = 1, 2, \dots, k,$$

где  $k$  — целое положительное число.

### 3.5. Нормальное распределение

Как уже отмечалось, при некоторых условиях ( $N \rightarrow \infty$ ,  $p$  имеет постоянное значение) биномиальное распределение сходится к дискретному *нормальному* (или *гауссовскому*) распределению. Однако нормальный закон распределения, впрочем, как и распределение Пуассона, применяется в статистике гораздо шире и среди других законов распределения играет особую роль.

Плотность вероятности нормально распределенной случайной величины в ненормированной форме имеет вид

$$p(x) = C e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$$

где  $m$  и  $\sigma^2$  — параметры распределения, имеющие конечные значения,  $C$  — некоторая постоянная.

Произведем нормировку плотности вероятности. Имеем

$$\int_{-\infty}^{\infty} p(x) dx = C \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = 1.$$

Полагая  $\frac{x-m}{\sqrt{2}\sigma} = t$ , приведем это соотношение к виду

$$C\sigma\sqrt{2} \int_{-\infty}^{\infty} e^{-t^2} dt = 1.$$

Так как

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi} \quad (\text{интеграл Пуассона}),$$

находим

$$C = \frac{1}{\sqrt{2\pi}\sigma}.$$

Теперь можно записать распределение Гаусса в обычной форме:

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

Остается выяснить смысл параметров  $m$  и  $\sigma^2$ . Покажем, что параметр  $m$  имеет смысл математического ожидания. Действительно,

$$\begin{aligned} M[X] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-m) e^{-\frac{(x-m)^2}{2\sigma^2}} dx + m \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx. \end{aligned}$$

Первый из записанных интегралов равен нулю, поскольку заменой переменных  $\frac{x-m}{\sqrt{2}\sigma} = t$  сводится к интегралу, у которого под знаком интеграла стоит нечетная функция, а интегрирование осуществляется в симметричных пределах. Второй интеграл равен единице по условию нормировки распределения.

Таким образом,

$$M[X] = m.$$

Найдем дисперсию нормального распределения:

$$D[X] = \sigma_x^2 = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Сделав замену переменных  $t = \frac{x-m}{\sqrt{2\sigma}}$ , получим

$$\sigma_x^2 = \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} 2t^2 e^{-t^2} dt.$$

Применяя интегрирование по частям ( $t = u$ ,  $2te^{-t^2} dt = dv$ ), получим окончательно

$$\sigma_x^2 = \sigma^2.$$

Следовательно, параметр  $\sigma^2$  имеет смысл дисперсии случайной величины.

Как видно из выражения для плотности вероятности нормально распределенной случайной величины, нормальное распределение симметрично относительно прямой  $x = m$ , имеет максимум, равный  $\frac{1}{\sqrt{2\pi}\sigma} \approx \frac{0,4}{\sigma}$ , и монотонно убывает при возрастании  $|x - m|$ , асимптотически приближаясь к нулю, причем весьма быстро. Графики нормального распределения при разных значениях параметров приведены на рис. 3.6.

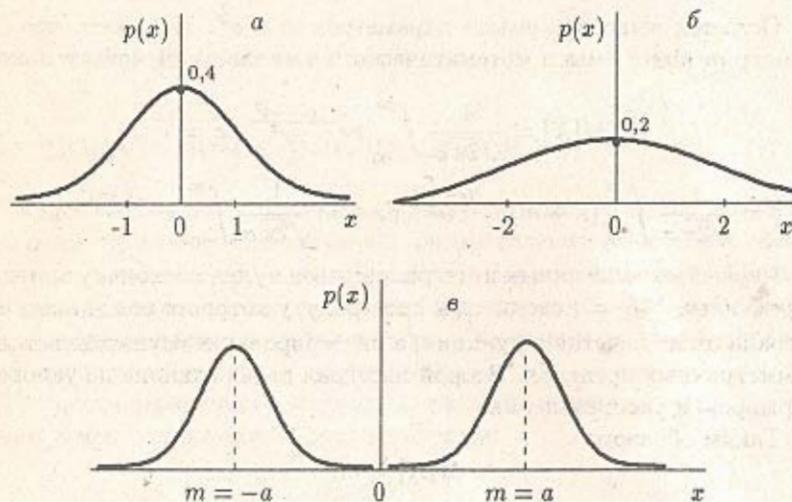


Рис. 3.6. Плотность распределения нормально распределенной случайной величины при различных значениях параметров:  
а —  $m = 0$ ,  $\sigma = 1$ ; б —  $m = 0$ ,  $\sigma = 2$ ; в —  $m = -a$  и  $m = a$

Нормальное распределение с нулевым математическим ожиданием и единичной дисперсией называют нормированным (стандартным) нормальным распределением. Его плотность вероятности имеет вид

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

Функция распределения нормально распределенной случайной величины  $X$  имеет вид

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(z-m)^2}{2\sigma^2}} dz.$$

Ее график приведен на рис. 3.7 для  $m = 0$  и  $\sigma^2 = 1$ .

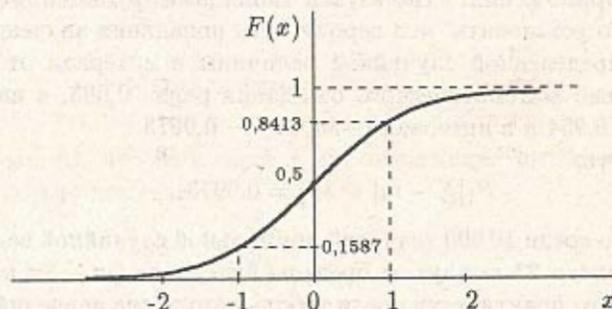


Рис. 3.7. Функция распределения нормально распределенной случайной величины

Найдем вероятность того, что случайная величина  $X$ , распределенная по нормальному закону, примет значение, лежащее в интервале  $[a, b]$ :

$$P(a \leq X < b) = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(x-m)^2}{2\sigma^2}} dx = F(b) - F(a).$$

Переходя к новой переменной  $t = \frac{x-m}{\sigma}$ , получим!

$$P(a \leq X < b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right),$$

где

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt$$

интеграл вероятности, представляющий собой функцию распределения нормированной нормально распределенной случайной величины.

Иногда интегралом вероятности называют функцию Лапласа

$$\Phi_1(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt.$$

Тогда

$$P(a \leq X < b) = \frac{1}{2} \left( \Phi_1 \left( \frac{b-m}{\sigma} \right) - \Phi_1 \left( \frac{a-m}{\sigma} \right) \right).$$

Интеграл вероятности табулирован. Его значения приведены в таблице 1 Приложения. Пользуясь таблицами функций  $\Phi(z)$  или  $\Phi_1(z)$ , можно установить, что вероятность попадания значений нормально распределенной случайной величины в интервал от  $-\sigma$  до  $\sigma$  относительно математического ожидания равна 0,683, в интервал  $(-2\sigma, 2\sigma)$  — 0,954 и в интервал  $(-3\sigma, 3\sigma)$  — 0,9973.

Из того, что

$$P(|X - m| < 3\sigma) = 0,9973,$$

вытекает, что среди 10 000 значений нормальной случайной величины в среднем только 27 выйдут за пределы интервала  $(m - 3\sigma, m + 3\sigma)$ . Это значит, что практически среди небольшого числа значений  $X$  нет таких, которые выходят за пределы указанного интервала. На этом основано "правило трех сигм", которое часто применяется в статистике.

Отметим, что среднеквадратическое отклонение  $\sigma$  в нормальном распределении не равно полуширине плотности вероятности на половине полувысоты. Полуширина на половине высоты равна  $1,176\sigma$ .

Найдем характеристическую функцию нормально распределенной случайной величины. Чтобы упростить вычисления, примем математическое ожидание равным нулю:

$$\theta(j\omega) = \int_{-\infty}^{\infty} e^{j\omega x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx.$$

Для вычисления интеграла в подынтегральной функции удобно показатель степени дополнить до полного квадрата разности:

$$j\omega x - \frac{x^2}{2\sigma^2} = - \left( \frac{x^2}{2\sigma^2} - j\omega x \right) =$$

$$= - \left( \left( \frac{x^2}{2\sigma^2} - j\omega x + d^2 \right) - d^2 \right) = - \left( \left( \frac{x}{\sqrt{2}\sigma} - d \right)^2 - d^2 \right),$$

где  $d$  определяется из условия

$$j\omega x = 2 \frac{x}{\sqrt{2}\sigma} d,$$

то есть  $d = \frac{j\omega\sigma}{\sqrt{2}}$ . Таким образом,

$$\theta(j\omega) = \frac{1}{\sqrt{2\pi}\sigma} e^{d^2} \int_{-\infty}^{\infty} e^{-\left(\frac{x}{\sqrt{2}\sigma} - d\right)^2} dx.$$

Переходя к новой переменной  $t = \frac{x}{\sqrt{2}\sigma} - d$ , получаем

$$\theta(j\omega) = \frac{1}{\sqrt{2\pi}\sigma} \sqrt{2}\sigma e^{d^2} \int_{-\infty}^{\infty} e^{-t^2} dt.$$

Учитывая то, что входящий в это выражение интеграл равен  $\sqrt{\pi}$ , окончательно получаем

$$\theta(j\omega) = e^{d^2} = e^{-\frac{\sigma^2}{2}\omega^2}.$$

Если математическое ожидание отлично от нуля, то

$$\theta(j\omega) = e^{(j\omega m - \frac{\sigma^2}{2}\omega^2)}.$$

Кроме рассмотренных распределений, при теоретическом и экспериментальном изучении случайных величин часто встречаются и другие распределения. Число возможных распределений очень велико, так как любую неотрицательную функцию, удовлетворяющую условию нормировки, можно рассматривать как плотность вероятности некоторой случайной величины.

## Глава 4

## Системы случайных величин

На практике часто приходится иметь дело с совокупностью (системой) двух или большего числа случайных величин. Систему  $n$  таких случайных величин можно рассматривать как точку в  $n$ -мерном евклидовом пространстве со случайными координатами  $X_1, X_2, \dots, X_n$ . Поэтому такую систему называют  $n$ -мерной случайной величиной или  $n$ -мерным случайным вектором

$$X = (X_1, X_2, \dots, X_n).$$

При  $n = 2$  двумерная случайная величина может рассматриваться как случайная точка на плоскости, а при  $n = 3$  — как случайная точка в пространстве. Такая трактовка совокупности двух или трех случайных величин дает возможность пользоваться наглядными геометрическими представлениями.

Системы случайных величин могут быть дискретными, непрерывными и смешанными в зависимости от типа случайных величин, образующих систему.

## 4.1. Законы распределения двумерных случайных величин

Для простоты изложения материала рассмотрим сначала *двумерные* случайные величины. Как и в случае одномерной случайной величины, полной вероятностной характеристикой двумерной случайной величины является закон распределения — соотношение, устанавливающее связь между областями возможных значений двумерной случайной величины и вероятностями ее появления в этих областях.

Закон распределения может быть задан в различных формах. Например, если  $X$  и  $Y$  — дискретные случайные величины, то закон распределения двумерной случайной величины  $(X, Y)$  в табличной форме имеет вид, представленный на рис. 4.1.

Y	X					
	$x_1$	$x_2$	...	$x_i$	...	$x_n$
$y_1$	$p_{11}$	$p_{21}$	...	$p_{i1}$	...	$p_{n1}$
$y_2$	$p_{12}$	$p_{22}$	...	$p_{i2}$	...	$p_{n2}$
...	...	...	...	...	...	...
$y_j$	$p_{1j}$	$p_{2j}$	...	$p_{ij}$	...	$p_{nj}$
...	...	...	...	...	...	...
$y_m$	$p_{1m}$	$p_{2m}$	...	$p_{im}$	...	$p_{nm}$

Рис. 4.1. Табличная форма закона распределения двумерной дискретной случайной величины

Здесь  $p_{ij} = P(X = x_i, Y = y_j)$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$  — вероятность того, что случайная величина  $X$  примет значение  $x_i$ , тогда как случайная величина  $Y$  примет значение  $y_j$ . При этом, поскольку события  $(X = x_i, Y = y_j)$  образуют полную группу, то сумма вероятностей, расположенных во всех клетках таблицы, равна единице.

$$\sum_{i,j} p_{ij} = 1,$$

Зная закон распределения двумерной дискретной случайной величины, можно найти закон распределения каждой из одномерных случайных величин:

$$P(X = x_i) = p_x(x_i) = p_i = \sum_{j=1}^m p_{ij};$$

$$P(Y = y_j) = p_y(y_j) = p_j = \sum_{i=1}^n p_{ij}.$$

Графическая форма задания закона распределения приведена на рис. 4.2.

Универсальной характеристикой многомерных случайных величин, пригодной для описания как дискретных, так и непрерывных

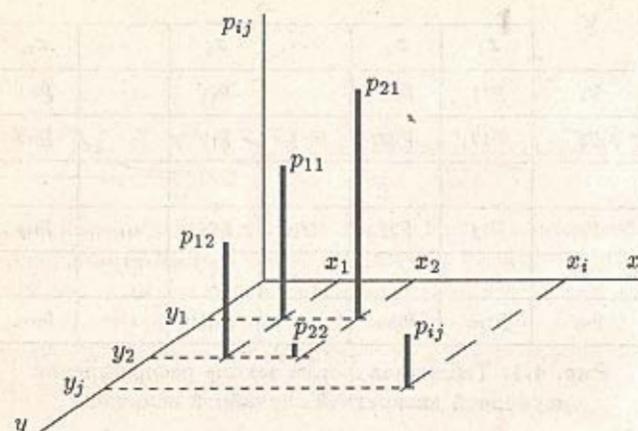


Рис. 4.2. Графическая форма закона распределения двумерной дискретной случайной величины

случайных величин, является функция распределения. В случае двумерной случайной величины функция распределения  $F_2(x, y)$  есть вероятность одновременного выполнения двух неравенств  $X < x$  и  $Y < y$ , рассматриваемая как функция переменных  $x$  и  $y$ . Индекс 2 указывает на то, что функция распределения двумерная. Часто этот индекс опускается.

В геометрической интерпретации функцию распределения  $F_2(x, y)$  можно трактовать как вероятность попадания случайной точки внутрь бесконечного левого нижнего квадранта с вершиной  $(x, y)$  (рис. 4.3).

Функция  $F_2(x, y)$  обладает следующими свойствами:

1.  $F_2(x, y)$  — неубывающая функция обоих своих аргументов, то есть

$$F_2(x_2, y) \geq F_2(x_1, y), \text{ если } x_2 > x_1;$$

$$F_2(x, y_2) \geq F_2(x, y_1), \text{ если } y_2 > y_1.$$

2.  $F_2(x, -\infty) = F_2(-\infty, y) = F_2(-\infty, -\infty) = 0.$

3.  $F_2(\infty, \infty) = 1.$

4.  $F_2(x, \infty) = F_x(x)$ ,  $F_2(\infty, y) = F_y(y)$ , где  $F_x(x)$  и  $F_y(y)$  — соответственно функции распределения случайных величин  $X$  и  $Y$ .

Докажем некоторые из приведенных свойств. Так  $F_2(x, -\infty)$  есть вероятность события  $X < x$  и  $Y < -\infty$ . Но такое событие невозмож-

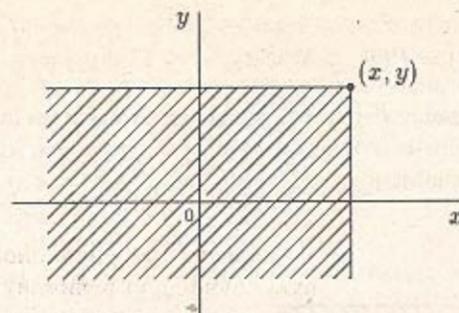


Рис. 4.3. Геометрическая интерпретация двумерной функции распределения  $F_2(x, y)$

но, поскольку невозможно событие  $Y = -\infty$ , и его вероятность равна нулю. Так как событие  $Y < \infty$  достоверно, то  $F_2(x, \infty)$  определяет вероятность события  $X < x$ , то есть представляет собой функцию распределения случайной величины  $X$ .

Если функция распределения  $F_2(x, y)$  непрерывна и имеет непрерывную смешанную производную второго порядка, то двумерная плотность вероятности системы двух случайных величин  $X$  и  $Y$  определяется формулой

$$p_2(x, y) = \frac{\partial^2 F_2(x, y)}{\partial x \partial y}.$$

Вместо плотности вероятности можно использовать двумерную характеристическую функцию

$$\theta_2(j\omega_1, j\omega_2) = M [e^{j(\omega_1 X + \omega_2 Y)}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(\omega_1 x + \omega_2 y)} p_2(x, y) dx dy.$$

Плотность вероятности  $p_2(x, y)$  обладает следующими свойствами:

1.  $p_2(x, y) \geq 0.$

2.  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_2(x, y) dx dy = 1$  (условие нормировки).

3.  $F_2(x, y) = \int_{-\infty}^x \int_{-\infty}^y p_2(u, v) du dv.$

4. Вероятность  $P(R)$  попадания случайной точки  $(X, Y)$  в прямоугольник  $R$  со сторонами, параллельными осям  $Ox$  и  $Oy$ , и с координатами вершин  $(x_1, y_1)$ ,  $(x_2, y_1)$ ,  $(x_2, y_2)$ ,  $(x_1, y_2)$ , равна (рис. 4.4):

$$P(R) = P(x_1 \leq X < x_2, y_1 \leq Y < y_2) = \\ = F_2(x_2, y_2) - F_2(x_1, y_2) - F_2(x_2, y_1) + F_2(x_1, y_1).$$

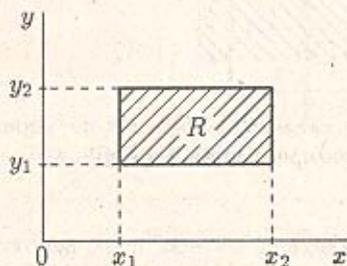


Рис. 4.4. К определению вероятности попадания случайной точки  $(X, Y)$  в прямоугольник  $R$

5. Вероятность  $P(D)$  попадания случайной точки  $(X, Y)$  в область  $D$  равна двойному интегралу по области  $D$  от плотности  $p_2(x, y)$ :

$$P(D) = \iint_{(D)} p_2(x, y) dx dy.$$

В геометрической интерпретации это можно трактовать следующим образом: вероятность того, что точка попадет в заданную область  $D$  плоскости  $xOy$  равна объему цилиндра под поверхностью  $p_2(x, y)$ , основанием которого служит заданная область  $D$ .

Одномерные функции распределения и плотности вероятности выражаются через двумерные с помощью следующих выражений:

$$F_x(x) = F_2(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{\infty} p_2(u, v) du dv;$$

$$p_x(x) = \frac{\partial F_x(x)}{\partial x} = \int_{-\infty}^{\infty} p_2(x, y) dy;$$

$$F_y(y) = F_2(\infty, y) = \int_{-\infty}^y \int_{-\infty}^{\infty} p_2(u, v) du dv;$$

$$p_y(y) = \frac{\partial F_y(y)}{\partial y} = \int_{-\infty}^{\infty} p_2(x, y) dx.$$

Плотности вероятности многомерных случайных величин удовлетворяют условию согласованности, в соответствии с которым из  $n$ -мерной плотности вероятности можно получить  $m$ -мерную ( $m < n$ ) путем интегрирования по "лишним" случайным величинам

$$p_m(x_1, \dots, x_m) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_n(x_1, \dots, x_m, x_{m+1}, \dots, x_n) dx_{m+1} \dots dx_n.$$

Рассмотрим пример. Пусть двумерная плотность вероятности двух случайных величин  $X$  и  $Y$  описывается выражением

$$p_2(x, y) = \begin{cases} \frac{1}{b-a} \cdot \frac{1}{d-c}, & a \leq x \leq b, \\ & c \leq y \leq d, \\ 0, & x < a, x > b, \\ & y < c, y > d. \end{cases}$$

Вид этой плотности вероятности показан на рис. 4.5.

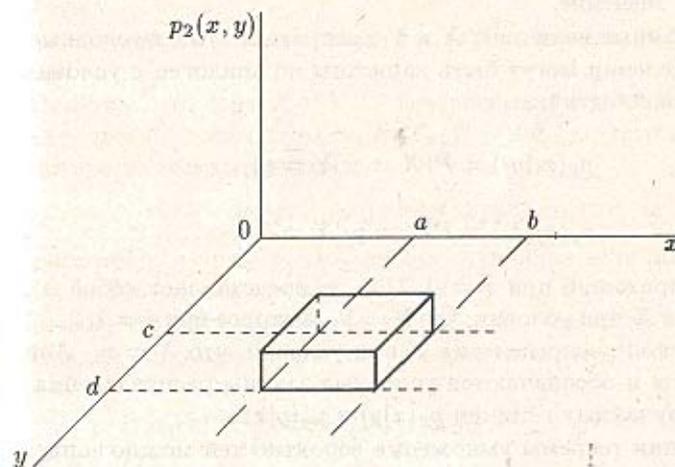


Рис. 4.5. Двумерная равномерная плотность вероятности

Тогда

$$F_x(x) = \int_a^x \int_c^d \frac{1}{b-a} \cdot \frac{1}{d-c} du dv = \int_a^x \frac{1}{b-a} du = \frac{x-a}{b-a};$$

$$p_x(x) = \int_c^d \frac{1}{b-a} \cdot \frac{1}{d-c} dy = \frac{1}{b-a}, \quad a \leq x \leq b;$$

$$F_y(y) = \int_c^y \int_a^b \frac{1}{b-a} \cdot \frac{1}{d-c} du dv = \int_c^y \frac{1}{d-c} dv = \frac{y-c}{d-c};$$

$$p_y(y) = \int_a^b \frac{1}{b-a} \cdot \frac{1}{d-c} dx = \frac{1}{d-c}, \quad c \leq y \leq d.$$

## 4.2. Условные распределения двух случайных величин

Закон распределения системы случайных величин  $X$  и  $Y$  определяется распределением каждой из величин, входящих в систему, и зависимостью между ними. Степень зависимости случайных величин  $X$  и  $Y$  характеризуется *условным законом распределения*, под которым понимается закон распределения одной случайной величины, найденный при условии, что другая случайная величина приняла определенное значение.

Если случайные величины  $X$  и  $Y$  дискретные, то их условные законы распределения могут быть записаны по аналогии с условными вероятностями событий:

$$p_x(x_i|y_j) = P(X = x_i|Y = y_j);$$

$$p_y(y_j|x_i) = P(Y = y_j|X = x_i).$$

Первое выражение при  $i = 1, 2, \dots, n$  представляет собой закон распределения  $X$  при условии, что  $Y = y_j$ , а второе при  $j = 1, 2, \dots, m$  определяет закон распределения  $Y$  при условии, что  $X = x_i$ . Аналогично вводятся и обозначаются условные законы распределения непрерывных случайных величин  $p_x(x|y)$  и  $p_y(y|x)$ .

На основании теоремы умножения вероятностей можно записать:

$$p_x(x_i|y_j) = \frac{P(X = x_i, Y = y_j)}{p_y(y_j)} = \frac{p_{ij}}{p_j};$$

$$p_y(y_j|x_i) = \frac{P(X = x_i, Y = y_j)}{p_x(x_i)} = \frac{p_{ij}}{p_i}.$$

Если вероятность события, состоящего в том, что случайная величина  $X$  примет значение  $x_i$ , не зависит от того, какое значение приняла случайная величина  $Y$ , то

$$p_x(x_i|y_j) = p_x(x_i) = p_i.$$

При аналогичных условиях для случайной величины  $Y$

$$p_y(y_j|x_i) = p_y(y_j) = p_j.$$

Тогда можно записать:

$$p_{ij} = p_i p_j.$$

Это равенство, если оно выполняется для всех  $i$  и  $j$ , является необходимым и достаточным условием независимости двух дискретных случайных величин.

В случае непрерывных случайных величин условие их независимости имеет вид:

$$p_2(x, y) = p_x(x)p_y(y),$$

то есть двумерная плотность вероятности должна быть представима в виде произведения одномерных плотностей вероятностей случайных величин.

Указанное условие обобщается и на большее число случайных величин.

Отметим, что если  $X$  и  $Y$  — независимые случайные величины с характеристическими функциями  $\theta_x(j\omega_1)$  и  $\theta_y(j\omega_2)$ , то их совместная характеристическая функция

$$\theta_2(j\omega_1, j\omega_2) = \theta_x(j\omega_1)\theta_y(j\omega_2).$$

*Рассмотрим пример.* Двумерная случайная величина  $(X, Y)$  имеет плотность вероятности  $p(x, y) = x + y$  для  $0 \leq x \leq 1$  и  $0 \leq y \leq 1$ . Покажем, что случайные величины  $X$  и  $Y$  зависимы.

Найдем одномерные плотности вероятности:

$$p_x(x) = \int_0^1 (x+y)dy = xy|_0^1 + \frac{y^2}{2}|_0^1 = x + \frac{1}{2};$$

$$p_y(y) = \int_0^1 (x+y)dx = y + \frac{1}{2}.$$

Отсюда вытекает очевидное неравенство

$$(x + \frac{1}{2})(y + \frac{1}{2}) \neq x + y.$$

### 4.3. Числовые характеристики двумерных законов распределения

Закон распределения является исчерпывающей характеристикой двумерной случайной величины. Однако если система включает в себя более двух-трех случайных величин, то экспериментальное определение ее закона распределения весьма затруднено, а проведение расчетов требует громоздких вычислений. В связи с этим при исследовании систем случайных величин широкое распространение нашли их числовые характеристики, которые в определенной степени могут дать представление о характере закона распределения. В основу получения таких числовых характеристик положено понятие *моментов* (начальных и центральных).

Различные моменты соответственно дискретных и непрерывных двумерных случайных величин определяются следующими формулами:

1) начальный момент  $m_{\nu_1\nu_2}$  порядка  $\nu_1 + \nu_2$ :

$$m_{\nu_1\nu_2} = M[X^{\nu_1}Y^{\nu_2}] = \sum_i \sum_j x_i^{\nu_1} y_j^{\nu_2} p_{ij};$$

$$m_{\nu_1\nu_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^{\nu_1} y^{\nu_2} p_2(x, y) dx dy;$$

2) центральный момент  $\mu_{\nu_1\nu_2}$  порядка  $\nu_1 + \nu_2$ :

$$\mu_{\nu_1\nu_2} = M[(X - m_x)^{\nu_1}(Y - m_y)^{\nu_2}] = \sum_i \sum_j (x_i - m_x)^{\nu_1} (y_j - m_y)^{\nu_2} p_{ij};$$

$$\mu_{\nu_1\nu_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)^{\nu_1} (y - m_y)^{\nu_2} p_2(x, y) dx dy.$$

Из начальных моментов на практике наиболее часто используются начальные моменты первого порядка:

$$m_{10} = M[X^1Y^0] = M[X] = m_x;$$

$$m_{01} = M[X^0Y^1] = M[Y] = m_y,$$

которые являются математическими ожиданиями случайных величин  $X$  и  $Y$ , входящих в систему, и которые определяют координаты точки, называемой центром рассеяния значений  $X$  и  $Y$ .

Из центральных моментов наиболее употребительны моменты второго порядка. Два из них представляют собой дисперсии случайных величин  $X$  и  $Y$ :

$$D_x = \mu_{20} = M[(X - m_x)^2(Y - m_y)^0] = M[(X - m_x)^2];$$

$$D_y = \mu_{02} = M[(X - m_x)^0(Y - m_y)^2] = M[(Y - m_y)^2].$$

$D_x$  и  $D_y$  характеризуют рассеяние случайной точки в направлении осей  $Ox$  и  $Oy$ .

### 4.4. Корреляция

Среди смешанных моментов особую роль играет центральный смешанный момент второго порядка

$$\mu_{11} = R_{xy} = M[(X - m_x)(Y - m_y)],$$

который называется *корреляционным моментом*.

Для дискретных случайных величин

$$R_{xy} = \sum_i \sum_j (x_i - m_x)(y_j - m_y) p_{ij},$$

а для непрерывных

$$R_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) p_2(x, y) dx dy.$$

Из определения корреляционного момента следует, что он имеет размерность, равную произведению размерностей величин  $X$  и  $Y$ .

Помимо рассеивания величин  $X$  и  $Y$  корреляционный момент  $R_{xy}$  характеризует вероятностную линейную зависимость между ними.

Часто вместо  $R_{xy}$  пользуются безразмерной величиной — *коэффициентом корреляции*

$$r_{xy} = \frac{R_{xy}}{\sqrt{D_x D_y}}.$$

Коэффициент корреляции  $r_{xy}$  удовлетворяет условию

$$-1 \leq r_{xy} \leq 1$$

и определяет степень линейной вероятностной зависимости между случайными величинами. Подчеркнем, что  $r_{xy}$  характеризует именно линейную зависимость, проявляющуюся в том, что при возрастании одной случайной величины другая имеет тенденцию возрастать (или убывать).

Две случайные величины, для которых коэффициент корреляции равен нулю, называют *некоррелированными*. В противном случае эти величины называются *коррелированными*. При  $r_{xy} = -1$  или  $r_{xy} = 1$  случайные величины связаны между собой линейной зависимостью.

Независимые случайные величины всегда не коррелированы. Зависимые случайные величины могут быть как некоррелированными, так и коррелированными. Пусть, например, случайные величины  $X$  и  $Y$  связаны между собой жесткой функциональной зависимостью, то есть  $Y = CX^k$ , где  $C$  — постоянный множитель,  $X$  — случайная величина с нулевым математическим ожиданием и симметричной относительно оси ординат плотностью вероятности  $p(x)$ . Тогда при любом четном  $k$  корреляционный момент равен нулю:

$$R_{xy} = M[X(Y - m_y)] = CM[X^{k+1}] = C \int_{-\infty}^{\infty} x^{k+1} p(x) dx = 0,$$

хотя случайные величины и связаны жесткой зависимостью. В данном случае между случайными величинами отсутствует *линейная* зависимость.

При помощи корреляционного момента можно определить свойства математического ожидания произведения и дисперсии суммы двух зависимых случайных величин:

$$\begin{aligned} M[XY] &= M[(X - m_x + m_x)(Y - m_y + m_y)] = \\ &= M[(X - m_x)(Y - m_y) + m_x(Y - m_y) + (X - m_x)m_y + m_x m_y] = \\ &= R_{xy} + m_x m_y; \end{aligned}$$

$$\begin{aligned} D[X + Y] &= M[(X - m_x + Y - m_y)^2] = \\ &= M[(X - m_x)^2] + 2M[(X - m_x)(Y - m_y)] + M[(Y - m_y)^2] = \\ &= D_x + D_y + 2R_{xy}. \end{aligned}$$

Аналогично можно показать, что

$$D[X - Y] = D_x + D_y - 2R_{xy}.$$

В некоторых случаях используются *условные моменты* случайной величины  $X$  относительно  $Y$ .

Для условных математического ожидания и дисперсии случайной величины  $X$  относительно  $Y$  формулы соответственно имеют вид

$$M[X|y] = \int_{-\infty}^{\infty} x p_x(x|y) dx = \frac{\int_{-\infty}^{\infty} x p_2(x, y) dx}{\int_{-\infty}^{\infty} p_2(x, y) dx};$$

$$D[X|y] = \int_{-\infty}^{\infty} (x - M[X|y])^2 p_x(x|y) dx = \frac{\int_{-\infty}^{\infty} (x - M[X|y])^2 p_2(x, y) dx}{\int_{-\infty}^{\infty} p_2(x, y) dx}.$$

#### 4.5. Числовые характеристики $n$ -мерных случайных величин

Основными числовыми характеристиками системы  $n$  случайных величин являются математические ожидания и дисперсии:

$$m_{x_k} = M[X_k];$$

$$D_{x_k} = D[X_k] = M[\{X_k - M[X_k]\}^2], \quad k = 1, 2, \dots, n,$$

а также корреляционные моменты или коэффициенты корреляции:

$$R_{x_i, x_j} = M[(X_i - M[X_i])(X_j - M[X_j])], \quad i \neq j;$$

$$r_{x_i, x_j} = \frac{R_{x_i, x_j}}{\sqrt{D_{x_i} D_{x_j}}}, \quad i \neq j.$$

Для удобства корреляционные моменты и коэффициенты корреляции записываются в виде *корреляционной матрицы* или нормированной корреляционной матрицы (матрицы, составленной из элементов  $r_{x_i, x_j}$ ). Например,

$$[R_{x_i, x_j}] = \begin{bmatrix} R_{x_1 x_1} & R_{x_1 x_2} & \dots & R_{x_1 x_n} \\ R_{x_2 x_1} & R_{x_2 x_2} & \dots & R_{x_2 x_n} \\ \dots & \dots & \dots & \dots \\ R_{x_n x_1} & R_{x_n x_2} & \dots & R_{x_n x_n} \end{bmatrix},$$

где  $R_{x_i, x_j}$  определяет корреляцию между переменными  $X_i$  и  $X_j$ .

Из определения корреляционного момента ясно, что  $R_{x_i x_j} = R_{x_j x_i}$ , поэтому часто заполняют только верхнюю или нижнюю половины матрицы, считая от ее главной диагонали, то есть

$$[R_{x_i x_j}] = \begin{bmatrix} R_{x_1 x_1} & R_{x_1 x_2} & \dots & R_{x_1 x_n} \\ & R_{x_2 x_2} & \dots & R_{x_2 x_n} \\ \dots & \dots & \dots & \dots \\ & & & R_{x_n x_n} \end{bmatrix}.$$

По главной диагонали корреляционной матрицы стоят дисперсии случайных величин  $D_{x_1} = R_{x_1 x_1}, \dots, D_{x_n} = R_{x_n x_n}$ .

В тех случаях, когда случайные величины попарно некоррелированы, все элементы корреляционной матрицы, кроме диагональных, равны нулю:

$$[R_{x_i x_j}] = \begin{bmatrix} D_{x_1} & 0 & \dots & 0 \\ & D_{x_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ & & & D_{x_n} \end{bmatrix}.$$

Такая матрица называется диагональной.

Начальный смешанный момент второго порядка называется ковариационным моментом:

$$m_{11} = K_{x_i x_j} = M[X_i X_j].$$

Отметим, что в иностранной литературе приняты другие определения для корреляционного и ковариационного моментов: корреляционный момент определяется как начальный смешанный момент второго порядка, а ковариационный момент — как центральный смешанный момент второго порядка.

## 4.6. Функциональные преобразования двумерных плотностей вероятностей

Законы распределения функционально преобразованных случайных величин в двумерном случае строятся по аналогии с одномерным. Пусть известен двумерный закон распределения случайных величин  $X_1$  и  $X_2$ , то есть известна совместная плотность вероятности

$p_{x_1 x_2}(x_1, x_2)$ . Известны также детерминированные связи:

$$Y_1 = g_1(X_1, X_2) \text{ и } Y_2 = g_2(X_1, X_2).$$

Будем считать, что обратные функции

$$X_1 = h_1(Y_1, Y_2) \text{ и } X_2 = h_2(Y_1, Y_2)$$

являются однозначными.

Плотность вероятности  $p_{y_1 y_2}(y_1, y_2)$  строится аналогично одномерному случаю:

$$p_{y_1 y_2}(y_1, y_2) = p_{x_1 x_2}(h_1(y_1, y_2), h_2(y_1, y_2)) |J(x_1, x_2)|,$$

где  $J(x_1, x_2)$  — якобиан преобразования от случайных величин  $X_1, X_2$  к величинам  $Y_1, Y_2$ :

$$J(x_1, x_2) = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_1(y_1, y_2)}{\partial y_1} & \frac{\partial h_1(y_1, y_2)}{\partial y_2} \\ \frac{\partial h_2(y_1, y_2)}{\partial y_1} & \frac{\partial h_2(y_1, y_2)}{\partial y_2} \end{vmatrix}.$$

В тех случаях, когда обратные функции не однозначны, следует взять сумму по всем ветвям обратной функции. Так, если обратные функции двузначные, то есть

$$X_{11} = h_{11}(Y_1, Y_2), \quad X_{12} = h_{12}(Y_1, Y_2),$$

$$X_{21} = h_{21}(Y_1, Y_2), \quad X_{22} = h_{22}(Y_1, Y_2),$$

то

$$p_{y_1 y_2}(y_1, y_2) = p_{x_1 x_2}(h_{11}(y_1, y_2), h_{21}(y_1, y_2)) \left| \frac{\partial(x_{11}, x_{21})}{\partial(y_1, y_2)} \right| + \\ + p_{x_1 x_2}(h_{12}(y_1, y_2), h_{22}(y_1, y_2)) \left| \frac{\partial(x_{12}, x_{22})}{\partial(y_1, y_2)} \right|.$$

На основании вышеизложенного можно получить выражения для плотностей вероятностей, соответствующих сумме, разности, произведению и частному двух случайных величин.

