

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

А. В. Аксенчик

***ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА***

*Рекомендовано УМО Республики Беларусь по образованию
в области информатики и радиоэлектроники
в качестве учебно-методического пособия для студентов учреждений,
обеспечивающих получение высшего образования*

Минск БГУИР 2011

УДК 519.2(075.8)

ББК 22.17я73

А42

Р е ц е н з е н т ы:

кафедра высшей математики учреждения образования
«Белорусский государственный
аграрный технический университет», заведующий кафедрой,
кандидат физико-математических наук, доцент И. М. Морозова

доцент кафедры естественно-научных дисциплин
учреждения образования «Белорусский национальный
технический университет», кандидат
физико-математических наук С. В. Чернявская

Аксенчик, А. В.

А42 Теория вероятностей и математическая статистика : учеб.-метод.
пособие / А. В. Аксенчик. – Минск : БГУИР, 2011. – 184 с. : ил.
ISBN 978-985-488-632-9.

Дается систематическое изложение основ теории вероятностей и математической статистики в соответствии с типовой учебной программой по данной дисциплине.

Пособие состоит из трех частей. Часть 1 «Теория вероятностей» включает семь глав, в которых рассмотрены случайные события, случайные величины, предельные теоремы. Часть 2 «Математическая статистика» содержит пять глав, где приведены основные определения математической статистики, точечные и интервальные оценки, теория статистической проверки гипотез, линейный регрессионный анализ, а также примеры решения задач. Часть 3 «Случайные процессы» выходит за рамки учебной программы, однако этот раздел будет полезен студентам радиотехнических специальностей при изучении специальных дисциплин, использующих основы теории случайных процессов.

УДК 519.2(075.8)

ББК 22.17я73

ISBN 978-985-488-632-9

© Аксенчик А. В., 2011

© УО «Белорусский государственный
университет информатики
и радиоэлектроники», 2011

ПРЕДИСЛОВИЕ

Изучение основных положений теории вероятностей и математической статистики является важным элементом технического образования. Это обусловлено тем, что реальные физические процессы, при более глубоком их изучении, не описываются классическими методами математической физики, а требуют вероятностного подхода, т. е. неоднозначного описания физических явлений. Так возникли статистическая физика, квантовая механика. Непредсказуемый, случайный характер шумов, помех, статистическая структура источников сообщений, искажение радиосигналов при прохождении турбулентностей в атмосфере и ионосфере потребовали применения вероятностных методов в теоретической радиотехнике – так появилась статистическая радиотехника. И если в середине прошлого века теорию вероятностей и математическую статистику изучали студенты лишь некоторых специальностей, то в настоящее время студенты большинства специальностей не могут обойтись без знания этой дисциплины.

Основу данного учебно-методического пособия составляет курс лекций по теории вероятностей, математической статистике и случайным процессам, который читался студентам второго курса радиотехнических специальностей Белорусского государственного университета информатики и радиоэлектроники автором данного издания.

Содержание учебно-методического пособия предполагает знание студентами таких разделов курса высшей математики, как «Ряды», «Множества и операции над ними», «Дифференциальное и интегральное исчисления». Особое внимание в учебном пособии уделяется разъяснению основных определений и теорем теории вероятностей.

В части 1 «Теория вероятностей» (главы 1–7) данного пособия рассмотрены все основные определения, теоремы теории вероятностей, которые базируются на основных понятиях теории множеств, на аксиоматическом подходе к определению вероятности; все теоремы, свойства приводятся с доказательствами. В главе 1 даются понятия о событиях, в главе 2 рассматривается определение случайной величины, формы законов описания распределений случайных величин. В главе 3 содержатся сведения о многомерных случайных величинах. Это несколько не традиционно, так как обычно после знакомства с одномерной случайной величиной рассматривают числовые характеристики случайных величин и их свойства. Однако не имея понятия о многомерных случайных величинах, привести доказательства свойств математического ожидания, дисперсии для суммы и произведения случайных величин не представляется возможным. Ввиду этого в третьей главе, для упрощения и лучшего понимания материала студентами, на примере двумерной случайной величины рассмотрены определения функции распределения и плотности распределения, а также приведены доказательства их свойств. Дано понятие общего определения математического ожидания, когда аргументом является функция случайной величины. Это об-

легчает запись выражения для определения дисперсии, начальных и центральных моментов случайных величин.

В части 2 «Математическая статистика» приведены основные понятия математической статистики, описаны методы получения точечных оценок параметров распределений. Рассмотрены основные законы распределений случайных величин, наиболее часто применяемых в математической статистике. Даны определения доверительной вероятности, доверительных интервалов для числовых характеристик. Изложена теория статистической проверки гипотез и линейный регрессионный анализ.

Включение третьей части в данное пособие вызвано тем, что ранее студентам радиотехнических специальностей разделы по случайным процессам читались, но в новой типовой программе они отсутствуют. Чтобы восполнить этот пробел, в пособие включена часть 3 «Случайные процессы» (глава 13) посвященная элементам теории случайных процессов. Даны основные понятия теории случайных процессов, доказаны теоремы Винера-Хинчина, Котельникова, рассмотрены марковские случайные процессы.

Содержание 1-й и 2-й частей данного учебно-методического пособия полностью соответствует требованиям типовой учебной программы «Теория вероятностей и математическая статистика», утвержденной Министерством образования Республики Беларусь 03. 06. 2008, № ТД-1.030/тип., и предназначено для подготовки студентов широкого ряда специальностей: 1-39 01 01 «Радиотехника», 1-39 01 02 «Радиоэлектронные системы», 1-39 01 03 «Радиоинформатика», 1-39 01 04 «Радиоэлектронная защита информации», 1-41 01 02 «Микро- и нанoeлектронные технологии и системы», 1-41 01 03 «Квантовые информационные системы», 1-39 02 01 «Моделирование и компьютерное проектирование радиоэлектронных средств», 1-39 02 02 «Проектирование и производство радиоэлектронных средств», 1-36 04 01 «Электронно-оптические системы и технологии», 1-40 02 02 «Электронные вычислительные средства» и др.

Часть 3 «Случайные процессы» данного учебно-методического пособия может использоваться для подготовки студентов следующих специальностей: 1-39 01 01 «Радиотехника», 1-39 01 02 «Радиоэлектронные системы», 1-39 01 03 «Радиоинформатика», 1-39 01 04 «Радиоэлектронная защита информации», 1-40 02 02 «Электронные вычислительные средства».

Данное учебно-методическое пособие может быть использовано для подготовки студентов высших и средних специальных учебных заведений по экономическим и другим специальностям.

ЧАСТЬ 1. ТЕОРИЯ ВЕРОЯТНОСТЕЙ

ГЛАВА 1. СЛУЧАЙНЫЕ СОБЫТИЯ

1.1. Основные понятия теории вероятностей

Теория вероятностей – это математическая наука, которая изучает закономерности случайных явлений.

До возникновения теории вероятностей объектом исследования науки были явления или эксперименты, в которых условия постановки эксперимента однозначно определяли его исход. В этом случае результат эксперимента может быть рассчитан заранее с использованием соответствующих законов физики.

*Эксперименты, результаты которых можно предвидеть заранее, используя естественно-научные законы и выполняя соответствующие условия его проведения, называются **детерминированными**.*

Примером детерминированного эксперимента является определение силы постоянного тока I через проводник по известному сопротивлению R провод-

ника и падению напряжения U на нем: $I = \frac{U}{R}$ (используется закон Ома). Здесь искомая величина I однозначно определяется условиями проведения эксперимента (сопротивлением R и напряжением U).

Однако условия, при которых проводится эксперимент, часто бывает трудно сделать фиксированными. И это может приводить к разным результатам. Например, если проводить измерение режима электрической схемы какого-либо устройства, собранного на печатной плате, то результаты измерений могут зависеть от разных факторов: входного сопротивления вольтметра (если цепи высокоомные, то происходит шунтирование их входным сопротивлением вольтметра, изменяется режим работы схемы, и это влияет на результаты измерений); температуры, влажности в помещении, где находится устройство (при повышенной влажности возникают токи утечки между печатными проводниками, особенно заметно их влияние в высокоомных цепях, и результаты измерений изменяются); напряжения сети (известно, что в жилых многоэтажных домах утром и днем напряжение сети повышенное – 240–250 В, а вечером пони-

жается – 215–230 В, т. к. люди приходят с работы включают различные потребители электрической энергии, нагрузка (ток) в сети возрастает и на подводящих электричество к дому кабелях увеличивается падение напряжения, что приводит к уменьшению напряжения на розетках в квартирах). Здесь, как видим, проявляется зависимость напряжения сети и от времени суток. Отметим, что, несмотря на то что для питания электронных схем используют стабилизаторы напряжения, не все они могут поддерживать стабильное напряжение, если напряжение сети меняется в широких пределах. Поэтому часто в документации на электронные устройства указывают условия измерения режима работы схемы устройства: напряжение сети, температура, влажность, входное сопротивление вольтметра.

Однако условия, при которых проводится эксперимент, трудно сделать фиксированными, и даже незначительное их изменение приводит к другому результату. Поэтому в отличие от детерминированных можно выделить другой класс экспериментов – вероятностные.

***Вероятностные эксперименты** (стохастические или случайные) – это эксперименты, которые при соблюдении одних и тех же (насколько это возможно) фиксированных условий можно повторять произвольное количество раз, каждый раз, но исход вероятностного эксперимента будет неоднозначен, случаен.*

Но при многократном повторении вероятностного эксперимента, соблюдая одни и те же (насколько это возможно) фиксированные условия, можно заметить, что совокупность исходов таких экспериментов все же подчиняется определенным закономерностям. Изучением этих закономерностей, а точнее их математических моделей, и занимается теория вероятностей. Приведем примеры вероятностных экспериментов.

1. Самолету для полета на определенной высоте отводится коридор – минимальная и максимальная высоты, выходить за пределы которых он не должен. Теоретически он должен лететь горизонтально, прямолинейно, равномерно на заданной высоте (середина коридора). Однако при полете на заданной высоте на самолет воздействуют вертикальные и горизонтальные потоки воздуха, т. е. влияет турбулентность атмосферы. Это вызывает отклонение центра масс самолета от полета по горизонтальной прямой и приводит к колебаниям самолета около центра масс. Поэтому здесь высота, направление и скорость полета не определяются однозначно, а являются случайными и изменяются от полета к полету.

2. Пусть проводится эксперимент с броском симметричной монеты. В результате возможны два взаимоисключающих друг друга исхода: выпадает герб или цифра. Если бы мы точно знали начальные условия данного эксперимента (скорость вращения, скорость поступательного движения, начальное положение монеты в момент броска), то мы смогли бы предвидеть исход этого эксперимента (т. е. точно рассчитать). Однако начальные условия трудно сделать фиксированными, и незначительное изменение их приводит к другому результату. Тем не менее если мы будем проводить такой опыт многократно, то

совокупное число его исходов будет подчиняться определенным закономерностям: число выпадений герба \approx числу выпадений цифры (в том случае, если монета симметричная).

3. В городе возникла эпидемия гриппа. Человек, заболевший гриппом в этот период, может выздороветь; получить осложнения различной степени тяжести; а может и умереть. Видим, что исход заболевания гриппом непредсказуем.

4. Предположим, что проводим измерения некоторой величины (измерение напряженности электромагнитного поля телевизионного сигнала в какой-либо точке города, дальности радиолокатором и т. д.). Заранее предвидеть результат трудно. Однако при аккуратном проведении измерений и после многократного их повторения наблюдается закономерность: среднее арифметическое результатов измерений приближается к некоторой постоянной величине.

Можно найти приближенное значение данной константы и, построив математическую модель эксперимента, с определенной степенью уверенности принять эту оценку за реальное значение измеряемой величины.

Следует также отметить, что теория вероятностей занимается изучением не любых случайных событий, а только тех, которые обладают определенными свойствами. Во-первых, таких, которые при фиксированных условиях могут осуществляться *произвольное число раз*, например – бросок монеты, бросок игральной кости. Во-вторых, таких, которые обладают *статистической устойчивостью*. Например, сделаем n бросков монеты (испытания) и будем следить за появлением герба (событие A). Пусть эти испытания проводятся в одинаковых условиях и каждое испытание не оказывает влияния на остальные, т. е. они являются независимыми. Пусть событие A появится при n испытаниях m раз. Тогда *относительная частота появления события A* , обозначим ее $P^*(A)$, будет определяться как $P^*(A) = m/n$. Проводя такой опыт, мы можем заметить, что при большом n относительная частота $P^*(A)$ близка к какой-то постоянной и лишь незначительно меняется в различных сериях из n испытаний. В дальнейшем мы покажем, что число, около которого колеблется относительная частота, и есть не что иное, как вероятность события A – $P(A)$. Заметим, что относительная частота появления события A может быть определена только после проведения опыта (когда становится известным число m), а вероятность $P(A)$ может быть рассчитана и до опыта. Из приведенных примеров логически следует определение *вероятности – как меры возможности наступления какого-либо события при проведении вероятностного эксперимента*. Это не точное определение вероятности (его мы рассмотрим позже при изучении аксиом), но смысл передается верно.

Таким образом, мы видим, что понятие вероятностного эксперимента достаточно широко, однако можно выделить следующие общие черты:

- 1) множество возможных исходов;
- 2) непредвиденность результата;
- 3) наличие определенных количественных закономерностей при многократном повторении эксперимента.

Эти закономерности изучаются *методом моделирования*. Для этого строится математическая модель вероятностного эксперимента, в котором исследуемые закономерности описываются различными математическими уравнениями, функциями. Насколько точна математическая модель и хорошо ли она согласуется с практикой, можно вычислить с помощью методов *математической статистики*. Таким образом, предметом теории вероятностей является количественный и качественный анализ математических моделей вероятностных экспериментов. Эксперименты проводят с целью получения научных выводов, рекомендаций, на основе которых можно принимать решения.

Первые работы, в которых зарождались основные понятия теории вероятностей, представляли собой попытки создания теории азартных игр. Именно они стимулировали построение математических моделей игровых ситуаций в работах Дж. Кардано (1501–1575), Х. Гюйгенса (1629–1695), Б. Паскаля (1623–1662), П. Ферма (1601–1665) и др. в XVI – XVII вв. Следующий этап развития теории вероятностей связан с именем Якова Бернулли (1654–1705). Доказанная им теорема, получившая впоследствии название «Закон больших чисел», была первым теоретическим обоснованием накопленных ранее фактов. Дальнейшими успехами теория вероятностей обязана П. Д. Бернулли (1700–1782), П. Лапласу (1749–1827), К. Гауссу (1777–1855), С. Пуассону (1781–1840) и др. Новый наиболее плодотворный период связан с именами П. Л. Чебышева (1821–1894) и его учеников А. А. Маркова (1856–1922) и А. М. Ляпунова (1857–1918). Современная теория вероятностей получила развитие в работах С. Н. Бернштейна, А. Н. Колмогорова (1903–1987), Б. В. Гнеденко и др. Однако строгой математической наукой теорию вероятностей сделала работа русского ученого А. Н. Колмогорова «Основные понятия теории вероятностей», опубликованная в 1933 г. В этой работе А. Н. Колмогоров сформулировал аксиомы теории вероятностей и для доказательства основных положений применил теорию множеств.

1.2. Пространство элементарных событий

Рассмотрим простые примеры.

1. Подбрасывание игральной кости один раз. В результате этого опыта могут наступить различные события: выпали 2, 6, 5... очков, число выпавших очков четно и т. д. Мы будем различать *элементарные* (неразложимые) *события*, *составные события* или просто *события*. Обозначим ω_k – событие состоит в том, что выпадает k очков. Элементарными событиями в данном опыте являются: $\omega_1, \omega_2, \omega_3, \dots, \omega_6$. Составные события или просто события могут быть описаны как *подмножества* множества элементарных событий $\Omega = \{\omega_1, \omega_2, \omega_3, \dots, \omega_6\}$. Так событие $A = \{\text{выпало четное число очков}\}$ через элементарные события выражается следующим образом: $A = \{\omega_2, \omega_4, \omega_6\}$.

2. Трехкратное подбрасывание монеты. При каждом подбрасывании будем записывать результат опыта, обозначая выпадение герба – «1», а цифры – «0» (010 – при втором подбрасывании выпал герб). Множество элементарных событий Ω состоит из 8 элементарных событий: $\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\}$ или $\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$. $A = \{\text{при первом подбрасывании выпал герб}\}$ является составным: $A = \{100, 101, 110, 111\}$. Любое подмножество множества Ω можно интерпретировать как некоторое событие реального опыта. Например, событие $B = \{110, 101, 011\}$ состоит в том, что выпало ровно 2 герба.

3. Семья с тремя детьми. Для того чтобы изучить распределение мальчиков и девочек во всех семьях, имеющих трех детей, составлен список этих семей. Какое пространство событий отвечает эксперименту, заключающемуся в выборе одной семьи? Множество элементарных событий Ω состоит из 8 элементарных событий $\Omega = \{МММ, ММД, МДН, МДД, ДММ, ДМД, ДДМ, ДДД\}$. Тройка ДММ представляет следующий исход: старший ребенок – девочка, средний и младший – мальчики. Задача легко решается, если построить «дерево» эксперимента (рис. 1.1).

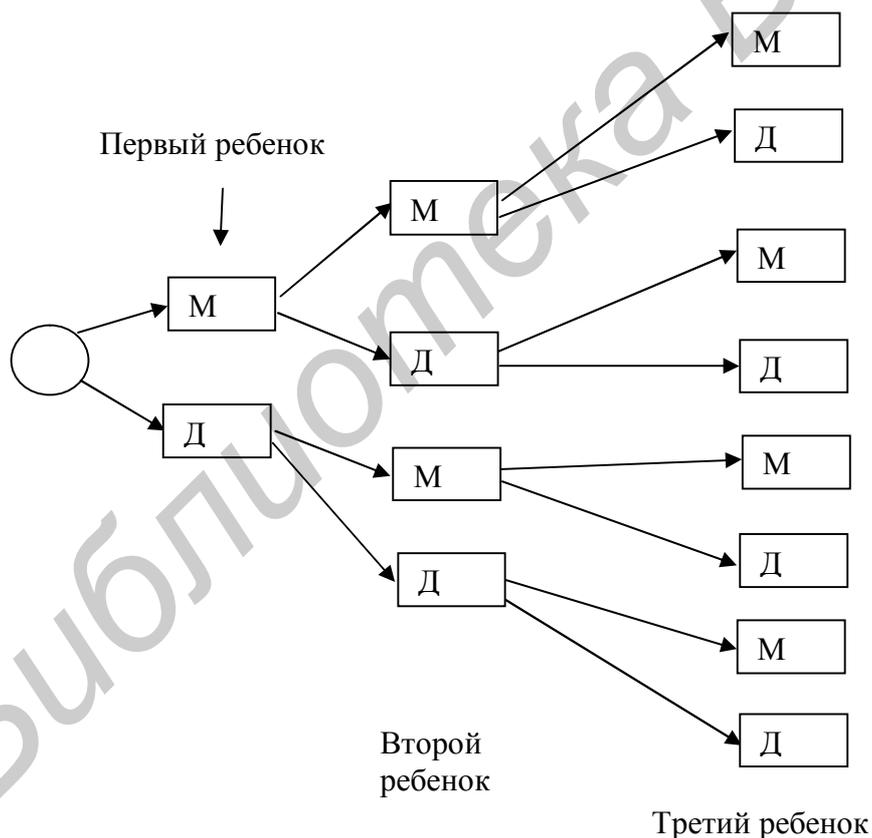


Рис. 1.1

Таким образом, можно дать следующие определения:

1. **Элементарным событием** ω называется любой мысленно возможный результат эксперимента. Совокупность всех элементарных событий ω образует некоторое множество Ω .

2. **Пространством элементарных событий** будем называть множество $\Omega = \{\omega\}$ всех мыслимых, взаимоисключающих друг друга исходов опыта или эксперимента.

По числу элементов $\omega \in \Omega$ пространство элементарных событий может быть конечным, счетным. Пространство Ω с конечным множеством элементов ω будем обозначать $\Omega = \{\omega_i \mid i = \overline{1, n}\}$, а число его элементов $|\Omega| = n$. Пространство со счетным множеством элементов ω обозначим $\Omega = \{\omega_i \mid i = \overline{1, \infty}\}$. Если Ω состоит из конечного или счетного множества ω , то Ω называют *дискретным пространством* элементарных событий, а если из несчетного множества, то Ω называют *непрерывным пространством* элементарных событий.

1.3. Операции над событиями

1. **Случайным событием** или просто **событием** назовем любое подмножество множества Ω . События будем обозначать заглавными буквами латинского алфавита: A, B, C, \dots – и записывать $A \subseteq \Omega$ (A есть подмножество множества Ω).

Приведем более простое определение события. Из приведенных выше примеров следует, что случайным событием называется всякий факт, который в результате опыта может произойти или не произойти. Если A влечет B , а B влечет A ($A \subset B$, а $B \subset A$), то A и B происходят лишь одновременно ($A = B$), множества A и B совпадают, и говорят, что A равносильно или тождественно B .

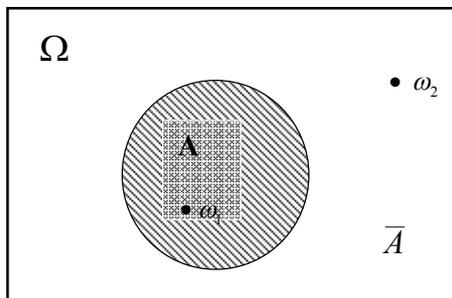
2. **Достоверным событием** называется событие, совпадающее с пространством элементарных событий Ω . Если записать $A = \Omega$, то A – достоверное событие. Достоверное событие происходит при каждом повторении эксперимента (опыта).

3. **Невозможным событием** называется событие, совпадающее с пустым множеством. Если записать $A = \emptyset$, тогда A является невозможным событием. Невозможное событие не появится ни в одном из исходов вероятностного эксперимента.

4. **Противоположным событием** \bar{A} называется дополнение множества A до Ω , т. е. событие \bar{A} состоит из таких элементарных событий, которые не входят в событие A : $\bar{A} = \{\omega \mid \omega \in \bar{A}\}$. Противоположное событие \bar{A} происходит тогда, когда событие A не происходит. Например, при броске игральной кости

событие \bar{A} – нечетное число очков. $\bar{A} = \{\omega_1, \omega_3, \omega_5\}$ является противоположным событию $A = \{\omega_2, \omega_4, \omega_6\}$.

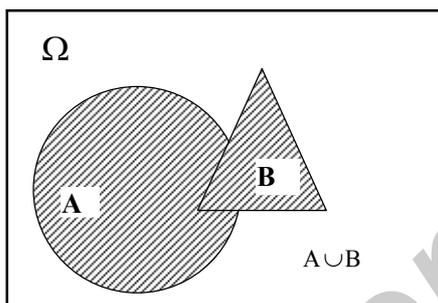
Очевидно, что $\bar{\bar{\Omega}} = \Omega$, $\bar{\bar{A}} = A$. Для иллюстрации операций над событиями удобно использовать *диаграмму Венна* (рис. 1.2).



ω_1 принадлежит событию A , $\omega_1 \in A$;
 ω_2 не принадлежит событию A , $\omega_2 \notin A$.
 Справедливы соотношения:
 $\bar{\bar{A}} = A$, $\bar{\bar{\Omega}} = \Omega$, $\bar{\bar{\emptyset}} = \emptyset$.

Рис. 1.2

5. **Объединением** или **суммой** двух событий A и B называют событие, содержащее те элементарные события $\omega \in \Omega$, которые принадлежат или событию A , или событию B , или A и B . Объединение или сумму двух событий будем обозначать $A \cup B$ (рис. 1.3):



$$A \cup B = \{\omega \mid \omega \in A, \text{ или } B, \text{ или } A \text{ и } B\}.$$

Для любого события A справедливы следующие соотношения:

$$A \cup A = A, \quad A \cup \emptyset = A, \quad A \cup \Omega = \Omega.$$

Рис. 1.3

6. **Пересечением** или **произведением** двух событий A и B называется событие, состоящее из тех элементарных событий $\omega \in \Omega$, которые принадлежат и событию A , и событию B одновременно или совместно. Пересечение или произведение двух событий будем обозначать $AB = A \cap B$. Согласно определению

$$A \cap B = \{\omega \mid \omega \in A, \omega \in B\}.$$

Справедливы следующие соотношения:

$$A \cap A = A, \quad A \cap \emptyset = \emptyset, \quad A \cap \Omega = A.$$

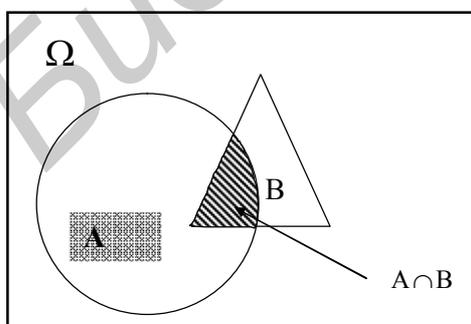


Рис. 1.4

Отметим, что выделенные слова **или** в определении для суммы, а также **и** в определении для произведения позволяют легко сориентироваться, какую операцию следует применить между событиями.

7. События A и B называются **несовместными**, если их пересечение является невозможным событием, т. е. если $A \cap B = \emptyset$. Для несовместных событий одновременное наступление в одном опыте невозможно.

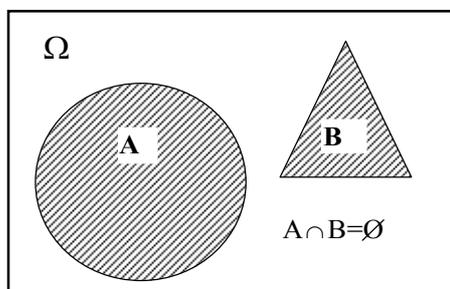


Рис. 1.5

Отметим, что элементарные события являются несовместными, т. к. при $i \neq j$, $\omega_i \cap \omega_j = \emptyset$.

8. **Разностью** двух событий A и B называется событие, состоящее из тех элементарных событий $\omega \in \Omega$, которые входят в событие A , но не входят в событие B (рис. 1.6).

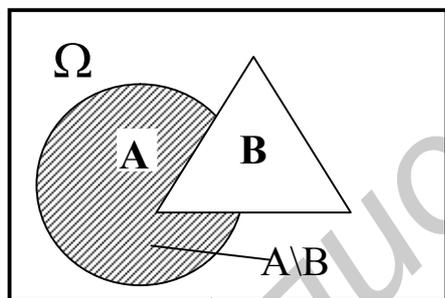


Рис. 1.6

Разность будем обозначать:
 $A \setminus B: A \setminus B = \{\omega | \omega \in A, \text{ но } \omega \notin B\}$.

Разность событий можно выразить через ранее введенные операции:
 $A \setminus B = A \cap \bar{B}$.

Основные законы алгебры событий.

Для алгебры событий справедливы следующие законы:

а) **дополнения:**

$$A \cup \bar{A} = \Omega; \quad A \cap \bar{A} = \emptyset; \quad \bar{\Omega} = \emptyset; \quad \bar{\emptyset} = \Omega;$$

принцип двойственности:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}; \quad \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

Докажем, что $\overline{A \cup B} = \bar{A} \cap \bar{B}$.

Событие $\overline{A \cup B}$ состоит в том, что не происходит событие $A \cup B$, состоящее в наступлении хотя бы одного из событий A или B (согласно пункту 5).

Но это буквально и означает, что не происходит ни событие A , ни B , т. е., согласно пунктам 4 и 7, происходит $\overline{A \cap B}$, значит, $\overline{A \cup B} \subset \overline{A \cap B}$. Наоборот, если происходит $\overline{A \cap B}$, то не происходит ни A ни B , следовательно не происходит событие $A \cup B$, т. е. происходит $\overline{A \cup B}$, следовательно, $\overline{A \cup B} \subset \overline{A \cap B}$, а значит, $\overline{A \cup B} = \overline{A \cap B}$. Кратко можно описать так:

$$(\omega \in \overline{A \cup B} \leftrightarrow \omega \notin A \cup B \leftrightarrow \omega \notin A, \omega \notin B \leftrightarrow \omega \in \overline{A}, \omega \in \overline{B} \leftrightarrow \omega \in \overline{A \cap B}),$$

здесь \leftrightarrow означает эквивалентно.

б) **коммутативности:**

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

в) **ассоциативности:**

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

г) **дистрибутивности:**

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Если последнее соотношение очевидно, то первое можно доказать так (начнем с правой части):

$$\begin{aligned} (A \cup B) \cap (A \cup C) &= A \cap A + A \cap C + B \cap A + B \cap C = | \text{учтем: } A \cap A = A | = \\ &= A + A \cap C + B \cap A + B \cap C = | \text{представим } A = A \cap \Omega | = \\ &= (A \cap \Omega + A \cap C + A \cap B) + B \cap C = | \text{применим закон ассоциативности} | = \\ &= A \cap (\Omega + C + B) + B \cap C = | \text{учтем: } \Omega + C + B = \Omega | = A + (B \cap C). \end{aligned}$$

Для удобства записей в теории вероятностей можно опускать символ пересечения \cap , т. е. пишется ABC вместо $A \cap B \cap C$ или $A+B$ вместо $A \cup B$.

1.4. Аксиомы теории вероятностей (Аксиомы Колмогорова А. Н.)

Рассмотренные ранее свойства операций над событиями носят алгебраический характер. В теории вероятностей для некоторого класса подмножеств F вводится понятие *алгебры событий*. Класс подмножеств F всех возможных событий вероятностного эксперимента должен удовлетворять следующим условиям.

1. Для каждой пары событий A и B из таких событий, как $A \in F$ и $B \in F$, следует, что $A \cup B \in F$. Иными словами, класс F вместе с каждой парой событий *содержит их объединение*.

2. Вместе с каждым событием A класс F *содержит противоположное событие* \overline{A} . Если класс F не пуст, то $\Omega \in F$, т. к. $\Omega = A + \overline{A}$. Следовательно, $\emptyset \in F$, поскольку $\emptyset = \overline{\Omega}$. Наконец, согласно принципу двойственности, $\overline{A \cup B} = \overline{A} \cap \overline{B}$, класс F содержит их пересечение $A \cap B$ и разность $A \setminus B = A \cap \overline{B}$.

Класс подмножеств F , удовлетворяющий условиям 1 и 2, называется алгеброй событий.

Конструкция алгебры событий позволяет охарактеризовать множество всех возможных результатов любого эксперимента со случайным исходом, если множество Ω его элементарных исходов конечно. Например, в эксперименте с игральной костью Ω состоит из шести элементарных событий, а F состоит из всех подмножеств Ω . F содержит пустое множество \emptyset , $6 = C_6^1$ – одноточечных подмножеств, $15 = C_6^2$ – двухточечных подмножеств и т. д. Значит F содержит $2^6 = 1 + C_6^1 + C_6^2 + \dots + C_6^6 = 64$ события.

Если Ω состоит из n элементарных событий, то F состоит из $2^n = C_n^0 + C_n^1 + \dots + C_n^n$ событий. Если $|\Omega| = n$, то $|F| = 2^n$.

Ниже перечислены аксиомы теории вероятностей (или аксиомы А. Н. Колмогорова).

1. **F является σ -алгеброй событий.** Алгебра событий F называется σ -алгеброй, если для каждой счетной последовательности событий $A_j \in F$, $j = 1, 2, \dots$, их объединение $B = A_1 \cup A_2 \cup \dots = \bigcup_1^\infty A_j$ также принадлежит F , т. е.

является событием. Согласно принципу двойственности, отсюда следует, что и $C = \bigcap_1^\infty A_j \in F$. Действительно, $C = \overline{\overline{A_1 \cup A_2 \cup \dots}} = \overline{A_1 \cap A_2 \cap \dots}$. Следует подчеркнуть, что речь идет лишь о счетных объединениях и пересечениях. Если A_α , $\alpha \in S$ – произвольная система событий, то, например, их объединение $\bigcup_{\alpha \in S} A_\alpha$ может и не быть событием.

Первая аксиома вводит понятие σ -алгебры событий. Можно отметить, что σ -алгебра событий – это расширение понятия алгебры событий на бесконечную последовательность событий.

2. На σ -алгебре F определяется **функция $P(\cdot)$** , принимающая числовые значения $P(A) \geq 0$ для $A \in F$, называемая **вероятностью** и обладающая свойствами, которые описывают нижеследующие аксиомы.

Отметим, что вторая аксиома вводит понятие вероятности.

3. Для всяких двух событий A и B , при которых $A \cap B = \emptyset$, т. е. несовместных, $P(A + B) = P(A) + P(B)$ – **аксиома сложения вероятностей для несовместных событий** (аксиома конечной аддитивности). Для произвольного конечного числа несовместных событий

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

4. Пусть A_j , $j = 1, 2, \dots$ – попарно несовместны: $A_i \cap A_j = \emptyset$, если $i \neq j$; $i, j = 1, 2, \dots$ и $A = A_1 + A_2 + \dots$, тогда $P(A) = P(A_1) + P(A_2) + \dots = \sum_{i=1}^\infty P(A_i)$ (за-

метим, что, согласно аксиоме 1, $A \in F$). Эта аксиома определяет *счетную аддитивность вероятности*, иногда ее называют *аксиомой непрерывности вероятности*.

Здесь следует сделать некоторые пояснения. Напомним, что аддитивные величины – это величины, связанные с геометрическими или физическими объектами таким образом, что величина, соответствующая целому объекту, равна сумме соответствующих его частей, каким бы образом мы не разбивали объект на части. Поэтому площади, объемы, массы тел удовлетворяют условиям аддитивности (аксиомам) и их можно принимать за меры тел. Диаметры, например, не удовлетворяют условиям аддитивности и за меру приняты быть не могут.

5. $P(\Omega) = 1$ – *аксиома нормировки*.

Числовую функцию $P(A)$, как мы отмечали, называют *вероятностью* или *вероятностной мерой* появления события A , и из аксиом следует, что **вероятность любого события A : $0 \leq P(A) \leq 1$** .

*Пространство элементарных событий Ω , выделенная σ -алгебра событий F и определенная на измеримом пространстве (Ω, F) вероятностная мера P появления события $A \in F$ образуют **вероятностное пространство**, которое принято обозначать (Ω, F, P) .*

Вероятностное пространство (Ω, F, P) совместно с аксиомами представляет собой математическую модель вероятностного эксперимента.

Рассмотрим следующий эксперимент: в коробке находятся три шара разных цветов – белый, черный, красный. Не глядя, вытаскиваем один шар.

Пространство элементарных событий этого эксперимента $\Omega = \{\omega_1, \omega_2, \omega_3\}$ и состоит из трех исходов, где ω_1 – {достали шар белого цвета}, ω_2 – {достали шар черного цвета}, ω_3 – {достали шар красного цвета}. Выделим в качестве σ -алгебры событий F этого эксперимента систему всех подмножеств Ω (ранее мы отметили, что количество подмножеств в F должно быть равно $|F| = 2^3 = 8$): $F = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \Omega\}$. Здесь событие $\{\omega_1, \omega_2\}$ состоит в том, что вытащили или белый, или черный шар; $\{\omega_1, \omega_3\}$ – состоит в том, что вытащили или белый, или красный шар; $\{\omega_2, \omega_3\}$ состоит в том, что вытащили или черный, или красный шар;

Определим для этих событий вероятностную меру P :

$$P(\emptyset) = 0; \quad P(\omega_1) = 1/3; \quad P(\omega_2) = 1/3; \quad P(\omega_3) = 1/3; \quad P(\omega_1, \omega_2) = 2/3;$$

$$P(\omega_1, \omega_3) = 2/3; \quad P(\omega_2, \omega_3) = 2/3; \quad P(\Omega) = 1.$$

Определенная таким образом функция $P(A)$ удовлетворяет аксиомам Колмогорова

$$(\Omega = \omega_1 \cup \omega_2 \cup \omega_3; \quad P(\Omega) = P(\omega_1) + P(\omega_2) + P(\omega_3) = 1; \quad \omega_1 \cap \omega_2 \cap \omega_3 = \emptyset),$$

а следовательно, она является вероятностной мерой наступления событий $A \in F$. Для тех же событий $A \in F$ можно записать другую функцию, которую обозначим P_1 :

$$P_1(\emptyset) = 0; P_1(\omega_1) = 1/4; P_1(\omega_2) = 1/2; P_1(\omega_3) = 1/4; P_1(\omega_1, \omega_2) = 3/4; \\ P_1(\omega_1, \omega_3) = 1/2; P_1(\omega_2, \omega_3) = 3/4; P_1(\Omega) = 1.$$

Эта функция также удовлетворяет аксиомам Колмогорова, а следовательно, она также является вероятностной мерой наступления событий $A \in F$. Отсюда видно, что система аксиом не полна, не устанавливает взаимно однозначного соответствия между событиями $A \in F$ и вероятностями $P(A)$. Только опыт и наблюдения могут показать правильность наших модельных предположений. Согласуется ли выбранная математическая модель с данными эксперимента, можно выяснить с помощью методов математической статистики.

Рассмотрим методы определения вероятностей для событий, удовлетворяющих аксиомам Колмогорова.

1.5. Классический метод определения вероятностей

Пусть Ω – конечное или бесконечное пространство элементарных событий некоторого случайного эксперимента. Предположим, что структура эксперимента такова, что на Ω можно указать n событий A_1, A_2, \dots, A_n , обладающих следующими свойствами:

1) события A_1, A_2, \dots, A_n **попарно несовместны** в том смысле, что никакие два из них не могут произойти одновременно. Иначе говоря, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j = 1, \dots, n$.

2) A_1, A_2, \dots, A_n образуют **полную группу событий** в том смысле, что при любом исходе эксперимента происходит хотя бы одно из них. Это означает, что $A_1 + A_2 + \dots + A_n = \Omega$, $P(A_1 + \dots + A_n) = P(\Omega) = 1$.

3) события A_1, A_2, \dots, A_n **равновероятны** (равновозможны), т. е. ни одно из них нельзя считать более предпочтительным, чем любое из остальных.

В классической схеме вероятности событий A_1, A_2, \dots, A_n , удовлетворяющих условиям 1–3, называются **полной группой попарно несовместных, равновероятных событий**. Вероятность в классической схеме определяется лишь для тех исходов эксперимента, которые могут быть представлены в виде объединения некоторых из событий A_i , $i = 1, \dots, n$. Пусть произвольное событие A можно представить в виде суммы или объединения событий:

$$A = A_{i_1} + A_{i_2} + \dots + A_{i_m}. \quad (1.1)$$

Если все слагаемые в формуле (1.1) различны, то вероятность события A определяется следующим равенством:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{m}{n}, \quad (1.2)$$

где m – число слагаемых в формуле (1.1).

Чтобы определение (1.2) было корректным, необходимо доказать единственность разложения (1.1). Для любого события A , согласно условию 2) и свойству дистрибутивности,

$$A = A \cap \Omega = A \cap (A_1 + \dots + A_n) = A \cap A_1 + \dots + A \cap A_n,$$

поэтому разложение (1.1) $A \cap A_j$ либо пусто, если $j \neq i_s$, $s = 1, \dots, m$, либо $A \cap A_j = A_{i_s}$, если $j = i_s$.

Для классического определения вероятности характерна своя терминология. Эксперимент называется испытанием, опытом; полную группу попарно несовместных, равновероятных событий называют полной группой возможных исходов испытания; а те исходы, из которых складывается событие A – называют исходами благоприятствующими появлению A .

Тогда вероятностью события A называется отношение числа m исходов, благоприятствующих появлению события A , к числу n всех возможных, равновероятных исходов опыта или эксперимента (1.2).

Отыскание вероятностей, основанное на классическом определении, часто сводится к комбинаторным вычислениям. Напомним простейшие комбинаторные формулы.

Размещения:

$$A_n^m = n(n-1)\dots(n-(m-1)).$$

Формула для размещений позволяет подсчитать количество комбинаций из n элементов по m , где комбинации будут отличаться как самими элементами, так и расположением элементов относительно друг друга.

Сочетания:

$$C_n^m = \frac{n(n-1)\dots(n-(m-1))}{1 \cdot 2 \cdot 3 \cdot \dots \cdot m} = \frac{n!}{(n-m)!m!}.$$

Формула для сочетаний позволяет подсчитать количество комбинаций из n элементов по m , где комбинации будут отличаться только самими элементами, т. е. каждая комбинация хотя бы одним элементом должна отличаться от другой.

Перестановки:

$$P_n = 1 \cdot 2 \cdot \dots \cdot n = n!, \quad P_0 = 0! = 1.$$

Формула для перестановок позволяет подсчитать количество комбинаций из n элементов, где комбинации будут отличаться только расположением элементов относительно друг друга.

Пример 1.

Проводится три броска симметричной монеты. Какова вероятность того, что герб появится два раза?

Задачу начинаем решать с определения события A . Здесь событие A состоит в том, что при трех бросках герб появится два раза. Определяем пространство элементарных событий Ω этого эксперимента. В разделе 1.2 в пункте

2 мы рассмотрели этот вопрос и получили (обозначили выпадение герба – 1, цифры – 0)

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

Здесь общее число исходов $n = |\Omega| = 8$. Выпишем благоприятствующие исходы, из которых состоит событие A : $A = \{011, 101, 110, 111\}$, тогда $m = |A| = 4$.

$$P(A) = \frac{m}{n} = \frac{4}{8} = 0,5.$$

Пример 2.

Какова вероятность того, что наудачу взятый телефонный номер из семи цифр имеет: 1) все цифры различные; 2) только нечетные цифры.

1) определим вначале событие A . Событие A состоит в том, что в семизначном номере все цифры различны. Так как номер семизначный, а цифр всего 10, то общее число исходов n (всего может быть 10^7 номеров) равно 10^7 . Для подсчета благоприятствующих исходов подходит формула для размещений –

$$m = A_{10}^7. \text{ Тогда } P(A) = \frac{m}{n} = \frac{A_{10}^7}{10^7} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{10^7} \approx 0,06.$$

2) определим событие B . Событие B состоит в том, что в семизначном номере все цифры нечетные. Имеем 5 нечетных цифр. Из них можно получить лишь 5^7 различных семизначных номеров, т. е. $m = 5^7$. Общее число исходов $n = 10^7$.

$$P(B) = \frac{m}{n} = \frac{5^7}{10^7} = \left(\frac{1}{2}\right)^7 \approx 0,008.$$

1.6. Геометрическое определение вероятностей

Пусть эксперимент описывается пространством элементарных событий Ω со счетным множеством элементарных событий ω : $|\Omega| = \infty$. Элементарные события ω будем рассматривать как координаты точки в n -мерном пространстве R^n ($n=1, 2, 3, \dots$), а события $A \in F$ как некоторые области этого пространства. Если задавать вероятности по аналогии с классическим методом, то получим $P(\omega) = 1/|\Omega| = 0$. Ранее мы указывали, что классическое определение вероятности справедливо не для любых событий, а для удовлетворяющих определенным требованиям. Поэтому в нашем случае можно приписать вероятности не каждому ω , а некоторому множеству элементарных событий $A \in F$, т. е. некоторой области пространства. При изучении аксиом мы указали, что геометрические меры множеств $A \in F$ удовлетворяют условию аддитивности, могут быть приняты за вероятностную меру этих множеств. Для этого геометрические меры нормируют, т. е. принимают меру Ω за 1, а вероятность произвольного события $A \in F$ назначают пропорционально области A :

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{\text{(мера области } A\text{)}}{\text{(мера области } \Omega\text{)}}. \quad (1.3)$$

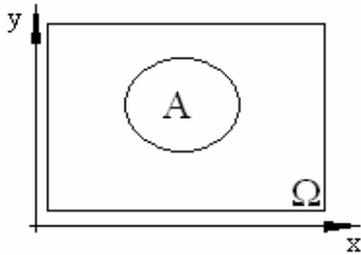


Рис. 1.7

В одномерном пространстве R^1 – это отношение длин, в двумерном R^2 – отношение площадей (рис. 1.7), в трехмерном R^3 – отношение объемов, соответствующих множествам A и Ω . Как видно из рис. 1.7 и формулы (1.3), вероятность события A не зависит ни от формы области A , ни от места ее расположения.

Такой метод задания вероятностей (1.3) называется *геометрическим* – это отношение меры области A к мере области Ω . Его особенность в том, что области A и Ω должны иметь конечную меру в пространстве R^n .

Пример 1.

Наудачу взяты два положительных числа x и y , причем $x \leq 5$, $y \leq 2$. Найти вероятность того, что $y + ax - b \leq 0$ и $y - cx \leq 0$, если $a = 2$, $b = 10$, $c = 2$.

Подставляя значения коэффициентов в неравенства, получаем

$$\begin{cases} y + 2x \leq 10, \\ y \leq 2x. \end{cases} \quad (a)$$

Строим на рис. 1.8 оси координат и область, которая определяет пространство элементарных событий Ω , она задается неравенствами $x \leq 5$, $y \leq 2$ и отображается на рисунке 1.8 прямоугольником.

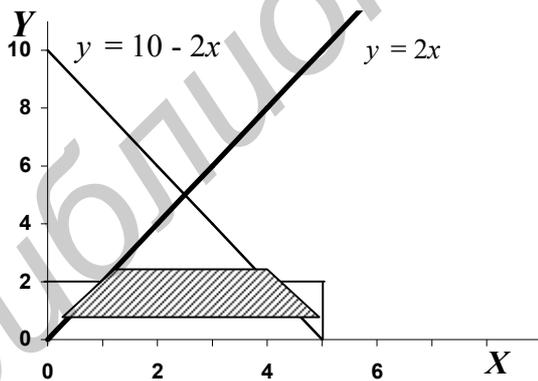


Рис. 1.8

Площадь прямоугольника $S_{\Omega} = 2 \cdot 5 = 10$ [условных единиц]. Область благоприятствующих исходов определяется неравенствами (а), поэтому строим на рисунке прямые, которые задаются из неравенств (а). Заштрихованная на рисунке 1.8 область описывает благоприятствующие исходы (с учетом всех воз-

можных значений) и является трапецией, площадь которой $S_A = \frac{5+3}{2} \cdot 2 = 8$ [условных единиц]. Тогда вероятность события A

$$P(A) = \frac{S_A}{S_\Omega} = \frac{8}{10} = 0,8.$$

Пример 2.

Найти вероятность того, что на экране радиолокатора отметка от цели появится в кружке радиусом $r = 1$ см, на азимуте φ ноль градусов, на расстоянии $L = 3$ см от центра экрана, если радиус экрана равен 30 см.

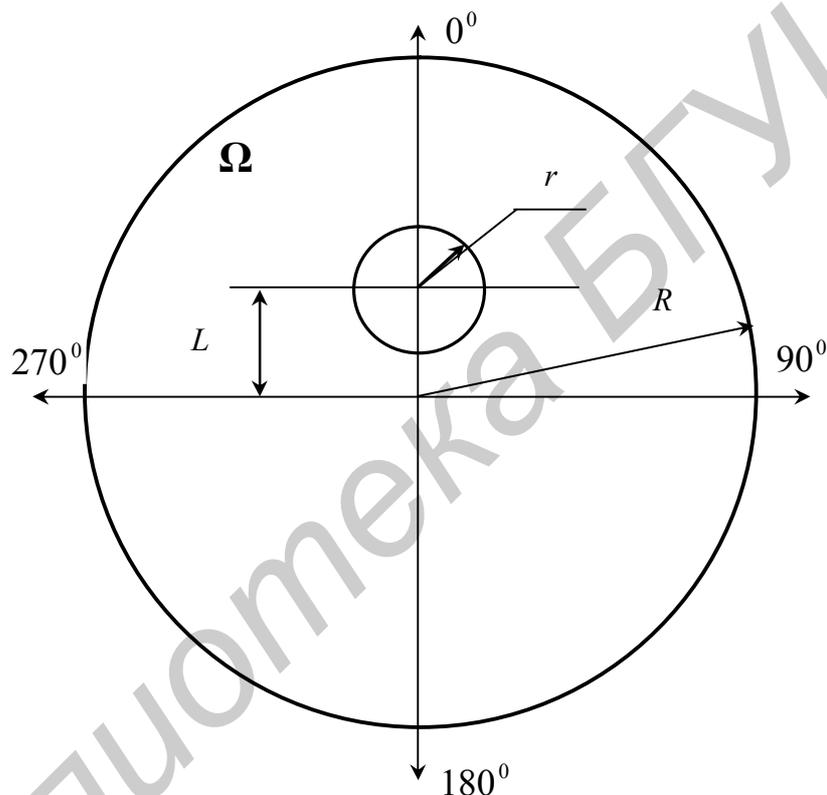


Рис. 1.9

Экран радиолокатора, рис. 1.9, представляет собой электронно-лучевую трубку с радиальной разверткой, в которой от центра до края экрана движется электронный луч и после достижения края движение луча опять начинается от центра к краю, но с некоторым смещением по азимуту. Это перемещение луча от центра экрана соответствует началу излучения радиоимпульса антенного радиолокатора, который укреплен на боковой стенке кабины с передающим устройством, а кабина, в свою очередь, вращается вокруг вертикальной оси, что соответствует смещению луча на экране по азимуту. И когда радиоимпульс отражается от цели, на экране радиолокатора вспыхивает яркая точка. По положению этой точки на экране легко определить расстояние до цели и ее азимут.

Зная геометрическое определение вероятности, можно сразу сказать, что вероятность появления отметки от цели в кружке радиусом r зависит только от отношения площадей и не зависит ни от формы области благоприятствующих исходов, ни от места ее расположения. Поэтому в этой задаче есть избыточная информация – расстояние L и азимут $\varphi=0^0$.

Определяем область благоприятствующих исходов, которой является кружок радиусом r и площадью – $S_A = \pi r^2$. Пространство элементарных событий Ω – это область экрана, его площадь $S = \pi R^2$. Тогда

$$P(A) = \frac{S_A}{S_\Omega} = \frac{\pi r^2}{\pi R^2} = \frac{r^2}{R^2} = \frac{1^2}{30^2} \approx 0,0011.$$

1.7. Свойства вероятности (основные теоремы)

Пусть Ω – пространство элементарных событий некоторого эксперимента и состоит из конечного или счетного множества элементарных событий $\omega \in \Omega$. Пусть событие A принадлежит σ -алгебре F . Тогда вероятность появления события $A \in F$ находится на основании следующей теоремы.

Теорема 1. Если в дискретном пространстве Ω , содержащем конечное или счетное множество возможных исходов $\omega_1, \omega_2, \dots, \omega_i$, $\Omega = \{\omega_1, \omega_2, \dots, \omega_i\}$, заданы вероятности элементарных событий $P(\omega_1) = P_1, P(\omega_2) = P_2, \dots, P(\omega_i) = P_i$, так $P_i > 0$, $\sum_{i=1}^{\infty} P_i = 1$, то вероятность произвольного события $A = \{\omega_1, \omega_2, \dots, \omega_m\}$, составленного из некоторых из событий $\omega_i \in \Omega$, равна сумме элементарных событий, входящих в событие A :

$$P(A) = \sum_{\omega_i \in A} P(\omega_i).$$

Доказательство следует непосредственно из 3-й аксиомы (т. к. ω_i – несовместные события).

Теорема 2. Если событие A содержится в событии B , то справедливы следующие соотношения:

$$P(A) \leq P(B); \quad P(B \setminus A) = P(B) - P(A).$$

Доказательство:

Представим событие $B = A \cup (B \setminus A)$. События A и $B \setminus A$ несовместны (рис. 1.10), поэтому на основании 3-й аксиомы имеем

$$P(B) = P(A) + P(B \setminus A) \Rightarrow P(B \setminus A) = P(B) - P(A).$$

Из второй аксиомы следует, что $P(B \setminus A) \geq 0$, следовательно, $P(B) - P(A) \geq 0$ или $P(A) \leq P(B)$.

Следствие 1. Вероятность появления невозможного события равна нулю.

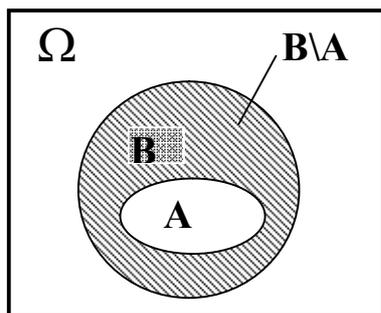


Рис. 1.10

Теорема 3. Теорема сложения вероятностей произвольных событий. Вероятность появления суммы двух произвольных событий равна сумме вероятностей этих событий без вероятности их совместного появления:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (1.4)$$

Эта теорема является обобщением 3-й аксиомы на случай, когда события $A \in F$ и $B \in F$ произвольные, т. е. могут быть совместными.

Доказательство:

Исходя из рис. 1.11 для произвольных событий A и B можно записать следующие соотношения:

$$A \cup B = A \cup (B \setminus A), \quad B = (A \cap B) \cup (B \setminus A).$$

Здесь события, входящие в правые части, – несовместны: A и $B \setminus A$, а также события $A \cap B$ и $B \setminus A$. Используя третью аксиому, для несовместных событий можно записать следующие равенства:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \setminus A), \\ P(B) &= P(A \cap B) + P(B \setminus A). \end{aligned}$$

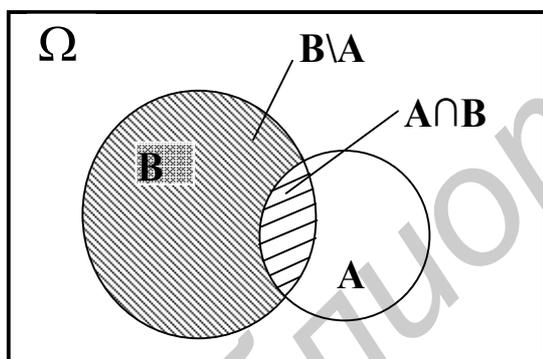


Рис. 1.11

Вычитая из первого уравнения второе и перенося $P(B)$ вправо, получаем формулу (1.4).