Предположим, что имеют место следующие детерминированные связи:

$$Y_1 = X_1, \quad Y_2 = g_2(X_1, X_2).$$

Соответствующие им обратные функции однозначны:

$$h_1(Y_1, Y_2) = X_1 = Y_1, \quad X_2 = h_2(Y_1, Y_2).$$

Найдем якобиан преобразования:

$$J(x_1, x_2) = \begin{vmatrix} 1 & 0 \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix} = \frac{\partial h_2}{\partial y_2}.$$

Поэтому

$$p_{y_1 y_2}(y_1, y_2) = p_{x_1 x_2}(y_1, h_2(y_1, y_2)) \left| \frac{\partial h_2}{\partial y_2} \right|.$$

Из этой формулы получаем одномерную плотность вероятности для одной случайной величины  $Y_2$ :

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1 x_2}(y_1, h_2(y_1, y_2)) \left| \frac{\partial h_2}{\partial y_2} \right| dy_1.$$

1. Пусть  $Y_2 = g_2(X_1, X_2) = X_1 + X_2$ . Тогда  $h_2 = X_2 = Y_2 - X_1$ ,  $\frac{\partial h_2}{\partial y_2} = 1$  и

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1 x_2}(x_1, y_2 - x_1) dx_1.$$

2. Пусть  $Y_2 = g_2(X_1, X_2) = X_2 - X_1$ . Тогда

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1 x_2}(x_1, y_2 + x_1) dx_1.$$

3. Пусть  $Y_2 = g_2(X_1, X_2) = X_1 X_2$ . Тогда  $h_2 = X_2 = Y_2 / X_1$ ,  $\frac{\partial h_2}{\partial y_2} = \frac{1}{X_1}$  и

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1 x_2}(x_1, y_2 / x_1) \frac{1}{|x_1|} dx_1.$$

4. Пусть  $Y_2 = g_2(X_1, X_2) = X_2 / X_1$ . Тогда  $h_2 = X_2 = Y_2 X_1$ ,  $\frac{\partial h_2}{\partial y_2} = X_1$  и

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1 x_2}(x_1, y_2 x_1) |x_1| dx_1.$$

Для независимых случайных величин  $X_1$  и  $X_2$  с плотностями вероятности  $p_{x_1}(x_1)$  и  $p_{x_2}(x_2)$  в предыдущих формулах нужно положить

$$p_{x_1 x_2}(x_1, x_2) = p_{x_1}(x_1) p_{x_2}(x_2).$$

При этом получим:

1) для  $Y_2 = X_1 + X_2$  2)  $p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1}(x_1) p_{x_2}(y_2 - x_1) dx_1$  (это интеграл свертки двух плотностей вероятностей);

2) для  $Y_2 = X_2 - X_1$

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1}(x_1) p_{x_2}(y_2 + x_1) dx_1;$$

3) для  $Y_2 = X_1 X_2$

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1}(x_1) p_{x_2}(y_2 / x_1) \frac{1}{|x_1|} dx_1;$$

4) для  $Y_2 = X_2 / X_1$

$$p_{y_2}(y_2) = \int_{-\infty}^{\infty} p_{x_1}(x_1) p_{x_2}(y_2 x_1) |x_1| dx_1.$$

Особое практическое значение имеет задача отыскания плотности вероятности суммы нескольких независимых случайных величин по известным плотностям вероятностей слагаемых, то есть так называемая композиция законов распределения. Эта задача может быть решена с помощью формулы для плотности вероятности суммы двух случайных величин путем последовательного вычисления интегралов свертки. Однако эта процедура достаточно сложна, поэтому при числе слагаемых больше двух пользуются аппаратом характеристических функций. Сначала находят характеристическую функцию  $\theta_n(j\omega)$  суммы  $n$  независимых случайных величин  $Y = X_1 + X_2 + \dots + X_n$ , которая равна произведению характеристических функций отдельных слагаемых

$$\theta_n(j\omega) = M[e^{j\omega Y}] = \prod_{i=1}^n \theta_{x_i}(j\omega),$$

а затем из обратного преобразования Фурье находят плотность вероятности

$$p_n(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega y} \theta_n(j\omega) d\omega.$$

Числовые характеристики функций  $Y_1 = g_1(X_1, X_2)$ ,  $Y_2 = g_2(X_1, X_2)$  системы двух случайных величин  $X_1$  и  $X_2$  можно найти, не производя предварительного определения плотности вероятности  $p_{y_1 y_2}(y_1, y_2)$ , а непосредственно используя совместную плотность вероятности  $p_{x_1 x_2}(x_1, x_2)$  этих величин.

В таком случае математические ожидания, дисперсии и корреляционный момент для непрерывных случайных величин определяются выражениями

$$m_{y_1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x_1, x_2) p_{x_1 x_2}(x_1, x_2) dx_1 dx_2;$$

$$m_{y_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_2(x_1, x_2) p_{x_1 x_2}(x_1, x_2) dx_1 dx_2;$$

$$D_{y_1} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (g_1(x_1, x_2) - m_{y_1})^2 p_{x_1 x_2}(x_1, x_2) dx_1 dx_2;$$

$$D_{y_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (g_2(x_1, x_2) - m_{y_2})^2 p_{x_1 x_2}(x_1, x_2) dx_1 dx_2;$$

$$R_{y_1 y_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (g_1(x_1, x_2) - m_{y_1}) \times \\ \times (g_2(x_1, x_2) - m_{y_2}) p_{x_1 x_2}(x_1, x_2) dx_1 dx_2.$$

## 4.7. Гамма-распределение

Гамма-распределение описывает распределение суммы независимых случайных величин, каждая из которых распределена по экспоненциальному закону с одинаковыми параметрами  $\lambda$ :

$$p_{x_i}(x_i) = \lambda e^{-\lambda x_i}, \quad x_i \geq 0.$$

Найдем распределение суммы двух независимых случайных величин  $X_1 + X_2 = Y$ :

$$p_y(y) = \int_{-\infty}^{\infty} p_{x_1}(x_1) p_{x_2}(y - x_1) dx_1 = \int_0^y \lambda e^{-\lambda x_1} \lambda e^{-\lambda(y-x_1)} dx_1 = \\ = \lambda^2 e^{-\lambda y} \int_0^y dx_1 = \lambda^2 y e^{-\lambda y}, \quad y \geq 0.$$

Получившееся распределение не является экспоненциальным, из чего можно заключить, что экспоненциальные законы не являются

устойчивыми по отношению к линейным комбинациям случайных величин, то есть к их композиции.

Пусть теперь имеется сумма трех случайных величин:

$$Z = X_1 + X_2 + X_3 = Y + X_3.$$

Тогда

$$p_z(z) = \int_0^z \lambda^2 y e^{-\lambda y} \lambda e^{-\lambda(z-y)} dy = \\ = \lambda^3 e^{-\lambda z} \int_0^z y dy = \lambda^3 \frac{z^2}{2} e^{-\lambda z}, \quad z \geq 0.$$

Если бы было четыре слагаемых, то получили бы

$$\lambda^4 \frac{u^3}{2 \cdot 3} e^{-\lambda u},$$

где  $U = X_1 + X_2 + X_3 + X_4$ . При пяти слагаемых получили бы

$$\lambda^5 \frac{v^4}{2 \cdot 3 \cdot 4} e^{-\lambda v}, \quad V = U + X_5.$$

Обозначим сумму  $k + 1$  случайных величин через  $X$ :

$$X = X_1 + X_2 + \dots + X_{k+1}.$$

Тогда для плотности вероятности  $p(x)$  получим

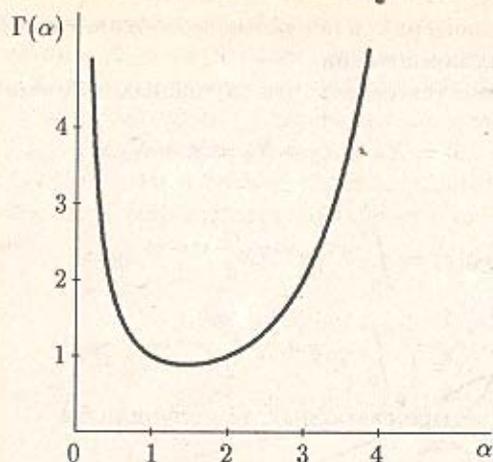
$$p(x) = \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x}, \quad \text{где } x \geq 0, \quad k > -1.$$

Если вместо  $\lambda$  использовать величину  $\frac{1}{\beta}$ , а вместо  $k$  — букву  $\alpha$ , то

$$p(x) = \frac{1}{\alpha! \beta^{\alpha+1}} x^\alpha e^{-\frac{x}{\beta}}, \quad \alpha > -1, \quad \beta > 0.$$

Это и есть гамма-распределение. Оно зависит от двух параметров —  $\alpha$  и  $\beta$ . Часто это распределение записывают в виде

$$p(x) = \frac{1}{\beta^{\alpha+1} \Gamma(\alpha+1)} x^\alpha e^{-\frac{x}{\beta}},$$

Рис. 4.6. Гамма-функция для  $\alpha > 0$ 

где  $\Gamma(\alpha)$  — гамма-функция Эйлера:

$$\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt, \quad \alpha > 0.$$

Гамма-функция для положительных  $\alpha$  имеет вид, показанный на рис. 4.6. Основные свойства гамма-функции:

- 1)  $\Gamma(1) = \Gamma(2) = 1$ ;
- 2)  $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha) = \alpha\Gamma(\alpha - 1 + 1) = \alpha(\alpha - 1)\Gamma(\alpha - 1) =$   
 $= \alpha(\alpha - 1)\Gamma(\alpha - 2 + 1) = \alpha(\alpha - 1)(\alpha - 2)\Gamma(\alpha - 2) = \dots = \alpha!$ ;

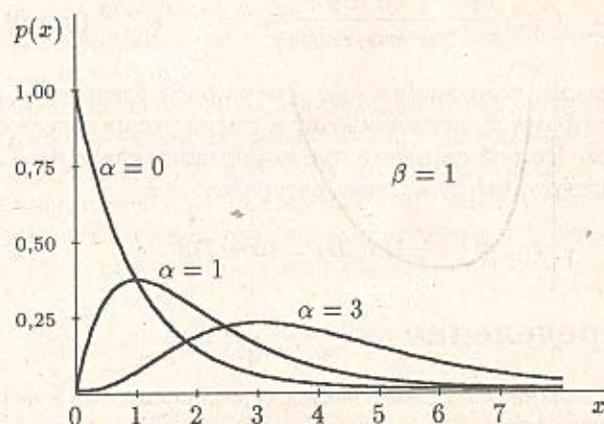
$$3) \Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} e^{-t} \frac{dt}{\sqrt{t}} = 2 \int_0^{\infty} e^{-u^2} du = \sqrt{\pi}.$$

Отсюда для любого целого  $\alpha > 0$

$$\Gamma\left(\alpha + \frac{1}{2}\right) = \left(\alpha - \frac{1}{2}\right) \Gamma\left(\alpha - \frac{1}{2}\right) = \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2\alpha - 1)}{2^\alpha} \sqrt{\pi}.$$

Таким образом, в гамма-распределении  $0 \leq x < \infty$ , а параметры  $\alpha$  и  $\beta$  имеют допустимые области  $\alpha > -1$ ,  $\beta > 0$ . Форма плотности вероятности  $p(x)$  определяется параметром  $\alpha$ , а параметр  $\beta$  носит характер масштабного множителя. При  $\alpha = 0$  гамма-распределение переходит в экспоненциальное.

Графики гамма-распределения при некоторых значениях параметра  $\alpha$  и  $\beta = 1$  приведены на рис. 4.7.

Рис. 4.7. Гамма-распределение при  $\beta = 1$  и  $\alpha = 0; 1; 3$ 

Композиция гамма-распределений снова дает гамма-распределение. Покажем это на примере двух случайных величин  $X_1$  и  $X_2$ , имеющих плотности вероятности с одинаковыми параметрами  $\beta$ :

$$p_{x_1}(x_1) = C_1 x_1^{\alpha-1} e^{-x_1/\beta};$$

$$p_{x_2}(x_2) = C_2 x_2^{\alpha-1} e^{-x_2/\beta},$$

где  $C_1$  и  $C_2$  — нормирующие множители распределений. Тогда для случайной величины  $X = X_1 + X_2$  плотность вероятности

$$\begin{aligned} p_x(x) &= \int_0^{x_1} p_{x_1}(x_1) p_{x_2}(x - x_1) dx_1 = \\ &= \int_0^{x_1} C_1 x_1^{\alpha-1} e^{-x_1/\beta} C_2 (x - x_1)^{\alpha-1} e^{-(x-x_1)/\beta} dx_1 = \\ &= C_1 C_2 e^{-x/\beta} \int_0^{x_1} x_1^{\alpha-1} (x - x_1)^{\alpha-1} dx_1. \end{aligned}$$

С помощью подстановки  $x_1 = xt$  преобразуем интеграл:

$$p_x(x) = C_1 C_2 e^{-x/\beta} \int_0^1 x^{\alpha-1} t^{\alpha-1} (x - xt)^{\alpha-1} x dt =$$

$$\begin{aligned}
 &= C_1 C_2 e^{-x/\beta} x^{\alpha_1 + \alpha_2 + 1} \int_0^1 t^{\alpha_1} (1-t)^{\alpha_2} dt = \\
 &= C_1 C_2 \frac{\Gamma(\alpha_1 + 1) \Gamma(\alpha_2 + 1)}{\Gamma(\alpha_1 + \alpha_2 + 2)} x^{\alpha_1 + \alpha_2 + 1} e^{-x/\beta} \quad (x > 0).
 \end{aligned}$$

Таким образом, композиция двух гамма-распределений с одинаковыми параметрами  $\beta$  снова приводит к гамма-распределению с тем же параметром. Второй параметр при композиции равен  $\alpha_1 + \alpha_2 + 1$ .

Можно показать, что для гамма-распределения

$$m_x = (\alpha + 1)\beta, \quad D_x = (\alpha + 1)\beta^2.$$

#### 4.8. Распределение $\chi^2$

Случайной величиной  $\chi_k^2$  называется сумма квадратов  $k$  независимых случайных величин  $X_1, X_2, \dots, X_k$ , каждая из которых распределена по нормальному закону с нулевым математическим ожиданием и единичной дисперсией. Число  $k$  называют *числом степеней свободы* случайной величины  $\chi^2$ .

Очевидно, что плотность вероятности случайной величины  $\chi^2$  зависит только от числа  $k$ , то есть числа слагаемых. Чтобы найти эту плотность, найдем сначала плотность вероятности величины  $Y = X^2$ , где  $X$  распределена по закону

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

Функция  $y = x^2$  является двузначной, поэтому существуют две обратные функции:

$$x_1 = \sqrt{y}, \quad x_2 = -\sqrt{y}.$$

Модули их производных одинаковы и равны

$$\left| \frac{dx_1(y)}{dy} \right| = \left| \frac{dx_2(y)}{dy} \right| = \frac{1}{2\sqrt{y}}.$$

Тогда плотность вероятности для случайной величины  $Y$

$$p_y(y) = p_x(\sqrt{y}) \frac{1}{2\sqrt{y}} + p_x(-\sqrt{y}) \frac{1}{2\sqrt{y}} =$$

$$= \frac{1}{\sqrt{2\pi}} \left( e^{-\frac{y}{2}} + e^{-\frac{y}{2}} \right) \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} y^{-\frac{1}{2}}, \quad y > 0.$$

Это распределение представляет собой гамма-распределение с параметрами  $\alpha = -\frac{1}{2}$  и  $\beta = 2$ .

Как было показано ранее, композиция гамма-распределений с одинаковыми значениями параметра  $\beta$  снова дает гамма-распределение с тем же значением  $\beta$ ; при этом второй параметр  $\alpha$  определяется суммой параметров  $\alpha$  распределений слагаемых. Следовательно, распределение  $\chi^2$  является гамма-распределением с параметрами  $\beta = 2$  и  $\alpha = k/2 - 1$ . Его плотность вероятности записывается в виде

$$p(\chi^2) = \frac{1}{2^{k/2} \Gamma(k/2)} (\chi^2)^{k/2-1} e^{-\chi^2/2}, \quad 0 < \chi^2 < \infty,$$

где  $k = 1, 2, \dots$  — целые положительные числа. Графики плотностей вероятностей для нескольких значений степеней свободы приведены на рис. 4.8.

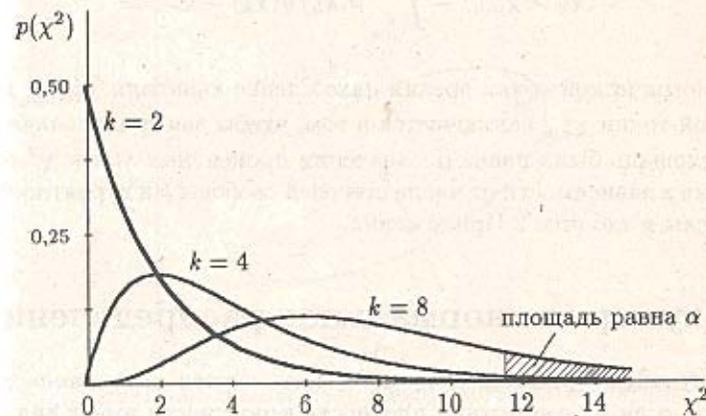


Рис. 4.8. Распределения  $\chi^2$

С увеличением числа степеней свободы  $k$  плотность вероятности стремится к нормальной. При  $k \leq 2$  плотность вероятности постоянно

убывает, а при  $k > 2$  имеет единственный максимум  $\chi^2 = k - 2$ . Математическое ожидание распределения  $\chi^2$  равно  $k$ , а дисперсия —  $2k$ .

На практике, как правило, используется не плотность вероятности, а *квантили* или *процентные точки* распределения  $\chi^2$ . Напомним, что квантилем порядка  $p$  закона распределения случайной величины  $X$  называется такое значение  $x_p$ , которое удовлетворяет равенству  $F(x_p) = p$ , где  $F(x)$  — функция распределения, а  $p$  — заданное значение вероятности. Часто в статистике задают не вероятность  $p$ , а вероятность  $\alpha = 1 - p$ . Тогда применительно к случайной величине  $\chi_k^2$  квантилем порядка  $1 - \alpha$  называется такое значение  $\chi_{1-\alpha;k}^2$ , при котором

$$F(\chi_{1-\alpha;k}^2) = P(\chi_k^2 < \chi_{1-\alpha;k}^2) = \int_0^{\chi_{1-\alpha;k}^2} p(\chi_k^2) d(\chi_k^2) = 1 - \alpha.$$

В свою очередь, процентная точка  $\chi_{\alpha;k}^2$  определяется как

$$P(\chi_k^2 > \chi_{\alpha;k}^2) = \int_{\chi_{\alpha;k}^2}^{\infty} p(\chi_k^2) d(\chi_k^2) = \alpha.$$

С геометрической точки зрения нахождение квантиля  $\chi_{1-\alpha;k}^2$  или процентной точки  $\chi_{\alpha;k}^2$  заключается в том, чтобы заштрихованная на рис. 4.8 площадь была равна  $\alpha$ . Значения процентных точек  $\chi^2$ -распределения в зависимости от числа степеней свободы  $k$  и вероятностей  $\alpha$  приведены в таблице 2 Приложения.

#### 4.9. Двумерное нормальное распределение

Две случайные величины  $X_1$  и  $X_2$  называются *совместно нормальными*, если их совместная плотность вероятности имеет вид

$$p_2(x_1, x_2) = \frac{1}{2\pi\sqrt{D_1 D_2 - R^2}} \times \exp\left(\frac{-D_2(x_1 - m_1)^2 + 2R(x_1 - m_1)(x_2 - m_2) - D_1(x_2 - m_2)^2}{2(D_1 D_2 - R^2)}\right) =$$

#### 4.9. Двумерное нормальное распределение

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \times \exp\left(\frac{-\sigma_2^2(x_1 - m_1)^2 + 2r\sigma_1\sigma_2(x_1 - m_1)(x_2 - m_2) - \sigma_1^2(x_2 - m_2)^2}{2\sigma_1^2\sigma_2^2(1-r^2)}\right),$$

где  $m_1$  и  $m_2$  — соответственно математические ожидания случайных величин  $X_1$  и  $X_2$ ,  $D_1 = \sigma_1^2$  и  $D_2 = \sigma_2^2$  — их дисперсии,  $R$  — корреляционный момент между случайными величинами  $X_1$  и  $X_2$ ,  $r$  — коэффициент корреляции.

Из приведенной плотности вероятности видно, что она имеет постоянное значение на так называемых эллипсах постоянной плотности (эллипсах рассеяния), которые определяются уравнением

$$\frac{(x_1 - m_1)^2}{\sigma_1^2} - 2r\frac{(x_1 - m_1)(x_2 - m_2)}{\sigma_1\sigma_2} + \frac{(x_2 - m_2)^2}{\sigma_2^2} = C^2,$$

где постоянная  $C > 0$  определяется значением плотности вероятности и параметрами распределения. Центр этого эллипса находится в точке с координатами  $m_1$  и  $m_2$  (рис. 4.9).

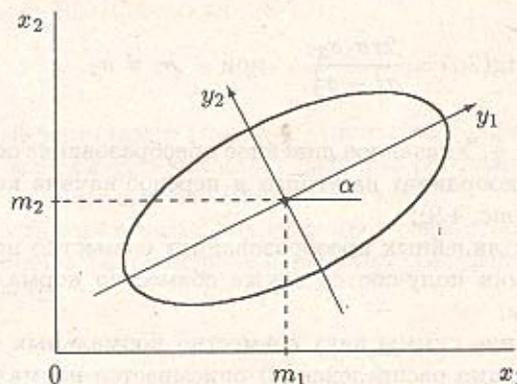


Рис. 4.9. Эллипс постоянной плотности вероятности

В центре эллипса плотность вероятности максимальна и равна

$$p_2(m_1, m_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}.$$

На примере двух случайных величин рассмотрим характерные свойства совместно нормальных случайных величин:

1) если случайные величины  $X_1$  и  $X_2$  не коррелированы, то есть  $r = 0$ , то их совместная плотность вероятности равна произведению плотностей вероятностей каждой из них:

$$p_2(x_1, x_2) = p_{x_1}(x_1)p_{x_2}(x_2).$$

Но такие величины по определению называются независимыми. Следовательно, если две совместно нормальные случайные величины не коррелированы, то они и независимы, то есть некоррелированность двух совместно нормальных случайных величин тождественна их независимости;

2) две коррелированные (зависимые) совместно нормальные случайные величины  $X_1$  и  $X_2$  всегда можно привести к двум некоррелированным (независимым) нормальным случайным величинам  $Y_1$  и  $Y_2$  при помощи линейного преобразования:

$$\begin{aligned} Y_1 &= (X_1 - m_1) \cos \alpha + (X_2 - m_2) \sin \alpha; \\ Y_2 &= -(X_1 - m_1) \sin \alpha + (X_2 - m_2) \cos \alpha, \end{aligned}$$

где угол  $\alpha$  определяется из условия  $M[Y_1 Y_2] = 0$  и может быть найден из равенства

$$\operatorname{tg}(2\alpha) = \frac{2r\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \quad \text{при} \quad \sigma_1 \neq \sigma_2.$$

При  $\sigma_1 = \sigma_2$   $\alpha = \frac{\pi}{4}$ . Указанное линейное преобразование осуществляет поворот осей координат на угол  $\alpha$  и перенос начала координат в точку  $(m_1, m_2)$  (рис. 4.9);

3) при любых линейных преобразованиях совместно нормальных случайных величин получаются также совместно нормальные случайные величины;

4) распределение суммы двух совместно нормальных случайных величин (композиция распределений) описывается нормальным распределением с параметрами

$$\begin{aligned} M[Y = X_1 + X_2] &= m_1 + m_2; \\ D[Y = X_1 + X_2] &= D_1 + D_2 + 2r\sqrt{D_1 D_2}; \end{aligned}$$

5) если две случайные величины являются совместно нормальными, то каждая из величин будет также нормальной. Обратное утверждение в общем случае неверно. Оно верно только для независимых случайных величин;

6) для совместно нормальных случайных величин условная плотность вероятности одной из них при фиксированном значении другой является нормальной. Так,

$$\begin{aligned} p(x_2|X_1 = x_1) &= \frac{p_2(x_1, x_2)}{p_{x_1}(x_1)} = \\ &= \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{1}{2\sigma_2^2(1-r^2)}(x_2 - m_2 - r\frac{\sigma_2}{\sigma_1}(x_1 - m_1))^2\right). \end{aligned}$$

Видно, что условная плотность вероятности является нормальной с математическим ожиданием и дисперсией, равными соответственно

$$\begin{aligned} M[X_2|X_1 = x_1] &= m_2 + r\frac{\sigma_2}{\sigma_1}(x_1 - m_1); \\ D[X_2|X_1 = x_1] &= \sigma_2^2(1-r^2). \end{aligned}$$

Условное математическое ожидание  $X_2$  при данном  $x_1$  зависит от  $x_1$ , а условная дисперсия от  $x_1$  не зависит.

Профиль условной плотности вероятности в плоскости  $x_1 = \text{const}$  описывается нормальной кривой с единственным максимумом, расположенным на прямой линии

$$x_2 = m_2 + r\frac{\sigma_2}{\sigma_1}(x_1 - m_1),$$

проходящей через центр  $(m_1, m_2)$  эллипса постоянной плотности.

Почти все приведенные выше свойства обобщаются на несколько совместно нормальных случайных величин.

Плотность вероятности для системы  $n$  совместно нормальных случайных величин (многомерное нормальное распределение) можно записать в виде

$$p_n(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sqrt{|R|}} \exp\left(\frac{-1}{2|R|} \sum_{i,j=1}^n A_{ij}(x_i - m_{x_i})(x_j - m_{x_j})\right),$$

где  $A_{ij}$  — алгебраическое дополнение элемента  $R_{ij}$  в определителе  $|R|$  корреляционной матрицы

$$R = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \dots & \dots & \dots & \dots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{bmatrix}.$$

Напомним, что если из определителя вычеркнуть  $i$ -ю строку и  $j$ -й столбец и умножить полученный минор на  $(-1)^{i+j}$ , то полученное произведение называется алгебраическим дополнением элемента  $R_{ij}$ .

Плотность вероятности системы совместно нормальных случайных величин наиболее компактно записывается в матричной форме. Определим вектор-столбцы

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{и} \quad \mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix}.$$

Тогда нормальная плотность вероятности случайных величин  $X_1, X_2, \dots, X_n$  примет вид

$$p_n(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\mathbf{R}|} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{R}^{-1}(\mathbf{x}-\mathbf{m})},$$

где  $\mathbf{R}^{-1}$  — матрица, обратная матрице  $\mathbf{R}$ ; символ  $T$  обозначает операцию транспонирования матрицы.

## Глава 5

# Закон больших чисел и предельные теоремы

Закон больших чисел составляет основу математической статистики и представляет собой совокупность теорем, связывающих между собой измеренные и истинные значения характеристик случайных величин. Этот закон позволяет найти пределы, к которым стремятся вероятностные количественные оценки случайных величин (например, среднее значение при увеличении числа измерений). Предельные теоремы дают возможность определить характер закона распределения суммы случайных величин при увеличении их числа.

Закон больших чисел составляют несколько положений, главными из которых применительно к независимым случайным величинам являются неравенство Чебышева, теорема Чебышева, теорема Бернулли, лемма Маркова и теорема Пуассона.

Из предельных теорем будет рассмотрена только центральная предельная теорема Ляпунова.

### 5.1. Неравенство Чебышева

Неравенство Чебышева записывается обычно в виде одной из следующих форм:

$$P(|X - m_x| < \varepsilon) \geq 1 - \frac{D_x}{\varepsilon^2}; \quad (5.1)$$

$$P(|X - m_x| \geq \varepsilon) \leq \frac{D_x}{\varepsilon^2}, \quad (5.2)$$

где  $\varepsilon > 0$  — произвольное число. Оно справедливо для любой случайной величины, имеющей конечную дисперсию.

Проведем доказательство неравенства Чебышева для непрерывной

случайной величины. Запишем выражение для дисперсии

$$D_x = \int_{-\infty}^{\infty} (x - m_x)^2 p(x) dx.$$

Разобьем область интегрирования на три части (рис. 5.1):

$$(-\infty, m_x - \varepsilon), (m_x - \varepsilon, m_x + \varepsilon), (m_x + \varepsilon, \infty).$$

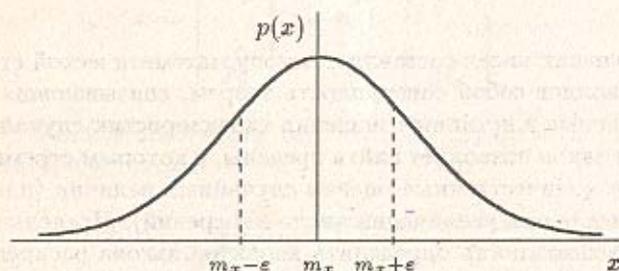


Рис. 5.1. К доказательству неравенства Чебышева

Тогда

$$D_x = \int_{-\infty}^{m_x - \varepsilon} (x - m_x)^2 p(x) dx + \int_{m_x - \varepsilon}^{m_x + \varepsilon} (x - m_x)^2 p(x) dx + \int_{m_x + \varepsilon}^{\infty} (x - m_x)^2 p(x) dx.$$

Опуская средний интеграл, приходим к неравенству:

$$D_x \geq \int_{-\infty}^{m_x - \varepsilon} (x - m_x)^2 p(x) dx + \int_{m_x + \varepsilon}^{\infty} (x - m_x)^2 p(x) dx.$$

В пределах каждой из этих областей интегрирования справедливо соотношение  $(x - m_x)^2 \geq \varepsilon^2$ . Значит,

$$\begin{aligned} D_x &\geq \varepsilon^2 \left( \int_{-\infty}^{m_x - \varepsilon} p(x) dx + \int_{m_x + \varepsilon}^{\infty} p(x) dx \right) = \\ &= \varepsilon^2 P(-\infty < X \leq m_x - \varepsilon, m_x + \varepsilon \leq X < \infty) = \varepsilon^2 P(|X - m_x| \geq \varepsilon). \end{aligned}$$

Отсюда следует

$$P(|X - m_x| \geq \varepsilon) \leq \frac{D_x}{\varepsilon^2}.$$

Неравенство Чебышева имеет прежде всего теоретическое значение. Однако в тех случаях, когда сведения о законе распределения незначительны, оно находит и практическое применение.

*Рассмотрим пример.* Оценим вероятность того, что отклонение какой-либо случайной величины от ее математического ожидания по модулю будет меньше трех среднеквадратических отклонений этой величины. Пусть  $X$  — случайная величина и  $D_x = \sigma_x^2$ . Тогда

$$P(|X - m_x| < 3\sigma_x) \geq 1 - \frac{\sigma_x^2}{(3\sigma_x)^2} = 1 - \frac{1}{9} = 0,8889.$$

В случае нормально распределенной случайной величины эта вероятность равна 0,9973, что не противоречит полученному результату.

По второй форме записи неравенство Чебышева дает верхнюю границу для вероятности того, что будет наблюдаться превышение некоторого заданного числа среднеквадратических отклонений независимо от вида закона распределения.

## 5.2. Теорема Чебышева

В соответствии с теоремой Чебышева при достаточно большом числе независимых испытаний  $n$  среднее арифметическое  $\bar{x}$  наблюдаемых значений  $x_1, x_2, \dots, x_n$  случайной величины  $X$  сходится по вероятности к ее математическому ожиданию  $m_x$ , то есть при любом  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\bar{x} - m_x| < \varepsilon) = 1.$$

Для доказательства теоремы обратим внимание на то, что наблюдаемые значения  $x_1, x_2, \dots, x_n$  можно рассматривать как значения независимых случайных величин  $X_1, X_2, \dots, X_n$  с одинаковыми математическими ожиданиями и дисперсиями:

$$M[X_i] = m_x; D[X_i] = D[X], \quad i = 1, \dots, n.$$

Для случайной величины среднего арифметического  $\bar{X}$ , используя свойства математического ожидания и дисперсии, получим:

$$M[\bar{X}] = M\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n M[X_i] = \frac{1}{n} n m_x = m_x,$$

$$D[\bar{X}] = D\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n^2} \sum_{i=1}^n D[X_i] = \frac{D[X]}{n}.$$

Применим к случайной величине  $\bar{X}$  неравенство Чебышева:

$$P(|\bar{X} - m_x| < \varepsilon) \geq 1 - \frac{D[X]}{\varepsilon^2 n}.$$

При  $n \rightarrow \infty \frac{D[X]}{\varepsilon^2 n} \rightarrow 0$  при любых  $\varepsilon > 0$ , что и доказывает теорему Чебышева.

Теорема Чебышева имеет большое практическое значение. Так, если нужно измерить некоторую величину, истинное значение которой равно  $a$ , проводят  $n$  измерений этой величины. Результат каждого измерения — это случайная величина  $X_i$ . Можно считать, что  $X_i$  — независимые случайные величины, а их дисперсии ограничены. Тогда среднее арифметическое значение результатов измерений с увеличением  $n$  приближается к истинному значению измеряемой величины  $a$ , поэтому можно положить  $\bar{X} \approx a$ .

Точность этого приближенного равенства можно характеризовать математическим ожиданием квадрата случайной ошибки  $\Delta$ :

$$M[\Delta^2] = M[(\bar{X} - a)^2] = D[\bar{X}] = \frac{D[X]}{n},$$

если точность всех измерений одна и та же.

Однако ошибочно думать, что, увеличивая число измерений, можно достигнуть сколь угодно большой точности. Дело в том, что сам измерительный прибор дает показания лишь с определенной точностью, поэтому каждый из результатов измерений, а следовательно, и их среднее арифметическое, будут получены лишь с точностью, не превосходящей точности прибора.

Выводы теоремы Чебышева можно распространить и на другие моменты распределения. Для дисперсии, например, получаем приближенную формулу, пригодную для практических вычислений:

$$D[X] = M[(X - a)^2] \approx \frac{\sum_{i=1}^n (x_i - a)^2}{n},$$

где вместо  $a$ , согласно теореме Чебышева, можно пользоваться  $X_{\text{ср}}$ :

$$a \approx \frac{1}{n} \sum_{i=1}^n X_i.$$

Неравенство Чебышева и теорема Чебышева для практических задач могут использоваться в тех случаях, когда по условию известна дисперсия случайной величины, причем эта дисперсия должна быть конечной величиной.

*Рассмотрим пример.* Применим теорему Чебышева для решения следующей задачи: сколько раз нужно измерять величину с истинным значением  $a$ , чтобы с вероятностью, не меньшей 0,95, можно было утверждать, что среднее арифметическое этих измерений отличается от  $a$  по абсолютной величине меньше, чем на 2, если среднеквадратическое отклонение измерений меньше 10?

В соответствии с теоремой нужно найти число  $n$ , которое удовлетворяет неравенству

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| < \varepsilon\right) \geq 1 - \frac{D[X]}{n\varepsilon^2},$$

где  $\varepsilon = 2$ ,  $D[X] = 100$ . Отсюда

$$1 - \frac{D[X]}{n\varepsilon^2} = 0,95, \quad n = \frac{D[X]}{\varepsilon^2(1 - 0,95)} = \frac{100}{4 \cdot 0,05} = 500.$$

Таким образом, измерять нужно не менее 500 раз.

### 5.3. Теорема Бернулли

В статистическом определении вероятности наступления события использовалось свойство устойчивости относительной частоты наступления этого события при увеличении числа испытаний. Теоретическое обоснование этого свойства дается теоремой Бернулли, в соответствии с которой при неограниченном увеличении числа  $n$  независимых испытаний, в каждом из которых интересующее нас событие имеет одну и ту же вероятность наступления  $p$ , вероятность отклонения относительной частоты  $\frac{m}{n}$  от вероятности  $p$  стремится к единице, как бы мало ни было заданное число  $\varepsilon$ :

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1.$$

Подчеркнем следующее: теорема Бернулли не позволяет утверждать, что неравенство

$$\left|\frac{m}{n} - p\right| < \varepsilon$$

будет выполняться для всех достаточно больших  $n$ . Она лишь утверждает, что выполнение такого неравенства при достаточно больших  $n$  будет очень вероятно.

## 5.4. Лемма Маркова

В соответствии с этой леммой, если случайная величина  $X$  принимает только положительные значения, то вероятность того, что она не превзойдет некоторой величины  $A$ , больше, чем  $(1 - \frac{m_x}{A})$ , то есть

$$P(X \leq A) > 1 - \frac{m_x}{A}.$$

Иногда лемму Маркова записывают в виде

$$P(X \leq km_x) > 1 - \frac{1}{k},$$

где  $km_x = A$ .

Последнее неравенство показывает, что вероятность того, что случайная величина не превзойдет своего математического ожидания в  $k$  раз, больше  $(1 - \frac{1}{k})$ .

В обоих неравенствах принимается, что  $k > 1$ , то есть  $m_x < A$ .

## 5.5. Теорема Пуассона

Пусть производится  $n$  независимых испытаний, причем вероятности появления события  $A$  в каждом испытании различны и равны  $P_1, P_2, \dots, P_n$ .

Можно ввести понятие средней вероятности

$$P_{\text{ср}} = \frac{P_1 + P_2 + \dots + P_n}{n}.$$

Теорема Пуассона утверждает, что

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{m}{n} - P_{\text{ср}} \right| < \epsilon \right) = 1,$$

где  $m$  — число испытаний, при которых имело место событие  $A$ .

## 5.6. Центральная предельная теорема Ляпунова

В соответствии с этой теоремой, плотность вероятности суммы независимых или слабо зависимых, равномерно малых (то есть играющих примерно одинаковую роль) слагаемых при неограниченном увеличении их числа как угодно близко приближается к нормальному закону распределения независимо от того, какие законы распределения имеют эти слагаемые. Докажем эту теорему для случая, когда законы распределения слагаемых одинаковы.

Пусть известно, что законы распределения  $n$  независимых случайных величин  $X_i$  одинаковы и определяются некоторой неопределенной функцией  $p_x(x)$ . Поставим задачу найти закон распределения суммы случайных величин  $p_y(y)$ , где  $Y = X_1 + X_2 + \dots + X_n$ , когда число слагаемых стремится к бесконечности.

Математическое ожидание и дисперсию каждого из распределений обозначим через  $a$  и  $\sigma^2$ . Тогда математическое ожидание и дисперсия суммы будут равны:

$$m_y = na, \quad \sigma_y^2 = n\sigma^2.$$

Будем искать закон распределения нормированной случайной величины

$$Z = \frac{Y - m_y}{\sigma_y} = \frac{Y - na}{\sqrt{n}\sigma}.$$

Характеристическая функция каждой случайной величины определяется формулой

$$\theta_x(j\omega) = M[e^{j\omega X}] = \int_{-\infty}^{\infty} e^{j\omega x} p_x(x) dx.$$

Характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций. В нашем случае

$$\theta_y(j\omega) = (\theta_x(j\omega))^n.$$

Логарифмическая характеристическая функция

$$\ln \theta_y(j\omega) = n \ln \theta_x(j\omega).$$

Разлагая экспоненту в выражении для  $\theta_x(j\omega)$  в ряд Маклорена и беря затем математическое ожидание от каждого слагаемого, получим

$$\theta_x(j\omega) = M \left[ \sum_{k=0}^{\infty} \frac{1}{k!} X^k (j\omega)^k \right] = 1 + \sum_{k=1}^{\infty} \frac{(j\omega)^k}{k!} m_k.$$

Если вместо случайной величины  $X$  рассматривать нормированную случайную величину  $Z = \frac{Y - m_y}{\sigma_y}$ , то в этом разложении вместо начальных моментов  $m_k$  нужно ввести центральные моменты  $\mu_k$  и, кроме того, вместо  $\omega$  писать  $\frac{\omega}{\sigma\sqrt{n}}$ . Учитывая это, для логарифмической характеристической функции получим выражение:

$$\ln \theta_z(j\omega) = n \ln \left( 1 + \sum_{k=1}^{\infty} j^k \frac{1}{k!} \mu_k \left( \frac{\omega}{\sigma\sqrt{n}} \right)^k \right).$$

Так как  $\mu_1 = 0$ , а  $\mu_2 = \sigma^2$ , можно записать

$$\ln \theta_z(j\omega) = n \ln \left( 1 - \frac{\omega^2}{2n} - j \frac{\mu_3 \omega^3}{6\sigma^3 n^{3/2}} + \frac{\mu_4 \omega^4}{24\sigma^4 n^2} + \dots \right).$$

При увеличении числа  $n$  роль последних слагаемых в скобках будет падать. Сохраняя первые два члена этого выражения и воспользовавшись при малых  $x$  приближенной формулой

$$\ln(1-x) \approx -x,$$

получим

$$\ln \theta_z(j\omega) = -\frac{\omega^2}{2}.$$

Следовательно,

$$\theta_z(j\omega) = e^{-\frac{\omega^2}{2}}.$$

Как известно, для нормального распределения в виде

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

характеристическая функция имеет вид

$$\theta(j\omega) = e^{-\frac{\sigma^2 \omega^2}{2}}.$$

Следовательно, закон распределения случайной величины  $Z$ , соответствующий характеристической функции  $\theta_z(j\omega)$ ,

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

Возвращаясь к случайной величине  $Y$ , получим:

$$p_y(y) = \frac{1}{\sqrt{2\pi n}\sigma} e^{-\frac{1}{2} \left( \frac{y - m_y}{\sigma\sqrt{n}} \right)^2}.$$

Это выражение показывает, что распределение вероятностей суммы независимых случайных величин, имеющих одинаковые законы распределения, стремится к нормальному при любом виде закона распределения слагаемых. С некоторыми ограничениями это заключение справедливо и для случая, когда законы распределения отдельных слагаемых не являются одинаковыми. Эти ограничения сводятся к тому, чтобы

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n M[|X_i - m_{x_i}|^3]}{(\sum_{i=1}^n \sigma_{x_i}^2)^{3/2}} = 0,$$

где  $m_{x_i}$  и  $\sigma_{x_i}^2$  — соответственно математические ожидания и дисперсии слагаемых  $X_i$ .

Центральная предельная теорема была доказана Ляпуновым в 1900 г. Она позволила объяснить широкое распространение нормального закона распределения в природе и технике в тех явлениях, где рассеяние изучаемых величин вызывается очень большим количеством случайных причин, влияние каждой из которых в отдельности ничтожно мало. В частности, эта теорема объясняет успешное применение нормального закона в теории ошибок измерений, при описании некоторых физических и биологических явлений и т.п.

## Глава 6

# Случайные процессы

### 6.1. Понятие случайного процесса

Часто приходится иметь дело с такими случайными величинами, которые в процессе одного наблюдения изменяются случайным образом с течением времени. Можно привести следующие примеры: осциллограммы собственных шумов электронного устройства; перемещение броуновской частицы; колебания давления температуры, влажности и скорости ветра в какой-либо точке атмосферы и др. Во всех подобных случаях говорят о *случайном процессе*. Конкретный вид случайного процесса (то есть его фактическая запись, например, в виде фотографии или осциллограммы) в определенном опыте называется *реализацией* случайного процесса.

Для формального обозначения зависимости случайного процесса (наблюдаемой величины) от аргументов применяются случайные функции. Будем обозначать случайные функции буквами греческого алфавита с указанием в скобках аргумента, а их реализации — малыми буквами латинского алфавита.

*Случайной функцией*  $\xi(t)$  называется такая функция неслучайного аргумента  $t$ , которая при любом фиксированном значении аргумента является случайной величиной. Это означает, что при неизменных условиях опыта значения  $\xi(t)$  в реализациях, полученных для нескольких полностью *идентичных систем*, будут различными. Под идентичными понимаются системы, в которых воспроизводятся одни и те же условия протекания случайного процесса и одни и те же условия его наблюдения и регистрации. Если, например, речь идет о флуктуациях, то одинаковыми должны быть макроскопические параметры всех систем, составляющих ансамбль. В этом состоит существенное различие между случайной и детерминированной функциями. Значение последней однозначно определяется значениями аргументов. Элемент случайности в совокупности реализаций может быть различным. Если он полностью отсутствует, что в действитель-

ности никогда не имеет места, то случайная функция переходит в детерминированную.

Очевидно, что задать случайный процесс аналитически (формулой) невозможно, а значит, будущие значения процесса не могут быть точно предсказаны на основе зарегистрированных ранее значений. Говорят, что такой процесс является *недетерминированным*. Почти все существующие в природе случайные процессы относятся к недетерминированным, так как физический механизм, лежащий в основе их возникновения, либо не наблюдаем, либо очень сложен.

Однако иногда имеется возможность определить случайные процессы, для которых будущие значения можно точно предсказать, зная прежние значения. Такие случайные процессы называются *квазидетерминированными*. В качестве примера можно привести случайный процесс вида

$$\xi(t) = A \cos(\omega t + \varphi),$$

$A$  и  $\omega$  — постоянные,  $\varphi$  — случайная фаза, имеющая определенный закон распределения. Для какой-то одной реализации величина  $\varphi$  имеет одно и то же значение, но для других членов ансамбля — другие значения. В этом случае имеют место изменения только по ансамблю реализаций, но не по времени.

Случайная функция может зависеть не от одного, а от нескольких аргументов (например, скорость ветра зависит от времени и от пространственных координат). При записи случайной функции обычно указывают область ее задания, то есть область возможного изменения аргументов. Наиболее часто приходится оперировать со случайными процессами  $\xi(t)$ , зависящими от одного аргумента — времени.

В дальнейшем элементарном изложении будут рассматриваться только случайные процессы с вещественными значениями.

Случайные процессы называют также *стохастическими*, или *вероятностными*.

### 6.2. Виды случайных процессов

В зависимости от того, непрерывное или дискретное множество значений принимают случайная величина  $\xi(t)$  и ее аргумент  $t$ , различают следующие пять видов случайных процессов:

1. Дискретная случайная последовательность (дискретный процесс с дискретным временем) — случайный процесс, у которого

область значений и область определения являются дискретными множествами. В данном случае время  $t$  пробегает дискретный ряд значений  $t_0, t_1, \dots, t_i, \dots, t_m$ , а случайная величина  $x_i = \xi(t_i)$  может принимать лишь дискретное множество значений  $x_0, x_1, \dots, x_k$ . Множества значений  $t_i$  и  $x_i$  могут быть конечными или бесконечными. Характер временной реализации такого случайного процесса показан на рис. 6.1.

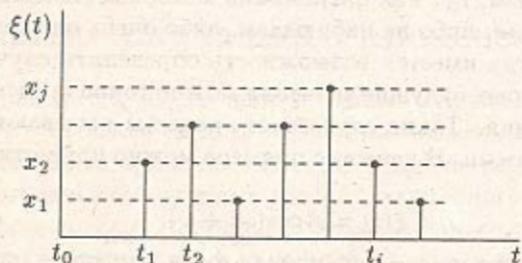


Рис. 6.1. Реализация дискретной случайной последовательности

2. Случайная последовательность (непрерывный процесс с дискретным временем) — случайный процесс, у которого область значений представляет собой непрерывное множество, а область определения — дискретное. Такой процесс отличается от процесса первого вида лишь тем, что теперь случайная величина  $\xi(t_i)$ ,  $i = 1, 2, \dots, m$  может принимать непрерывный ряд значений. Реализация одного из такого вида процессов приведена на рис. 6.2.

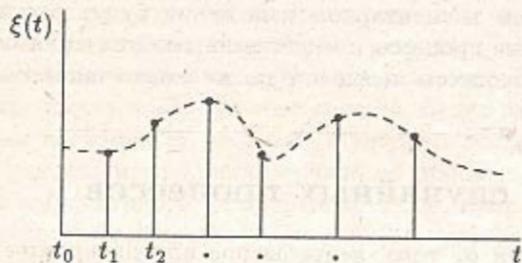


Рис. 6.2. Реализация случайной последовательности

3. Дискретный случайный процесс (дискретный процесс с непрерывным временем) — случайный процесс, у которого область зна-

чений — дискретное, а область определения — непрерывное множество. В этом случае  $\xi(t)$  может принимать дискретные значения  $x_j$ ,  $j = 1, 2, \dots, k$ , а время  $t$  — непрерывные значения на интервале задания процесса  $\xi(t)$  (рис. 6.3).

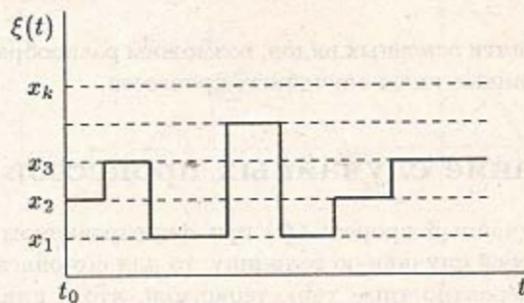


Рис. 6.3. Реализация дискретного случайного процесса

4. Непрерывнозначный случайный процесс — случайный процесс, область значений и область определения которого — непрерывные множества. В данном случае  $\xi(t)$  принимает значения из некоторого непрерывного пространства и аргумент  $t$  изменяется также непрерывно, причем реализации могут иметь разрывы первого рода. Если подобные скачки отсутствуют, то такой процесс называется непрерывным. Пример реализации такого процесса показан на рис. 6.4.

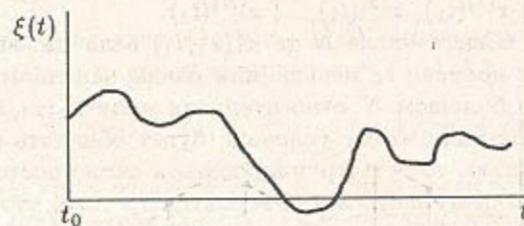


Рис. 6.4. Реализация непрерывнозначного случайного процесса

5. Случайный точечный процесс — случайный процесс, представляющий собой последовательность точек, расположенных случайным образом на оси времени (рис. 6.5). Эти точки могут соответствовать различным событиям. С таким видом случайного процесса мы уже встречались при выводе распределения Пуассона.



Рис. 6.5. Реализация случайного точечного процесса

Кроме этих пяти основных видов, возможны разнообразные, более сложные, смешанные виды случайных процессов.

### 6.3. Описание случайных процессов

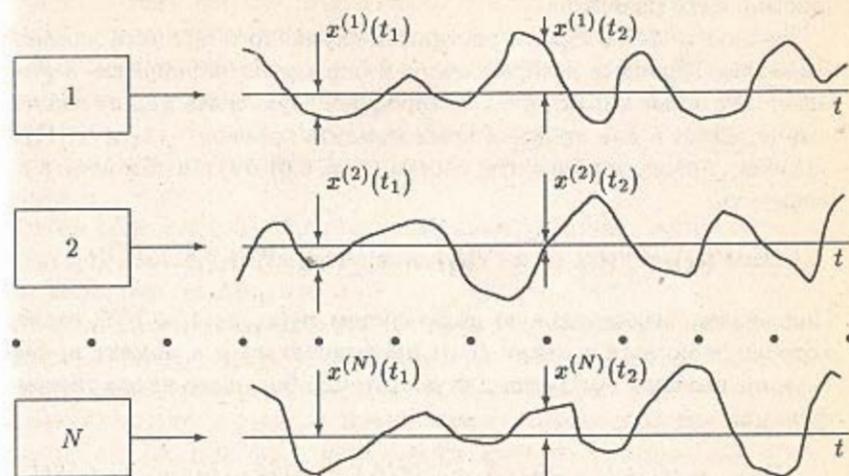
Так как случайный процесс  $\xi(t)$  при фиксированном значении  $t$  представляет собой случайную величину, то для его описания применяются те же вероятностные характеристики, что и для случайных величин, а именно: плотности вероятности (законы распределения), функции распределения, характеристические функции, моментные и корреляционные функции.

Пусть имеется большое число  $N$  полностью одинаковых систем (рис. 6.6), образующих некоторый "ансамбль". Допустим, что все системы работают одновременно при одинаковых условиях. На выходах этих систем наблюдается случайный процесс  $\xi(t)$ . Если к каждой системе подключить одинаковые регистрирующие приборы (например, осциллографы) и на всех приборах в один и тот же момент времени  $t_1$  измерить мгновенные значения, то получим отличающиеся друг от друга величины  $x^{(1)}(t_1), x^{(2)}(t_1), \dots, x^{(N)}(t_1)$ .

Выделим из общего числа  $N$  те  $n_1(x_1, t_1)$  величин, значения которых в момент времени  $t_1$  меньше или равны заданному числу  $x_1$ . При достаточно большом  $N$  относительная доля  $n_1(x_1, t_1)/N$  величин, удовлетворяющих этому условию, будет обладать статистической устойчивостью, то есть группироваться около постоянного числа, и может рассматриваться как вероятность того, что при  $t = t_1$  случайная функция  $\xi(t)$  находится ниже уровня  $x_1$ :

$$P(\xi(t_1) < x_1) = F(x_1; t_1) \simeq n_1(x_1; t_1)/N, \quad \text{при } N \rightarrow \infty.$$

Функция  $F(x_1; t_1)$  является *одномерной функцией распределения* случайного процесса. Слово "одномерная" подчеркивает тот факт, что рассматриваются значения случайной функции в один фиксированный момент времени.

Рис. 6.6. Ансамбль  $N$  одинаковых систем

Производная от функции распределения, если она существует, является одномерной плотностью вероятности случайного процесса:

$$p(x_1; t_1) = \frac{\partial}{\partial x_1} F(x_1; t_1).$$

Безразмерная величина  $p(x_1; t_1) dx_1$  равна вероятности того, что значения  $\xi(t_1)$  попадут в интервал  $x_1 \leq \xi(t_1) < x_1 + dx_1$ :

$$p(x_1; t_1) dx_1 = P(x_1 \leq \xi(t_1) < x_1 + dx_1).$$

Правую часть этого равенства можно рассматривать как относительную долю систем  $\Delta n_1(x_1; t_1)/N$ , отсчеты которых в момент времени  $t_1$  попадают в горизонтальное окно  $[x_1, x_1 + dx_1)$ .

*Одномерная плотность вероятности*, как и функция распределения, является важной, но неполной характеристикой случайного процесса. Она дает представление о процессе лишь в отдельные, фиксированные моменты времени, не указывая, например, на то, как значения  $\xi(t_1)$  в момент времени  $t_1$  влияют на дальнейшее поведение процесса при  $t_2 > t_1$ . Можно сказать, что одномерная плотность вероят-

ности характеризует процесс "статически" и не дает представления о динамике его развития.

Более полными характеристиками случайного процесса являются *двумерная функция распределения* и *двумерная плотность вероятности*, которые характеризуют вероятностную связь между значениями процесса в два произвольных момента времени —  $t_1$  и  $t_2$ . Пусть значения процесса на выходе систем (рис. 6.6) в эти два момента времени есть

$$x^{(1)}(t_1), x^{(2)}(t_1), \dots, x^{(N)}(t_1) \text{ и } x^{(1)}(t_2), x^{(2)}(t_2), \dots, x^{(N)}(t_2).$$

Подсчитаем относительную долю систем  $n_2(x_1, x_2; t_1, t_2)/N$ , отсчеты которых в момент времени  $t_1$  не превышают  $x_1$  и в момент времени  $t_2$  не превышают  $x_2$ . Тогда для достаточно большого числа систем  $N$  функция

$$F_2(x_1, x_2; t_1, t_2) = P[\xi(t_1) < x_1; \xi(t_2) < x_2] \simeq n_2(x_1, x_2; t_1, t_2)/N$$

является двумерной функцией распределения случайного процесса. Производная от этой функции

$$p_2(x_1, x_2; t_1, t_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_2(x_1, x_2; t_1, t_2)$$

называется двумерной плотностью вероятности случайного процесса.

Безразмерная величина  $p_2(x_1, x_2; t_1, t_2) dx_1 dx_2$  определяет вероятность совместного выполнения двух неравенств:

$$p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 = P(x_1 \leq \xi(t_1) < x_1 + dx_1, x_2 \leq \xi(t_2) < x_2 + dx_2).$$

Здесь правая часть при больших  $N$  представляет собой относительную долю систем  $\Delta n_2(x_1, x_2; t_1, t_2)/N$ , отсчеты которых в момент времени  $t_1$  попадают в горизонтальное окно  $[x_1, x_1 + dx_1)$  и в момент времени  $t_2$  — в окно  $[x_2, x_2 + dx_2)$ .

В общем случае двумерная функция распределения или двумерная плотность вероятности также не дают исчерпывающего описания случайного процесса. Они позволяют судить о связи между вероятностными значениями случайного процесса лишь в два момента времени. Более полное и детальное описание случайного процесса дается многомерными плотностями вероятности или функциями распределения.

Плотность вероятности  $p_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ , называемая  $n$ -мерной, определяет вероятность того, что значения случайного процесса  $\xi(t)$  в  $n$  моментов времени  $t_1, t_2, \dots, t_n$  заключены соответственно в малых интервалах  $[x_1, x_1 + dx_1)$ , ...,  $[x_n, x_n + dx_n)$ . Эта вероятность равна  $p_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2, \dots, dx_n$ . Плотность вероятности  $p_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$  позволяет судить о связи между вероятностными значениями процесса в  $n$  произвольных моментах времени.

Таким образом, случайный процесс в общем случае описывается с помощью  $n$ -мерной плотности вероятности (функции распределения), и тем детальнее, чем больше  $n$ .

Из приведенных определений следует, что плотности вероятности  $p_n(x_1, \dots, x_n; t_1, \dots, t_n)$  и функции распределения  $F_n(x_1, \dots, x_n; t_1, \dots, t_n)$  случайного процесса  $\xi(t)$ , не считая небольшой разницы в обозначениях, полностью аналогичны совместным плотностям вероятности  $p_n(x_1, \dots, x_n)$  и функциям распределения  $F_n(x_1, \dots, x_n)$  совокупности  $n$  случайных величин  $x_1, \dots, x_n$ . Плотности вероятности  $p_n(x_1, \dots, x_n; t_1, \dots, t_n)$  и функции распределения  $F_n(x_1, \dots, x_n; t_1, \dots, t_n)$  по-прежнему связаны однозначными зависимостями, причем плотность вероятности случайного процесса  $p_n(x_1, \dots, x_n; t_1, \dots, t_n)$  удовлетворяет прежним условиям неотрицательности, нормировки, симметрии и согласованности. В частности, последнее условие с учетом разницы в обозначениях принимает вид

$$\begin{aligned} p_m(x_1, \dots, x_m; t_1, \dots, t_m) &= \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_n(x_1, \dots, x_m; x_{m+1}, \dots, x_n; t_1, \dots, t_n) dx_{m+1} \dots dx_n, \quad m < n. \end{aligned}$$

Для совместного вероятностного описания двух или нескольких случайных процессов вводят *совместные* функции распределения и плотности вероятности. Так, для двух процессов  $\xi(t)$  и  $\eta(t)$  их определяют при помощи следующих соотношений:

$$\begin{aligned} F_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) &= \\ &= P(\xi(t_1) < x_1, \dots, \xi(t_n) < x_n, \eta(t'_1) < y_1, \dots, \eta(t'_m) < y_m); \\ p_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) dx_1 \dots dx_n dy_1 \dots dy_m &= \\ &= P(x_1 \leq \xi(t_1) < x_1 + dx_1, \dots, x_n \leq \xi(t_n) < x_n + dx_n, \\ & \quad y_1 \leq \eta(t'_1) < y_1 + dy_1, \dots, y_m \leq \eta(t'_m) < y_m + dy_m), \end{aligned}$$

где  $n, m$  — целые неотрицательные числа.

Два случайных процесса  $\xi(t)$  и  $\eta(t)$  называются независимыми, если совокупность значений первого процесса  $\xi(t_1), \dots, \xi(t_n)$  не зависит от совокупности значений второго процесса  $\eta(t_1), \dots, \eta(t_m)$  при любых  $t_1, \dots, t_n, t'_1, \dots, t'_m$ . Необходимым и достаточным условием независимости процессов является то, чтобы их совместная плотность вероятности распадалась на произведение плотностей вероятностей каждого из процессов:

$$p_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) = p_n(x_1, \dots, x_n, t_1, \dots, t_n) p_m(y_1, \dots, y_m, t'_1, \dots, t'_m).$$

Для случайных процессов можно ввести условные плотности вероятности. Так, например, случайное значение процесса  $\xi(t_1)$  при известном значении его в другой момент времени  $\xi(t_2) = x_2$  описывается условной плотностью вероятности

$$p(x_1; t_1 | x_2; t_2) = \frac{p_2(x_1, x_2; t_1, t_2)}{p_1(x_2; t_2)},$$

где

$$p_1(x_2; t_2) = \int_{-\infty}^{\infty} p_2(x_1, x_2; t_1, t_2) dx_1.$$

Условная плотность вероятности  $p(x_1; t_1 | x_2; t_2)$  содержит больше (по крайней мере не меньше) сведений о  $\xi(t_1)$ , чем безусловная плотность вероятности  $p_1(x_1; t_1)$ . Насколько именно увеличилась информация о  $\xi(t_1)$  в результате того, что стало известным значение  $\xi(t_2) = x_2$ , зависит от конкретных условий. В некоторых случаях информации о  $\xi(t_1)$  вообще не прибавляется, каким бы ни оказалось значение  $x_2$ . Это значит, что

$$p(x_1; t_1 | x_2; t_2) = p(x_1; t_1);$$

при этом

$$p_2(x_1, x_2; t_1, t_2) = p(x_1; t_1) p(x_2; t_2).$$

Эта формула выражает необходимое и достаточное условие независимости значений случайного процесса  $\xi(t)$  в два момента времени —  $t_1$  и  $t_2$ .

Вместо плотностей вероятностей для описания случайного процесса можно использовать характеристическую функцию:

$$\begin{aligned} \theta_n(j\omega_1, \dots, j\omega_n; t_1, \dots, t_n) &= M[e^{j\omega_1 \xi_1 + \dots + j\omega_n \xi_n}] = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{j(\omega_1 x_1 + \dots + \omega_n x_n)} p_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n, \end{aligned}$$

где  $\xi_i = \xi(t_i)$ .

Напомним, что между функцией распределения, плотностью вероятности и характеристической функцией существует однозначная связь.

## 6.4. Моментные функции

Хотя полная характеристика случайного процесса дается его многомерным законом распределения, на практике в большинстве случаев достаточно оперировать только моментными функциями. Моментные функции в теории случайных процессов играют такую же роль, какую выполняют моменты в теории случайных величин. И те и другие являются результатом некоторого усреднения. Различают начальные и центральные моментные функции.

Начальные моментные функции определяются следующим образом: начальная моментная функция порядка  $\nu_1$

$$m_{\nu_1}(t) = M[\xi^{\nu_1}(t)] = \int_{-\infty}^{\infty} x^{\nu_1}(t) p(x; t) dx;$$

начальная моментная функция порядка  $\nu_1 + \nu_2$

$$\begin{aligned} m_{\nu_1 \nu_2}(t_1, t_2) &= M[\xi^{\nu_1}(t_1) \xi^{\nu_2}(t_2)] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^{\nu_1} x_2^{\nu_2} p_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

Моментная функция  $m_{\nu_1 \nu_2 \dots \nu_n}(t_1, t_2, \dots, t_n)$ , зависящая от  $n$  несовпадающих аргументов  $t_1, t_2, \dots, t_n$ , называется  $n$ -мерной моментной функцией  $\nu_1 + \nu_2 + \dots + \nu_n$ -го порядка.

Центральные моментные функции определяются формулой

$$m_{\nu_1 \dots \nu_n}(t_1, \dots, t_n) = M[(\xi(t_1) - m_1(t_1))^{\nu_1} \dots (\xi(t_n) - m_1(t_n))^{\nu_n}] =$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_1 - m_1(t_1))^{\nu_1} \dots (x_n - m_1(t_n))^{\nu_n} p_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n.$$

Особую роль играют следующие моментные функции.

1. Начальная моментная функция первого порядка  $m_1(t)$ :

$$m_1(t) = M[\xi(t)] = \int_{-\infty}^{\infty} xp(x; t) dx,$$

которая представляет собой математическое ожидание случайного процесса. Ее часто обозначают как  $m_\xi(t)$ .

2. Центральная моментная функция второго порядка вида

$$\mu_{20}(t) = M[(\xi(t) - m_\xi(t))^2] = \int_{-\infty}^{\infty} (x - m_\xi(t))^2 p(x; t) dx,$$

которая является дисперсией случайного процесса. Ее часто обозначают как

$$D_\xi(t) = \sigma_\xi^2(t).$$

3. Начальная моментная функция второго порядка  $m_{11}(t_1, t_2)$ :

$$m_{11}(t_1, t_2) = M[\xi(t_1)\xi(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p_2(x_1, x_2; t_1, t_2) dx_1 dx_2.$$

Эта функция называется *ковариационной функцией* случайного процесса и обозначается как  $K_\xi(t_1, t_2)$ .

4. Центральная моментная функция второго порядка  $\mu_{11}(t_1, t_2)$ :

$$\begin{aligned} \mu_{11}(t_1, t_2) &= M[(\xi(t_1) - m_\xi(t_1))(\xi(t_2) - m_\xi(t_2))] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_\xi(t_1))(x_2 - m_\xi(t_2)) p_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

Эта функция носит название *корреляционной функции* случайного процесса и обозначается как  $R_\xi(t_1, t_2)$ .

Заметим, что в иностранной литературе обычно используется обратная терминология:  $K_\xi(t_1, t_2)$  называется корреляционной, а  $R_\xi(t_1, t_2)$  — ковариационной функцией случайного процесса.

Из выражений для  $K_\xi(t_1, t_2)$  и  $R_\xi(t_1, t_2)$  следует связь между этими функциями:

$$K_\xi(t_1, t_2) = R_\xi(t_1, t_2) + m_\xi(t_1)m_\xi(t_2).$$

Таким образом, математическое ожидание  $m_\xi(t)$  и дисперсия  $D_\xi(t)$  определяются одномерным законом распределения случайного процесса, а ковариационная и корреляционная функции — двумерным законом. Они являются неслучайными функциями. Корреляционная функция характеризует статистическую (вероятностную) связь между значениями случайного процесса в два различных момента времени.

*Рассмотрим пример.* Случайный процесс  $\xi(t)$  задан в виде  $\xi(t) = X \cos(\omega t)$ , где  $X$  — случайная величина с математическим ожиданием  $a$  и дисперсией  $\sigma^2$ . Необходимо найти математическое ожидание, дисперсию и корреляционную функцию процесса.

Используя свойства математического ожидания и дисперсии случайных величин, получим:

$$\begin{aligned} m_\xi(t) &= M[X \cos \omega t] = \cos \omega t \cdot M[X] = a \cos \omega t; \\ D_\xi(t) &= D[X \cos \omega t] = \cos^2 \omega t \cdot D[X] = \sigma^2 \cos^2 t. \end{aligned}$$

Для корреляционного момента имеем:

$$\begin{aligned} R_\xi(t_1, t_2) &= M[(X \cos \omega t_1 - a \cos \omega t_1)(X \cos \omega t_2 - a \cos \omega t_2)] = \\ &= \cos \omega t_1 \cos \omega t_2 M[(X - a)(X - a)] = \\ &= \cos \omega t_1 \cos \omega t_2 M[(X - a)^2] = \cos \omega t_1 \cos \omega t_2 \sigma^2. \end{aligned}$$

## 6.5. Нестационарные и стационарные случайные процессы

Важным классом случайных процессов являются *стационарные* случайные процессы. Случайный процесс  $\xi(t)$  называется стационарным в узком (строгом) смысле, если все конечномерные функции распределения вероятностей любого порядка инвариантны относительно сдвига по времени, то есть при любых  $n$  и  $t_0$  справедливо равенство

$$F_n(x_1, \dots, x_n; t_1 - t_0, \dots, t_n - t_0) = F_n(x_1, \dots, x_n; t_1, \dots, t_n).$$

Это означает, что процессы  $\xi(t)$  и  $\xi(t - t_0)$  имеют одинаковые вероятностные характеристики при любом  $t_0$ . Случайные процессы, не удовлетворяющие этому условию, называются нестационарными в узком смысле.

Разумеется, аналогичное равенство должно выполняться и для плотностей вероятностей

$$p_n(x_1, \dots, x_n; t_1 - t_0, \dots, t_n - t_0) = p_n(x_1, \dots, x_n; t_1, \dots, t_n),$$

а также для характеристических, моментных и корреляционных функций.

Стационарный в узком смысле случайный процесс в отличие от нестационарного ведет себя однородно (однообразно) во времени. Характер реализаций таких процессов показан соответственно на рис. 6.7, а и б.

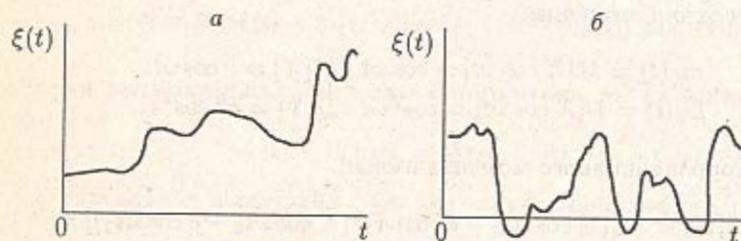


Рис. 6.7. Реализация нестационарного (а) и стационарного (б) случайных процессов

Стационарные в узком смысле случайные процессы аналогично установившимся детерминированным процессам получают в установившемся режиме работы системы при неизменных внешних условиях.

Понятие стационарности в узком смысле обобщается на два и более случайных процесса. Два случайных процесса  $\xi(t)$  и  $\eta(t)$  называются совместно стационарными в узком смысле, если их совместные функции распределения любого порядка инвариантны относительно сдвига по времени:

$$F_{\xi\eta}(x_1, \dots, x_n; y_1, \dots, y_m; t_1, \dots, t_n; t'_1, \dots, t'_m) = F_{\xi\eta}(x_1, \dots, x_n; y_1, \dots, y_m; t_1 - t_0, \dots, t_n - t_0; t'_1 - t_0, \dots, t'_m - t_0).$$

Если каждый из процессов  $\xi(t)$  и  $\eta(t)$  является стационарным, то отсюда вовсе не следует, что они будут стационарными в узком смысле.

Из определения стационарности случайного процесса, в частности, следует, что

$$p(x; t_1) = p(x; t_1 - t_1) = p_1(x);$$

$$p_2(x_1, x_2; t_1, t_2) = p_2(x_1, x_2; t_1 - t_1, t_2 - t_1) = p_2(x_1, x_2; \tau),$$

где  $\tau = t_2 - t_1$ ,

$$p_3(x_1, x_2, x_3; t_1, t_2, t_3) = p_3(x_1, x_2, x_3; \tau, t_3 - t_1);$$

...

$$p_n(x_1, \dots, x_n; t_1, \dots, t_n) = p_n(x_1, \dots, x_n; \tau, t_3 - t_1, \dots, t_n - t_1).$$

Таким образом, для стационарного в узком смысле случайного процесса  $n$ -мерная плотность вероятности,  $n$ -мерные моменты и корреляционные функции зависят не от  $n$ , а от  $n - 1$  моментов времени, так как один из этих моментов можно всегда принять за начало отсчета (например, положить  $t_1 = 0$ ).

Из приведенных формул видно, что одномерная плотность вероятности стационарного в узком смысле случайного процесса вообще не зависит от времени. Поэтому одномерная плотность вероятности и одномерные моменты не учитывают временных характеристик стационарного процесса. Так, процесс, протекающий в  $\nu$  раз быстрее или медленнее, будет иметь одну и ту же одномерную плотность вероятности. Грубо говоря, описание случайного процесса одномерной плотностью вероятности подобно указанию амплитуды гармонического колебания  $A \cos(\omega t + \varphi)$  без задания его частоты. Отсюда ясно, что описание процесса с помощью одномерной плотности вероятности является неполным.

Математическое ожидание (среднее значение) стационарного в узком смысле случайного процесса не зависит от времени:

$$m_\xi = M[\xi(t)] = \int_{-\infty}^{\infty} x p_1(x) dx.$$

Ковариационная  $K_\xi(t_1, t_2)$  и корреляционная  $R_\xi(t_1, t_2)$  функции зависят лишь от разности аргументов  $\tau = t_2 - t_1$ , причем

$$\begin{aligned} R_\xi(\tau) &= M[(\xi(t) - m_\xi)(\xi(t + \tau) - m_\xi)] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_\xi)(x_2 - m_\xi) p_2(x_1, x_2; \tau) dx_1 dx_2 = \end{aligned}$$

$$= M[\xi(t)\xi(t+\tau)] - m_\xi^2 = K_\xi(\tau) - m_\xi^2.$$

Дисперсия стационарного процесса

$$D_\xi = \sigma_\xi^2 = M[(\xi(t) - m_\xi)^2] = R_\xi(0) = \int_{-\infty}^{\infty} (x - m_\xi)^2 p_1(x) dx = M[\xi^2(t)] - m_\xi^2.$$

Как видно, дисперсия постоянна и равна значению корреляционной функции при нулевом значении аргумента.