Для несовместных событий A и B : $A \cap B = \emptyset$ и $P(A \cap B) = 0$, тогда формула (1.4) совпадает с аксиомой 3: $P(A \cup B) = P(A) + P(B)$.

Теорема 4. Вероятность появления суммы трех произвольных событий A, B, C определяется по формуле

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - \\ &- P(B \cap C) + P(A \cap B \cap C). \end{aligned} \quad (1.5)$$

Доказательство:

$$\begin{aligned}
 P(A \cup B \cup C) &= \left| \text{обозначим } E = B \cup C \right| = P(A \cup E) = P(A) + P(E) - P(A \cap E) = \\
 &= \left| \begin{array}{l} P(E) = P(B \cup C) = P(B) + P(C) - P(B \cap C), \\ P(A \cap E) = P(A \cap (B \cup C)) = \left| \begin{array}{l} \text{применяем закон} \\ \text{дистрибутивности} \end{array} \right| = \\ = P((A \cap B) + (A \cap C)) = P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)) = \\ = \left| \text{т. к. } A \cap A = A \right| = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C); \\ = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{array} \right| =
 \end{aligned}$$

Теорема 5. Вероятность суммы конечного множества произвольных событий равна разности между единицей и вероятностью произведения противоположных событий.

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n). \quad (1.6)$$

Доказательство:

Используем принцип двойственности и применим метод математической индукции.

$$\overline{A \cup B} = \bar{A} \cap \bar{B},$$

Для двух событий имеем

$$P(A_1 \cup A_2) = 1 - P(\overline{A_1 \cup A_2}) = 1 - P(\bar{A}_1 \cap \bar{A}_2).$$

Для трех событий

$$\begin{aligned}
 P(A_1 \cup A_2 \cup A_3) &= P((A_1 \cup A_2) \cup A_3) = 1 - P(\overline{(A_1 \cup A_2) \cup A_3}) = \\
 &= 1 - P(\overline{(A_1 \cup A_2)} \cap \bar{A}_3) = 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3).
 \end{aligned}$$

Продолжая далее, получим для n событий

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n).$$

Докажем формулу (1.6) для $n + 1$ события. Обозначим

$$C = A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n, \quad \bar{C} = \bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \cap \dots \cap \bar{A}_n.$$

Тогда для $n + 1$ события получим

$$\begin{aligned}
 P(A_1 \cup A_2 \cup \dots \cup A_n \cup A_{n+1}) &= \left| \text{применяем закон ассоциативности} \right| = \\
 &= P((A_1 \cup A_2 \cup \dots \cup A_n) \cup A_{n+1}) = P(C \cup A_{n+1}).
 \end{aligned}$$

Используем следствие 2 теоремы 2:

$$\begin{aligned}
 P(C \cup A_{n+1}) &= 1 - P(\overline{C \cup A_{n+1}}) = 1 - P(\bar{C} \cap \bar{A}_{n+1}) = \\
 &= 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n \cap \bar{A}_{n+1}).
 \end{aligned}$$

Раз формула (1.6) справедлива для $n + 1$ события, значит, она справедлива для любого n .

Отметим, что иногда эта теорема называется теоремой о появлении хотя бы одного из событий $A_1 \cup \dots \cup A_n$.

1.8. Условная вероятность. Независимость событий

Пусть (Ω, F, P) – математическая модель вероятностного эксперимента. В этой математической модели каждому событию $A \in F$ соответствует вполне определенная вероятность $0 \leq P(A) \leq 1$, которая в данной модели (Ω, F, P) остается неизменной. Предположим, что в ходе эксперимента произошло событие $B \in F$. Полученная дополнительная информация может изменить вероятности других событий, связанных с событием B . Для того чтобы пересчитать эти вероятности, необходимо построить новую математическую модель эксперимента с учетом дополнительной информации (событие B произошло) и пересчитать вероятности для интересующих нас событий.

Пример 1.

Допустим, что студент из 30 билетов успел выучить билеты с 1-го по 3-й и с 28-го по 30-й. На экзамен он пришел десятым, и оказалось, что к его приходу остались только билеты с 1-го по 21-й (событие A). Вероятность события B – студент получит выученный билет без дополнительной информации о том, что событие A произошло, может быть вычислена по классическому определению с $\Omega = \{1, 2, \dots, 30\}$, $B = \{1, 2, 3, 28, 29, 30\}$, $P(B) = \frac{|B|}{|\Omega|} = 1/5$. При дополнительной информации (событие A произошло) мно-

жество возможных событий уже состоит из $|A| = 21$ элементарного события. А событие B вместе с A $|A \cap B|$ наступает в 3-х случаях. Поэтому после выполнения события A примем новое пространство $\Omega^* = A$ и назовем его приведенным. Все возможные подмножества (события) этого пространства будем обозначать символом $B|A$ и читать: «событие B при условии появления события A ». Вероятность появления события $B|A$ обозначают $P(B|A)$ и называют условной вероятностью события B при условии появления события A . Тогда искомая вероятность равна отношению числа элементарных событий, входящих в событие $B \cap A$, к числу элементарных событий, входящих в приведенное пространство $\Omega^* = A$: $P(B|A) = \frac{|B \cap A|}{|A|} = 3/21 = 1/7$.

Множества $B \cap A = B|A$ совпадают. Эту же вероятность можно выразить через вероятности событий $B \cap A$ и A в исходном пространстве Ω (рис. 1.12);

так как по классическому определению $P(B \cap A) = \frac{|B \cap A|}{|\Omega|}$ и $P(A) = \frac{|A|}{|\Omega|}$, полу-

чаем:

$$P(B|A) = \frac{P(B \cap A)}{P(A)}. \quad (1.7)$$

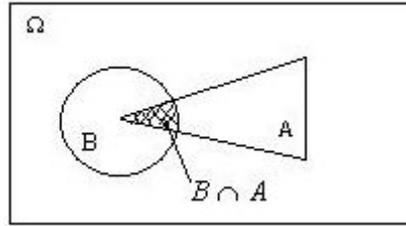


Рис. 1.12

Можно сформулировать следующее определение.

Пусть (Ω, F, P) – вероятностное пространство, и пусть A и B – произвольные события, причем $P(A) \neq 0$. Условной вероятностью наступления события B в предположении, что событие A произошло, называют число $P(B|A)$, определяемое по формуле (1.7).

Аналогично определяется условная вероятность события A в предположении, что событие B произошло ($P(B) \neq 0$):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1.8)$$

Следует отметить, что числа $P(B)$ и $P(B|A)$ определяют вероятности события B в двух различных пространствах Ω и $\Omega^* = A$. Новая σ -алгебра событий F^* пространства Ω^* строится пересечением событий F и A , а вероятности появления событий $B|A$ находится по формуле (1.8). Все аксиомы, обязательные для вероятностей, справедливы для величин $P(B|A)$. Из определения условия вероятности следует, что $0 \leq P(B|A) \leq 1$; $P(A|A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$.

С учетом изложенного и дополнительной информации о протекании эксперимента уточненной моделью этого эксперимента является новое пространство $(\Omega^* = A, F^*, P^* = \frac{P(A \cap B)}{P(A)})$. Вероятность отдельных событий $B \in F^*$ изменяется тем сильнее, чем больше информации получено о вероятностном эксперименте. Если бы мы имели в примере полную информацию о том, что, когда студент пришел, осталось 3 билета с номерами 1, 2, 3, то в этом случае вероятностный эксперимент превратился бы в детерминированный, т. е. мы могли бы утверждать, что он вытащит нужный билет.

Формулы (1.7), (1.8) позволяют сформулировать **теорему умножения вероятностей событий**.

Теорема. Вероятность произведения двух произвольных событий равна произведению вероятности одного из этих событий на условную вероятность второго события, вычисленную в предположении, что первое событие произошло:

$$P(A \cap B) = P(A) \cdot P(B | A) = P(B) \cdot P(A | B). \quad (1.9)$$

Доказательство:

Из формул (1.7), (1.8) следует (1.9).

Теорема умножения для трех событий – A, B, C:

$$\begin{aligned} P(A \cap B \cap C) &= P((A \cap B) \cap C) = P(A \cap B) \cdot P(C | A \cap B) = \\ &= P(A) \cdot P(B | A) \cdot P(C | A \cap B). \end{aligned} \quad (1.10)$$

Теорема умножения для n произвольных событий:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}). \quad (1.11)$$

С условной вероятностью тесно связано понятие *независимости событий*. Например, при броске двух игральных костей число очков, выпавших на одной кости, не зависит от числа очков, выпавших на другой кости; два блока в приборе работают независимо друг от друга (например в телевизоре – блок строчной развертки и блок УНЧ) – это физическая независимость, при которой объекты не оказывают влияния друг на друга. Физическую независимость можно сформулировать на математическом языке и дать следующее определение.

*Пусть проводится эксперимент (Ω, F, P) – вероятностное пространство этого эксперимента A и B – два произвольных события $(A \in F$ и $B \in F)$. В этом случае **событие A не зависит от события B**, если появление события B не изменяет вероятности появления события A, т. е. если условная вероятность события A равна его безусловной вероятности, то*

$$P(A|B) = P(A). \quad (1.12)$$

Если же $P(A|B) \neq P(A)$, то событие A является зависимым от события B.

Для независимых событий теорема умножения принимает вид

$$P(A \cap B) = P(A) \cdot P(B). \quad (1.13)$$

Пример 1.

В ящике 2 диода и 8 транзисторов. Какова вероятность того, что вынимая не глядя два раза подряд по одной детали, мы оба раза достанем транзисторы?

Решение.

Введем: событие A – из ящика первый раз достали транзистор, событие B – из ящика второй раз достали транзистор. Тогда вероятность события A определяется так: $P(A) = \frac{|A|}{|\Omega|} = 8/10 = 0,8$. Условная вероятность события B при условии, что произошли события A (здесь благоприятствующих событий будет $|B|A| = 7$ и приведенное пространство элементарных событий Ω^* будет состо-

ять из $|\Omega^*| = 9$ событий. Тогда условную вероятность события B при условии, что событие A произошло найдем так: $P(B|A) = |B \cap A| / |\Omega^*| = 7/9$.

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{8}{10} \cdot \frac{7}{9} = 56/90.$$

Теорема. Пусть (Ω, F, P) – вероятностное пространство. События A, B, C ($A \in F \dots$) называются независимыми в целом, если они попарно независимы и если каждое из этих событий независимо от двух других:

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

$$P(AC) = P(A) \cdot P(C),$$

$$P(BC) = P(B) \cdot P(C).$$

Данные события являются попарно независимыми. Тогда все три события являются независимыми, т. е. любое из этих событий независимо от двух других:

$$P(ABC) = P(A) \cdot P(B) \cdot P(C).$$

Пример 2.

Составлена схема электрической цепи (рис. 1.13).

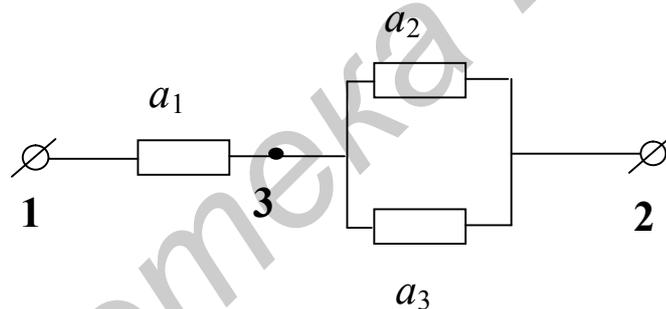


Рис. 1.13

Заданы вероятности работы элементов схемы a_1, a_2, a_3 , соответственно $p_1 = 0,2, p_2 = 0,3, p_3 = 0,4$. Найти вероятность того, что ток пройдет с клеммы 1 на клемму 2.

Решение.

Введем события: A_1 – работает элемент a_1 , A_2 – работает элемент a_2 , A_3 – работает элемент a_3 . Тогда $P(A_1) = p_1, P(A_2) = p_2, P(A_3) = p_3$. Введем в схему точку 3, которая разбивает схему на два последовательно соединенных участка, через которые проходит ток – от точки 1 до точки 3 и от точки 3 до точки 2. Рассмотрим событие B – ток пройдет из точки 1 в точку 3 схемы. Поскольку между точками 1 и 3 включен только один элемент a_1 , то событие B выполнится тогда, когда будет работать элемент a_1 , значит $B = A_1$. Введем событие C , состоящее в том, что ток пройдет из точки 3 в точку 2. Оно выполнится тогда, когда будет работать или элемент a_2 , или a_3 , или a_2 и a_3 . Мы видим, что для описа-

ния события C надо использовать теоремы о сумме или объединении двух *произвольных* событий A_2 и A_3 . Запишем $C = A_2 \cup A_3$.

Теперь рассмотрим событие E – ток пройдет из точки 1 в точку 2. Оно выполнится тогда, когда выполнится *и* событие B , *и* событие C . Здесь надо использовать теоремы об умножении или пересечении двух событий, запишем $E = B \cap C$. Тогда вероятность события E определяется по формуле (1.9), а с учетом того, что события B и C независимы применяем формулу (1.13):

$$P(E) = P(B) \cdot P(C).$$

Найдем $P(B) = P(A_1) = p_1$. Определим $P(C)$, применяя теорему о вероятности суммы двух *произвольных* событий (т. к. события A_2 и A_3 могут происходить *совместно*) и используя формулы (1.4) или (1.6):

$$\begin{aligned} P(C) &= P(A_2 \cup A_3) = 1 - P(\bar{A}_2 \cap \bar{A}_3) = \left| \text{события } A_2 \text{ и } A_3 \text{ – независимы} \right| = \\ &= 1 - P(\bar{A}_2) \cdot P(\bar{A}_3) = \left| P(\bar{A}_2) = 1 - P(A_2) = 1 - p_2, P(\bar{A}_3) = 1 - P(A_3) = 1 - p_3 \right| = \\ &= 1 - (1 - p_2)(1 - p_3). \end{aligned}$$

Тогда вероятность того, что ток пройдет из точки 1 в точку 2

$$P(E) = P(B) \cdot P(C) = p_1 \cdot (1 - (1 - p_2)(1 - p_3)) = 0,2 \cdot (1 - 0,7 \cdot 0,6) = 0,116.$$

1.9. Формула полной вероятности

Проводится эксперимент, в результате которого происходит событие A . Пусть событие A может наступить вместе с одним из несовместных событий $H_i, i = \overline{1, n}$, объединение которых совпадает с пространством элементарных событий Ω , т. е. $\Omega = \{ H_1 \cup H_2 \cup \dots \cup H_n \}$, $H_i \cap H_j = \emptyset$ при $i \neq j$. События H_i обычно называют *гипотезами*. Их вероятности $P(H_i), i = \overline{1, n}$ предполагаются известными (H_i составляют полную группу событий и обязательным условием должно быть $\sum_{i=1}^n P(H_i) = 1$), и заданы условные вероятности $P(A | H_i), i = \overline{1, n}$ наступления события A при выполнении каждой из гипотез. Тогда справедлива следующая теорема:

Теорема. Вероятность события A , которое может произойти вместе с одной из несовместных гипотез H_1, H_2, \dots, H_n равна сумме попарных произведений вероятностей каждой из этих гипотез на соответствующие им условные вероятности наступления события A при H_i -й гипотезе:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A | H_i). \quad (1.14)$$

Эта формула и называется **формулой полной вероятности**.

Доказательство:

Для доказательства приведем рис. 1.14. Для упрощения рисунка отобразим три гипотезы H_1, H_2, H_3 , но доказательство будем проводить для n гипотез.

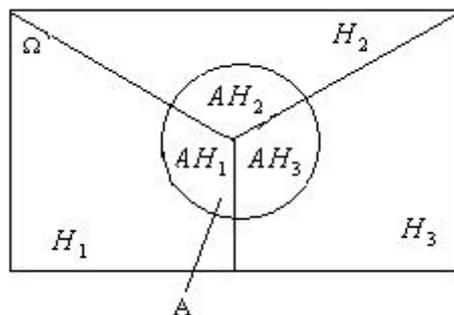


Рис. 1.14

Из рисунка видно, что $A = \{ (AN_1) \cup (AN_2) \cup \dots \cup (AN_n) \}$, причем $AN_i \cap AN_j = \emptyset$ при $i \neq j$, т. е. события $AN_i = A \cap H_i, i = \overline{1, n}$ – несовместные. На основании аксиомы 3 и теоремы умножения вероятности (1.9) имеем

$$P(A) = \sum_{i=1}^n P(A \cap H_i) = \sum_{i=1}^n P(H_i) \cdot P(A|H_i). \quad (1.15)$$

Пример.

В двух одинаковых коробках имеется по 100 резисторов. В 1-й – 60 резисторов по 100 КОм, во 2-й 30 – резисторов по 100 КОм. Определить вероятность того, что взятый наугад из какой-либо коробки резистор будет 100 КОм.

Решение.

Пусть событие A – достали резистор 100 КОм, гипотезы: H_1 – выбрали 1-ю коробку, H_2 – выбрали 2-ю коробку. Так как коробки выбирали произвольно, то $P(H_1) = 0,5, P(H_2) = 0,5$. Условная вероятность того, что взяли резистор 100 КОм, при условии, что выбрана 1-я коробка – $P(A|H_1) = 0,6$, соответственно $P(A|H_2) = 0,3$. Тогда, применяя формулу (1.14) для $n = 2$, получаем

$$P(A) = \sum_{i=1}^2 P(H_i) \cdot P(A|H_i) = 0,45.$$

1.10. Формула Байеса

В формуле полной вероятности используются вероятности гипотез H_i до проведения эксперимента. Но если в ходе эксперимента происходит событие A , то получаем дополнительную информацию о протекании эксперимента, которая может изменить вероятности гипотез H_i . Сформулируем следующую задачу.

Пусть имеется полная группа несовместных гипотез $H_i, i = \overline{1, n}$, объединение которых совпадает с пространством элементарных событий $\Omega = \{H_1 \cup H_2 \cup \dots \cup H_n\}, H_i \cap H_j = \emptyset$ при $i \neq j$. До проведения эксперимента

известны вероятности каждой из гипотез $P(H_i)$, $i = \overline{1, n}$, $\sum_{i=1}^n P(H_i) = 1$. Известны также до проведения эксперимента условные вероятности события A при каждой гипотезе: $P(A|H_i)$, $i = \overline{1, n}$. Проводится эксперимент, в результате которого появляется событие A . Требуется найти новые вероятности гипотез с учетом того, что произошло событие A .

Теорема Байеса. Условная вероятность гипотезы H_i , $i = \overline{1, n}$ после проведения эксперимента равна произведению вероятности этой гипотезы до проведения эксперимента на соответствующую ей условную вероятность события A , которое произошло при испытании, деленному на полную вероятность события A :

$$P(H_i | A) = \frac{P(H_i) \cdot P(A | H_i)}{\sum_{i=1}^n P(H_i) \cdot P(A | H_i)}. \quad (1.16)$$

Доказательство:

По определению условной вероятности (1.8) $P(A | H_i) = \frac{P(H_i \cap A)}{P(A)}$, но по теореме умножения (1.10) $P(H_i \cap A) = P(H_i) \cdot P(A | H_i)$, а $P(A)$ определим из (1.14).

Формулы Байеса позволяют пересматривать, пересчитать вероятности гипотез после получения добавочной информации о ходе вероятностного эксперимента (событие A появилось).

Пример.

Вероятность дождливого дня в городе равна 0,2. Известно, что вероятность выиграть футбольный матч команде этого города в дождливый день равна 0,4, а в сухой – 0,7. Известно, что команда выиграла матч. Определить, что в этот день шел дождь.

Решение.

Событие A состоит в том, что команда выиграла матч. Гипотезы: H_1 – шел дождь, H_2 – дождя не было.

$P(H_1) = 0,2$, $P(H_2) = 0,8$, $P(A | H_1) = 0,4$, $P(A | H_2) = 0,7$. Чтобы ответить на вопрос, пересмотрим вероятность 1-й гипотезы с учетом результата опыта – появилось событие A . Определим апостериорную вероятность гипотезы H_1 с учетом результата опыта (появилось событие A):

$$P(H_1 | A) = \frac{P(H_1) \cdot P(A | H_1)}{P(A)} = \frac{0,2 \cdot 0,4}{0,4 \cdot 0,2 + 0,8 \cdot 0,7} = 0,125.$$

Из полученного результата видим, что с учетом события A вероятность дождя в городе уменьшилась ($0,125 < 0,2$), значит скорее всего дождя не было.

1.11. Последовательность независимых испытаний. Формула Бернулли

При решении научных и практических задач часто приходится проводить многократно повторяющиеся испытания в примерно одинаковых условиях. При этом результаты предыдущих испытаний никак не сказываются на последующих.

Очевидно, что n раз независимо повторенным испытаниям отвечает вероятностное пространство $(\Omega_n, F_n, P_n) = [\times(\Omega, F, P)]^n$, в котором $\Omega_n = [\times\Omega]^n$, $F_n = [\times F]^n$, а вероятность P_n задана равенствами

$$P_n(\{\omega_i \omega_j \dots \omega_k\}) = P_i P_j P_k, \quad i, j, k = 1, 2. \quad (1.17)$$

Определение. Пусть (Ω, F, P) – дискретное вероятностное пространство эксперимента E , $\Omega = \{\omega_1 \omega_2 \dots\}$, $P_n(\{\omega_j\}) = P_j$, $j=1, 2, \dots$, $\sum_{j=1}^{\infty} P_j = 1$. Пусть эксперимент проводился n раз, тогда последовательностью независимых испытаний называется вероятностное пространство (Ω_n, F_n, P_n) , элементарными событиями в котором являются последовательности $(\omega_{i_1} \omega_{i_2} \dots \omega_{i_k})$ и вероятность для каждого элементарного события определена равенством (1.17).

Последовательность независимых испытаний называется схемой Бернулли, если $\Omega = \{\omega_1 \omega_2\}$, т. е. если эксперимент имеет лишь два исхода. Обычно в схеме Бернулли исход ω_1 называют успехом и обозначают соответствующую вероятность $P(\omega_1) = P$; ω_2 – называют неудачей, $P(\omega_2) = 1 - P = q$. В серии из n независимых испытаний в схеме Бернулли $|\Omega_n|$ состоит из 2^n элементарных событий.

Пусть проводится n независимых испытаний, в результате каждого испытания может появиться событие A с определенной вероятностью P .

Если вероятность события A в каждом испытании не зависит от исходов других испытаний, то такие испытания называются **независимыми испытаниями** относительно события A .

Определим вероятность того, что в результате проведения n независимых испытаний событие A наступит ровно m раз, если в каждом из этих испытаний данное событие наступает с постоянной вероятностью $P(A) = p$. Вероятность противоположного события \bar{A} : $P(\bar{A}) = 1 - p = q$. Обозначим через A_i ($i = \overline{1, n}$) наступление события A в i -ом испытании, тогда $P(A_i) = p$, $P(\bar{A}_i) = 1 - p = q$, $i = \overline{1, n}$.

Найдем вероятность того, что событие A при n испытаниях наступает ровно m раз, а в оставшихся $n - m$ испытаниях наступает противоположное ему событие \bar{A} . Учтем, что событие A в n испытаниях наступит ровно m раз в разных последовательностях или комбинациях, число которых равно числу сочетаний из n по m т. е. C_n^m . Рассмотрим комбинацию B_i , в которой событие A на-

ступает подряд m раз начиная с 1-го испытания, а оставшиеся $n-m$ раз наступает противоположное событие \bar{A} :

$$C_n^m \begin{cases} B_i = A_1 \cap A_2 \cap \dots \cap A_m \cap \bar{A}_{m+1} \cap \dots \cap \bar{A}_n, \\ B_{i+1} = A_1 \cap A_2 \cap \dots \cap A_{m-1} \cap \bar{A}_m \cap A_{m+1} \cap \bar{A}_{m+2} \cap \dots \cap \bar{A}_n, \\ \dots \end{cases} \quad (1.18)$$

В комбинации B_{i+1} (1.18) событие A наступает также m раз, а оставшиеся $n-m$ раз наступает событие \bar{A} , только в другой последовательности. И таких комбинаций можно выписать как число сочетаний C_n^m . Так как по условию испытания независимы, а значит, независимы и события, входящие в комбинации (1.18), используя теорему умножения, находим

$$C_n^m \begin{cases} P(B_i) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_m) \cdot P(\bar{A}_{m+1}) \cdot \dots \cdot P(\bar{A}_n) = p^m q^{n-m}, \\ P(B_{i+1}) = P(A_1) \cdot \dots \cdot P(A_{m-1}) \cdot P(\bar{A}_m) \cdot P(A_{m+1}) \cdot P(\bar{A}_{m+2}) \cdot \dots \cdot P(\bar{A}_n) = p^m q^{n-m}, \\ \dots \end{cases} \quad (1.19)$$

Видим, что вероятности комбинаций B_i одинаковы. Поскольку все комбинации B_i (1.18) являются несовместными, и нам все равно, в какой последовательности появляются события A и \bar{A} , то, применяя теорему сложения вероятностей для несовместных событий B_i , получим

$$P(B_1 \cup B_2 \cup \dots) = \sum_{i=1}^{C_n^m} P(B_i) = C_n^m \cdot p^m \cdot q^{n-m}. \quad (1.20)$$

Это и есть формула Бернулли, ее обычно записывают в следующем виде:

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m} = \frac{n!}{m!(n-m)!} \cdot p^m \cdot q^{n-m}. \quad (1.21)$$

Совокупность вероятностей $P_{m,n}$, $m = 0, 1, \dots, n$ называется *биномиальным распределением*, т. к. вероятность (1.21) является общим членом разложения бинома

$$1 = (p + q)^n = \sum_{j=0}^n C_n^j p^j q^{n-j}.$$

Это равенство в данном случае является отражением того факта, что в серии из n независимых испытаний по схеме Бернулли исходы, содержащие $0, 1, \dots, n$ успехов, образуют полную группу попарно несовместных событий.

Пример.

Проводятся испытания 10-ти телевизоров на надежность в течение времени T . Вероятность выхода телевизора из строя в течении времени T равна 0,2. Считая, что телевизоры выходят из строя независимо друг от друга, найти:

- а) вероятность выхода из строя 3-х телевизоров из 10;
- б) вероятность выхода из строя не менее 3-х телевизоров и не более 6;
- в) вероятность выхода из строя хотя бы одного телевизора.

Решение.

а) применим формулу Бернулли (1.21): $n = 10$; $m = 3$; $p = 0,2$; $q = 0,8$;
 $P_{10}(3) = C_{10}^3 \cdot 0,2^3 \cdot 0,8^7 \approx 0,201$.

б) используем теорему сложения независимых событий:

$$P_{10}(3 \dots 6) = P_{10}(3) + P_{10}(4) + P_{10}(5) + P_{10}(6) = \dots$$

в) здесь вероятность того, что выйдет из строя или 1, или 2, ..., или 10 телевизоров, удобно определить через противоположное событие – $P_{10}(1, \dots, 10) = 1 - P_{10}(0) = 1 - C_{10}^0 \cdot 0,2^0 \cdot 0,8^{10} = \dots$

1.11.1. Наивероятнейшее число наступлений события при повторении испытаний

Определение. Наивероятнейшим числом m_0 появления события A в n независимых испытаниях называется число, для которого вероятность $P_n(m_0)$ больше или, по крайней мере, не меньше вероятности каждого из остальных возможных исходов испытаний.

Пусть наивероятнейшему числу m_0 соответствует вероятность

$$P_n(m_0) = C_n^{m_0} \cdot p^{m_0} \cdot q^{n-m_0} = \frac{n!}{m_0!(n-m_0)!} \cdot p^{m_0} \cdot q^{n-m_0}.$$

Согласно определению наивероятнейшего числа, вероятность наступления события A $m_0 + 1$ и $m_0 - 1$ раз не должна превышать вероятность $P_n(m_0)$, т. е.

$$P_n(m_0) \geq P_n(m_0 + 1), \quad (1.22)$$

$$P_n(m_0) \geq P_n(m_0 - 1). \quad (1.23)$$

Решим неравенство (1.22) относительно m_0 :

$$\frac{n!}{m_0!(n-m_0)!} \cdot p^{m_0} \cdot q^{n-m_0} \geq \frac{n!}{(m_0+1)!(n-m_0-1)!} \cdot p^{m_0+1} \cdot q^{n-m_0-1},$$

$$\frac{q}{n-m_0} \geq \frac{p}{m_0+1}, \quad m_0 \geq np - q.$$

Аналогично решая неравенство (1.23), получаем: $m_0 \leq np + p$.

Объединяя два последних неравенства, имеем

$$np - q \leq m_0 \leq np + p. \quad (1.24)$$

Пример.

Вероятность поражения самолета одним осколком при подрыве вблизи него зенитной ракеты равна 0,01. Боеголовка ракеты при подрыве разлетается на тысячу осколков. Найти наивероятнейшее число попавших в самолет осколков и вероятность поражения самолета.

Решение.

Обозначим $p = 0,01$, $q = 1 - p = 0,99$, $n = 1000$. Подставляя данные значения в (1.24), имеем

$$1000 \cdot 0,01 - 0,99 \leq m_0 \leq 1000 \cdot 0,01 + 0,01; \quad 9,01 \leq m_0 \leq 10,01; \quad m_0 = 10.$$

Вероятность поражения самолета

$$P_{1000}(10) = C_{1000}^{10} \cdot (0,01)^{10} \cdot (0,99)^{990} \approx 0,126.$$

ГЛАВА 2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

2.1. Определение случайной величины

Вы уже знакомы с тем, что основой научного исследования в теории вероятности является эксперимент (опыт) и наблюдение. Эксперименты могут давать различные результаты в зависимости от того комплекса условий, в которых они происходят. Результаты эксперимента (для краткости опыта) можно охарактеризовать *качественно и количественно*. *Качественная характеристика случайного результата опыта есть событие*. Например, попадание в цель при выстреле является событием; выход прибора из строя за время t – тоже событие; интервал времени – тоже событие; выпадение какого-либо числа при бросании кости – событие. Но случайный результат опыта можно охарактеризовать и количественно. Например, *число попаданий* в цель при трех выстрелах; *число телевизоров*, выходящих из строя за время t , и др. *Количественной характеристикой случайного результата опыта является случайная величина*.

Случайной величиной называется величина, которая в результате эксперимента может принять то или иное (но только одно) значение, причем заранее неизвестно, какое именно.

Случайные величины будем обозначать заглавными буквами латинского алфавита – X, Y, Z, \dots , а их возможные значения – малыми x, y, z, \dots .

Так как случайные величины количественно характеризуют событие ω , т. е. ее значение зависит от события, то говорят, что *случайная величина есть функция от события, и обозначают $X(\omega)$* . В качестве аргумента здесь выступает ω – элементарное событие, которому соответствует значение функции $X(\omega) = x$.

Случайные величины можно разделить на три типа: дискретные, непрерывные, смешанные.

Дискретной случайной величиной (ДСВ) называется величина, которая принимает отделенные друг от друга значения, число возможных значений которой конечно или счетно, и которые могут быть перечислены, перенумерованы одно за другим.

Примеры дискретных случайных величин:

1. Пусть случайная величина X – число выстрелов до первого попадания в цель.

Здесь случайная величина X может принимать счетное множество значений: $x_1 = 1, x_2 = 2, x_3 = 3, \dots$.

2. Пусть случайная величина Y – число телефонных звонков, поступающих в вашу квартиру в течение суток. Тогда случайная величина Y может принимать следующие значения: $y_1 = 0, y_2 = 1, y_3 = 2, \dots$.

3. Если случайная величина Z – число дефектных транзисторов в партии из n штук, тогда возможные значения этой случайной величины могут принимать значения: $z_0 = 0, z_1 = 1, \dots, z_n = n$.

Непрерывной случайной величиной (НСВ) называется величина, возможные значения которой непрерывно заполняют некоторый интервал числовой оси (конечный или бесконечный).

Число возможных значений непрерывной случайной величины бесконечно. Примеры непрерывной случайной величины:

1. Случайное отклонение по дальности точки падения снаряда от цели (рис. 2.1).

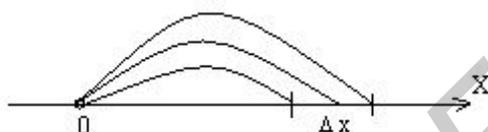


Рис. 2.1

Данный интервал, рис. 2.1, Δx представляет собой интервал рассеивания, все числа этого интервала будут возможными значениями случайной величины X – отклонения точки падения от цели. Отделить и перечислить возможные значения случайной величины X здесь невозможно.

2. Ошибка при измерении высоты радиолокационным высотомером.

3. Время безотказной работы телевизора.

Смешанной случайной величиной называется величина, которая на одних участках числовой оси принимает дискретные значения, а на других непрерывные.

Например, специальные радиотехнические сигналы, изображенные на рис. 2.2.

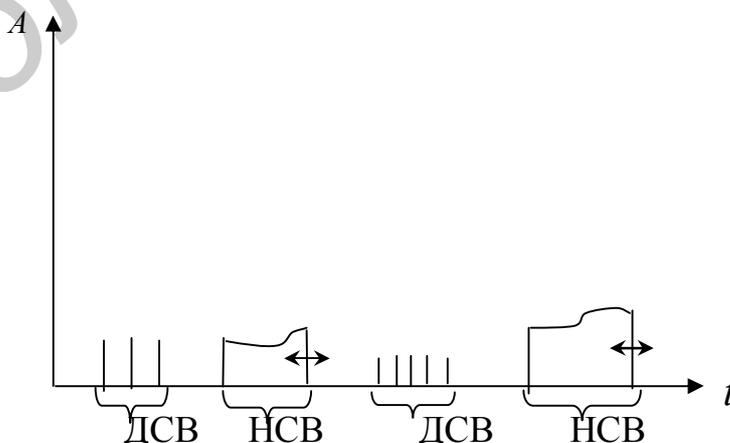


Рис. 2.2

Здесь на одних участках числовой оси появляются пачки импульсов – количество импульсов описывает дискретная случайная величина. На других участках числовой оси появляются импульсы, длительность которых меняется случайным образом – длительность импульсов описывает непрерывная случайная величина.

2.2. Законы распределения случайных величин

Пусть нам известны значения x_1, x_2, \dots, x_n дискретной случайной величины. Однако знание возможных значений случайной величины не позволяет полностью описать случайную величину, т. к. мы не знаем, как часто следует ожидать появления тех или иных возможных значений случайной величины в результате повторения опыта при одном и том же комплексе условий. В результате опыта, случайная величина X может принять только одно из возможных значений $\{X = x_1\}, \{X = x_2\}, \dots, \{X = x_n\}$. Эти события являются несовместными и образуют полную группу событий, т. к. никаких других событий в результате опыта произойти не может. Обозначим вероятности этих событий: $P(X = x_1) = p_1, P(X = x_2) = p_2, \dots, P(X = x_n) = p_n$. Так как события $\{X = x_1\}, \{X = x_2\}, \dots, \{X = x_n\}$ образуют полную группу несовместных событий, то $\sum_{i=1}^n P(X = x_i) = \sum_{i=1}^n p_i = 1$. Эта

суммарная вероятность каким-то образом распределена между отдельными значениями случайной величины. Дискретная случайная величина будет полностью описана с вероятностной точки зрения, если будет указано, какую вероятность имеет каждое из событий.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.

Способы или формы представления закона распределения случайной величины могут быть различны. Простейшей формой задания закона распределения дискретной случайной величины X является *ряд распределения*. Это таблица, в которой перечислены возможные значения случайной величины и соответствующие им вероятности.

X	$X = x_1$	$X = x_2$...	$X = x_n$
P	p_1	p_2	...	p_n

Для любой таблицы всегда должно выполняться $\sum_{i=1}^n p_i = 1$.

Графическое изображение ряда распределения называется многоугольником распределения (рис. 2.3).

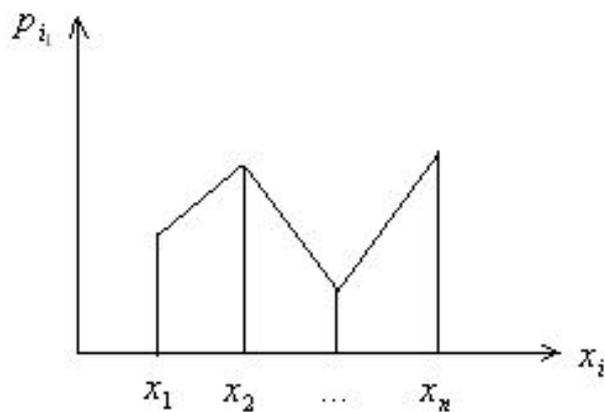


Рис. 2.3

Вершины ординат соединяются ломаной линией только для наглядности, т. к. в промежутках x_1 и x_2 , x_2 и x_3 , ... дискретная случайная величина никаких значений принять не может и вероятность ее появления в этих промежутках равна нулю.

Многоугольники распределения (см. рис. 2.3) могут иметь различную форму, но имеют общее свойство – сумма ординат многоугольника распределения равна единице. Это следует из того, как отмечалось ранее, что все возможные значения дискретной случайной величины X образуют полную группу несовместных событий.

2.3. Функция распределения

Для дискретной случайной величины закон распределения чаще задается в виде ряда распределения. Но для непрерывной случайной величины ряд распределения построить нельзя. Поскольку непрерывная случайная величина имеет бесчисленное множество возможных значений, которые сплошь заполняют некоторый интервал числовой оси, отделить и перечислить их в какой-либо таблице невозможно. Кроме того, каждое отдельное значение непрерывной случайной величины не обладает никакой, отличной от нуля, вероятностью. Поэтому желательно иметь наиболее общую форму закона распределения случайной величины X , каковой является *функция распределения*.

Функцией распределения $F(x)$, или интегральным законом распределения случайной величины X , называется вероятность того, что случайная величина X примет значение меньше чем x – аргумент функции $F(x)$:

$$F(x) = P(X < x). \quad (2.1)$$

Отсюда следует, что она существует для любых случайных величин: как дискретных, так и непрерывных.

Определение функции распределения имеет простую геометрическую интерпретацию. Если рассматривать случайную величину как случайную точку X на оси Ox , которая в результате опыта может занять то или иное положение (рис. 2.4), то функция распределения $F(x)$ есть вероятность того, что случайная точка X в результате опыта попадает левее точки x .

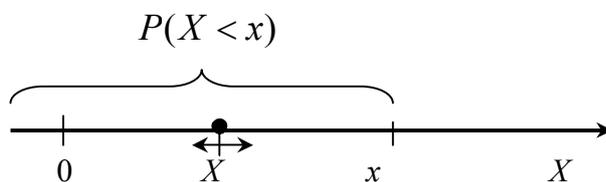


Рис. 2.4

Для дискретной случайной величины X функция распределения определяется так:

$$F(x) = \sum_{x_i < x} P(X = x_i). \quad (2.2)$$

Из выражения (2.2) видно, что суммируются вероятности для тех значений случайной величины x_i , которые по своей величине меньше аргумента x . Формула (2.2) показывает, что функция распределения дискретной случайной величины X разрывна и возрастает скачками при переходе через точки возможных ее значений x_1, \dots, x_n , причем величина скачка равна вероятности соответствующего значения. Характерный вид $F(x)$ для дискретной случайной величины приведен на рис. 2.5, $F(x)$ для непрерывной случайной величины – на рис. 2.6.

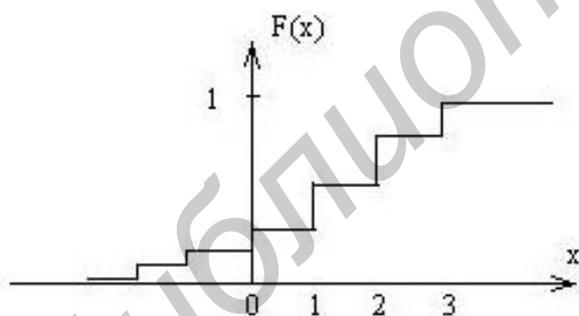


Рис. 2.5

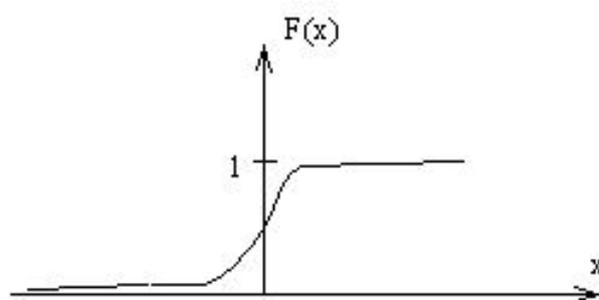


Рис. 2.6

Свойства функции распределения одномерной случайной величины.

1. Функция распределения $F(x)$ есть неотрицательная функция и принимает значения $0 \leq F(x) \leq 1$.

Доказательство.

Так как по определению (2.1) $F(x)$ – это вероятность случайного события $\{X < x\}$, а вероятность любого события $0 \leq P(\cdot) \leq 1$, значит и $0 \leq F(x) \leq 1$.

2. Вероятность попадания случайной величины в интервал $[\alpha, \beta)$, полузамкнутый слева, равна разности значений функции распределения на концах интервала:

$$P(\alpha \leq X < \beta) = F(\beta) - F(\alpha). \quad (2.3)$$

Доказательство:

На рис. 2.7. изображены события: событие A состоит в том, что случайная величина X принимает значение меньше α : $A = \{X < \alpha\}$; событие B состоит в том, что X принимает значение лежащее между α и β : $B = \{\alpha \leq X < \beta\}$; событие C состоит в том, что случайная величина X принимает значение меньше β : $C = \{X < \beta\}$.

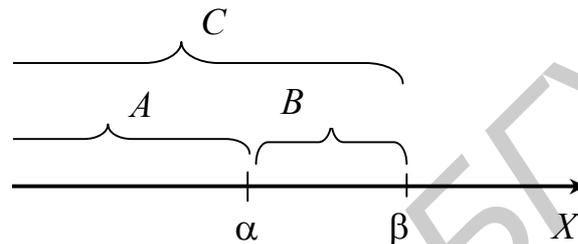


Рис. 2.7

Можно записать: $C = A \cup B$, события A и B несовместны, т. е. $A \cap B = \emptyset$. Применяя третью аксиому для несовместных событий получаем

$$P(C) = P(A) + P(B). \quad (2.4)$$

Учитываем определение функции распределения (2.1) и получаем: $P(C) = P(X < \beta) = F(\beta)$; $P(A) = P(X < \alpha) = F(\alpha)$; $P(B) = P(\alpha \leq X < \beta)$. Подставляя эти выражения в (2.4), получим: $F(\beta) = F(\alpha) + P(\alpha \leq X < \beta) \Rightarrow P(\alpha \leq X < \beta) = F(\beta) - F(\alpha)$.

Замечание. Вероятность того, что непрерывная случайная величина примет в результате опыта определенное значение, равна нулю. Покажем это.

Определим вероятность того, что непрерывная случайная величина принимает отдельно взятое значение α , для этого найдем предел формулы (2.3), когда $\beta \rightarrow \alpha$: $P(X = \alpha) = \lim_{\beta \rightarrow \alpha} P(\alpha \leq X < \beta) = \lim_{\beta \rightarrow \alpha} [F(\beta) - F(\alpha)]$. При $\beta \rightarrow \alpha$, если $F(x)$ непрерывна в точке α , предел равен нулю. Вероятность, равная нулю, свидетельствует о том, что частота этого события неограниченно убывает при увеличении числа опытов и не означает, что данное событие невозможно.

3. *Функция распределения случайной величины – неубывающая функция:* если $\beta > \alpha$, то $F(\beta) \geq F(\alpha)$. Из свойства 2 имеем $F(\beta) = F(\alpha) + P(\alpha \leq X < \beta)$, но т. к. для любого события $P(\cdot) \geq 0$ всегда, значит $F(\beta) \geq F(\alpha)$, если $\beta > \alpha$.

4. *Значение функции распределения на $-\infty$ равно нулю, а на $+\infty$ равно единице:* $F(-\infty) = 0$, $F(+\infty) = 1$.

При неограниченном перемещении точки x влево, когда $x \rightarrow -\infty$, попадание случайной точки X левее x в пределе становится невозможным событием. То есть никакая случайная величина принять значение меньше $-\infty$ не может. Невозможное событие обозначается пустым множеством \emptyset , а значит и вероятность этого события равна нулю: $F(-\infty) = P(X < -\infty) = P(\emptyset) = 0$. При неограниченном перемещении точки x вправо, когда $x \rightarrow +\infty$, попадание точки X левее x в пределе становится достоверным событием, т. е. любая случайная величина в результате опыта принимает значение меньше $+\infty$. Достоверное событие обозначается Ω , и вероятность этого события равна единице: $F(+\infty) = P(X < +\infty) = P(\Omega) = 1$.

Все свойства функции распределения можно сформулировать так: каждая функция распределения является неотрицательной, неубывающей функцией, удовлетворяющей условиям $F(-\infty) = 0$, $F(+\infty) = 1$.

Отметим, что свойства функции распределения 1, 3, 4 называются характеристическими, т. е. функцией распределения может быть только та функция, которая удовлетворяет этим свойствам.

Пример.

Для дискретной случайной величины заданной рядом распределения построить функцию распределения.

Ряд распределения имеет следующий вид.

X	$x_1 = 1$	$x_2 = 2$	$x_3 = 4$
P	0,2	0,5	0,3

Для построения функции распределения используем формулу (2.2):

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

Рассчитаем значения функции распределения для фиксированных значений x_i , взятых из таблицы.

$$1. \ x_1 = 1, F(1) = \sum_{x_i < 1} P(X = x_i) = 0.$$

$$2. \ x_2 = 2, F(2) = \sum_{x_i < 2} P(X = x_i) = P(X = 1) = 0,2.$$

$$3. \ x_3 = 4, F(4) = \sum_{x_i < 4} P(X = x_i) = P(X = 1) + P(X = 2) = 0,2 + 0,5 = 0,7.$$

4. Обязательно делаем четвертый шаг, полагаем $x_4 = +\infty$, и суммируем вероятности для $x_i < +\infty$:

$$x_4 = +\infty, F(+\infty) = \sum_{x_i < +\infty} P(X = x_i) = P(X = 1) + P(X = 2) + P(X = 4) = 0,2 + 0,5 + 0,3 = 1.$$

Строим оси координат и делаем разметку осей (рис. 2.8).

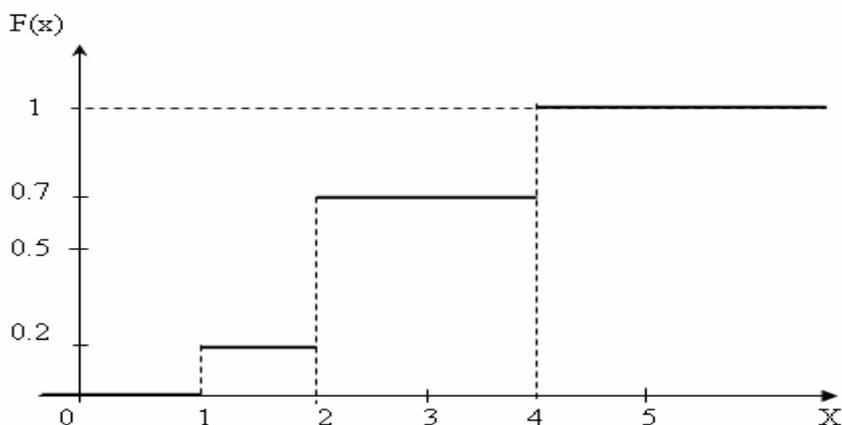


Рис. 2.8.

Опишем построение функции распределения $F(x)$. Рассмотрим первый промежуток по оси X (от $-\infty$ до 1): согласно пункту 1, значение $F(x) = 0$ и линия идет по оси X от $-\infty$ до 1 включительно. На втором промежутке по оси X (от 1 до 2), согласно пункту 2, значение $F(x) = 0,2$ – проводим ступеньку высотой 0,2. На третьем промежутке (от 2 до 4), согласно пункту 3, значение $F(x) = 0,7$ – проводим ступеньку высотой 0,7. На четвертом промежутке (от 4 до $+\infty$), согласно пункту 4 значение $F(x)=1$ – проводим ступеньку высотой 1. В точках $X = 1, 2, 4$ функция распределения $F(x)$ имеет разрыв.

2.4. Плотность распределения непрерывной случайной величины

Недостатком функции распределения непрерывной случайной величины является то, что по ней трудно судить о характере распределения случайной величины в окрестности той или иной точки числовой оси. Более наглядно описывает распределение непрерывной случайной величины функция, которая называется *плотностью распределения вероятностей* или *дифференциальным законом* распределения случайной величины.

Пусть непрерывная случайная величина X задана функцией распределения $F(x)$. Найдем вероятность попадания этой случайной величины на элементарный участок $(x, x + \Delta x)$, используя свойство 2 функции распределения: $P(x < X < x + \Delta x) = F(x + \Delta x) - F(x)$. Разделим обе части этого соотношения на Δx :

$$\frac{P(x < X < x + \Delta x)}{\Delta x} = \frac{F(x + \Delta x) - F(x)}{\Delta x}. \quad (2.5)$$

Пусть функция распределения $F(x)$ непрерывная и дифференцируемая, вычислим предел от обеих частей (2.5), когда $\Delta x \rightarrow 0$:

$$\lim_{\Delta x \rightarrow 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx} = F'(x).$$

Определение. Предел отношения вероятности попадания непрерывной случайной величины на элементарный участок от x до $x + \Delta x$ к длине этого уча-

стка Δx , когда $\Delta x \rightarrow 0$, называется **плотностью распределения** (или **плотностью вероятности**) случайной величины в точке x , обозначается $f(x)$ и вычисляется как производная от функции распределения $F(x)$:

$$f(x) = \frac{dF(x)}{dx} = F'(x). \quad (2.6)$$

Смысл плотности распределения $f(x)$ состоит в том, что она показывает, как часто появляется случайная величина X в окрестности той или иной точки числовой оси при повторении опытов. Возможный вид $f(x)$ показан на рис. 2.9.

На рис. 2.9 видно, что значение $f(x_1)$ – мало, значит в окрестности точки x_1 случайная величина появляется редко, а значение $f(x_2)$ – значительно больше $f(x_1)$, значит, в окрестности точки x_2 случайная величина появляется гораздо чаще.

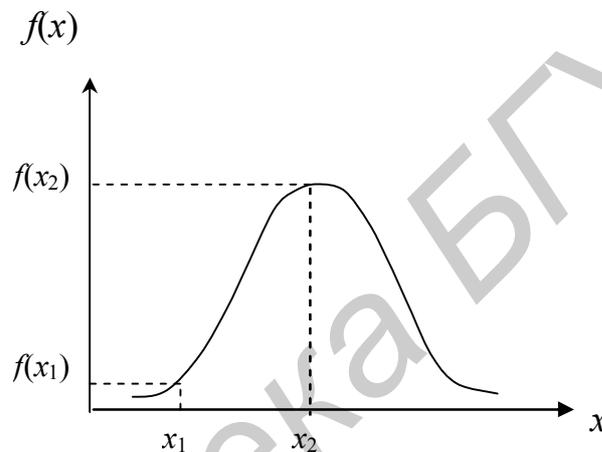


Рис. 2.9

Свойства плотности распределения:

1. Плотность распределения – положительная функция: $f(x) \geq 0$.

Это следует из того, что $f(x)$ есть производная от неубывающей функции распределения $F(x)$.

2. Функция распределения непрерывной случайной величины равна интегралу от плотности в интервале от $-\infty$ до x :

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (2.7)$$

Доказательство:

Из формулы (2.6) можно записать $dF(x) = f(x)dx$. Интегрируя обе части, получаем

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

согласно свойству 4 функции распределения $F(-\infty) = 0$.

Геометрически функция распределения представляется площадью заштрихованной фигуры на графике плотности распределения (рис. 2.10).

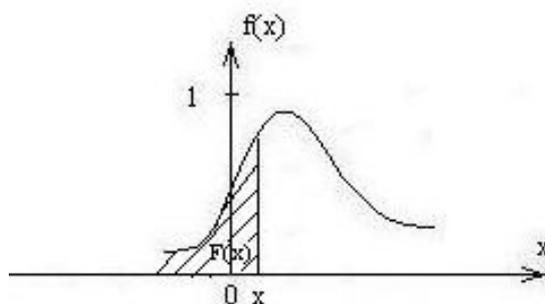


Рис. 2.10

3. Вероятность попадания непрерывной случайной величины X в интервал (α, β) равна определенному интегралу от плотности распределения, взятому по этому интервалу:

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

Доказательство:

На основании свойства 2 функции распределения и замечания к нему имеем

$$\begin{aligned} P(\alpha \leq X < \beta) &= P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \int_{-\infty}^{\beta} f(x) dx - \int_{-\infty}^{\alpha} f(x) dx = \\ &= \left| \int_{-\infty}^{\beta} \dots = \int_{-\infty}^{\alpha} \dots + \int_{\alpha}^{\beta} \dots \right| = \int_{-\infty}^{\alpha} f(x) dx + \int_{\alpha}^{\beta} f(x) dx - \int_{-\infty}^{\alpha} f(x) dx = \int_{\alpha}^{\beta} f(x) dx. \end{aligned}$$

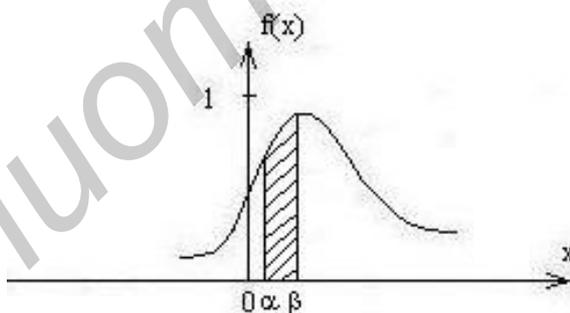


Рис. 2.11

Геометрически (рис. 2.11) это означает: вероятность того, что непрерывная случайная величина примет значение, принадлежащее интервалу (α, β) , равна площади заштрихованной фигуры.