При решении некоторых практических задач многомерные плотности вероятности не рассматривают, а оперируют только с математическим ожиданием и ковариационной или корреляционной функциями. В связи с этим вводят понятие стационарности в широком смысле.

Случайный процесс  $\xi(t)$  с конечной дисперсией называется стационарным в широком смысле, если его математическое ожидание и ковариационная функция инвариантны относительно сдвига по времени, то есть математическое ожидание постоянно (не зависит от времени), а ковариационная функция зависит только от разности аргументов  $t_2 - t_1$ :

$$m_\xi = \text{const}; K_\xi(t_1, t_2) = K_\xi(t_2 - t_1).$$

Легко видеть, что случайные процессы, стационарные в узком смысле, всегда стационарны в широком смысле. Однако обратное утверждение в общем случае неверно.

Два случайных процесса  $\xi(t)$  и  $\eta(t)$  называются стационарно связанными в широком смысле, если их взаимная ковариационная функция инвариантна относительно сдвига по времени:

$$K_{\xi\eta}(t_1, t_2) = M[\xi(t_1)\eta(t_2)] = M[\xi(t_1 - t_1)\eta(t_2 - t_1)] = K_{\xi\eta}(\tau),$$

где  $\tau = t_2 - t_1$ . Однако если каждый из процессов  $\xi(t)$  и  $\eta(t)$  является стационарным в широком смысле, то это вовсе не означает, что они являются стационарно связанными в широком смысле.

*Рассмотрим пример.* Пусть имеется случайный процесс вида

$$\xi(t) = A \cos(\omega t + \varphi),$$

где  $A$  и  $\omega$ , соответственно, — постоянная амплитуда и частота,  $\varphi$  — случайная фаза, распределенная равномерно в интервале  $(0, 2\pi)$ .

Посмотрим, является ли такой процесс стационарным в широком смысле.

Используя формулу для преобразования моментов функционально связанных случайных величин, найдем:

$$m_\xi(t) = M[A \cos(\omega t + \varphi)] = A \int_0^{2\pi} \cos(\omega t + \varphi) \frac{1}{2\pi} d\varphi = 0.$$

Найдем корреляционную функцию, учитывая, что  $m_\xi(t) = 0$ :

$$\begin{aligned} R_\xi(t_1, t_2) &= M[A \cos(\omega t_1 + \varphi) A \cos(\omega t_2 + \varphi)] = \\ &= A^2 M \left[ \frac{\cos(\omega t_2 - \omega t_1) + \cos(\omega t_2 + \omega t_1 + 2\varphi)}{2} \right] = \\ &= \frac{A^2}{2} \cos \omega(t_2 - t_1), \end{aligned}$$

поскольку  $M[\cos(\omega t_2 + \omega t_1 + 2\varphi)] = 0$ .

Итак, математическое ожидание рассматриваемого случайного процесса постоянно при всех значениях аргумента  $t$ , а корреляционная функция зависит только от разности аргументов. Следовательно, случайный процесс является стационарным в широком смысле.

## 6.6. Эргодические стационарные процессы

Плотности вероятности случайных процессов имеют статистический смысл. Это относительные частоты в ансамбле одинаковых систем, в каждой из которых воспроизведены одни и те же условия протекания данного случайного процесса  $\xi(t)$  и одни и те же способы его регистрации или наблюдения. Если, например, речь идет о флуктуациях, то одинаковыми должны быть макроскопические характеристики всех систем, составляющих ансамбль. Имея ансамбль систем, мы располагаем большим набором реализаций рассматриваемого случайного процесса и с помощью соответствующих вероятностей, то есть распределения систем ансамбля по возможным значениям  $\xi(t)$ , можем находить  $M[\xi(t)]$ ,  $R_\xi(t_1, t_2)$  и т.д.

Теоретики предпочитают оперировать с ансамблями, но у экспериментаторов обычно только одна лаборатория и только одна установка, а не, скажем,  $10^6$  или  $10^9$ . За данный промежуток времени  $(0, T)$

экспериментатор может получить лишь одну реализацию интересующего его случайного процесса и поэтому предпочитает усреднять по времени, пользуясь одной реализацией  $\xi(t)$ , одной осциллограммой.

Спрашивается: в каком соотношении находятся эти два способа усреднения — по времени и по ансамблю? Оказывается, для большинства случайных процессов, являющихся стационарными в узком смысле, эти способы усреднения эквивалентны, то есть моментные функции можно получить путем усреднения соответствующих величин “вдоль процесса”, пользуясь одной реализацией достаточно большой длительности. Представляется, например, естественным за оценку математического ожидания  $m_\xi$  стационарного процесса  $\xi(t)$  принять величину

$$\hat{m}_\xi = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt,$$

а в качестве оценок дисперсии  $D_\xi$  и корреляционной функции  $R_\xi(\tau)$  взять соответственно

$$\hat{D}_\xi = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (\xi(t) - m_\xi)^2 dt;$$

$$\hat{R}_\xi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (\xi(t + \tau) - m_\xi)(\xi(t) - m_\xi) dt.$$

На практике временной интервал усреднения  $T$  берут конечным, но по возможности большим.

Возможность усреднения по времени может быть оправдана тем, что стационарный случайный процесс протекает однородно во времени. В связи с этим одна реализация достаточно большой длительности может содержать все сведения о свойствах случайного процесса. Это можно пояснить следующим образом. Представим себе, что длинная реализация стационарного случайного процесса разбита на “кусочки” примерно одинаковой длительности. Для ряда стационарных процессов каждый из таких “кусочков” можно рассматривать в качестве “полномочного представителя” отрезка реализации на выходе отдельного члена статистического ансамбля одинаковых систем. Сказанное схематически иллюстрируется рис. 6.8.

Стационарные случайные процессы, для которых вероятностные характеристики процессов можно получить при надлежащей обработке лишь одной реализации достаточно большой длительности, назы-

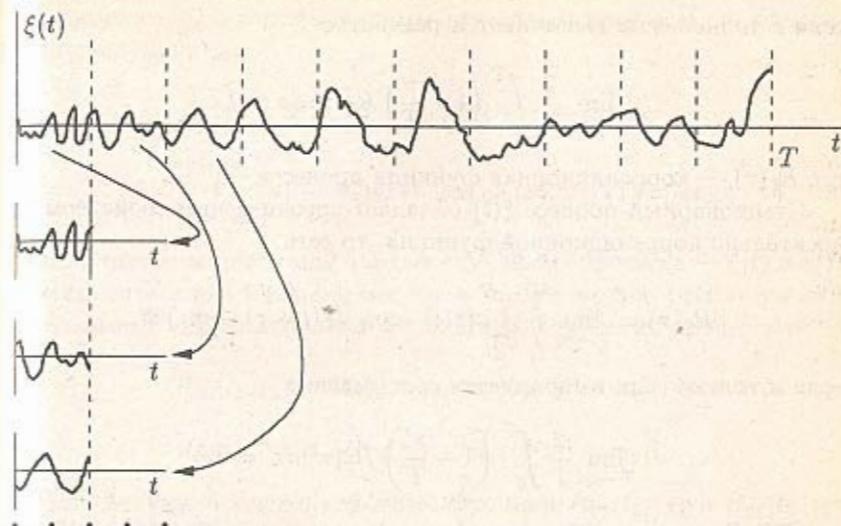


Рис. 6.8. Формирование реализаций ансамбля одинаковых систем из одной реализации стационарного случайного процесса

ваются *эргодическими* случайными процессами. Часто говорят, что такие случайные процессы обладают эргодическим свойством.

Стационарный случайный процесс называется эргодическим в строгом смысле, если с вероятностью единица все его вероятностные характеристики могут быть найдены по одной реализации процесса. На практике часто интересуются не всеми, а только отдельными характеристиками процесса — например, математическим ожиданием, корреляционной функцией и одномерной плотностью вероятности. Ясно, что процесс может быть эргодическим относительно одной характеристики (параметра) и неэргодическим для других. В связи с этим вводят понятие эргодичности относительно отдельных характеристик процесса.

Стационарный случайный процесс  $\xi(t)$  обладает эргодическим свойством относительно математического ожидания, то есть

$$M[\xi(t)] = m_\xi = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt,$$

если и только если выполняется равенство

$$\lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R_{\xi}(\tau) d\tau = 0,$$

где  $R_{\xi}(\tau)$  — корреляционная функция процесса.

Стационарный процесс  $\xi(t)$  обладает эргодическим свойством относительно корреляционной функции, то есть

$$R_{\xi}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (\xi(t) - m_{\xi})(\xi(t + \tau) - m_{\xi}) dt,$$

если и только если выполняется соотношение

$$\lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T \left(1 - \frac{\tau'}{T}\right) R_{\eta}(\tau') d\tau' = 0,$$

где  $R_{\eta}(\tau')$  — корреляционная функция процесса  $\eta(t) = \xi(t + \tau)\xi(t)$ .

Заметим, что для эргодичности относительно математического ожидания необходима стационарность процесса в широком смысле, а для эргодичности относительно корреляционной функции — стационарность в узком смысле. Не существует приборов, позволяющих определить эргодичность случайного процесса. В связи с этим на практике эргодичность предполагается, а затем убеждаются, что это предположение справедливо.

## 6.7. Корреляционные функции и их свойства

Корреляционная функция случайного процесса представляет собой математическое ожидание произведения значений централизованного случайного процесса для двух моментов времени —  $t_1$  и  $t_2$ :

$$\begin{aligned} R_{\xi}(t_1, t_2) &= M[(\xi(t_1) - m_{\xi}(t_1))(\xi(t_2) - m_{\xi}(t_2))] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_{\xi}(t_1)][x_2 - m_{\xi}(t_2)] p(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

Эту корреляционную функцию обычно называют *автокорреляционной функцией* случайного процесса.

Применительно к стационарным случайным процессам автокорреляционная функция

$$\begin{aligned} R_{\xi}(\tau) &= M[(\xi(t) - m_{\xi})(\xi(t + \tau) - m_{\xi})] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_{\xi})(x_2 - m_{\xi}) p_2(x_1, x_2; \tau) dx_1 dx_2. \end{aligned}$$

Если имеются два стационарных случайных процесса —  $\xi(t)$  и  $\eta(t)$  с математическими ожиданиями  $m_{\xi}$  и  $m_{\eta}$ , то можно рассматривать две взаимные корреляционные функции между этими процессами:

$$R_{\xi\eta}(t_1, t_2) = M[(\xi(t_1) - m_{\xi})(\eta(t_2) - m_{\eta})];$$

$$R_{\eta\xi}(t_1, t_2) = M[(\eta(t_1) - m_{\eta})(\xi(t_2) - m_{\xi})].$$

Если взаимные корреляционные функции  $R_{\xi\eta}(t_1, t_2)$  и  $R_{\eta\xi}(t_1, t_2)$  зависят от разности  $\tau = t_2 - t_1$ , то процессы  $\xi(t)$  и  $\eta(t)$  называются стационарно связанными.

Значения случайного процесса  $\xi(t)$  в рассматриваемые моменты времени  $t_1$  и  $t_2$  называются некоррелированными, если  $R_{\xi}(t_1, t_2) = 0$ . Два случайных процесса —  $\xi(t)$  и  $\eta(t)$  называются некоррелированными, если их взаимная корреляционная функция для двух произвольных моментов времени равна нулю:

$$R_{\xi\eta}(t_1, t_2) = 0.$$

Корреляционные функции показывают качественную линейную связь между значениями одного или двух случайных процессов в выбранные моменты времени. Для количественной характеристики степени линейной зависимости случайных процессов вводят нормированные автокорреляционные и взаимные корреляционные функции, называемые коэффициентами автокорреляции и взаимной корреляции. Они определяются соответственно формулами

$$r_{\xi}(t_1, t_2) = \frac{R_{\xi}(t_1, t_2)}{\sqrt{D_{\xi}(t_1)D_{\xi}(t_2)}},$$

$$r_{\xi\eta}(t_1, t_2) = \frac{R_{\xi\eta}(t_1, t_2)}{\sqrt{D_{\xi}(t_1)D_{\eta}(t_2)}}.$$

Если случайные процессы стационарны и стационарно связаны, то

$$r_{\xi}(\tau) = \frac{R_{\xi}(\tau)}{D_{\xi}} = \frac{R_{\xi}(\tau)}{\sigma_{\xi}^2},$$

$$r_{\xi\eta}(\tau) = \frac{R_{\xi\eta}(\tau)}{\sqrt{D_{\xi}D_{\eta}}} = \frac{R_{\xi\eta}(\tau)}{\sigma_{\xi}\sigma_{\eta}}.$$

Как и для случайных величин, если случайные процессы связаны детерминированной линейной зависимостью, то коэффициент корреляции между ними в любой момент времени равен  $\pm 1$ . Если же случайные процессы независимы, то коэффициент корреляции равен нулю. Поэтому можно сказать, что коэффициент корреляции характеризует степень линейной зависимости между значениями одного или двух случайных процессов в выбранные моменты времени.

Независимые случайные процессы всегда являются некоррелированными. Однако обратное утверждение неверно, так как условие независимости является более жестким. Для нормальных случайных процессов некоррелированность значений процесса тождественна их независимости.

Рассмотрим свойства автокорреляционной функции стационарного случайного процесса  $\xi(t)$ :

1. Функция  $R_{\xi}(\tau)$  является четной, то есть

$$R_{\xi}(\tau) = R_{\xi}(-\tau).$$

Это следует из определения стационарного случайного процесса, то есть из условия независимости его характеристик от начала отсчета времени. Поэтому

$$R_{\xi}(\tau) = M[\xi(t)\xi(t+\tau)] - m_{\xi}^2 = M[\xi(t-\tau)\xi(t)] - m_{\xi}^2 = R_{\xi}(-\tau).$$

2. Абсолютное значение автокорреляционной функции при любом  $\tau$  не может превышать ее значения при  $\tau = 0$ , то есть

$$|R_{\xi}(\tau)| \leq R_{\xi}(0) = D_{\xi}.$$

Этот результат следует из того очевидного неравенства, что математическое ожидание положительной функции не может быть отрицательным:

$$M[(\xi(t) - m_{\xi}) \pm (\xi(t+\tau) - m_{\xi})]^2 \geq 0.$$

Отсюда имеем

$$M[(\xi(t) - m_{\xi})^2] \pm 2M[(\xi(t) - m_{\xi})(\xi(t+\tau) - m_{\xi})] + M[(\xi(t+\tau) - m_{\xi})^2] = 2D_{\xi} \pm 2R_{\xi}(\tau) \geq 0$$

и, следовательно,

$$D_{\xi} \geq |R_{\xi}(\tau)|.$$

3. Если  $\xi(t)$  — периодический процесс, то  $R_{\xi}(\tau)$  также будет периодической функцией с таким же периодом. Например, процесс  $\xi(t) = A \cos(\omega t + \varphi)$ ; где  $A$  и  $\omega$  — постоянные, а  $\varphi$  — случайная величина, распределенная равномерно в интервале  $(0, 2\pi)$ , имеет  $R_{\xi}(\tau) = \frac{A^2}{2} \cos \omega \tau$ .

4. Для многих стационарных процессов, представляющих практический интерес, справедливо соотношение

$$\lim_{\tau \rightarrow \infty} R_{\xi}(\tau) = 0,$$

которое заведомо гарантирует эргодичность процесса относительно математического ожидания. Физически этот результат объясняется тем, что устойчиво работающие системы обычно имеют конечное время затухания (конечную постоянную времени). Именно поэтому для случайных процессов, наблюдаемых в устойчиво работающих системах, последующее значение процесса оказывается практически независимым и некоррелированным с предыдущим значением, если они разделены достаточно большим интервалом времени.

5. Преобразование Фурье от корреляционной функции есть неотрицательная функция:

$$\int_{-\infty}^{\infty} R_{\xi}(\tau) e^{-j\omega\tau} d\tau \geq 0.$$

Этим свойством можно воспользоваться для решения вопроса о том, может ли какая-либо функция  $R(\tau)$ , удовлетворяющая предыдущим условиям, представлять корреляционную функцию стационарного в широком смысле случайного процесса. Смысл такого ограничения будет очевиден после рассмотрения спектральной плотности случайного процесса. Кроме всего прочего это ограничение отрицает возможность существования корреляционных функций с плоскими вершинами, вертикальными боковыми сторонами или какими-либо разрывами в их графических изображениях.

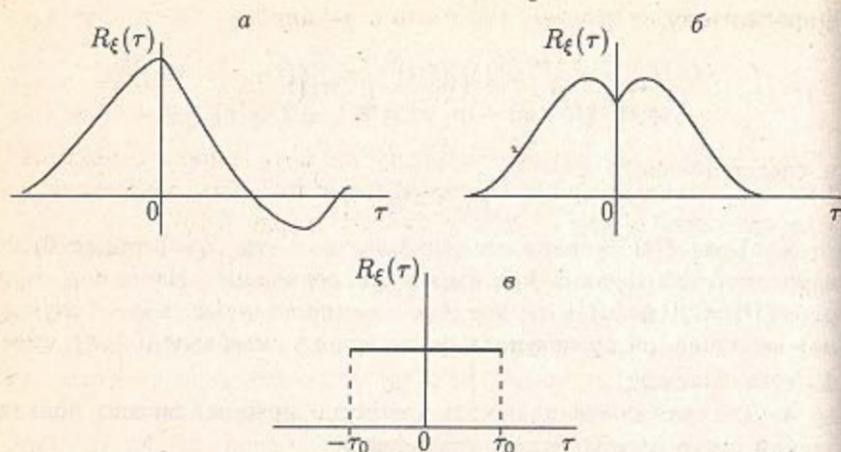


Рис. 6.9. Функции, которые не могут быть корреляционными функциями стационарного случайного процесса: а — не выполняется первое свойство; б — не выполняется второе свойство; в — не выполняется пятое свойство

Таким образом, автокорреляционная функция стационарного случайного процесса является четной функцией аргумента, имеет максимум, равный дисперсии при  $\tau = 0$ , и, как правило, убывает до нуля при  $\tau \rightarrow \infty$ .

Функции, которые не могут быть корреляционными функциями стационарного процесса, показаны на рис. 6.9.

Аналогичные свойства присущи и коэффициенту корреляции:

$$r_{\xi}(\tau) = r_{\xi}(-\tau); |r_{\xi}(\tau)| \leq r_{\xi}(0) = 1;$$

$$\lim_{\tau \rightarrow \infty} r_{\xi}(\tau) = 0; \int_{-\infty}^{\infty} r_{\xi}(\tau) e^{-j\omega\tau} d\tau \geq 0.$$

Вместо точного аналитического задания вида коэффициента корреляции на практике часто ограничиваются указанием лишь *интервала*, или *времени корреляции*  $\tau_k$ , который дает ориентировочное представление о том, на каком интервале времени в среднем имеет место заметная коррелированность между значениями случайного процесса, существенная для решаемой задачи.

Аналогично тому, как оценивается длительность импульса, время корреляции можно определить по-разному. Так, можно условиться

под временем корреляции  $\tau_k$  понимать величину

$$\tau_k = \frac{1}{2} \int_{-\infty}^{\infty} r_{\xi}(\tau) d\tau = \int_0^{\infty} r_{\xi}(\tau) d\tau.$$

Геометрически  $\tau_k$  равно основанию прямоугольника с высотой  $r_{\xi}(0) = 1$ , имеющего ту же площадь, что и площадь, заключенная между кривой  $r_{\xi}(\tau)$  при  $\tau > 0$  и осью абсцисс (рис. 6.10).

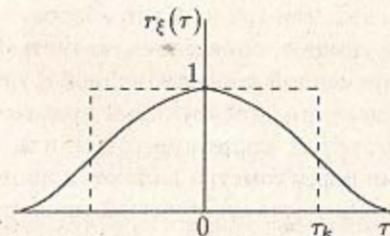


Рис. 6.10. К определению времени корреляции  $\tau_k$

## 6.8. Экспериментальное определение параметров процесса

Для стационарного случайного процесса, обладающего эргодическим свойством, можно указать простые методы экспериментального определения его основных характеристик. При этом используется тот факт, что эти характеристики могут быть определены путем временного усреднения одной достаточно длинной реализации.

Предположим, что время наблюдения  $T$  за стационарным эргодическим процессом  $\xi(t)$  достаточно велико и значительно превышает время корреляции  $\tau_k$  ( $T \gg \tau_k$ ). Тогда

$$m_{\xi} = m_{\xi T} = \frac{1}{T} \int_0^T \xi(t) dt, \quad \vdots$$

$$\sigma_{\xi}^2 = \sigma_{\xi T}^2 = \frac{1}{T} \int_0^T [\xi(t) - m_{\xi}]^2 dt,$$

$$R_{\xi}(\tau) = R_{\xi T}(\tau) = \frac{1}{T} \int_0^T (\xi(t) - m_{\xi})(\xi(t + \tau) - m_{\xi}) dt.$$

Если  $\xi(t)$  представляет собой флуктуационное напряжение (ток) в каком-либо радиоэлектронном устройстве, находящемся в стационарном состоянии, то в соответствии с первой формулой среднее значение равно постоянной составляющей напряжения (тока), которая экспериментально легко может быть определена при помощи соответствующих приборов магнитоэлектрической системы.

Дисперсия  $\sigma_\xi^2$  равна квадрату эффективного значения переменной составляющей напряжения (тока) и может быть определена при помощи термоэлектрических или тепловых приборов.

Корреляционную функцию, определяемую третьей формулой, часто называют кратковременной корреляционной функцией. Для экспериментального определения этой функции применяются специальные устройства, называемые коррелометрами или корреляторами. Основными элементами коррелометра являются линия задержки, перемножитель, интегратор и регистрирующий прибор. В зависимости от того, выполняется умножение цифровым методом или аналоговым, различают коррелометры дискретного или непрерывного действия. Простейшая функциональная схема коррелометра непрерывного действия приведена на рис. 6.11.

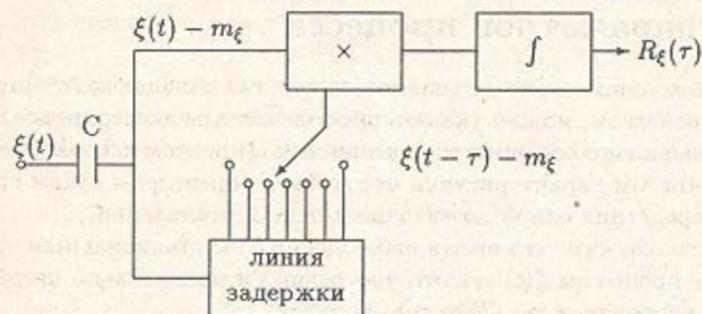


Рис. 6.11. Схема коррелометра непрерывного действия

Определение корреляционной функции выполняется по формуле

$$R_\xi(\tau) = \frac{1}{T - \tau} \int_{\tau}^T [\xi(t) - m_\xi][\xi(t - \tau) - m_\xi] dt.$$

Вычисление корреляционной функции производят последовательно, начиная с малых значений  $\tau$ , и продолжают до таких значений  $\tau$ ,

при которых она становится практически равной нулю. Общий вид функции  $R_\xi(\tau)$  воспроизводят по отдельным точкам (рис. 6.12).

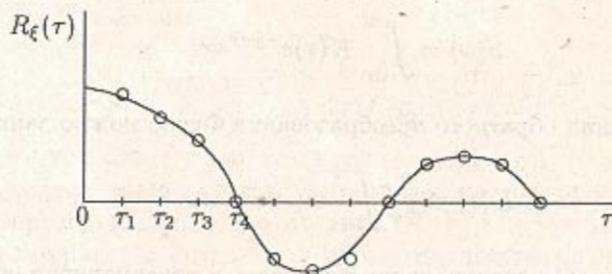


Рис. 6.12. Экспериментальная корреляционная функция

При подборе аналитической кривой для корреляционной функции следует не только руководствоваться необходимой точностью аппроксимации, но и иметь в виду, что корреляционная функция стационарного процесса должна удовлетворять рассмотренным ранее свойствам.

## 6.9. Спектральная плотность

При изучении детерминированных сигналов и реакции на них линейных систем с постоянными параметрами широко используются спектральные представления, базирующиеся на возможности представления (при определенных условиях) сигналов рядом или интегралом Фурье. При этом математически сравнительно просто и физически наглядно можно найти форму сигнала на выходе линейной системы простым пересчетом отдельных спектральных составляющих входного сигнала через комплексный коэффициент передачи системы и последующим применением принципа суперпозиции.

Представляется естественным желание распространить гармонический анализ на случайные процессы для решения, в принципе, одних типовых задач, хотя физический смысл результатов будет при этом несколько другим. Наиболее просто эта задача решается для стационарных случайных процессов путем введения спектральной плотности процесса.

Спектральная плотность  $S(\omega)$  стационарного в широком смысле случайного процесса  $\xi(t)$  определяется как преобразование Фурье от ковариационной функции:

$$S(\omega) = \int_{-\infty}^{\infty} K(\tau) e^{-j\omega\tau} d\tau. \quad (6.1)$$

На основании обратного преобразования Фурье можно записать:

$$K(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau} d\omega. \quad (6.2)$$

Таким образом, спектральная плотность и ковариационная функция стационарного случайного процесса представляют собой пару взаимных преобразований Фурье.

Аналогичным образом связаны между собой спектральная плотность  $S_0(\omega)$  и корреляционная функция  $R(\tau)$  стационарного в широком смысле центрированного случайного процесса  $\xi_0(t) = \xi(t) - m_\xi$ :

$$S_0(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau; \quad (6.3)$$

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) e^{j\omega\tau} d\omega. \quad (6.4)$$

Подставив в (6.1) выражение ковариационной функции через корреляционную, получим

$$S(\omega) = S_0(\omega) + m_\xi^2 \delta(\omega),$$

где  $\delta(\omega)$  —  $\delta$ -функция. Видно, что спектральная плотность стационарного процесса с отличным от нуля математическим ожиданием отличается от спектральной плотности соответствующего центрированного процесса лишь наличием дискретной линии на нулевой частоте, которая отражает наличие постоянной составляющей (математического ожидания) процесса. Имея это в виду, можно сказать, что формулы (6.1) — (6.4) являются равнозначными. Эти формулы были получены независимо друг от друга русским ученым А.Я.Хинчиным и американским ученым Н. Винером, и потому называются формулами Винера — Хинчина.

Поясним физический смысл спектральной плотности. Если под  $\xi(t)$  понимать случайный (флуктуационный) ток или напряжение, то величины  $S(\omega)$  и  $S_0(\omega)$  будут иметь размерность энергии. Полагая в (6.4)  $\tau = 0$ , получим

$$R(0) = D_\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) d\omega.$$

Если  $\xi(t)$  — флуктуационный ток (напряжение), то дисперсию  $D_\xi$  можно рассматривать как среднюю мощность, выделяемую этим током (напряжением) на сопротивлении в 1 Ом и численно равную площади под кривой спектральной плотности центрированного случайного процесса. Выражение  $\frac{1}{2\pi} S_0(\omega) d\omega$  можно трактовать как долю средней мощности, сосредоточенную в малом интервале частот — от  $\omega - \frac{d\omega}{2}$  до  $\omega + \frac{d\omega}{2}$ . Функции  $S(\omega)$  и  $S_0(\omega)$  иногда называют спектром мощности, или энергетическим спектром.

Перечислим основные свойства спектральной плотности:

1. Спектральная плотность — функция неотрицательная:

$$S(\omega) \geq 0.$$

Это свойство отражает тот факт, что мощность, приходящаяся на интервал частот  $(\omega - \frac{d\omega}{2}, \omega + \frac{d\omega}{2})$ , не может быть отрицательной, поскольку мощность — всегда положительная величина.

2. Спектральная плотность всегда является вещественной функцией, причем эта функция четная. Четность и вещественность спектральной плотности следуют из четности корреляционной функции  $R(\tau)$ :

$$S_0(-\omega) = \int_{-\infty}^{\infty} R(\tau) e^{j\omega\tau} d\tau = S_0(\omega).$$

Учитывая четность спектральной плотности, можно записать

$$S_0(\omega) = \int_{-\infty}^{\infty} R(\tau) \cos \omega\tau d\tau = 2 \int_0^{\infty} R(\tau) \cos \omega\tau d\tau;$$

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) \cos \omega\tau d\omega = \frac{2}{2\pi} \int_0^{\infty} S_0(\omega) \cos \omega\tau d\omega.$$

3. Корреляционная функция  $R(\tau)$  и спектральная плотность  $S_0(\omega)$  обладают всеми свойствами, характерными для пары взаимных преобразований Фурье. В частности, чем “шире” спектральная

плотность  $S_0(\omega)$ , тем "уже" корреляционная функция  $R(\tau)$ , и наоборот (рис. 6.13).

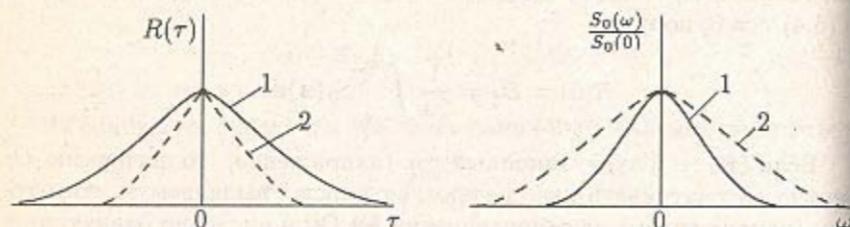


Рис. 6.13. Связь между "ширинами" корреляционной функции и спектральной плотности случайных процессов 1 и 2

Заметим, что во всех предыдущих формулах спектральная плотность  $S_0(\omega)$  определена для положительных и отрицательных значений частоты, причем  $S_0(\omega) = S_0(-\omega)$ . В отличие от такого "математического" спектра вводят односторонний "физический" спектр, отличный от нуля лишь при положительных частотах (рис. 6.14):

$$S_0(\omega)^+ = S_0(\omega) + S_0(-\omega) = 2S_0(\omega).$$

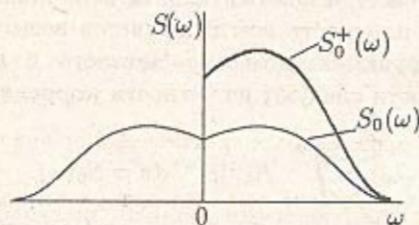


Рис. 6.14. "Математический"  $S_0(\omega)$  и "физический"  $S_0^+(\omega)$  спектры.

Тогда получим следующие формулы Винера — Хинчина:

$$S^+(\omega) = 4 \int_0^{\infty} K(\tau) \cos \omega \tau d\tau, \quad \omega \geq 0;$$

$$K(\tau) = \frac{1}{2\pi} \int_0^{\infty} S^+(\omega) \cos \omega \tau d\omega, \quad \omega \geq 0.$$

Формулами с бесконечными пределами целесообразно пользоваться при выполнении вычислений, так как интегралы в бесконечных пределах, как правило, находятся проще, чем интегралы с одним конечным пределом. При физическом рассмотрении и проведении экспериментов следует оперировать последними формулами.

"Протяженность" спектральной плотности по оси частот часто характеризуют *эффективной шириной спектра*. Ее можно определять по-разному. Например,

$$\Delta\omega_s = \frac{1}{S_0(\omega_0)} \int_0^{\infty} S_0(\omega) d\omega,$$

где  $S_0(\omega_0)$  — значение спектральной плотности при некоторой характерной частоте. Обычно берут за  $S_0(\omega_0)$  максимум спектральной плотности; иногда указывают ширину  $\Delta\omega_{0,5}$  спектральной плотности на уровне  $0,5S_0^+(\omega_0)$  (рис. 6.15).

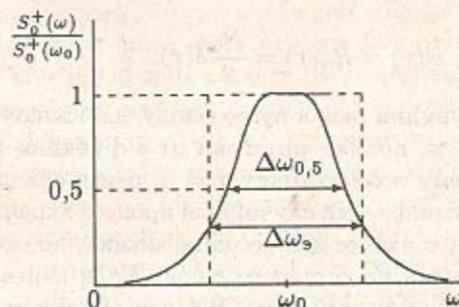


Рис. 6.15. К определению эффективной ширины спектра

В отличие от спектрального анализа детерминированных сигналов спектральная плотность случайного процесса не дает возможности восстановить какую-либо реализацию процесса, так как она является усредненной характеристикой и не содержит сведений о фазах отдельных спектральных составляющих. Можно указать несколько различных по характеру случайных процессов, имеющих одинаковую спектральную плотность и корреляционную функцию. Таким образом, эти две характеристики случайный процесс описывают явно неполно.

Аналогично спектральной плотности  $S_{\xi}(\omega)$  одного процесса  $\xi(t)$  взаимная спектральная плотность двух стационарно связанных

случайных процессов  $\xi(t)$  и  $\eta(t)$  определяется как прямое преобразование Фурье от взаимных ковариационных функций:

$$S_{\xi\eta}(\omega) = \int_{-\infty}^{\infty} K_{\xi\eta}(\tau) e^{-j\omega\tau} d\tau,$$

$$S_{\eta\xi}(\omega) = \int_{-\infty}^{\infty} K_{\eta\xi}(\tau) e^{-j\omega\tau} d\tau.$$

Обратное преобразование Фурье позволяет записать выражения для  $K_{\xi\eta}(\tau)$  и  $K_{\eta\xi}(\tau)$  через спектральные плотности  $S_{\xi\eta}(\omega)$  и  $S_{\eta\xi}(\omega)$ .

## 6.10. Белый шум

Рассмотрим центрированный случайный процесс  $n(t)$ , корреляционная функция которого равна  $\delta$ -функции, умноженной на некоторую постоянную величину  $\frac{N_0}{2}$  (рис. 6.16):

$$R(\tau) = R_n(\tau) = \frac{N_0}{2} \delta(\tau).$$

Как известно,  $\delta$ -функция равна нулю всюду, за исключением точки  $\tau = 0$ , где  $\delta(0) = \infty$ , причем интеграл от  $\delta$ -функции по любому интервалу, содержащему особую точку  $\tau = 0$ , равен единице. Отсюда следует, что рассматриваемый случайный процесс характеризуется тем, что значения  $n(t)$  в любые два несовпадающих, но сколь угодно близких момента времени не коррелированы. Если считать процесс  $n(t)$  нормальным, то они и независимы. В связи с этим такой процесс условно можно назвать абсолютно случайным процессом.

Найдем спектральную плотность рассматриваемого случайного процесса

$$S_0(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau = \frac{N_0}{2} = \text{const},$$

$$S_0^+(\omega) = N_0.$$

Таким образом, спектральная плотность постоянна на всех частотах (рис. 6.16).

Случайный процесс, обладающий равномерным спектром в очень широком диапазоне частот, принято называть *белым шумом* по аналогии с белым светом, имеющим в видимой части равномерный сплошной спектр.

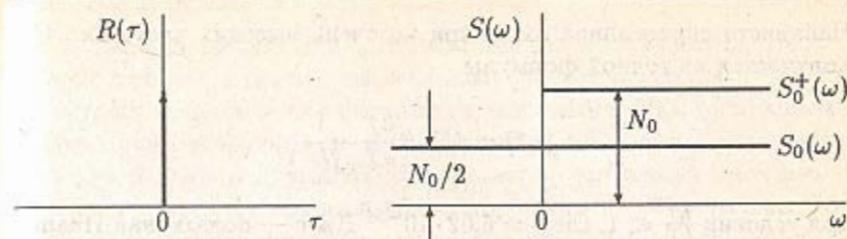


Рис. 6.16. Корреляционная функция и спектральная плотность белого шума

Для белого шума формула

$$D = \sigma^2 = R(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) d\omega$$

дает непонятный физический результат: дисперсия (средняя мощность) такого шума  $D = \infty$ , и для него нельзя аналитически записать даже одномерную плотность вероятности. Этот результат объясняется тем, что белый шум следует рассматривать как математическую идеализацию, так как все реальные процессы имеют спектральную плотность, убывающую при высоких частотах, и, следовательно, обладают конечным временем корреляции  $\tau_k \neq 0$  и ограниченной средней мощностью. Белый шум является полезной математической идеализацией, применяемой в тех случаях, когда в пределах амплитудно-частотной характеристики интересующей нас системы спектральную плотность внешнего воздействия реального шума можно приближенно считать постоянной или, иначе, когда время корреляции шума много меньше всех существенных постоянных времени системы, на которую воздействует шум. Рассмотрим пример флуктуационного шума, который можно рассматривать как белый шум.