4. *Условие нормировки.* Интеграл в бесконечных пределах от плотности распределения равен единице:

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (2.8)$$

Доказательство:

Из свойства 2 и 4 функции распределения имеем

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

Пример.

Задана плотность распределения

$$f(x) = \begin{cases} 0 & x \leq a \\ k & a < x < b \\ 0 & x \geq b \end{cases}$$

Определить коэффициент k и функцию распределения $F(x)$.

Решение.

$$\int_{-\infty}^{\infty} f(x) dx = \int_a^b k dx = k \cdot (b - a) = 1; \text{ Отсюда } k = \frac{1}{b - a}.$$

Построим график $f(x)$ (рис. 2.12).

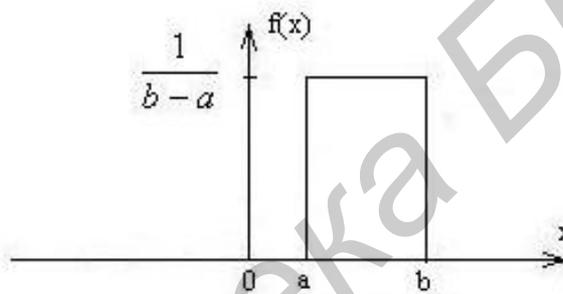


Рис. 2.12

Найдем функцию распределения, используя (2.7):

$$F(x) = \int_{-\infty}^x f(x) dx = \begin{cases} 0 & x \leq a \\ \frac{x - a}{b - a} & a < x < b \\ 1 & x \geq b \end{cases}$$

Построим график $F(x)$ (рис. 2.13).

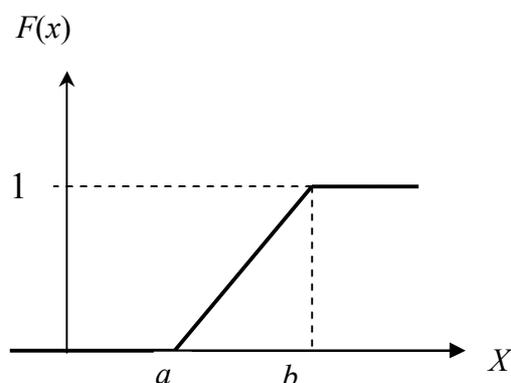


Рис. 2.13

ГЛАВА 3. МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

3.1. Понятие о многомерных (векторных) случайных величинах

При решении многих практических задач каждому элементарному событию $\omega \in \Omega$ можно поставить в соответствие не одну случайную величину, а несколько: $x_1(\omega), x_2(\omega) \dots$.

Пример 1.

Пусть вероятностный эксперимент состоит в рождении ребенка. Тогда каждому элементарному событию ω (каждому новорожденному) можно поставить в соответствие следующие числа: $X_1(\omega)$ – рост, $X_2(\omega)$ – вес, $X_3(\omega)$ – пол (0 или 1). Таким образом, эксперимент E описывается трехмерной случайной величиной, или системой трех случайных величин, или трехмерным вектором $(X_1(\omega), X_2(\omega), X_3(\omega))$.

Пример 2.

Пусть эксперимент состоит в измерении коэффициента усиления транзистора K_u на высокой частоте. Тогда элементарное событие ω – это K_u одного транзистора и ему можно поставить в соответствие следующие случайные величины: $X_1(\omega) \rightarrow |\beta|$ – коэффициент передачи тока на высокой частоте, $X_2(\omega) \rightarrow I_{ко}$ – обратный ток коллекторного перехода, $X_3(\omega) \rightarrow y_{11}$ – входная проводимость, $X_4(\omega) \rightarrow C_E$ – емкость эмиттерного перехода, $X_5(\omega) \rightarrow C_k$ – емкость коллекторного перехода, $X_6(\omega) \rightarrow y_{22}$ – выходная проводимость. Все перечисленные параметры для каждого транзистора при массовом производстве полупроводниковых приборов имеют отклонения от некоторого среднего значения (ввиду сложности поддержания технологического процесса стабильным для каждого кристалла при напылении, диффузии, травлении и т. д.). Поэтому их можно считать случайными величинами, и от значения каждого из них зависит коэффициент усиления транзистора K_u на высокой частоте. Тогда в данном случае эксперимент описывается шестимерной случайной величиной, или шестимерным вектором, или системой шести случайных величин.

При изучении многомерных случайных величин удобно пользоваться следующей геометрической интерпретацией. Например (рис. 3.1), систему двух случайных величин (X, Y) можно рассматривать как случайную точку на плоскости с координатами X, Y или как случайный вектор на плоскости со случайными составляющими X, Y .

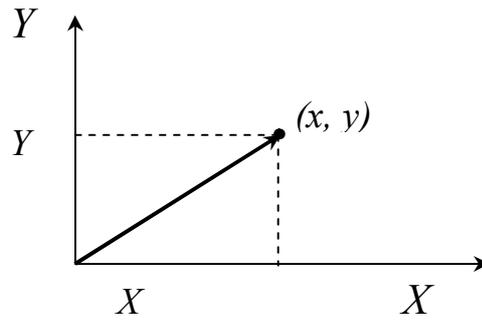


Рис. 3.1

Систему трех случайных величин (X, Y, Z) можно рассматривать как случайный вектор в трехмерном пространстве, а при n -составляющих – как n -мерный случайный вектор.

При изучении многомерных случайных величин ограничимся рассмотрением двумерной случайной величины, т. к. все положения, касающиеся двумерных случайных величин, легко распространить на n -мерные случайные величины.

3.2. Закон распределения многомерной случайной величины

Законом распределения многомерной случайной величины называется всякое соотношение, устанавливающее связь между областями возможных значений случайной величины и вероятностями появления многомерной случайной величины в этих областях.

Формы законов различны. Например, для дискретной двумерной случайной величины (X, Y) закон распределения часто задают в виде *таблицы распределения двумерной случайной величины*.

$X \backslash Y$	x_1	x_2	...	x_n
y_1	P_{11}	P_{21}	...	P_{n1}
y_2	P_{12}	P_{22}	...	P_{n2}
...
y_m	P_{1m}	P_{2m}	...	P_{nm}

В таблице P_{ij} – это вероятность совместного появления случайных величин X и Y , где $P_{ij} = P(X = x_i \cap Y = y_j)$ соответствует вероятности пересечения двух событий, при которых случайная величина X примет значение x_i и одно-

временно с этим случайная величина Y примет значение y_j :
 $P_{ij} = P(\{X = x_i\} \cap \{Y = y_j\})$. Вероятности P_{ij} сведены в таблицу.

Все возможные события, $\{X = x_i\}, \{Y = y_j\}$, $i = \overline{1, n}, j = \overline{1, m}$, составляют полную группу несовместных событий:

$$\sum_{i=1}^n \sum_{j=1}^m P(\{X = x_i\} \cap \{Y = y_j\}) = 1.$$

3.3. Функция распределения двумерной случайной величины

Как отмечали ранее, для непрерывной одномерной случайной величины задать закон распределения в виде таблицы невозможно, поэтому и для непрерывных двумерных случайных величин закон распределения задается в виде функции распределения.

Функцией распределения двумерной случайной величины называется функция двух аргументов $F(x, y)$, равная вероятности совместного появления двух событий, при которых $\{X < x\}$ и $\{Y < y\}$:

$$F(x, y) = P(X < x \cap Y < y). \quad (3.1)$$

Геометрически функция распределения двумерной случайной величины представляет собой вероятность попадания случайной точки (X, Y) в левый нижний бесконечный квадрант плоскости с вершиной в точке (x, y) (рис. 3.2).

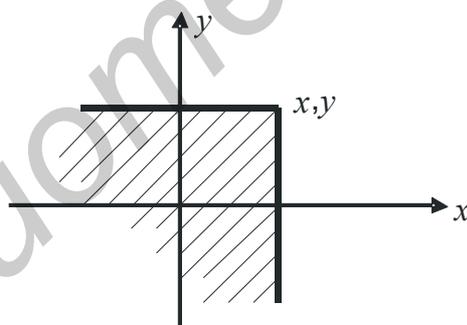


Рис. 3.2

Свойства двумерной функции распределения.

1. $0 \leq F(x, y) \leq 1$.

Это следует из определения функции распределения (3.1), т. к. $F(x, y)$ – это вероятность, а вероятность любого события лежит в пределах от нуля до единицы.

2. Если один из аргументов стремится к плюс бесконечности, то двумерная функция распределения стремится к функции распределения одной СВ, соответствующей другому аргументу.

а) пусть $x = \infty$, тогда

$$F(\infty, y) = P(X < \infty \cap Y < y) = P(\Omega \cap Y < y) = P(Y < y) = F_2(y).$$

Геометрически $F_2(y)$ – это вероятность попадания случайной точки (Y) в заштрихованную область (рис. 3.3).

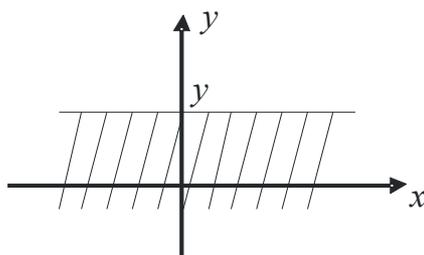


Рис. 3.3

б) пусть $y = +\infty$, тогда

$$F(x, \infty) = P(X < x \cap Y < \infty) = P(X < x \cap \Omega) = P(X < x) = F_1(x).$$

Геометрически $F_1(x)$ – это вероятность попадания случайной точки (X) в заштрихованную область (рис. 3.4).

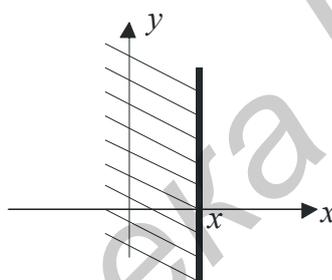


Рис. 3.4

Функции $F_1(x)$ и $F_2(y)$ называются *маргинальными функциями распределения* составляющих X и Y (название произошло от латинского слова *margo* – край, граница).

3. Если оба аргумента стремятся к $+\infty$, то функция распределения стремится к 1:

$$F(\infty, \infty) = P(X < \infty \cap Y < \infty) = P(\Omega) = 1.$$

Квадрант с вершиной (x, y) (см. рис. 3.2) обращается во всю координатную плоскость xOy , попадание случайной точки в которую есть достоверное событие.

4. При стремлении одного или обоих аргументов к $-\infty$ функция распределения стремится к 0:

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

Пусть $x = -\infty$, тогда

$$F(-\infty, y) = P(X < -\infty \cap Y < y) = P(\emptyset \cap Y < y) = P(\emptyset) = 0.$$

5. Функция распределения является *неубывающей функцией по каждому аргументу*:

$$F(x_2, y) \geq F(x_1, y), x_2 > x_1$$

$$F(x, y_2) \geq F(x, y_1), y_2 > y_1$$

6. Вероятность попадания случайной точки (X, Y) в произвольный прямоугольник со сторонами, параллельными координатным осям, вычисляется по формуле (рис. 3.5):

$$P(\alpha \leq x < \beta \cap \gamma \leq Y < \delta) = F(\alpha, \gamma) + F(\beta, \delta) - F(\alpha, \delta) - F(\beta, \gamma).$$

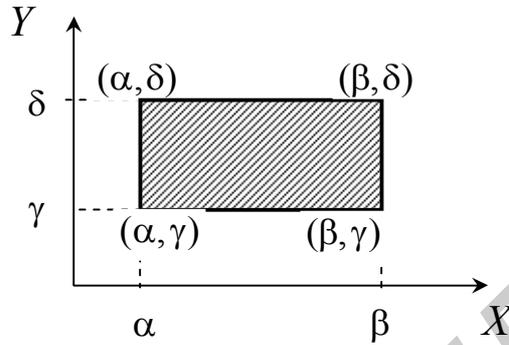


Рис. 3.5

Свойства 1–5 называются характеристическими, т. е. любая функция, удовлетворяющая этим свойствам является функцией распределения двумерной случайной величины.

3.4. Плотность распределения непрерывной двумерной случайной величины

Введение плотности распределения для непрерывных многомерных случайных величин позволяет упростить расчет вероятности попадания в различные области и делает более наглядным описание распределения многомерных случайных величин.

Плотностью распределения непрерывной двумерной случайной величины называется предел отношения вероятности попадания случайной величины (X, Y) в прямоугольник со сторонами Δx и Δy к площади этого прямоугольника, когда длины обеих сторон его стремятся к нулю, и вычисляется как вторая смешанная частная производная от функции распределения :

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(x \leq X \leq x + \Delta x \cap y \leq Y \leq y + \Delta y)}{\Delta x \Delta y} = f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \quad (3.2)$$

На рис. 3.6 изображен элементарный прямоугольник.

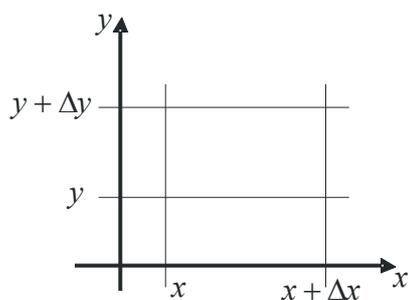


Рис. 3.6

Геометрически двумерная плотность распределения $f(x,y)$ изображается поверхностью (рис. 3.7), которую называют *поверхностью распределения*.

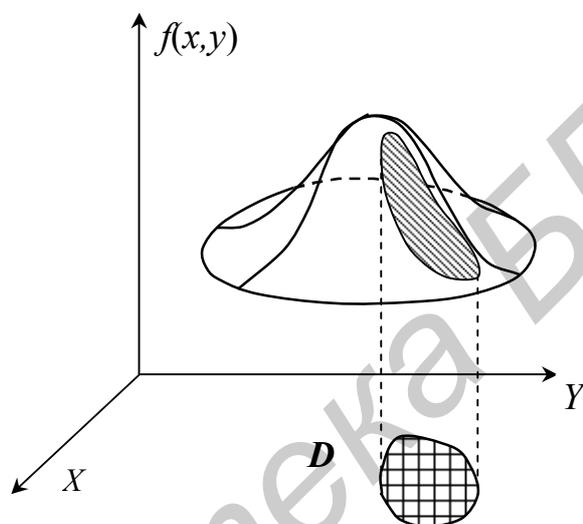


Рис. 3.7

Свойства плотности распределения двумерной случайной величины.

1. Плотность распределения – положительная функция $f(x,y) \geq 0$, т. к. это предел отношения вероятности к площади – обе величины >0 .

2. Функция распределения двумерной случайной величины через плотность распределения записывается следующим образом:

$$F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(x,y) dx dy. \quad (3.3)$$

3. Вероятность попадания в произвольную область D получается путем разбиения D на n элементарных областей $\Delta x \Delta y$ и суммирования $f(x,y) \Delta x \Delta y$ по D при $\Delta x \Delta y \rightarrow 0$ и выражается, как двойной интеграл, по заданной области D :

$$P((X,Y) \in D) = \iint_D f(x,y) dx dy. \quad (3.4)$$

Геометрически эта вероятность представляет собой объем цилиндрического тела, ограниченного сверху поверхностью распределения и опирающегося на эту область (рис. 3.7).

4. На основании свойства 2 двумерной функции распределения имеем *маргинальные плотности распределений каждой составляющей*:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx. \quad (3.5)$$

Доказательство:

Из свойства 2 двумерной функции распределения имеем

$$F_1(x) = F(x, +\infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(x, y) dx dy,$$

учитывая, что плотность распределения одномерной случайной величины равна $F'(x)$, получим после дифференцирования по x :

$$f_1(x) = F'(x) = \int_{-\infty}^{+\infty} f(x, y) dy.$$

Аналогично получаем $f_2(y)$ для другой составляющей.

Замечание. Плотность распределения одной составляющей двумерной случайной величины определяется путем интегрирования $f(x, y)$ в бесконечных пределах по аргументу, соответствующему другой составляющей случайной величины.

5. *Условие нормировки*

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1. \quad (3.6)$$

Доказательство:

Из свойства 2 двумерной плотности при $x = +\infty$ и $y = +\infty$ имеем

$$F(+\infty, +\infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = | \text{но } F(+\infty, +\infty) = 1, \text{ значит} | = 1.$$

3.5. Условные законы распределения

Если известны законы распределения отдельных составляющих двумерной случайной величины (см. свойство 2 формулы (3.5)), то для зависимых случайных величин X и Y закон распределения двумерной случайной величины (X, Y) не может быть выражен через законы распределения отдельных составляющих, т. к. не известно, как они связаны между собой.

Поэтому в теории вероятностей вводят понятие условных законов распределения.

Условным законом распределения называется распределение одной составляющей двумерной случайной величины, найденное при условии, что дру-

гая составляющая двумерной случайной величины приняла определенное значение.

Условный закон распределения может выражаться как условной функцией распределения $F(x/y)$, так и условной плотностью распределения $f(x/y)$ (для случайной величины X при условии, что другая случайная величина Y приняла определенное значение). Рассмотрим условные плотности распределения. По определению условной плотности распределения для случайной величины X при условии, что случайной величины Y приняла определенное значение, получим

$$f(x|y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(x < X < x + \Delta x | y < Y < y + \Delta y)}{\Delta x}. \quad (3.7)$$

Ранее мы вводили определение условной вероятности: $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Применим это определение к числителю формулы (3.7):

$$P(x < X < x + \Delta x | y < Y < y + \Delta y) = \frac{P(x < X < x + \Delta x \cap y < Y < y + \Delta y)}{P(y < Y < y + \Delta y)}.$$

Подставляя это выражения в числитель формулы (3.7) и разделив числитель и знаменатель (3.7) на Δy , получим

$$f(x|y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\frac{P(x < X < x + \Delta x \cap y < Y < y + \Delta y)}{P(y < Y < y + \Delta y) \Delta x \Delta y}}{\Delta y} = \frac{f(x, y)}{f_2(y)}. \quad (3.8)$$

Аналогично для составляющей Y условной плотности распределения имеем

$$f(y|x) = \frac{f(x, y)}{f_1(x)}. \quad (3.9)$$

С учетом свойства 4 плотности двумерной случайной величины формулы (3.8) и (3.9) можно переписать в следующем виде :

$$f(x|y) = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dx}, \quad f(y|x) = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy}. \quad (3.10)$$

Используя формулы (3.8) и (3.9), сформулируем следующую теорему.

Теорема умножения законов распределений. Плотность распределения двумерной случайной величины равна произведению плотности распределения одной составляющей двумерной случайной величины и условной плотности распределения другой составляющей двумерной случайной величины.

Доказательство:

Из формул (3.8) и (3.9) имеем

$$f(x, y) = f_1(x)f(y|x) = f_2(y)f(x|y). \quad (3.11)$$

Условная плотность распределения обладает всеми свойствами безусловной плотности распределения.

1. $f(x|y) \geq 0$.

2. $F(x|y) = \int_{-\infty}^x f(x|y)dx$ и $f(x|y) = \frac{dF(x|y)}{dx}$.

3. Условие нормировки:

$$F(+\infty|y) = \int_{-\infty}^{\infty} f(x|y)dx = 1.$$

3.6. Зависимые и независимые случайные величины

Предположим, что измеряемые в эксперименте случайные величины X и Y – координаты точки падения снаряда при стрельбе из орудия по цели в виде прямоугольника, обозначенного на земле. Тогда координаты точки падения одного снаряда не зависят от координат точки падения другого снаряда. Понятие независимости случайных величин является одним из важнейших в теории вероятностей.

Случайная величина X называется **независимой** от случайной величины Y , если закон распределения случайной величины X не зависит от того, какое значение приняла случайная величина Y .

$$f(x|y) = f_1(x). \quad (3.12)$$

Видим, что для независимых случайных величин условная плотность распределения равна безусловной плотности.

Если же одна случайная величина зависит от другой случайной величины, то

$$f(x|y) \neq f_1(x).$$

Покажем, что если случайная величина X не зависит от Y , то и случайная величина Y не зависит от X , т. е. зависимость и независимость случайных величин взаимны. Для доказательства используем теорему умножения законов распределений выраженную формулой (3.11).

Пусть X не зависит от Y : $f(x|y) = f_1(x)$.

Тогда из (3.11) имеем $f_1(x)f(y|x) = f_2(y)f(x|y)$, подставляя в левую часть этого равенства $f_1(x) = f(x|y)$, получим $\Rightarrow f(y|x) = f_2(y)$. Это и есть условие независимости случайной величины Y от X .

Сформулируем и докажем следующую теорему.

Для того чтобы непрерывные случайные величины X и Y были независимы, необходимо и достаточно, чтобы плотность распределения двумерной случайной величины (X, Y) была равна произведению плотностей распределения отдельных составляющих двумерной случайной величины.

$$f(x, y) = f_1(x)f_2(y). \quad (3.13)$$

Доказательство:

1. Условие *необходимости*. (Покажем что необходимость того, чтобы X и Y были независимы; тогда выполнится (3.13)). Пусть X и Y – независимые случайные величины, т. е.

$$f_{x|y} = f_1(x), \quad f(y|x) = f_2(y).$$

Подставив эти выражения в правую часть теоремы умножения (формула 3.11), получаем

$$f(x, y) = f_1(x)f_2(y).$$

2. Условие *достаточности*. (Покажем, что достаточно, чтобы выполнялось условие (3.13), тогда X и Y будут независимы). Пусть $f(x, y) = f_1(x)f_2(y)$. Тогда, используя (3.8), получим

$$f(x|y) = \frac{f(x, y)}{f_2(y)} = \left| \text{подставляя в числитель формулу (3.13), получим} \right| = f_1(x).$$

Это и есть условие независимости X от Y . Аналогично доказательство для независимости Y от X . Эта теорема имеет следствие.

Следствие. *Если двумерная плотность распределения $f(x, y)$ представима в виде произведения двух сомножителей, один из которых содержит только X , а другой только Y составляющие, то случайные величины X и Y будут независимы.*

Пример 1.

Пусть двумерная случайная величина (X, Y) задана плотностью распределения.

$$f(x, y) = \begin{cases} ax^2y & \text{если } 0 < x < 1 \cap 0 < y < 1, \\ 0 & \text{вне этой области.} \end{cases}$$

Найти коэффициент a и показать, что случайные величины X и Y независимы.

Решение.

Для определения коэффициента a используем условие нормировки

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ax^2 y dx dy = 1.$$

Учтем область определения $f(x, y)$:

$$\int_0^1 \int_0^1 ax^2 y dx dy = a \int_0^1 x^2 dx \int_0^1 y dy = a \left(\frac{x^3}{3} \Big|_0^1 \times \frac{y^2}{2} \Big|_0^1 \right) = \frac{a}{6} = 1,$$

откуда $a = 6$.

1-й способ проверки независимости X и Y .

Найдем плотности распределений отдельных составляющих $f_1(x)$ и $f_2(y)$, применяя (3.4):

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 6x^2 y dy = 6x^2 \frac{y^2}{2} \Big|_0^1 = 3x^2,$$

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^1 6x^2 y dx = 6y \frac{x^3}{3} \Big|_0^1 = 2y.$$

Подставим эти выражения в условие (3.13):

$$f(x, y) = f_1(x) f_2(y) = 6x^2 y.$$

Значит случайные величины X и Y независимы.

2-й способ проверки независимости X и Y .

Найдем условную плотность распределения:

$$f(x|y) = \frac{f(x, y)}{f_2(y)} = \frac{6x^2 y}{2y} = 3x^2 = f_1(x).$$

Видим, что выполняется условие (3.12), значит случайные величины X и Y независимы.

3-й способ проверки независимости X и Y .

Все предыдущие вычисления можно не делать, а сослаться на следствие к доказанной выше теореме. Видим, что двумерная плотность $6x^2 y$ представима в виде произведения двух сомножителей, один из которых содержит только X , а другой только Y составляющие. Тогда на основании следствия из доказанной выше теоремы случайные величины X и Y независимы.

Пример 2.

Двумерная случайная величина равномерно распределена внутри области D , выделенной на рис. 3.8 жирными прямыми:

$$f(x, y) = \begin{cases} c, & (x, y) \in D; \\ 0, & (x, y) \notin D. \end{cases}$$

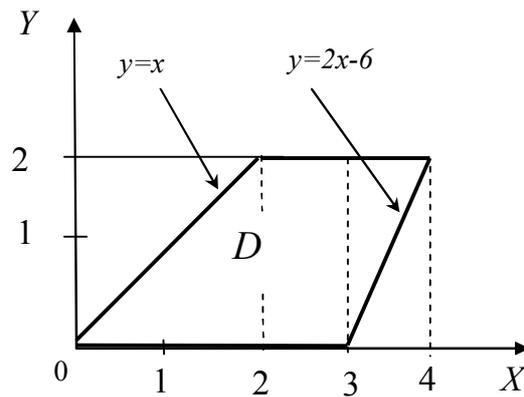


Рис. 3.8

Найти константу c .

Решение.

Запишем уравнения прямых (см. рис. 3.7): $y = x$, $y = 2x - 6$.

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

Этот интеграл вычислим в виде суммы трех интегралов, разбивая область D на три части (пунктирные прямые на рис. 3.8):

$$\begin{aligned}
 c \left[\int_0^2 dx \int_0^x dy + \int_2^3 dx \int_0^2 dy + \int_3^4 dx \int_{2x-6}^2 dy \right] &= c \left[\int_0^2 x dx + \int_2^3 2 dx + \int_3^4 [2 - (2x - 6)] dx \right] = \\
 &= c \left[\frac{x^2}{2} \Big|_0^2 + 2x \Big|_2^3 - \frac{2x^2}{2} \Big|_3^4 + 8x \Big|_3^4 \right] = c5 = 1; \Rightarrow c = 1/5.
 \end{aligned}$$

ГЛАВА 4. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

4.1. Числовые характеристики одномерной случайной величины

При решении многих практических задач нет необходимости указывать полностью закон распределения случайной величины, а можно указать наиболее характерные его черты. Поэтому для общей характеристики случайных величин используют некоторые величины, которые называются *числовыми характеристиками*. Они позволяют в сжатой форме выразить наиболее существенные особенности того или иного распределения. Важнейшими из них являются: математическое ожидание, дисперсия, среднее квадратическое отклонение.

4.1.1. Математическое ожидание

Одной из важнейших характеристик случайной величины является математическое ожидание. *Математическое ожидание* – это среднее значение случайной величины – это не точное определение, а смысл математического ожидания. Математическое ожидание легко вычисляется и обладает полезным свойством: среднее арифметическое одинаково распределенной случайной величины близко к математическому ожиданию.

Пусть $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$ – дискретное пространственное элементарных событие. Рассмотрим событие ω_k состоящее в том, что случайная величина X принимает значение x_k , т. е. $\omega_k = \{X = x_k\}$. Пусть события ω_k – равновероятны, тогда вероятности этих событий: $P(\omega_k) = \frac{1}{N}$, $k = \overline{1, N}$.

Тогда среднее значение случайной величины можно определить так:

$$(x_1 + \dots + x_N) / N = \frac{1}{N}x_1 + \dots + \frac{1}{N}x_N = \sum_{k=1}^N x_k \frac{1}{N} = \left| \text{учтем, что } \frac{1}{N} = P(\omega_k) \right| = \sum_{k=1}^N x_k P(\omega_k).$$

В общем случае вероятности $P(\omega_k) = p_k$ могут и не быть одинаковы.

Тогда среднее значение вычисляется по формуле

$$p_1x_1 + p_2x_2 + \dots + p_Nx_N = \sum_{k=1}^N p_kx_k.$$

Будем обозначать математическое ожидание случайной величины X как $M[X]$ или m_x .

Математическим ожиданием дискретной случайной величины X называется сумма произведений всех возможных значений случайной величины на их вероятности:

$$M[X] = \sum_{k=1}^N x_k p_k. \quad (4.1)$$

Рассмотрим непрерывную случайную величину X , все возможные значения которой принадлежат интервалу $[a, b]$. Пусть $f(x)$ – плотность распределения. Разобьем интервал $[a, b]$ на n отрезков $\Delta x_1, \Delta x_2, \dots, \Delta x_n$. Вероятность попадания случайной величины X на элементарный участок Δx_k равна $p_k = f(x_k) \Delta x_k$, ($k = \overline{1, n}$). Тогда по аналогии с определением математического ожидания для дискретной случайной величины запишем

$$\sum x_k p_k = \sum_{k=1}^n x_k f(x_k) \Delta x_k.$$

Перейдем к пределу при $\Delta x_k \rightarrow 0$, получим определенный интеграл:

$$\lim_{\Delta x_k \rightarrow 0} \sum_{k=1}^n x_k f(x_k) \Delta x_k = \int_a^b x f(x) dx.$$

Математическим ожиданием непрерывной случайной величины X , возможные значения которой принадлежат отрезку $[a, b]$ называется определенный интеграл:

$$M[X] = \int_a^b x f(x) dx. \quad (4.2)$$

Если плотность распределения $f(x)$ определена на всей числовой оси, тогда:

$$M[X] = \int_{-\infty}^{\infty} x f(x) dx. \quad (4.3)$$

Физический смысл математического ожидания. Математическое ожидание – это абсцисса центра тяжести стержня, имеющего массу, равную единице, с плотностью $f(x)$. Равномерное распределение плотности $f(x)$ по длине стержня показано на рис. 4.1. Неравномерное распределение плотности $f(x)$ показано на рис. 4.2.

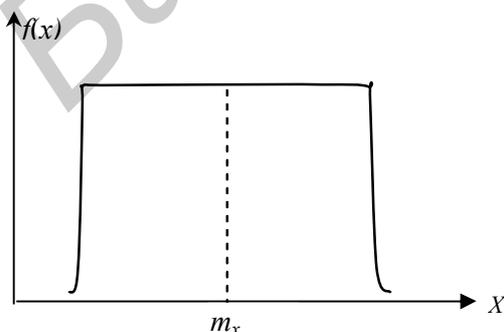


Рис. 4.1

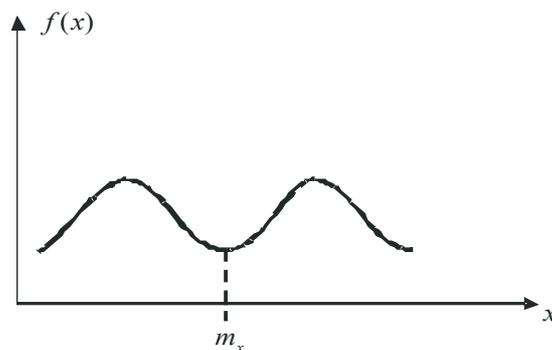


Рис. 4.2

4.1.2. Общее определение математического ожидания

При решении практических задач часто используются функции случайных величин, например: $Y = X^2$, $Y = \sin(X)$ и т. д. Очевидно, что если X случайная величина, то и Y тоже случайная величина, но с определенной функциональной зависимостью от X . Пусть задана функциональная зависимость $Y = \varphi(X)$. Требуется найти математическое ожидание случайной величины Y , зная плотность распределения случайной величины X . Для простоты предположим, что $\varphi(X)$ – монотонная непрерывная и обратная функция $X = \varphi^{-1}(Y)$ – однозначна (рис. 4.3).

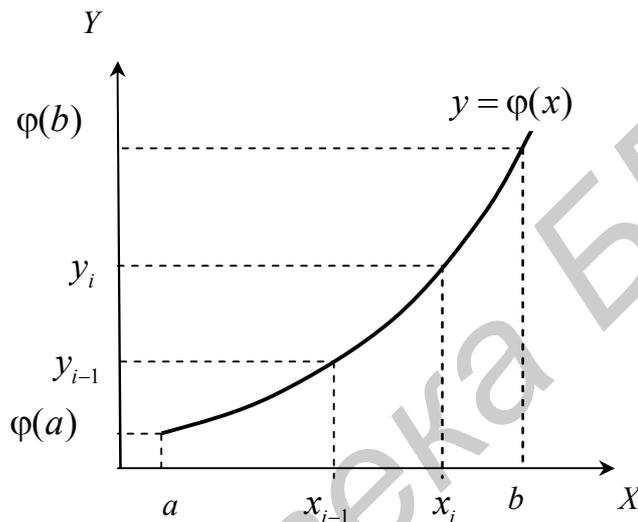


Рис. 4.3

Рассмотрим событие состоящее в том, что случайная величина X попадет в промежуток $[x_{i-1}, x_i)$: $(x_{i-1} \leq X < x_i)$, в этом случае выполнится событие, состоящее в том, что случайная величина Y попадет в промежуток $[y_{i-1}, y_i)$: $(y_{i-1} \leq Y < y_i)$. Эти события равносильны, т. е. событие $(x_{i-1} \leq X < x_i)$ влечет событие $(y_{i-1} \leq Y < y_i)$ и наоборот. Это означает, что между точками промежутка $[x_{i-1}, x_i)$ и $[y_{i-1}, y_i)$ существует взаимно однозначное соответствие.

Раз события равносильны, то вероятности этих событий одинаковы:

$$P(x_{i-1} \leq X < x_i) = P(y_{i-1} \leq Y < y_i). \quad (4.4)$$

Обозначим плотность распределения случайной величины Y как $\tilde{f}(y)$ и положим $\varphi(a) = \alpha$ и $\varphi(b) = \beta$, тогда, используя (4.2) математическое ожидание случайной величины Y , определим так:

$M[Y] = \int_{\alpha}^{\beta} y \tilde{f}(y) dy =$ | ранее мы показали, что этот интеграл является пре-

делом следующей суммы $= \lim_{\Delta y_k \rightarrow 0} \sum_{k=1}^n y_k \tilde{f}(y_k) \Delta y_k =$

$=$ | *учтем, что $P(y_{i-1} \leq Y < y_i) \approx \tilde{f}(y) \Delta y$* | $=$

$= \lim_{\Delta y_k \rightarrow 0} \sum_k y_{k-1} P(y_{k-1} \leq Y < y_k) =$ | *с учетом (4.4) перейдем к переменным X* | $=$

$= \lim_{\Delta x_k \rightarrow 0} \sum_k \varphi(x_{k-1}) P(x_{k-1} \leq X < x_k) = \sum_{k=1}^n \varphi(x_{k-1}) f(x_k) \Delta x_k = \int_a^b \varphi(x) f(x) dx.$

Таким образом, общее определение математического ожидания для непрерывной случайной величины запишется так:

$$M[\varphi(X)] = \int_a^b \varphi(x) f(x) dx. \quad (4.5)$$

Если случайная величина X определена на всей числовой оси: $a = -\infty$, $b = \infty$ и, соответственно, несобственный интеграл существует, то

$$M[\varphi(x)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx. \quad (4.6)$$

Для дискретной случайной величины X общее определение математического ожидания запишется так:

$$M[\varphi(x)] = \sum_{k=1}^n \varphi(x_k) p_k. \quad (4.7)$$

Аналогично можно получить формулы общего определения математического ожидания для двумерной случайной величины (X, Y) , если задана функциональная зависимость $Z = \varphi(X, Y)$ и известен закон распределения двумерной случайной величины (X, Y) .

Для непрерывной случайной величины

$$M[\varphi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) f(x, y) dx dy.$$

Для дискретной случайной величины

$$M[\varphi(X, Y)] = \sum_{j=1}^m \sum_{i=1}^n \varphi(x_i, y_j) p_{ij}.$$

4.1.3. Свойства математического ожидания

1. Если C не случайная величина (т. е. постоянная), то математическое ожидание от C равно самой величине C :

$$M[C] = C.$$

Доказательство:

Неслучайную величину C можно рассматривать как дискретную случайную величину с единственным значением C , которому соответствует вероятность $P = 1$. Тогда по определению математического ожидания (4.1)

$$M[C] = 1 * C = C.$$

2. *Свойство аддитивности.* Математическое ожидание суммы случайных величин равно сумме математических ожиданий:

$$M[X + Y] = M[X] + M[Y].$$

Доказательство:

Из определения математического ожидания имеем

$$\begin{aligned} M[X + Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \\ &= \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x, y) dy \right) dx + \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f(x, y) dx \right) dy = \\ &= \left| \begin{array}{l} \text{с учетом свойства 4 двумерной плотности,} \\ \text{формулы (3.5), получим} \end{array} \right| = \int_{-\infty}^{\infty} x f_1(x) dx + \\ &+ \int_{-\infty}^{\infty} y f_2(y) dy = M[X] + M[Y]. \end{aligned}$$

Аналогично доказывается

$$M[X - Y] = M[X] - M[Y].$$

3. Если C – постоянный множитель, то его можно выносить за знак математического ожидания:

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

Доказательство:

Воспользуемся общим определением математического ожидания для $Y = \varphi(x)$ (4.6).

$$\text{Пусть } Y = CX, \text{ тогда } M[CY] = \int_{-\infty}^{\infty} Cx f(x) dx = C \int_{-\infty}^{\infty} x f(x) dx = CM[X].$$

4. *Мультипликативное свойство.* Для независимых случайных величин X и Y математическое ожидание произведения этих величин равно произведению их математических ожиданий:

$$M[XY] = M[X] \times M[Y].$$

Доказательство:

$$\begin{aligned} M[X \cdot Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx dy = \\ &= \left| \begin{array}{l} \text{с учетом независимости случайных величин} \\ \text{формула (3.13): } f(x, y) = f_1(x) f_2(y), \text{ получим} \end{array} \right| = \int_{-\infty}^{\infty} x f_1(x) dx \times \\ &\times \int_{-\infty}^{\infty} y f_2(y) dy = M[X] \times M[Y]. \end{aligned}$$

5. Величина $X - M[X]$ называется *центрированной случайной величиной*, обозначим ее $\overset{0}{X}$. Тогда справедливо следующее соотношение

$$M[X - M[X]] = 0.$$

Доказательство:

$$M[X - M[X]] = M[X] - M[M[X]] = M[X] - M[X] = 0.$$

Кроме математического ожидания, которое является основной характеристикой положения случайных величин на числовой оси, на практике применяют и другие числовые характеристики, описывающие центр распределения случайной величины, в частности моду, медиану.

Модой дискретной случайной величины X называют ее наиболее вероятное значение (рис. 4.4). **Модой непрерывной** случайной величины называется такое ее значение x_{mod} , при котором плотность распределения имеет максимум (рис. 4.5).

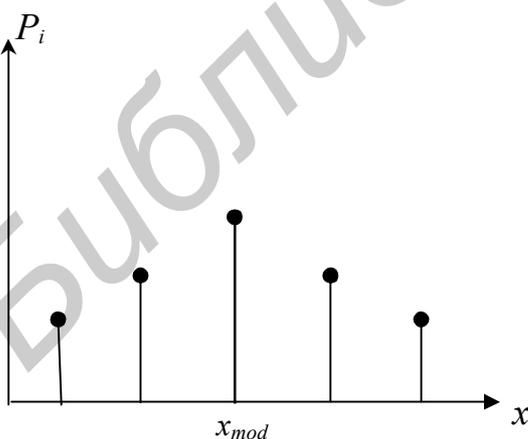


Рис. 4.4

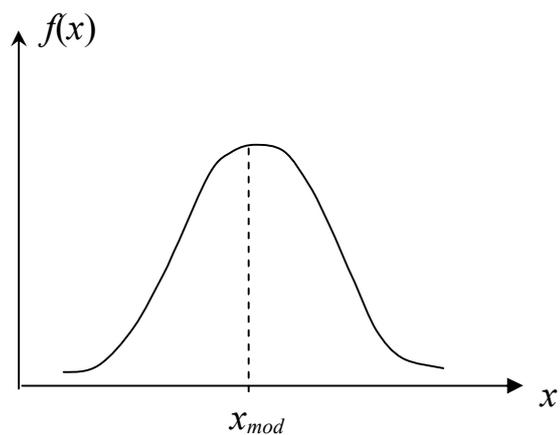


Рис. 4.5

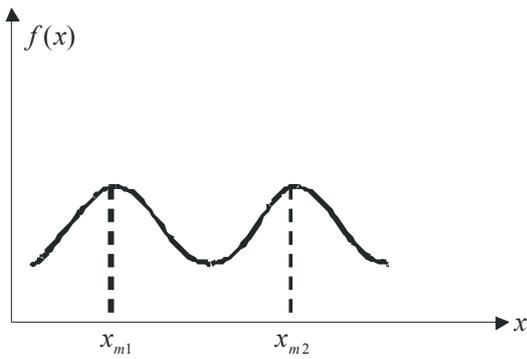


Рис. 4.6

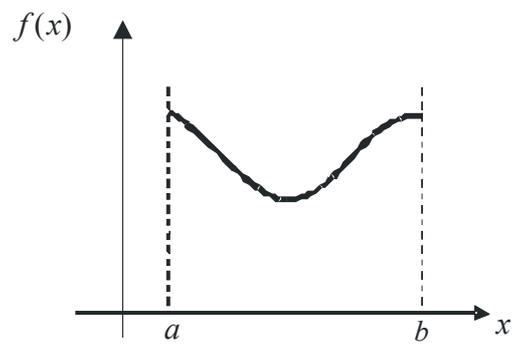


Рис. 4.7

Распределения случайных величин бывают одномодальные (рис. 4.5), двумодальные (рис. 4.6), полимодальные и антимодальные – моды не имеют (рис. 4.7).

Для определения моды необходимо продифференцировать плотность распределения – $f'(x)$, приравнять к нулю производную: $f'(x) = 0$, и решить это уравнение относительно x .

Медианой x_{med} произвольной случайной величины X называют такое ее значение, относительно которого равновероятно получение большего или меньшего значения случайной величины:

$$P(X < x_{med}) = P(X > x_{med}) = 0,5.$$

С геометрической точки зрения медиана непрерывной случайной величины – это абсцисса точки, в которой площадь фигуры, ограниченной сверху кривой распределения, делится пополам $p_1 = p_2 = 0,5$ (рис. 4.8).

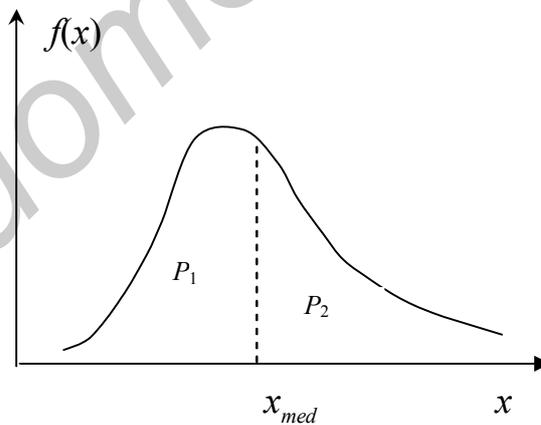


Рис. 4.8

Если распределение одномодальное и симметрическое, то все три характеристики положения случайной величины: математическое ожидание, мода и медиана – совпадают.

Квантилем порядка p называется значение случайной величины x_p , определяемое из уравнения

$$F(x_p) = P. \quad (4.8)$$

Очевидно, что квантиль порядка 0,5 – это медиана, которая определяется из уравнения $F(x_{med}) = 0,5$.

4.1.4. Дисперсия, среднее квадратическое отклонение

При проведении вероятностных экспериментов можно заметить, что значения случайных величин всегда колеблются около среднего значения. Это явление называется *рассеиванием случайной величины* около ее среднего значения. Числовые характеристики, описывающие рассеивание случайной величины, показывают, насколько тесно сгруппированы возможные значения случайной величины около ее среднего значения – математического ожидания.

Основными из них являются *дисперсия* и *среднее квадратическое отклонение*.

Для определения этих характеристик используется разность между случайной величиной X и ее математическим ожиданием. Ранее в свойстве 5 математического ожидания мы дали определение, в котором центрированная случайная величина обозначается $\overset{0}{X}$ и определяется как $X - M[X]$. В свойстве 5 мы показали, что $M[\overset{0}{X}] = 0$, т. е. математическое ожидание центрированной случайной величины никак не характеризует ее рассеивание, а указывает лишь на то, что значения отклонения – числа разного знака. Поэтому в качестве меры рассеивания случайной величины берут математическое ожидание квадрата центрированной случайной величины.

Дисперсией случайной величины X называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания и обозначается

$$D[X] = M[(X - M[X])^2]. \quad (4.9)$$

Используя общее определение математического ожидания (4.6), легко записать дисперсию для непрерывной случайной величины, если обозначить $\varphi(x) = (X - m_x)^2$:

$$D[X] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx. \quad (4.10)$$

Аналогично, применив формулу (4.7), дисперсия для дискретной случайной величины

$$D[X] = \sum_{i=1}^n (x_i - m_x)^2 p_i. \quad (4.11)$$

Подчеркнем еще раз смысл дисперсии: *дисперсия – мера рассеивания (или отклонения) случайной величины относительно ее математического ожидания.*

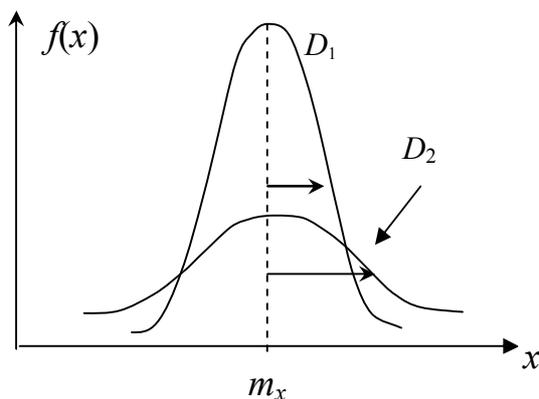


Рис. 4.9

На рис. 4.9 показаны два варианта распределения для двух разных случайных величин. Видим, что у более плоской кривой отклонение от среднего значения m_x (можно взять по среднему уровню кривой) больше, чем у острокопечной кривой. Значит, дисперсия D_2 будет больше дисперсии D_1 .

Использование дисперсии хотя и удобно для характеристики рассеивания случайной величины X , но лишено наглядности, т. к. имеет размерность квадрата случайной величины. Поэтому на практике для рассеивания (или отклонения) чаще используют характеристику, совпадающую по размерности со случайной величиной, – это среднее квадратическое отклонение.

Средним квадратическим отклонением называется положительное значение квадратного корня из дисперсии случайной величины X , которое обозначается как σ_x :

$$\sigma_x = \sqrt{D[X]}. \quad (4.12)$$

Свойства дисперсии

1. *Дисперсия постоянной величины ($C = \text{const}$) равна нулю:*

$$D[C] = 0.$$

Доказательство:

Так как $M[C] = C$ (свойство 1 математического ожидания) из определения дисперсии (4.9), подставляя $X = C$, получаем

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

2. *Постоянный множитель выносится за знак дисперсии и возводится в квадрат:*

$$D[CX] = C^2 D[X].$$

Доказательство:

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

3. Дисперсия случайной величины равна математическому ожиданию квадрата случайной величины минус квадрат ее математического ожидания:

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

Доказательство:

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

4. Для **независимых** случайных величин дисперсия суммы или разности случайной величины равна сумме дисперсии:

$$D[X \pm Y] = D[X] + D[Y].$$

Доказательство:

$$\begin{aligned} D[X \pm Y] &= M[((X \pm Y) - M[X \pm Y])^2] = M[((X - M[X]) \pm (Y - M[Y]))^2] = \\ &= \left| \text{раскроем квадрат двух величин, стоящих в круглых скобках} \right| = \\ &= M[(X - M[X])^2 \pm 2(X - M[X])(Y - M[Y]) + (Y - M[Y])^2] = \\ &= \left| \text{используя свойства математического ожидания,} \right. \\ &\quad \left. \text{найдем математическое ожидание от каждого слагаемого} \right| = \\ &= M[(X - M[X])^2] \pm 2M[(X - M[X])(Y - M[Y])] + M[(Y - M[Y])^2] = \\ &= D[X] + D[Y] \pm 2M[(X - M[X])(Y - M[Y])]. \end{aligned}$$

Если случайной величины независимы, то

$$M[(X - M[X])(Y - M[Y])] = M[X - M[X]]M[Y - M[Y]] = 0.$$

Окончательно получаем:

$$D[X \pm Y] = D[X] + D[Y].$$

5. Дисперсия произведения **независимых** случайных величин

$$D[X \cdot Y] = D[X] \cdot D[Y] + m_x^2 D[Y] + m_y^2 D[X]. \quad (4.13)$$

Доказательство:

Обозначим $X \cdot Y = Z$.

По определению дисперсии

$$D[X \cdot Y] = D[Z] = M[Z^2] = M[(Z - m_z)^2]. \quad (4.14)$$

Поскольку X и Y независимые – $m_z = m_x \cdot m_y$,

$$\begin{aligned} D[X \cdot Y] &= M[(X \cdot Y - m_x m_y)^2] = \\ &= M[X^2 \cdot Y^2] - 2m_x m_y M[X \cdot Y] + m_x^2 m_y^2. \end{aligned} \quad (4.15)$$

Если X и Y независимые, то X^2, Y^2 тоже независимы, тогда

$$M[X^2 \cdot Y^2] = M[X^2] \cdot M[Y^2], \quad (4.16)$$

$$M[X \cdot Y] = m_x \cdot m_y.$$

Подставим (4.16) в выражение (4.15) и получим

$$D[X \cdot Y] = M[X^2] \cdot M[Y^2] - m_x^2 \cdot m_y^2. \quad (4.17)$$

Найдем $M[X^2]$, используя свойство 3 дисперсии:

$$D[X] = M[X^2] - m_x^2.$$

Отсюда получим

$$M[X^2] = D[X] + m_x^2. \quad (4.18)$$

Аналогично для случайной величины Y

$$M[Y^2] = D[Y] + m_y^2. \quad (4.19)$$

Подставляя (4.18, 4.19) в (4.17), получим (4.13):

$$\begin{aligned} D[X \cdot Y] &= (D[X] + m_x^2)(D[Y] + m_y^2) - m_x^2 \cdot m_y^2 = \\ &= D[X] \cdot D[Y] + m_x^2 D[Y] + m_y^2 D[X]. \end{aligned}$$

4.1.5. Моменты распределения случайных величин

Рассмотренные ранее числовые характеристики случайных величин являются частным случаем более общего понятия числовых характеристик – моментов распределений случайных величин. Здесь название «момент» взято из механики, где оно употребляется для описания распределений масс тел, а в теории вероятностей употребляется для описания распределений случайных величин.

В теории вероятностей моменты бывают двух видов: начальные и центральные.

Начальным моментом k -го порядка случайной величины X называется математическое ожидание случайных величин X^k .

Обозначать начальный момент k -го порядка будем α_k и определим так:

$$\alpha_k = M[X^k]. \quad (4.20)$$

Используя общее определение математического ожидания (4.6), легко записать начальный момент k -го порядка для непрерывной случайной величины, если обозначить $f(x) = X^k$:

$$\alpha_k = \int_{-\infty}^{\infty} x^k f(x) dx. \quad (4.21)$$

Аналогично, применяя формулу (4.7), начальный момент k -го порядка для дискретной случайной величины запишется так:

$$\alpha_k = \sum_{i=1}^n x_i^k P_i. \quad (4.22)$$

Особо важное значение имеет начальный момент первого порядка. Видим, что при $k = 1$ $\alpha_1 = M[X]$ и представляет математическое ожидание. Начальные моменты высших порядков используются для вычисления центральных моментов.

Центральным моментом k -го порядка случайной величины X называется математическое ожидание величины $(X - m_x)^k$.

Обозначать центральный момент k -го порядка будем μ_k и определять следующим образом:

$$\mu_k = M[(X - m_x)^k]. \quad (4.23)$$

При использовании общего определения математического ожидания (4.6), если обозначить $f(x) = (X - m_x)^k$, центральный момент k -го порядка для непрерывной случайной величины запишется так:

$$\mu_k = \int_{-\infty}^{\infty} (x_i - m_x)^k f(x) dx. \quad (4.24)$$

Применяя формулу (4.7), центральный момент k -го порядка для дискретной случайной величины рассчитывается так:

$$\mu_k = \sum_{i=1}^n (x_i - m_x)^k P_i. \quad (4.25)$$

С учетом свойства 5 математического ожидания отметим, что центральный момент первого порядка всегда равен нулю. Центральный момент второго порядка является дисперсией случайной величины X и через 1-й и 2-й начальные моменты (используем свойство 3 дисперсии) определяется так:

$$\mu_2 = D[X] = M[X^2] - (M[X])^2 = \alpha_2 - \alpha_1^2.$$

Третий центральный момент μ_3 служит характеристикой асимметрии («скошенности») распределения. Если случайная величина X распределена симметрично относительно своего математического ожидания, то $\mu_3 = M[(X - m_x)^3] = 0$. Третий центральный момент имеет размерность куба случайной величины, поэтому при решении практических задач для характеристики асимметрии чаще пользуются безразмерной величиной – *коэффициентом асимметрии*: $A = \mu_3 / \sigma^3$. На рис. 4.10 и 4.11 показана положительная и отрицательная асимметрия соответственно.

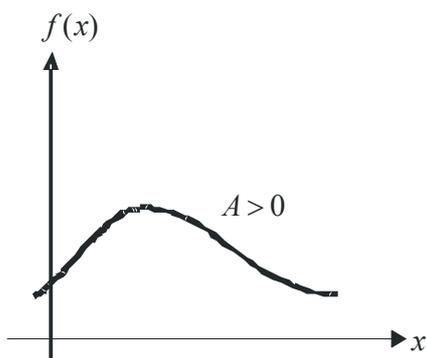


Рис. 4.10

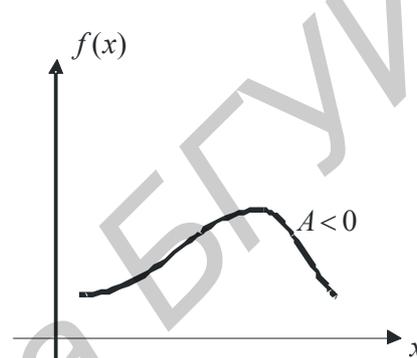


Рис. 4.11

Четвертый центральный момент μ_4 используется для характеристики *плосковершинности*. Введя безразмерный коэффициент плосковершинности $B = \mu_4 / \sigma^4$, можно показать, что для нормального закона распределения $B = \mu_4 / \sigma^4 = 3$. Тогда в качестве характеристики плосковершинности применяют безразмерный коэффициент, называемый *эксцессом*: $\mathcal{E} = \mu_4 / \sigma^4 - 3$.

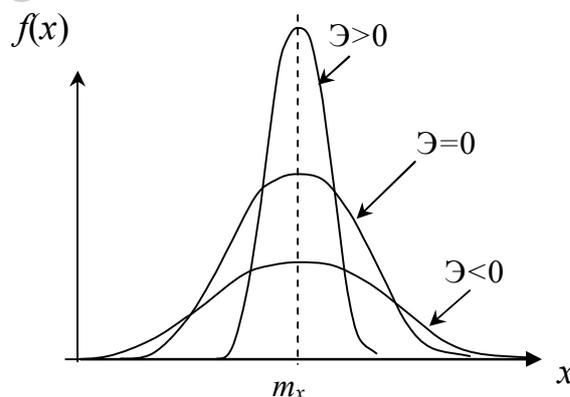


Рис. 4.12

**4.1.6. Числовые характеристики двумерной случайной величины.
Ковариация. Коэффициент корреляции**

Основной характеристикой, описывающей связь между составляющими X и Y двумерной случайной величины (X, Y) , является *ковариация*, называемая также *корреляционным моментом* K_{xy} или *моментом связи* K_{xy} , и определяется по формуле

$$\begin{aligned} K_{xy} &= \text{cov}(X, Y) = M[(X - m_x)(Y - m_y)] = \\ &= M[XY] - M[Xm_y] - M[Ym_x] + M[m_x m_y] = \\ &= M[XY] - 2M[X]M[Y] + M[X]M[Y] = \\ &= M[XY] - M[X]M[Y], \end{aligned} \quad (4.26)$$

где математическое ожидание составляющей X для непрерывной двумерной случайной величины (X, Y) вычисляется так:

$$m_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dx dy = \int_{-\infty}^{\infty} xf_1(x) dx. \quad (4.27)$$

Для составляющей X дискретной двумерной случайной величины (X, Y)

$$m_x = \sum x_i P_i = \sum_i \sum_j x_i P_{ij}. \quad (4.28)$$

Аналогично записываются математические ожидания для составляющей Y .

Отметим, что из определения корреляционного момента (4.26) следует, что корреляционный момент показывает рассеивание (отклонение) случайных величин X, Y относительно среднего значения и одновременно характеризует влияние одной случайной величины на другую.

Для дискретной двумерной случайной величины (X, Y) корреляционный момент запишется так:

$$K_{xy} = \sum_i \sum_j (x_i - m_x)(y_j - m_y) P_{ij} = \sum_{i,j} x_i y_j P_{ij} - m_x m_y. \quad (4.29)$$

Корреляционный момент для непрерывной двумерной случайной величины (X, Y) имеет следующий вид:

$$K_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx dy - m_x m_y. \quad (4.30)$$

Покажем, что если составляющие X и Y двумерной случайной величины (X, Y) независимы, то $K_{xy} = 0$. *Необходимое и достаточное условие независимости* случайных величин X и Y определено ранее (3.13):

$$f(x, y) = f_1(x)f_2(y),$$

тогда (4.24) принимает следующий вид:

$$K_{X,Y} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x,y) dx dy = \int_{-\infty}^{\infty} (x - m_x) f_1(x) dx \int_{-\infty}^{\infty} (y - m_y) f_2(y) dy =$$

$$= \int_{-\infty}^{\infty} (x - m_x) f(x) dx = \int_{-\infty}^{\infty} x f(x) dx - m_x \int_{-\infty}^{\infty} f(x) dx = m_x - m_x = 0.$$

Обратное утверждение, если $K_{X,Y} = 0$, то случайные величины X и Y независимы, не всегда выполняется.

Корреляционный момент имеет размерность квадрата случайной величины, но на практике для характеристики связи чаще используют безразмерные величины. Поэтому составляющие нормируют. Если положить $\bar{X} = (X - m_x) / \sigma_x$, $\bar{Y} = (Y - m_y) / \sigma_y$, то корреляционный момент нормированных случайных величин X и Y называется **коэффициентом корреляции** и обозначается

$$\rho_{xy} = M[\bar{X}\bar{Y}] = M\left[\left(\frac{X - m_x}{\sigma_x}\right)\left(\frac{Y - m_y}{\sigma_y}\right)\right] = \frac{m_{xy} - m_x m_y}{\sigma_x \sigma_y} = \frac{K_{xy}}{\sigma_x \sigma_y}. \quad (4.31)$$

Коэффициент корреляции является безразмерной величиной и характеризует степень линейной связи между составляющими X и Y двумерной случайной величины (X, Y) .

Свойства коэффициента корреляции

1. Абсолютная величина коэффициента корреляции не превосходит единицу $|\rho_{xy}| \leq 1$; при доказательстве свойства 4 дисперсии мы получили

$$0 \leq D[\bar{X} \pm \bar{Y}] = D[\bar{X}] + D[\bar{Y}] \pm 2 \text{cov}[\bar{X}\bar{Y}] =$$

$$= \left| \text{но: } D[\bar{X}] = D\left[\frac{X - m_x}{\sigma_x}\right] = \frac{1}{\sigma_x^2} M[(X - m_x)^2] = 1 \right| =$$

$$= 2(1 \pm \rho_{xy}) \Rightarrow -1 \leq \rho_{xy} \leq 1.$$

2. В зависимости от значения коэффициента корреляции ρ_{xy} между составляющими X и Y двумерной случайной величины (X, Y) могут существовать следующие зависимости (рис. 4.13, $a - d$):

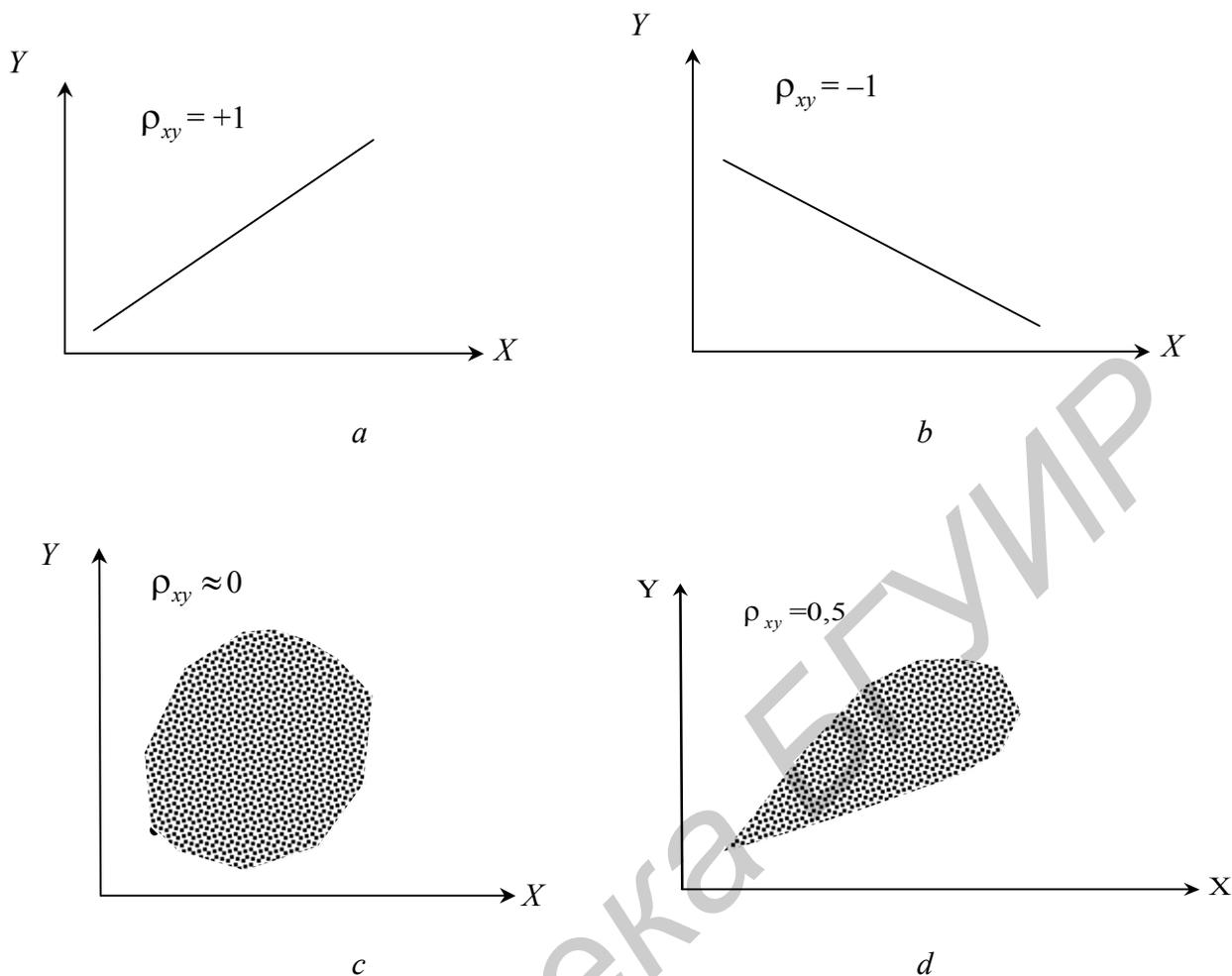


Рис. 4.13

a – при $\rho_{xy} = 1$ между X и Y существует линейная положительная зависимость (рис. 4.13, *a*);

b – при $\rho_{xy} = -1$ между X и Y существует линейная отрицательная зависимость (рис. 4.13, *b*);

c – при $\rho_{xy} = 0$ между X и Y отсутствует зависимость, и говорят, что X и Y некоррелированы (рис. 4.13, *c*);

d – когда $0 < \rho_{xy} < 1$, то между X и Y существует положительная вероятностная линейная зависимость, или между X и Y существует положительная корреляция (рис. 4.13, *d*).

ГЛАВА 5. ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

При решении многих практических задач приходится рассматривать функции одной или нескольких случайных величин: ($Y = X^2$, $Y = X^3 + \cos X$, $Z = X^3 + Y^3 \dots$). Такие функции тоже являются случайными величинами, но с известной функциональной зависимостью. Задача может быть сформулирована следующим образом.

Дана многомерная случайная величина (X_1, X_2, \dots, X_n) , закон распределения которой известен. Рассматривается некоторая случайная величина Y как функция от случайных величин X_1, X_2, \dots, X_n :

$$Y = \varphi(X_1, X_2, \dots, X_n).$$

Требуется найти закон распределения случайных величин Y , зная вид функции φ и закон распределения случайных величин (X_1, X_2, \dots, X_n) .

5.1. Одномерное приближение

Сформулируем задачу для *дискретных случайных величин*. Пусть задана функциональная зависимость $Y = \varphi(X)$. Закон распределения дискретной случайной величины X задается в виде ряда распределения:

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Требуется найти закон распределения случайной величины Y .

Тогда $Y = \varphi(X)$ также дискретная случайная величина с возможными значениями $y_i = \varphi(x_i)$. Если значения y_1, \dots, y_n различны, то для каждого $i = \overline{1, n}$ события $\{X = x_i\}$ и $\{Y = y_i = \varphi(x_i)\}$ равносильны. То есть в результате опыта случайная величина Y примет значение $y_i = \varphi(x_i)$ тогда и только тогда, когда случайная величина X примет значение x_i , а вероятность последнего события $\{X = x_i\}$ равна p_i . Следовательно, вероятности этих событий одинаковы:

$$P(\{Y = y_i = \varphi(x_i)\}) = P(\{X = x_i\}) = p_i.$$

Тогда искомый ряд распределения запишется в следующем виде

X	$y_1 = \varphi(x_1)$	$y_2 = \varphi(x_2)$...	$y_n = \varphi(x_n)$
P	p_1	p_2	...	p_n

Для непрерывной случайной величины задача ставится так: зная плотность распределения $f(x)$ случайной величины X и функциональную зависимость $Y = \varphi(X)$, найти плотность распределения $f_y(y)$ случайной величины Y .

Пусть X – непрерывная случайная величина с плотностью распределения $f(x)$, а $Y = \varphi(X)$ – непрерывная и дифференцируемая случайная величина. Рассмотрим общий случай, когда обратная функция $X = \varphi^{-1}(Y)$ – неоднозначна, т. е. одному значению y соответствует несколько значений $x_i = \varphi_i^{-1}(y)$, $i = \overline{1, n}$.

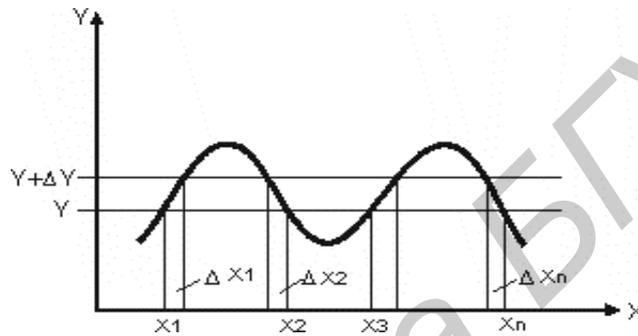


Рис. 5.1

Рассмотрим малое приращение функции Δy . Событие $\{Y \in (y, y + \Delta y)\}$ произойдет тогда, когда случайная величина X попадет в интервал или Δx_1 , или Δx_2 , ..., т. е. событие равносильно сумме несовместных событий $\{X \in (x_i, x_i \pm \Delta x_i)\}$, $i = \overline{1, n}$, где знак $\langle + \rangle$ или $\langle - \rangle$ берется в зависимости от того, возрастает или убывает φ_i^{-1} в интервале $(y, y + \Delta y)$ (рис. 5.1):

$$\{y < Y < y + \Delta y\} = \sum_{i=1}^n \{x_i < X < x_i \pm \Delta x_i\}. \quad (5.1)$$

Раз события, записанные в (5.1), тождественны, то и вероятности их равны. Применяя 3 аксиому к событиям в правой части (5.1) (несовместные события), получаем

$$P(\{y < Y < y + \Delta y\}) = \sum_{i=1}^n P(\{x_i < X < x_i \pm \Delta x_i\}). \quad (5.2)$$

Учитывая, что вероятность попадания случайной точки в малый интервал приближенно равна произведению плотности распределения в какой-либо точке этого интервала на его длину, найдем

$$f_y(y)\Delta y \approx \sum_{i=1}^n f(x_i)\Delta x_i. \quad (5.3)$$

Разделив (5.3) на Δy , переходя к пределу при $\Delta y \rightarrow 0$ и подставив $\varphi_i^{-1}(y)$ вместо x_i , получим

$$f_y(y) = \sum_{i=1}^n f[\varphi_i^{-1}(y)] \left| \frac{d\varphi_i^{-1}(y)}{dy} \right|. \quad (5.4)$$

Здесь производную берем по модулю потому, что плотность распределения не может быть отрицательной.

Пример.

Пусть X – нормированная нормальная случайная величина с плотностью распределения $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ и задана функциональная зависимость $Y = \sqrt[3]{|X|}$. Найти $f_y(y)$.

Строим график функции $Y = \sqrt[3]{|X|}$ (рис. 5.2).

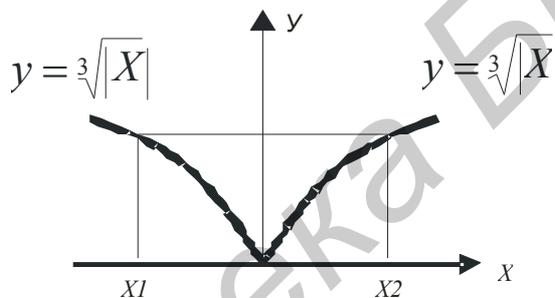


Рис. 5.2

Видим, что обратная функция $x = \varphi^{-1}(y) = \pm y^3$ – двузначная, $n = 2$:

$$x_1 = \varphi_1^{-1}(y) = y^3, \quad x_2 = \varphi_2^{-1}(y) = -y^3.$$

Подставляем эти значения в (5.4):

$$\begin{aligned} f_y(y) &= f(y^3) \cdot \left| \frac{dy^3}{dy} \right| + f(-y^3) \cdot \left| \frac{-dy^3}{dy} \right| = \frac{1}{\sqrt{2\pi}} \cdot e^{-y^6/2} \cdot 3y^2 + \frac{1}{\sqrt{2\pi}} \cdot e^{-y^6/2} \cdot 3y^2 = \\ &= \frac{6y^2}{\sqrt{2\pi}} \cdot e^{-y^6/2}. \end{aligned}$$

5.2. Двумерное приближение

Рассмотрим двумерную случайную величину (X_1, X_2) с плотностью распределения $f(x_1, x_2)$. Пусть задана для двумерной случайной величины (Y_1, Y_2) функциональная зависимость в следующем виде:

$$Y_1 = \varphi_1(X_1, X_2), \quad Y_2 = \varphi_2(X_1, X_2). \quad (5.5)$$

Требуется найти плотность распределения $f_y(y_1, y_2)$ двумерной случайной величины (Y_1, Y_2) .

Найдем обратные функции:

$$X_1 = \varphi_1^{-1}(Y_1, Y_2), \quad X_2 = \varphi_2^{-1}(Y_1, Y_2). \quad (5.6)$$

Пусть эти функции – многозначные или n -значные. То есть попадание случайной точки (Y_1, Y_2) в элементарную область ΔG на плоскости (Y_1, Y_2) равносильно попаданию случайной точки (X_1, X_2) или в область ΔG_1 , или в область $\Delta G_2 \dots$, или в область ΔG_n , где $(\Delta G_i, i = \overline{1, n})$ – непересекающаяся область на плоскости $X_1 X_2$ (рис. 5.3).

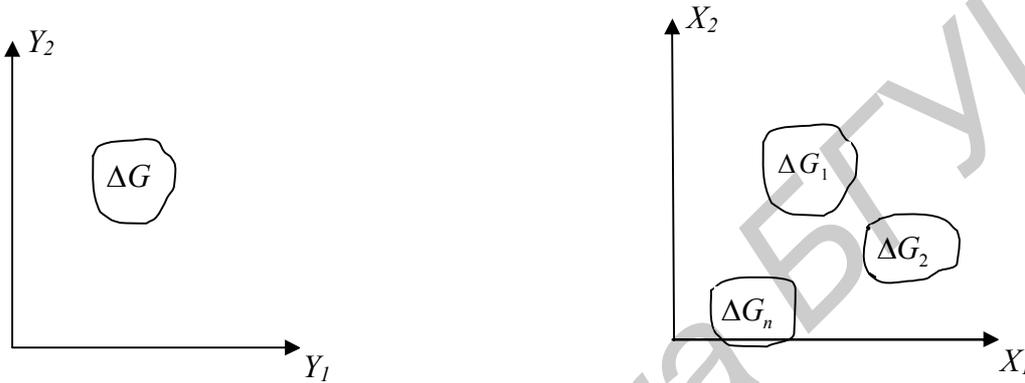


Рис. 5.3

Обозначим $\{(Y_1, Y_2) \in \Delta G\}$ – событие состоящее в том, что двумерная случайная величина (Y_1, Y_2) попадет в область ΔG . Это событие будет *равносильно* сумме *несовместных* событий, состоящих в том, что двумерная случайная величина (X_1, X_2) попадет *или* в область ΔG_1 , *или* в область $\Delta G_2 \dots$, *или* в область ΔG_n , т. е. можно записать, что

$$\{(Y_1, Y_2) \in \Delta G\} = \sum_{i=1}^n \{(X_1, X_2) \in \Delta G_i\}. \quad (5.7)$$

Так как события, записанные в (5.7), равносильны, то и вероятности этих событий одинаковы. Применяя к правой части (5.7) третью аксиому, получаем

$$P(\{(Y_1, Y_2) \in \Delta G\}) = \sum_{i=1}^n P(\{(X_1, X_2) \in \Delta G_i\}). \quad (5.8)$$

Если площадки ΔG_i – достаточно малы, то вероятность попадания двумерной случайной величины (X_1, X_2) на площадку ΔG_i можно приближенно вычислить, используя определение двумерной плотности (см. раздел 3.4, формула (3.2)). Тогда формула (5.8) запишется в следующем виде:

$$f_y(y_1 y_2) \Delta G \approx \sum_{i=1}^n f(x_{1i}, x_{2i}) \Delta G_i, \quad (5.9)$$

$$\begin{aligned} \text{где } x_{1i} &= \varphi_{1i}^{-1}(y_1, y_2), \\ x_{2i} &= \varphi_{2i}^{-1}(y_1, y_2), \\ i &= \overline{1, n} \end{aligned} \quad (5.10)$$

представляют собой соответствующие n ветвей преобразования (5.6).

Но, как известно, отношение элементарных площадей $\frac{\Delta G_i}{\Delta G}$ в пределе при преобразовании переменных равно соответствующему Якобиану (определителю 2-го порядка из частных производных). Следовательно, разделив (5.9) на ΔG и перейдя к пределу при $\Delta G \rightarrow 0$, получим

$$f_y(y_1 y_2) = \sum_{i=1}^n f(x_{1i}, x_{2i}) \left| \frac{\partial(x_{1i}, x_{2i})}{\partial(y_1 y_2)} \right|, \quad (5.11)$$

где величины x_{1i} и x_{2i} определяются из (5.10), а Якобиан есть определитель 2-го порядка:

$$\frac{\partial(x_{1i}, x_{2i})}{\partial(y_1 y_2)} = \begin{vmatrix} \frac{\partial x_{1i}}{\partial y_1} & \frac{\partial x_{1i}}{\partial y_2} \\ \frac{\partial x_{2i}}{\partial y_1} & \frac{\partial x_{2i}}{\partial y_2} \end{vmatrix}. \quad (5.12)$$

Если обратные функции (5.6) – однозначны, то в (5.11) – только одно слагаемое.

При решении практических задач часто используется следующая функциональная зависимость: $Y = \varphi(X_1 X_2)$. Тогда перейти к формулам (5.6) – (5.12) можно следующим образом, введя фиктивную переменную Y_1 :

$$\begin{aligned} Y_1 &= X_1, \\ Y &= \varphi(X_1 X_2). \end{aligned} \quad (5.13)$$

Положим для простоты, что обратная функция φ^{-1} – однозначна ($n = 1$), тогда обратные функции запишутся так:

$$\begin{aligned} X_1 &= Y_1, \\ X_2 &= \varphi^{-1}(Y_1, Y). \end{aligned} \quad (5.14)$$

Якобиан принимает следующий вид:

$$\frac{\partial(x_1, x_2)}{\partial(y_1 y)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ \frac{\partial \varphi^{-1}(y_1, y)}{\partial y_1} & \frac{\partial \varphi^{-1}(y_1, y)}{\partial y} \end{vmatrix} = \frac{\partial \varphi^{-1}(y_1, y)}{\partial y}.$$

Подставляем эти выражения в (5.11):

$$f_y(y_1, y) = f(y_1, \varphi^{-1}(y_1, y)) \left| \frac{\partial \varphi^{-1}(y_1, y)}{\partial y} \right|.$$

Интегрируя это выражение по y_1 (и заменяя y_1 на x_1), получаем

$$f_y(y) = \int_{-\infty}^{\infty} f_y(y_1, y) dy_1 = \int_{-\infty}^{\infty} f_y(x_1, \varphi^{-1}(x_1, y)) \left| \frac{\partial \varphi^{-1}(x_1, y)}{\partial y} \right| dx_1. \quad (5.15)$$

5.3. Распределение суммы независимых случайных величин

Пусть для независимых случайных величин X_1 и X_2 известны плотности распределений $f_1(x_1)$ и $f_2(x_2)$. Задана функциональная зависимость $Y = X_1 + X_2$. Требуется найти плотность распределения случайной величины Y . Эта задача рассматривается в качестве примера на использование формулы (5.15).

Запишем случайные величины вида (5.13):

$$Y_1 = X_1, \quad Y = X_1 + X_2.$$

Определим обратные функции:

$$X_1 = Y_1, \quad X_2 = \varphi^{-1}(x_1, y) = y - x_1.$$

Найдем $\frac{\partial \varphi^{-1}}{\partial y} = \frac{\partial (y - x_1)}{\partial y} = 1$.

Подставляя в формулу (5.15), получим

$$f_y(y) = \int_{-\infty}^{\infty} f_y(x_1, y - x_1) dx_1 = \left| \text{т. к. } X_1 \text{ и } X_2 \text{ независимы, то} \right.$$

$$f_y(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) \left. \right| = \int_{-\infty}^{\infty} f_1(x_1) \cdot f_2(y - x_1) dx_1. \quad (5.16)$$

Распределение суммы независимых случайных величин X_1 и X_2 называется *композицией распределений этих величин*.

5.4. Характеристические функции

Характеристической функцией случайной величины X называется математическое ожидание случайной величины e^{jtx} :

$$g(t) = M[e^{jtx}], \quad (5.17)$$

где $j = \sqrt{-1}$, t – вещественная переменная, x – случайная величина.

Используя общее определение математического ожидания, запишем характеристическую функцию для непрерывной случайной величины, если задана плотность распределения $f(x)$:

$$g(t) = \int_{-\infty}^{\infty} e^{jtx} f(x) dx. \quad (5.18)$$

Если для дискретной случайной величины X заданы вероятности возможных ее значений: $P(X = x_k) = p_k$, $k = \overline{1, n}$, то характеристическая функция принимает вид:

$$g(t) = \sum_{k=1}^n e^{jtx_k} p_k. \quad (5.19)$$

Характеристическая функция существует для любой случайной величины, поскольку ввиду равенства $|e^{jtx}| = 1$ ряд (5.19) и интеграл (5.18) сходятся абсолютно.

Свойства характеристической функции

1. $g(0) = 1$, например из (5.17) при $t = 0$; $g(0) = \int_{-\infty}^{\infty} f(x) dx = 1$.

2. $|g(t)| \leq 1$.

3. Если известна характеристическая функция $g(t)$, то плотность распределения случайной величины X находится с помощью обратного преобразования Фурье:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} g(t) dt. \quad (5.20)$$

4. Если $Y = ax + b$, где a и b – const, то $g_y(t) = e^{itb} g(at)$.

Покажем это, исходя из определения, формула (5.17), получим

$$g_y(t) = M[e^{it(ax+b)}] = e^{itb} M[e^{itax}] = e^{itb} g(at).$$

5. Если $Y = X_1 + X_2 + \dots + X_n$ – независимы, то $g_y(t) = \prod_{k=1}^n g_k(t)$.

Из независимости X_i следует независимость e^{jtx_i} . Применяя свойство мультипликативности математического ожидания, получим

$$g_y(t) = M[e^{jt \sum x_i}] = M[\prod_{i=1}^n e^{jtx_i}] = \prod_{i=1}^n M[e^{jtx_i}] = \prod_{i=1}^n g_i(t).$$

6. Пусть у случайной величины X существует начальный момент k -го порядка α_k . Тогда, зная характеристическую функцию, можно найти начальный момент k -го порядка:

$$\alpha_k = \frac{g^{(k)}(0)}{i^k}, \quad (5.21)$$

где $g^{(k)}(0)$ – k -я производная при $t = 0$.

Доказательство:

Дифференцируя (5.18) по t k раз, получим:
первую производную

$$\frac{\partial}{\partial t} g(t) = \int_{-\infty}^{\infty} (ix) e^{itx} f(x) dx,$$

вторую производную

$$g''(t) = \int_{-\infty}^{\infty} (ix)^2 e^{itx} f(x) dx,$$

.....

k -ю производную

$$g^{(k)}(t) = \int_{-\infty}^{\infty} (ix)^k e^{itx} f(x) dx.$$

Полагая $t = 0$,

$$g^{(k)}(0) = i^k \int_{-\infty}^{\infty} x^k f(x) dx = i^k \alpha_k,$$

откуда получаем $\alpha_k = \frac{g^{(k)}(0)}{i^k}$.

ГЛАВА 6. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН В ТЕОРИИ ВЕРОЯТНОСТЕЙ

6.1. Биномиальный закон распределения

Пусть производится n опытов, в каждом из которых наступает событие A с одной и той же вероятностью P . Обозначим X как число наступлений события A в этих n опытах. Тогда X – дискретная случайная величина и вероятности возможных ее значений записываются с использованием формулы Бернулли:

$$P_n(X = x_k = k) = C_n^k p^k q^{n-k}.$$

Ряд распределения, в котором вероятности определяются с использованием формулы Бернулли, называется биномиальным распределением.

Закон описывает вероятность во многих задачах, таких, как число элементов радиоаппаратуры, выходящих из строя за время t ; число приборов, выходящих из строя за время t при испытаниях на надежность и др.

Ряд распределения имеет вид:

X	0	1	...	k	...	n
P	q^n	npq^{n-1}	...	$C_n^k p^k q^{n-k}$...	p^n

Для вероятностей p и q , входящих в формулу Бернулли, можно записать следующее разложение в ряд:

$$1 = (p + q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k}.$$

В правой части стоит бином Ньютона, откуда и произошло название биномиального распределения.

Пространство элементарных событий Ω здесь дискретно и состоит из $n + 1$ элементарного события: $\omega_0, \omega_1, \omega_2, \dots, \omega_n$, где ω_k – элементарное событие, состоящее в том, что при n опытах событие A наступит k раз.

На рис. 6.1 изображено распределение вероятностей биномиального распределения. На рис. 6.2 приведено изображение функции распределения $F(x)$.

Это ступенчатая функция со скачками в точках разрыва $X = k$, $F(x) = \sum_{x_k < x} P_k$.

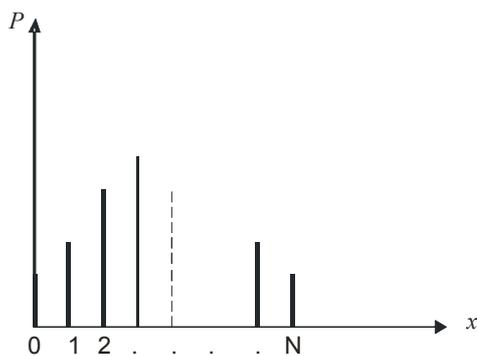


Рис. 6.1

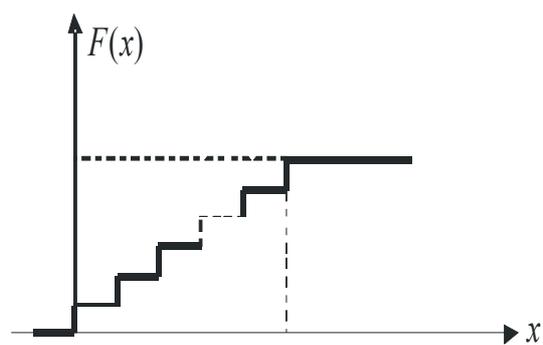


Рис. 6.2

Определим числовые характеристики случайной величины X , распределенной по биномиальному закону. Для этого используем аппарат характеристических функций. Приведем **алгоритм определения числовых характеристик**, который используется для любых законов распределения.

1. Определяем характеристическую функцию $g(t)$, применяя формулы (5.18), (5.19).

2. Находим первую $g'(t)$ и вторую $g''(t)$ производные от характеристической функции.

3. Определяем значения производных при $t = 0$:

$$g'(0), g''(0).$$

4. Используем свойство 6 характеристической функции и, применяя (5.21), определим 1-й и 2-й начальные моменты:

$$\alpha_1 = \frac{g'(0)}{i}, \quad \alpha_2 = \frac{g''(0)}{i^2}.$$

5. Определяем математическое ожидание $M[X] = \alpha_1$ и дисперсию $D[X] = \alpha_2 - \alpha_1^2$.

Найдем характеристическую функцию для биномиального распределения:

$$g(t) = \sum_{k=0}^n e^{itx_k} P_k = \sum_{k=0}^n e^{itk} \cdot C_n^k p^k q^{n-k} = \sum_{k=0}^n C_n^k (pe^{it})^k q^{n-k} = (pe^{it} + q)^n.$$

Продифференцируем ее 2 раза по t :

$$g'(t) = n(pe^{it} + q)^{n-1} \cdot pe^{it} \cdot i,$$

$$g''(t) = (n-1)n(pe^{it} + q)^{n-2} \cdot pe^{it} \cdot i \cdot pe^{it} \cdot i + n(pe^{it} + q)^{n-1} \cdot p \cdot i \cdot e^{it} \cdot i.$$

Найдем значения производных при $t = 0$:

$$g'(0) = n(p + q)^{n-1} \cdot p \cdot i = n \cdot p \cdot i,$$

$$g''(0) = (n-1)n \cdot p^2 \cdot i^2 + n \cdot p \cdot i^2 = n \cdot p \cdot i^2 [(n-1)p + 1].$$

Определяем 1-й и 2-й начальные моменты:

$$\alpha_1 = \frac{g'(0)}{i} = np,$$

$$\alpha_2 = \frac{g''(0)}{i^2} = -g''(0) = n^2 p^2 - np^2 + np = n^2 p^2 + np(1-p) = n^2 p^2 + npq.$$

Найдем числовые характеристики:

$$M[X] = \alpha_1 = np,$$

$$D[X] = \alpha_2 - \alpha_1^2 = n^2 p^2 + npq - n^2 p^2 = npq.$$

6.2. Закон распределения Пуассона

Распределение Пуассона описывает вероятности во многих задачах, таких, как число отказов радиоэлектронной аппаратуры за время t , число радиоактивных атомов, распавшихся к моменту t , число вызовов на телефонной станции и др.

Дискретная случайная величина X имеет распределение Пуассона, если ее закон распределения задан формулой

$$P_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad (6.1)$$

где $k = \overline{0, \infty}$,

λ – параметр закона Пуассона,

$\lambda = np$.

Формула Пуассона получается как *предельная* для биномиального распределения, когда число испытаний n стремится к бесконечности, а вероятность успеха p стремится к 0. Вычисления по формулам биномиального распределения приводят при больших n к очень громоздким вычислениям. Поэтому важно иметь приближенные, но достаточно простые формулы для вычисления соответствующих вероятностей.

Пусть вероятность p наступления события A при каждом испытании равна $\frac{\lambda}{n}$, т. е. $\lambda = np$. Найдем предел формулы Бернулли при $n \rightarrow \infty$:

$$\begin{aligned} \lim_{n \rightarrow \infty} P_n(k) &= \lim_{n \rightarrow \infty} \frac{n!}{(n-k)!k!} p^k q^{n-k} = \lim_{n \rightarrow \infty} \frac{n!}{(n-k)!k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left\{ \frac{n(n-1)\dots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^{-k} \right\} = \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \times \\ &\times \lim_{n \rightarrow \infty} \left\{ 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{-k} \right\} = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \cdot \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Ряд распределения дискретной случайной величины X записан с применением формулы (6.1):

X	0	1	...	k	...
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$...	$\frac{\lambda^k e^{-\lambda}}{k!}$...

Найдем $\sum_{k=0}^{\infty} P_k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$. Видим, что события $X = x_k$ составляют полную группу.

Пространство элементарных событий здесь дискретно и состоит из счетного множества элементарных событий $\omega_k : P(\omega_k) = P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$. На рис. 6.3 приведено распределение вероятностей, а на рис. 6.4 график функции распределения.

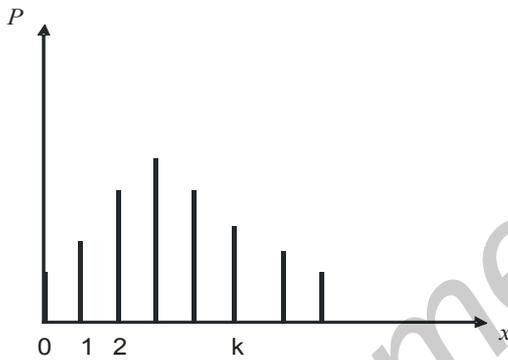


Рис. 6.3

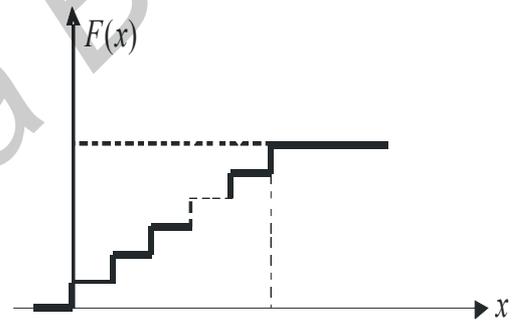


Рис. 6.4

Для определения числовых характеристик составим характеристическую функцию:

$$g(t) = \sum_{k=0}^{\infty} e^{itk} P_k = \sum_{k=0}^{\infty} e^{itk} \cdot \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it} \cdot \lambda)^k}{k!} =$$

$$= \left\| e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \right\| = e^{-\lambda} e^{\lambda \cdot \exp(it)} = e^{\lambda \cdot (\exp(it) - 1)}.$$

Продифференцируем 2 раза $g(t)$:

$$g'(t) = e^{\lambda(\exp(it)-1)} \cdot \lambda e^{it} \cdot i,$$

$$g''(t) = e^{\lambda(\exp(it)-1)} \cdot \lambda e^{it} \cdot i \cdot \lambda e^{it} \cdot i + e^{\lambda(\exp(it)-1)} \cdot \lambda i \cdot e^{it} \cdot i.$$

Найдем значения производных при $t = 0$:

$$g'(0) = \lambda \cdot i,$$

$$g''(0) = \lambda^2 \cdot i^2 + \lambda \cdot i^2.$$

Определяем 1-й и 2-й начальные моменты:

$$\alpha_1 = \frac{\lambda i}{i} = \lambda; \quad \alpha_2 = \frac{i^2(\lambda^2 + \lambda)}{i^2} = \lambda^2 + \lambda.$$

Теперь найдем числовые характеристики:

$$M[X] = \lambda; \quad D[X] = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Видим, что для Пуассоновского закона $M[X]$ и $D[X]$ совпадают. Это позволяет сделать *замечание*: если для некоторой случайной величины $M[X]$ и $D[X]$ совпадают, то такая случайная величина имеет, *скорее всего*, закон распределения Пуассона.

Закон распределения Пуассона служит приближением для биномиального закона, если $n \geq 100$; $0 \leq np < 10$.

6.3. Равномерный закон распределения

Непрерывная случайная величина X имеет равномерный закон распределения на отрезке $[a, b]$, если на этом отрезке плотность распределения вероятностей $f(x)$ сохраняет постоянное значение:

$$f(x) = \begin{cases} c, & \forall x \in [a, b], \\ 0, & \forall x \notin [a, b]. \end{cases}$$

На рис. 6.5 приведен график плотности распределения для равномерного закона.

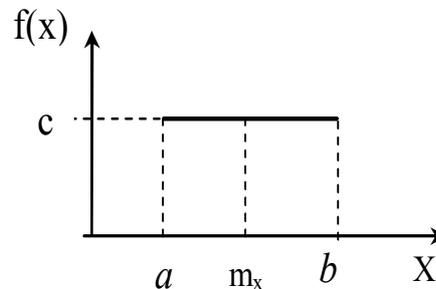


Рис. 6.5

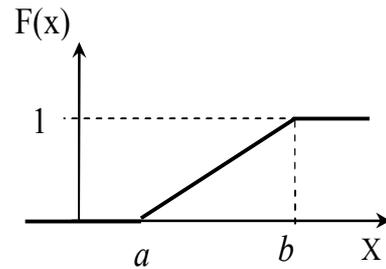
Определим c из условия нормировки:

$$\int_{-\infty}^{\infty} f(x) dx = \int_a^b f(x) dx = c \int_a^b dx = c(b-a) = 1 \Rightarrow c = \frac{1}{b-a}.$$

Найдем функцию распределения:

$$F(x) = \int_{-\infty}^x f(x) dx = \int_a^b \frac{dx}{b-a} = \frac{x}{b-a} \Big|_a^x = \frac{x-a}{b-a}.$$

$$F(x) = \begin{cases} 0, & \forall x \leq a, \\ \frac{x-a}{b-a}, & \forall x \in [a, b], \\ 1, & \forall x > b. \end{cases}$$



На рис. 6.6 приведен график $F(x)$.

Рис. 6.6

Определим числовые характеристики:

$$M[x] = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b x \frac{dx}{b-a} = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{a+b}{2},$$

$$M[x^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_a^b x^2 \frac{dx}{b-a} = \frac{1}{b-a} \left. \frac{x^3}{3} \right|_a^b = \frac{1}{3} \frac{b^3 - a^3}{b-a} = \frac{a^2 + ab + b^2}{3},$$

$$D[x] = M[x^2] - (M[x])^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2} \right)^2 = \frac{(b-a)^2}{12}.$$

6.4. Показательный (экспоненциальный) закон распределения

Показательный (экспоненциальный) закон распределение имеет важное значение и широко используется в теории надежности.

Случайная величина имеет показательный (экспоненциальный) закон распределения, если плотность распределения определяется так:

$$f(x) = \begin{cases} 0 & \forall x < 0, \\ \theta e^{-\theta x} & \forall x \geq 0. \end{cases}$$

На рис. 6.7 приведен график функции $f(x)$, на рис. 6.8 приведен график функции распределения.

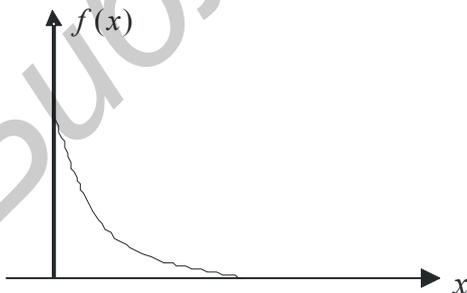


Рис. 6.7

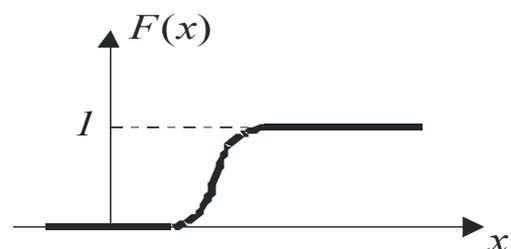


Рис. 6.8

Определим числовые характеристики, используя характеристические функции. Найдем характеристическую функцию:

$$g(t) = \int_0^{\infty} \theta e^{-\theta x} e^{itx} dx = \theta \int_0^{\infty} e^{-x(-it+\theta)} dx = \theta \frac{e^{-x(-it+\theta)}}{-it+\theta} \Big|_0^{\infty} = \lambda \frac{e^{-x(\theta-it)}}{it-\theta} \Big|_0^{\infty} = -\frac{\theta}{it-\theta} = \theta(\theta-it)^{-1}.$$

Вычислим производные $g'(t)$ и $g''(t)$:

$$g'(t) = \frac{\theta i}{(\theta-it)^2}, \quad g''(t) = 2 \frac{\theta i^2}{(\theta-it)^3}.$$

Найдем значения производных при $t = 0$:

$$g'(0) = \frac{i}{\theta}, \quad g''(0) = \frac{2i^2}{\theta^2}.$$

Вычислим 1-й и 2-й начальные моменты:

$$\alpha_1 = \frac{1}{\theta}, \quad \alpha_2 = \frac{2}{\theta^2}.$$

Определим $M[X]$ и $D[X]$:

$$M[X] = \alpha_1 = \frac{1}{\theta}, \quad D[x] = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}, \quad \sigma_x = \frac{1}{\theta}.$$

Отметим, что $M[X]$ и σ_x совпадают; это позволяет сделать *замечание*: если для случайной величины $M[X]$ и σ_x совпадают, то такая случайная величина имеет *скорее всего* экспоненциальный закон распределения.

6.5. Нормальный (гауссовский) закон распределения

Случайная величина имеет нормальный закон распределения, если плотность распределения вероятностей имеет следующий вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}. \quad (6.2)$$

Здесь $-\infty < x < \infty$,

σ и m – параметры.

На рис. 6.9 приведен график нормального закона.

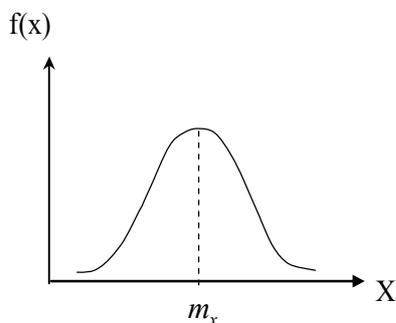


Рис. 6.9

Проверим характеристические свойства $f(x)$:

1. $f(x) > 0$.

2. Проверим условие нормировки.

При вычислениях по нормальному закону широко применяется *интеграл Пуассона*:

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

Вычислим интеграл:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \text{используем подстановку: } \frac{x-m}{\sigma} = u \right| = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} \sigma \cdot du = 1. \end{aligned}$$

Видим, что условие нормировки выполняется.

Пространство элементарных событий является непрерывным и представляет собой всю числовую ось $(-\infty; \infty)$.

Функция распределения для нормального закона записывается так:

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

Уточним вероятностный смысл параметров σ и m . Найдем характеристическую функцию для нормального распределения:

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} e^{itx} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left| \begin{array}{l} \frac{x-m}{\sigma} = z \\ \sigma \cdot dz = dx \\ x = \sigma \cdot z + \mu \end{array} \right| = \frac{1 \cdot \sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it(z\sigma + \mu) - \frac{z^2}{2}} dz = \\ &= \left| \begin{array}{l} it(z\sigma + m) - \frac{z^2}{2} = itz\sigma + itm - \frac{z^2}{2} = \\ = itm - \frac{1}{2}(z^2 - 2itz\sigma + (it\sigma)^2) + \frac{(it\sigma)^2}{2} = \\ = itm + \frac{(it\sigma)^2}{2} - \frac{1}{2}(z - it\sigma)^2 \end{array} \right| = \frac{1}{\sqrt{2\pi}} e^{itm - \frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-it\sigma)^2} dz = e^{itm - \frac{t^2\sigma^2}{2}}. \end{aligned}$$

Продифференцируем характеристическую функцию 2 раза:

$$\begin{aligned} g'(t) &= e^{itm - \frac{t^2\sigma^2}{2}} (im - t\sigma^2), \\ g''(t) &= -\sigma^2 e^{-itm - \frac{t^2\sigma^2}{2}} + (im - t\sigma^2)^2 e^{itm - \frac{t^2\sigma^2}{2}}. \end{aligned}$$

Определим 1-й, 2-й начальные моменты:

$$\alpha_1 = \frac{g'(0)}{i} = m,$$

$$\alpha_2 = \frac{g''(0)}{i^2} = \frac{-\sigma^2 + i^2 m^2}{i^2} = i^2 \sigma^2 + m^2.$$

Тогда

$$M[x] = \alpha_1 = m, \quad D[x] = \alpha_2 - \alpha_1^2 = \sigma^2 + m^2 - m^2 = \sigma^2.$$

Таким образом, мы установили вероятностный смысл параметров m и σ . Построим график плотности нормального распределения (кривой Гаусса). Найдем моду случайной величины с нормальным законом, для чего решим уравнение

$$f'(x) = \frac{-1}{\sigma\sqrt{2\pi}} \frac{x-m}{\sigma^2} e^{-\frac{(x-m)^2}{2\sigma^2}} = 0.$$

При $x = m$ $f(x)$ достигает максимума, m является модой. Для определения точек перегиба функции необходимо решить уравнение

$$f''(x) = \frac{-1}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \left[1 - \frac{(x-m)^2}{\sigma^2} \right] = 0.$$

$$1 - \frac{(x-m)^2}{\sigma^2} = 0, \quad (x-m)^2 = \sigma^2.$$

Видим, что функция имеет две точки перегиба: $x = m \pm \sigma$. На рис. 6.10 приведены кривые Гаусса для разных значений σ . На рис. 6.11 приведены кривые Гаусса для разных значений параметра m .

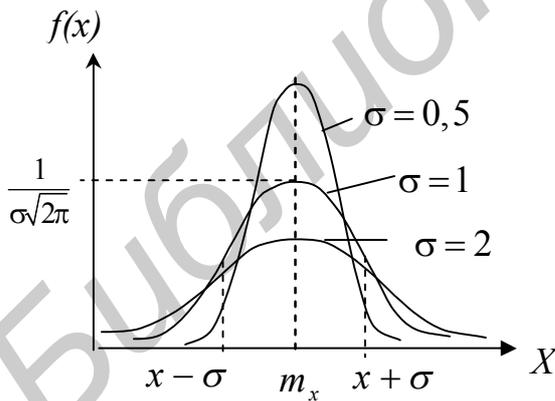


Рис. 6.10

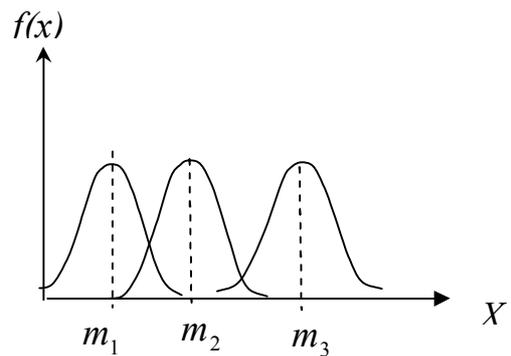


Рис. 6.11

Площадь криволинейной трапеции, ограниченной осью абсцисс, кривой Гаусса и прямыми $x' = m \pm \sigma$, приблизительно равна ($S' \approx 0,68$); если прямые раздвинуть на $2\sigma - x'' = m \pm 2\sigma$, то $S'' \approx 0,95$; при $x''' = m \pm 3\sigma$, $S''' \approx 0,997$. Эти соотношения позволяют сформулировать **правило трех сигм** (3σ):

Практически все возможные значения случайной величины, распределенной по нормальному закону, лежат в пределах от -3σ до 3σ относительно m .

$$\text{Коэффициент асимметрии } A = \frac{\mu_3}{\sigma^3} = 0; \text{ эксцесс } \mathcal{E} = \frac{\mu_4}{\sigma^4} - 3 = 0.$$

Нормальное распределение имеет исключительно важное значение в теории вероятностей и часто встречается на практике в самых различных областях. Например: отклонения в величине параметров полупроводниковых приборов при массовом производстве, вес и размер деталей, вес и рост людей имеют нормальное распределение.

6.5.1. Функция Лапласа

Определим вероятность того, что случайная величина, распределенная по нормальному закону, попадет в интервал $[\alpha, \beta]$.

$$P(\alpha < x < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \left. \begin{array}{l} \frac{x-m}{\sigma} = t \\ dx = \sigma \cdot dt \\ t_{\beta} = \frac{\beta-m}{\sigma} \\ t_{\alpha} = \frac{\alpha-m}{\sigma} \end{array} \right| = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\beta-m}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\alpha-m}{\sigma}} e^{-\frac{t^2}{2}} dt =$$

$$= \frac{1}{2} \left[\Phi\left(\frac{\beta-m}{\sigma}\right) - \Phi\left(\frac{\alpha-m}{\sigma}\right) \right]. \quad (6.3)$$

Функция $\Phi(x)$ называется интегралом вероятностей или *функцией Лапласа* (рис. 6.12).

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt. \quad (6.4)$$

Свойства функции Лапласа

1. Функция нечетная
 $\Phi(-x) = -\Phi(x)$.
2. $\forall x_2 > x_1 \Rightarrow \Phi(x_2) > \Phi(x_1)$ — свойство монотонности.
3. $\lim_{x \rightarrow \infty} \Phi(x) = 1$, $\lim_{x \rightarrow -\infty} \Phi(x) = -1$.

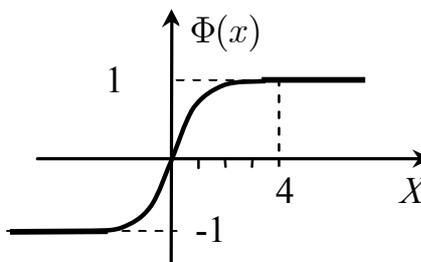


Рис. 6.12

4. Значения $\Phi(x)$, применяющиеся наиболее часто:

$$\Phi(0) = 0, \quad \Phi(1) = 0,6827, \quad \Phi(2) = 0,9545, \quad \Phi(3) = 0,9373, \quad \Phi(4) = 0,99994, \\ \forall x > 4, \quad \Phi(x) \approx 1.$$

Пример.

Цех на заводе выпускает транзисторы с емкостью коллекторного перехода $C_k = 1,5$ пФ. Сколько транзисторов попадет в группу «Б», если в нее попадают транзисторы с емкостью коллекторного перехода от 1,80 до 2,00 пФ. Цех выпустил партию в 1000 штук.

Решение.

Статистическими исследованиями в цеху установлено, что C_k можно трактовать как случайную величину, подчиняющуюся нормальному закону.