Известно, что спектральная плотность напряжения теплового шума резистора  $R$  определяется формулой Найквиста:

$$S_0^+(f) = 4kTR = N_0,$$

где  $k = 1,38 \cdot 10^{-23}$  Дж/К — постоянная Больцмана;  $T$  — температура резистора в градусах Кельвина,  $f = \omega/2\pi$  — частота. Из этой формулы видно, что тепловой шум является, казалось бы, идеальным примером белого шума. Однако следует иметь в виду, что формула

Найквиста справедлива лишь при не очень высоких частотах. Она получается из точной формулы

$$S_0^+(f) = 4kTR \frac{hf}{kT} \frac{1}{e^{\frac{hf}{kT}} - 1}$$

при условии  $\frac{hf}{kT} \ll 1$ , где  $h = 6,62 \cdot 10^{-34}$  Дж·с — постоянная Планка.

При нормальной (комнатной) температуре даже на миллиметровых волнах неравенство  $\frac{hf}{kT} \ll 1$  практически выполняется. Поэтому в радиоэлектронике оправдано применение приближенной формулы для  $S_0^+(f)$ . Однако при вычислении дисперсии теплового шума необходимо пользоваться точной формулой. В результате получаем

$$D = 4kTR \int_0^\infty \frac{hf}{kT} \frac{1}{e^{\frac{hf}{kT}} - 1} df = 4kTR \frac{kT}{h} \int_0^\infty \frac{x dx}{e^x - 1} = \frac{2\pi^2}{3h} (kT)^2 R.$$

## 6.11. Нормальные случайные процессы

При рассмотрении способов описания случайных процессов указывалось, что случайный процесс в общем случае задается  $n$ -мерной плотностью вероятности, и тем детальнее, чем больше  $n$ .

Однако, многие наиболее часто встречающиеся на практике процессы оказываются не столь сложными, чтобы для получения о них полной информации требовалось знание  $n$ -мерных плотностей вероятностей. Часто бывает достаточно знать лишь одномерную  $p_1(x; t)$  или двумерную  $p_2(x_1, x_2; t_1, t_2)$  плотность вероятности. Если эти плотности вероятности содержат все сведения о процессе, то по ним могут быть найдены и многомерные плотности вероятности.

Примером таких случайных процессов являются нормальные случайные процессы.

Случайный процесс  $\xi(t)$  называется нормальным (гауссовским), если для любого конечного множества моментов времени  $t_1, t_2, \dots, t_n$  случайные величины  $\xi(t_1), \xi(t_2), \dots, \xi(t_n)$  имеют совместную нормальную плотность вероятностей.

Нормальные случайные процессы обладают рядом важных свойств, среди которых отметим следующие.

1. Нормальный случайный процесс полностью описывается указанием математического ожидания  $m_\xi(t)$  и корреляционной функции

$R_\xi(t)$ . Если на основании каких-либо соображений известно, что случайный процесс является нормальным, то его  $n$ -мерная плотность вероятности однозначно определяется математическим ожиданием и корреляционной функцией. Нормальные случайные процессы могут отличаться друг от друга только характером временной зависимости математического ожидания и видом корреляционной функции. Следовательно, корреляционная теория дает полное описание нормальных процессов.

2. Для нормальных случайных процессов некоррелированность значений процесса тождественна их независимости.

3. Для нормальных процессов понятия стационарности в широком и узком смыслах совпадают.

4. Условные плотности вероятности значений нормального случайного процесса являются нормальными.

5. При линейных преобразованиях нормальных случайных процессов получается также нормальный процесс.

какой-то продукции). В этих случаях наилучшим способом является тщательный отбор из генеральной совокупности части ее элементов (измерений), называемой *выборочной совокупностью* (или просто *выборкой*), и исследование ее свойств, а затем обобщение полученных результатов на всю генеральную совокупность. Число отобранных элементов (измерений) из генеральной совокупности называется *объемом выборки*.

Основное требование к выборке — хорошо представлять генеральную совокупность, то есть быть *представительной (репрезентативной)*. Это обеспечивается следующими условиями: случайностью выбора объектов (результатов измерений) из генеральной совокупности, когда каждому из них обеспечивается одинаковая возможность быть отобранным (включенным в выборку); независимостью результатов наблюдений в выборке; правильным определением объема выборки с учетом всех конкретных условий.

Различают *повторные* и *бесповторные* случайные выборки. При повторной выборке каждый ее элемент следует возвращать в генеральную совокупность перед извлечением следующего. В случае бесповторной выборки выбранные элементы не возвращаются в генеральную совокупность и, следовательно, могут появиться в ней не более одного раза. Естественно, что при ограниченном объеме генеральной совокупности абсолютно случайная выборка может быть образована только при повторной выборке. Однако если генеральная совокупность бесконечна или велика по сравнению с объемом выборки, то оба вида выборки дают случайную выборку.

На практике наиболее часто используется бесповторная выборка, поскольку повторная выборка может быть, в принципе, нереализуемой в связи с тем, что элемент ее в процессе наблюдений может разрушаться.

## 7.2. Основные задачи математической статистики

Методы математической статистики применяются в самых различных областях науки (физике, биологии, медицине, экономике и др.) и могут преследовать различные цели. Однако как основные можно выделить следующие три задачи.

## Глава 7

# Основные понятия математической статистики

Математическая статистика — раздел математики, посвященный установлению закономерностей случайных явлений или процессов на основании систематизации обработки экспериментальных результатов. Подобно любой отрасли знаний математическая статистика имеет свою терминологию. Так, например, физики говорят “определить”, а статистики — “оценить”, физики говорят “оценить”, а статистики — “угадать”. Таким образом, слово “оценить” в физике и статистике имеет различный смысл. В дальнейшем это слово будет применяться так, как это принято в статистике.

## 7.1. Генеральная и выборочная совокупности

*Генеральной совокупностью (популяцией)* называется вся совокупность подлежащих изучению объектов или возможных результатов наблюдений над одним объектом.

Число изучаемых объектов, составляющих генеральную совокупность, или число возможных результатов наблюдений над одним объектом называется *объемом генеральной совокупности*. Объем генеральной совокупности может быть конечным или бесконечным. Если объем генеральной совокупности конечен, то наилучшим способом отыскания ее свойств является рассмотрение каждого объекта (элемента) этой совокупности. Например, при переписи населения требуется учет всех жителей страны.

Однако в большинстве задач математической статистики приходится рассматривать бесконечные генеральные совокупности (например, при проведении измерений) или конечные, но с большими объемами (например, при массовом промышленном производстве

1. Оценка неизвестной функции распределения или плотности вероятности. Эта задача обычно формулируется так. В результате независимых наблюдений случайной величины  $X$  получены следующие ее конкретные значения  $x_1, x_2, \dots, x_n$ . Требуется оценить неизвестную функцию распределения  $F(x)$  случайной величины  $X$  или ее плотность вероятности  $p(x)$ .

2. Оценка неизвестных параметров закона распределения. В этой задаче предполагается, что на основании физических или общетеоретических соображений можно считать, что случайная величина  $X$  имеет плотность вероятности определенного вида, зависящую от одного или нескольких параметров, значения которых неизвестны. По результатам наблюдений случайной величины  $X$  нужно оценить значения этих параметров.

3. Проверка статистических гипотез. Задачи, связанные с проверкой *статистических гипотез*, весьма обширны. Например, по результатам экспериментальных наблюдений (выборке) следует проверить заключение о том, что они принадлежат предполагаемой генеральной совокупности, или о том, что полученная оценка неизвестного параметра закона распределения не противоречит выдвинутой гипотезе о значении данного параметра.

Из определений и смысла основных величин, фигурирующих в теории вероятностей, следуют и методы их определения по экспериментальным результатам. Так, за вероятность события можно принять его относительную частоту при большом числе наблюдений, за математическое ожидание случайной величины — ее среднее арифметическое значение и т.д. Однако число наблюдений практически всегда ограничено, причем результаты наблюдений случайны. В связи с этим в математической статистике говорят о подходящих *оценках* интересующих величин.

Все параметры (числовые характеристики), подлежащие определению по результатам наблюдений, принято называть статистическими параметрами (характеристиками), а любая функция результатов наблюдений, которая может быть принята за подходящее значение неизвестного статистического параметра (характеристики), называется оценкой этого статистического параметра (характеристики).

Разработка методов нахождения оценок и исследования точности их приближения к неизвестным статистическим параметрам (характеристикам) является важным разделом математической статистики.

## Глава 8

# Оценка закона распределения

## 8.1. Гистограмма распределения

Чтобы получить представление о законе распределения случайной величины на основе наблюдения ее значений, поступают следующим образом. Если изучается непрерывная случайная величина, то область ее экспериментальных значений разбивается на  $m$  обычно одинаковых интервалов длиной  $\Delta x$  и вычисляется относительная плотность точек в каждом интервале (отношение относительной частоты попадания в этот интервал к его длине  $\Delta x$ ):

$$p_k^* = \frac{v_k^*}{\Delta x} = \frac{n_k}{n \Delta x}, \quad k = 1, 2, \dots, m,$$

где  $n_k$  — число экспериментальных точек в  $k$ -м интервале,  $n$  — общее число точек.

Подсчитанные таким образом значения можно представить графически в виде ступенчатой кривой: по оси абсцисс откладывают соответствующие интервалы и на каждом из них, как на основании, строится прямоугольник, высота которого равна относительной плотности  $p_k^*$ . Полученная ступенчатая кривая называется *гистограммой*.

Пусть, например, по результатам наблюдений случайной величины  $X$  оказалось, что 12 значений попали в интервал (3, 4), 14 — в интервал (4, 5), 17 — в интервал (5, 6), 20 — в интервал (6, 7), 14 — в интервал (7, 8), 12 — в интервал (8, 9), 6 — в интервал (9, 10) и 5 — в интервал (10, 11). Тогда, вычисляя соответствующие относительные плотности с учетом того, что  $\Delta x = 1$ , а  $n = 100$ , получим гистограмму выборки (рис. 8.1).

Общее правило построения гистограммы распределения выглядит следующим образом:

отмечаются наименьшее  $x_{\min}$  и наибольшее  $x_{\max}$  значения в выборке объема  $n$ ;

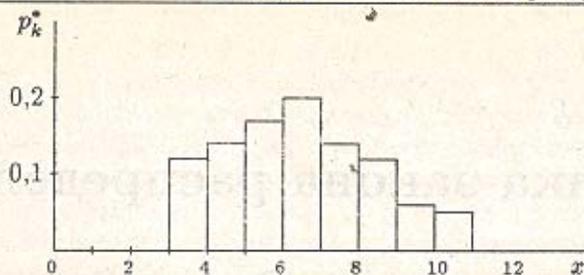


Рис. 8.1. Гистограмма выборки

весь диапазон значений от  $x_{\min}$  до  $x_{\max}$  разбивается на  $m$  равных интервалов, (при этом количество интервалов  $m$  не должно быть меньше 8 — 10 и больше 20 — 25);

отмечаются крайние точки каждого из интервалов  $c_0, c_1, c_2, \dots, c_m$  в порядке их возрастания, а также их середины  $x_1^0, x_2^0, \dots, x_m^0$ ;

подсчитываются числа выборочных данных, попавших в каждый из интервалов:  $n_1, n_2, \dots, n_m$  (очевидно, что  $n_1 + n_2 + \dots + n_m = n$ ).

Выборочные данные, попавшие на границы интервалов, либо равномерно распределяют по двум соседним интервалам, либо относят их только к какому-либо одному из них — например, к левому.

Выбор количества интервалов существенно зависит от объема выборки. В литературе приводятся рекомендации по выбору числа интервалов, однако они существенно отличаются одна от другой. Так, многие руководства по статистике рекомендуют использовать формулу Старджеса:

$$m \simeq \log_2 n + 1 \simeq 3,32 \ln n + 1.$$

Есть также и другие формулы:

$$m \simeq 5 \lg n;$$

$$m \simeq \sqrt{n}.$$

Все эти формулы следует рассматривать скорее как оценку снизу для  $m$  (особенно при больших  $n$ ). Так, для рассмотренного выше примера построения гистограммы по формуле Старджеса будем иметь:

$$m = 3,32 \ln 100 + 1 = 3,32 \cdot 4,6052 + 1 \approx 16;$$

две другие формулы дают  $m = 10$ .

В зависимости от конкретного содержания задачи в данную схему построения гистограммы могут быть внесены некоторые изменения. Например, в некоторых случаях целесообразно отказаться от требования равной длины интервалов разбиения выборочных данных.

Аналогично строится гистограмма в двумерном случае, когда рассматривается распределение выборочных данных на плоскости. Разбив часть плоскости, занятую экспериментальными точками, на прямоугольники и подсчитав число точек в каждом из них, можно определить соответствующие относительные плотности точек как относительные частоты попадания точек в прямоугольник, деленные на его площадь.

На практике часто возникает необходимость аппроксимации полученной гистограммы подходящим аналитическим выражением, представляющим собой некоторый теоретический закон распределения или плотность вероятности. При этом стремятся к тому, чтобы такая аппроксимация была в определенном смысле наилучшей.

Существует много разнообразных способов и приемов подбора распределений для экспериментальных данных, и выделить какой-либо из них невозможно. Успех в значительной степени зависит от накопленного в этом отношении опыта. Однако некоторые общие рекомендации можно дать.

Обычно аппроксимация гистограммы не является самоцелью, а производится для получения каких-либо выводов о физическом механизме изучаемого явления или выполнения последующих расчетов. Исходя из этого, прежде всего необходимо принять решение — аппроксимировать гистограмму дискретным или непрерывным распределением (плотностью вероятности). После этого производят качественное сопоставление характера построенной гистограммы с графиками различных теоретических распределений (дискретных или непрерывных) и исходя из близости их поведения, останавливаются на каком-либо одном из наиболее подходящих.

В зависимости от условий решаемой задачи характер и степень близости поведения следует понимать по-разному. Иногда можно ограничиться хорошим совпадением в центральной области гистограммы (в области больших вероятностей), а иногда нужно стремиться к хорошему совпадению на «хвостах» распределения (в области малых вероятностей).

Чтобы оценить, насколько хорошо выбранный теоретический закон распределения (плотность вероятности) согласуется с результатами наблюдений, пользуются так называемыми критериями согласия. Таких критериев существует несколько. Наиболее часто применяют критерий согласия  $\chi^2$ , предложенный Пирсоном.

## 8.2. Критерий согласия $\chi^2$

Пусть результаты  $n$  независимых наблюдений случайной величины  $X$  сгруппированы по  $m$  непересекающимся интервалам  $(c_0, c_1)$ ,  $(c_1, c_2)$ , ...,  $(c_{m-1}, c_m)$  и подсчитаны относительные частоты  $\nu_k^*$  попадания результатов в соответствующие интервалы. Пусть случайная величина  $X$  имеет теоретическую функцию распределения  $F(x)$ . В качестве меры расхождения между теоретическим и эмпирическим распределениями Пирсон предложил взять величину

$$\chi^2 = \sum_{k=1}^m \frac{(n_k - np_k)^2}{np_k} = n \sum_{k=1}^m \left( \frac{n_k}{n} - p_k \right)^2 \frac{1}{p_k} = n \sum_{k=1}^m \frac{(\nu_k^* - p_k)^2}{p_k},$$

где  $p_k$  — теоретическая вероятность нахождения значений случайной величины  $X$  в  $k$ -м интервале группировки. Эта вероятность определяется по формуле

$$p_k = F(c_{k+1}) - F(c_k),$$

где  $F(c_{k+1})$  и  $F(c_k)$  — соответственно значения теоретической функции распределения в точках  $c_{k+1}$  и  $c_k$ .

При достаточно большом числе  $n$  практически независимо от вида закона распределения исследуемой случайной величины  $X$  плотность вероятности случайной величины  $\chi^2$  описывается законом распределения  $\chi^2$  с числом степеней свободы  $m - 1$ , так как на частоты  $n_k$  наложена одна связь:  $n = \sum_{k=1}^m n_k$ . Поскольку на практике параметры теоретического распределения обычно неизвестны, то их приходится оценивать по результатам наблюдений. Если оценке подлежат  $s$  параметров (например, в случае нормального распределения их два), то число степеней свободы величины  $\chi^2$  будет равным  $l = m - 1 - s$ .

Поясним сущность изложенного. Значения  $x_1, x_2, \dots, x_n$ , полученные в выборке, следует рассматривать как наблюдаемое «значение»  $n$ -мерной случайной величины  $(X_1, X_2, \dots, X_n)$ , где каждая величина  $X_i$  представляет как бы одно значение величины  $X$ , с которым мы встречаемся в  $i$ -м наблюдении. В связи с этим выборку можно

рассматривать как испытание, в котором осуществляется конкретная реализация величины  $(X_1, X_2, \dots, X_n)$ . Естественно, что, повторяя выборки, мы будем получать различные значения этой величины. В дальнейшем следует иметь в виду, что любые статистические характеристики, определяемые по результатам выборки, являются случайными, в то время как вероятностные характеристики — фиксированными (хотя и неизвестными) постоянными.

В соответствии с этим при случайных результатах опыта относительные частоты  $\nu_1^*, \nu_2^*, \dots, \nu_m^*$  следует трактовать как случайные величины. Если некоторые параметры теоретического распределения оцениваются по данным наблюдений, то и вероятности  $p_1, p_2, \dots, p_m$  будут случайными величинами. Поэтому величину

$$\chi^2 = n \sum_{k=1}^m \frac{(\nu_k^* - p_k)^2}{p_k}$$

следует рассматривать как случайную, а ее значение — как конкретное значение случайной величины, реализованное в фактически полученной выборке.

Методика применения критерия  $\chi^2$  для оценки расхождения теоретического и эмпирического распределений сводится к следующему:

- 1) на основании выборочных данных  $x_1, x_2, \dots, x_n$  находят оценки параметров теоретического распределения;
- 2) вычисляются, исходя из теоретического распределения, вероятности  $p_k$  попадания значений случайной величины в интервалы  $\Delta c_k = c_{k+1} - c_k$ ;
- 3) рассчитывается значение  $\chi^2$ ;
- 4) определяется число степеней свободы  $l = m - s - 1$ ;
- 5) в зависимости от характера задачи назначается достаточно малая вероятность  $\alpha$ , называемая *уровнем значимости*. Считается, что событие с такой вероятностью является практически невозможным. Обычно берут  $\alpha = 0,05$  или  $0,01$  (5 или 1%);
- 6) по таблицам процентных точек распределения  $\chi^2$  находят значение процентной точки  $\chi_{l, \alpha}^2$ , определяемое уравнением

$$P(\chi^2 > \chi_{l, \alpha}^2) = \int_{\chi_{l, \alpha}^2}^{\infty} p(\chi^2) d(\chi^2);$$

- 7) если вычисленное значение  $\chi^2$  больше  $\chi_{l, \alpha}^2$ , то теоретическое распределение считается плохо согласующимся с результатами наблюде-

ний при уровне значимости  $\alpha$ . Если же вычисленное значение  $\chi^2$  оказывается меньше  $\chi_{\alpha}^2$ , то принимается, что выбранное теоретическое распределение согласуется с результатами наблюдений.

График, поясняющий определение процентной точки распределения  $\chi^2$ , приведен на рис. 8.2.

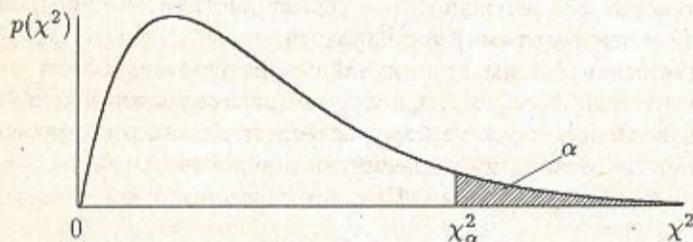


Рис. 8.2. К определению процентной точки распределения  $\chi^2$

Разумеется, решение зависит от объема выборки  $n$ , поскольку он определяет число степеней свободы, а также от числа интервалов группировки  $m$  и уровня значимости  $\alpha$ .

При использовании критериев согласия положительный ответ нельзя рассматривать как утвердительный о правильно выбранной модели распределения. Определенным является лишь отрицательный ответ.

Для того, чтобы введенная величина  $\chi^2$  имела распределение  $\chi^2$ , теоретические вероятности должны быть не слишком малыми. Для всех  $\Delta s_k$  должно выполняться условие  $np_k \geq 5$ . Если для некоторых  $\Delta s_k$  это условие нарушается, то соседние интервалы следует объединить в один.

Проиллюстрируем применение критерия  $\chi^2$  на двух примерах. В первом примере рассмотрим случай, когда выдвигается гипотеза о том, что исследуемая случайная величина подчиняется закону Пуассона, во втором — нормальному закону распределения.

**Пример 1.** Радиоактивное вещество наблюдалось в течение 60 равных интервалов времени по 5 с каждый. Для каждого из этих интервалов регистрировалось число частиц, попавших в счетчик. Числа интервалов времени  $n_k$ , в течение которых в счетчик попало ровно  $k$  частиц, приведены в следующей таблице:

$k$	0	1	2	3	4	5	6	7
$n_k$	8	17	16	10	6	2	0	1

Таким образом, общее число зарегистрированных частиц

$$n = \sum_k n_k = 60.$$

Используя критерий  $\chi^2$ , необходимо проверить гипотезу о согласии наблюдаемых данных с законом Пуассона

$$P(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda},$$

взяв уровень значимости равным 0,05.

Поскольку в распределении Пуассона параметр  $\lambda$  неизвестен, его требуется оценить по имеющимся данным. Этот параметр является математическим ожиданием, а его оценкой, как будет показано далее, может служить величина

$$\hat{\lambda} = \frac{1}{n} \sum_{k=0}^7 kn_k = 2.$$

Вычисляем теоретические вероятности  $p_k$  попадания в счетчик  $k$  частиц при справедливости закона Пуассона:

$$\begin{aligned} P(0) = p_0 &= \frac{2^0}{0!} \cdot e^{-2} = 0,1353; & P(1) = p_1 &= \frac{2^1}{1!} \cdot e^{-2} = 0,2707; \\ P(2) = p_2 &= \frac{2^2}{2!} \cdot e^{-2} = 0,2707; & P(3) = p_3 &= \frac{2^3}{3!} \cdot e^{-2} = 0,1804; \\ P(4) = p_4 &= 0,0902; & P(5) = p_5 &= 0,0361; & P(6) = p_6 &= 0,0120; \\ & & P(7) = p_7 &= 0,0034. \end{aligned}$$

Вычисляем значение  $\chi^2$ :

$$\chi^2 = \sum_{k=0}^7 \frac{(n_k - np_k)^2}{np_k} = \frac{(8 - 60 \cdot 0,1353)^2}{60 \cdot 0,1353} + \frac{(17 - 60 \cdot 0,2707)^2}{60 \cdot 0,2707} + \dots \approx 0,2.$$

Вычисляем число степеней свободы  $l = m - 1 - s$ . В нашем случае число интервалов  $m = 8$ , а число параметров распределения, оцененных по наблюдаемым данным,  $s = 1$ . Тогда  $l = 8 - 1 - 1 = 6$ .

По таблицам процентных точек распределения  $\chi^2$  с числом степеней свободы 6 и  $\alpha = 0,05$  находим  $\chi_{\alpha}^2 = 12,59$ . Поскольку  $\chi^2 < \chi_{\alpha}^2$ ,

то наблюдаемые данные согласуются с распределением Пуассона при уровне значимости 0,05.

*Пример 2.* Измерены диаметры 100 изготовленных деталей и определены их отклонения от заданного размера. Минимальное и максимальное отклонения составили соответственно -2,5 и 5,0. В соответствии с формулой Старджеса диапазон отклонений разбит на 8 равных интервалов и подсчитано количество наблюдений, попавших в каждый интервал. Данные представлены в таблице:

Границы интервалов	-3 ÷ -2	-2 ÷ -1	-1 ÷ 0	0 ÷ 1	1 ÷ 2	2 ÷ 3	3 ÷ 4	4 ÷ 5
$n_k$	3	10	15	24	25	13	7	3

Требуется при уровне значимости  $\alpha = 0,01$  проверить гипотезу о том, что отклонения размеров деталей от заданных согласуются с нормальным законом распределения.

Обратим внимание на то, что число наблюдений в крайних интервалах мало. Поэтому эти интервалы следует объединить с соседними. В результате объединения получим следующую таблицу:

Границы интервалов	-3 ÷ -1	-1 ÷ 0	0 ÷ 1	1 ÷ 2	2 ÷ 3	3 ÷ 5
$n_k$	13	15	24	25	13	10

Чтобы найти теоретические вероятности  $p_k$  попадания случайных отклонений  $X$  в соответствующие интервалы, необходимо знать параметры предполагаемого нормального распределения: математическое ожидание и дисперсию. Как будет показано позже, оценки этих параметров могут быть сделаны по формулам:

$$m_x^* = \frac{1}{n} \sum_{k=1}^m x_k^0 n_k;$$

$$D_x^* = \frac{1}{n} \sum_{k=1}^m (x_k^0 - m_x^*)^2 n_k,$$

где  $x_k^0$  — середины соответствующих интервалов. В формуле предполагается, что интервалы имеют одинаковую длину, поэтому следует воспользоваться первоначальной таблицей. Расчеты дают:

$$m_x^* = 0,6; D_x^* = 2,53; \sigma_x^* = 1,6.$$

Рассмотрим вычисление вероятностей  $p_k$ . Для первого интервала имеем:

$$\begin{aligned} p_1 &= P(-3 < x < -1) = P\left(\frac{-3-0,6}{1,6} < \frac{x-m_x}{\sigma} < \frac{-1-0,6}{1,6}\right) = \\ &= P(-2,25 < \frac{x-m_x}{\sigma} < -1,00) = \Phi(-1,00) - \Phi(-2,25). \end{aligned}$$

По таблице интеграла вероятности (таблица 1 Приложения) находим:

$$\Phi(-1,00) = 1 - \Phi(1,00) = 1 - 0,8413;$$

$$\Phi(-2,25) = 1 - \Phi(2,25) = 1 - 0,9878,$$

отсюда  $p_1 = 0,1465$ .

Аналогично вычисляются другие значения  $p_k$ :

$$p_2 = 0,1933; p_3 = 0,2467; p_4 = 0,2119;$$

$$p_5 = 0,1296; p_6 = 0,0638.$$

Теперь вычисляем величину  $\chi^2$ :

$$\chi^2 = \frac{13 - 14,64}{14,64} + \frac{14 - 19,33}{19,33} + \dots + \frac{10 - 6,38}{6,38} = 5,53.$$

Количество интервалов равно 6. По выборке оценивались два параметра, следовательно, число степеней свободы  $l = 3$ . По таблице процентных точек распределения  $\chi^2$  находим

$$\chi_{3,0,01}^2 = 11,34.$$

Поскольку  $\chi^2 < \chi_{3,0,01}^2$ , то нет оснований отвергнуть проверяемую гипотезу.

## Глава 9

# Оценка моментов и параметров распределения

Допустим, что по результатам выборочных значений  $x_1, x_2, \dots, x_n$  случайной величины  $X$  нужно оценить ее числовые характеристики (например, математическое ожидание, дисперсию или другие моменты) или неизвестный параметр  $a$ , от которого зависит закон распределения (плотность вероятности)  $p(x; a)$  этой случайной величины. Подлежащих оценке параметров распределения может быть несколько. Так, в законе Пуассона достаточно оценить один параметр, а в нормальном законе может понадобиться оценка и двух параметров — математического ожидания и дисперсии.

### 9.1. Виды оценок и их характеристики

Пусть из генеральной совокупности с плотностью вероятности  $p(x; a)$ , где  $a$  — неизвестный параметр, извлечена выборка объема  $n$  и получены результаты  $x_1, x_2, \dots, x_n$ . Получить оценку по существу означает, что каждой выборке нужно поставить в соответствие некоторое значение параметра  $a$ , то есть необходимо образовать некоторую функцию результатов наблюдений  $U(x_1, \dots, x_n)$ , значение которой принимается за оценку  $\hat{a}_n$  параметра  $a$ . Индекс  $n$  указывает объем выборки.

Любая функция, зависящая от результатов наблюдений, называется *статистикой*. Так как результаты наблюдений являются случайными величинами, то статистика тоже будет случайной величиной. Например, одной из оценок математического ожидания является арифметическое среднее:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Если повторить наблюдение (опыт)  $k$  раз, извлекая из генеральной совокупности случайные выборки одного и того же объема, то по полученным результатам можно найти ряд чисел  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ , которые, вообще говоря, будут отличаться друг от друга.

Таким образом, оценку  $\hat{a} = U(x_1, \dots, x_n)$  неизвестного параметра  $a$  следует рассматривать как случайную величину, а ее значения, вычисленные по данной выборке объема  $n$ , — как одну из реализаций этой случайной величины, то есть как одно из множества ее возможных значений.

Оценки параметров распределений (числовых характеристик случайной величины) подразделяются на точечные и интервальные. *Точечная оценка* параметра  $a$  определяется одним числом  $\hat{a}$ , и ее точность характеризуют дисперсией оценки. *Интервальной оценкой* называют оценку, которая определяется двумя числами  $\hat{a}_1$  и  $\hat{a}_2$  — концами интервала, накрывающего оцениваемый параметр  $a$  с заданной вероятностью.

Рассмотрим требования, предъявляемые к точечным оценкам.

#### 9.1.1. Состоятельность оценки

Оценка  $\hat{a}_n$  параметра  $a$  называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру при увеличении объема выборки, то есть

$$\lim_{n \rightarrow \infty} P(|\hat{a}_n - a| < \varepsilon) = 1,$$

где  $\varepsilon > 0$  — сколь угодно малое число.

На основании неравенства Чебышева можно показать, что достаточное условие выполнения приведенного соотношения заключается в том, что

$$\lim_{n \rightarrow \infty} M[(\hat{a}_n - a)^2] = 0.$$

Состоятельность оценки означает, что при достаточно большом объеме выборки со сколь угодно большой достоверностью отклонение оценки от истинного значения параметра меньше наперед заданной величины. Состоятельность является лишь асимптотической характеристикой оценки при  $n \rightarrow \infty$ .

## 9.1.2. Несмещенность оценки

Несмещенность оценки означает, что для всех  $n$  математическое ожидание оценки должно быть равно оцениваемому параметру, то есть

$$M[\hat{a}_n] = a. \quad (9.1)$$

Если это равенство не удовлетворяется, то оценка называется *смещенной*. Разность  $\Delta\hat{a} = M[\hat{a}_n] - a$  называется смещением, или систематической ошибкой оценки. Если равенство (9.1) выполняется лишь в пределе при  $n \rightarrow \infty$ , то соответствующая оценка называется *асимптотически несмещенной*.

В общем случае точность оценки некоторого параметра  $a$  на основании выборки обычно характеризуется средним значением квадрата ошибки

$$\varepsilon^2 = M[(\hat{a}_n - a)^2].$$

Запишем это выражение в виде

$$\begin{aligned} M[(\hat{a}_n - a)^2] &= M[(\hat{a}_n - M[\hat{a}_n] + M[\hat{a}_n] - a)^2] = \\ &= M[(\hat{a}_n - M[\hat{a}_n])^2] + 2M[(\hat{a}_n - M[\hat{a}_n])(M[\hat{a}_n] - a)] + M[(M[\hat{a}_n] - a)^2]. \end{aligned}$$

Если параметр  $a$  в процессе наблюдений остается постоянным, то перед последним слагаемым в правой части операцию математического ожидания можно опустить, а среднее слагаемое равно нулю, так как

$$M[\hat{a}_n - M[\hat{a}_n]] = M[\hat{a}_n] - M[\hat{a}_n] = 0.$$

Поэтому

$$\varepsilon^2 = M[(\hat{a}_n - M[\hat{a}_n])^2] + (M[\hat{a}_n] - a)^2.$$

Как видно, среднее значение квадрата ошибки состоит из двух слагаемых. Первое представляет собой дисперсию оценки

$$D_{\hat{a}} = M[(\hat{a}_n - M[\hat{a}_n])^2] = M[\hat{a}_n^2] - (M[\hat{a}_n])^2,$$

характеризующую долю "случайности" в величине ошибки, второе является квадратом смещения оценки

$$\Delta_{\hat{a}}^2 = (M[\hat{a}_n] - a)^2.$$

Итак, среднее значение квадрата ошибки равно сумме дисперсии оценки и квадрата смещения оценки:

$$\varepsilon^2 = M[(\hat{a}_n - a)^2] = D_{\hat{a}} + (\Delta\hat{a})^2.$$

Если оценка несмещенная, то

$$\varepsilon^2 = D_{\hat{a}} = M[(\hat{a}_n - a)^2].$$

## 9.1.3. Достаточность оценки

Оценка является *достаточной статистикой*, если вся полученная из выборки информация относительно параметра  $a$  содержится в  $\hat{a}_n$ . Если известна достаточная статистика, то никакая иная статистика, вычисленная по той же выборке, не может дать дополнительной информации, касающейся параметра  $a$ .

## 9.1.4. Эффективность оценки

При ограниченном объеме выборки оценки могут различаться средним квадратом ошибки  $\varepsilon^2$ . Очевидно, чем меньше эта ошибка, тем теснее сгруппированы значения оценки около оцениваемого параметра. Поэтому всегда желательно, чтобы ошибка оценки была по возможности малой, то есть выполнялось условие

$$\varepsilon^2 = M[(\hat{a} - a)^2] = \min_a.$$

Оценку  $\hat{a}$ , удовлетворяющую этому условию, можно назвать оценкой с минимальным средним квадратом ошибки. Если оценка  $\hat{a}$  несмещенная, то величина  $\varepsilon^2$  совпадает с дисперсией оценки  $D_{\hat{a}}$ . Несмещенная оценка с минимальной дисперсией является наиболее важной.

Доказано, что дисперсия любой несмещенной оценки не может быть меньше нижней границы, определяемой неравенством Рао — Крамера:

$$D_{\hat{a}} \geq \frac{1}{M\left[\frac{\partial}{\partial a} \ln p(x_1, \dots, x_n | a)\right]} = \frac{-1}{M\left[\frac{\partial^2}{\partial a^2} \ln p(x_1, \dots, x_n | a)\right]},$$

где  $p(x_1, \dots, x_n | a)$  — совместная плотность вероятности выборочных значений при истинном значении параметра  $a$ . Несмещенная оценка,

имеющая нижнюю граничную дисперсию, называется *эффективной*, то есть для эффективной оценки неравенство Рао — Крамера превращается в равенство.

Для существования эффективной оценки необходимо выполнение двух условий. Во-первых, должно выполняться условие достаточности оценки; во-вторых, производная от  $\ln p(x_1, \dots, x_n | a)$  по параметру  $a$  должна удовлетворять равенству

$$\frac{\partial}{\partial a} \ln p(x_1, \dots, x_n | a) = k(a)(\hat{a} - a),$$

где  $k(a)$  — функция, не зависящая от выборки и, следовательно, от  $\hat{a}$ . В неравенстве Рао — Крамера математическое ожидание берется по переменной  $x = (x_1, \dots, x_n)$ , а производные — в точке истинного значения параметра  $a$ .

Поскольку истинное значение параметра  $a$  неизвестно, то с точностью до величин более высокого порядка на практике можно пользоваться следующим выражением для дисперсии эффективной оценки:

$$D_{\hat{a}} \approx \frac{-1}{\frac{\partial^2}{\partial a^2} \ln p(x_1, \dots, x_n | a)} \Big|_{a=\hat{a}}$$

Точность выражения возрастает с увеличением объема выборки.