Чтобы вычислить количество транзисторов, попадающих в группу «Б», необходимо учитывать, что вся партия транзисторов имеет разброс параметров, покрывающий всю (условно говоря) числовую ось. То есть кривая Гаусса охватывает всю числовую ось, центр ее совпадает с $X = m = 1,5$ пФ (т. к. все установки в цеху настроены на выпуск транзисторов именно с этой емкостью). Вероятность попадания отклонений параметров всех транзисторов на всю числовую ось равна 1. Поэтому нам необходимо фактически определить вероятность попадания случайной величины X в интервал $[1,8; 2,0]$, а затем пересчитать количество пропорциональной вероятности.

Для расчета этой вероятности надо построить математическую модель. Экспериментальные данные говорят о том, что нормальное распределение можно принять в качестве математической модели. Эмпирическая оценка (установлена статистическими исследованиями в цеху) среднего значения C_k дает $\bar{m} = \bar{x}_{cp} \approx 1,5$ пФ, оценка среднего квадратического отклонения $\sigma = 0,15$ пФ. Обозначая $\alpha = 1,8$ пФ, $\beta = 2,0$ пФ, подставим приведенные значения в (6.3):

$$P(1,8 < X < 2,0) = \frac{1}{2} \left[\Phi \left(\frac{2,0 - 1,5}{0,15} \right) - \Phi \left(\frac{1,8 - 1,5}{0,15} \right) \right] = \\ = \frac{1}{2} [\Phi(3,3) - \Phi(2)] \approx \frac{1}{2} [0,998 - 0,9545] \approx \frac{1}{2} 0,044 \approx 0,022.$$

Тогда количество транзисторов n , попавших в интервал $[1,8; 2,0]$ пФ, можно найти так: $n = 1000 \cdot 0,022 = 22$ шт. Таким образом можно планировать и рассчитывать количество транзисторов, попадающих в ту или иную группу.

ГЛАВА 7. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Группа теорем, устанавливающих соответствие между теоретическими и экспериментальными характеристиками случайных величин и случайных событий при большом числе повторения эксперимента в одинаковых условиях, а также касающихся предельных законов распределения, под общим названием *предельных теорем*.

Закон больших чисел занимает важнейшее место в теории вероятностей. Он устанавливает связь между математическими моделями эксперимента и практикой. Этот закон выражается рядом теорем (Чебышева, Бернулли, Ляпунова и др.). В этих теоремах указывают условия, при выполнении которых совокупное действие многих случайных причин приводят к устойчивым результатам, практически независимым от случая.

7.1. Неравенство Чебышева

Пусть X произвольная случайная величина и ε – любое число > 0 , тогда справедливо следующее неравенство:

$$P(|x| \geq \varepsilon) \leq \frac{M[x^2]}{\varepsilon^2}. \quad (7.1)$$

Докажем это для непрерывной случайной величины X с $f(x)$.

$$\begin{aligned} M[x^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx \geq \int_{|x| \geq \varepsilon} x^2 f(x) dx = \left. \begin{array}{l} \text{т. к. подынтегральная функция} \\ \text{положительная, то уменьшение} \\ \text{пределов интегрирования ведет} \\ \text{к уменьшению значения интеграла} \end{array} \right| = \\ &= \left. \begin{array}{l} \text{т. к. } |x| \geq \varepsilon, \text{ то положим} \\ x = \varepsilon \text{ — мин значению, т. е.} \\ x = \varepsilon \text{ — это также не} \\ \text{увеличивает интеграл} \end{array} \right| \geq \int_{|x| \geq \varepsilon} \varepsilon^2 f(x) dx = \varepsilon^2 \int_{|x| \geq \varepsilon} f(x) dx = \varepsilon^2 P(|x| \geq \varepsilon) \Rightarrow \\ &\Rightarrow P(|x| \geq \varepsilon) \leq \frac{M[x^2]}{\varepsilon^2}. \end{aligned}$$

Для дискретной случайной величины доказательство аналогично.

Полагая в (7.1) вместо $X \rightarrow X - M[X]$ (т. к. X – была произвольна) и переходя к противоположному событию, получим

$$P(|X - M[X]| < \varepsilon) \geq 1 - \frac{D[X]}{\varepsilon^2}. \quad (7.2)$$

Это и есть **неравенство Чебышева**. Оно имеет большое теоретическое значение, т. к. *позволяет оценить вероятность отклонения случайной величины от ее среднего значения, когда неизвестен закон распределения, но известны только первые два начальных момента.*

Пример.

Полагая в (2) $\varepsilon = 3\sigma$, получим

$$P(|X - M[X]| < 3\sigma) \geq 1 - \frac{\sigma^2}{9\sigma^2} = 1 - \frac{1}{9} \approx 0,899,$$

где $\sigma^2 = D[X]$.

Если X – нормальная случайная величина, то $P(|X - M[X]| < 3\sigma) \approx 0,997$, т. к. $0,899 \ll 0,997$ (с вероятностной точки зрения), видно, что неравенство Чебышева дает грубую оценку.

Все же неравенство Чебышева имеет большое значение в практике, т. к. позволяет действовать с неизвестными законами распределения.

Пусть дана последовательность случайных величин $X_1, X_2, \dots, X_n, \dots$, определенных на одном и том же вероятностном пространстве (Ω, F, P) . Тогда *последовательность случайных величин $X_n, n = \overline{1, \infty}$ сходится к случайной или не случайной величине X по вероятности, если для $\forall \varepsilon > 0$ выполняется соотношение*

$$\lim_{n \rightarrow \infty} P\{|X_n - X| < \varepsilon\} = 1. \quad (7.3)$$

В обычном анализе неравенство $|X_n - X| \leq \varepsilon$ всегда выполняется, начиная с какого-то значения n_0 . Если же $X_n - X$ при $n \rightarrow \infty$ по вероятности, то это неравенство для отдельных n может и не выполняться.

7.2. Теорема Чебышева

Пусть $X_1, X_2, \dots, X_n, \dots$ последовательность независимых случайных величин с ограниченной $D[X]$, т. е.

$$\left(\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i,j=1}^n \text{cov}(X_i, X_j) = 0, \quad D[X_i] \leq c, \text{ т. е. ограничена } \right)^*.$$

Тогда среднее арифметическое этих величин сходится по вероятности к среднему арифметическому их математических ожиданий:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M[X_i] \right| < \varepsilon \right) = 1. \quad (7.4)$$

где ε – любое, сколь угодно малое положительное число.

Доказательство:

Используем неравенство Чебышева (7.2). Положим вместо $X \rightarrow \frac{1}{n} \sum_{i=1}^n X_i$.

Определим $M[X] = \frac{1}{n} \sum_{i=1}^n M[X_i]$ (применили свойство аддитивности).

Вычислим дисперсию для $\frac{1}{n} \sum_{i=1}^n X_i$:

$$D \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n^2} D \left[\sum_{i=1}^n X_i \right] = \frac{1}{n^2} \sum_{i=1}^n D[X_i] = \left| \begin{array}{l} \text{с учетом условия (*)} \\ D[X_i] \leq c \end{array} \right| \leq \frac{cn}{n^2} = \frac{c}{n}.$$

Подставляем полученные выражения в (7.2):

$$P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M[X_i] \right| < \varepsilon \right) \geq 1 - \frac{c}{n\varepsilon^2}. \quad (7.5)$$

Переходя к пределу и учитывая, что $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n\varepsilon^2} \right) = 1$, а также то, что вероятность не может быть больше 1, получаем:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n M[X_i] \right| < \varepsilon \right) = 1.$$

На практике нередко используется следующий **частный случай теоремы Чебышева**.

Если $x_1, x_2, \dots, x_n, \dots$ – независимые случайные величины с одинаковым математическим ожиданием, равным a , и равномерно ограниченной дисперсией (т. е. $M[x_i] = a$, $D[X_i] \leq c = \text{const}$, $\forall i = \overline{1, n}$), то среднее арифметическое этих случайных величин сходится по вероятности к a .

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - a \right| < \varepsilon \right) = 1. \quad (7.6)$$

Доказательство:

Это следует из (7.5), где $\frac{1}{n} \sum M[X_i] = \frac{na}{n} = a$. Тогда правая часть в пределе равна $\lim_{n \rightarrow \infty} (1 - \frac{c}{n\varepsilon^2}) = 1$.

Таким образом, согласно теореме Чебышева, среднее арифметическое большого числа слабозависимых или независимых (условие *) случайных величин является практически неслучайной величиной, т. е. постоянной, обладающей сколь угодно малой дисперсией. Причина этого явления – взаимное погашение ошибок. Если прогноз каждого отдельного значения случайной величины ненадежен, то прогноз среднего арифметического является достаточно надежным.

Пример.

Пусть x_k – результат k -го измерения длины, например комнаты.

1. Случайные величины x_1, x_2, \dots, x_n – независимы.

2. $M[x_i] = a$ – т. к. длина комнаты – существует.

3. $D[x_i] \leq c$ – т. к. ошибка не накапливается.

Тогда вероятность при $n \rightarrow \infty$, $P \rightarrow 1$ для $\left| \frac{x_1 + \dots + x_n}{n} - a \right| < \varepsilon$.

7.3. Теорема Бернулли

Пусть производится n независимых опытов, в каждом из которых наступает событие A или \bar{A} и $P(A) = p$. Событие A происходит в n опытах m раз. Тогда относительная частота $\frac{m}{n}$ события A сходится по вероятности к вероятности p :

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \quad (7.7)$$

Доказательство:

Рассмотрим двоичную случайную величину X_k , равную числу наступлений события A в каждом опыте. Закон распределения ее задан в виде таблицы.

X_k	$x_1 = 0$	$x_2 = 1$
P	$1-p$	p

Покажем, что для случайной величины X_k выполняются условия частного случая теоремы Чебышева:

$$M[X_k] = \sum_{i=1}^2 x_i p_i = 0 \cdot (1-p) + 1 \cdot p = p,$$

$$M[X_k^2] = \sum_i x_i^2 p_i = 0^2 \cdot (1-p) + 1^2 \cdot p = p,$$

$$D[X_k] = M[X_k^2] - (M[X_k])^2 = p - p^2 = p(1-p).$$

Таким образом, случайные величины X_1, X_2, \dots, X_n независимы, имеют одинаковое математическое ожидание, равное p , и ограниченную дисперсию. Можно применить частный случай теоремы Чебышева. Учитывая, что $M[X_k] = p = a$, $D[X_k] = p(1-p) = c$, $\frac{1}{n} \sum_{k=1}^n X_k = \frac{m}{n}$, т. к. в n опытах событие A наступает m раз, тогда из (7.6) получим

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1.$$

Формула (7.7) объясняется свойство устойчивости относительной частоты, состоящей в том, что в разных опытах, проведенных в одинаковых условиях, она колеблется около некоторого числа, которое есть не что иное, как вероятность; (7.7) позволяет вместо неизвестной вероятности использовать известную относительную частоту.

Пример.

Пусть проводится эксперимент с броском монеты. Обозначим через событие A появление герба. Пусть при n бросках герб появится m раз. Найдем относительную частоту появления герба: $\nu = \frac{m}{n}$. Повторяя такой опыт много раз, будем откладывать результаты на графике. (рис. 7.1).

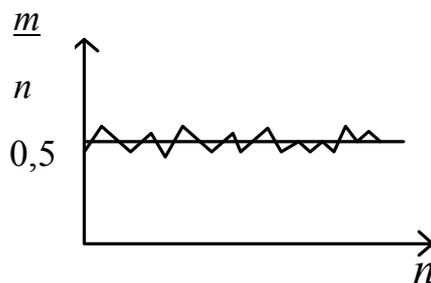


Рис. 7.1

Мы увидим, что относительная частота будет колебаться около некоторого постоянного числа (в случае с симметричной монетой это $0,5$). И это значение можно принимать за неизвестную вероятность события A . Но надо огово-

речь, что $\frac{m}{n} \rightarrow p$ при $n \rightarrow \infty$ по вероятности, т. к. возможность отклонения всегда остается.

7.4. Центральная предельная теорема (Теорема Ляпунова)

Эта теорема указывает условия, при которых закон распределения суммы большого числа независимых случайных величин близок к нормальному закону. Впервые она была сформулирована и доказана великим русским математиком академиком А. М. Ляпуновым (1857 – 1918).

Распределение суммы независимых случайных величин $X_1, X_2, \dots, X_n, \dots$, имеющих произвольный, но одинаковый закон распределения при $n \rightarrow \infty$, стремится к нормальному распределению. (Эти ограничения существенно упрощают доказательство).

Доказательство:

Основная идея – найдем характеристическую функцию, а она окажется характеристической функцией нормальной случайной величины.

Рассмотрим случайную величину $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_k$.

Пусть случайные величины X_k имеют следующие числовые характеристики:

$$\mu = M[X_k], \quad \sigma^2 = D[X_k].$$

Нормируем случайную величину Y :

$$Y_0 = \frac{Y - M[Y]}{\sqrt{D[Y]}} = \left| \begin{array}{l} D[Y] = D\left[\sum_{i=1}^n X_k\right] = \sum_{i=1}^n D[X_k] = nD[X_k] = n\sigma^2, \\ M[Y] = \sum_{i=1}^n M[X_k] = n\mu \end{array} \right| = \sum_{k=1}^n \frac{X_k - \mu}{\sqrt{n}\sigma}.$$

Найдем характеристическую функцию $g(t)$ вспомогательной величины Y_0 .

$$g(t) = M\left[e^{itY_0}\right] = M\left[e^{it\sum_{k=1}^n \frac{X_k - \mu}{\sqrt{n}\sigma}}\right] = M\left[\prod_{k=1}^n e^{it\frac{X_k - \mu}{\sqrt{n}\sigma}}\right] =$$

$$= \left| \begin{array}{l} \text{т. к. случайные величины } X_k - \\ \text{независимы, то используем} \\ \text{свойство мультипликативности} \\ \text{математического ожидания} \end{array} \right| = \prod_{k=1}^n M\left[e^{it\frac{X_k - \mu}{\sqrt{n}\sigma}}\right] =$$

$$\begin{aligned}
&= \left[\begin{array}{l} \text{разложим } e^x \text{ в ряд:} \\ e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \end{array} \right] = \prod_{k=1}^n M \left[1 + it \frac{x_k - \mu}{\sqrt{n}\sigma} + i^2 t^2 \frac{(x_k - \mu)^2}{2!(\sqrt{n}\sigma)^2} + \dots \right] = \\
&= \left[\begin{array}{l} \prod_{k=1}^n M = (M)^n \\ \text{ограничимся 3 - мя} \\ \text{членами разложения} \\ \text{и учтём, что} \\ M[X - m_x] = 0, \\ M[(X - m_x)^2] = \sigma^2, \end{array} \right] = \left(1 - \frac{t^2 \sigma^2}{2\sigma^2 n} + \frac{c}{n^{3/2}} + \dots \right)^n.
\end{aligned}$$

Здесь c – некоторая постоянная, которая не зависит от n и содержит факториалы. Переходя к логарифмам и используя формулу $\ln(1 + \varepsilon) \approx \varepsilon$, получаем

$$\ln g(t) = n \ln \left(1 - \frac{t^2}{2n} + \frac{c}{n^{3/2}} + \dots \right) \approx -\frac{t^2}{2} + \frac{c}{n^{1/2}} + \dots$$

Найдем предел этого выражения при $n \rightarrow \infty$ и увидим, что $\lim_{n \rightarrow \infty} \frac{c}{\sqrt{n}} = 0$, тогда

$$\ln g(t) = -\frac{t^2}{2}.$$

Потенцируя это выражение, получаем $g(t) = e^{-\frac{t^2}{2}}$. Отсюда следует, что характеристическая функция нормированной последовательности Y_0 сходится к характеристической функции нормированной нормальной случайной величины.

Таким образом, при достаточно большом n распределение суммы случайных величин приближенно совпадает с нормальным распределением с математическим ожиданием $n\mu$ и дисперсией $n\sigma^2$.

Доказанная теорема, как установил Ляпунов, справедлива также и для случая, когда случайные величины X_1, X_2, \dots, X_k слабо зависимы и имеют произвольные разные законы распределения. При этом требуется, чтобы никакое X_k не доминировало над остальными слагаемыми в $\sum_{k=1}^n X_k$. Этим и объясняется широкое распространение в природе и технике нормального закона распределения.

Пример.

Показать, что характеристическая функция $g(t) = e^{-\frac{t^2}{2}}$ имеет плотность распределения нормированной нормальной случайной величины.

Решение.

Используем обратное преобразование Фурье:

$$f(y_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iy_0 t} g(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iy_0 t} e^{-\frac{t^2}{2}} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2} - iy_0 t} dt = \left. \begin{array}{l} \text{для краткости} \\ \text{записываем } y \\ \text{вместо } y_0 \end{array} \right| =$$

$$= \left. \begin{array}{l} -\frac{t^2}{2} - iy_0 t = \\ = -\frac{1}{2}(t^2 + 2ity_0 + (iy_0)^2) - \\ - (iy_0)^2 = -\frac{1}{2}(t + iy_0)^2 + \\ + \frac{1}{2}(iy_0)^2 \end{array} \right| = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(iy_0)^2} e^{-\frac{1}{2}(t+iy_0)^2} dt = \left. \begin{array}{l} U = t + iy_0 \\ dU = dt \end{array} \right| = \frac{e^{-\frac{y_0^2}{2}}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{U^2}{2}} dU =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y_0^2}{2}}.$$

7.5. Формулы Муавра – Лапласа

Рассмотрим частный случай центральной предельной теоремы, когда слагаемые в $Y = \sum_{k=1}^n X_k$ являются дискретными случайными величинами. Пусть производится n независимых испытаний по схеме Бернулли, в каждом из которых событие A появляется с вероятностью p : $P(A) = p$, $P(\bar{A}) = 1 - p = q$. Тогда

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

Введем случайную величину X_k , равную числу наступлений событий A в k -м опыте. Рассмотрим случайную величину $Y = X_1 + X_2 + \dots + X_n$, т. е. равную общему числу наступлений события A в n опытах. Случайные величины X_k независимы и имеют одинаковый закон распределения.

X_k	0	1
P_k	q	p

Ранее мы показали (см. раздел 7.3), что $M[X_k] = p$, $D[X_k] = p \cdot q$.

Сформируем центрированную случайную величину $Z_n = \frac{Y - np}{\sqrt{npq}}$.

Тогда при $n \rightarrow \infty$ распределение суммы случайных величин стремится к нормальному распределению с параметрами $M[Z] = 0$ и $D[Z] = 1$. Обозначим

функцию распределения $F_n(z) = P(Z_n < z)$. Тогда последовательность функции распределения $F_n(z)$ сходится по распределению к функции распределения нормированного нормального закона $N(0,1)$:

$$\lim_{n \rightarrow \infty} F_n(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}\sqrt{npq}} \int_{-\infty}^Y e^{-\left(\frac{Y-np}{\sqrt{npq}}\right)^2 \frac{1}{2}} dY.$$

Вероятность попадания случайной величины Y в интервал $[a, b]$ (на основании свойства 2 функции распределения):

$$P(a \leq Y \leq b) \approx \frac{1}{2} \left[\Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right) \right].$$

Это интегральная теорема Муавра – Лапласа, где $\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{z^2}{2}} dz$ – функция Лапласа.

Вероятность того, что при n независимых испытаниях события A наступят ровно k раз (при больших n), удобно считать на основании **локальной теоремы Лапласа**:

$$P(Y = k) \approx \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{k-np}{\sqrt{npq}}\right)^2 \frac{1}{2}} = \frac{1}{\sqrt{npq}} \varphi(x),$$

где $x = \frac{k-np}{\sqrt{npq}}$.

Пример:

Вероятность успешной работы телевизора после сборки равна 0,75. Найти вероятность того, что из 10 телевизоров 8 заработают. Используем локальную теорему Лапласа и получаем $P_{10}(Y = 8) \approx 0,27$ – по таблицам, а по формуле Бернулли – $P_{10}(8) = 0,28$.

ЧАСТЬ 2. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ГЛАВА 8. ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

8.1. Выборка, вариационный ряд, гистограмма

Если теория вероятностей оперирует с известными законами распределения и их параметрами (числовыми характеристиками), то математическая статистика по результатам экспериментов проверяет, правильно ли подобрано распределение (нормальное, биномиальное, экспоненциальное и т. д.), оценивает параметры этого распределения, проверяет гипотезы о параметрах принятого распределения. Это позволяет заменить большое число экспериментальных данных небольшим числом параметров распределения, которые в сжатом виде характеризуют случайную величину и позволяют прогнозировать результаты эксперимента при известном комплексе условий.

Пусть проводится n измерений. В результате измерений получено n чисел x_1, x_2, \dots, x_n . Если повторить еще раз n измерений, то получатся другие n чисел, отличные от первого набора. Процесс из n измерений можно описать как n независимых случайных величин.

*Результат n наблюдений x_1, x_2, \dots, x_n случайной величины X называется **выборкой**, n – объем выборки, а сама случайная величина X – называется **генеральной случайной величиной**.*

Результат эксперимента x_i может быть интерпретирован либо апостериорной величиной, либо априорной. В первом случае это результат опыта. Во втором случае x_i является случайной величиной (т. к. до опыта неизвестна), которая получит свое конкретное значение в результате какого-то i -го опыта. В этом случае можно предполагать, что закон распределения x_i совпадает с законом распределения генеральной случайной величиной X и x_i можно рассматривать как экземпляр генеральной случайной величины X .

Далее мы будем считать выборки *априорными*. При этом будем полагать, что элементы выборки – независимые случайные величины с одинаковым законом распределения, т. е. мы можем широко использовать теоремы независимых случайных величинах.

Упорядоченная в порядке возрастания последовательность выборочных значений образует **вариационный ряд**:

$$x^{(1)} < x^{(2)} < \dots < x^{(n)},$$

члены вариационного ряда $x^{(i)}$ называются *порядковыми статистиками*. Если объем выборки n – велик, то выборка позволяет приблизительно оценить закон распределения случайной величиной X . Для этого необходимо построить гистограмму. Есть два способа построения гистограммы – *равноинтервальный* и *равновероятностный*.

Рассмотрим *равноинтервальный* способ.

1. Разобьем весь диапазон выборочных значений от x_{\max} до x_{\min} на k равных частей. Величину k выбирают достаточно произвольно, можно так: $k = \sqrt{n}$, где n – объем выборки.

2. Определяем длину каждого интервала: $h = \frac{x_{\max} - x_{\min}}{k}$.

3. Находим границы каждого интервала:

для первого: $a_1 = x_{\min}, b_1 = a_1 + h$;

для второго: $a_2 = b_1, b_2 = a_2 + h$;

...

для k -го: $a_k = b_{k-1}, b_k = a_k + h$.

Определим середины каждого интервала: $x_0^i = \frac{a_i + b_i}{2}, i = \overline{1, k}$.

4. Подсчитываем (используя вариационный ряд) количество выборочных значений, попадающих в i -й интервал – m_i .

5. Находим относительную частоту $v_i = \frac{m_i}{n}$ попадания случайной величиной X в i -й интервал.

Полученные данные заносим в таблицу.

x_0^1 [a_1, b_1]	x_0^2 [a_2, b_2]	...	x_0^i [a_i, b_i]	...	x_0^k [a_k, b_k]
m_1	m_2	...	m_i	...	m_k
$\frac{m_1}{n}$	$\frac{m_2}{n}$...	$\frac{m_i}{n}$...	$\frac{m_k}{n}$

Эта таблица называется *статистическим рядом*.

Графическое изображение статистического ряда – это гистограмма.

Рисуем оси координат, делаем разметку осей, наносим на ось X границы интервалов и их середины. После этого строим на каждом отрезке прямоугольники высотой v_i . Аппроксимируем фигуру из прямоугольников пунктирной линией (рис. 8.1). По виду этой кривой можно выдвинуть предположение (гипотезу) о виде закона распределения генеральной случайной величины X (на рис. 8.1. видно, что пунктирная линия похожа на кривую Гаусса, которая относится к нормальному закону).

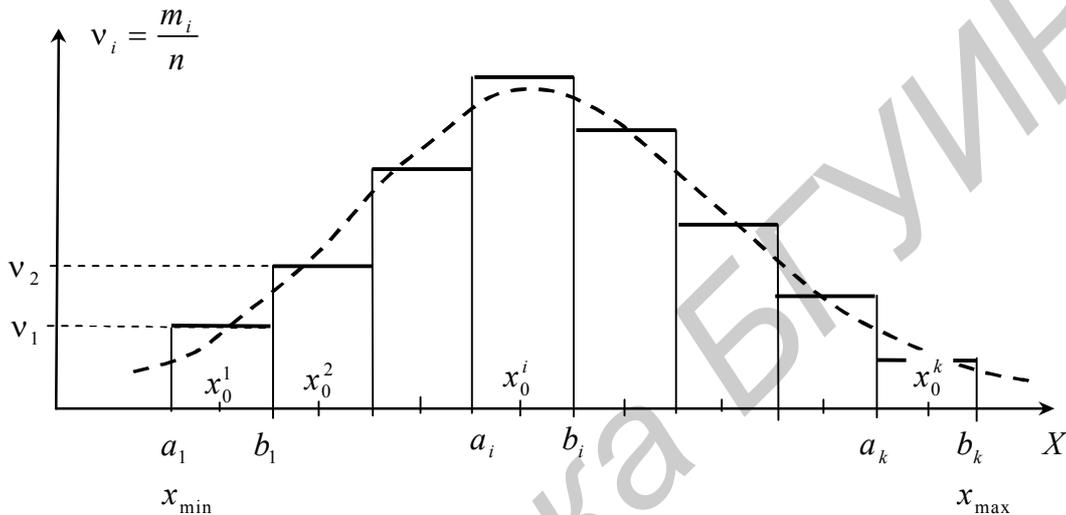


Рис. 8.1

Имея статистический ряд можно оценить числовые характеристики генеральной случайной величиной X :

$$M[x] = \sum_{i=1}^k \frac{m_i}{n} x_0^i = \frac{1}{n} \sum m_i x_0^i = a,$$

$$D[x] = \sum_{i=1}^k \frac{m_i}{n} (x_0^i - a)^2 = \frac{1}{n} \sum m_i (x_0^i - a)^2.$$

8.2. Оценки и методы их получения

*Приближенные значения параметров, входящих в законы распределения, определяемые каким-либо способом по выборкам, называются оценками или статистиками. Оценки бывают точечными и интервальными. Точечные оценки представляются одним числом, интервальные – двумя числами (θ_1 и θ_2): началом и концом интервала, накрывающего оцениваемый параметр. Самыми распространенными методами получения точечных оценок являются *метод моментов* и *метод наибольшего правдоподобия*.*

8.2.1. Метод моментов

Пусть генеральная случайная величина X имеет плотность распределения $f(x)$. Начальный и центральный моменты можно найти так (см. раздел 4.1.6):

$$\alpha_K = \int_{-\infty}^{+\infty} x^k f(x) dx, \quad (8.1)$$

$$\mu_K = \int_{-\infty}^{+\infty} (x - \alpha_1)^k f(x) dx. \quad (8.2)$$

По выборке x_1, x_2, \dots, x_n определяем *выборочные* начальные и центральные моменты:

$$\bar{\alpha}_K = \frac{1}{n} \sum_{i=1}^n x_i^K, \quad (8.3)$$

$$\bar{\mu}_K = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{\alpha}_1)^K. \quad (8.4)$$

Метод моментов состоит в том, что генеральные моменты (8.1, 8.2), в которые входят оцениваемые параметры, приблизительно приравниваются к соответствующим выборочным моментам (8.3), (8.4). Составляется система уравнений:

$$\alpha_k \approx \bar{\alpha}_k, \quad (8.5)$$

$$\mu_k \approx \bar{\mu}_k. \quad (8.6)$$

Решая систему (8.5), (8.6), находим оцениваемые параметры.

Особо важную роль играет $\bar{\alpha}_1$ – выборочный начальный момент 1-го порядка, он называется *выборочным средним* и обозначается \bar{x} :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (8.7)$$

Следующим по важности выборочным моментом является выборочный центральный момент 2-го порядка $\bar{\mu}_2$, который называется *выборочной дисперсией* и обозначается S^2 :

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (8.8)$$

Наиболее часто используются две формулы метода моментов.

$$M[X] \approx \bar{x}, \quad (8.9)$$

$$D[X] \approx S^2. \quad (8.10)$$

Сформулируем метод моментов в общем виде.

Пусть $f(x, \theta_1, \theta_2, \dots, \theta_r)$ плотность распределения случайной величины X , где θ_i – неизвестные параметры. Чтобы найти оценки $\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_r$, выражаем первые r начальных или центральных моментов случайной величины X через параметры $\theta_1, \theta_2, \dots, \theta_r$, затем генеральные моменты аппроксимируем соответствующими выборочными. В результате имеем систему из r уравнений с r неизвестными, откуда и получаем $\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_r$.

Пример.

Пусть генеральная случайная величина X имеет показательный закон распределения с плотностью $f(x) = \theta e^{-\theta x}$, $x > 0$, $\theta > 0$. По выборке x_1, x_2, \dots, x_n методом моментов найти оценку параметра θ .

1. Определяем α_1 , используя (8.1):

$$\begin{aligned} \alpha_1 = M[X] &= \int_0^{\infty} x \theta e^{-\theta x} dx = \theta \int_0^{\infty} x e^{-\theta x} dx = \left. \begin{array}{l} u = x, \quad dv = e^{-\theta x} dx, \\ du = dx, \quad v = -\frac{1}{\theta} e^{-\theta x} \end{array} \right| = \\ &= -\theta \frac{1}{\theta} e^{-\theta x} x \Big|_0^{\infty} - \frac{\theta}{\theta^2} e^{-\theta x} \Big|_0^{\infty} = \frac{1}{\theta}. \end{aligned}$$

2. По (8.3) или (8.7) находим выборочный начальный момент 1-го порядка или \bar{x} и составляем выражение вида (8.5) или (8.9):

$$\frac{1}{\theta} \approx \bar{x}.$$

3. Заменяя в п. 2 θ на оценку $\tilde{\theta}$, составим уравнение: $\frac{1}{\tilde{\theta}} = \bar{x}$.

4. Откуда определим оценку параметра $\tilde{\theta}$:

$$\tilde{\theta} = \frac{1}{\bar{x}} = \frac{1}{\frac{1}{n} \sum_{i=1}^n x_i}.$$

8.2.2. Метод наибольшего правдоподобия

Этот метод предложен математиком Фишером в 1912 г.

Пусть $f(x, \theta_1, \theta_2, \dots, \theta_r)$ – плотность распределения генеральной случайной величины X , где $\theta_1, \theta_2, \dots, \theta_r$ – неизвестные параметры. Согласно методу, наилучшими оценками $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$ параметров $\theta_1, \theta_2, \dots, \theta_r$ являются такие, для которых функция правдоподобия L принимает наибольшее значение.

Для непрерывной случайной величины

$$L(\theta_1, \dots, \theta_r, x_1, \dots, x_n) = \prod_{i=1}^n f(x_i, \theta_1, \dots, \theta_r). \quad (8.11)$$

Для дискретной случайной величины

$$L(\theta_1, \dots, \theta_r, x_1, \dots, x_n) = \prod_{i=1}^n P(x_i, \theta_1, \dots, \theta_r). \quad (8.12)$$

Здесь x_1, x_2, \dots, x_n – выборка из генеральной случайной величины X .

Априорные выборочные значения x_1, x_2, \dots, x_n – являются независимыми случайными величинами, закон распределения которых совпадает с законом распределения генеральной случайной величины X . Тогда правую часть (8.11) на основании теоремы умножения законов распределений (см. раздел 3.5) можно рассматривать как плотность распределения вероятности n -мерного вектора (X_1, X_2, \dots, X_n) . Согласно методу, для наилучших оценок $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$ случайный вектор (X_1, X_2, \dots, X_n) будет иметь наибольшую плотность распределения. То есть надо найти такие оценки $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$, для которых функция правдоподобия L – максимальна. Для этого составляют и решают такую систему уравнений:

$$\frac{\partial L}{\partial \theta_i} = 0, \quad (i = \overline{1, r}). \quad (8.13)$$

Так как функция и ее логарифм достигают экстремума в одной точке, то часто для упрощения решения задачи используют логарифмическую функцию правдоподобия. В случае логарифмической функции правдоподобия составляется система следующих уравнений:

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \quad (i = \overline{1, r}). \quad (8.14)$$

Пример.

Пусть генеральная случайная величина X имеет показательный закон распределения с плотностью $f(x, \theta) = \theta e^{-\theta x}$, $x > 0$, $\theta > 0$. По выборке x_1, x_2, \dots, x_n методом наибольшего правдоподобия найти оценку параметра θ .

1. Так как нам необходимо оценить один параметр θ , то надо составить и решить одно уравнение. Найдем функцию правдоподобия, используя (8.11):

$$L(\theta, x_1, \dots, x_n) = \prod_{i=1}^n f(x_i, \theta) = \prod_{i=1}^n \theta e^{-\theta x_i} = \theta^n \prod_{i=1}^n e^{-\theta x_i} = \theta^n e^{-\theta \sum_{i=1}^n x_i}.$$

2. Составим логарифмическую функцию правдоподобия:

$$\ln L = n \ln \theta - \theta \sum_{i=1}^n x_i.$$

3. Для определения максимума логарифмической функции правдоподобия составляем и решаем следующее уравнение:

$$\frac{d \ln(L)}{d \theta} = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0.$$

Откуда оценка $\hat{\theta}$ параметра θ определяется так:

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n x_i}$$

При сравнении это выражение с оценкой $\tilde{\theta}$, полученной по методу моментов (см. раздел 8.1), мы понимаем, что они одинаковы. Методы, рассмотренные нами, как видим, абсолютно разные. Это свидетельствует о их достоверности.

8.3. Свойства оценок

Пусть x_1, x_2, \dots, x_n – выборка из генеральной совокупности. Обозначим оценку параметра θ через $\tilde{\theta}$. Ранее мы показали, что эта оценка определяется с помощью различных методов по полученной выборке x_1, x_2, \dots, x_n , т. е. является функцией от x_1, x_2, \dots, x_n .

$$\tilde{\theta} = \varphi(x_1, x_2, \dots, x_n).$$

Так как любая выборка типа x_1, x_2, \dots, x_n – случайна, то и выборочные функции $\tilde{\theta} = \varphi(x_1, x_2, \dots, x_n)$ – тоже являются случайными. Следовательно, она тоже имеет свои характеристики.

1. Оценка $\tilde{\theta}$ называется **несмещенной**, если ее математическое ожидание совпадает с самим оцениваемым параметром:

$$M[\tilde{\theta}] = \theta.$$

В противном случае оценка называется **смещенной**:

$$M[\tilde{\theta}] \neq \theta.$$

Полную погрешность $\theta - \tilde{\theta}$, возникшую от замены θ на $\tilde{\theta}$, можно представить так:

$$\underbrace{\theta - \tilde{\theta}}_{\text{полная погрешность}} = \underbrace{(\theta - M[\tilde{\theta}])}_{\text{систематическая погрешность}} + \underbrace{(M[\tilde{\theta}] - \tilde{\theta})}_{\text{случайная погрешность}}.$$

Таким образом, если оценка несмещенная, то систематическая погрешность равна нулю, т. е. $M[\tilde{\theta}] = \theta$.

Наиболее опасна систематическая ошибка, если она заранее неизвестна или среднее квадратичное отклонение не очень большое. Среднее значение случайной ошибки $M[M[\tilde{\theta}] - \tilde{\theta}] = 0$.

Мы уже отмечали, что x_1, x_2, \dots, x_n – независимые случайные величины, имеющие тот же закон распределения, что и X , генеральная случайная величина, в частности, выборочное математическое ожидание и дисперсия имеет те же числовые характеристики, т. е. справедливы тождества:

$$\begin{aligned} M[x_i] &\equiv M[X], \\ D[x_i] &\equiv D[X]. \end{aligned} \quad (*)$$

Проверим смещенность оценки математического ожидания выборочной средней \bar{x} . Используя обычные свойства математического ожидания, найдем $M[\bar{x}]$:

$$\begin{aligned} M[\bar{x}] &= M\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n M[x_i] = \left. \begin{array}{l} \text{на основании тождества } (*) \\ \text{имеем } M[x_i] = M[X] \end{array} \right| = \\ &= \frac{1}{n} n M[X] = M[X]. \end{aligned}$$

Обозначим $\theta = M[X]$, $\tilde{\theta} = \bar{x}$; видим, что $M[\tilde{\theta}] = \theta$, значит, *выборочное среднее \bar{x} является несмещенной оценкой математического ожидания.*

Проверим смещенность оценки дисперсии выборочной дисперсией S^2 . Найдем математическое ожидание от выборочной дисперсии:

$$\begin{aligned}
M[S^2] &= M\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right] = M\left[\frac{1}{n} \sum_{i=1}^n \{(x_i - M[X]) - (\bar{x} - M[X])\}^2\right] = \\
&= M\left[\frac{1}{n} \left\{ \sum_{i=1}^n (x_i - M[X])^2 - 2(\bar{x} - M[X]) \sum_{i=1}^n (x_i - M[X]) + n(\bar{x} - M[X])^2 \right\}\right] = \\
&= \left| \sum_{i=1}^n (x_i - M[X]) = \sum_{i=1}^n x_i - \sum_{i=1}^n M[X] = n\bar{x} - nM[X] \right| = \\
&= M\left[\frac{1}{n} \left\{ \sum_{i=1}^n (x_i - M[X])^2 - 2n(\bar{x} - M[X])^2 + n(\bar{x} - M[X])^2 \right\}\right] = \\
&= M\left[\frac{1}{n} \left\{ \sum_{i=1}^n (x_i - M[X])^2 - n(\bar{x} - M[X])^2 \right\}\right] = \\
&= \frac{1}{n} \sum_{i=1}^n M[(x_i - M[X])^2] - M[(\bar{x} - M[X])^2] = \\
&= \frac{1}{n} \sum_{i=1}^n D[x_i] - D[\bar{x}] = \left| \text{применим тождества (*)} \right| = \frac{1}{n} nD[X] - D[\bar{x}], \\
\text{но } D[\bar{x}] &= D\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n D[x_i] = \left| \text{используем (*)} \right| = \frac{1}{n} D[X].
\end{aligned}$$

То есть дисперсия выборочной средней в n раз меньше дисперсии генеральной случайной величины. Тогда

$$M[S^2] = D[X] - \frac{1}{n} D[X] = \left(1 - \frac{1}{n}\right) D[X].$$

Обозначим $\theta = D[X]$, $\tilde{\theta} = S^2$. $M[\tilde{\theta}] \neq \theta$, значит, выборочная дисперсия S^2 является *смещенной* оценкой дисперсии. Можно отметить, что выборочная дисперсия S^2 является асимптотически несмещенной оценкой, т. к. при n , стремящемся к бесконечности, смещение стремится к нулю.

При решении практических задач часто используется *несмещенная* оценка дисперсии – это *модифицированная выборочная дисперсия*:

$$\bar{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Найдем математическое ожидание от \bar{S}^2 :

$$\begin{aligned}
M[\bar{S}^2] &= M\left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2\right] = M\left[\frac{n}{n-1} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right] = M\left[\frac{n}{n-1} S^2\right] = \\
&= \frac{n}{n-1} M[S^2] = \frac{n}{n-1} \left(1 - \frac{1}{n}\right) D[X] = D[X].
\end{aligned}$$

Обозначим $\theta = D[X]$, $\tilde{\theta} = \tilde{S}^2$; как видим, $M[\tilde{\theta}] = \theta$, значит, оценка \tilde{S}^2 уже несмещенная. При малых n этой формулой пользоваться лучше (при $n > 30$ оценки совпадают). На практике используют еще одну *несмещенную* оценку дисперсии – когда известно математическое ожидание:

$$\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2.$$

Найдем $M[\tilde{S}^2]$:

$$\begin{aligned} M[\tilde{S}^2] &= M\left[\frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2\right] = \frac{1}{n} \sum_{i=1}^n M[(x_i - m_x)^2] = \frac{1}{n} \sum_{i=1}^n D[x_i] = \\ &= \left| \begin{array}{l} \text{используем тождества} \\ \text{со звездочкой (*): } D[x_i] = D[X] \end{array} \right| = \frac{1}{n} n D[X] = D[X]. \end{aligned}$$

Обозначим $\theta = D[X]$, $\tilde{\theta} = \tilde{S}^2$. $M[\tilde{\theta}] = \theta$, значит, оценка \tilde{S}^2 несмещенная.

2. Оценка $\tilde{\theta}$ параметра θ называется **состоятельной**, если она сходится по вероятности к параметру θ , т. е. если $\forall \varepsilon > 0$ выполняется:

$$\lim_{n \rightarrow \infty} P(|\tilde{\theta} - \theta| < \varepsilon) = 1. \quad (\alpha)$$

Условие (α) на практике проверить трудно. Поэтому для проверки состоятельности оценок применяют более простые условия:

- а) $\lim_{n \rightarrow \infty} (M[\tilde{\theta}] - \theta) = 0$,
- б) $\lim_{n \rightarrow \infty} (D[\tilde{\theta}]) = 0$.

Как видим, оценка $\tilde{\theta}$ будет состоятельной, если при $n \rightarrow \infty$ смещение устраняется и дисперсия оценки стремится к нулю.

Пример.

Проверим состоятельность оценки математического ожидания выборочной средней \bar{x} . Ранее мы показали, что \bar{x} является несмещенной оценкой математического ожидания, т. е. условие а) выполняется и без вычисления предела. Проверим условие б), найдем $D[\bar{x}]$:

$$\begin{aligned} D[\bar{x}] &= D\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n D[x_i] = \left| \begin{array}{l} \text{применяя} \\ \text{тождества (*)} \end{array} \right| = \\ &= \frac{1}{n^2} \sum_{i=1}^n D[X] = \frac{D[X]}{n}. \end{aligned} \quad (\beta)$$

Видим, что при $n \rightarrow \infty$ предел $\frac{D[X]}{n}$ будет стремиться к нулю, значит условие б) выполняется. Следовательно, \bar{x} является *состоятельной* оценкой математического ожидания.

3. Несмещенная оценка $\tilde{\theta}$ параметра θ называется *эффективной*, если она имеет наименьшую дисперсию среди всех оценок при одном и том же объеме выборки n .

Для определения наименьшей дисперсии эффективной оценки $\tilde{\theta}^*$ параметра θ применяется формула Рао-Крамера:

$$D[\tilde{\theta}^*] = \frac{1}{nM \left[\left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 \right]}, \quad (8.15)$$

где $f(x, \theta)$ – плотность распределения генеральной случайной величины X .

Отметим, если оценка $\tilde{\theta}$ смещенная, то малость ее дисперсии еще не говорит о ее эффективности. Например, если в качестве оценки $\tilde{\theta}$ взять любую постоянную величину c , то ее дисперсия будет равна нулю, а ошибка может быть какой угодно большой.

Пример.

Задана нормальная случайная величина X с плотностью распределения

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

Проверим эффективность оценки математического ожидания выборочной средней \bar{x} .

Найдем дисперсию эффективной оценки параметра m . Обозначим эффективную оценку \tilde{m}^* . Чтобы воспользоваться формулой Рао-Крамера (8.15), вычислим

$$\ln f(x, m, \sigma) = -\ln \sqrt{2\pi}\sigma - \frac{(x-m)^2}{2\sigma^2}.$$

Найдем производную:

$$\frac{\partial \ln f(x, m, \sigma)}{\partial m} = \frac{x-m}{\sigma^2}.$$

Подставим полученное выражение в (8.15):

$$D[\tilde{m}^*] = \frac{1}{nM \left[\left(\frac{x-m}{\sigma^2} \right)^2 \right]} = \frac{\sigma^4}{nM \left[(x-m)^2 \right]} = \frac{\sigma^2}{n} = \frac{D[X]}{n}. \quad (8.16)$$

Ранее мы показали, что такую же дисперсию имеет \bar{x} (см. формулу (β)). Видим, что правые части формул (8.16) и (β) совпадают, следовательно, выборочное среднее \bar{x} является эффективной оценкой параметра m .

Отметим, что оценки, полученные методом наибольшего правдоподобия, являются состоятельными. Если существует эффективная оценка, то метод наибольшего правдоподобия позволяет найти ее, но не всегда оценки, полученные этим методом, являются несмещенными.

Библиотека БГУИР

ГЛАВА 9. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН В МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

В математической статистике наиболее часто применяются такие распределения:

1. Нормальное (Гауссовское) распределение.
2. Распределение Пирсона, распределение χ^2 (хи-квадрат).
3. Распределение Стьюдента (t – распределение).
4. Распределение Фишера (F – распределение).

Нормальный закон распределения мы подробно рассмотрели при изучении раздела 6.5 теории вероятностей и здесь рассматриваться не будет.

Отметим, что в законы распределений математической статистики входит гамма-функция, поэтому необходимо познакомиться с этой функцией и рассмотреть ее свойства.

9.1. Гамма-функция и ее свойства

Гамма-функцией или интегралом Эйлера второго рода называется функция следующего вида:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad (9.1)$$

где α – параметр, от которого зависит значение интеграла.

Свойства гамма-функции:

1. $\Gamma(1) = \Gamma(2) = 1$.

Доказательство:

Подставим $\alpha=1$ в (9.1): $\Gamma(1) = \int_0^{\infty} x^{1-1} e^{-x} dx = -e^{-x} \Big|_0^{\infty} = 1$.

2. $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ при $\alpha > 0$.

Доказательство:

Вычислим интеграл в (9.1), используя интегрирование по частям:

$$\Gamma(\alpha+1) = \int_0^{\infty} x^{\alpha} e^{-x} dx = -x^{\alpha} e^{-x} \Big|_0^{\infty} + \alpha \int_0^{\infty} x^{\alpha-1} e^{-x} dx = \alpha\Gamma(\alpha).$$

$$3. \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Доказательство:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} x^{-\frac{1}{2}} e^{-x} dx = \left. \begin{array}{l} \text{сделаем замену} \\ \text{переменных:} \\ x = \frac{u^2}{2} \\ dx = u du \end{array} \right| = \sqrt{2} \int_0^{\infty} u^{-1} e^{-\frac{u^2}{2}} u du = \sqrt{2} \int_0^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{\pi}.$$

Значит, если значение α кратно $\frac{1}{2}$, то $\Gamma(\alpha)$ легко вычисляется с использованием свойства 2:

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(\frac{1}{2} + 1\right) = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1}{2} \sqrt{\pi},$$

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

Для целых α – гамма-функция это факториал:

$$\Gamma(\alpha) = (\alpha - 1)!.$$

Например,

$$\Gamma(3) = 2! = 2, \quad \Gamma(4) = 3! = 6, \quad \Gamma(5) = 4! = 24.$$

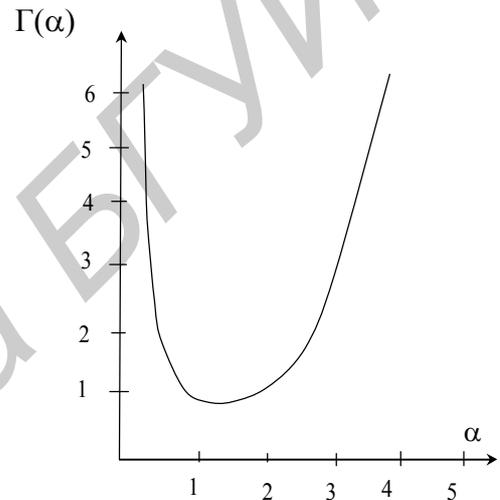


Рис. 9.1

Отметим, что смысл гамма-функции – распространение понятия факториала на нецелые значения. На рис. 9.1 приведен график гамма-функции.

9.2. Распределение χ^2 (хи-квадрат)

Случайная величина имеет закон распределения χ^2 , если она определяется так:

$$\chi^2 = \sum_{k=1}^n X_k^2, \quad (9.2)$$

где X_1, \dots, X_n – независимые нормированные нормальные случайные величины,

т. е. $X_k = \frac{x_k - m_x}{\sigma_x}$ с плотностью распределения $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Распределение случайной величины, определенной по формуле (9.2), называется *распределением Пирсона*.

Покажем, что плотность распределения случайной величины χ^2 , формула (9.2), определяется следующим равенством:

$$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}. \quad (9.3)$$

Здесь для краткости записи $X = \chi^2$;

$$0 \leq X \leq \infty;$$

$$\Gamma(\alpha) = \int_0^{\infty} U^{\alpha-1} e^{-U} dU \text{ – гамма-функция.}$$

Для доказательства используем аппарат характеристических функций. Найдем характеристическую функцию случайной величины X_k^2 , которая входит в формулу (9.2), учитывая, что X_k^2 имеет нормированное нормальное распределение:

$$g_0(t) = M \left[e^{itX_k^2} \right] = \int_{-\infty}^{\infty} e^{itx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}(1-2it)} dx.$$

Используем подстановку $z = x\sqrt{1-2it}$, тогда

$$g_0(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \frac{dz}{\sqrt{1-2it}} = \frac{1}{\sqrt{2\pi}\sqrt{1-2it}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \left| \begin{array}{l} \text{это интеграл} \\ \text{Пуассона} = \sqrt{2\pi} \end{array} \right| = (1-2it)^{-\frac{1}{2}}.$$

Согласно 5-му свойству характеристической функции (для суммы независимых случайных величин) найдем характеристическую функцию $g_1(t)$ случайной величины χ^2 :

$$g_1(t) = M \left[e^{it\chi^2} \right] = M \left[e^{it \sum_{k=1}^n X_k^2} \right] = M \prod_{k=1}^n e^{itX_k^2} = \prod_{k=1}^n g_{0k}(t) = (g_0(t))^n = (1-2it)^{-\frac{n}{2}}.$$

Найдем характеристическую функцию $g_2(t)$ случайной величины χ^2 , плотность распределения которой определяется по формуле (9.3):

$$\begin{aligned} g_2(t) &= \int_{-\infty}^{\infty} e^{itx} f(x) dx = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} e^{itx} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{n}{2}-1} e^{-\frac{x}{2}(1-2it)} dx = \\ &= \left| \begin{array}{l} U = \frac{x}{2}(1-2it) \\ dU = \frac{1}{2} dx(1-2it) \\ x = \frac{2u}{1-2it} \end{array} \right| = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} \left(\frac{2u}{1-2it} \right)^{\frac{n}{2}-1} e^{-u} \frac{2du}{(1-2it)} = \frac{2^{\frac{n}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) (1-2it)^{\frac{n}{2}}} \times \\ &\times \int_0^{\infty} U^{\frac{n}{2}-1} e^{-U} dU = \left| \text{м. к. } \int_0^{\infty} U^{\frac{n}{2}-1} e^{-U} dU = \Gamma\left(\frac{n}{2}\right) \right| = (1-2it)^{-\frac{n}{2}}. \end{aligned}$$

При сравнении правых частей характеристических функций $g_1(t)$ и $g_2(t)$ мы увидим, что они совпадают. Значит случайная величина, определяемая по формуле (9.2), действительно имеет плотность распределения вероятностей, определяемую формулой (9.3).

Найдем математическое ожидание и дисперсию случайной величины χ^2 . Продифференцируем 2 раза $g_2(t)$:

$$g_2'(t) = -\frac{n}{2}(1-2it)^{-\frac{n}{2}-1}(-2i),$$

$$g_2''(t) = -\frac{n}{2}(-2i)\left(-\frac{n}{2}-1\right)(1-2it)^{-\frac{n}{2}-1}(-2i).$$

Значения производных при $t=0$:

$$g_2'(0) = in,$$

$$g_2''(0) = 4i^2 \frac{n}{2} \left(\frac{n}{2} + 1\right) = i^2 (n^2 + 2n),$$

$$\alpha_1 = n, \quad \alpha_2 = n(n+2),$$

$$M[\chi^2] = n,$$

$$D[\chi^2] = \alpha_2 - \alpha_1^2 = n^2 + 2n - n^2 = 2n.$$

Таким образом, математическое ожидание χ^2 равно числу степеней свободы $\nu = n$, а дисперсия – удвоенному числу степеней свободы. *Числом степеней свободы называется параметр ν , равный числу независимых случайных величин в (9.2) и который записывают $\nu = n$.*

С ростом n распределение становится симметричным относительно $\nu = n - 2$, т. к. с увеличением n по центральной теореме закон распределения должен стремиться к нормальному закону (рис. 9.2).

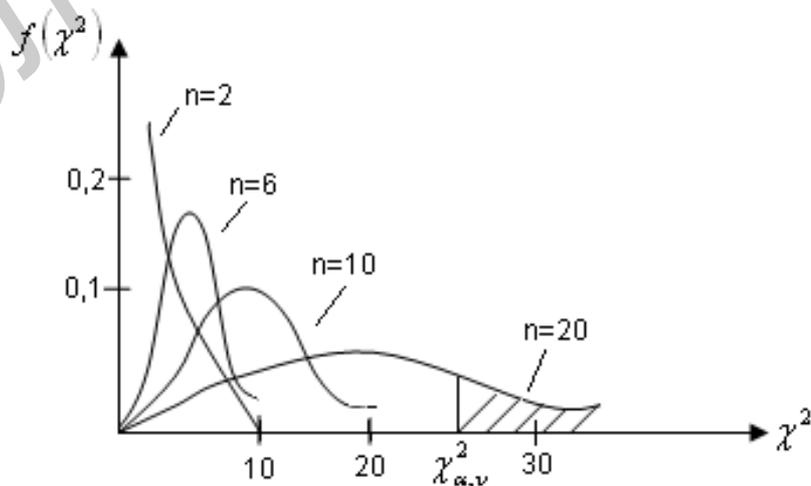


Рис. 9.2

При $n \geq 25$ закон распределения χ^2 практически совпадает с нормальным законом.