Когда оценивается несколько параметров  $a_1, a_2, \dots, a_k$  распределения, то оценки являются несмещенными, если нулевое смещение имеет каждый параметр. Условием совместно эффективных оценок является выполнение равенства

$$\hat{a}_i - a = \sum_{j=1}^k K_{ij}(a_1, \dots, a_k) \frac{\partial}{\partial a_j} \ln p(x_1, \dots, x_n | a_1, \dots, a_k),$$

где функции  $K_{ij}(a_1, \dots, a_k)$  не зависят от оценок  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$ . При оценке нескольких параметров величиной, аналогичной дисперсии, является корреляционная матрица  $R$  ошибок несмещенных оценок, состоящая из элементов

$$R_{ij} = M[(\hat{a}_i - a_i)(\hat{a}_j - a_j)], \quad i, j = 1, 2, \dots, k.$$

Нижняя граница корреляционной матрицы несмещенных оценок имеет вид

$$R = J^{-1},$$

где  $J^{-1}$  — матрица, обратная так называемой *информационной матрице Фишера* с элементами

$$\begin{aligned} J_{ij} &= M \left[ \frac{\partial \ln p(x_1, \dots, x_n | a_1, \dots, a_k)}{\partial a_i} \cdot \frac{\partial \ln p(x_1, \dots, x_n | a_1, \dots, a_k)}{\partial a_j} \right] = \\ &= -M \left[ \frac{\partial^2 \ln p(x_1, \dots, x_n | a_1, \dots, a_k)}{\partial a_i \partial a_j} \right]. \end{aligned}$$

Диагональные элементы матрицы  $R$  представляют собой дисперсии оценок соответствующих параметров. Применительно к оценке двух параметров  $a_1$  и  $a_2$  информационная матрица Фишера имеет вид

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \begin{bmatrix} -M \left[ \frac{\partial^2 \ln p(x_1, \dots, x_n | a_1, a_2)}{\partial a_1^2} \right] & -M \left[ \frac{\partial^2 \ln p(x_1, \dots, x_n | a_1, a_2)}{\partial a_1 \partial a_2} \right] \\ -M \left[ \frac{\partial^2 \ln p(x_1, \dots, x_n | a_1, a_2)}{\partial a_1 \partial a_2} \right] & -M \left[ \frac{\partial^2 \ln p(x_1, \dots, x_n | a_1, a_2)}{\partial a_2^2} \right] \end{bmatrix}.$$

Находя матрицу, обратную матрице  $J$ , и считая оценки параметров совместно эффективными, получим

$$\begin{aligned} D_{\hat{a}_1} = R_{11} &= \frac{J_{22}}{J_{11}J_{22} - J_{12}^2} = \frac{1}{J_{11}} \frac{1}{1 - r^2(\hat{a}_1, \hat{a}_2)}, \\ D_{\hat{a}_2} = R_{22} &= \frac{J_{11}}{J_{11}J_{22} - J_{12}^2} = \frac{1}{J_{22}} \frac{1}{1 - r^2(\hat{a}_1, \hat{a}_2)}, \end{aligned}$$

где  $r(\hat{a}_1, \hat{a}_2) = \frac{J_{12}}{\sqrt{J_{11}J_{22}}}$  — нормированный корреляционный момент (коэффициент корреляции) между оценками  $\hat{a}_1$  и  $\hat{a}_2$ .

Из сравнения двух последних формул с неравенством Рао — Крамера следует, что первый множитель в правой части этих формул совпадает с дисперсией эффективной оценки одного параметра. Поскольку  $0 \leq r^2(a_1, a_2) \leq 1$ , то видно, что наличие конечной корреляции между оценками всегда приводит к увеличению дисперсии совместно эффективных оценок по сравнению с дисперсией эффективной оценки одного параметра.

Дисперсии оценок являются важнейшими характеристиками качества оценок, и во многих практических задачах ограничиваются только их вычислением. Однако в математической статистике, как отмечалось ранее, точность оценки определяют также величиной доверительного интервала, соответствующего заданной доверительной вероятности. Для построения доверительного интервала необходимо знать закон, по которому распределена оценка.

## 9.2. Точечные оценки моментов случайной величины

Часто на практике плотность вероятности случайной величины неизвестна, поэтому остается довольствоваться только оценками моментов случайной величины. Рассмотрим оценки математического ожидания  $m_x$  и дисперсии  $D_x$ . Один из возможных и естественных способов получения оценок этих числовых характеристик на основе независимых наблюдений случайной величины  $X$  заключается в том, что в качестве статистических оценок  $\hat{m}_x$  и  $\hat{D}_x$  берут соответственно выборочное среднее  $m_x^*$  и выборочную дисперсию  $D_x^*$ :

$$m_x^* = \frac{1}{n} \sum_{i=1}^n x_i;$$

$$D_x^* = \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (m_x^*)^2.$$

Аналогичные выражения можно записать для оценок начальных и центральных моментов более высоких порядков. Если математическое ожидание  $m_x$  известно, то при оценке дисперсии и других центральных моментов в формулу для  $D_x^*$  следует вместо  $m_x^*$  подставлять истинное значение.

При вычислении выборочного среднего и дисперсии по сгруппированным данным пользуются формулами

$$m_x^* = \frac{1}{n} \sum_{k=1}^m x_k^0 n_k;$$

$$D_x^* = \frac{1}{n} \sum_{k=1}^m (x_k^0 - m_x^*)^2 n_k = \frac{1}{n} \sum_{k=1}^m x_k^{02} n_k - m_x^{*2},$$

где  $m$  — число интервалов группировки,  $x_k^0$  — середины интервалов,  $n_k$  — число наблюдений, попавших в  $k$ -й интервал.

Рассмотрим некоторые характеристики оценок математического ожидания и дисперсии. На основании закона больших чисел (теорема Чебышева) имеем

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{1}{n} \sum_{i=1}^n x_i - m_x \right| < \varepsilon \right) = 1,$$

то есть оценка  $m_x^*$  является состоятельной.

Найдем математическое ожидание выборочного среднего:

$$M[m_x^*] = \frac{1}{n} M \left[ \sum_{i=1}^n x_i \right] = \frac{1}{n} \sum_{i=1}^n M[x_i] = \frac{1}{n} \cdot n \cdot m_x = m_x.$$

Следовательно, оценка  $m_x^*$  является несмещенной.

Найдем дисперсию выборочного среднего:

$$D_{m_x^*} = M[(m_x^* - m_x)^2] = M \left[ \left( \frac{1}{n} \sum_{i=1}^n x_i - m_x \right)^2 \right] =$$

$$= M \left[ \left( \frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n m_x \right)^2 \right] = \frac{1}{n^2} M \left[ \sum_{i=1}^n (x_i - m_x)^2 \right] =$$

$$= \frac{1}{n^2} \sum_{i=1}^n M[(x_i - m_x)^2] = \frac{1}{n^2} n D_x = \frac{D_x}{n}.$$

Как видно, дисперсия оценки математического ожидания уменьшается обратно пропорционально объему выборки.

Рассмотрим теперь оценку дисперсии  $D_x$ . Если математическое ожидание случайной величины  $X$  известно, то

$$D_x^* = \frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2;$$

$$M[D_x^*] = \frac{1}{n} M \left[ \sum_{i=1}^n (x_i - m_x)^2 \right] = \frac{1}{n} \sum_{i=1}^n M[(x_i - m_x)^2] = \frac{1}{n} n D_x = D_x,$$

то есть такая оценка является несмещенной.

Найдем теперь математическое ожидание выборочной дисперсии, когда используется выборочное среднее:

$$M[D_x^*] = M \left[ \frac{1}{n} \sum_{i=1}^n (x_i - m_x^*)^2 \right] = \frac{1}{n} M \left[ \sum_{i=1}^n (x_i - m_x^*)^2 \right].$$

Представим сумму в другом виде:

$$\sum_{i=1}^n (x_i - m_x^*)^2 = \sum_{i=1}^n (x_i - m_x + m_x - m_x^*)^2 =$$

$$= \sum_{i=1}^n (x_i - m_x)^2 - 2(m_x^* - m_x) \sum_{i=1}^n (x_i - m_x) + \sum_{i=1}^n (m_x^* - m_x)^2 =$$

$$= \sum_{i=1}^n (x_i - m_x)^2 - 2n(m_x^* - m_x)^2 + n(m_x^* - m_x)^2 =$$

$$= \sum_{i=1}^n (x_i - m_x)^2 - n(m_x^* - m_x)^2.$$

Так как  $M[(x_i - m_x)^2] = D_x$  и  $M[(m_x^* - m_x)^2] = \frac{D_x}{n}$ , то

$$M[D_x^*] = \frac{1}{n} (n D_x - n \frac{D_x}{n}) = \frac{n-1}{n} D_x = D_x - \frac{D_x}{n}.$$

Следовательно, если за оценку дисперсии  $D_x$  принять выборочную дисперсию  $D_x^*$ , когда оценка математического ожидания производится по выборке, то эта оценка окажется смещенной. Несмещенную оценку в этом случае можно получить, вычисляя несколько отличную от  $D_x^*$  выборочную дисперсию

$$s^2 = \frac{n}{n-1} D_x^* = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x^*)^2,$$

которую называют исправленной дисперсией. В самом деле,

$$M[s^2] = M\left[\frac{n}{n-1} D_x^*\right] = \frac{n}{n-1} M[D_x^*] = \frac{n}{n-1} \frac{n-1}{n} D_x = D_x.$$

Эта оценка удовлетворяет условиям состоятельности и эффективности. Отметим, что при большом объеме выборки  $n$  практически безразлично, по какой из формул вычислять дисперсию. Однако при малом  $n$  следует пользоваться формулой для исправленной дисперсии.

Для оценки среднеквадратического отклонения случайной величины используют так называемое "исправленное" среднеквадратическое отклонение, которое равно квадратному корню из исправленной дисперсии:  $s = \sqrt{s^2}$ . Однако  $s$  не является несмещенной оценкой, то есть  $M[s] \neq \sqrt{D_x}$ .

### 9.3. Методы оценки параметров распределения

Предположим, что на основании анализа физического механизма, порождающего случайную величину  $X$ , можно сделать достаточно обоснованное заключение о виде закона распределения или плотности вероятности интересующей нас случайной величины. Однако параметры этого распределения неизвестны, и их нужно оценить по результатам наблюдений, представленных в виде конечной выборки  $x_1, x_2, \dots, x_n$ .

Для решения этой задачи могут применяться различные методы. Ниже будут рассмотрены два из них: метод моментов и метод максимального правдоподобия.

#### 9.3.1. Метод моментов

Метод моментов является наиболее простым методом оценки параметров. Он был предложен английским статистиком Пирсоном в 1894 г. Метод моментов заключается в том, что определенное количество моментов выборочного распределения приравнивается к соответствующим моментам теоретического распределения, которые являются функциями от неизвестных параметров.

Рассматриваемое количество моментов должно быть равно числу подлежащих оценке параметров. Решая систему уравнений относительно этих параметров, можно найти искомые оценки. Так, для нахождения оценок параметров нормальной плотности вероятности  $p(x, a_1, a_2)$ , содержащей два неизвестных параметра  $a_1$  и  $a_2$ , составляется система двух уравнений:

$$m_1(\hat{a}_1, \hat{a}_2) = m_1^*,$$

$$\mu_2(\hat{a}_1, \hat{a}_2) = \mu_2^*,$$

где  $m_1$  — начальный момент первого порядка,  $\mu_2$  — центральный момент второго порядка,  $m_1^*$  — выборочный начальный момент первого порядка (выборочное среднее),  $\mu_2^*$  — выборочный центральный момент второго порядка (выборочная дисперсия). Решая эту систему, находят оценки  $\hat{a}_1$  и  $\hat{a}_2$  параметров плотности вероятности. Можно использовать также моменты других порядков, если они выражаются через параметры распределения.

*Рассмотрим пример.* Пусть случайная величина  $X$  имеет плотность вероятности вида одностороннего нормального закона:

$$p(x, \sigma) = \frac{2}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0.$$

Пусть наблюдение за случайной величиной  $X$  дает результаты  $x_1, x_2, \dots, x_n$ . По этим результатам методом моментов необходимо дать оценку единственного параметра  $\sigma$  распределения  $p(x, \sigma)$ .

Найдем начальный теоретический момент первого порядка:

$$\begin{aligned} m_1 &= \int_0^{\infty} \frac{2x}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = \int_0^{\infty} \frac{2\sigma^2}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} d\left(\frac{x^2}{2\sigma^2}\right) = \\ &= \int_0^{\infty} \frac{2\sigma}{\sqrt{2\pi}} e^{-t} dt = \frac{2\sigma}{\sqrt{2\pi}} (-e^{-t}) \Big|_0^{\infty} = \sigma \sqrt{\frac{2}{\pi}} = 0,8\sigma. \end{aligned}$$

Выборочный начальный момент первого порядка (выборочное среднее)

$$m_1^* = \frac{1}{n} \sum_{i=1}^n x_i.$$

Отсюда следует уравнение

$$0,8\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n x_i,$$

то есть

$$\hat{\sigma} = \frac{1}{0,8n} \sum_{i=1}^n x_i.$$

Легко убедиться, что полученная оценка является несмещенной.

Метод моментов отличается простотой, не требует сложных вычислений. Однако получаемые с его помощью оценки не являются наилучшими из возможных, поскольку имеют не наименьшую возможную дисперсию.

### 9.3.2. Метод максимального правдоподобия

Рассмотрим теперь оценку параметров распределения методом максимального правдоподобия. Сущность этого метода можно сформулировать словами его основоположника Фишера: "Наилучшим описанием исследуемого явления является то, при котором максимальна вероятность получить фактически измеренные значения наблюдаемых величин". Чтобы воспользоваться этим принципом, необходимо каким-то образом математически описать вероятность получения при наблюдениях каждого конкретного набора результатов (которые уже получены) и далее подогнать параметры этого математического описания так, чтобы при имеющихся конкретных результатах эта функция имела максимум.

Пусть распределение случайной величины  $X$  зависит от одного параметра  $a$ , а  $x_1, x_2, \dots, x_n$  — результат выборки. Распределение выборки при фиксированном значении параметра  $a$  определяется условной плотностью вероятности  $p(x_1, \dots, x_n | a)$ . Эта плотность вероятности, рассматриваемая только как функция параметра  $a$  при фиксированных  $x_1, \dots, x_n$ , называется *функцией правдоподобия*, относя-

щейся к данной выборке. Функция правдоподобия обычно обозначается через  $L(a)$ :

$$L(a) = p(x_1, \dots, x_n | a).$$

Оценка  $\hat{a}$  параметра  $a$ , построенная по выборочным значениям случайной величины и максимизирующая функцию правдоподобия, называется *максимально правдоподобной оценкой*. Таким образом, максимально правдоподобная оценка находится из условия

$$L(\hat{a}) = \max_a L(a),$$

то есть является решением уравнения

$$\left. \frac{dL(a)}{da} \right|_{a=\hat{a}} = 0,$$

которое явным образом зависит от результатов выборки.

Поскольку логарифм — функция монотонная, то расположение максимума можно искать по функции  $\ln L(a)$ , что часто бывает более удобно. В связи с этим вместо приведенного уравнения можно отыскивать соответствующий корень уравнения

$$\left. \frac{d \ln L(a)}{da} \right|_{a=\hat{a}} = 0,$$

которое называется *уравнением правдоподобия*.

Если оценке подлежат несколько параметров  $a_1, a_2, \dots, a_k$  распределения  $p(x, a_1, \dots, a_k)$ , то функция правдоподобия будет зависеть от этих параметров. В данном случае совместные оценки  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$  параметров определяются как решения системы уравнений правдоподобия вида

$$\left. \frac{\partial}{\partial a_i} \ln L(a_1, a_2, \dots, a_k) \right|_{a_i=\hat{a}_i} = 0, \quad i = 1, 2, \dots, k.$$

Следует отметить, что во многих важных случаях системы уравнений правдоподобия решаются очень сложно и параметры обычно находятся с помощью численных методов.

Основные достоинства оценок, получаемых методом максимального правдоподобия, заключаются в следующем:

1) максимально правдоподобные оценки являются состоятельными, то есть сходятся по вероятности к соответствующим параметрам;

2) при большом объеме выборки оценки распределены приближенно по нормальному закону с центром в точке истинного значения параметра и дисперсией, равной нижней границе, задаваемой неравенством Рао — Крамера. Иначе говоря, максимально правдоподобная оценка является асимптотически эффективной, то есть не существует другой оценки с меньшей дисперсией;

3) если параметр допускает эффективную оценку, то эта оценка получается как единственное в этом случае решение уравнения правдоподобия;

4) оценка является достаточной статистикой, так как полностью использует всю информацию относительно параметра, которая содержится в выборке;

5) оценка обладает свойством “инвариантности” относительно замены переменной, то есть если вместо параметра  $a$  нужно оценить некоторую его однозначную функцию  $\varphi(a)$ , то оценкой для нее будет  $\varphi(\hat{a})$ . Однако оценка  $\varphi(\hat{a})$  может не сохранять свойств оценки  $\hat{a}$ . Примером может служить уже упоминавшаяся ранее оценка дисперсии  $s^2$  и оценка среднеквадратического отклонения  $\sqrt{s^2}$ .

Получаемые методом максимального правдоподобия оценки часто оказываются смещенными. Однако эта смещенность не имеет существенного значения и во многих случаях может быть устранена при надлежащем исправлении оценки.

Рассмотрим ряд примеров. Пусть произведено  $n$  независимых измерений какой-то величины  $a$ . Предполагается, что все измерения равноточны с дисперсией  $\sigma^2$ , а ошибки измерений распределены по нормальному закону с нулевым средним. Результат  $i$ -го измерения можно представить в виде

$$x_i = a + \varepsilon_i.$$

Поскольку измерения независимы, то функция правдоподобия будет иметь вид

$$\begin{aligned} L(a) &= p(x_1, \dots, x_n | a) = p(x_1 | a) p(x_2 | a) \dots p(x_n | a) = \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - a)^2}{2\sigma^2}} = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - a)^2}. \end{aligned}$$

Логарифмируя это выражение, имеем

$$\ln L(a) = -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - a)^2.$$

Дифференцируя результат по  $a$ , получим

$$\frac{d \ln L(a)}{da} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - a) = \frac{1}{\sigma^2} \left( \sum_{i=1}^n x_i - \sum_{i=1}^n a \right) = \frac{1}{\sigma^2} \left( \sum_{i=1}^n x_i - na \right).$$

Приравнявая  $\frac{d \ln L(a)}{da}$  к нулю, получаем уравнение правдоподобия, а из него находим

$$\hat{a} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Полученная оценка представляет собой просто арифметическое среднее из результатов измерений (выборочное среднее). Как известно, эта оценка является несмещенной.

Проверим оценку на эффективность:

$$\frac{d \ln L(a)}{da} = \frac{1}{\sigma^2} \left( \sum_{i=1}^n x_i - na \right) = \frac{1}{\sigma^2} (\hat{a}n - na) = \frac{n}{\sigma^2} (\hat{a} - a),$$

то есть  $\frac{d \ln L(a)}{da}$  представима в виде  $k(a)(\hat{a} - a)$ . Значит, оценка  $\hat{a}$  эффективна.

Дисперсия оценки

$$D_{\hat{a}} = \frac{-1}{M \left[ \frac{d^2}{da^2} \ln L(a) \right]} = \frac{-1}{M \left[ -\frac{n}{\sigma^2} \right]} = \frac{\sigma^2}{n}.$$

Легко видеть, что если производится одно измерение величины  $a$ , то за ее оценку следует принять измеренное значение  $x_1$ . При этом дисперсия оценки будет равна  $\sigma^2$ .

Пусть теперь в той же задаче дисперсия ошибок  $\sigma^2$  неизвестна. Тогда по результатам измерений необходимо получить совместную оценку двух параметров —  $a$  и  $\sigma^2$ . В этом примере функция правдоподобия имеет аналогичный вид, но зависит от двух параметров:

$$L(a, \sigma^2) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - a)^2}.$$

Ее логарифм

$$\ln L(a, \sigma^2) = -n \ln(\sqrt{2\pi} \sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - a)^2.$$

Дифференцируя это выражение сначала по  $a$ , а затем по  $\sigma^2$ , получим систему двух уравнений правдоподобия для нахождения параметров  $\hat{a}$  и  $\hat{\sigma}^2$ :

$$\left. \frac{\partial \ln L(a, \sigma^2)}{\partial a} \right|_{a=\hat{a}, \sigma^2=\hat{\sigma}^2} = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{a}) = 0;$$

$$\left. \frac{\partial \ln L(a, \sigma^2)}{\partial(\sigma^2)} \right|_{a=\hat{a}, \sigma^2=\hat{\sigma}^2} = \frac{-1}{2\hat{\sigma}^2} \left[ n - \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{a})^2 \right] = 0.$$

Решениями этой системы являются

$$\hat{a} = \frac{1}{n} \sum_{i=1}^n x_i;$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{a})^2,$$

представляющие собой выборочное среднее и выборочную дисперсию. Последняя является смещенной, но смещение можно устранить.

Определим дисперсии оценок и нормированную взаимную корреляцию между ними. Сначала найдем элемент  $J_{11}$  информационной матрицы Фишера:

$$-M \left[ \frac{\partial^2}{\partial a^2} \ln L(a, \sigma^2) \right] = \frac{n}{\sigma^2}.$$

Вместо  $\sigma^2$  подставим оценку  $\hat{\sigma}^2$ . Тогда

$$J_{11} = \frac{n}{\hat{\sigma}^2}.$$

Для элемента  $J_{22}$  матрицы Фишера получим:

$$\begin{aligned} J_{22} &= -M \left[ \frac{\partial^2}{\partial(\sigma^2)^2} \ln L(a, \sigma^2) \right] = \\ &= -M \left[ \frac{\partial}{\partial(\sigma^2)} \left( -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - a)^2 \right) \right] = \\ &= M \left[ -\frac{n}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^3} \sum_{i=1}^n (x_i - a)^2 \right]. \end{aligned}$$

Произведя операцию математического ожидания, получим

$$J_{22} = \frac{n}{2(\sigma^2)^2}.$$

Вычислим элемент  $J_{12}$  матрицы Фишера:

$$\begin{aligned} J_{12} &= -M \left[ \frac{\partial \ln L(a, \sigma^2)}{\partial a} \cdot \frac{\partial \ln L(a, \sigma^2)}{\partial(\sigma^2)} \right] = \\ &= -M \left[ -\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - a) \left( -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - a)^2 \right) \right] = \\ &= M \left[ \frac{n}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - a) - \frac{1}{2(\sigma^2)^3} \sum_{i=1}^n (x_i - a)^3 \right] = 0, \end{aligned}$$

так как все нечетные центральные моменты нормально распределенной случайной величины равны нулю.

Итак, оценки  $\hat{a}$  и  $\hat{\sigma}^2$  являются асимптотически эффективными оценками с дисперсиями, равными соответственно

$$D_{\hat{a}} = \frac{\hat{\sigma}^2}{n} \quad \text{и} \quad D_{\hat{\sigma}^2} = \frac{2(\hat{\sigma}^2)^2}{n},$$

причем корреляционная функция между ними равна нулю, то есть эти оценки асимптотически независимы. Заметим, что в рассмотренном примере оценить по одному измерению величину  $a$  и дисперсию  $\sigma^2$  нельзя: требуются, как минимум, два измерения.

## 9.4. Интервальная оценка параметров

Выше был рассмотрен вопрос нахождения точечных оценок параметров распределений. Точность таких оценок характеризовалась их дисперсией. При этом отсутствовали сведения о том, насколько близки полученные оценки к истинным значениям параметров.

Более полный и надежный способ оценивания параметров распределений заключается в определении не единственного точечного значения, а интервала, который с заданной вероятностью накрывает истинное значение оцениваемого параметра. Пусть найденная по результатам выборки объема  $n$  оценка  $\hat{a} = \hat{a}(x_1, x_2, \dots, x_n)$  является точечной оценкой неизвестного параметра  $a$ . Чем меньше разность  $|a - \hat{a}|$ , тем лучше качество оценки и точнее оценка. Таким образом, положительное число  $\delta$  характеризует точность оценки:

$$|a - \hat{a}| < \delta.$$

Понятно, что точность  $\delta$  зависит от объема выборки. Каков должен быть объем выборки  $n$ , чтобы обеспечить заданную точность  $\delta$ , или как определить точность  $\delta$  при заданном объеме выборки? На эти вопросы нельзя ответить, непосредственно используя написанное неравенство. Можно говорить только о вероятности  $\beta$ , с которой это неравенство выполняется.

Вероятность  $\beta$ , называемая *доверительной вероятностью*, задается до осуществления выборки и зависит не от условий опыта, а от необходимой степени надежности выводов, следующих из данной статистической обработки. Значение  $\beta$  обычно выбирается из ряда 0,95; 0,99; 0,999.

Пусть вероятность того, что  $|a - \hat{a}| < \delta$ , равна  $\beta$ :

$$P(|a - \hat{a}| < \delta) = \beta.$$

Преобразуем эту формулу:

$$P(|a - \hat{a}| < \delta) = P(-\delta < a - \hat{a} < \delta) = P(\hat{a} - \delta < a < \hat{a} + \delta).$$

Интервал длиной  $(\hat{a} - \delta, \hat{a} + \delta)$ , накрывающий с вероятностью  $\beta$  истинное значение параметра, называется *доверительным интервалом*, соответствующим доверительной вероятности  $\beta$ . Отметим, что нельзя говорить о том, что значение параметра лежит внутри доверительного интервала с вероятностью  $\beta$ . Хотя обе формулировки на первый взгляд одинаковы, они имеют существенное различие. Используемая формулировка означает, что, хотя оцениваемый параметр и неизвестен, он имеет постоянное значение, а следовательно, не имеет разброса, поскольку это не случайная величина.

Смысл доверительного интервала состоит в том, что при многократном повторении выборки объема  $n$  в относительной доле случаев, равной  $\beta$ , доверительный интервал, соответствующий доверительной вероятности  $\beta$ , накрывает истинное значение оцениваемого параметра.

Задача определения доверительного интервала может быть решена только тогда, когда удастся найти закон распределения случайной величины, используемой в качестве оценки, то есть плотность вероятности  $p(\hat{a})$ . В общем случае этот закон зависит от самого неизвестного параметра. Однако иногда удается перейти от оценки  $\hat{a}$  к таким функциям выборочных значений, закон распределения которых зависит только от объема выборки и закона распределения случайной величины  $X$  и не зависит от неизвестных параметров.

Рассмотрим, каким образом находятся доверительные интервалы для оценок параметров, когда выборка произведена из генеральной совокупности с нормальным законом распределения с параметрами  $m_x$  и  $D_x$ .

## 9.5. Доверительные интервалы для параметров распределения

### 9.5.1. Доверительный интервал для математического ожидания при известной дисперсии

Если случайная величина  $X$  распределена по нормальному закону с математическим ожиданием  $m_x$  и дисперсией  $D_x$ , то выборочное среднее

$$m_x^* = \frac{1}{n} \sum_{i=1}^n x_i$$

также будет распределено по нормальному закону с параметрами  $M[m_x^*] = m_x$  и  $D_{m_x^*} = \frac{D_x}{n}$ :

$$p(m_x^*) = \frac{1}{\sqrt{2\pi \frac{D_x}{n}}} e^{-\frac{(m_x^* - m_x)^2}{2 \frac{D_x}{n}}}.$$

Как видно, плотность вероятности оцениваемого параметра зависит от самого этого параметра. Введем случайную величину

$$z = \frac{m_x^* - m_x}{\sqrt{D_x}} \sqrt{n},$$

которая распределена по нормальному закону с нулевым математическим ожиданием и единичной дисперсией. Тогда вероятность  $\beta$  того, что случайная величина  $z$  (рис. 9.1) не отклонится от своего математического ожидания на величину  $z_\beta$ , находится по формуле

$$P(-z_\beta < z < z_\beta) = \frac{1}{\sqrt{2\pi}} \int_{-z_\beta}^{z_\beta} e^{-\frac{z^2}{2}} dz = 2\Phi(z_\beta) - 1 = \beta,$$

где  $\Phi(z)$  — интеграл вероятности.

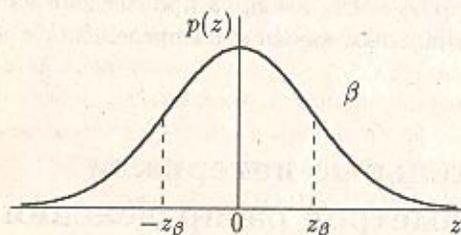


Рис. 9.1. К определению доверительного интервала для случайной величины  $z$

Отсюда получаем уравнение для нахождения границ доверительного интервала для величины  $z$ :

$$\Phi(z_\beta) = \frac{1 + \beta}{2},$$

где  $z_\beta$  — квантиль порядка  $\frac{1+\beta}{2}$  нормального распределения. Его значение можно найти по таблицам интеграла вероятности.

Пусть квантиль, соответствующий заданной доверительной вероятности  $\beta$ , известен. Тогда

$$\begin{aligned} P(-z_\beta < z < z_\beta) &= P(-z_\beta < \frac{m_x^* - m_x}{\sqrt{D_x}} \sqrt{n} < z_\beta) = \\ &= P(-z_\beta \frac{\sigma_x}{\sqrt{n}} < m_x^* - m_x < z_\beta \frac{\sigma_x}{\sqrt{n}}) = \\ &= P(m_x^* - z_\beta \frac{\sigma_x}{\sqrt{n}} < m_x < m_x^* + z_\beta \frac{\sigma_x}{\sqrt{n}}) = \beta. \end{aligned}$$

Следовательно, доверительный интервал  $(m_x^* - z_\beta \frac{\sigma_x}{\sqrt{n}}, m_x^* + z_\beta \frac{\sigma_x}{\sqrt{n}})$  покрывает неизвестное математическое ожидание с заданной вероятностью  $\beta$ . Точность оценки математического ожидания  $\delta = z_\beta \frac{\sigma_x}{\sqrt{n}}$ .

Анализируя полученные формулы, легко видеть, что увеличение объема выборки  $n$  приводит к уменьшению длины доверительного интервала; увеличение доверительной вероятности  $\beta$  приводит к увеличению длины доверительного интервала, то есть к уменьшению точности  $\delta$ ; если задать точность  $\delta$  и доверительную вероятность  $\beta$ , то из соотношения

$$\delta = z_\beta \frac{\sigma_x}{\sqrt{n}}$$

можно найти минимальный объем выборки, который обеспечит заданную точность:

$$n = \left( \frac{z_\beta \sigma_x}{\delta} \right)^2.$$

Если величина  $X$  распределена не по нормальному закону, то так как величина  $m_x^*$  представляет собой сумму независимых одинаково распределенных случайных величин, то, согласно центральной предельной теореме, при достаточно больших  $n$  ( $n \geq 30$ ) ее закон распределения близок к нормальному.

### 9.5.2. Доверительный интервал для дисперсии

Рассмотрим сначала случай, когда математическое ожидание случайной величины  $X$  известно. Оценка дисперсии в этом случае дается выражением

$$D_x^* = \frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2.$$

Введем случайную величину

$$\chi^2 = \frac{nD_x^*}{D_x} = \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2.$$

Если случайная величина  $X$  распределена по нормальному закону, то случайная величина  $\chi^2$  представляет собой сумму квадратов независимых случайных величин, имеющих нулевое математическое ожидание и единичную дисперсию, и распределена по закону  $\chi^2$  с  $n$  степенями свободы.

Пусть теперь математическое ожидание случайной величины  $X$  неизвестно и оценивается по выборке. Тогда несмещенная оценка дисперсии дается выражением

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x^*)^2.$$

Введем случайную величину

$$\begin{aligned} \chi^2 &= \frac{(n-1)s^2}{D_x} = \sum_{i=1}^n \left( \frac{x_i - m_x^*}{\sqrt{D_x}} \right)^2 = \sum_{i=1}^n \left( \frac{x_i - m_x + m_x - m_x^*}{\sqrt{D_x}} \right)^2 = \\ &= \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2 - \sum_{i=1}^n \frac{(x_i - m_x)(m_x^* - m_x)}{D_x} + \sum_{i=1}^n \left( \frac{m_x^* - m_x}{\sqrt{D_x}} \right)^2 = \\ &= \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2 - 2n \left( \frac{m_x^* - m_x}{\sqrt{D_x}} \right)^2 + n \left( \frac{m_x^* - m_x}{\sqrt{D_x}} \right)^2 = \\ &= \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2 - n \left( \frac{m_x - m_x^*}{\sqrt{D_x}} \right)^2 = \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2 - n \left( \frac{m_x - m_x^*}{\sqrt{n D_{m_x^*}}} \right)^2 = \\ &= \sum_{i=1}^n \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2 - \left( \frac{m_x - m_x^*}{\sqrt{D_{m_x^*}}} \right)^2. \end{aligned}$$

Поскольку случайные величины  $\frac{x_i - m_x}{\sqrt{D_x}}$  и  $\frac{m_x - m_x^*}{\sqrt{D_{m_x^*}}}$  имеют нулевое математическое ожидание и единичную дисперсию, то

$$\chi^2 = \sum_{i=1}^{n-1} \left( \frac{x_i - m_x}{\sqrt{D_x}} \right)^2.$$

Если случайная величина  $X$  распределена по нормальному закону, то величина  $\chi^2$  распределена по закону  $\chi^2$  с  $l = n - 1$  степенями свободы.

Таким образом, если математическое ожидание случайной величины  $X$  неизвестно, то случайная величина  $\chi^2 = \frac{(n-1)s^2}{D_x}$  распределена по закону  $\chi^2$  с  $l = n - 1$  степенями свободы. Уменьшение числа степеней свободы в данном случае обусловлено тем, что выборочные значения связаны между собой через оценку математического ожидания.

Вероятность того, что случайная величина  $\chi_i^2$  попадает в интервал  $(c_1, c_2)$  (рис. 9.2), равна

$$P(c_1 < \chi_i^2 < c_2) = \int_{c_1}^{c_2} p(\chi_i^2) d(\chi_i^2) = F_{\chi_i^2}(c_2) - F_{\chi_i^2}(c_1),$$

где  $F_{\chi_i^2}$  — функция распределения величины  $\chi_i^2$ , которая обычно не табулируется. Поэтому целесообразно перейти к процентным точкам  $\chi^2$ -распределения, для которых имеются таблицы.

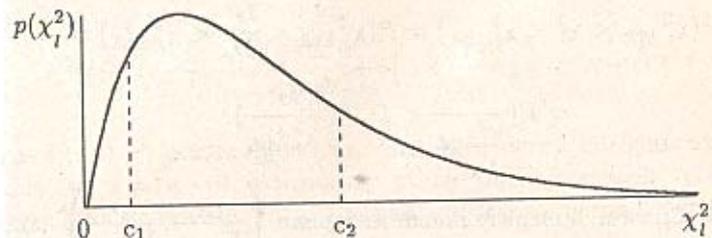


Рис. 9.2. К определению вероятности попадания случайной величины  $\chi_i^2$  в интервал  $[c_1, c_2]$

Поскольку распределение  $\chi^2$  асимметрично, то обычно доверительный интервал строят так, чтобы площади под плотностью вероятности от 0 до  $c_1$  и от  $c_2$  до бесконечности были равны  $\frac{1-\beta}{2}$ :

$$P(\chi_i^2 < c_1) = \int_0^{c_1} p(\chi_i^2) d(\chi_i^2) = \frac{1-\beta}{2};$$

$$P(\chi_i^2 > c_2) = \int_{c_2}^{\infty} p(\chi_i^2) d(\chi_i^2) = \frac{1-\beta}{2}.$$

Из второго написанного уравнения при заданном значении  $\beta$  по таблицам процентных точек  $\chi^2$ -распределения можно найти  $c_2 = \chi_{l, \frac{1-\beta}{2}}^2$ . Чтобы выразить вероятность  $P(\chi_i^2 < c_1)$  через процентную точку, перепишем ее в виде

$$P(\chi_i^2 < c_1) = 1 - \int_{c_1}^{\infty} p(\chi_i^2) d(\chi_i^2) = \frac{1-\beta}{2}.$$

Отсюда

$$\int_{c_1}^{\infty} p(\chi_i^2) d(\chi_i^2) = \frac{1+\beta}{2}.$$

Таким образом,

$$c_1 = \chi_{l, \frac{1+\beta}{2}}^2.$$

Если процентные точки, соответствующие заданной доверительной вероятности  $\beta$ , известны, то для случая неизвестного математического ожидания

$$\begin{aligned} P(\chi_{l; \frac{1+\beta}{2}}^2 < \chi_l^2 < \chi_{l; \frac{1-\beta}{2}}^2) &= P(\chi_{l; \frac{1+\beta}{2}}^2 < \frac{ls^2}{D_x} < \chi_{l; \frac{1-\beta}{2}}^2) = \\ &= P\left(\frac{ls^2}{\chi_{l; \frac{1-\beta}{2}}^2} < D_x < \frac{ls^2}{\chi_{l; \frac{1+\beta}{2}}^2}\right). \end{aligned}$$

Таким образом, доверительный интервал  $\left(\frac{ls^2}{\chi_{l; \frac{1-\beta}{2}}^2}, \frac{ls^2}{\chi_{l; \frac{1+\beta}{2}}^2}\right)$  покрывает неизвестную дисперсию с заданной вероятностью  $\beta$ .

Процентные точки распределения  $\chi^2$  приведены в таблице 2 Приложения.

### 9.5.3. Доверительный интервал для математического ожидания при неизвестной дисперсии

При построении доверительного интервала математического ожидания при известной дисперсии была введена нормированная случайная величина

$$Z = \frac{m_x^* - m_x}{\sqrt{D_x}} \sqrt{n},$$

которая распределена по нормальному закону.

Введем по аналогии с  $Z$  случайную величину

$$T = \frac{m_x^* - m_x}{\sqrt{s^2}} \sqrt{n}.$$

Перепишем выражение для  $T$  в виде

$$T = \frac{m_x^* - m_x}{\sqrt{\frac{D_x}{n}} \sqrt{\frac{s^2}{D_x}}} = \frac{m_x^* - m_x}{\sqrt{\frac{D_x}{n}}} : \sqrt{\frac{s^2}{D_x}} = \frac{Z}{\sqrt{U}}.$$

Случайная величина  $U$  в знаменателе, если случайная величина  $X$  распределена по нормальному закону, имеет распределение  $\frac{\chi_{n-1}^2}{n-1}$ .

Таким образом, случайная величина  $T$  есть функция двух случайных величин  $Z$  и  $U$ , распределения которых известны, и ее распределение можно найти по известным правилам преобразования плотностей вероятностей. Можно показать, что величина  $T$  распределена по закону

$$p(t) = \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi(n-1)}\Gamma(\frac{n-1}{2})} \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}, \quad -\infty < t < \infty,$$

где  $\Gamma(\dots)$  — гамма-функция. Число степеней свободы этого распределения  $l = n - 1$ , поскольку такое число степеней свободы имела исходная величина  $\chi_{n-1}^2$ . Тогда плотность вероятности случайной величины  $T$  можно записать в виде

$$p(t) = \frac{\Gamma(\frac{l+1}{2})}{\sqrt{\pi l}\Gamma(\frac{l}{2})} \left(1 + \frac{t^2}{l}\right)^{-\frac{l+1}{2}}, \quad -\infty < t < \infty.$$

Это распределение было получено в 1908 г. англичанином Госсетом и опубликовано им под псевдонимом "Стьюдент", поэтому оно называется  $t$ -распределением Стьюдента.

На рис. 9.3 приведены плотности вероятностей случайной величины, распределенной по закону Стьюдента при  $l = 1$ , и нормированной нормальной случайной величины с нулевым математическим ожиданием. Видно, что распределение Стьюдента имеет более протяженные хвосты, чем нормальное распределение. Однако при  $n \geq 30$  кривые почти сливаются.

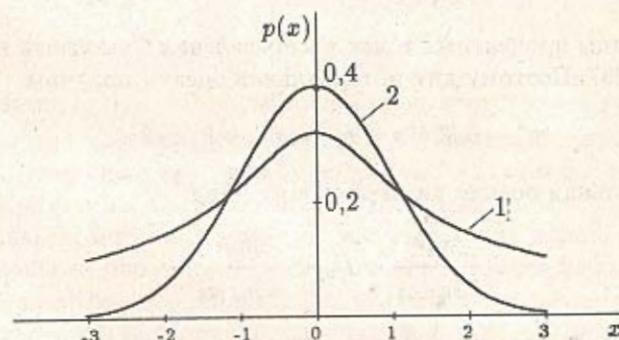


Рис. 9.3. Распределение Стьюдента (1) при одной степени свободы и нормальное распределение (2)

Рассуждая аналогично случаю известной дисперсии, можно показать, что при неизвестной дисперсии интервальная оценка математического ожидания будет иметь вид

$$m_x^* - t_{1-\frac{\beta}{2}} \cdot \frac{s}{\sqrt{n}} < m_x < m_x^* + t_{1-\frac{\beta}{2}} \cdot \frac{s}{\sqrt{n}},$$

где  $t_{1-\frac{\beta}{2}}$  —  $(\frac{1-\beta}{2})$ -процентная точка распределения Стьюдента. Значения процентных точек  $t$ -распределения Стьюдента приведены в таблице 3 Приложения. Полученный доверительный интервал обладает теми же свойствами, что и доверительный интервал для математического ожидания при известной дисперсии.

*Рассмотрим пример.* Пусть произведено  $n = 31$  независимых измерений какой-то величины. Ошибки измерений одинаковы, но неизвестны. Требуется определить 90%-ные доверительные интервальные оценки истинного значения измеряемой величины (математического ожидания) и истинной дисперсии ошибок измерения. Ошибки измерений считать распределенными по нормальному закону. Результаты измерений имеют следующий вид: 60, 61, 47, 56, 61, 63, 65, 69, 54, 59, 43, 61, 55, 61, 56, 48, 67, 65, 60, 58, 57, 62, 57, 58, 53, 59, 58, 61, 57, 62, 54.

Интервальная оценка математического ожидания  $m_x$ , соответствующая доверительной вероятности  $\beta = 0,9$  при выборочном среднем значении  $m_x^*$  и выборочной дисперсии  $s^2$ , вычисленным по выборке объема  $n = 31$ , имеет вид

$$m_x^* - t_{30;0,05} \frac{s}{\sqrt{31}} < m_x < m_x^* + t_{30;0,05} \frac{s}{\sqrt{31}}.$$

Из таблицы процентных точек распределения Стьюдента находим  $t_{30;0,05} = 1,697$ . Поэтому для интервальной оценки получим

$$m_x^* - 0,3048s < m_x < m_x^* + 0,3048s.$$

Интервальная оценка дисперсии имеет вид

$$\frac{30s^2}{\chi_{30;0,05}^2} < D_x < \frac{30s^2}{\chi_{30;0,95}^2}.$$

Из таблицы процентных точек распределения  $\chi^2$  находим

$$\chi_{30;0,05}^2 = 43,77; \chi_{30;0,95}^2 = 18,49.$$

Поэтому интервальная оценка дается неравенством

$$0,6854s^2 < D_x < 1,622s^2.$$

Теперь остается только вычислить выборочное среднее и выборочную дисперсию и подставить их в интервальные оценки:

$$m_x^* = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{31} \sum_{i=1}^{31} x_i = 58,29;$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x^*)^2 = \frac{1}{30} \left( \sum_{i=1}^{31} x_i^2 - 31(m_x^*)^2 \right) = 31,12.$$

Следовательно, 90%-ные доверительные интервальные оценки для математического ожидания (истинного значения) и дисперсии ошибок измерений имеют вид:

$$56,59 < m_x < 59,99;$$

$$21,33 < D_x < 50,47.$$

Часто экспериментаторы предпочитают приводить результат в виде величины  $m_x^* \pm$  одно среднее квадратическое отклонение, то есть

$$m_x = m_x^* \pm \sqrt{D_{m_x^*}} = m_x^* \pm \frac{s}{\sqrt{n}}.$$

По существу, такая запись содержит указание доверительного интервала при  $t_{\beta; \frac{1-\beta}{2}} = 1$ . Так, для рассмотренного примера получим

$$58,29 - 1,00 < m_x < 58,29 + 1,00$$

или

$$57,29 < m_x < 59,29.$$

Из равенства  $t_{30; \frac{1-\beta}{2}} = 1$  по таблицам процентных точек распределения Стьюдента можно определить  $\beta = 0,68$ . Таким образом, приведенный доверительный интервал будет покрывать истинное значение изменяемой величины с вероятностью  $\beta$  всего 0,68. !

Интервальные оценки можно строить также при одновременной оценке нескольких параметров, когда эти оценки коррелированы. В этом случае пространство параметров, то есть совокупность возможных допустимых комбинаций их значений, не будет одномерным: например, при совместной оценке двух параметров пространство параметров будет представлять собой некоторую область на плоскости.

## Глава 10

# Оценка параметров общей линейной модели измерений

## 10.1. Характеристика общей линейной модели

Пусть имеется  $n$  наблюдаемых  $x_1, x_2, \dots, x_n$  значений случайных величин  $X_1, X_2, \dots, X_n$ . Предположим, что математическое ожидание каждой случайной величины  $X_i$  линейно зависит от  $k$  неизвестных параметров  $a_1, a_2, \dots, a_k$ , так что

$$M[X_i] = a_1 f_{1i} + a_2 f_{2i} + \dots + a_k f_{ki}, \quad i = 1, 2, \dots, n,$$

где  $f_{1i}, f_{2i}, \dots, f_{ki}$  — известные постоянные. В этом случае каждое наблюдаемое значение  $x_i$  можно записать в виде суммы

$$x_i = a_1 f_{1i} + a_2 f_{2i} + \dots + a_k f_{ki} + \varepsilon_i,$$

где  $\varepsilon_1, \dots, \varepsilon_n$  — ошибки измерений.

Общая линейная модель задается последними соотношениями при дополнительных предположениях:

$$M[\varepsilon_i] = 0; \quad D[\varepsilon_i] = \sigma^2,$$

однако дисперсия не предполагается известной; ошибки  $\varepsilon_i$  независимы в совокупности. Задача состоит в том, чтобы по результатам измерений  $x_1, x_2, \dots, x_n$  оценить параметры  $a_1, a_2, \dots, a_k$  и  $\sigma^2$ .

К простейшей задаче, решаемой в рамках общей линейной модели измерений, относится задача измерения какой-то величины  $A$ . В этом случае результат  $i$ -го измерения

$$x_i = A + \varepsilon_i = a \cdot f + \varepsilon_i,$$

где  $f$  — мера измерения (эталон),  $a$  — коэффициент, выражающий долю меры в измеряемой величине.

В последующих разделах будут рассмотрены два примера использования общей линейной модели для решения, казалось бы, разных задач.

## 10.2. Оценка амплитуды сигнала при наличии шума методом максимального правдоподобия

Пусть реализация  $\xi(t)$  на интервале наблюдения  $(0, T)$  представляет собой сумму полезного сигнала заданной формы и стационарного нормального шума с дисперсией  $\sigma^2$ :

$$\xi(t) = s(t, a) + \varepsilon(t) = a s_0(t) + \varepsilon(t),$$

где  $a$  — “амплитуда” сигнала, под которой будем понимать коэффициент, связывающий между собой мгновенные значения сигнала с формой этого сигнала  $s_0(t)$ . Ставится задача оценки амплитуды сигнала по имеющейся реализации (рис. 10.1).

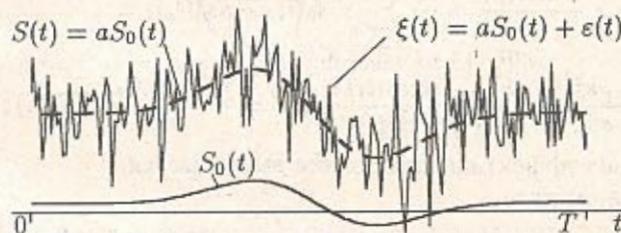


Рис. 10.1. К оценке амплитуды сигнала при наличии шума

Пусть  $x_1, x_2, \dots, x_n$  образуют систему значений, полученных в результате взятия  $n$  отсчетов  $\xi(t)$  в моменты времени  $t_1, t_2, \dots, t_n$ . Отсчеты взяты таким образом, что расстояние между ними превышает время корреляции шума, то есть отсчеты не коррелированы между собой. Поскольку шум нормальный, это означает также независимость

отсчетов. Тогда функция правдоподобия будет иметь вид:

$$L(a) = \left( \frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - as_0(t_i))^2},$$

а уравнение правдоподобия —

$$\left. \frac{\partial \ln L(a)}{\partial a} \right|_{a=\hat{a}} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - as_0(t_i)) s_0(t_i) \Big|_{a=\hat{a}} = 0.$$

Отсюда находим оценку амплитуды  $a$ :

$$\hat{a} = \frac{\sum_{i=1}^n x_i s_0(t_i)}{\sum_{i=1}^n s_0^2(t_i)}.$$

Легко убедиться, что полученная оценка является несмещенной. В самом деле,

$$\begin{aligned} M[\hat{a}] &= \frac{1}{\sum_{n=1}^n s_0^2(t_i)} M\left[\sum_{i=1}^n x_i s_0(t_i)\right] = \frac{1}{\sum_{n=1}^n s_0^2(t_i)} \sum_{i=1}^n M[x_i] s_0(t_i) = \\ &= \frac{1}{\sum_{n=1}^n s_0^2(t_i)} \sum_{i=1}^n as_0(t_i) s_0(t_i) = a. \end{aligned}$$

Эта оценка является также эффективной, поскольку

$$\begin{aligned} \frac{\partial \ln L(a)}{\partial a} &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i s_0(t_i) - as_0^2(t_i)) = \\ &= \frac{\sum_{i=1}^n s_0^2(t_i)}{\sigma^2} \left( \frac{\sum_{i=1}^n x_i s_0(t_i)}{\sum_{i=1}^n s_0^2(t_i)} - a \right) = \frac{\sum_{i=1}^n s_0^2(t_i)}{\sigma^2} (\hat{a} - a), \end{aligned}$$

то есть условие эффективности оценки выполняется.

Дисперсия оценки

$$D_{\hat{a}} = \frac{-1}{M\left[\frac{\partial^2}{\partial a^2} \ln L(a)\right]} = \frac{\sigma^2}{\sum_{i=1}^n s_0^2(t_i)}.$$

Отметим, что  $D_{\hat{a}}$  можно найти непосредственно, применив к оценке  $\hat{a}$  операцию дисперсии. Получим:

$$D[\hat{a}] = D\left[\frac{\sum_{i=1}^n x_i s_0(t_i)}{\sum_{i=1}^n s_0^2(t_i)}\right] = \frac{1}{\left(\sum_{i=1}^n s_0^2(t_i)\right)^2} D\left[\sum_{i=1}^n x_i s_0(t_i)\right] =$$

$$= \frac{1}{\left(\sum_{i=1}^n s_0^2(t_i)\right)^2} \sum_{i=1}^n s_0^2(t_i) D[X_i] = \frac{\sigma^2}{\sum_{i=1}^n s_0^2(t_i)}.$$

Можно найти также интервальную оценку амплитуды. Для этого нужно определить сначала закон распределения оценки  $\hat{a}$ . Эта оценка представляет собой линейную комбинацию взаимно независимых случайных величин  $X_i$ , распределенных по нормальному закону. Следовательно, она тоже имеет нормальное распределение. Математическое ожидание оценки  $\hat{a}$  равно истинному значению параметра  $a$ , а дисперсия дается последним выражением. В связи с этим интервальная оценка запишется в виде

$$\hat{a} - z_{\beta} \frac{\sigma}{\sqrt{\sum_{i=1}^n s_0^2(t_i)}} < a < \hat{a} + z_{\beta} \frac{\sigma}{\sqrt{\sum_{i=1}^n s_0^2(t_i)}},$$

где  $z_{\beta}$  находится из уравнения

$$\Phi(z_{\beta}) = \frac{1+\beta}{2}.$$

Если дисперсия шума неизвестна, то ее можно оценить, решив систему уравнений правдоподобия:

$$\left. \frac{\partial \ln L(a, \sigma^2)}{\partial a} \right|_{a=\hat{a}, \sigma^2=\hat{\sigma}^2} = 0,$$

$$\left. \frac{\partial \ln L(a, \sigma^2)}{\partial (\sigma^2)} \right|_{a=\hat{a}, \sigma^2=\hat{\sigma}^2} = 0.$$

В результате после устранения смещения получим

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{a}s_0(t_i))^2.$$

### 10.3. Совместная оценка амплитуд нескольких сигналов

Пусть на интервале  $(0, T)$  на фоне нормального шума присутствует несколько сигналов известных форм, амплитуды которых подлежат оценке, то есть

$$\xi(t) = a_j s_{0j}(t) + \varepsilon(t), \quad j = 1, 2, \dots, k.$$

Тогда функция правдоподобия будет иметь вид:

$$L(a_1, \dots, a_k) = \left( \frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - a_1 s_{01}(t_i) - \dots - a_k s_{0k}(t_i))^2\right).$$

Оценки амплитуд сигналов  $a_1, a_2, \dots, a_k$  находятся в результате решения системы уравнений правдоподобия следующего вида:

$$\begin{aligned} \sum_{i=1}^n [\hat{a}_1 s_{01}(t_i) + \hat{a}_2 s_{02}(t_i) + \dots + \hat{a}_k s_{0k}(t_i)] s_{01}(t_i) &= \sum_{i=1}^n x_i s_{01}(t_i), \\ \sum_{i=1}^n [\hat{a}_1 s_{01}(t_i) + \hat{a}_2 s_{02}(t_i) + \dots + \hat{a}_k s_{0k}(t_i)] s_{02}(t_i) &= \sum_{i=1}^n x_i s_{02}(t_i), \\ &\dots \\ \sum_{i=1}^n [\hat{a}_1 s_{01}(t_i) + \hat{a}_2 s_{02}(t_i) + \dots + \hat{a}_k s_{0k}(t_i)] s_{0k}(t_i) &= \sum_{i=1}^n x_i s_{0k}(t_i). \end{aligned}$$

Эта система уравнений является линейной и может быть представлена в матричном виде:

$$SA = H.$$

Здесь  $A$  — матрица-столбец, составленная из оценок параметров  $\hat{a}_j$  ( $j = 1, 2, \dots, k$ );  $H$  — матрица-столбец, составленная из  $\sum_{i=1}^n x_i s_{0j}(t_i)$  ( $j = 1, 2, \dots, k$ ):

$$A = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \dots \\ \hat{a}_k \end{bmatrix}, \quad H = \begin{bmatrix} \sum_{i=1}^n x_i s_{01}(t_i) \\ \sum_{i=1}^n x_i s_{02}(t_i) \\ \dots \\ \sum_{i=1}^n x_i s_{0k}(t_i) \end{bmatrix}.$$

Матрица  $S$  имеет вид:

$$S = \begin{bmatrix} \sum_{i=1}^n s_{01}(t_i) s_{01}(t_i) & \sum_{i=1}^n s_{01}(t_i) s_{02}(t_i) & \dots & \sum_{i=1}^n s_{01}(t_i) s_{0k}(t_i) \\ \sum_{i=1}^n s_{02}(t_i) s_{01}(t_i) & \sum_{i=1}^n s_{02}(t_i) s_{02}(t_i) & \dots & \sum_{i=1}^n s_{02}(t_i) s_{0k}(t_i) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n s_{0k}(t_i) s_{01}(t_i) & \sum_{i=1}^n s_{0k}(t_i) s_{02}(t_i) & \dots & \sum_{i=1}^n s_{0k}(t_i) s_{0k}(t_i) \end{bmatrix}.$$

Решение системы уравнений правдоподобия имеет вид:

$$A = S^{-1}H,$$

где  $S^{-1}$  — матрица, обратная матрице  $S$ . Тогда оценка  $j$ -й амплитуды  $\hat{a}_j$  находится по формуле

$$\hat{a}_j = \frac{\Delta_j}{\Delta},$$

где  $\Delta$  — определитель матрицы  $S$ , а  $\Delta_j$  — определитель, полученный из определителя  $\Delta$  заменой  $j$ -го столбца на столбец  $H$ .

Для случая двух сигналов можно получить выражения для оценок амплитуд  $a_1$  и  $a_2$  в виде:

$$\begin{aligned} \hat{a}_1 &= \frac{\sum_{i=1}^n s_{01}(t_i) x_i \sum_{i=1}^n s_{02}^2(t_i) - \sum_{i=1}^n s_{02}(t_i) x_i \sum_{i=1}^n s_{01}(t_i) s_{02}(t_i)}{\sum_{i=1}^n s_{01}^2(t_i) \sum_{i=1}^n s_{02}^2(t_i) - (\sum_{i=1}^n s_{01}(t_i) s_{02}(t_i))^2}, \\ \hat{a}_2 &= \frac{\sum_{i=1}^n s_{02}(t_i) x_i \sum_{i=1}^n s_{01}^2(t_i) - \sum_{i=1}^n s_{01}(t_i) x_i \sum_{i=1}^n s_{01}(t_i) s_{02}(t_i)}{\sum_{i=1}^n s_{01}^2(t_i) \sum_{i=1}^n s_{02}^2(t_i) - (\sum_{i=1}^n s_{01}(t_i) s_{02}(t_i))^2}. \end{aligned}$$

Можно показать, что полученные оценки являются несмещенными и эффективными. Найдем дисперсии оценок амплитуд для случая двух сигналов. Элементами информационной матрицы Фишера в этом случае будут

$$\begin{aligned} J_{11} &= \frac{1}{\sigma^2} \sum_{i=1}^n s_{01}^2(t_i); \quad J_{22} = \frac{1}{\sigma^2} \sum_{i=1}^n s_{02}^2(t_i); \\ J_{12} &= J_{21} = \frac{1}{\sigma^2} \sum_{i=1}^n s_{01}(t_i) s_{02}(t_i). \end{aligned}$$

На основании выражений

$$D_{\hat{a}_1} = R_{11} = \frac{J_{22}}{J_{11}J_{22} - (J_{12})^2};$$

$$D_{\hat{a}_2} = R_{22} = \frac{J_{11}}{J_{11}J_{22} - (J_{12})^2};$$

получим

$$D_{\hat{a}_1} = \frac{\sigma^2 \sum_{i=1}^n s_{02}^2(t_i)}{\sum_{i=1}^n s_{01}^2(t_i) \sum_{i=1}^n s_{02}^2(t_i) - (\sum_{i=1}^n s_{01}(t_i) s_{02}(t_i))^2},$$

$$D_{\hat{a}_2} = \frac{\sigma^2 \sum_{i=1}^n s_{01}^2(t_i)}{\sum_{i=1}^n s_{01}^2(t_i) \sum_{i=1}^n s_{02}^2(t_i) - (\sum_{i=1}^n s_{01}(t_i) s_{02}(t_i))^2}.$$

Отметим, что полученные оценки являются коррелированными.

Можно показать, что при неизвестной дисперсии  $\sigma^2$  ее несмещенная оценка находится из выражения

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n (x_i - \hat{a}_1 s_{01}(t_i) - \dots - \hat{a}_k s_{0k}(t_i))^2.$$

#### 10.4. Задачи, связанные с оценкой амплитуд

Выше был рассмотрен случай сигналов, являющихся функцией времени. Однако под сигналом можно понимать и изменение какой-то величины в функции произвольного параметра — например, в функции пространственных координат, частоты и т.д. Так, регистрируемая спектрометром спектральная характеристика какого-либо источника излучения представляет собой зависимость интенсивности от пространственных координат, которые обычно переводятся в длины волн.

Любой оптический спектр записывается при наличии шумов на выходе регистрирующего прибора. В связи с этим можно ставить, например, задачу оценки амплитуды спектральной линии, связанной с концентрацией частиц, вызывающих излучение, при условии, что известна форма спектральной линии, отвечающая известной концентрации частиц и полученная при той же геометрии измерений.

В качестве сигнала, но не в реальном масштабе времени можно рассматривать и спектр источника  $\gamma$ -излучения, зарегистрированный соответствующей спектрометрической аппаратурой (многоканальным амплитудным анализатором). При этом статистика отсчетов в каждом канале может отличаться от нормальной, или, если она все же близка к нормальной, дисперсии отсчетов в каждом канале могут быть различными.

Если имеет место второй случай и анализу подвергается спектр только одного источника излучения, то на основе проведенных ранее выкладок можно получить следующее выражение для оценки амплитуды сигнала:

$$\hat{a} = \frac{\sum_{i=1}^n \frac{1}{\sigma_i^2} s_{0i} x_i}{\sum_{i=1}^n \frac{1}{\sigma_i^2} s_{0i}^2}.$$

Дисперсия этой оценки будет равна

$$D_{\hat{a}} = \frac{1}{\sum_{i=1}^n \frac{s_{0i}^2}{\sigma_i^2}}.$$

Обратим внимание на то, что в рассмотренной общей линейной модели оценки параметров  $\hat{a}_j$  представляют собой линейные функции от элементов выборки (результатов наблюдений). Такие оценки называются *линейными*.

#### 10.5. Аппроксимация экспериментальных данных по методу наименьших квадратов

В рамках общей линейной модели оценки параметров чаще всего находят так называемым *методом наименьших квадратов* (МНК). При использовании этого метода не требуется знания закона распределения ошибок.

Применительно к рассмотренному ранее примеру оценки амплитуды сигнала на фоне шума в МНК должна минимизироваться следующая функция:

$$\sum_{i=1}^n (x_i - a s_0(t_i))^2 = \min_a. \quad (10.1)$$

Таким образом, следует найти такое  $a$ , чтобы сумма квадратов отклонений выборочных значений от истинных была минимальной. Отсюда и происходит название "метод наименьших квадратов".

Чтобы выражение (10.1) имело минимум, нужно найти производную по  $a$  и приравнять ее к нулю. Тогда получим следующее уравнение:

$$\sum_{i=1}^n (x_i - a s_0(t_i)) s_0(t_i) = 0.$$

Это уравнение, как легко видеть, совпадает с уравнением правдоподобия при условии, что ошибки распределены по нормальному закону. Таким образом, МНК следует из метода максимального правдоподобия, если результаты наблюдений (измерений) имеют

совместное нормальное распределение. Оценки нескольких параметров  $a_1, a_2, \dots, a_k$  находятся из уравнений, которые были получены ранее при рассмотрении метода максимального правдоподобия при нормальном законе распределения ошибок.

Применительно к линейной модели вида

$$x_i = a_1 f_{1i} + a_2 f_{2i} + \dots + a_k f_{ki} + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

оценки, получаемые методом наименьших квадратов, обладают следующими тремя важными свойствами, которые не зависят от закона распределения ошибок:

- 1) оценки  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$  являются несмещенными оценками параметров  $a_1, a_2, \dots, a_k$ ;
- 2) в соответствии с теоремой Гаусса — Маркова среди существующих несмещенных линейных оценок оценки, полученные методом наименьших квадратов, имеют минимальную дисперсию;
- 3) оценка дисперсии ошибок дается выражением

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n [x_i - a_1 f_{1i} - \dots - a_k f_{ki}]^2.$$

Наиболее важное применение метода наименьших квадратов для нахождения параметров линейной модели связано с задачей аппроксимации экспериментальной зависимости. Ее можно сформулировать следующим образом. Пусть задана детерминированная функция

$$y = a_0 f_0(x) + a_1 f_1(x) + \dots + a_m f_m(x),$$

зависящая от  $m$  параметров  $a_1, a_2, \dots, a_m$ , значения которых неизвестны. По результатам эксперимента, то есть значениям  $y_1, y_2, \dots, y_n$ , в выбранных точках  $x_1, x_2, \dots, x_n$  нужно оценить неизвестные параметры  $a_1, a_2, \dots, a_m$ . Предполагается, что значения аргумента  $x = x_i$  могут быть установлены точно. Искомая функция, например, может иметь вид многочлена

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m.$$

Рассмотрим в качестве примера оценку параметров линейной функции. Пусть в результате эксперимента получена совокупность

значений  $(x_i, y_i)$ ,  $i = 1, \dots, n$  (рис. 10.2). Необходимо с помощью МНК подобрать параметры  $a_0$  и  $a_1$  линейной функции

$$y = a_0 + a_1 x,$$

характеризующей данную экспериментальную зависимость.

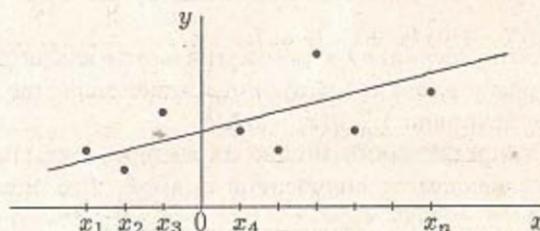


Рис. 10.2. Линейная аппроксимация

Применительно к рассматриваемому случаю требуется минимизировать по  $a_0$  и  $a_1$  выражение

$$\sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2.$$

Приравнивая к нулю результат дифференцирования по  $a_0$  и  $a_1$ , получаем систему двух линейных уравнений:

$$\sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0;$$

$$\sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i) x_i = 0.$$

Из первого уравнения находим

$$\hat{a}_0 = m_y^* - \hat{a}_1 m_x^*,$$

где  $m_y^* = \frac{1}{n} \sum_{i=1}^n y_i$ ,  $m_x^* = \frac{1}{n} \sum_{i=1}^n x_i$ .

Подставив найденное значение  $\hat{a}_0$  во второе уравнение и решив его относительно  $\hat{a}_1$ , получим

$$\hat{a}_1 = \frac{\sum_{i=1}^n x_i y_i - n m_x^* m_y^*}{\sum_{i=1}^n x_i^2 - n (m_x^*)^2} = \frac{\sum_{i=1}^n (x_i - m_x^*)(y_i - m_y^*)}{\sum_{i=1}^n (x_i - m_x^*)^2}.$$

Таким образом, линейная функция

$$y = m_y^* + \frac{\sum_{i=1}^n (x_i - m_x^*)(y_i - m_y^*)}{\sum_{i=1}^n (x_i - m_x^*)^2} (x - m_x^*)$$

наилучшим образом (среди всех линейных функций) выражает зависимость  $y$  от  $x$ .

Отметим, что, поскольку  $x$  рассматривается как неслучайная величина,  $m_x^*$  представляет собой просто среднее значений  $x_i$ . Не имеет смысла также величина  $\sum_{i=1}^n (x_i - m_x^*)^2$ .

Качество аппроксимации можно характеризовать среднеквадратическим отклонением от полученной прямой. Его оценкой служит величина

$$s = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i)^2}{n - 2}},$$

где цифра 2 в знаменателе — число оцененных параметров.

Отметим, что задачи, связанные с подбором аппроксимирующих функций, решаются в разделе математической статистики, называемом регрессионным анализом.

## Глава 11

# Проверка статистических гипотез

### 11.1. Нулевая и альтернативная гипотезы

Статистическая гипотеза — это утверждение относительно значений одного или нескольких параметров распределения или о самом виде распределения. Гипотеза, которая проверяется, называется *нулевой гипотезой* и обозначается  $H_0$ . *Альтернативной гипотезой*  $H_1$  называется гипотеза, конкурирующая с нулевой, то есть противоречащая ей.

Рассмотрим примеры нулевых и альтернативных статистических гипотез. Пусть проверяется гипотеза о равенстве параметра  $a$  некоторому значению  $a_0$ , то есть нулевая гипотеза  $H_0: a = a_0$ . Тогда в качестве альтернативной гипотезы  $H_1$  можно рассматривать одну из следующих гипотез:

$$H_1: a > a_0;$$

$$H_1: a < a_0;$$

$$H_1: a \neq a_0;$$

$$H_1: a = a_1.$$

Простой называют гипотезу, содержащую только одно предположение. Так, приведенная выше гипотеза  $H_0: a = a_0$  — простая. Сложной называют гипотезу, которая состоит из конечного или бесконечного числа простых гипотез. Например, сложной будет гипотеза  $H_0: a > 5$ .

Выбор альтернативной гипотезы определяется конкретной формулировкой задачи. Причина выделения нулевой гипотезы состоит в том, что она обычно рассматривается как утверждение, которое более важно, если оно отвергнуто. Это основано на общем принципе, в соответствии с которым теория должна быть отвергнута, если есть

противоречащий ей факт, но не обязательно должна быть принята, если противоречащих ей фактов на текущий момент не установлено. Правило, по которому принимается решение принять или отклонить гипотезу  $H_0$ , называется *статистическим критерием*.

Проверку статистических гипотез осуществляют на основании результатов выборки, из которых формируют функцию выборки (функцию результатов наблюдений), называемую *проверочной статистикой*. Следовательно, статистический критерий устанавливает, при каких значениях проверочной статистики проверяемая гипотеза принимается, а при каких она отвергается. Значения проверочной статистики, при которых гипотеза  $H_0$  отвергается, образуют критическую область проверяемой гипотезы.

*Рассмотрим пример.* Пусть на выходе приемного устройства присутствует случайное напряжение  $\xi(t)$ , которое представляет собой либо сумму известного постоянного напряжения величиной  $a$ , вызванного наличием на входе сигнала и шума, либо один шум. В какой-то момент времени  $t_0$  производится измерение величины  $X = \xi(t_0)$ , то есть фиксируется принятое ею значение. По этому значению нужно решить, присутствовал ли на входе приемного устройства сигнал, то есть выбрать одну из двух возможностей.

В качестве нулевой гипотезы  $H_0$  возьмем гипотезу, соответствующую отсутствию сигнала, а в качестве альтернативной  $H_1$  — гипотезу о его наличии, то есть будем считать, что отсутствие сигнала является более важным утверждением (именно так поступают, например, в радиолокации).

Пусть шумовое напряжение распределено по нормальному закону с нулевым математическим ожиданием и известной дисперсией. Изобразим графики плотностей вероятностей при нулевой и альтернативной гипотезах (рис. 11.1).

Предположим, что измеренное значение случайной величины  $X$  оказалось равным  $x_n$ . Экспериментатор должен решить, с какой из гипотез следует связать это значение. На основании определенных соображений он должен выбрать критическое значение  $x_{кр}$ , задав тем самым критическую область проверяемой гипотезы. Если измеренное значение  $x_n$  превосходит  $x_{кр}$  (попадает в критическую область), то он отвергает гипотезу  $H_0$ . Если же  $x_n$  не превышает  $x_{кр}$  (не попадает в критическую область), то гипотеза  $H_0$  не отвергается.

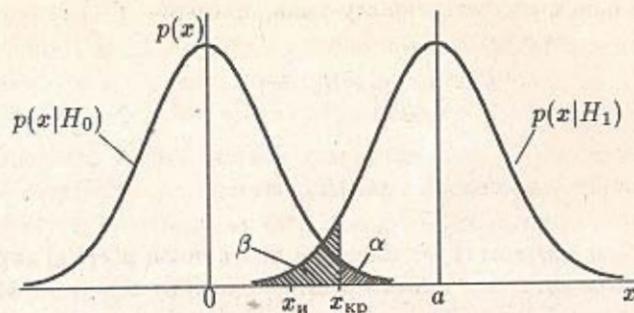


Рис. 11.1. Плотность вероятности  $p(x)$  при нулевой и альтернативной гипотезах

Следует иметь в виду, что решение об отклонении или неотклонении гипотезы  $H_0$  основано на выборочном наблюдении и может быть ошибочным. Если измеренное значение не попадает в критическую область, это не является доказательством справедливости нулевой гипотезы, а лишь свидетельствует о том, что измеренное значение ей не противоречит. Принятие же нулевой гипотезы осуществляется с определенным уровнем значимости.

## 11.2. Уровень значимости и мощность критерия

При проверке статистических гипотез в общем случае возможны два типа ошибок, связанных с принятием решения. Если в действительности гипотеза  $H_0$  верна, а принято решение ее отвергнуть, то говорят, что допущена *ошибка первого рода*. С другой стороны, если в действительности верна гипотеза  $H_1$ , а принято решение принять гипотезу  $H_0$ , то говорят, что допущена *ошибка второго рода*.

Обозначим вероятность ошибки первого рода через  $\alpha$ , а вероятность ошибки второго рода через  $\beta$ . Тогда

$$\alpha = P(\text{отвергнуть } H_0 | H_0 \text{ верна});$$

$$\beta = P(\text{принять } H_0 | H_0 \text{ ложна}).$$

Применительно к рассмотренному выше примеру

$$\alpha = \int_{x_{кр}}^{\infty} p(x|H_0) dx;$$

$$\beta = \int_{-\infty}^{x_{кр}} p(x|H_1) dx.$$

Вероятность  $\alpha$  соответствует площади под кривой  $p(x|H_0)$  справа от  $x_{кр}$ , а  $\beta$  — площади под кривой  $p(x|H_1)$  слева от  $x_{кр}$  (рис. 11.1). Легко видеть, что  $\alpha$  — это вероятность принять за сигнал шум, а  $\beta$  — вероятность принять неверное решение, отвергающее наличие сигнала. Применительно к радиолокации говорят, что  $\alpha$  — это вероятность ложной тревоги, а  $\beta$  — вероятность пропуска сигнала.

Статистическая задача состоит в том, чтобы найти правило (*критерий*), по которому принимается решение принять или отклонить гипотезу и которое минимизирует ошибки первого и второго рода.

Вероятность  $\alpha$  отвергнуть нулевую гипотезу  $H_0$ , когда она является справедливой, в математической статистике называют уровнем значимости критерия. Вероятность  $\gamma = 1 - \beta$  не допустить ошибку второго рода (отклонить нулевую гипотезу) называют *мощностью критерия*. Для нахождения мощности критерия необходимо знать плотность вероятности проверочной статистики при альтернативной гипотезе. Простые критерии с заданным уровнем значимости контролируют лишь ошибки первого рода и не учитывают мощность критерия.

### 11.3. Проверка гипотез о математическом ожидании

Предположим, что случайная величина  $X$  распределена по нормальному закону с неизвестным математическим ожиданием  $m_x$  и известной дисперсией  $D_x = \sigma^2$ . Пусть значения  $x_1, x_2, \dots, x_n$  — случайная выборка из этого распределения.