Квантилем $\chi_{\alpha, \nu}^2$ (где α – заданный уровень вероятности, ν – число степеней свободы) называется такое значение $\chi^2 = \chi_{\alpha, \nu}^2$, при котором

$$P(\chi^2 > \chi_{\alpha, \nu}^2) = \int_{\chi_{\alpha, \nu}^2}^{\infty} f(\chi^2) d\chi^2 = \alpha, \quad (9.4)$$

т. е. это то значение χ^2 , при котором площадь заштрихованной фигуры на рис. 9.2 равна α . Для определения квантилей $\chi_{\alpha, \nu}^2$ составлены таблицы хи-квадрат распределения. Чтобы воспользоваться ими, необходимо задать уровень вероятности α и число степеней свободы ν .

9.3. Распределение Стьюдента (t – распределение)

Случайная величина t имеет распределение Стьюдента, если она определяется так:

$$t = X \sqrt{\frac{n}{Y}} = X \frac{1}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}} = \frac{X}{\sqrt{\frac{1}{n} \chi_n^2}}, \quad (9.5)$$

где X – нормированная нормальная случайная величина,

Y – величина χ^2 с $\nu = n$ степенями свободы,

X и Y – независимые случайные величины.

Случайная величина t является функцией нормально распределенных нормированных случайных величин и называется безразмерной дробью Стьюдента. Плотность распределения случайной величины t определяется равенством

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad (9.6)$$

где $-\infty < t < \infty$.

Числовые характеристики случайной величины t :

$$M[t] = 0; \quad D[t] = \frac{n}{n-2}.$$

На рис. 9.3 приведены кривые распределения Стьюдента. Кривые на рис. 9.3 качественно напоминают кривые нормального закона распределения с математическим ожиданием, равным нулю, и при $n \rightarrow \infty$ они стремятся к нормальному закону.

Квантили распределения Стьюдента $t_{\alpha/2, \nu}$ в зависимости от числа степеней ν свободы и заданного уровня вероятности α находятся из уравнения:

$$P(|t| > t_{\alpha/2, \nu}) = 2 \int_{t_{\alpha/2, \nu}}^{\infty} f(t) dt = \alpha.$$

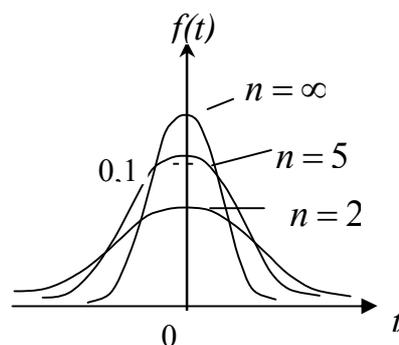


Рис. 9.3

Рис. 9.4 иллюстрирует процесс определения квантилей, т. е. необходимо так выбрать $t_{\alpha/2, \nu}$, чтобы суммарная площадь заштрихованных фигур была равна α .

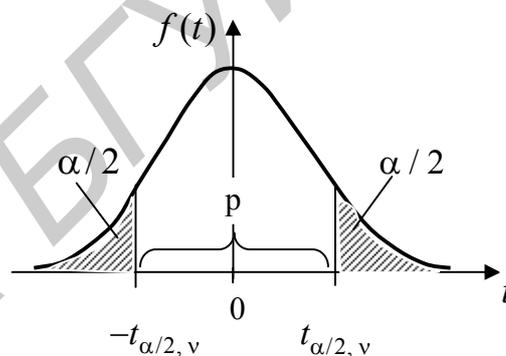


Рис. 9.4

9.4. Распределение Фишера (F-распределение)

Случайная величина F имеет распределение Фишера, если она определяется так:

$$F = \frac{U_1/n_1}{U_2/n_2}, \quad (9.7)$$

где U_1 и U_2 — независимые случайные величины, имеющие распределение χ^2 с n_1 и n_2 степенями свободы, т. е. F можно записать в следующем виде:

$$F = \frac{\frac{1}{n_1} \sum_{i=1}^{n_1} U_{1i}^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} U_{2i}^2}. \quad (9.8)$$

Безразмерная случайная величина F (9.8) имеет плотность распределения, определяемую следующей формулой:

$$f(F) = \frac{n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right)}{(n_2 + n_1 F)^{\frac{n_1 + n_2}{2}}} F^{\frac{n_1}{2} - 1}. \quad (9.9)$$

Распределение случайной величины F зависит от двух параметров $\nu_1 = n_1$, $\nu_2 = n_2$ – степеней свободы. График плотности распределения случайной величины F для разного числа степеней свободы приведен на рис. 9.5.

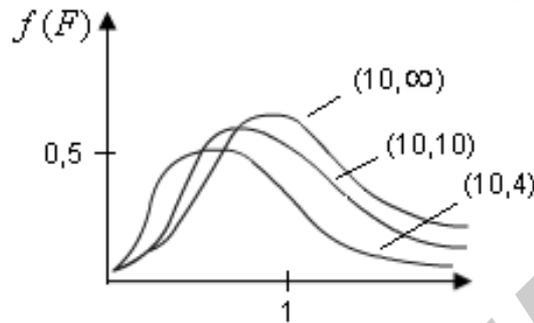


Рис. 9.5

Квантили распределения Фишера F_α для заданного уровня вероятности α и числа степеней свободы ν_1 и ν_2 определяются из условия

$$P(F > F_\alpha) = \int_{F_\alpha}^{\infty} f(F) dF = \alpha.$$

На рис. 9.6 показано, что надо так выбрать F_α , чтобы площадь заштрихованной фигуры была равна заданной вероятности α .

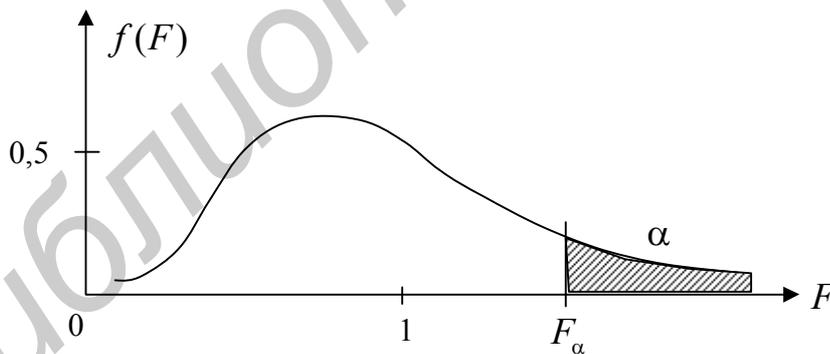


Рис. 9.6

Как правило, квантили F_α находят по таблицам распределения Фишера и для их определения необходимо задать три параметра: уровень вероятности α и число степеней свободы ν_1 и ν_2 .

ГЛАВА 10. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЙ

10.1. Доверительный интервал, доверительная вероятность

Точечная оценка неизвестного параметра θ , найденная по выборке объема n из генеральной совокупности, не позволяет непосредственно узнать ошибку, которая получается, когда вместо точного значения неизвестного параметра θ принимается некоторое его приближение (оценка) $\tilde{\theta} = \varphi(x_1, \dots, x_n)$. Поэтому чаще пользуются интервальной оценкой, основанной на определении некоторого интервала, накрывающего неизвестное значение параметра θ с определенной вероятностью. На рис. 10.1 изображен интервал длиной 2ε , внутри которого в любом месте может находиться неизвестное значение параметра θ .

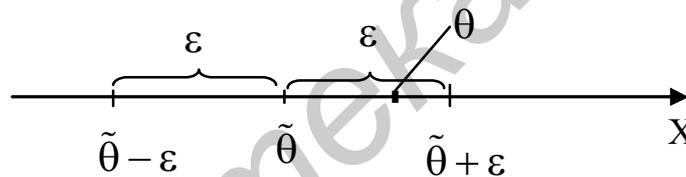


Рис. 10.1

Чем меньше разность $|\theta - \tilde{\theta}|$, тем лучше качество оценки. И если записать $|\theta - \tilde{\theta}| < \varepsilon$, то ε будет характеризовать точность оценки.

Доверительной вероятностью оценки называется вероятность $p = 1 - \alpha$ выполнения неравенства $|\theta - \tilde{\theta}| < \varepsilon$. Доверительную вероятность p обычно задают заранее: 0,9; 0,95; 0,9973. И доверительная вероятность показывает, что с вероятностью p параметр θ будет накрываться данным интервалом

$$p(|\theta - \tilde{\theta}| < \varepsilon) = p = 1 - \alpha \quad \text{или} \\ p(\tilde{\theta} - \varepsilon < \theta < \tilde{\theta} + \varepsilon) = 1 - \alpha. \quad (10.1)$$

Из (10.1) видно, что неизвестный параметр θ находится внутри интервала $]\tilde{\theta} - \varepsilon; \tilde{\theta} + \varepsilon[$.

Доверительным интервалом называется интервал $]\tilde{\theta} - \varepsilon; \tilde{\theta} + \varepsilon[$, накрывающий неизвестный параметр θ с заданной доверительной вероятностью $p = 1 - \alpha$. Длина его (см. рис. 10.1) 2ε . Параметр α – уровень значимости.

10.2. Доверительный интервал для математического ожидания случайной величины X при известной дисперсии (или σ)

Пусть эксперимент E описывается нормальной случайной величиной X .

Плотность распределения $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Предположим, что известна

дисперсия $D[X] = \sigma^2$, а $M[X] = m$ – неизвестна. Тогда точечную оценку математического ожидания можно получить из выборки объемом n : x_1, x_2, \dots, x_n – и она

определится так: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. Рассматривая выборку x_1, \dots, x_n как n независимых

случайных величин, имеющих одно и тоже нормальное распределение, определим числовые характеристики \bar{x} :

$$M[\bar{x}] = M\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} M\left[\sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n M[x_i] = \frac{1}{n} \sum_{i=1}^n m = m,$$

$$D[\bar{x}] = D\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n^2} \sum_{i=1}^n D[x_i] = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n} = \sigma_{\bar{x}}^2,$$

откуда получим

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}. \quad (10.2)$$

Для определения доверительного интервала рассмотрим разность между оценкой и параметром: $\bar{x} - m$. Нормируем ее (сделаем безразмерной), т. е. разделим на $\sqrt{D[\bar{x}]}$ и обозначим как случайную величину U :

$$U = \frac{\bar{x} - m}{\sqrt{D[\bar{x}]}} = \frac{\bar{x} - m}{\sigma_{\bar{x}}} = \frac{\bar{x} - m}{\sigma/\sqrt{n}}. \quad (10.3)$$

Покажем, что случайная величина U имеет нормированный нормальный закон распределения. Найдем ее числовые характеристики:

$$M[U] = M\left[\frac{\bar{x} - m}{\sigma/\sqrt{n}}\right] = \frac{m - m}{\sigma/\sqrt{n}} = 0,$$

$$D[U] = M\left[\left(\frac{\bar{x} - m}{\sigma/\sqrt{n}}\right)^2\right] = \frac{1}{\sigma^2/n} M[(\bar{x} - m)^2] = \frac{\sigma^2/n}{\sigma^2/n} = 1.$$

Таким образом $M[U] = 0$, $D[U] = 1$ – это значит, что U имеет нормированное нормальное распределение, график которого изображен на рис. 10.2.

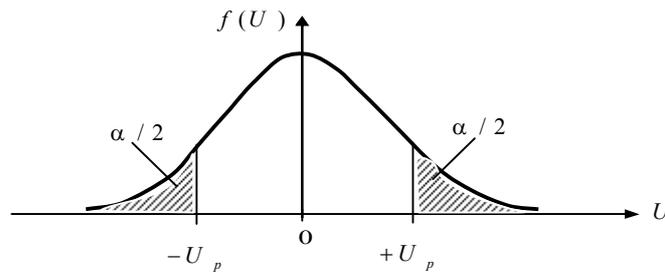


Рис. 10.2

Зная плотность распределения случайной величины U , легко найти вероятность попадания случайной величины U в интервал $[-U_p; U_p]$ (см. рис. 10.2):

$$P(-U_p < U < U_p) = \frac{1}{\sqrt{2\pi}} \int_{-U_p}^{U_p} e^{-\frac{u^2}{2}} du = \frac{2}{\sqrt{2\pi}} \int_0^{U_p} e^{-\frac{u^2}{2}} du = \Phi(U_p). \quad (10.4)$$

Левая часть этого уравнения представляет собой доверительную вероятность p :

$$P\left(-U_p < \frac{\bar{x} - m}{\delta/\sqrt{n}} < U_p\right) = p = 1 - \alpha. \quad (10.5)$$

Тогда из (10.4) и (10.5) следует уравнение

$$\Phi(U_p) = p. \quad (10.6)$$

Решая уравнение (10.6), по таблицам функции Лапласа для заданной доверительной вероятности $p = 1 - \alpha$ можно найти границы доверительного интервала для U , т. е. квантили U_p . Считая, что квантили U_p известны, преобразуем правую часть уравнения (10.5), подставляя в нее (10.3):

$$\begin{aligned} P\left(-U_p \frac{\sigma}{\sqrt{n}} < \bar{x} - m < U_p \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha, \\ P\left(\bar{x} - U_p \frac{\sigma}{\sqrt{n}} < m < \bar{x} + U_p \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha. \end{aligned} \quad (10.7)$$

Считая, что σ^2 — известна, из (10.7) следует, что **доверительный интервал** $[\bar{x} - U_p \frac{\sigma}{\sqrt{n}}; \bar{x} + U_p \frac{\sigma}{\sqrt{n}}]$ **накрывает неизвестное математическое ожидание m с заданной доверительной вероятностью $p = 1 - \alpha$.** Точность оценки математического ожидания или длина доверительного интервала

$$\varepsilon = U_p \frac{\sigma}{\sqrt{n}}. \quad (10.8)$$

Замечания по формуле (10.8):

1) при увеличении объема выборки n из (10.8) видим, что ε уменьшается, значит, уменьшается длина доверительного интервала, а точность оценки увеличивается;

2) увеличение доверительной вероятности $p=1-\alpha$ приводит к увеличению длины доверительного интервала (см. рис. 10.2, где квантили U_p увеличиваются), т. е. ε увеличивается, а точность оценки падает;

3) если задать точность ε и доверительную вероятность $p=1-\alpha$, то можно найти объем выборки, который обеспечит заданную точность:

$$n = \frac{U_p^2 \sigma^2}{\varepsilon^2}. \quad (10.9)$$

Пример.

Сколько конденсаторов одного номинала надо измерить, чтобы с вероятностью 0,95 можно было утверждать, что мы с точностью 1 % определили их среднее значение – математическое ожидание.

Обозначим $\varepsilon = 0,01$, $p = 0,95$; по таблицам функции Лапласа найдем квантиль для заданной доверительной вероятности 0,95: $U_p = 1,96$. Для проведения расчетов положим $\sigma = 0,05$. Подставляя эти значения в (10.9), получим

$$n = \frac{1,96^2 \cdot 0,05^2}{0,01^2} \approx 96.$$

10.3. Доверительный интервал для математического ожидания нормальной случайной величины X при неизвестной дисперсии или σ

Пусть эксперимент описывается случайной величиной X с нормальным распределением с неизвестными параметрами m и σ . Для определения точечных оценок этих параметров из генеральной совокупности извлечена выборка x_1, \dots, x_n объемом n . Тогда точечные оценки этих параметров определяются так:

$$\tilde{m} = \bar{x}, \quad \tilde{\sigma} = \bar{S} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Здесь использовали для оценки дисперсии \bar{S}^2 – модифицированную выборочную дисперсию, несмещенную оценку. Для построения доверительного интервала рассмотрим разность между оценкой и параметром: $\bar{x} - m$. Нормируем ее, т. е. разделим на $\sqrt{D[\bar{x}]}$ и обозначим результат как случайную величину t . Ранее мы показали, что $D[\bar{x}] = \sigma^2 / n$, но т. к. здесь σ неизвестна, возьмем ее оценку $\tilde{\sigma}$, и тогда $D[\bar{x}] = \tilde{\sigma}^2 / n$. Тогда случайная величина t принимает вид

$$t = \frac{\bar{x} - m}{\bar{S}/\sqrt{n}}. \quad (10.10)$$

Умножим числитель и знаменатель в (10.10) на σ :

$$t = \frac{\bar{x} - m}{\frac{1}{\sqrt{n}} \frac{\sigma}{\sigma} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}} = \frac{\bar{x} - m}{\frac{\sigma}{\sqrt{n}} \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma}\right)^2}} = \frac{X}{\sqrt{\frac{1}{n-1} \chi_{n-1}^2}}. \quad (10.11)$$

Здесь X – нормированная нормальная случайная величина, знаменатель – распределение χ^2 с $n-1$ степенями свободы. Поэтому, согласно определению (см. раздел 9.3, формула (9.5)), можно утверждать, что случайные величины t , определяемые по формулам (10.10) и (10.11), имеют закон распределения Стьюдента с $n-1$ степенями свободы.

Зная закон распределения случайной величины t и задавая доверительную вероятность $p = 1 - \alpha$, можно найти вероятность попадания ее в интервал $]-t_p, t_p[$ (рис. 10.3).

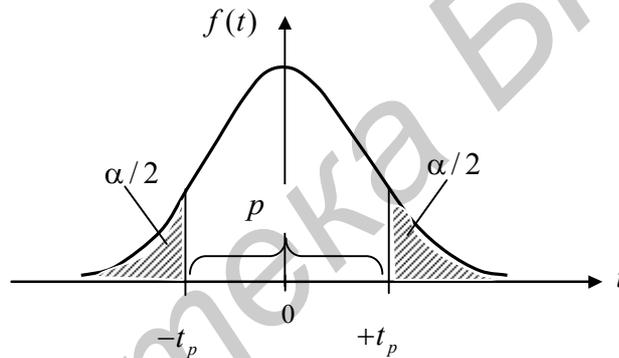


Рис. 10.3

$$P(-t_p < t < t_p) = \int_{-t_p}^{t_p} f(t) dt = p = 1 - \alpha. \quad (10.12)$$

Из таблиц распределений Стьюдента по заданной доверительной вероятности $p = 1 - \alpha$ и числу степеней свободы $n-1$ находим квантили $t_p = t_{\alpha/2, n-1}$, удовлетворяющие условию

$$P(-t_p < t < t_p) = p = 1 - \alpha. \quad (10.13)$$

Подставляя в (10.13) вместо t равенство (10.10), получаем

$$P\left(-t_p < \frac{\bar{x} - m}{\bar{S}/\sqrt{n}} < t_p\right) = p = 1 - \alpha. \quad (10.14)$$

Разрешим неравенство в левой части формулы (10.14) относительно m :

$$P\left(\bar{x} - t_p \frac{\bar{S}}{\sqrt{n}} < m < \bar{x} + t_p \frac{\bar{S}}{\sqrt{n}}\right) = p = 1 - \alpha. \quad (10.15)$$

Отсюда непосредственно следует, что **доверительный интервал** $\left[\bar{x} - t_p \frac{\bar{S}}{\sqrt{n}}, \bar{x} + t_p \frac{\bar{S}}{\sqrt{n}}\right]$ *накрывает неизвестный параметр m – (математическое ожидание) с доверительной вероятностью p .*

Интервал (10.15) несколько шире интервала (10.7), определенного для той же выборки и той же доверительной вероятности. Зато в (10.15) используется меньшая априорная информация – σ знать не надо.

Можно обозначить ширину доверительного интервала или точность через ε_0 , и из (10.15) следует

$$\varepsilon_0 = t_p \frac{\bar{S}}{\sqrt{n}}. \quad (10.16)$$

Все замечания, сделанные по формуле (10.8), справедливы и для формулы (10.16).

Пример.

Даны результаты четырех измерений напряжения сети (значения приведены в [В], $n = 4$):

$$x_1 = 220, \quad x_2 = 226, \quad x_3 = 224, \quad x_4 = 222.$$

Считаем, что X – напряжение сети – является нормальной случайной величиной. Построить доверительный интервал с вероятностью 0,95 для истинного напряжения сети – m .

Найдем точечную оценку m :

$$\tilde{m} = \bar{x} = \frac{1}{n}(220 + 226 + 224 + 222) = 223 \text{ [В]}.$$

Из таблиц распределения Стьюдента для $p = 0,95$; $\alpha = 1 - p = 0,05$; $\nu = n - 1 = 3$ – число степеней свободы; находим квантиль $t_{0,95} = t_{0,025, 3} = 3,18$. Вычислим модифицированную выборочную дисперсию \bar{S}^2 :

$$\bar{S}^2 = \frac{1}{n-1} \sum_{i=1}^4 (x_i - \bar{x})^2 = \frac{1}{3}(9 + 9 + 1 + 1) = \frac{20}{3} [\text{В}^2].$$

Тогда $\bar{S}^2 \approx 2,6$ [В].

Полученные значения подставим в формулу (10.16):

$$\varepsilon_0 = t_p \frac{\bar{S}}{\sqrt{n}} = 3,18 \frac{2,6}{2} \approx 4,2 \text{ [В]}.$$

Найдем левую и правую границы доверительного интервала для m :

$$\bar{x} - \varepsilon_0 = 223 - 4,2 = 218,8 \text{ [В]},$$

$$\bar{x} + \varepsilon_0 = 223 + 4,2 = 227,2 \text{ [В]}.$$

Таким образом, истинное напряжение сети с вероятностью 0,95 покрывается доверительным интервалом]218,8, 227,2[[В].

Найдем минимальное число измерений, чтобы с вероятностью 0,95 точность определения истинного напряжения сети не превышала 0,5 В, т. е. $\varepsilon_0 = 0,5$. Из (10.16) имеем

$$n = \frac{t_p^2 \bar{S}^2}{\varepsilon^2} = \frac{2,6^2 \cdot 3,18^2}{0,5^2} \approx 273 \text{ измерения.}$$

Видим, что число измерений n велико. Следует отметить, что значение квантиля $t_{0,95} = t_{0,025, 3} = 3,18$ зависит от n и при увеличении n будет убывать. При больших n ($n > 100$) значение квантиля стремится к постоянной величине и равно $t_p = t_{0,95} = 1,96$. Тогда после коррекции значения квантиля вычисляем по формуле (10.16) скорректированное значение n :

$$n = \frac{t_p^2 \bar{S}^2}{\varepsilon^2} = \frac{2,6^2 \cdot 1,96^2}{0,5^2} \approx 104 \text{ измерения.}$$

10.4. Доверительный интервал для дисперсии или σ нормальной случайной величины X

Рассмотрим вероятностный эксперимент с нормальной моделью, где параметры m и σ неизвестны. Предположим, что по выборке x_1, \dots, x_n найдены точечные оценки этих параметров:

$$\tilde{m} = \bar{x}, \quad \tilde{\sigma} = \bar{S} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Составим вспомогательную случайную величину

$$V = \frac{(n-1)\bar{S}^2}{\sigma^2}. \quad (10.17)$$

Эта случайная величина имеет распределение χ^2 с $\nu = n-1$ степенями свободы. Покажем это, подставив в (10.17) выражение для \bar{S}^2 :

$$V = \frac{n-1}{\sigma^2} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2 = \chi^2.$$

Это и есть распределение хи-квадрат с $n-1$ степенью свободы. На рис. 10.4 приведен график этого распределения.

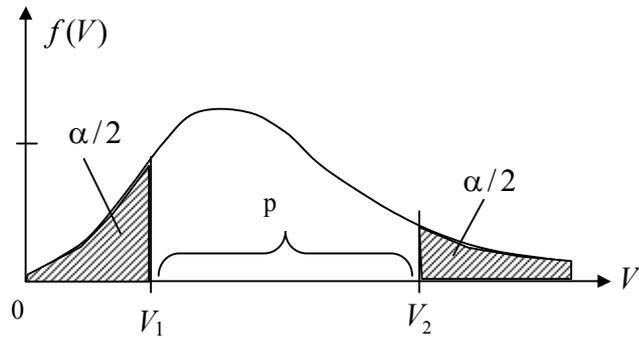


Рис. 10.4

Зная закон распределения случайной величины V , определим вероятность того, что случайная величина V попадет в интервал $]V_1, V_2[$:

$$P(V_1 < V < V_2) = \int_{V_1}^{V_2} f(V) dV = p = 1 - \alpha. \quad (10.18)$$

Здесь $f(V)$ плотность распределения χ^2 с $n-1$ степенями свободы. Из рис. 10.4 видно, что кривая для плотности распределения χ^2 несимметрична относительно центра распределения, поэтому границы доверительного интервала или квантили V_1 и V_2 для данной вероятности p не определяются однозначно. Чтобы избежать неопределенности будем их находить из условия

$$P(V \leq V_1) = \frac{1-p}{2} = \frac{\alpha}{2}, \quad P(V \geq V_2) = \frac{1-p}{2} = \frac{\alpha}{2}. \quad (10.19)$$

Это означает, что площади заштрихованных фигур равны. Задавая доверительную вероятность $p = 1 - \alpha$, по таблицам распределения χ^2 для числа степеней свободы $n-1$, используя условия (10.19), находим квантили V_1 и V_2 .

Считая V_1 , V_2 и p известными, перепишем (10.18) в следующем виде:

$$P(V_1 < V < V_2) = p. \quad (10.20)$$

Подставим в (10.20) значение V , определяемое формулой (10.17):

$$P\left(V_1 < \frac{(n-1)\bar{S}^2}{\sigma^2} < V_2\right) = p. \quad (10.21)$$

Решаем неравенство в левой части (10.21) относительно σ^2 :

$$P\left(\frac{(n-1)\bar{S}^2}{V_2} < \sigma^2 < \frac{(n-1)\bar{S}^2}{V_1}\right) = p. \quad (10.22)$$

Из (10.22) записываем **доверительный интервал** для σ^2 :

$$\left] \frac{(n-1)\bar{S}^2}{V_2}, \frac{(n-1)\bar{S}^2}{V_1} \left[.$$

Для среднего квадратического отклонения σ **доверительный интервал** имеет следующий вид:

$$\left] \bar{S} \sqrt{\frac{n-1}{V_2}}, \bar{S} \sqrt{\frac{n-1}{V_1}} \left[.$$

Можно ввести коэффициенты γ_1 и γ_2 :

$$\gamma_1 = \sqrt{\frac{n-1}{V_2}} = \sqrt{\frac{n-1}{\chi_{\alpha/2, n-1}^2}}, \quad \gamma_2 = \sqrt{\frac{n-1}{V_1}} = \sqrt{\frac{n-1}{\chi_{1-\alpha/2, n-1}^2}}. \quad (10.23)$$

Тогда **доверительный интервал** для σ определится следующим образом:

$$\left] \bar{S}\gamma_1, \bar{S}\gamma_2 \left[.$$

Коэффициенты γ_1 и γ_2 , соответствующие доверительной вероятности $p = 1 - \alpha$ и числу степеней свободы $\nu = n - 1$, находятся по таблицам распределения χ^2 .

Пример.

В предыдущем разделе (10.3) приведен пример для измеренных значений напряжения сети. Продолжим и найдем доверительный интервал для среднего квадратического отклонения σ .

Найдена точечная оценка для σ : $\tilde{\sigma} = \bar{S} = 2,6$. Задавая доверительную вероятность $p = 1 - \alpha = 0,95$, зная число степеней свободы $\nu = n - 1 = 4 - 1 = 3$, по таблицам распределения χ^2 , используя (10.23), находим коэффициенты $\gamma_1 = 0,566$, $\gamma_2 = 3,73$.

Тогда нижняя граница для σ : $\gamma_1 \cdot \bar{S} = 0,566 \cdot 2,6 \approx 1,47$ [В].

Верхняя граница для σ : $\gamma_2 \cdot \bar{S} = 3,73 \cdot 2,6 \approx 9,7$ [В].

И окончательно: $1,47 < \sigma < 9,7$.

ГЛАВА 11. ТЕОРИЯ СТАТИСТИЧЕСКОЙ ПРОВЕРКИ ГИПОТЕЗ

11.1. Основные понятия

Пусть имеется выборка x_1, x_2, \dots, x_n объемом n . Если по выборке построить гистограмму, то можно по виду гистограммы сделать предположение (выдвинуть гипотезу) о виде закона распределения генеральной случайной величины X .

Тогда **нулевой гипотезой** H_0 называют основную (проверяемую) гипотезу, которая утверждает, что различие между сравниваемыми величинами отсутствует.

Альтернативной (конкурирующей, противоположной) гипотезой H называется гипотеза, которая принимается тогда, когда отвергается нулевая.

Целью статистической проверки гипотез является выбор критерия по выборке x_1, \dots, x_n , на основании которого принимается гипотеза H_0 или отклоняется в пользу альтернативной. При этом возможны ошибки двух видов:

1. Отклонение H_0 , когда она на самом деле верна – *ошибка первого рода*. Вероятность этой ошибки обозначается α и называется *уровнем значимости*.
2. Принятие H_0 , когда она на самом деле не верна – *ошибка второго рода*, вероятность ошибки – β .

Чем серьезнее будут последствия ошибки первого рода, тем меньше надо выбирать уровень значимости α . Обычно выбирают $\alpha \approx 0,01 \div 0,05$.

Статистической характеристикой Z гипотезы H_0 называется некоторая случайная величина, определяемая по выборке, для которой известен закон распределения.

Областью отклонения (критической областью) G_0 называется область, при попадании в которую статистической характеристики Z гипотеза H_0 отклоняется.

Дополнение области отклонения до всех возможных значений статистической характеристики Z называется **областью принятия** G .

При попадании статистической характеристики Z в область принятия гипотеза H_0 принимается. На рис. 11.1 изображены область отклонения G_0 и область принятия G . Разделяет их точка на числовой оси z_α .

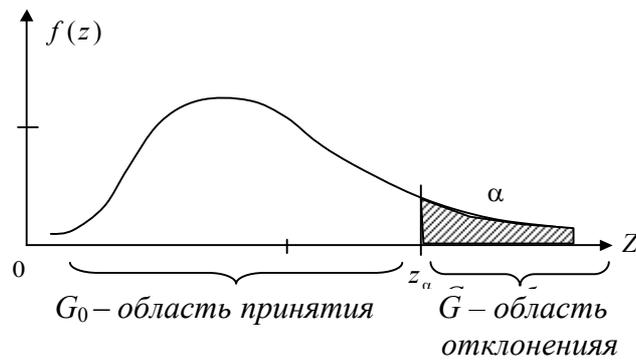


Рис. 11.1

При попадании Z в область принятия гипотеза H_0 принимается. По существу область принятия есть доверительный интервал для статистической характеристики Z с доверительной вероятностью $p = 1 - \alpha$.

Область отклонения G_0 выбирается таким образом, чтобы вероятность попадания в нее статистической характеристики Z при условии, что H_0 верна, равнялась уровню значимости α . То есть область отклонения удовлетворяет условию:

$$P(Z \in G_0 | H_0) = \alpha. \quad (11.1)$$

С другой стороны, для того чтобы уменьшить вероятность ошибки второго рода при выбранном α , область отклонения G_0 , удовлетворяющую условию 1, нужно выбрать таким образом, чтобы вероятность попадания в нее статистической характеристики Z при условии, что верна альтернативная гипотеза H , была максимальной, т. е.

$$P(Z \in G_0 | H) = \gamma \rightarrow \max.$$

Вероятность γ называется **мощностью критерия проверки гипотез**. Так как события $Z \in G_0$ и $Z \notin G_0$ – противоположны, то можно написать

$$P(Z \in G_0 : H) + P(Z \notin G_0 : H) = 1.$$

Таким образом, имеем

$$\gamma + \beta = 1, \quad (11.2)$$

где $\beta = P(Z \notin G_0 : H)$ (β – вероятность совершения ошибки второго рода).

Отметим, что ошибка первого рода существенней, поэтому α мы выбираем, а β – нет (принимаем полученное значение).

Из (11.2) следует, что между γ и β существует простая зависимость и чтобы уменьшить β , надо увеличить мощность критерия γ . Если $\gamma \rightarrow \max$, то $\beta \rightarrow \min$.

Между α и β простой функциональной связи не существует, можно только сказать, что с увеличением одной, другая уменьшается и наоборот.

На рис. 11.2 приведены две кривые плотности распределения: одна кривая $y = f(Z|H_0)$ – когда верна гипотеза H_0 , другая кривая $y = f(Z|H)$ – когда верна альтернативная гипотеза H .

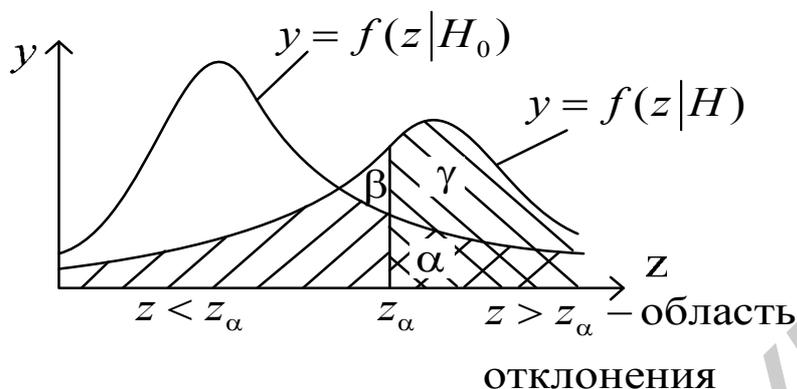


Рис. 11.2

Из рис. 11.2 видно, что при уменьшении α z_α возрастает, область отклонения сужается и, следовательно, уменьшается вероятность отклонения гипотезы H_0 , если она верна. Вместе с тем при сужении области отклонения G_0 расширяется область принятия G и увеличивается вероятность принятия гипотезы H_0 , если она на самом деле не верна. Поэтому нельзя брать α слишком малой.

Гипотезы бывают двух видов – параметрические и непараметрические.

Параметрические гипотезы – это гипотезы о проверке параметров законов распределения.

Непараметрические – это гипотезы о виде закона распределения.

11.2. Проверка гипотезы равенства математических ожиданий при неизвестной дисперсии (критерий Стьюдента)

Пусть X и Y – независимые нормальные случайные величины. Введем обозначения:

$$\begin{aligned} M[X] &= \mu_x, & M[Y] &= \mu_y, \\ D[X] &= \sigma_x^2, & D[Y] &= \sigma_y^2. \end{aligned}$$

Пусть дисперсии этих случайных величин равны и неизвестны:

$$\sigma_x^2 = \sigma_y^2 = \sigma^2,$$

где σ^2 – не предполагается известным.

Пусть даны выборки

$$(x_1, \dots, x_{n_1}) \in X,$$

$$(y_1, \dots, y_{n_2}) \in Y.$$

По выборкам найдем критерий проверки гипотезы H_0 , состоящей в том, что математические ожидания этих случайных величин одинаковы:

$$H_0 : \mu_x = \mu_y .$$

При альтернативной гипотезе $H : \mu_x \neq \mu_y$.

Известно, что случайные величины

$$\chi_1^2 = \frac{n_1 S_1^2}{\sigma^2}, \quad \chi_2^2 = \frac{n_2 S_2^2}{\sigma^2}$$

имеют распределение χ^2 с $n_1 - 1$ и $n_2 - 1$ степенями свободы, где

$$S_1^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i,$$

$$S_2^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} (y_i - \bar{y})^2, \quad \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i.$$

Сумма независимых случайных величин с распределением χ^2 имеет то же распределение χ^2 с суммарным числом степеней свободы:

$$W = \chi_1^2 + \chi_2^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{\sigma^2}. \quad (11.3)$$

Случайная величина W имеет распределение χ^2 с $n_1 + n_2 - 2$ степенями свободы, (этот факт не очевиден, но несложно показать с помощью характеристических функций).

Ранее мы показывали, что несмещенной оценкой математического ожидания является выборочное среднее. Поэтому для проверки гипотезы H_0 возьмем разность между оценками математических ожиданий: $\bar{x} - \bar{y}$. Нормируем эту разность, т. е. сделаем безразмерной. Для этого разделим ее на $D[\bar{x} - \bar{y}]$ и обозначим как U :

$$U = \frac{\bar{x} - \bar{y}}{\sqrt{D[\bar{x} - \bar{y}]}} = \frac{\bar{x} - \bar{y}}{\sqrt{D[\bar{x}] + D[\bar{y}]}} = \left. \begin{array}{l} \text{но: } D[\bar{x}] = \frac{\sigma_x^2}{n_1}, D[\bar{y}] = \frac{\sigma_y^2}{n_2}, \\ \text{учтем, что } \sigma_x^2 = \sigma_y^2 = \sigma^2 \end{array} \right| =$$

$$= \frac{\bar{x} - \bar{y}}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (11.4)$$

Очевидно, что случайная величина U имеет нормальное распределение, т. к. X и Y нормально распределены. Если проверяемая гипотеза H_0 о равенстве математических ожиданий выполняется ($\mu_x = \mu_y$), то имеем:

$$M[U] = M\left[\frac{\bar{x} - \bar{y}}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right] = \frac{M[\bar{x}] - M[\bar{y}]}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\mu_x - \mu_y}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 0.$$

$$\begin{aligned} D[U] &= D\left[\frac{\bar{x} - \bar{y}}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right] = \frac{1}{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \{D[\bar{x}] + D[\bar{y}]\} = \frac{1}{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \cdot \\ &\cdot \left\{D\left[\frac{1}{n_1} \sum_{i=1}^{n_1} x_i\right] + D\left[\frac{1}{n_2} \sum_{i=1}^{n_2} y_i\right]\right\} = \frac{1}{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \left\{\frac{1}{n_1^2} D\left[\sum_{i=1}^{n_1} x_i\right] + \frac{1}{n_2^2} D\left[\sum_{i=1}^{n_2} y_i\right]\right\} = \\ &= \frac{1}{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \left\{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}\right\} = \left| \text{т. к. } \sigma_x^2 = \sigma_y^2 = \sigma^2 \right| = 1. \end{aligned}$$

Таким образом, если гипотеза H_0 верна, то случайная величина U имеет нормированный нормальный закон распределения.

Рассмотрим случайную величину t :

$$t = \frac{\bar{x} - \bar{y}}{S_0 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad (11.5)$$

где $S_0^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$ — объединенная выборочная дисперсия.

Случайную величину t можно представить в следующем виде через ранее введенные U и W :

$$t = U \sqrt{\frac{n_1 + n_2 - 2}{W}}. \quad (11.6)$$

Действительно:

$$t = U \sqrt{\frac{n_1 + n_2 - 2}{W}} = \frac{\bar{x} - \bar{y}}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sqrt{\frac{(n_1 + n_2 - 2)\sigma^2}{n_1 S_1^2 + n_2 S_2^2}} = \frac{\bar{x} - \bar{y}}{S_0 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad (11.7)$$

т. е. правые части (11.5) и (11.6 или 11.7) совпадают.

Но величина t (11.6) имеет распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы. Это следует из того, что U имеет нормированное нормальное распределение при условии, что H_0 – верна. W – имеет распределение χ^2 с $(n_1 + n_2 - 2)$ степенями свободы, кроме того величины U и W независимы. Таким образом, величина t определяется по (11.5) и имеет распределение Стьюдента с $(n_1 + n_2 - 2)$ степенями свободы, если верна проверяемая гипотеза H_0 .

Эту величину t (11.5) примем за статистическую характеристику Z . Проверка гипотезы о равенстве математических ожиданий состоит в следующем.

По таблицам распределения Стьюдента для заданного уровня значимости α или доверительной вероятности $p = 1 - \alpha$ и числу степеней свободы $\nu = n_1 + n_2 - 2$ находим квантиль $z_\alpha = t_{\frac{\alpha}{2}, n_1 + n_2 - 2}$, удовлетворяющий условию (на рис. 11.3 изображена кривая распределения Стьюдента и заштрихована область отклонения G_0):

$$P(\underbrace{|t| > z_\alpha}_{t \in G_0} : H_0) = \alpha.$$

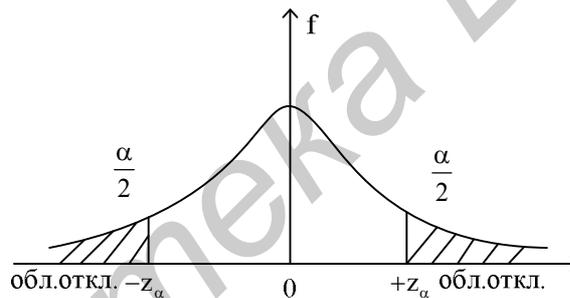


Рис. 11.3

Тогда если фактически найденное по выборкам значение статистической характеристики t (11.5) удовлетворяет условию $\{|t| > z_\alpha\} \in G_0$, то проверяемую гипотезу H_0 о равенстве математических ожиданий отклоняем как несогласующуюся с результатами выборочных данных; при этом вероятность ошибки равна α . Если $|t| < z_\alpha$, то гипотеза H_0 принимается, математические ожидания случайных величин X и Y одинаковы.

11.3. Проверка гипотезы о равенстве дисперсий (критерий Фишера)

Пусть X и Y – нормальные независимые случайные величины. Обозначим их дисперсии:

$$D[X] = \sigma_x^2, \quad D[Y] = \sigma_y^2.$$

По выборкам $x_1, \dots, x_{n_1} \in X$, $y_1, \dots, y_{n_2} \in Y$ найдем критерий проверки гипотезы H_0 состоящей в том, что дисперсии этих случайных величин равны

$$H_0 : \sigma_x^2 = \sigma_y^2 = \sigma^2.$$

При альтернативной гипотезе $H : \sigma_x^2 > \sigma_y^2$.

Такая гипотеза выбирается, например, при $\bar{S}_1^2 > \bar{S}_2^2$, где $\bar{S}_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2$, $\bar{S}_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2$ – модифицированные выборочные дисперсии.

В качестве статистической характеристики возьмем случайную величину

$$F = \frac{\bar{S}_1^2}{\bar{S}_2^2}. \quad (11.8)$$

Если гипотеза H_0 о равенстве дисперсии верна, то случайная величина F имеет распределение Фишера с $[n_1 - 1, n_2 - 1]$ степенями свободы. Покажем это, представляя числитель и знаменатель (11.8) в следующем виде:

$$F = \frac{\bar{S}_1^2}{\bar{S}_2^2} = \frac{\frac{(n_1 - 1)\bar{S}_1^2}{\sigma^2}}{\frac{(n_2 - 1)\bar{S}_2^2}{\sigma^2}} = \frac{n_1 - 1}{n_2 - 1} \cdot \frac{\frac{(n_1 - 1)\bar{S}_1^2}{\sigma^2}}{\frac{(n_2 - 1)\bar{S}_2^2}{\sigma^2}}.$$

Видим, что величина $\frac{(n_1 - 1)\bar{S}_1^2}{\sigma^2}$ имеет распределение χ^2 с $(n_1 - 1)$ степенью свободы, а $\frac{(n_2 - 1)\bar{S}_2^2}{\sigma^2} - \chi^2$ с $(n_2 - 1)$ степенями свободы. Следовательно, согласно определению (см. раздел 9.5, формула (9.7)), случайная величина F имеет распределение Фишера с $[n_1 - 1, n_2 - 1]$ степенями свободы.

Проверка гипотезы H_0 состоит в следующем:

Из таблицы распределения Фишера по выбранному уровню значимости α и числу степеней свободы $[n_1 - 1, n_2 - 1]$ находим квантиль F_α , который удовлетворяет

условию $P(F > F_\alpha : H_0) = \int_{F_\alpha}^{\infty} f(F) dF = \alpha$. На рис. 11.4 изображена кривая распределения Фишера с числом степеней свободы $[n_1 - 1, n_2 - 1]$ и заштрихована область отклонения G_0 , площадь которой области равна α , отмечен квантиль F_α .

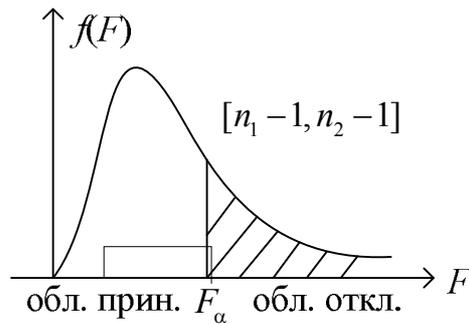


Рис. 11.4

По выборкам, используя (11.8), определяем значение статистической характеристики F . Если фактически вычисленное по формуле (11.8) значение F окажется больше табличного F_α (как видно из рис. 11.4, мы попадаем в область отклонения), то гипотезу о равенстве дисперсий отклоняем как не согласующуюся с выборкой. При этом вероятность ошибки равна α . В противном случае, когда $F < F_\alpha$, принимается гипотеза H_0 , т. е. дисперсии случайных величин X и Y равны.

Пример.

Пусть X – чувствительность телевизоров марки «Горизонт», Y – чувствительность телевизоров марки «Витязь». Проведены выборочные измерения чувствительности телевизоров для $n_1 = 7$ телевизоров марки «Горизонт» и $n_2 = 6$ телевизоров марки «Витязь». Результаты измерений чувствительности в $[\mu\text{В}]$ представлены в таблицах.

x_1	x_2	x_3	x_4	x_5	x_6	x_7
66	72	62	68	76	61	71

y_1	y_2	y_3	y_4	y_5	y_6
62	75	64	70	72	71

Определить лучшую марку телевизора, если лучшим будет тот, у которого чувствительность в $[\mu\text{В}]$ будет меньше.

Найдем по результатам измерений средние значения чувствительности, вычисляя \bar{x} и \bar{y} :

$$\bar{x} = \frac{1}{7} \sum_{i=1}^7 x_i = 68 \text{ [}\mu\text{В]}, \quad \bar{y} = \frac{1}{6} \sum_{i=1}^6 y_i = 69 \text{ [}\mu\text{В]}.$$

Можно ли сказать, что чувствительность телевизоров марки «Горизонт» лучше? Нет, т. к. выборки, выборочные средние \bar{x} , \bar{y} и разность между ними – элементы случайные.

Сначала убедимся в равенстве дисперсий по критерию Фишера – гипотеза $H_0: \sigma_x^2 = \sigma_y^2$.

Вычислим несмещенные оценки дисперсий X и Y :

$$\bar{S}_1^2 = \frac{1}{6} \sum_{i=1}^7 (x_i - 68)^2 \approx 29,67 [\mu\text{B}]^2, \quad \bar{S}_2^2 = \frac{1}{5} \sum_{i=1}^6 (y_i - 69)^2 = 24,8 [\mu\text{B}]^2.$$

Используя (11.8), найдем значение статистической характеристики F :

$$F = \frac{29,67}{24,8} \approx 1,196.$$

По таблицам распределения Фишера для $[6;5]$ степеней свободы, задавая уровень значимости $\alpha = 0,05$, найдем квантиль $F_\alpha = 4,95$. Сравнивая F_α с F , видим, что $1,196 < 4,95$. Значит, гипотеза H_0 принимается, т. е. дисперсии случайных величин X и Y равны.

Теперь проверим гипотезу о равенстве математических ожиданий случайных величин X и Y , применяя критерий Стьюдента.

Гипотеза $H_0: M[X] = M[Y]$, т. е. чувствительность телевизоров марки «Горизонт» и «Витязь» одинакова.

Найдем объединенную выборочную дисперсию:

$$S_0^2 = \frac{\sum_{i=1}^7 (x_i - 68)^2 + \sum_{i=1}^6 (y_i - 69)^2}{7 + 6 - 2} = 27,46 [\mu\text{B}]^2.$$

По формуле (11.5) вычислим статистическую характеристику t :

$$t = \frac{68 - 69}{\sqrt{27,46} \sqrt{\frac{1}{7} + \frac{1}{6}}} = -0,343.$$

Задавая уровень значимости $\alpha = 0,05$ для числа степеней свободы $\nu = 7 + 6 - 2 = 11$, по таблицам распределения Стьюдента находим квантиль $t_{0,025,11} = 2,201$. Сравнивая t с $t_{0,025,11}$, видим, что $|0,343| < 2,201$, значит, гипотезу о равенстве чувствительности телевизоров марки «Горизонт» и «Витязь» принимаем.

11.4. Проверка гипотезы о законе распределения генеральной случайной величины. Критерий Пирсона. (Критерий согласия χ^2)

Пусть задана генеральная случайная величина X и выборка x_1, x_2, \dots, x_n .

Если по выборке построить гистограмму, то по виду гистограммы можно выдвинуть гипотезу о виде закона распределения генеральной случайной величины X . Тогда в качестве нулевой гипотезы H_0 будет предположение, что случайная величина X имеет плотность распределения $f_0(x)$:

$$H_0: f(x) = f_0(x).$$

При альтернативной гипотезе $H: f(x) \neq f_0(x)$.

Обычно для построения гистограммы равноинтервальным способом разбивают весь диапазон выборочных значений случайной величины X на k одинаковых интервалов. Если m_i – число выборочных значений, попавших в i -й интервал, то $\sum_{i=1}^n m_i = n$ – объем выборки. Введем случайную величину $Z_i = \frac{m_i}{n}$ – относительную частоту попадания случайной величины X в i -й интервал. Теоретическая вероятность p_i попадания значений случайной величины X в i -й интервал может быть определена как $p_i = \int_{X \in \Delta x_i} f_0(x) dx$, где $\Delta x_i = a_i - b_i$ – длина i -го интервала, a_i ; b_i – границы i -го интервала.

Рассмотрим событие, состоящее в том, что случайная величина X попадет в интервал Δx_i m_i раз. Тогда введем случайную величину Y , равную числу попаданий случайной величины в i -й интервал ($Y = m_i$). Вероятности возможных ее значений определяются по формуле Бернулли, случайная величина Y имеет биномиальный закон распределения, и ее числовые характеристики имеют вид $M[Y] = np$, $D[Y] = np_i q_i$, $q_i = 1 - p_i$.

Для введенной ранее случайной величины $Z_i = \frac{m_i}{n}$ определим числовые характеристики:

$$M[Z_i] = M\left[\frac{m_i}{n}\right] = \frac{1}{n} M[m_i] = \frac{1}{n} M[Y] = \frac{1}{n} np_i = p_i,$$

$$D[Z_i] = D\left[\frac{m_i}{n}\right] = \frac{1}{n^2} D[m_i] = \frac{1}{n^2} D[Y] = \frac{1}{n^2} np_i q_i = \frac{p_i q_i}{n}.$$

Проведем нормировку случайной величины Z_i , для этого мы ее центрируем, сделаем безразмерной, разделив на $\sqrt{D[Z_i]}$, и обозначим W_i :

$$W_i = \frac{Z_i - M[Z_i]}{\sqrt{D[Z_i]}} = \frac{\frac{m_i}{n} - p_i}{\sqrt{p_i q_i / n}}. \quad (11.9)$$

Эта величина распределена по биномиальному закону, т. к. в нее входит случайная величина Y . образуем сумму квадратов случайных величин W_i :

$$\sum_{i=1}^k W_i^2 = \sum_{i=1}^k \left(\frac{\frac{m_i}{n} - p_i}{\sqrt{p_i q_i / n}} \right)^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i q_i} = \left. \begin{array}{l} \text{при } n \rightarrow \infty, k \rightarrow \infty \text{ биномиальный} \\ \text{закон распределения по централь-} \\ \text{ной предельной теореме} \\ \text{переходит в нормальный закон,} \\ \text{(учтем: если } p_i \rightarrow 0, \text{ то } q_i \rightarrow 1) \end{array} \right| =$$

$$= \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}.$$

Сумма квадратов нормированных нормальных случайных величин (как было показано ранее) имеет распределение χ^2 , обозначим

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}. \quad (11.10)$$

Эта случайная величина имеет закон распределения χ^2 с числом степеней свободы

$$v = k - r - 1, \quad (11.11)$$

где r – число параметров закона распределения, оцениваемых по выборочным данным.

Анализируя правые части формул (11.9) и (11.10), можно отметить, что в критерии согласия χ^2 фактически сравниваются эмпирические и теоретические частоты распределения.

Проверка гипотезы состоит в следующем. Задаем уровень значимости α . По таблицам χ^2 – распределения для заданных α и числу степеней свободы $v = k - r - 1$ находим квантиль $\chi_{\alpha, v}^2$, удовлетворяющий условию $P(\chi^2 \geq \chi_{\alpha, v}^2 : H_0) = \alpha$. По формуле (11.10) вычисляем значение χ^2 . Сравнивая рассчитанное значение χ^2 с квантилем $\chi_{\alpha, v}^2$, найденным по таблицам, принимаем одно из двух решений:

1. Если $\chi^2 > \chi_{\alpha, v}^2$, то нулевая гипотеза H_0 отвергается в пользу альтернативной H , т. е. $f(x)$ не согласуется с результатами эксперимента.

2. Если $\chi^2 < \chi_{\alpha, v}^2$, то H_0 принимается, т. е. $f(x)$ согласуется с экспериментальными данными, закон распределения $f_0(x)$ подтверждается. При этом вероятность ошибки равна α .

11.5. Критерий Романовского

Рассмотрим неравенство

$$\frac{|\chi^2 - v|}{\sqrt{2v}} \geq 3, \quad (11.12)$$

где χ^2 вычисляется по формуле (11.10);
 $v = k - r - 1$ по (11.11).

Проверка гипотезы состоит в следующем: если это неравенство выполняется (≥ 3), то расхождение теоретических и экспериментальных данных неслучайно, т. е. закон распределения не подтверждается, гипотеза H_0 отклоняется. В противном случае гипотеза H_0 подтверждается, действительно случайная величина X имеет плотность распределения $f_0(x)$. Этот критерий хорош тем, что для проверки гипотезы не требуются таблицы χ^2 – распределения.