Рассмотрим проверку гипотезы  $H_0: m_x = a$ , то есть предположения о равенстве неизвестного математического ожидания  $m_x$  значению  $a$  при альтернативной гипотезе  $H_1: m_x = a_1$  ( $a_1 > a$ ) при заданном уровне значимости  $\alpha$ .

Критическая область для нахождения критерия определяется условием  $m_x^* \geq m_{кр}^*$ , где  $m_x^*$  — выборочное среднее, а  $m_{кр}^*$  выбрано так, что

$$P(m_x^* \geq m_{кр}^* | H_0) = \alpha.$$

Как известно, выборочное среднее  $m_x^* = \frac{1}{n} \sum_{i=1}^n x_i$  имеет нормальное распределение с математическим ожиданием  $M[m_x^*] = a$  и дисперсией  $D[m_x^*] = \frac{\sigma^2}{n}$ . Поэтому  $m_{кр}^*$  находится из условия

$$\int_{m_{кр}^*}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} e^{-\frac{(m_{кр}^* - a)^2}{2\sigma^2/n}} dm_x^* = \alpha.$$

В качестве проверочной статистики удобно ввести случайную величину

$$Z = \frac{(m_x^* - a)\sqrt{n}}{\sigma}.$$

Тогда при справедливости гипотезы  $H_0$  случайная величина  $Z$  имеет нормальное распределение с нулевым математическим ожиданием и дисперсией, равной 1, а при справедливости гипотезы  $H_1$  — нормальное распределение с математическим ожиданием  $m_{z|H_1} = \frac{a_1 - a}{\sigma} \sqrt{n}$  и единичной дисперсией (рис. 11.2).

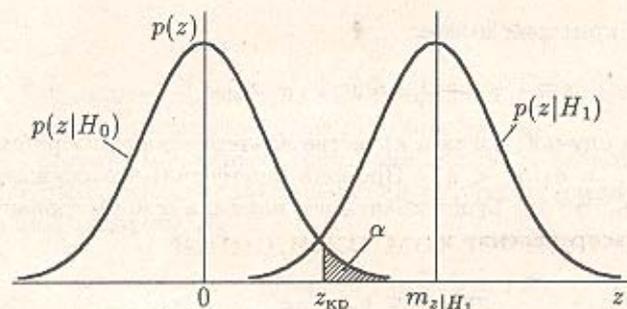


Рис. 11.2. Плотность вероятности  $p(z)$  при нулевой и альтернативной гипотезах в случае правосторонней критической области

Критическое значение  $z_{кр}$  находится из условия

$$\int_{z_{кр}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 - \int_{-\infty}^{z_{кр}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 - \Phi(z_{кр}) = \alpha.$$

Таким образом, критическим значением случайной величины  $z$  для проверяемой гипотезы является  $(1 - \alpha)$ -й квантиль нормированного нормального распределения  $z_{1-\alpha}$ . Если рассчитанное значение  $z > z_{1-\alpha}$ , то нулевая гипотеза отклоняется с таким же уровнем значимости. В противном случае нулевая гипотеза принимается с уровнем значимости  $\alpha$ .

Переходя к случайной величине  $m_x^*$ , получим, что ее критическое значение

$$m_{кр}^* = a + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}.$$

Итак, критической областью для проверки гипотезы  $H_0$  является правый хвост распределения  $p(z|H_0)$  или распределения  $p(m_x^*|H_0)$ .

Вычислим вероятность совершить ошибку второго рода при принятии гипотезы  $H_0$ :

$$\begin{aligned} \beta &= P(z < z_{кр} | H_1) = \int_{-\infty}^{z_{кр}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z - m_{z|H_1})^2}{2}} dz = \\ &= \int_{-\infty}^{z_{кр} - m_{z|H_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \Phi(z_{кр} - m_{z|H_1}) = \Phi(z_{1-\alpha} - m_{z|H_1}). \end{aligned}$$

Для мощности критерия имеем

$$\gamma = 1 - \beta = 1 - \Phi(z_{1-\alpha} - m_{z|H_1}).$$

Рассмотрим случай, когда в качестве альтернативной берется гипотеза  $H_1: m_x = a_1 (a_1 < a)$ . Проводя аналогичные рассуждения, легко получить, что  $z_{кр}$  будет квантилем порядка  $\alpha$  нормированного нормального распределения и

$$m_{кр}^* = a + z_{\alpha} \frac{\sigma}{\sqrt{n}}.$$

Соответствующие распределения  $p(z|H_0)$  и  $p(z|H_1)$  показаны на рис. 11.3. В качестве критической области в данном случае берется левый хвост распределения  $p(z|H_0)$ .

*Рассмотрим пример.* Из нормальной генеральной совокупности с дисперсией  $\sigma^2 = 5$  взята выборка, состоящая из девяти наблюдений. Требуется при уровне значимости  $\alpha = 0,05$  проверить гипотезу

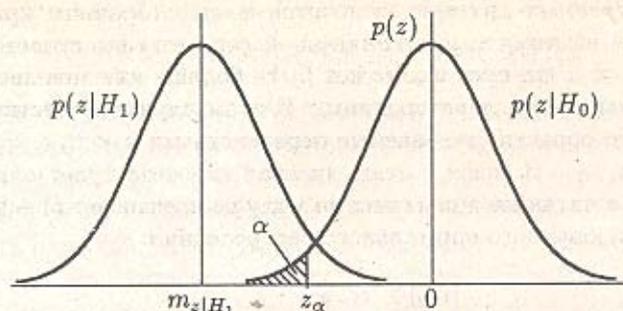


Рис. 11.3. Плотности вероятности при нулевой и альтернативной гипотезах в случае левосторонней критической области

$H_0: m_x = 2$  при альтернативной гипотезе  $H_1: m_x = 3$ . Оценка математического ожидания по выборке дала значение 2,3.

Поскольку значение математического ожидания при альтернативной гипотезе больше, чем при нулевой, то критическая область является правосторонней. По таблице интеграла вероятностей находим квантиль порядка  $1 - \alpha = z_{1-\alpha} = z_{0,95} = 1,65$ .

Вычисляем значение проверочной статистики

$$z = \frac{2,3 - 2}{\sqrt{5}} \sqrt{9} = 0,4.$$

Поскольку это значение не превышает критического 1,65, то нет оснований отвергнуть проверяемую гипотезу.

Рассчитаем мощность критерия. Сначала вычислим математическое ожидание случайной величины  $z$  при справедливости альтернативной гипотезы  $H_1$ :

$$m_{z|H_1} = \frac{3 - 2}{\sqrt{5}} \sqrt{9} = 1,33.$$

Тогда мощность критерия

$$\begin{aligned} \gamma &= 1 - \Phi(z_{1-\alpha} - 1,33) = 1 - \Phi(1,65 - 1,33) = \\ &= 1 - \Phi(0,32) = 1 - 0,6255 = 0,3745. \end{aligned}$$

Нетрудно убедиться, что для большего объема выборки мощность критерия была бы больше.

Оба рассмотренных критерия относятся к односторонним критериям. Если же в качестве альтернативной гипотезы выступает гипотеза  $H_1: m_x \neq a$  (то есть  $m_x$  может быть больше или меньше  $a$ ), то критерий становится двусторонним. В этом случае существуют две критические области, задаваемые неравенствами  $z < z_{кр1} = z_{\frac{\alpha}{2}}$  и  $z > z_{кр2} = z_{1-\frac{\alpha}{2}}$ . В связи с этим нулевая гипотеза будет отклоняться, если рассчитанное значение  $z$  по модулю превышает  $(1 - \frac{\alpha}{2})$ -й квантиль нормированного нормального распределения:

$$|z_{кр}| > z_{1-\frac{\alpha}{2}}.$$

Переходя к случайной величине  $m_x^*$ , получим для нее два критических значения:

$$m_{кр1}^* = a - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}};$$

$$m_{кр2}^* = a + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}.$$

Эти критические значения совпадают с границами доверительного интервала для неизвестного математического ожидания при известной дисперсии, поскольку  $z_{1-\frac{\alpha}{2}} = z_{1-\frac{1-\beta}{2}} = z_{\frac{1+\beta}{2}}$ , где  $\beta$  — доверительная вероятность.

Рассмотрим теперь случай, когда случайная величина  $X$  распределена по нормальному закону с математическим ожиданием  $a$ , но неизвестной дисперсией. В связи с этим для проверки гипотезы необходимо найти предварительно оценку дисперсии:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x^*)^2.$$

В качестве проверочной статистики возьмем случайную величину

$$T = \frac{m_x^* - a}{s} \sqrt{n},$$

которая, как известно, имеет  $t$ -распределение Стьюдента с  $l = n - 1$  степенями свободы. Тогда получим следующие критерии проверки гипотезы  $H_0: m_x = a$ . При  $H_1: m_x > a$  гипотеза  $H_0$  принимается, если  $t < t_{l;\alpha}$ ; при  $H_1: m_x < a$ , если  $t > t_{l;1-\alpha}$ ; при  $H_1: m_x \neq a$ , если  $|t| < t_{l;\frac{\alpha}{2}}$ . Здесь  $t_{l;\alpha}$  —  $\alpha$ -процентная точка распределения Стьюдента с  $l$  степенями свободы.

## 11.4. Проверка гипотезы о равенстве двух выборочных средних

В исследовательской работе большое значение имеет воспроизводимость результатов измерений. Серии экспериментов неоднократно повторяют. Иногда оказывается, что среднее значение результата в одной серии экспериментов заметно отличается от такого же показателя в другой серии. При этом возникает вопрос, можно ли обнаруженное расхождение средних объяснить случайными ошибками или оно вызвано какими-либо незамеченными или даже неизвестными закономерностями.

Пусть проведены две серии независимых наблюдений над случайной величиной  $X$ , то есть получены две выборки. Выборка 1 имеет объем  $n_1$ :  $x_{11}, x_{21}, \dots, x_{n_11}$ , выборка 2 — объем  $n_2$ :  $x_{12}, x_{22}, \dots, x_{n_22}$ . Предположим, что эти выборки взяты из генеральных совокупностей, имеющих нормальные распределения с математическими ожиданиями  $m_1$  и  $m_2$  и дисперсиями  $\sigma_1^2$  и  $\sigma_2^2$ .

По результатам выборки рассчитаны выборочные средние значения  $m_1^*$  и  $m_2^*$ . Ставится задача проверки гипотезы  $H_0: m_1 = m_2$  при альтернативной гипотезе  $H_1: m_1 \neq m_2$ .

Рассмотрим сначала случай, когда дисперсии  $\sigma_1^2$  и  $\sigma_2^2$  известны. Тогда можно воспользоваться тем, что при гипотезе  $H_0$  выборочные средние  $m_1^*$  и  $m_2^*$  имеют нормальные распределения с параметрами соответственно  $(m_1, \frac{\sigma_1^2}{n_1})$  и  $(m_2, \frac{\sigma_2^2}{n_2})$ . Так как выборочные средние независимы, то их разность тоже имеет нормальное распределение с параметрами  $m_1 - m_2$  и  $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ . Поэтому нормированная разность выборочных средних, то есть величина

$$Z = \frac{m_1^* - m_2^*}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

при нулевой гипотезе имеет нормальное распределение с нулевым математическим ожиданием и единичной дисперсией. Эта величина и берется в качестве проверочной статистики.

Критическими для проверяемой гипотезы являются значения

$$|z| > z_{1-\frac{\alpha}{2}},$$

где  $z_{1-\frac{\alpha}{2}}$  — квантиль порядка  $(1 - \frac{\alpha}{2})$  нормального распределения.

Теперь рассмотрим случай, когда проверяется такая же гипотеза, но дисперсии  $\sigma_1^2$  и  $\sigma_2^2$  неизвестны. Эти дисперсии необходимо сначала оценить по выборкам. Предположим, что оценки оказались равными  $s_1^2$  и  $s_2^2$ . Кроме того, пусть различие этих оценок оказалось незначимым (проверка гипотез относительно дисперсий будет рассмотрена дальше). Тогда полагаем, что  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , и для получения более точной оценки уже одной генеральной совокупности используем результаты обеих выборок. Такую оценку называют объединенной оценкой дисперсии:

$$s_{об}^2 = \frac{\sum_{i=1}^{n_1} (x_{i1} - m_1^*)^2 + \sum_{j=1}^{n_2} (x_{j2} - m_2^*)^2}{n_1 - 1 + n_2 - 1} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

Эта оценка имеет распределение  $\chi^2$  с  $l = n_1 + n_2 - 2$  степенями свободы.

В качестве статистики для проверки гипотезы возьмем случайную величину

$$T = \frac{m_1^* - m_2^*}{s_{об} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

которая имеет  $t$ -распределение Стьюдента с  $n_1 + n_2 - 2$  степенями свободы.

Критическими для проверяемой гипотезы являются значения

$$|t| > t_{n_1+n_2-2; \frac{\alpha}{2}},$$

где  $t_{n_1+n_2-2; \frac{\alpha}{2}}$  —  $\frac{\alpha}{2}$ -процентная точка распределения Стьюдента с  $n_1 + n_2 - 2$  степенями свободы.

Если в качестве альтернативных к гипотезе  $H_0: m_1 = m_2$  выступают гипотезы  $H_1: m_1 > m_2$  или  $H_1: m_1 < m_2$ , то критическими значениями являются, соответственно,

$$t_{кр} = t_{n_1+n_2-2; \alpha} \quad \text{и} \quad t_{кр} = -t_{n_1+n_2-2; \alpha}.$$

*Рассмотрим пример.* Из двух нормальных генеральных совокупностей 1 и 2 взяты выборки объемами, соответственно,  $n_1 = 10$  и  $n_2 = 15$ , по которым оценены математические ожидания совокупностей  $m_1^* = 5$  и  $m_2^* = 7,5$ . Дисперсии генеральных совокупностей известны:  $\sigma_1^2 = 20$  и  $\sigma_2^2 = 30$ . Требуется при уровне значимости 0,1 проверить нулевую гипотезу  $H_0: m_1 = m_2$  при альтернативной гипотезе  $H_1: m_1 \neq m_2$ .

Найдем значение проверочной статистики  $Z$ :

$$z = \frac{5 - 7,5}{\sqrt{\frac{20}{10} + \frac{30}{15}}} = -\frac{2,5}{\sqrt{4}} = -1,25.$$

По таблице интеграла вероятности находим значение квантиля

$$z_{1-\frac{\alpha}{2}} = z_{1-0,05} = z_{0,95} = 1,64.$$

Так как  $|z| < z_{1-\frac{\alpha}{2}}$ , то нулевая гипотеза не отвергается, то есть математические ожидания генеральных совокупностей не отличаются с уровнем значимости больше 0,1.

Пусть теперь в этом же примере дисперсии  $\sigma_1^2$  и  $\sigma_2^2$  неизвестны, а их оценки по выборкам оказались равными:  $s_1^2 = 18$  и  $s_2^2 = 26$ . Предположим, что эти оценки дисперсий незначимо отличаются друг от друга.

Тогда вычисляем объединенную оценку дисперсии:

$$s_{об}^2 = \frac{10 \cdot 17 + 15 \cdot 27}{10 + 15 - 2} = 25.$$

Рассчитываем значение статистики  $T$ :

$$t = \frac{5 - 7,5}{s \sqrt{\frac{1}{10} + \frac{1}{15}}} \approx -1,225.$$

По таблице процентных точек распределения Стьюдента находим критические значения статистики  $T$ :

$$t_{n_1+n_2-2; \frac{\alpha}{2}} = t_{23; 0,05} = 1,714.$$

Итак, и в этом случае гипотеза о равенстве математических ожиданий не отвергается.

## 11.5. Проверка гипотезы о равенстве дисперсий двух совокупностей

Гипотезы о дисперсиях играют в экспериментах исключительно важную роль, поскольку только знание степени рассеяния полученных результатов и позволяет судить о надежности измерений. Рассмотрим задачу сравнения двух дисперсий на основании их оценок

по выборкам. Предположим, что имеются две выборки объемами  $n_1$  и  $n_2$ , извлеченные из генеральных совокупностей, имеющих нормальные распределения с математическими ожиданиями  $m_1$  и  $m_2$  и дисперсиями  $\sigma_1^2$  и  $\sigma_2^2$ . Выборки независимы. По выборкам найдены исправленные выборочные дисперсии  $s_1^2$  и  $s_2^2$ . Требуется решить, лежит ли различие между ними в границах возможных случайных изменений, то есть можно ли оба значения  $s_1^2$  и  $s_2^2$  рассматривать как оценку одной и той же дисперсии  $\sigma^2$  генеральной совокупности с нормальным распределением. Таким образом, следует проверить гипотезу  $H_0: \sigma_1^2 = \sigma_2^2$ . В качестве альтернативной возьмем гипотезу  $H_1: \sigma_1^2 \neq \sigma_2^2$ .

При использовании распределения Стьюдента для проверки различия между двумя выборочными средними целесообразно было полагать, что величина  $T$  пропорциональна разности этих величин:  $T \sim (m_1^* - m_2^*)$ . В результате проверочная статистика, на основании которой строится критерий проверки гипотезы, не зависела от неизвестного математического ожидания.

При сравнении двух выборочных дисперсий в качестве проверочной статистики удобно ввести случайную величину

$$F = \frac{s_1^2}{s_2^2}.$$

Найдем распределение случайной величины  $F$ , воспользовавшись тем, что при справедливости нулевой гипотезы  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ . Перепишем отношение  $\frac{s_1^2}{s_2^2}$  в виде

$$\frac{s_1^2}{s_2^2} = \frac{\frac{s_1^2}{\sigma^2}}{\frac{s_2^2}{\sigma^2}} = \frac{\chi_{n_1-1}^2}{\chi_{n_2-1}^2},$$

где  $n_1$  и  $n_2$  — размеры выборок. При этом, как видно, данное отношение не зависит от параметра  $\sigma^2$ . Поскольку обе величины  $\chi^2$  в этом выражении делятся на соответствующее число степеней свободы, то числитель и знаменатель оказываются порядка единицы. Следовательно, плотность вероятности величины  $F$  будет группироваться в окрестности единицы.

Плотности вероятности случайных величин  $\chi^2$  описываются распределением  $\chi^2$  с соответствующим числом степеней свободы. Для на-

хождения плотности вероятности величины  $F$  можно применить известное правило преобразования плотностей вероятностей. Эта плотность вероятности описывается выражением:

$$p(f) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} f^{\frac{n_1}{2}-1} \left(1 + \frac{n_1}{n_2}f\right)^{-\frac{n_1+n_2}{2}}, \quad 0 < f < \infty,$$

и называется *распределением Фишера (F-распределением)* с  $l_1 = n_1 - 1$  и  $l_2 = n_2 - 1$  степенями свободы. Вид  $F$ -распределения при некоторых значениях степеней свободы показан на рис. 11.4.

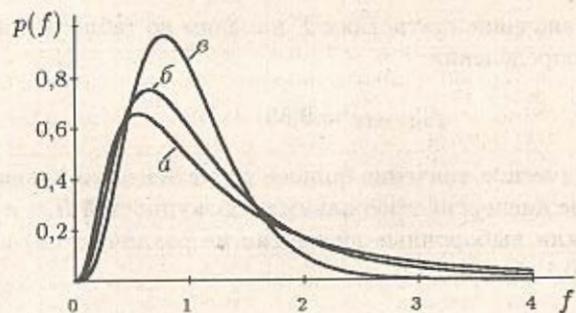


Рис. 11.4. Распределение Фишера: а —  $n_1 = 10, n_2 = 5$ ; б —  $n_1 = 10, n_2 = 10$ ; в —  $n_1 = 10, n_2 = 300$

Если гипотеза  $H_0: \sigma_1^2 = \sigma_2^2$  верна, то с вероятностью  $1 - \alpha$  должно выполняться соотношение

$$P(f_{1-\frac{\alpha}{2}; l_1; l_2} \leq \frac{s_1^2}{s_2^2} \leq f_{\frac{\alpha}{2}; l_1; l_2}) = 1 - \alpha,$$

где  $f_{\frac{\alpha}{2}}$  и  $f_{1-\frac{\alpha}{2}}$  — соответствующие процентные точки  $F$ -распределения. Критическими для проверяемой гипотезы являются значения

$$\frac{s_1^2}{s_2^2} < f_{1-\frac{\alpha}{2}; l_1; l_2} \quad \text{или} \quad \frac{s_1^2}{s_2^2} > f_{\frac{\alpha}{2}; l_1; l_2}.$$

Аналогичным образом могут быть найдены критические значения при альтернативных гипотезах  $H_1: \sigma_1^2 < \sigma_2^2$  и  $H_1: \sigma_1^2 > \sigma_2^2$ .

Для того, чтобы сократить объем таблиц процентных точек  $F$ -распределения, за значение  $s_1^2$  принимается большая из вычисленных оценок дисперсии. Процентные точки  $F$ -распределения для уровня значимости  $\alpha = 0,05$  приведены в таблице 4 Приложения.

Рассмотрим пример. Из двух нормальных генеральных совокупностей 1 и 2 взяты выборки объемом  $n_1 = 10$  и  $n_2 = 15$ . Рассчитанные по выборкам оценки дисперсий оказались равными:  $s_1^2 = 18$ ,  $s_2^2 = 26$ .

Требуется проверить гипотезу о равенстве дисперсий генеральных совокупностей при альтернативной гипотезе об их неравенстве. Уровень значимости принять равным 0,1.

Находим значение проверочной статистики  $F$ :

$$f = \frac{s_1^2}{s_2^2} = \frac{26}{18} = 1,44.$$

Критическое значение статистики  $F$  находим по таблице критических точек  $F$ -распределения:

$$f_{9;14;0,05} = 2,65.$$

Так как критическое значение больше расчетного, то нулевая гипотеза о равенстве дисперсий генеральных совокупностей  $H_0: \sigma_1^2 = \sigma_2^2$  не отвергается (или выборочные дисперсии не различаются) с уровнем значимости 0,1.

## Приложение

Таблица 1. Значения интеграла вероятности

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{x^2}{2}} dx \text{ для } 0,00 \leq z \leq 4,99; \Phi(-z) = 1 - \Phi(z)$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0,1	5396	5438	5478	5517	5557	5596	5636	5675	5714	5753
0,2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0,3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0,7	7580	7611	7642	7673	7703	7734	7764	7794	7823	7852
0,8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1,1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1,2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9014
1,3	90230	90490	90658	90824	90988	91149	91309	91466	91621	91774
1,4	91924	92073	92220	92364	92507	92647	92785	92922	93056	93189
1,5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1,7	95543	95637	95728	95818	95907	95994	96080	96164	96246	96327
1,8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1,9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2,1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2,2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2,3	98928	98956	98983	99009	99035	99061	99086	99110	99134	99167
2,4	99180	99202	99224	99245	99266	99287	99308	99324	99341	99361
2,5	99379	99396	99413	99429	99445	99461	99476	99491	99506	99520
2,6	99539	99547	99560	99573	99585	99597	99609	99620	99631	99642
2,7	99653	99663	99673	99683	99692	99702	99711	99719	99728	99736
2,8	99744	99752	99759	99767	99774	99781	99788	99794	99801	99807
2,9	99813	99819	99825	99830	99835	99841	99846	99851	99855	99860
3,0	99865	99869	99873	99877	99881	99885	99889	99893	99896	99899
3,1	99903	99906	99909	99912	99915	99918	99921	99923	99926	99928
3,2	99931	99933	99935	99938	99940	99942	99944	99946	99948	99949
3,3	99951	99953	99954	99956	99958	99959	99961	99962	99963	99965
3,4	99966	99967	99968	99969	99970	99971	99972	99973	99974	99975
3,5	99976	99977	99978	99979	99980	99981	99982	99983	99984	99985
3,6	99986	99987	99988	99989	99990	99991	99992	99993	99994	99995
3,7	99996	99997	99998	99999	99999	99999	99999	99999	99999	99999
3,8	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
3,9	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,0	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,1	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,2	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,3	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,4	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,5	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,6	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,7	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,8	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999
4,9	99999	99999	99999	99999	99999	99999	99999	99999	99999	99999

Пример.  $\Phi(3,57) = 0,98215 = 0,9998215$ .

Таблица 2. Процентные точки распределения  $\chi^2$ 

$$P[\chi_k^2 > \chi_{k,\alpha}^2] = \alpha$$

k	$\alpha$							
	0,990	0,975	0,950	0,900	0,10	0,05	0,025	0,010
1	0,00016	0,00098	0,0039	0,158	2,71	3,84	5,02	6,63
2	0,0201	0,0506	0,103	0,211	4,61	5,99	7,38	9,21
3	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34
4	0,297	0,484	0,711	1,06	7,78	9,49	11,14	13,28
5	0,554	0,831	1,15	1,61	9,24	11,07	12,83	15,09
6	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09
9	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21
11	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73
12	4,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00
17	6,41	7,56	8,67	10,08	24,77	27,59	30,19	33,41
18	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81
19	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	35,56	38,88	41,92	45,64
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96
28	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89
40	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69
60	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38
120	86,92	91,58	95,70	100,62	140,23	146,57	152,21	158,95

Таблица 3. Процентные точки  $t$ -распределения Стьюдента

$$P[t_k > t_{k,\alpha}] = \alpha$$

k	$\alpha$				
	0,10	0,050	0,025	0,010	0,005
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,553	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,455
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
60	1,296	1,671	2,000	2,390	2,660
120	1,289	1,658	1,980	2,358	2,617

Примечание. Значения  $\alpha = 0,995; 0,990; 0,975; 0,950$  и  $0,900$  получают, пользуясь соотношением  $t_{k;1-\alpha} = -t_{k;\alpha}$ .

Таблица 4. Процентные точки  $F$ -распределения для  $\alpha = 0,05$ 

n	m									
	1	2	3	4	5	6	7	8	9	10
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,19
$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

n	m								
	12	15	20	24	30	40	60	120	$\infty$
1	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,38	2,31	2,23	2,19	2,16	2,10	2,06	2,01	1,96
18	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
$\infty$	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

## Литература

1. Герасимович А.И., Матвеева Я.И. Математическая статистика. Мн.: Высшая школа, 1978.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1999.
3. Гурский Е.И. Теория вероятностей с элементами математической статистики. М.: Высшая школа, 1971.
4. Коваленко И.И., Филиппова А.А. Теория вероятностей и математическая статистика. М.: Высшая школа, 1973.
5. Румшинский Л.З. Элементы теории вероятностей. М.: Наука, 1970.
6. Худсон Д. Статистика для физиков. М.: Мир, 1970.
7. Тихонов В.И. Статистическая радиотехника. М.: Радио и связь, 1982.

## Предметный указатель

- Автокорреляционная функция  
случайного процесса,  
116
- Ансамбль, 98
- Белый шум, 128
- Вероятность  
безусловная, 16  
геометрическая, 12  
доверительная, 149, 160  
классическое определение, 7  
накопленная (кумуля-  
тивная), 27  
события, 7  
статистическое определение,  
10  
условная, 16
- Время корреляции, 120
- Выборка, 133  
бесповторная, 133  
повторная, 133
- Генеральная совокупность, 132
- Гипотеза, 18  
статистическая, 134
- Гистограмма, 135
- Двумерная плотность вероятности, 65
- Двумерная функция распределения, 64
- Дисперсия, 30, 35
- Доверительный интервал, 160
- Достаточность оценки, 147
- Закон больших чисел, 89
- Закон распределения, 23, 68
- Инвариантность, 41
- Интеграл вероятности, 60
- Информационная матрица Фишера,  
149
- Испытание, 5
- Квантиль, 34
- Ковариационная функция, 108
- Ковариационный момент, 74
- Композиция законов распределения, 77
- Корреляционная матрица, 73
- Корреляционная функция, 108
- Корреляционный момент, 71
- Коэффициент  
асимметрии, 37  
корреляции, 71  
экссесса, 37
- Критерий согласия  $\chi^2$ , 138
- Критерий статистический, 182
- Лемма Маркова, 94
- Математическое ожидание, 30
- Медиана, 34
- Метод максимального правдоподобия, 154
- Метод моментов, 153
- Метод наименьших квадратов (МНК), 177
- Мода, 34
- Момент, 30  
начальный, 36, 70  
смешанный, 71  
условный, 73  
центральный, 36, 70
- Моментные функции, 107
- Мощность критерия, 184
- Некоррелированные случайные величины, 72
- Неравенство  
Рао — Крамера, 147  
Чебышева, 89
- Несмещенность оценки, 146
- Нормировка закона распределения, 24

- Объем выборки, 133
- Объем генеральной совокупности, 132
- Опыт, 5
- Ординарность, 51
- Относительная частота, 9
- Отсутствие последствия, 51
- Оценка
  - амплитуды сигнала, 171
  - асимптотически несмещенная, 146
  - линейная, 177
  - максимально правдоподобная, 155
  - параметра, 134
  - смещенная, 146
  - точечная и интервальная, 145
  - эффективная, 148
- Оценка амплитуд нескольких сигналов, 173
- Оценка дисперсии несмещенная, 164
- Ошибки первого и второго рода, 183
- Плотность вероятности, 24
- Полная группа (система) событий, 6
- Популяция, 132
- Поток событий, 51
- Произведение событий, 13
- Процентные точки распределения  $\chi^2$ , 84
- Равновозможные события, 6
- Распределение
  - Бернулли, 46
  - Пуассона, 49
  - биномиальное, 46
  - гамма, 78
  - дискретное нормальное, 56
  - дискретное равномерное, 56

- одностороннее показательное, 54
- экспоненциальное, 54
- Реализация, 98
- Случайная величина, 22
  - n-мерная, 62
  - дискретная, 22
  - коррелированная, 72
  - непрерывная, 22
  - смешанного типа, 28
  - центрированная, 35
- Случайная последовательность, 99
- Случайный процесс, 98
  - вероятностный, 99
  - дискретный, 100
  - квазидетерминированный, 99
  - непрерывнозначный, 101
  - стационарный, 109
  - стохастический, 99
  - точечный, 101
  - эргодический, 115
- Событие, 5
  - достоверное, 5
  - невозможное, 5
  - несовместное, 6
  - противоположное, 6
  - случайное, 5
- Совместные события, 6
- Совокупность
  - выборочная, 133
  - генеральная
    - представительная (репрезентативная), 133
- Среднее арифметическое, 30
- Среднее квадратическое отклонение, 35
- Статистика, 144
  - проверочная, 182
- Статически зависимые события, 16

- Стационарность, 51
- Степени свободы, 82
- Сумма событий, 13
- Теорема
  - Бернулли, 93
  - Ляпунова, 94
  - Пуассона, 94
  - Чебышева, 91
  - сложения вероятностей, 14
  - умножения вероятностей, 16
- Уровень значимости, 10
- Формула Байеса, 20
- Формула полной вероятности, 19
- Функция
  - Лапласа, 60
- моментная
  - начальная, 107
  - центральная, 107
- правдоподобия, 154
- распределения случайной величины, 26
- случайная, 98
- характеристическая, 37
- Характеристика
  - интервальная, 145
  - точечная, 145
- Центр распределения, 31
- Частота, 9
- Элемент вероятности, 24
- Эллипс рассеяния, 85

# Оглавление

Введение	3
Глава 1. Основные понятия теории вероятностей	5
1.1. Случайные события	5
1.2. Классическое определение вероятности	7
1.3. Статистическое определение вероятности	9
1.4. Аксиоматическое построение теории вероятностей	10
1.5. Сумма и произведение событий	13
1.6. Теорема сложения вероятностей	14
1.7. Теорема умножения вероятностей	16
1.8. Формула полной вероятности	18
1.9. Формула Байеса	20
Глава 2. Случайные величины	22
2.1. Понятие случайной величины	22
2.2. Описание случайных величин	22
2.3. Функция распределения	26
2.4. Числовые характеристики случайных величин	30
2.5. Характеристические функции	37
2.6. Преобразование законов распределения и моментов	39
Глава 3. Основные законы распределения случайных величин	45
3.1. Биномиальное распределение	45
3.2. Распределение Пуассона	49
3.3. Экспоненциальное распределение	54
3.4. Равномерное распределение	55
3.5. Нормальное распределение	56
Глава 4. Системы случайных величин	62
4.1. Законы распределения двумерных случайных величин	62
4.2. Условные распределения двух случайных величин	68
4.3. Числовые характеристики двумерных законов распределения	70
4.4. Корреляция	71
4.5. Числовые характеристики $n$ -мерных случайных величин	73

4.6. Функциональные преобразования двумерных плотностей вероятностей	74
4.7. Гамма-распределение	78
4.8. Распределение $\chi^2$	82
4.9. Двумерное нормальное распределение	84
Глава 5. Закон больших чисел и предельные теоремы	89
5.1. Неравенство Чебышева	89
5.2. Теорема Чебышева	91
5.3. Теорема Бернулли	93
5.4. Лемма Маркова	94
5.5. Теорема Пуассона	94
5.6. Центральная предельная теорема Ляпунова	95
Глава 6. Случайные процессы	98
6.1. Понятие случайного процесса	98
6.2. Виды случайных процессов	99
6.3. Описание случайных процессов	102
6.4. Моментные функции	107
6.5. Нестационарные и стационарные случайные процессы	109
6.6. Эргодические стационарные процессы	113
6.7. Корреляционные функции и их свойства	116
6.8. Экспериментальное определение параметров процесса	121
6.9. Спектральная плотность	123
6.10. Белый шум	128
6.11. Нормальные случайные процессы	130
Глава 7. Основные понятия математической статистики	132
7.1. Генеральная и выборочная совокупности	132
7.2. Основные задачи математической статистики	133
Глава 8. Оценка закона распределения	135
8.1. Гистограмма распределения	135
8.2. Критерий согласия $\chi^2$	138
Глава 9. Оценка моментов и параметров распределения	144
9.1. Виды оценок и их характеристики	144
9.1.1. Состоятельность оценки	145
9.1.2. Несмещенность оценки	146
9.1.3. Достаточность оценки	147
9.1.4. Эффективность оценки	147
9.2. Точечные оценки моментов случайной величины	150
9.3. Методы оценки параметров распределения	152

9.3.1. Метод моментов . . . . .	153
9.3.2. Метод максимального правдоподобия . . . . .	154
9.4. Интервальная оценка параметров . . . . .	159
9.5. Доверительные интервалы	
для параметров распределения . . . . .	161
9.5.1. Доверительный интервал	
для математического ожидания	
при известной дисперсии . . . . .	161
9.5.2. Доверительный интервал для дисперсии . . . . .	163
9.5.3. Доверительный интервал при неизвестной дисперсии . . . . .	166
<b>Глава 10. Оценка параметров общей</b>	
<b>линейной модели измерений</b> . . . . .	<b>170</b>
10.1. Характеристика общей линейной модели . . . . .	170
10.2. Оценка амплитуды сигнала при наличии шума методом	
максимального правдоподобия . . . . .	171
10.3. Совместная оценка амплитуд нескольких сигналов . . . . .	173
10.4. Задачи, связанные с оценкой амплитуд . . . . .	176
10.5. Аппроксимация экспериментальных данных по МНК . . . . .	177
<b>Глава 11. Проверка статистических гипотез</b> . . . . .	<b>181</b>
11.1. Нулевая и альтернативная гипотезы . . . . .	181
11.2. Уровень значимости и мощность критерия . . . . .	183
11.3. Проверка гипотез о математическом ожидании . . . . .	184
11.4. Проверка гипотезы о равенстве	
двух выборочных средних . . . . .	189
11.5. Проверка гипотезы о равенстве	
дисперсий двух совокупностей . . . . .	191
<b>Приложение</b> . . . . .	<b>195</b>
<b>Литература</b> . . . . .	<b>200</b>
<b>Предметный указатель</b> . . . . .	<b>201</b>

*Учебное издание*

**Фигурин Владимир Алексеевич,  
Оболонкин Владимир Викторович**

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ  
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА**

Редактор *Н.Г. Баранова*  
Корректор *С.Н. Климович*  
Художник *С.В. Ковалевский*

Подписано в печать 30.06.2000. Формат 60×84<sup>1/8</sup>. Бумага газетная. Гарнитура Computer Modern Roman. Печать офсетная. Усл. печ. л. 12,09. Уч.-изд. л. 8. Тираж 5000 экз. Зак. 1686.

Налоговая льгота — Общегосударственный классификатор Республики Беларусь  
ОКРБ 007-98, Ч. 1; 22.11.20.100

ООО «Новое знание». ЛВ № 310 от 15.04.99.  
220099, Минск, ул. Брестская, 72-39.  
Тел.: (10-375-17) 284-03-23, факс: (10-375-17) 284-01-79.  
E-mail: [publish@nk.com.by](mailto:publish@nk.com.by)  
<http://www.nk.com.by>

Отпечатано с готовых диапозитивов заказчика в типографии издательства  
«Белорусский Дом печати». 220013, Минск, пр. Скорины, 79.