11.6. Критерий согласия Колмогорова

В критерии согласия А. Н. Колмогорова проводится сравнение эмпирической и теоретической функций распределения. Укажем этапы проверки гипотез этим критерием.

1. По выборке x_1, x_2, \dots, x_n строится вариационный ряд и график эмпирической функции распределения.

2. По виду графика функции распределения выдвигается гипотеза о виде закона распределения генеральной случайной величины X . Тогда в качестве нулевой гипотезы H_0 будет предположение, что генеральная случайная величина X имеет функцию распределения $F_0(x)$:

$$H_0 : F(x) = F_0(x).$$

При альтернативной гипотезе $H : F(x) \neq F_0(x)$.

3. По выборке x_1, x_2, \dots, x_n находят точечные оценки параметров теоретической функции распределения $F_0(x)$, используя метод моментов или метод наибольшего правдоподобия.

4. На графике эмпирической функции распределения строится график теоретической функции распределения $F_0(x)$.

5. Путем сравнения графиков вычисляется максимальное значение модуля отклонения значений эмпирической функции распределения от теоретической функции распределения $F_0(x)$:

$$\Delta F = \max_x |F^*(x) - F(x)|.$$

6. Рассчитывают значение λ -критерия Колмогорова:

$$\lambda = \Delta F \sqrt{n} = \max_x |F^*(x) - F(x)| \sqrt{n}. \quad (11.13)$$

7. Задавая уровень значимости α , определяем квантиль λ_α из условия

$$P(\lambda \geq \lambda_\alpha) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda_\alpha^2} = \alpha.$$

Отметим, что самостоятельно решать это уравнение не надо, поскольку составлены таблицы квантилей распределения Колмогорова, из которых по заданному уровню значимости α определяем квантиль λ_α .

Сравнивая значение λ , рассчитанное по формуле (11.13) с квантилем λ_α , делаем следующие выводы:

а) если $\lambda \geq \lambda_\alpha$, то гипотеза H_0 отклоняется;

б) если $\lambda < \lambda_\alpha$, то гипотеза H_0 принимается, закон распределения подтверждается, т. е. действительно генеральная случайная величина X имеет функцию распределения $F_0(x)$.

Следует отметить, что критерий Колмогорова применяется тогда, когда полностью известен закон распределения функции распределения $F(x)$ и значения ее параметров. При решении практических задач это не всегда удается выполнить. Для этого прибегают к некоторым дополнительным исследованиям: применяют вероятностные бумаги, строят гистограммы и т. д. Это помогает правильно подобрать теоретический закон распределения для функции распределения $F(x)$. Но в этом случае неизвестны ее параметры. И если их оценивать по этой же выборке, то это может привести к ошибочным выводам в отношении принятой гипотезы. В этом случае следует использовать другие критерии согласия, например χ^2 .

Пример.

Проведено 100 измерений расстояния радиодальномером до цели. Результаты представлены в виде статистического ряда ($[x_i, x_{i+1}[$ – границы интервалов в [км], m_i – число выборочных значений, попавших в i -й интервал). Оценить закон распределения ошибки измерения дальности радиодальномером.

$[x_i, x_{i+1}[$ [км]	m_i
$-\infty - 450$	0
450 – 500	12
500 – 550	15
550 – 600	14
600 – 650	15
650 – 700	13
700 – 750	16
750 – 800	15

Занесем в таблицу значения относительных частот $v_i = \frac{m_i}{n}$.

Анализ значений относительных частот позволяет выдвинуть гипотезу о равномерном законе распределения. Теоретическая функция распределения для этого закона имеет вид

$$F_0(x) = \frac{x-a}{b-a}.$$

Принимаем $a = 450$, $b = 800$. Полагая $x = x_{i+1}$ для каждого интервала, рассчитываем $F_0(x)$ в этих точках и заносим результат в таблицу. Зная m_i , рассчитываем эмпирическую функцию распределения $F^*(x)$ в точках x_{i+1} для каждого интервала: $F^*(x) = \frac{n_x}{n}$, где n_x – число значений x_i меньших заданного x , n – объем выборки. Рассчитаем разность: $\Delta = |F^*(x) - F_0(x)|$. Данные заносим в таблицу.

$[x_i, x_{i+1}[$ [км]	m_i	v_i	$F_0(x)$	$F^*(x)$	Δ
$-\infty - 450$	0	0	0	0	
450 – 500	12	0,12	0,14 3	0,12	0,02 3
500 – 550	15	0,15	0,28 6	0,27	0,01 6
550 – 600	14	0,14	0,42 9	0,41	0,01 9
600 – 650	15	0,15	0,57 1	0,56	0,01 1
650 – 700	13	0,13	0,71 4	0,69	0,02 4
700 – 750	16	0,16	0,85 7	0,85	0,00 7
750 – 800	15	0,15	1	1	0

Вычисляем критерий Колмогорова по формуле (11.13), учитывая, что из таблицы $\Delta F = \max_x |F^*(x) - F(x)| = 0,024$, тогда $\lambda = 0,024 \cdot \sqrt{n} = 0,24$. Задавая уровень значимости $\alpha = 0,05$, по таблице квантилей Колмогорова находим квантиль $\lambda_\alpha = 1,358$. Поскольку $0,24 \leq 1,358$, то гипотеза H_0 принимается, т. е. действительно генеральная случайная величина X имеет функцию распределения $F_0(x)$ с равномерным законом распределения.

ГЛАВА 12. ЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ

12.1. Уравнение линейной регрессии

Регрессия – это оценка зависимости одной случайной величины от другой случайной величины.

Уравнением регрессии Y на X называется условное математическое ожидание составляющей Y двумерной случайной величины (X, Y) , вычисленное при условии, что составляющая X приняла определенное значение $X = x$:

$$M[Y / X = x] = \int_{-\infty}^{\infty} yf(y/x)dy = \varphi(x), \quad (12.1)$$

где $f(y/x) = \frac{f(x, y)}{f_1(x)}$ – условная плотность распределения.

Функцию $\varphi(x)$ называют *модельной функцией регрессии Y на X* , а ее график – *модельной линией регрессии Y на X* . Уравнение (12.1) называется *уравнением регрессии 1-го рода*. Функцией $\varphi(x)$ может представляться полином k -ой степени $\varphi(x) = \alpha + \beta x + \gamma x^2 + \delta x^3 + \dots$, где $\alpha, \beta, \gamma, \delta, \dots$ – коэффициенты уравнения регрессии.

Пусть заданы две генеральные случайные величины X и Y и выборочные пары их значений: $\{(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)\}$. Если эти выборочные значения нанести на плоскость в декартовой системе координат X, Y , то получим диаграмму в виде точек (диаграмму рассеивания), которая называется *корреляционным полем*.

Пример.

Пусть изучаем зависимость веса человека от его роста для определенной группы людей. Для этого, например, в группе студентов проводим измерения веса, пусть это будет случайная величина Y , и роста – случайная величина X для каждого студента. Результаты занесли в таблицу в виде выборочных пар их значений: $\{(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)\}$ и нанесли на плоскость в системе координат X, Y . В результате получим корреляционное поле (рис. 12.1). Изобразим на рисунке предполагаемую теоретическую зависимость между Y и X в виде жирной линии – это и есть модельная линия регрессии Y на X . Она, допустим, описывается определенной аналитической зависимостью $Y = \varphi(X)$, т. е. модельной функцией регрессии Y на X .

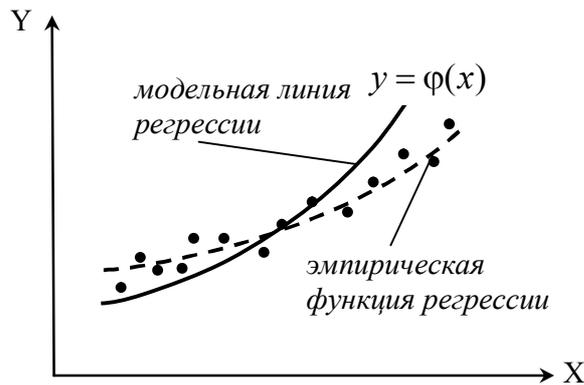


Рис. 12.1

Аппроксимируем корреляционное поле (см. на рис. 12.1) пунктирной линией – это будет эмпирическая линия регрессии, которая может описываться несколько другой аналитической зависимостью. Понятно, что вид эмпирической линии регрессии (зависимость веса человека от его роста) зависит от многих факторов: возраста, национальности, пола и т. д. Сравнение модельной и эмпирической линий регрессии позволяет выявить справедливость наших теоретических предположений.

Рассмотрим линейную регрессию. Предположим, что между X и Y существует линейная зависимость: $y = \alpha + \beta x$. Допустим, что при любом значении X мы можем измерить значение Y с некоторой ошибкой ε . Тогда выборочное значение y_i можно представить в следующем виде:

$$y_i = \alpha + \beta(x_i - \bar{x}) + \varepsilon_i, \quad (12.2)$$

где $\alpha + \beta(x_i - \bar{x})$ – точная линейная зависимость,
 ε_i – ошибка.

Будем предполагать, что ε_i , величина ошибки, – это случайная величина с нормальным законом распределения с $M[\varepsilon_i] = 0$ и $D[\varepsilon_i] = \sigma^2$; обозначим $\tilde{x} = x_i - \bar{x}$.

Исходя из выборочных значений можно каким-либо методом найти точечные оценки коэффициентов α и β уравнения регрессии (12.2).

Уравнение, в которое входят оценки $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}, \dots$ коэффициентов уравнения регрессии $\alpha, \beta, \gamma, \delta, \dots$ и которое является приближенным выражением модельной функции регрессии Y на X , называется **эмпирической функцией регрессии** или уравнением регрессии 2-го рода: $Y = \hat{\alpha} + \hat{\beta}(\hat{x}) + \hat{\gamma}\hat{x}^2 + \hat{\delta}\hat{x}^3 \dots$

Сформулируем суть регрессионного анализа.

Регрессионный анализ – это анализ функций регрессий первого и второго рода, состоящий в следующем:

1. Нахождение точечных и интервальных оценок параметров функции регрессии 1-го рода.
2. Осуществление точечного и интервального оценивания условных математических ожиданий, необходимого для предсказания средних значений од-

ной случайной величины, соответствующих определенным фиксированным значениям другой случайной величины.

3. Проверка согласованности найденной эмпирической функции регрессии с экспериментальными данными.

Для определения точечных оценок параметров функции регрессии чаще используется метод наименьших квадратов.

12.2. Метод наименьших квадратов

Метод наименьших квадратов (МНК) позволяет так выбрать параметры $\hat{\alpha}, \hat{\beta}, \hat{\gamma} \dots$ эмпирической функции регрессии, что она будет наилучшей оценкой модельной функции регрессии в том смысле, что сумма квадратов отклонений наблюдаемых значений переменной Y от соответствующих ординат эмпирической функции регрессии будет наименьшей.

Параметры $\hat{\alpha}, \hat{\beta}, \hat{\gamma} \dots$ эмпирической функции регрессии $y = \hat{\alpha} + \hat{\beta}\tilde{x} + \hat{\gamma}\tilde{x}^2 + \dots$ находятся методом наименьших квадратов из условия

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}\tilde{x}_i - \hat{\gamma}\tilde{x}_i^2 - \dots)^2 = \min. \quad (12.3)$$

МНК обеспечивает наилучшее согласование теоретической зависимости $\varphi(x) = \alpha + \beta x + \gamma x^2 + \delta x^3 + \dots$ и экспериментальных данных.

Подробнее рассмотрим применение МНК для определения точечных оценок параметров функции регрессии. В результате проведения n опытов получаем двумерную выборку $\{(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)\}$. Пусть эмпирическая функция регрессии линейна, т. е. $y = \alpha + \beta\tilde{x}$, (значок \wedge над коэффициентами α, β опускаем для упрощения записи), тогда (12.3) принимает вид

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \alpha - \beta\tilde{x}_i)^2 = \min. \quad (12.4)$$

Для определения минимума функции S необходимо найти производные по интересующим нас параметрам, приравнять их к нулю и решить полученные уравнения. Вычисляем частные производные и приравниваем их к нулю:

$$\begin{cases} \frac{\partial S}{\partial \alpha} = 2 \sum_{i=1}^n (y_i - \alpha - \beta\tilde{x}_i) = 0, \\ \frac{\partial S}{\partial \beta} = 2 \sum_{i=1}^n (y_i - \alpha - \beta\tilde{x}_i)\tilde{x}_i = 0. \end{cases} \quad (12.5)$$

Раскрываем знак суммы:

$$\begin{cases} \sum_{i=1}^n y_i - n\alpha - \beta \sum_{i=1}^n \tilde{x}_i = 0, \\ \sum_{i=1}^n y_i \tilde{x}_i - \alpha \sum_{i=1}^n \tilde{x}_i - \beta \sum_{i=1}^n \tilde{x}_i^2 = 0. \end{cases} \quad (12.6)$$

Учтем, что $\sum_{i=1}^n \tilde{x}_i = \sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = \sum_{i=1}^n x_i - n \frac{1}{n} \sum_{i=1}^n x_i = 0$.

Тогда из (12.6) получаем точечную оценку коэффициента α :

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}. \quad (12.7)$$

Из второго уравнения (12.6) имеем

$$\sum_{i=1}^n y_i \tilde{x}_i - \hat{\alpha} \sum_{i=1}^n \tilde{x}_i - \hat{\beta} \sum_{i=1}^n \tilde{x}_i^2 = 0, \Rightarrow \hat{\beta} = \frac{\sum_{i=1}^n y_i \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2} \quad (12.8)$$

После определения точечных оценок обычно проверяют предположение о виде эмпирической функции регрессии: линейная она или нелинейная. На практике условное математическое ожидание $M[Y/X] = \varphi(X)$ (или уравнение регрессии 1-го рода) часто считают линейной функцией, т. е. $\varphi(x) = \alpha + \beta x$. Это предположение является гипотезой, которая проверяется путем оценки коэффициента корреляции.

Корреляционный анализ – это анализ оценок коэффициента корреляции $\rho = M[(X - m_x)(Y - m_y)] / \sigma_x \sigma_y$, который позволяет ответить на вопрос, существует ли линейная функциональная зависимость между математическим ожиданием случайной величины X и случайной величины Y . Если ответ положительный, то метод корреляционного анализа позволяет измерить степень близости статистической зависимости (т. е. экспериментальных данных) к функциональной.

12.3. Коэффициент корреляции (оценки)

При изучении курса теории вероятностей в разделе 4.2 мы определили ковариацию или корреляционный момент

$$K_{xy} = M[(x - m_x)(y - m_y)] \quad (12.9)$$

и коэффициент корреляции

$$\rho_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} = \frac{M[(X - m_x)(Y - m_y)]}{\sigma_x \sigma_y}. \quad (12.10)$$

Для оценки связи между случайными величинами X и Y применяют точечные оценки $\hat{K}_{xy}, \hat{\rho}_{xy}$, которые называются *эмпирическим корреляционным моментом* и *эмпирическим коэффициентом корреляции*:

$$\hat{K}_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad (12.11)$$

$$\rho_{xy} = \frac{\hat{K}_{xy}}{\bar{\sigma}_x \bar{\sigma}_y} = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum (y_i - \bar{y})^2}} \quad (12.12)$$

Найдем связь между точечной оценкой коэффициента β линейного уравнения регрессии (12.2) и оценкой коэффициента корреляции $\hat{\rho}_{xy}$.

Преобразуем \hat{K}_{xy} (12.11):

$$\begin{aligned} \hat{K}_{xy} &= \frac{1}{n} \sum (x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \bar{y}) = \frac{1}{n} (\sum x_i y_i - \bar{y} \sum x_i - \bar{x} \sum y_i + \sum \bar{x} \bar{y}) = \\ &= \frac{1}{n} (\sum x_i y_i - n \bar{y} \bar{x} - n \bar{x} \bar{y} + n \bar{x} \bar{y}) = \frac{1}{n} \sum x_i y_i - \bar{x} \bar{y} = \overline{xy} - \bar{x} \bar{y}. \end{aligned} \quad (12.13)$$

Рассмотрим оценку $\hat{\beta}$ (12.8). Преобразуем выражение:

$$\begin{aligned} \sum y_i \tilde{x}_i &= \sum y_i (x_i - \bar{x}_i) = \sum y_i x_i - \bar{x} \sum y_i = \sum y_i x_i - n \bar{x} \bar{y} = \sum (y_i x_i - \bar{x} \bar{y}) = \\ &= \left| \sum x_i y_i = n \frac{1}{n} \sum x_i y_i = n \overline{xy} \right| = n \overline{xy} - n \bar{x} \bar{y} = n(\overline{xy} - \bar{x} \bar{y}). \end{aligned}$$

Данное выражение определяется через выборочную дисперсию S_x^2 :

$$\sum_{i=1}^n \tilde{x}_i^2 = n \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = n S_x^2.$$

Тогда оценка $\hat{\beta}$ (12.8) определится через K_{xy}, ρ_{xy} так:

$$\hat{\beta} = \frac{n(\overline{xy} - \bar{x} \bar{y})}{n S_x^2} = \frac{\hat{K}_{xy} S_y}{S_x^2 S_y} = \frac{\hat{K}_{xy}}{S_x S_y} \frac{S_y}{S_x} = \hat{\rho}_{xy} \frac{S_y}{S_x}, \quad (12.14)$$

где $S_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$, $S_y^2 = \frac{1}{n} \sum (y_i - \bar{y})^2$ – выборочные дисперсии.

Уравнение линейной регрессии (12.2) будет иметь вид

$$y_i = \hat{\alpha} + \hat{\rho}_{xy} \frac{S_y}{S_x} \tilde{x}_i. \quad (12.15)$$

Влияние значений коэффициента корреляции ρ_{xy} на функциональную зависимость между X и Y иллюстрирует рис. 12.2.

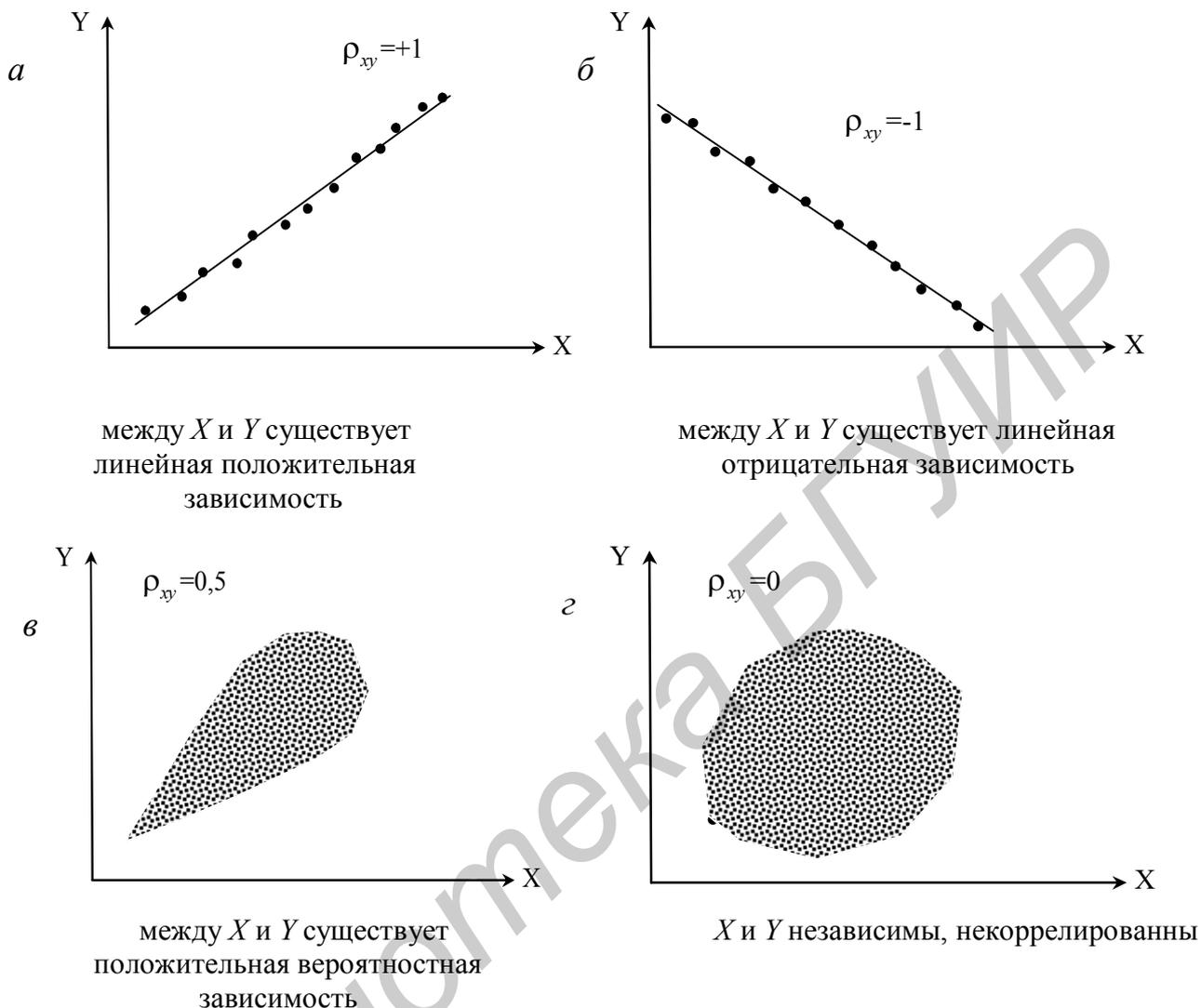


Рис. 12.2

Чем ближе по модулю коэффициент корреляции к единице, тем точнее линейная зависимость между случайными величинами X и Y , (тем ближе к линии регрессии располагаются точки на диаграмме рассеивания). Чем ближе $\hat{\rho}_{xy}$ к нулю, тем слабее эта зависимость. Говорят, что при $\hat{\rho}_{xy} > 0$ (см. рис. 12.2, а, в) между X и Y существует положительная корреляция. При $\hat{\rho}_{xy} < 0$ (см. рис. 12.2, б) между X и Y существует отрицательная корреляция. При $\hat{\rho}_{xy} = 0$ (см. рис. 12.2, г) X и Y некоррелированы.

Поскольку коэффициент корреляции $\hat{\rho}_{xy}$ характеризует степень линейной зависимости, то при $\hat{\rho}_{xy} = 0$ может оказаться, что между X и Y существует нелинейная связь.

12.4. Построение доверительных интервалов для коэффициентов уравнения регрессии

Доверительные интервалы для коэффициентов уравнения регрессии позволяют для заданной доверительной вероятности $p = 1 - \alpha$ выяснить, насколько сильно могут отклоняться коэффициенты эмпирической функции регрессии от соответствующих коэффициентов модельной функции регрессии. Это позволит оценить точность определения коэффициентов уравнения регрессии, корректировать объемы выборки для проведения теоретических исследований.

В разделе 12.2 были определены точечные оценки коэффициентов уравнения регрессии по выборочным данным:

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \hat{\beta} = \frac{\sum_{i=1}^n y_i \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2}. \quad (12.16)$$

Вычислим дисперсии точечных оценок $\hat{\alpha}$ и $\hat{\beta}$, т. е. определим рассеивание случайной величины Y относительно линии регрессии $\hat{\alpha} + \hat{\beta} \tilde{x}$. Ранее мы положили (см. раздел 12.1), что ε_i – величина ошибки, которая распределена по нормальному закону $M[\varepsilon_i] = 0$ и $D[\varepsilon_i] = \sigma^2$; ($\varepsilon_i = y_i - \alpha - \beta \tilde{x}_i$, $\tilde{x}_i = x_i - \bar{x}$). Полагая, что дисперсия Y постоянная и не зависит от X , можно \tilde{x}_i рассматривать в последующих преобразованиях как постоянную величину. Тогда

$$\begin{aligned} D[\hat{\alpha}] &= D\left[\frac{1}{n} \sum_{i=1}^n y_i\right] = D\left[\frac{1}{n} \sum_{i=1}^n (\alpha + \beta \tilde{x}_i + \varepsilon_i)\right] = \frac{1}{n^2} \sum_{i=1}^n D[(\alpha + \beta \tilde{x}_i + \varepsilon_i)] = \\ &= \frac{1}{n^2} \sum_{i=1}^n D[\varepsilon_i] = \frac{\sigma^2}{n}. \end{aligned} \quad (12.17)$$

$$\begin{aligned} D[\hat{\beta}] &= D\left[\frac{\sum_{i=1}^n y_i \tilde{x}_i}{\sum_{i=1}^n \tilde{x}_i^2}\right] = \frac{\sum_{i=1}^n D[y_i \tilde{x}_i]}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \frac{\sum_{i=1}^n \tilde{x}_i^2 D[y_i]}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \frac{\sum_{i=1}^n \tilde{x}_i^2 D[y_i]}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \\ &= \frac{\sum_{i=1}^n \tilde{x}_i^2 D[\varepsilon_i]}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \frac{\sum_{i=1}^n \tilde{x}_i^2 \sigma^2}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \frac{\sigma^2 \sum_{i=1}^n \tilde{x}_i^2}{\left(\sum_{i=1}^n \tilde{x}_i^2\right)^2} = \frac{\sigma^2}{\sum_{i=1}^n \tilde{x}_i^2}. \end{aligned} \quad (12.18)$$

Заменяя дисперсию σ^2 на оценку S^2 , определяемую как $S^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \alpha - \beta \tilde{x}_i)^2 = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2$, можно показать, что величина $\frac{nS^2}{\sigma^2}$ имеет закон распределения χ^2 с $(n-2)$ степенями свободы.

Для получения доверительного интервала для коэффициента α возьмем разность между точечной оценкой $\hat{\alpha}$ и коэффициентом α : $\hat{\alpha} - \alpha$. Нормируем ее, разделим на $D[\hat{\alpha}]$ и обозначим как V :

$$V = \frac{\hat{\alpha} - \alpha}{\sqrt{D[\hat{\alpha}]}} \quad (12.19)$$

Аналогично для коэффициента β :

$$W = \frac{\hat{\beta} - \beta}{\sqrt{D[\hat{\beta}]}} \quad (12.20)$$

Подставляя в (12.19), (12.20) выражения (12.17), (12.18) и разделив (12.19), (12.20) на $\sqrt{\frac{nS^2}{\sigma^2}/(n-2)}$, получим

$$V = \frac{\frac{\hat{\alpha} - \alpha}{\sigma/\sqrt{n}}}{\sqrt{\frac{nS^2}{\sigma^2}/(n-2)}}, \quad W = \frac{\frac{\hat{\beta} - \beta}{\sigma/\sum \tilde{x}_i^2}}{\sqrt{\frac{nS^2}{\sigma^2}/(n-2)}}. \quad (12.21)$$

Эти случайные величины имеют закон распределения Стьюдента с $\nu = n - 2$ степенями свободы.

Зная закон распределения случайной величины V (плотность распределения $f(\nu)$ закона распределения Стьюдента), можно найти вероятность ее попадания в интервал $]-V_p, +V_p[$:

$$P(-V_p < V < V_p) = \int_{-V_p}^{V_p} f(V) dV = p = 1 - \alpha. \quad (12.22)$$

Из условия (12.22) для заданной доверительной вероятности $p = 1 - \alpha$ и числа степеней свободы $\nu = n - 2$ по таблицам распределения Стьюдента находим квантили $V_p = t_{\alpha/2, n-2}$. Считая V_p известными и подставляя вместо V выражение (12.21) в правую часть выражения (12.22), получим неравенство

$$-V_p < \frac{\hat{\alpha} - \alpha}{S\sqrt{1/(n-2)}} < V_p.$$

Решая неравенство относительно α , находим *доверительный интервал для коэффициента α* :

$$\hat{\alpha} - V_p \frac{S}{\sqrt{n-2}} < \alpha < \hat{\alpha} + V_p \frac{S}{\sqrt{n-2}}. \quad (12.23)$$

Аналогично поступим для случайной величины W , записывая условие

$$P(-W_p < W < W_p) = \int_{-W_p}^{W_p} f(W) dW = p = 1 - \alpha. \quad (12.24)$$

Определяя квантили W_p по таблицам распределения Стьюдента, решаем неравенство в правой части (12.24):

$$-W_p < \frac{\hat{\beta} - \beta}{S \sqrt{1/\sum \tilde{x}_i^2}} \sqrt{\frac{n-2}{n}} < W_p. \quad (12.25)$$

Из (12.25) получаем *доверительный интервал для коэффициента β* :

$$\hat{\beta} - W_p \frac{S}{\sqrt{\sum \tilde{x}_i^2}} \sqrt{\frac{n-2}{n}} < \beta < \hat{\beta} + W_p \frac{S}{\sqrt{\sum \tilde{x}_i^2}} \sqrt{\frac{n-2}{n}}. \quad (12.26)$$

ЧАСТЬ 3. СЛУЧАЙНЫЕ ПРОЦЕССЫ

ГЛАВА 13. ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

13.1. Основные понятия

Пусть задано вероятностное пространство (Ω, F, P) , где Ω – пространство элементарных событий, F – σ -алгебра событий, P – вероятностная мера. Рассмотрим функцию $\varphi(t, \omega)$ двух аргументов, где ω элемент Ω и t – свободный аргумент, обычно время.

Случайным процессом называется случайная величина $\varphi(t, \omega)$, заданная на вероятностном пространстве (Ω, F, P) . Так как параметр t интерпретируется как время, то если $t = (0, 1, 2, \dots)$, то говорят что $\varphi(t, \omega)$ – процесс с дискретным временем; если $t \in [0; T]$, то $\varphi(t, \omega)$ – процесс с непрерывным временем.

Пусть ω приняло какое-то конкретное значение ω_1 . Тогда $x(t) = \varphi(t, \omega_1)$ будет функцией только аргумента t и называется *реализацией случайного процесса*. Другими словами, реализация – это тот вид, который принимает случайный процесс в результате какого-то конкретного опыта (рис. 13.1). В каждом опыте наблюдается своя реализация $x(t)$ случайного процесса (СП). Совокупность реализаций $\{x(t)\}$ носит название *ансамбля*. Значение случайного процесса в некоторый момент времени t называется *отсчетом*.

Рассмотрим отсчет x_1 случайного процесса $x(t)$ в момент времени t_1 : $x_1 = x(t_1)$, (см. рис. 13.1).

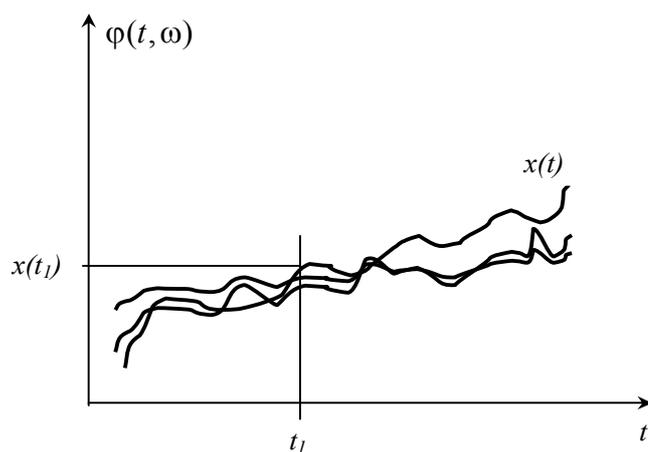


Рис. 13.1

Если нас интересует только x_1 , то это будет одномерная случайная величина, свойства которой полностью описываются одномерной плотностью распределений $f_1(x_1; t_1)$. Отметим, что в отличие от теории вероятностей $f_1(x_1; t_1)$ зависит не только от x_1 , но и от t_1 . Одномерная плотность распределения $f_1(x_1; t_1)$ дает некоторое представление о свойствах случайного процесса, но не полное. Более подробное представление о случайном процессе получается, если рассматривать два отсчета x_1 и x_2 , берущихся в моменты времени t_1 и t_2 : $x_1 = x(t_1)$, $x_2 = x(t_2)$. И характеризовать их двумерной плотностью распределения: $f_2(x_1, x_2; t_1, t_2)$. Тогда n -мерная плотность распределения запишется так:

$$f_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n). \quad (13.1)$$

Чем больше берется отсчетов и чем ближе они расположены друг к другу, тем подробнее описывается случайный процесс.

В пределе, когда $n \rightarrow \infty$, $\max(t_{i+1} - t_i) \rightarrow 0$, мы получаем совершенно точное представление о свойствах случайного процесса, т. е. случайный процесс описывается полностью, если задать все f_n для всех n и всех t_1, t_2, \dots, t_n .

13.2. Числовые характеристики случайного процесса

При экспериментальном получении случайного процесса можно сравнительно легко указать $f_1(x_1; t_1)$, трудно узнать $f_2(x_1, x_2; t_1, t_2)$, а измерить плотность распределения более высокого порядка практически невозможно. Поэтому на практике ограничиваются изучением некоторых характеристик случайного процесса, менее полных, но все же дающих представление об его основных свойствах.

Математическое ожидание случайного процесса (среднее по ансамблю) вводится следующим образом:

$$m(t) = M[x(t)] = \int_{-\infty}^{+\infty} x f_1(x; t) dx. \quad (13.2)$$

В отличие от теории вероятностей $m(t)$ не число, а функция от времени. Оно дает некоторую кривую, около которой группируются все реализации случайного процесса (рис. 13.2).

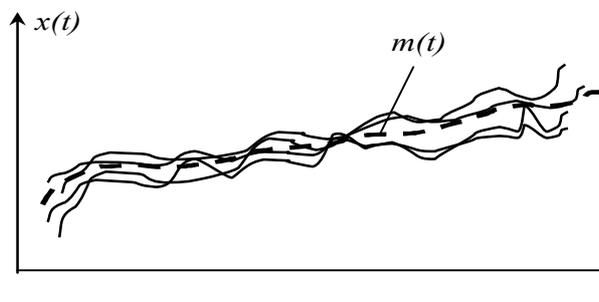


Рис. 13.2

Дисперсия случайного процесса (средняя по ансамблю):

$$D(t) = D[x(t)] = \int_{-\infty}^{+\infty} (x - m(t))^2 f_1(x; t) dx. \quad (13.3)$$

Дисперсия определяет, насколько сильно отдельные реализации могут отклоняться от математического ожидания. В отличие от теории вероятностей это также не число, а функция времени.

Функция корреляции случайного процесса (средняя по ансамблю):

$$R(t_1, t_2) = M[x(t_1), x(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (13.4)$$

Эта функция является важнейшей характеристикой случайного процесса. Функция корреляции характеризует степень зависимости между отсчетами случайного процесса, взятыми в разные моменты времени.

Наряду с $R(t_1, t_2)$ употребляется еще одна функция корреляции — это *функция корреляции флуктуаций*:

$$R_0(t_1, t_2) = \iint (x_1 - m(t_1))(x_2 - m(t_2)) f_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (13.5)$$

И нормированная функция корреляции:

$$\rho(t_1, t_2) = \frac{R_0(t_1, t_2)}{\sqrt{D(t_1)D(t_2)}}. \quad (13.6)$$

Последняя ничем по смыслу не отличается от коэффициента корреляции и определяет степень линейной зависимости отсчетов случайного процесса в моменты времени t_1 и t_2 .

13.3. Стационарные случайные процессы

Случайный процесс называется **стационарным** в узком смысле, если для любых τ и любых целых n имеет место равенство

$$f_n(x_1, \dots, x_n; t_1, \dots, t_n) = f_n(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau). \quad (13.7)$$

Смысл определения в следующем.

В левой части равенства стоят отсчеты, берущиеся в моменты времени t_1, t_2, \dots, t_n , и плотность распределений $f_n(x_1, \dots, x_n; t_1, \dots, t_n)$ полностью описывает свойства случайного процесса в моменты времени t_1, t_2, \dots, t_n . В правой части равенства стоят отсчеты, берущиеся в моменты времени $t_1 + \tau, t_2 + \tau, \dots, t_n + \tau$, и соответствующая плотность распределений $f_n(x_1, \dots, x_n; t_1 + \tau, \dots, t_n + \tau)$ описывает свойства случайного процесса в эти сдвинутые моменты времени. Равенство этих плотностей распределений означает, что свойства случайного процесса одинаковы как в моменты времени t_1, t_2, \dots, t_n , так и в моменты времени $t_1 + \tau, t_2 + \tau, \dots, t_n + \tau$. Таким образом, **стационарность в узком смысле означает, что все свойства, характеристики и т. п. случайного процесса не зависят от начала отсчета времени.** Грубо говоря, какими свойствами обладает случайный процесс сегодня, такие же свойства он имел год тому назад и такие же свойства он будет иметь через 1000 лет. Но реально это несколько не так. Все меняется, но нас интересуют отрезки времени, малые по сравнению со временем существования самого процесса.

Следствия стационарности.

1. Предположим в (13.7) $n = 1$, получим

$$f_1(x_1; t_1) = f_1(x_1; t_1 + \tau).$$

Поскольку τ произвольно, то полагая $\tau = -t_1$, получим:

$$f_1(x_1; t_1) = f_1(x_1; 0) = f_1(x_1).$$

Поэтому

$$m(t_1) = \int_{-\infty}^{+\infty} x_1 f_1(x_1; t_1) dx_1 = \int_{-\infty}^{+\infty} x_1 f_1(x_1; 0) dx_1 = m(0) = m,$$

$$D(t_1) = \int_{-\infty}^{+\infty} (x_1 - m)^2 f_1(x_1; t_1) dx_1 = \int_{-\infty}^{+\infty} (x_1 - m)^2 f_1(x_1; 0) dx_1 = D(0) = D.$$

Отметим, что у стационарных случайных процессов математическое ожидание, дисперсия и плотность распределения от времени не зависят (m и D – числа, плотность распределения $f_1(x_1)$ – функция только от x_1).

2. Полагая в (13.7) $n = 2$ и $\tau = -t_1$ получим

$$f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; 0, t_2 - t_1).$$

Тогда функция корреляции принимает вид

$$R(t_1, t_2) = \iint x_1 x_2 f_2(x_1, x_2; t_1, t_2) dx_1 dx_2 =$$

$$= \iint x_1 x_2 f_2(x_1, x_2; 0, t_2 - t_1) dx_1 dx_2 = R(t_1 - t_1) = R(\tau).$$

Видим, что у стационарных случайных процессов функция корреляции зависит лишь от разности $t_1 - t_2$ моментов времени, т. е. является функцией одного аргумента.

Случайный процесс называется **стационарным в широком смысле**, если его математическое ожидание и дисперсия от времени не зависят, а функция корреляции зависит лишь от разности моментов времени $t_1 - t_2$.

На рис. 13.3 показан вид стационарного случайного процесса.

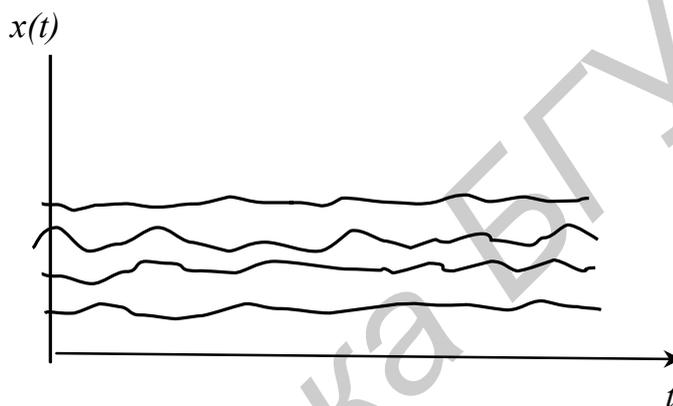


Рис. 13.3

13.4. Числовые характеристики случайного процесса – средние по времени

Ранее мы рассмотрели числовые характеристики случайного процесса, средние по ансамблю, общий вид которых

$$\langle \text{среднее по ансамблю} \rangle = \iint \left\langle \begin{array}{c} \text{усредненная} \\ \text{величина} \end{array} \right\rangle f_n(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n. \quad (13.8)$$

Они называются *средними по ансамблю* потому, что фиксируются моменты времени t_1, t_2, \dots, t_n и перебираются все возможные значения $\{x_i(t)\}$, т. е. весь ансамбль реализаций случайного процесса, на рис. 13.4 жирной линией изображены кривые плотностей распределений $f(x_i(t_1); t_1)$ и $f(x_i(t_2); t_2)$ в моменты времени t_1, t_2 .

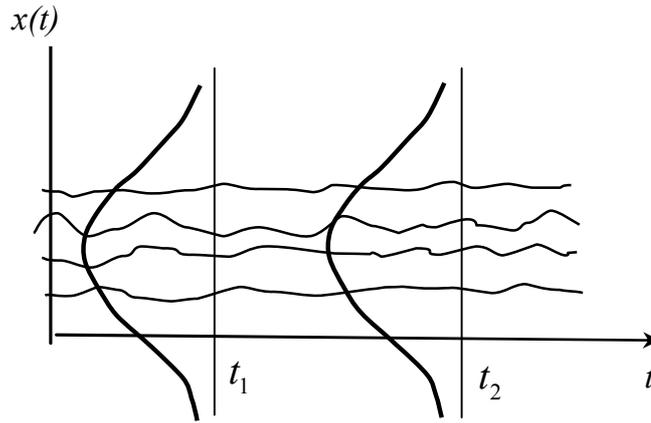


Рис. 13.4

Для *стационарных* случайных процессов кроме средних по ансамблю можно ввести еще так называемые *средние по времени*. В общем виде среднее по времени представляет собой:

$$\langle \text{среднее по времени} \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \langle \text{усредненная величина} \rangle dt. \quad (13.9)$$

Характерным для *средних по времени* является то, что фиксируются реализации, а «перебираются» все моменты времени. В средних по времени характерно также отсутствие плотности распределения $f_n(x_1, \dots, x_n; t_1, \dots, t_n)$, т. к. для стационарных процессов все моменты времени равноправны.

Тогда числовые характеристики *средние по времени* определяются так:

1. Математическое ожидание случайного процесса среднее по времени:

$$\tilde{m} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt. \quad (13.10)$$

2. Дисперсия случайного процесса средняя по времени:

$$\tilde{D} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (x(t) - \tilde{m})^2 dt. \quad (13.11)$$

3. Функция корреляции, средняя по времени (в ней фигурирует произведение значений случайного процесса в два различных момента времени t_1 и $t_2 = t_1 + \tau$). Учитывая, что усредненная величина имеет вид $x(t)x(t + \tau)$, получим

$$\tilde{R}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t + \tau) dt. \quad (13.12)$$

13.5. Свойства функций корреляции случайного процесса

1. Функция корреляции является *симметричной* функцией:

$$R(t_1, t_2) = R(t_2, t_1),$$

$$R(t_1, t_2) = M[x(t_1)x(t_2)] = M[x(t_2)x(t_1)] = R(t_2, t_1).$$

Для стационарного случайного процесса – *четная* функция:

$$R(\tau) = R(t_2 - t_1) = R(t_1 - t_2) = R(-\tau).$$

2. Функция корреляции – *ограниченная* функция. Воспользуемся неравенством Шварца (неравенство Коши-Буняковского):

$$M[U^2]M[V^2] \geq M^2[UV]. \quad (13.13)$$

Обозначим отсчеты $x(t_1), x(t_2)$, как $U = x(t_1), V = x(t_2)$, и подставим в (13.13):

$$M[x^2(t_1)]M[x^2(t_2)] \geq M^2[x(t_1)x(t_2)].$$

С учетом определения (13.4), получаем

$$R(t_1, t_1)R(t_2, t_2) \geq R^2(t_1, t_2).$$

Тогда

$$|R(t_1, t_2)| \leq \sqrt{R(t_1, t_1)R(t_2, t_2)}.$$

Когда $R(t_1, t_1) = R(t_2 - t_1)$, получаем:

$$|R(\tau)| \leq R(0).$$

3. Для *стационарного* случайного процесса дисперсия и математическое ожидание могут быть получены через функцию корреляции:

$$\lim_{\tau \rightarrow \infty} R(\tau) = M[x(t)x(t + \tau)] = \left. \begin{array}{l} \text{при } \tau \rightarrow \infty, \text{ т. е. когда интервал времени} \\ t_2 - t_1 \text{ между отсчетами неограниченно} \\ \text{увеличивается, то отсчеты становятся} \\ \text{независимыми, учтем: } M[x(t)] = m(0) = m \end{array} \right| =$$

$$= M[x(t)]M[x(t + \tau)] = m^2.$$

Откуда следует, что

$$m^2 = R(\infty).$$

Найдем дисперсию:

$$D = M[x^2(t)] - M^2[x(t)] = \left. \begin{array}{l} M[x(t)x(t + \tau) = R(\tau)] \\ R(0) = M[x^2(t)] \\ M^2[x(t)] = m^2 = R(\infty) \end{array} \right| = R(0) - R(\infty).$$

Если рассматривать функцию корреляции флуктуаций (13.5), то $D = R_0(0)$, т. к. $R_0(\infty) = M[x(t) - m]M[x(t) - m] = 0$.

4. Функция корреляции является *положительно определенной функцией*. Рассмотрим моменты времени t_1, \dots, t_n и произвольные величины ξ_1, \dots, ξ_n , поскольку:

$$M\left[\left\{\sum_{i=1}^n \xi_i x(t_i)\right\}^2\right] \geq 0, \Rightarrow M\left[\sum_{i=1}^n \sum_{j=1}^n x(t_i) \xi_i \xi_j\right] \geq 0,$$

вычисляя математическое ожидание, получим

$$\sum_{i,j=1}^n R(t_i, t_j) \xi_i \xi_j \geq 0.$$

Следствие: для любой положительно определенной функции $R(t_1, t_2)$ можно построить такой случайный процесс, для которого $R(t_1, t_2)$ будет функцией корреляции.

5. *Время корреляции* для стационарного случайного процесса определяется как

$$\tau_K = \frac{1}{R_0(0)} \int_0^{\infty} |R_0(\tau)| d\tau. \quad (13.14)$$

Время корреляции определяет, насколько далеко по времени наблюдается корреляционная связь между отсчетами в случайном процессе.

6. При изучении реальных случайных процессов для $R_0(\tau)$ используют следующие аппроксимации:

- а) $R_0(\tau) = D \cdot e^{-\frac{|\tau|}{\tau_0}}$.
Зависимость $R_0(\tau)$ приведена на рис. 13.5.

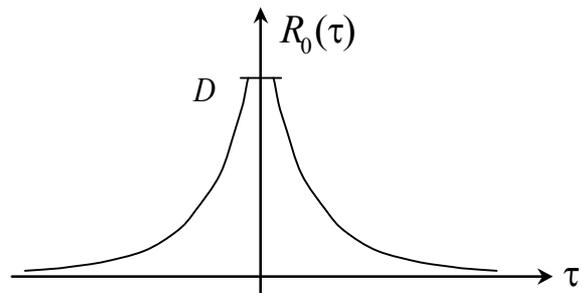


Рис. 13.5

- б) $R_0(\tau) = D \cdot e^{-\frac{|\tau|}{\tau_0}} \cos \beta \tau$.
Зависимость $R_0(\tau)$ приведена на рис. 13.6.

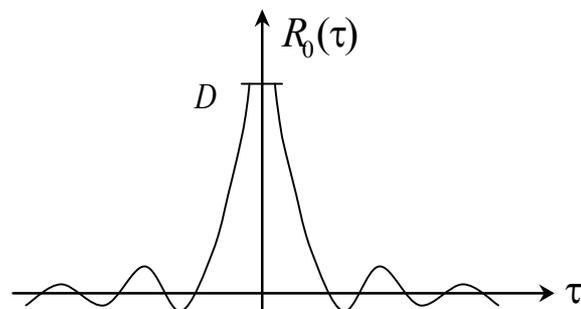


Рис. 13.6

$$в) \quad R_0(\tau) = D \cdot e^{-\tau^2/\tau_0^2}. \quad г) \quad R_0(\tau) = D \cdot e^{-\tau^2/\tau_0^2} \cos \beta \tau.$$

13.6. Эргодические случайные процессы

Случайный процесс называется *эргодическим*, если для него временные средние с вероятностью, равной 1, совпадают с соответствующими средними по ансамблю:

$$m = \tilde{m}, \quad R(\tau) = \tilde{R}(\tau), \quad D = \tilde{D}.$$

Эргодическим может быть только стационарный случайный процесс, но не всякий стационарный случайный процесс может быть эргодичен. Эргодичность имеет большое значение: она позволяет заменить изучение ансамбля реализаций изучением одной длинной реализации, т. к. каждая реализация (у эргодического случайного процесса) с вероятностью 1 имеет те же характеристики, что и весь ансамбль.

13.7. Спектр мощности случайного процесса

Из курса математического анализа известно, что любая достаточно гладкая функция, квадрат которой интегрируем $\int_{-\infty}^{\infty} x^2(t) dt$, меньшая $+\infty$, может быть представлена в виде интеграла Фурье (прямое и обратное преобразование Фурье):

$$x(t) = \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega, \quad (13.15)$$

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt, \quad (13.16)$$

где $X(\omega)$ – спектр функции $x(t)$, причем
 $|X(\omega)|$ – амплитудный спектр и
 $\arg X(\omega)$ – фазовый спектр функции $x(t)$.

Для стационарного случайного процесса $\int_{-\infty}^{\infty} x^2(t) dt = \infty$, поэтому преобразование Фурье можно ввести так: определим *урезанный* случайный процесс $x_T(t)$:

$$x_T(t) = \begin{cases} x(t) & |t| < T, \\ 0 & |t| > T. \end{cases}$$

Тогда $\int_{-\infty}^{\infty} x_T^2(t) dt = \int_{-T}^T x^2(t) dt < +\infty$, и для процесса $x_T(t)$ можно написать разложение в интеграл Фурье:

$$x_T(t) = \int_{-\infty}^{\infty} X_T(\omega) e^{i\omega t} d\omega, \quad (13.17)$$

$$X_T(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x_T(t) e^{-i\omega t} dt. \quad (13.18)$$

Рассмотрим величину $\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T x^2(t) dt = W$. В большинстве физических формул $\int_{-T}^T x^2(t) dt$ – величина пропорциональна энергии случайного процесса на интервале $[-T; T]$, а величина W имеет смысл средней мощности случайного процесса.

$$\begin{aligned} W &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T}^T x^2(t) dt = \left| \begin{array}{l} \text{в общем случае случайный процесс} \\ x(t) - \text{комплексная величина} \end{array} \right| = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} x_T(t) x_T^*(t) dt = \\ &= \left| \begin{array}{l} x^*(t) - \text{сопряженная} \\ \text{комплексная величина} \end{array} \right| = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_T(\omega_1) X_T^*(\omega_2) e^{i(\omega_1 - \omega_2)t} d\omega_1 d\omega_2 dt = \\ &= \left| \begin{array}{l} \text{учтем, что } \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it(\omega_1 - \omega_2)} dt = \delta(\omega_1 - \omega_2) - \delta\text{-функция Дирака,} \\ \text{ее свойство: } \int_{-\infty}^{\infty} f(\omega_1) \delta(\omega_1 - \omega_2) d\omega_1 d\omega_2 = f(\omega_2) \end{array} \right| = \\ &= \lim_{T \rightarrow \infty} \frac{2\pi}{2T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_T(\omega_1) X_T^*(\omega_2) \delta(\omega_1 - \omega_2) d\omega_1 d\omega_2 = \\ &= \lim_{T \rightarrow \infty} \frac{2\pi}{2T} \int_{-\infty}^{\infty} X_T(\omega) X_T^*(\omega) d\omega = \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{|X_T(\omega)|^2 2\pi}{2T} d\omega. \quad (13.19) \end{aligned}$$

Обозначим $S(\omega) = \lim_{T \rightarrow \infty} \frac{|X_T(\omega)|^2 2\pi}{2T}$ – эта функция называется **спектром мощности случайного процесса**.

Тогда запишем среднюю мощность так:

$$W = \int_{-\infty}^{\infty} S(\omega) d\omega, \quad (13.20)$$

где $S(\omega)d\omega$ – мощность, выделяемая случайным процессом в полосе частот $[\omega, \omega + d\omega]$.

Спектр мощности не является независимой характеристикой случайного процесса. Спектр мощности $S(\omega)$ можно связать с функцией корреляции $R(\tau)$, эту связь определяет следующая теорема, изложенная в разделе 13.8.

13.8. Теорема Винера-Хинчина

Эта теорема устанавливает связь между спектром мощности и функцией корреляции. Запишем функцию корреляции (13.12) через урезанный случайный процесс:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} x_T(t) x_T^*(t - \tau) dt = \left| \begin{array}{l} \text{в общем случае случайный процесс } x(t) - \\ \text{комплексная величина, поэтому запишем } R(\tau) \text{ в виде комплексно сопряженных} \\ \text{величин, используя (13.17): } x_T(t) = \int_{-\infty}^{\infty} X_T(\omega_1) e^{i\omega_1 t} d\omega_1, \text{ и учтем сопряженную вели-} \\ \text{чину } x_T^*(t) = \int_{-\infty}^{\infty} X_T^*(\omega_2) e^{-i\omega_2 t} d\omega_2 \end{array} \right. =$$

комплексная величина, поэтому запишем $R(\tau)$ в виде комплексно сопряженных величин, используя (13.17): $x_T(t) = \int_{-\infty}^{\infty} X_T(\omega_1) e^{i\omega_1 t} d\omega_1$, и учтем сопряженную вели-

$$\begin{aligned} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega_1 t - i\omega_2(t-\tau)} X_T(\omega_1) X_T^*(\omega_2) d\omega_1 d\omega_2 dt = \\ &= \left| \text{используем свойство } \delta\text{-функции} \right| = \\ &= \int_{-\infty}^{\infty} e^{i\omega\tau} \frac{1}{2T} \lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} X_T(\omega_1) X_T^*(\omega_2) 2\pi \delta(\omega_1 - \omega_2) d\omega_1 d\omega_2 = \\ &= \int_{-\infty}^{\infty} e^{i\omega\tau} \frac{1}{2T} \lim_{T \rightarrow \infty} \frac{|X_T(\omega)|^2}{2T} 2\pi d\omega = \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} d\omega. \end{aligned} \quad (13.21)$$

Таким образом, видим, что функция корреляции является прямым преобразованием Фурье от спектра мощности $S(\omega)$:

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} d\omega. \quad (13.22)$$

Тогда спектр мощности определится через обратное преобразование Фурье:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) e^{-i\omega\tau} d\tau. \quad (13.23)$$

Запишем спектр мощности через тригонометрические функции, используя формулу Эйлера: $e^{-i\omega\tau} = \cos \omega\tau - i \sin \omega\tau$.

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) \cos \omega\tau d\tau - \frac{i}{2\pi} \int_{-\infty}^{\infty} R(\tau) \sin \omega\tau d\tau. \quad (13.24)$$

Последний интеграл в (13.24) равен нулю, т. к. $R(\tau)$ четная функция. Тогда

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) \cos \omega\tau d\tau. \quad (13.25)$$

Отсюда следует, что спектр мощности $S(\omega)$ также четная функция. С учетом выше сказанного перепишем функцию корреляции (13.22):

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) \cos \omega\tau d\omega. \quad (13.26)$$

В формулы (13.25), (13.26) входит круговая частота $\omega = 2\pi f$, где f – линейная частота. Перепишем эти формулы через линейную частоту, вводя обозначения $S(\omega) d\omega = \bar{S}(f) df$, $\bar{S}(f) = 2\pi S(2\pi f)$. Тогда

$$R(\tau) = \int_{-\infty}^{\infty} \bar{S}(f) \cos(2\pi f \tau) df, \quad (13.27)$$

$$\bar{S}(f) = \int_{-\infty}^{\infty} R(\tau) \cos(2\pi f \tau) d\tau. \quad (13.28)$$

В этих формулах фигурируют отрицательные частоты от $-\infty$ до 0, т. к. они ничем не отличаются от положительных, то часто рассматривают только частотный интервал $[0; +\infty]$ и определяют $S(\omega)$ (для $\pm\omega$) так:

$$S(\omega) = \lim_{T \rightarrow \infty} \frac{[|X_T(\omega)|^2 + |X_T(-\omega)|^2] 2\pi}{2T} = 2 \lim_{T \rightarrow \infty} \frac{[|X_T(\omega)|^2] 2\pi}{2T} = 2S'(\omega). \quad (13.29)$$

где $S'(\omega)$ – только для $+\omega$.

В (13.29) мощность отрицательных частот прибавили к мощности положительных, тогда

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega) \cos(\omega\tau) d\omega = 2 \int_0^{\infty} S(\omega) \cos(\omega\tau) d\omega = 4 \int_0^{\infty} S'(\omega) \cos(\omega\tau) d\omega. \quad (13.30)$$

$$S'(\omega) = \frac{1}{2\pi} \int_0^{\infty} R(\tau) \cos(\omega\tau) d\tau. \quad (13.31)$$

Если вместо $R(\tau)$ взять функцию корреляции флуктуаций, то получится спектр мощности флуктуаций:

$$S_0(\omega) = \frac{1}{2\pi} \int_0^{\infty} R_0(\tau) \cos(\omega\tau) d\tau, \quad (13.32)$$

$$R_0(\tau) = 4 \int_0^{\infty} S_0(\omega) \cos(\omega\tau) d\omega. \quad (13.33)$$

13.9. Соотношение неопределенности для стационарных случайных процессов

Шириной спектра случайного процесса называется величина

$$\Delta\omega = \frac{1}{S_0(0)} \int_0^{\infty} |S_0(\omega)| d\omega = \left| \begin{array}{l} \text{из (13.33) при } \tau = 0 \\ R_0(0) = 4 \int_0^{\infty} S_0(\omega) d\omega, \text{ отсюда} \\ \int_0^{\infty} S_0(\omega) d\omega = \frac{R_0(0)}{4} \end{array} \right| = \frac{R_0(0)}{4S_0(0)}. \quad (13.34)$$

Время корреляции

$$\tau_0 = \frac{1}{R_0(0)} \int_0^{\infty} |R_0(\tau)| d\tau \geq \frac{1}{R_0(0)} \int_0^{\infty} R_0(\tau) d\tau = \left| \begin{array}{l} \text{из (13.32) при } \omega = 0 \\ S_0(0) = \frac{1}{2\pi} \int_0^{\infty} R_0(\tau) d\tau \end{array} \right| = \frac{2\pi S_0(0)}{R_0(0)}. \quad (13.35)$$

Умножая (13.34) на (13.35) получаем

$$\Delta\omega \cdot \tau_k \geq \frac{\pi}{2}.$$

Подставляя в $\Delta\omega$ круговую частоту, получим $2\pi \cdot \Delta f \cdot \tau_k \geq \frac{\pi}{2}$ и окончательно

$$\Delta f \cdot \tau_k \geq \frac{1}{4}. \quad (13.36)$$

13.10. Разложение случайного процесса в ряд Котельникова

Пусть $R(t_1, t_2)$ – функция корреляции случайного процесса $x(t)$. Учитывая, что $\tau = t_2 - t_1$ и $t_2 = t_1 + \tau$, можно записать $R(t_1, t_1 + \tau)$. Здесь функция корреляции зависит от момента времени t_1 и τ , поэтому запишем функцию корреляции как

функцию двух аргументов $R(\tau, t)$. Тогда на основании теоремы Винера-Хинчина можно записать

$$\begin{cases} S(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau, t) e^{-i\omega\tau} d\tau \\ R(\tau, t) = \int_{-\infty}^{\infty} S(\omega, t) e^{i\omega\tau} d\omega \end{cases} \quad (13.37)$$

Теорема: Если $S(\omega, t) \equiv 0$ для всех ω при $|\omega| > \Omega$, то имеет место разложение

$$x(t) = \lim_{n \rightarrow \infty} \sum_{k=-n}^n x\left(\frac{\pi k}{\Omega}\right) \frac{\sin(\Omega t - \pi k)}{\Omega t - \pi k}, \quad (13.38)$$

где $x\left(\frac{\pi k}{\Omega}\right)$ – отсчет случайного процесса в момент времени $\frac{\pi k}{\Omega}$.

На рис. 13.7 показана заштрихованная область частот, для которой спектр мощности $S(\omega, t) = 0$.

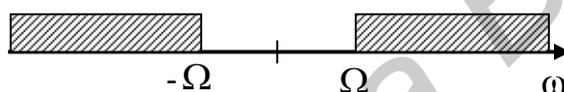


Рис. 13.7

Доказательство:

Предположим, что спектр мощности является периодической функцией с периодом 2Ω . Тогда спектр мощности можно представить в виде ряда Фурье:

$$S(\omega, \tau) = \sum_{k=-\infty}^{\infty} C_k e^{-i2\pi k \frac{\omega}{2\Omega}}. \quad (13.39)$$

Коэффициенты C_k определяются как

$$C_k = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} S(\omega, t) e^{i2\pi k \frac{\omega}{2\Omega}} d\omega = \left| \begin{array}{l} \text{из (13.37) получим} \\ R(\tau, t) = \int_{-\Omega}^{\Omega} S(\omega, t) e^{i\omega\tau} d\omega, \text{ обозначим } \tau = \frac{\pi k}{\Omega} \end{array} \right| =$$

$$= \frac{1}{2\Omega} R\left(\frac{\pi k}{\Omega}, t\right).$$

Подставим C_k в (13.39):

$$S(\omega, t) = \frac{1}{2\Omega} \sum_{k=-\infty}^{\infty} R\left(\frac{\pi k}{\Omega}, t\right) \cdot e^{-i2\pi k \frac{\omega}{2\Omega}}.$$

Определим функцию корреляции, используя (13.37):

$$\begin{aligned}
R(\tau, t) &= \int_{-\Omega}^{\Omega} S(\omega, t) e^{i\omega\tau} d\omega = \int_{-\Omega}^{\Omega} \sum_{k=-\infty}^{\infty} \frac{1}{2\Omega} R\left(\frac{\pi k}{2\Omega}, t\right) e^{-i2\pi k \frac{\omega}{2\Omega} + i\omega\tau} d\omega = \\
&= \frac{1}{2\Omega} \sum_{k=-\infty}^{\infty} R\left(\frac{\pi k}{2\Omega}, t\right) \int_{-\Omega}^{\Omega} e^{i\omega(\tau - \frac{\pi k}{\Omega})} d\omega = \\
&= \left[\int_{-\Omega}^{\Omega} e^{i\omega(\tau - \frac{\pi k}{\Omega})} d\omega = \frac{1}{i(\tau - \frac{\pi k}{\Omega})} e^{i\omega(\tau - \frac{\pi k}{\Omega})} \right]_{-\Omega}^{\Omega} = \left| \begin{array}{l} \text{учтем, что} \\ e^{i\theta} = \cos \theta + i \sin \theta \end{array} \right| = \\
&= \frac{2}{\tau - \frac{\pi k}{\Omega}} \cdot \sin(\Omega\tau - \pi k) = \sum_{k=-\infty}^{\infty} R\left(\frac{\pi k}{\Omega}, t\right) \cdot \frac{\sin(\Omega\tau - \pi k)}{\Omega\tau - \pi k}.
\end{aligned}$$

Поскольку функция корреляции симметрична, то можно записать

$$R(\tau, t) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} R\left(\frac{\pi k}{\Omega}, \frac{\pi l}{\Omega}\right) \frac{\sin(\Omega\tau - \pi k)}{\Omega\tau - \pi k} \frac{\sin(\Omega t - \pi l)}{\Omega t - \pi l}. \quad (13.40)$$

Докажем основное положение теоремы. Используем следующее определение сходимости: *Последовательность случайных величин X_n , $n=1, \infty$ сходится к случайной величине X в среднем квадратическом, если $\lim_{n \rightarrow \infty} M[(X_n - X)^2] = 0$.* Это определение заложено в основу метода наименьших квадратов.

Используя (13.38), можно записать

$$\begin{aligned}
\lim_{n \rightarrow \infty} \varepsilon_n &= \lim_{n \rightarrow \infty} M \left[\left\{ x(t) - \sum_{k=-n}^n x\left(\frac{\pi k}{\Omega}\right) \frac{\sin(\Omega t - \pi k)}{\Omega t - \pi k} \right\}^2 \right] = \\
&= \lim_{n \rightarrow \infty} \left\{ M[x(t)^2] - 2M \left[x(t) \cdot \sum_{k=-n}^n x\left(\frac{\pi k}{\Omega}\right) \frac{\sin(\Omega t - \pi k)}{\Omega t - \pi k} \right] + \right. \\
&\quad \left. + M \left[\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x\left(\frac{\pi k}{\Omega}\right) x\left(\frac{\pi l}{\Omega}\right) \frac{\sin(\Omega\tau - \pi k)}{\Omega\tau - \pi k} \frac{\sin(\Omega t - \pi l)}{\Omega t - \pi l} \right] \right\} = \\
&= \left| \begin{array}{l} \text{учитывая, что } M[x(t)x(t)] = R(t, t), \text{ а также (13.38), (13.40) получим} \end{array} \right| = \\
&= \lim_{n \rightarrow \infty} \{ R(t, t) - 2R(t, t) + R(t, t) \} = 0,
\end{aligned}$$

что и является доказательством нашей теоремы.

13.11. Классификация случайных процессов

Ранее, в разделе 13.1 мы отмечали, что все случайные процессы делятся на два основных класса: *процессы с дискретным временем* и *процессы с непрерывным временем*. После описания вероятностных характеристик случайного процесса, в частности n -мерной плотности распределения (раздел 13.1), мы далее ввели следующее разделение случайных процессов с учетом ограничений их вероятностных характеристик: *стационарные в узком смысле* и *стационарные в широком смысле* (раздел 13.3). Затем для стационарных случайных процессов в узком смысле мы ввели понятие *эргодических случайных процессов* (раздел 13.6). Далее мы рассмотрели энергетические характеристики случайных процессов, в частности спектр мощности случайного процесса (раздел 13.7). Иногда стационарные случайные процессы классифицируют по спектральным характеристикам: различают *стационарные случайные процессы в широком смысле – узкополосные*, когда спектр мощности сосредоточен в узкой полосе частот возле определенной фиксированной частоты ω_0 . В разделе 13.9 мы ввели понятие ширины спектра $\Delta\omega$: если $\Delta\omega \ll \omega_0$, то стационарный случайный процесс будет называться *узкополосным*. На рис. 13.8 изображен узкополосный спектр мощности $S(\omega)$ случайного процесса. На рис. 13.9 функция корреляции $R(\tau)$ узкополосного случайного процесса.

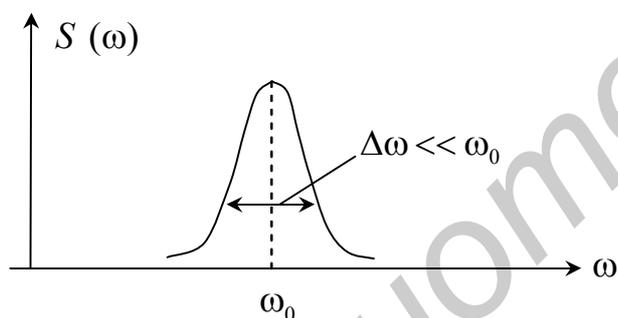


Рис. 13.8

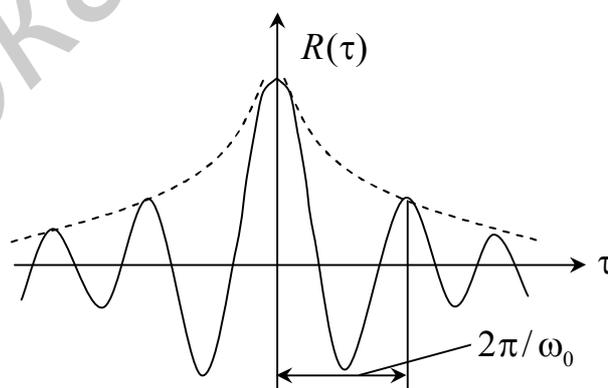


Рис. 13.9

Из рис. 13.9 видим, что функция корреляции узкополосного случайного процесса – это быстро осциллирующая функция с частотой ω_0 , но с медленно меняющейся огибающей. Ширина спектра $\Delta\omega$ и время корреляции определяются правыми частями формул (13.34) и (13.35).

Следующим классом будут *стационарные в широком смысле случайные процессы – широкополосные*. Здесь спектр мощности сохраняет постоянное значение W_0 на всех частотах (рис. 13.10):

$$S(\omega) = W_0, \quad -\infty < \omega < \infty. \quad (13.41)$$

Тогда стационарный в широком смысле случайный процесс с *равномерным спектром мощности на всех частотах* называется **белым шумом**. Функция корреляции белого шума представляет собой δ -функцию, расположенную в начале координат, (рис. 13.11):

$$R(\tau) = \delta(\tau)W_0. \quad (13.42)$$

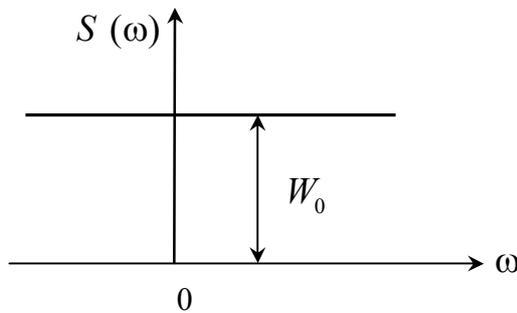


Рис. 13.10

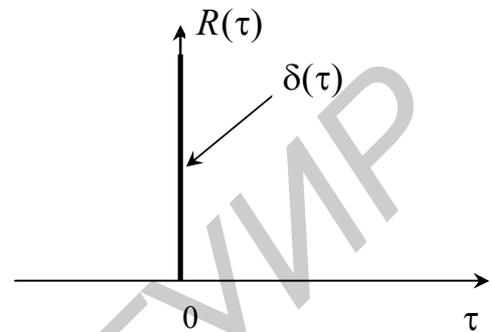


Рис. 13.11

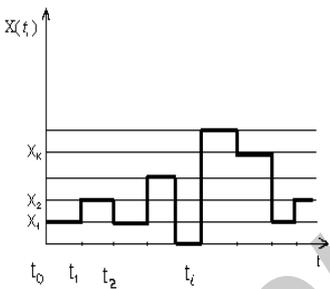
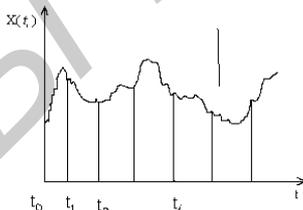
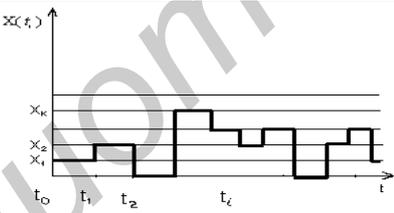
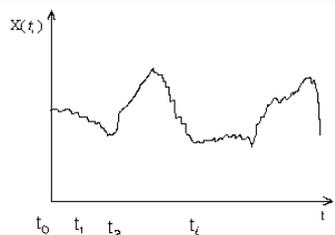
Для белого шума характерно то, что два его отсчета, взятые в сколь угодно близкие моменты времени, некоррелированы. Однако на практике белый шум реализовать невозможно, т. к. полная мощность белого шума бесконечна, а реальные случайные процессы имеют конечную мощность. И для реального случайного процесса рядом стоящие отсчеты коррелированы. Поэтому белый шум используют в качестве математической модели случайного процесса, что позволяет упростить анализ прохождения через линейные радиотехнические устройства с конечной полосой пропускания. Отметим, что любой случайный процесс, у которого спектр мощности сохраняет постоянное значение в некоторой полосе частот, может рассматриваться как белый шум. Но таких случайных процессов в природе не существует. Тем не менее такая модель случайного процесса оказывается удобной для анализа различных радиотехнических систем.

Еще одной моделью случайных процессов, которая находит широкое применение в приложениях теории случайных процессов, являются марковские случайные процессы.

13.12. Марковские случайные процессы

Марковские случайные процессы делятся на четыре основных типа в зависимости от того, какое множество значений (дискретное или непрерывное) принимает случайный процесс $X(t)$ и его параметр t в области задания процесса $[0, T]$: *марковские цепи* (дискретный процесс с дискретным временем), *марковские последовательности* (непрерывный процесс с дискретным временем), *дискретный марковский процесс* (дискретный процесс с непрерывным временем), *непрерывнозначный марковский процесс* (непрерывный процесс с непрерывным временем). Ниже в таблице приведены временные реализации для этих процессов.

Таблица

Значения аргумента t	Значения процессов	
	Дискретный	Непрерывный
Дискретные	 <p>Цепь Маркова</p>	 <p>Марковская последовательность</p>
Непрерывные	 <p>Дискретный марковский процесс</p>	 <p>Непрерывнозначный марковский процесс</p>

Марковскими процессами называются такие, в которых будущее не зависит от прошлого, а определяется только настоящим.

Цепи Маркова

Пусть некоторая физическая система может находиться в состояниях $A_1, A_2, \dots, A_i, \dots$ и переходить случайным образом из состояния в состояние в фиксированные моменты времени $t_1, t_2, \dots, t_s, \dots$. Эволюция системы образует *дискретную цепь Маркова с дискретным временем*, если выполняется следующее

условие (13.43) (условная вероятность в состоянии A_j в момент времени t_l определяется при условии, что в момент t_{l-1} система была в состоянии A_i):

$$P(A_j(t_l) | A_i(t_{l-1}), A_{i_1}(t_{l-2}), A_{i_2}(t_{l-3}), \dots, A_{i_k}(t_1)) = P(A_j(t_l) | A_i(t_{l-1})). \quad (13.43)$$

Дискретной цепью Маркова называется такой случайный процесс, в котором вероятность того, что система окажется в некотором состоянии A_j в момент времени t_l , зависит лишь от того, в каком состоянии находилась система в предыдущий момент времени t_{l-1} , и не зависит от эволюции системы до момента времени t_{l-1} .

Можно отметить, что в цепях Маркова зависимость между состояниями простирается лишь на один шаг назад.

Введем понятие *переходной вероятности* – это вероятность перехода системы из одного состояния в другое:

$$P_{ij}(t_l, t_{l-1}) = P(A_j(t_l) | A_i(t_{l-1})). \quad (13.44)$$

Если P_{ij} со временем не меняется, т. е. от t_{l-1} не зависит, то цепь Маркова называют *однородной*. Рассмотрим свойства таких цепей. Переходные вероятности однородной цепи Маркова образуют *матрицу переходных вероятностей*.

$$P = \begin{pmatrix} P_{11} & P_{12} \dots & P_{1j} \dots & P_{1n} \\ P_{21} & P_{22} \dots & P_{2j} \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ P_{i1} & P_{i2} \dots & P_{ij} \dots & P_{in} \\ P_{n1} & P_{n2} \dots & P_{nj} \dots & P_{nn} \end{pmatrix}. \quad (13.45)$$

Свойства матрицы переходных вероятностей:

1. $P_{ij} \geq 0$,

2. Сумма по строке – $\sum_{j=1}^N P_{ij} = 1$, ($i = \overline{1, N}$), N – количество состояний.

Матрицы, удовлетворяющие этим свойствам, называются *стохастическими*. Величины P_{ij} дают вероятности перехода из состояния A_i в состояние A_j за один шаг. Заметим, что здесь P_{ii} – вероятность того, что система, перешедшая к данному шагу в состоянии A_i , в нем же и задерживается на очередном шаге.

Вероятность перехода за n шагов. Рассмотрим вероятность перехода из состояния A_i , которое реализовано в s -м испытании, в состояние A_j за n шагов, т. е. в состояние A_j в $(s+n)$ -м испытании. Эта вероятность зависит только от n

(и не зависит от s). Обозначим ее P_{ij}^n . Тогда P_{ik}^m – вероятность перехода за m шагов из состояния A_j в состояние A_k и P_{kj}^{n-m} – вероятность перехода за $n-m$ шагов из состояния A_k в A_j (рис. 13.12). Используя формулу полной вероятности и учитывая, что промежуточные состояния A_1, A_2, \dots, A_n в $(s+m)$ -м испытании образуют полную группу попарно несовместимых событий, получим

$$P_{ij}^n = \sum_{k=1}^N P_{ik}^m P_{kj}^{n-m}. \quad (13.46)$$

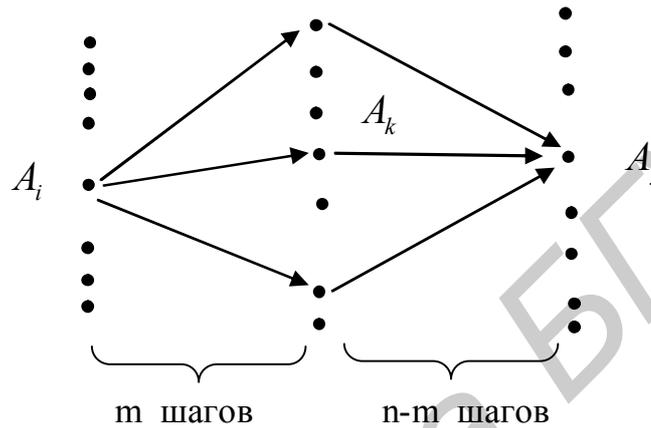


Рис. 13.12

Формула Маркова для цепей

Обозначим через π_n матрицу, составленную из вероятностей P_{ij}^n ; $i, j = \overline{1, n}$; таким образом, π_n – матрица перехода через n испытаний. Используя формулу для перемножения квадратных матриц (13.46), π_n можно записать в матричном виде:

$$\pi_n = \pi_m \pi_{n-m}, \quad m = \overline{1, n-1}. \quad (13.47)$$

Применяя последовательно формулу (13.47), получим

$$\pi_n = \pi_1 \pi_{n-1} = \pi_1 \pi_2 \pi_{n-2} = \dots = \pi_1^n.$$

Можно ожидать, что при переходах в n шагов влияние начального распределения с ростом n должно ослабевать в том смысле, что $P_{ij}^n \rightarrow P_j$ при $n \rightarrow \infty$ независимо от i . То есть если существует предел $\lim_{n \rightarrow \infty} P_{ij}^n = P_j$, то это свойство цепей Маркова называется эргодичностью.

Пусть P_j^k – вероятность того, что в k -м испытании осуществится событие A_j , назовем P_j^k – абсолютной вероятностью.

Пусть существует предел

$$\lim_{n \rightarrow \infty} P_j^n = P_j, \quad j = \overline{1, N}. \quad (13.48)$$

Тогда говорят, что существует предельное, *финальное*, распределение вероятностей P_1, \dots, P_j состояний A_1, \dots, A_n , не зависящее от начального распределения P_1 . Финальные вероятности удовлетворяют следующей системе уравнений:

$$P_j = \sum_{l=1}^N P_l P_{lj}, \quad j = \overline{1, N},$$

$$P_j \geq 0, \quad \sum_{j=1}^N P_j = 1.$$

Пример.

Дана цепь Маркова, которая описывается матрицей переходных вероятностей.

$$\begin{array}{c} j \rightarrow \\ i \left(\begin{array}{cc} P_{11} & P_{12} \\ P_{21} & P_{22} \end{array} \right) = \left(\begin{array}{cc} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{array} \right) \quad (\text{сумма вероятностей по строке равна } 1). \\ \downarrow \end{array}$$

Необходимо определить вероятность P_1 системы в 1-м состоянии.

Решение.

$$P_1 = \lim_{n \rightarrow \infty} P_{i1}^n.$$

Запишем вероятность перехода за n шагов, применяя формулу (13.46), получим для $m = 1$

$$\begin{aligned} P_{i1}^n &= P_{i1}^{n-1} P_{11}^1 + P_{i2}^{n-1} P_{21}^1 = \left| \begin{array}{l} P_{ij}^1 = P_{ij}, \quad P_{i2}^{n-1} = 1 - P_{i1}^{n-1} \\ \left| = P_{i1}^{n-1} P_{11} + (1 - P_{i1}^{n-1}) P_{21} = P_{21} + (P_{11} - P_{21}) P_{i1}^{n-1} = \right. \\ \left. = \right| \text{ введем переменные } \alpha = P_{21}, \quad \beta = P_{11} - P_{21} \quad \left| = \alpha + \beta \cdot z_{n-1}. \end{array} \right. \end{aligned}$$

Обозначая $z_n = P_{i1}^n$, запишем вероятность перехода за n шагов в новых переменных:

$$z_n = \alpha + \beta \cdot z_{n-1},$$

$$z_1 = \alpha + \beta \cdot z_0,$$

$$z_2 = \alpha + \beta \cdot z_1 = \alpha + \beta(\alpha + \beta \cdot z_0) = \alpha(1 + \beta) + \beta^2 \cdot z_0,$$

.....

$$z_n = \alpha(1 + \beta + \beta^2 + \dots + \beta^{n-1}) + \beta^n \cdot z_0.$$

Учтем, что $1 + \beta + \beta^2 + \dots + \beta^{n-1}$ – сумма геометрической прогрессии, тогда

$$z_n = \frac{\alpha(1 - \beta^{n-1})}{1 - \beta} + \beta^n \cdot z_0.$$

Найдем предел при $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} z_n = \text{т. к. } \beta < 1, z_n = P_{i1}^n \mid = \lim_{n \rightarrow \infty} \left(\frac{\alpha(1 - \beta^{n-1})}{1 - \beta} + \beta^n \cdot z_0 \right) = \frac{\alpha}{1 - \beta}.$$

Тогда искомая вероятность вычисляется так:

$$P_1 = \lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} P_{i1}^n = \frac{\alpha}{1 - \beta} = \frac{P_{21}}{1 - P_{11} + P_{21}} = \frac{\frac{1}{4}}{1 - \frac{1}{2} + \frac{1}{4}} = \frac{1}{3}.$$

13.12.1. Марковские процессы с непрерывным временем

Случайный процесс $X(t)$ называется **марковским**, если для любого момента времени t_1 при известном значении $X(t_1)$ случайные величины $X(t)$ с $t > t_1$ не зависят от случайных величин $X(S)$ с $S < t_1$, т. е. марковские процессы характеризуются тем, что вероятностные свойства процесса в момент $t > t_1$ определяются состоянием в момент t_1 и не зависят от состояний процесса до момента t_1 .

Среди марковских процессов с непрерывным множеством состояний наиболее важными являются процессы, которые имеют n -мерную плотность распределения. Если $X(t)$ – случайный процесс, $t \in T$, то пусть для каждого набора моментов времени $t_1, t_2, \dots, t_n \in T$ n -мерная случайная величина $(X(t_1), X(t_2), \dots, X(t_n))$ имеет n -мерную плотность распределения $f(t_1, x_1; \dots; t_n, x_n)$. Эта плотность обладает следующими свойствами:

1. $f(t_1, x_1; \dots; t_n, x_n)$ симметрична относительно любых перестановок пар аргументов (t_i, x_i) , т. к. $f(t_1, x_1; \dots; t_n, x_n) dx_1 \dots dx_n$ выражает вероятность совместного осуществления событий $\{x_i \leq X(t_i) \leq x_i + dx_i\}$ $i = \overline{1, n}$ и, значит, не зависит от порядка их перечисления.

2. Плотность любого k -мерного распределения при $k < n$ определяется с помощью n -мерного распределения:

$$f(t_1, x_1; \dots; t_k, x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_n(t_1, x_1; \dots; t_k, x_k; t_{k+1}, x_{k+1}; \dots; t_n, x_n) d_{k+1} \dots dx_n. \quad (13.49)$$

Условная вероятность для марковского процесса

$$\begin{aligned} P(X(t_i) < x_i \mid X(t_1) = x_1, X(t_2) = x_2, \dots, X(t_{i-1}) = x_{i-1}) = \\ = P(X(t_i) < x_i \mid X(t_{i-1}) = x_{i-1}). \end{aligned} \quad (13.50)$$

Поскольку вероятностные свойства процесса в момент t_n определяются состоянием в момент t_{n-1} и не зависят от протекания процессов в предшествующие моменты времени, тогда условная плотность распределения

$$g_n(t_n, x_n | t_1, x_1; \dots; t_{n-1}, x_{n-1}) = g_n(t_n, x_n | t_{n-1}, x_{n-1}). \quad (13.51)$$

Условную плотность распределения $g_n(t, x | \tau, y)$ называют *переходной плотностью распределения* в состояние t_n, x_n при условии, что процесс находится в состоянии $t_{n-1}, x_{n-1}(\tau, y)$.

Зная, что $P(AB) = P(A)P(B/A)$, для условной плотности распределения

$$f_n(t_1, x_1; \dots; t_n, x_n) = f_{n-1}(t_1, x_1; \dots; t_{n-1}, x_{n-1}) \cdot g_n(t_n, x_n | t_1, x_1; \dots; t_{n-1}, x_{n-1}). \quad (13.52)$$

Учитываем свойство марковских цепей:

$$g_n(t_n, x_n | t_1, x_1; \dots; t_{n-1}, x_{n-1}) = g(t_n, x_n | t_{n-1}, x_{n-1}). \quad (13.53)$$

Тогда $f_n(t_1, x_1; \dots; t_n, x_n) = f_{n-1}(t_1, x_1; \dots; t_{n-1}, x_{n-1}) \cdot g(t_n, x_n | t_{n-1}, x_{n-1})$, и применяя эту формулу в правой части для f_{n-1} , а затем для $f_{n-2} \dots f_2$, получим

$$f_n(t_1, x_1; \dots; t_n, x_n) = f_1(t_1, x_1) \cdot g(t_2, x_2 | t_1, x_1) g(t_3, x_3 | t_2, x_2) \dots g(t_n, x_n | t_{n-1}, x_{n-1}). \quad (13.54)$$

Отсюда видно, что для задания n -мерной плотности распределения марковского процесса достаточно знать лишь две функции: одномерную плотность $f_1(t, x)$ и переходную плотность распределения $g(t, x | \tau, y)$.

Рассмотрим 3 момента времени: $t_0 < \tau < t$; $t_0; \tau$; ($t \in T$). Согласно (13.54), имеем

$$f_3(t_0, x_0; \tau, y; t, x) = f_1(t_0, x_0) g(\tau, y | t_0, x_0) g(t, x | \tau, y), \quad (13.55)$$

$$f_2(t_0, x_0; t, x) = f_1(t_0, x_0) g(t, x | t_0, x_0). \quad (13.56)$$

Используя (13.49) для f_2 , получим

$$f_2(t_0, x_0; t, x) = \int_{-\infty}^{\infty} f_3(t_0, x_0; \tau, y; t, x) dy. \quad (13.57)$$

Подставляя в (13.57) уравнения (13.55) и (13.56), получим

$$f_1(t_0, x_0) \cdot g(t, x | t_0, x_0) = \int_{-\infty}^{\infty} f_1(t_0, x_0) \cdot g(\tau, y | t_0, x_0) \cdot g(t, x | \tau, y) dy,$$

$$g(t, x | t_0, x_0) = \int_{-\infty}^{\infty} g(\tau, y | t_0, x_0) \cdot g(t, x | \tau, y) dy. \quad (13.58)$$

Это уравнение Смолуковского (или Колмогорова-Чепмена) является основным в теории непрерывных марковских процессов.

13.12.2. Уравнения Колмогорова

Введем следующие обозначения:

$$x \rightarrow x_0, \quad t - \Delta t \rightarrow t_0,$$

$$y \rightarrow y, \quad t \rightarrow \tau,$$

$$z \rightarrow x, \quad T \rightarrow t.$$

Запишем переходные вероятности через уравнение Смолуковского в моменты $t - \Delta t$ и t :

$$g(T, z | t - \Delta t, x) = \int_{-\infty}^{\infty} g(t, y | t - \Delta t, x) \cdot g(T, z | t, y) dy, \quad (13.59)$$

$$g(T, z | t, x) = \int_{-\infty}^{\infty} g(T, z | t, y) \cdot g(t, y | t - \Delta t, x) dy. \quad (13.60)$$

Вычтем из формулы (13.60) формулу (13.59):

$$\begin{aligned} g(T, z | t, x) - g(T, z | t - \Delta t, x) &= \\ &= \int_{-\infty}^{\infty} [g(T, z | t, x) - g(T, z | t, y)] \cdot g(t, y | t - \Delta t, x) dy. \end{aligned} \quad (13.61)$$

Разложим в ряд Тейлора функцию

$$g(T, z | t, y) - g(T, z | t, x) = (y - x) \frac{\partial g}{\partial x} + \frac{(y - x)^2}{2} \frac{\partial^2 g}{\partial x^2} + \dots$$

Подставляя это разложение в интеграл (13.61), разделим левую и правую часть на Δt и перейдем к пределу при $\Delta t \rightarrow 0$:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{g(T, z | t, y) - g(T, z | t - \Delta t, x)}{\Delta t} &= \\ &= \lim_{\Delta t \rightarrow 0} \left(-\frac{1}{\Delta t} \int_{-\infty}^{\infty} [(y - x) \frac{\partial g}{\partial x} + \frac{(y - x)^2}{2} \frac{\partial^2 g}{\partial x^2}] \cdot g(t, y | t - \Delta t, x) dy, \end{aligned} \quad (13.62)$$

$$\frac{\partial g(T, z | t, x)}{\partial t} + a(t, x) \frac{\partial g(T, z | t, x)}{\partial x} + \frac{1}{2} b(t, x) \frac{\partial^2 g(T, z | t, x)}{\partial x^2} = 0. \quad (13.63)$$

Это первое уравнение Колмогорова, где

$$a(t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (y - x) g(t, y | t - \Delta t, x) dy, \quad (13.64)$$

$$b(t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (y - x)^2 \cdot g(t, y | t - \Delta t, x) dy. \quad (13.65)$$

Аналогично выводится *второе уравнение Колмогорова*:

$$\frac{\partial g(T, z | t, x)}{\partial T} + \frac{\partial}{\partial z} [a(T, z) g(T, z | t, x)] - \frac{1}{2} \frac{\partial^2}{\partial z^2} [b(T, z) g(T, z | t, x)] = 0. \quad (13.66)$$

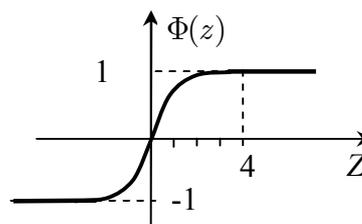
Здесь $a(T, z)$ и $b(T, z)$ – те же функции, что и $a(t, x)$ (13.64) и $b(t, x)$ (13.65), но взятые для T и z .

Непрерывный марковский процесс, который описывается уравнениями (13.63) и (13.66) называется диффузионным. Коэффициент $a(T, z)$ называется *коэффициентом сноса*, а коэффициент $b(T, z)$ – *коэффициентом диффузии*. Уравнение (13.66) называется *прямым уравнением Колмогорова*, а уравнение (13.63) называется *обратным уравнением Колмогорова*. Уравнения Колмогорова относятся к классу параболических дифференциальных уравнений в частных производных.

ПРИЛОЖЕНИЕ

1. Таблица значений функции Лапласа

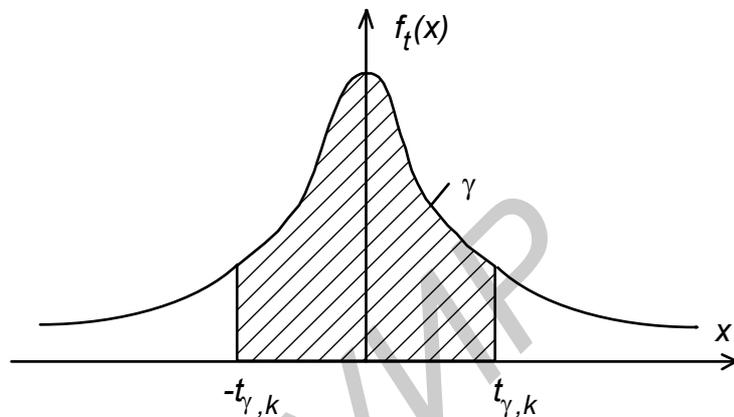
$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt.$$



z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
0,00	0,0000	0,66	0,4907	1,32	0,8132	1,98	0,9523
0,02	0,0160	0,68	0,5035	1,34	0,8198	2,00	0,9545
0,04	0,0319	0,70	0,5161	1,36	0,8262	2,05	0,9596
0,06	0,0478	0,72	0,5285	1,38	0,8324	2,10	0,9643
0,08	0,0638	0,74	0,5407	1,40	0,8385	2,15	0,9684
0,10	0,0797	0,76	0,5527	1,42	0,8444	2,20	0,9722
0,12	0,0955	0,78	0,5646	1,44	0,8501	2,25	0,9756
0,14	0,1113	0,80	0,5763	1,46	0,8557	2,30	0,9786
0,16	0,1271	0,82	0,5878	1,48	0,8611	2,35	0,9812
0,18	0,1428	0,84	0,5991	1,50	0,8664	2,40	0,9836
0,20	0,1585	0,86	0,6102	1,52	0,8715	2,45	0,9857
0,22	0,1741	0,88	0,6211	1,54	0,8764	2,50	0,9876
0,24	0,1897	0,90	0,6319	1,56	0,8812	2,55	0,9892
0,26	0,2051	0,92	0,6424	1,58	0,8859	2,60	0,9907
0,28	0,2205	0,94	0,6528	1,60	0,8904	2,66	0,9920
0,30	0,2358	0,96	0,6629	1,62	0,8948	2,70	0,9931
0,32	0,2510	0,98	0,6729	1,64	0,8990	2,75	0,9940
0,34	0,2661	1,00	0,6827	1,66	0,9031	2,80	0,9949
0,36	0,2812	1,02	0,6923	1,68	0,9070	2,85	0,9956
0,38	0,2961	1,04	0,7017	1,70	0,9109	2,90	0,9963
0,40	0,3108	1,06	0,7109	1,72	0,9146	2,95	0,9968
0,42	0,3255	1,08	0,7199	1,74	0,9181	3,00	0,9973
0,44	0,3401	1,10	0,7287	1,76	0,9216	3,10	0,9981
0,46	0,3545	1,12	0,7373	1,78	0,9249	3,20	0,9986
0,48	0,3688	1,14	0,7457	1,80	0,9281	3,30	0,9990
0,50	0,3859	1,16	0,7540	1,82	0,9312	3,40	0,9993
0,52	0,3969	1,18	0,7620	1,84	0,9342	3,50	0,9995
0,54	0,4108	1,20	0,7699	1,86	0,9371	3,60	0,9997
0,56	0,4245	1,22	0,7775	1,88	0,9399	3,70	0,9998
0,58	0,4381	1,24	0,7850	1,90	0,9426	3,80	0,9999
0,60	0,4515	1,26	0,7923	1,92	0,9451	3,90	0,9999
0,62	0,4647	1,28	0,7995	1,94	0,9476	4,00	0,9999
0,64	0,4778	1,30	0,8064	1,96	0,9500		

2. Таблица квантилей распределения Стьюдента

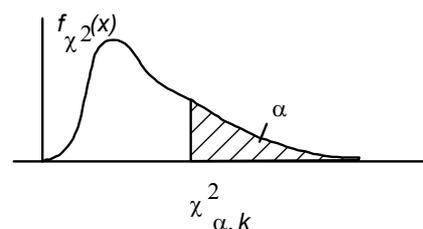
$$\gamma = \int_{-t_{\gamma,k}}^{t_{\gamma,k}} f_t(x) dx.$$



k	γ			
	0,90	0,95	0,98	0,99
1	6,31	12,71	31,8	63,7
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,77	3,75	4,60
5	2,02	2,57	3,36	4,03
6	1,943	2,45	3,14	4,71
7	1,895	2,36	3,00	3,50
8	1,860	2,31	2,90	3,36
9	1,833	2,26	2,82	3,25
10	1,812	2,23	2,76	3,17
12	1,782	2,18	2,68	3,06
14	1,761	2,14	2,62	2,98
16	1,746	2,12	2,58	2,92
18	1,734	2,10	2,55	2,88
20	1,725	2,09	2,53	2,84
22	1,717	2,07	2,51	2,82
24	1,711	2,06	2,49	2,80
30	1,697	2,04	2,46	2,75
40	1,684	2,02	2,42	2,70

3. Таблица квантилей $\chi^2_{\alpha,k}$ распределения «Хи-квадрат»

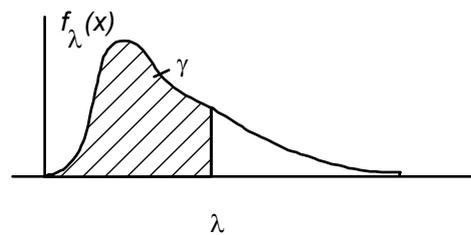
$$P(\chi^2 > \chi^2_{\alpha,k}) = \alpha.$$



k	α					
	0,01	0,02	0,05	0,95	0,98	0,99
1	6,64	5,41	3,84	0,004	0,001	0,000
2	9,21	7,82	5,99	0,103	0,040	0,020
3	11,34	9,84	7,82	0,352	0,185	0,115
4	13,28	11,67	9,49	0,711	0,429	0,297
5	15,09	13,39	11,07	1,145	0,752	0,554
6	16,81	15,03	12,59	1,635	1,134	0,872
7	18,48	16,62	14,07	2,17	1,564	1,239
8	20,10	18,17	15,51	2,73	2,03	1,646
9	21,07	19,68	16,92	3,32	2,53	2,09
10	23,20	21,2	18,31	3,94	3,06	2,56
12	26,2	24,1	21,0	5,23	4,18	3,57
14	29,1	26,9	23,7	6,57	5,37	4,66
16	32,0	29,6	26,3	7,96	6,61	5,81
18	34,8	32,3	28,9	9,39	7,91	7,02
20	37,6	35,0	31,4	10,85	9,24	8,26
22	40,3	37,7	33,9	12,34	10,60	9,54
24	43,0	40,3	36,4	13,85	11,99	10,86
26	45,6	42,9	38,9	15,38	13,41	12,20
28	48,3	45,4	41,3	16,93	14,85	13,56
30	50,9	48,0	43,8	18,49	16,31	14,95

4. Таблица значений функции Колмогорова

$$P(\lambda_\gamma) = \gamma = 1 - \alpha.$$



λ_γ	γ	λ_γ	γ	λ_γ	γ
0,50	0,0361	1,02	0,7500	1,54	0,9826
0,54	0,0675	1,06	0,7889	1,58	0,9864
0,58	0,1104	1,10	0,8223	1,62	0,9895
0,62	0,1632	1,14	0,8514	1,66	0,9918
0,66	0,2236	1,18	0,8765	1,70	0,9938
0,70	0,2888	1,22	0,8981	1,74	0,9953
0,74	0,3560	1,26	0,9164	1,78	0,9965
0,78	0,4230	1,30	0,9319	1,82	0,9973
0,82	0,4880	1,34	0,9449	1,86	0,9980
0,86	0,5497	1,38	0,9557	1,90	0,9985
0,90	0,6073	1,42	0,9646	1,94	0,9989
0,94	0,6601	1,46	0,9718	1,98	0,9992
0,98	0,7079	1,50	0,9778		

ЛИТЕРАТУРА

1. Вентцель, Е. С. Теория вероятностей и математическая статистика : учеб. пособие / Е. С. Вентцель; 5-е изд., стереотип. – М. : Высш. школа, 1998. – 576 с.
2. Вентцель, Е. С. Теория вероятностей и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. – М. : Наука, 1988. – 416 с.
3. Гнеденко, Б. В. Курс теории вероятностей : учебник / Б. В. Гнеденко; 6-е изд., перераб. и доп. – М. : Наука, 1988. – 448 с.
4. Пытьев, Ю. А. Курс теории вероятностей и математической статистики для физиков : учеб. пособие / Ю. А. Пытьев, И. А. Шишмарев. – М. : Изд-во Моск. ун-та, 1983. – 256 с.
5. Герасимович, А. И. Математическая статистика / А. И. Герасимович. – Минск : Выш. школа, 1983. – 279 с.
6. Левин, Б. Р. Теоретические основы статистической радиотехники / Б. Р. Левин. – М. : Радио и связь, 1989. – 656 с.
7. Феллер, В. Введение в теорию вероятностей и ее приложения. В 2 т. / В. Феллер; пер. с англ. – М. : Мир, 1984. – Т.1 – 528 с. ; Т.2 – 738 с.
8. Жевняк, Р. М. Теория вероятностей и математическая статистика : учеб. пособие / Р. М. Жевняк, А. А. Карпук, В. Т. Унукович. – Минск : Харвест, 2000. – 384 с.
9. Гмурман, В. Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов / В. Е. Гмурман; 9-е изд., стереотипное. – М. : Высшая школа, 2003. – 479 с.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ.....	3
ЧАСТЬ 1. ТЕОРИЯ ВЕРОЯТНОСТЕЙ.....	5
ГЛАВА 1. СЛУЧАЙНЫЕ СОБЫТИЯ.....	5
1.1. Основные понятия теории вероятностей.....	5
1.2. Пространство элементарных событий.....	8
1.3. Операции над событиями.....	10
1.4. Аксиомы теории вероятностей (Аксиомы Колмогорова А. Н.).....	13
1.5. Классический метод определения вероятностей.....	16
1.6. Геометрическое определение вероятностей.....	18
1.7. Свойства вероятности (основные теоремы).....	21
1.8. Условная вероятность. Независимость событий.....	24
1.9. Формула полной вероятности.....	28
1.10. Формула Байеса.....	29
1.11. Последовательность независимых испытаний. Формула Бернулли.....	31
1.11.1. Наивероятнейшее число наступлений события при повторении испытаний.....	33
ГЛАВА 2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ.....	35
2.1. Определение случайной величины.....	35
2.2. Законы распределения случайных величин.....	37
2.3. Функция распределения.....	38
2.4. Плотность распределения непрерывной случайной величины.....	42
ГЛАВА 3. МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ.....	46
3.1. Понятие о многомерных (векторных) случайных величинах.....	46
3.2. Закон распределения многомерной случайной величины.....	47
3.3. Функция распределения двумерной случайной величины.....	48
3.4. Плотность распределения непрерывной двумерной случайной величины.....	50
3.5. Условные законы распределения.....	52
3.6. Зависимые и независимые случайные величины.....	54
ГЛАВА 4. ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН.....	58
4.1. Числовые характеристики одномерной случайной величины.....	58
4.1.1. Математическое ожидание.....	58
4.1.2. Общее определение математического ожидания.....	60
4.1.3. Свойства математического ожидания.....	62
4.1.4. Дисперсия, среднее квадратическое отклонение.....	65
4.1.5. Моменты распределения случайных величин.....	68
4.1.6. Числовые характеристики двумерной случайной величины. Ковариация. Коэффициент корреляции.....	71
ГЛАВА 5. ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН.....	74
5.1. Одномерное приближение.....	74
5.2. Двумерное приближение.....	76
5.3. Распределение суммы независимых случайных величин.....	79
5.4. Характеристические функции.....	79
ГЛАВА 6. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН В ТЕОРИИ ВЕРОЯТНОСТЕЙ.....	82
6.1. Биномиальный закон распределения.....	82
6.2. Закон распределения Пуассона.....	84
6.3. Равномерный закон распределения.....	86
6.4. Показательный (экспоненциальный) закон распределения.....	87

6.5. Нормальный (гауссовский) закон распределения.....	88
6.5.1. Функция Лапласа.....	91
ГЛАВА 7. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ. ЗАКОН БОЛЬШИХ ЧИСЕЛ.....	93
7.1. Неравенство Чебышева.....	93
7.2. Теорема Чебышева.....	94
7.3. Теорема Бернулли.....	96
7.4. Центральная предельная теорема (Теорема Ляпунова).....	98
7.5. Формулы Муавра – Лапласа.....	100
ЧАСТЬ 2. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА.....	102
ГЛАВА 8. ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ.....	102
8.1. Выборка, вариационный ряд, гистограмма.....	102
8.2. Оценки и методы их получения.....	104
8.2.1. Метод моментов.....	105
8.2.2. Метод наибольшего правдоподобия.....	106
8.3. Свойства оценок.....	108
ГЛАВА 9. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН В МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ.....	114
9.1. Гамма-функция и ее свойства.....	114
9.2. Распределение χ^2 (хи-квадрат).....	115
9.3. Распределение Стьюдента (t -распределение).....	118
9.4. Распределение Фишера (F -распределение).....	119
ГЛАВА 10. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЙ.....	121
10.1. Доверительный интервал, доверительная вероятность.....	121
10.2. Доверительный интервал для математического ожидания случайной величины X при известной дисперсии (или σ).....	122
10.3. Доверительный интервал для математического ожидания нормальной случайной величины X при неизвестной дисперсии или σ	124
10.4. Доверительный интервал для дисперсии или σ нормальной случайной величины X	127
ГЛАВА 11. ТЕОРИЯ СТАТИСТИЧЕСКОЙ ПРОВЕРКИ ГИПОТЕЗ.....	130
11.1. Основные понятия.....	130
11.2. Проверка гипотезы равенства математических ожиданий при неизвестной дисперсии (критерий Стьюдента).....	132
11.3. Проверка гипотезы о равенстве дисперсий (критерий Фишера).....	135
11.4. Проверка гипотезы о законе распределения генеральной случайной величины. Критерий Пирсона. (Критерий согласия χ^2).....	138
11.5. Критерий Романовского.....	140
11.6. Критерий согласия Колмогорова.....	141
ГЛАВА 12. ЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ.....	144
12.1. Уравнение линейной регрессии.....	144
12.2. Метод наименьших квадратов.....	146
12.3. Коэффициент корреляции (оценки).....	147
12.4. Построение доверительных интервалов для коэффициентов уравнения регрессии.....	150
ЧАСТЬ 3. СЛУЧАЙНЫЕ ПРОЦЕССЫ.....	153
ГЛАВА 13. ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ.....	153
13.1. Основные понятия.....	153
13.2. Числовые характеристики случайного процесса.....	154
13.3. Стационарные случайные процессы.....	156
13.4. Числовые характеристики случайного процесса – средние по времени.....	157

13.5. Свойства функций корреляции случайного процесса	159
13.6. Эргодические случайные процессы.....	161
13.7. Спектр мощности случайного процесса.....	161
13.8. Теорема Винера-Хинчина	163
13.9. Соотношение неопределенности для стационарных случайных процессов	165
13.10. Разложение случайного процесса в ряд Котельникова.....	165
13.11. Классификация случайных процессов.....	168
13.12. Марковские случайные процессы.....	170
13.12.1. Марковские процессы с непрерывным временем	174
13.12.2. Уравнения Колмогорова	176
ПРИЛОЖЕНИЕ.....	178
ЛИТЕРАТУРА	181
ОГЛАВЛЕНИЕ.....	182

Библиотека БГУИР

Учебное издание

Аксенчик Анатолий Владимирович

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА**

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Редактор *И. П. Острикова*

Корректор *А. В. Тюхай*

Компьютерная верстка *В. М. Задоля*

Подписано в печать 01.07.2011. Формат 60x84 1/16. Бумага офсетная. Гарнитура «Таймс». Отпечатано на ризографе. Усл. печ. л. 10,81. Уч.-изд. л. 10,0. Тираж 100 экз. Заказ 439.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
ЛИ №02330/0494371 от 16.03.2009. ЛП №02330/0494175 от 03.04.2009.
220013, Минск, П. Бровки, 6