



Уральский
федеральный
университет

имени первого Президента
России Б.Н.Ельцина

Институт радиоэлектроники
и информационных
технологий — РТФ

А. Л. КРОХИН

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА С ЭЛЕМЕНТАМИ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Учебное пособие



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
УРАЛЬСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ ПЕРВОГО ПРЕЗИДЕНТА РОССИИ В. Н. ЕЛЬЦИНА

А. Л. Крохин

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА С ЭЛЕМЕНТАМИ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Учебное пособие

Рекомендовано методическим советом

*Уральского федерального университета в качестве учебного пособия
для студентов, обучающихся по направлениям подготовки*

09.03.01 – Информатика и вычислительная техника,

09.03.04 – Программная инженерия, 09.03.03 – Прикладная информатика,

10.03.01 – Информационная безопасность, 11.03.01 – Радиотехника,

11.03.02 – Инфокоммуникационные технологии и системы связи,

11.03.03 – Конструирование и технология электронных средств,

26.03.04 – Управление в технических системах,

29.03.03 – Технология полиграфического и упаковочного производства,

10.05.02 – Информационная безопасность телекоммуникационных систем

Екатеринбург

Издательство Уральского университета

2022

УДК 519.1(042.4)

ББК 22.176я73

К83

Рецензенты:

А. В. Наумов, доктор физико-математических наук
(Московский авиационный институт

(национальный исследовательский университет));

Ю. Б. Мельников, кандидат физико-математических наук, доцент
(Уральский государственный экономический университет);

С. С. Ланно, кандидат технических наук, доцент
(АО «Ведение реестров компаний»)

Крохин, А. Л.

К83

Теория вероятностей и математическая статистика с элементами дополненной реальности : учебное пособие / А. Л. Крохин ; Министерство науки и высшего образования Российской Федерации, Уральский федеральный университет. — Екатеринбург : Изд-во Урал. ун-та, 2022. — 190 с. : ил. — 30 экз. — ISBN 978-5-7996-3577-0. — Текст : непосредственный.

ISBN 978-5-7996-3577-0

В учебном пособии изложены основные разделы курса «Теория вероятностей и математическая статистика», в том числе рассмотрены функции случайных величин и свойства случайных функций, а также элементы математической статистики. Подробно разобраны решения основных математических задач, важных для инженерной практики и способствующих более глубокому усвоению материала. Для интенсификации работы с учебным материалом пособие снабжено QR-интерфейсом с веб-платформой, на которой располагаются дополнительные учебно-методические и информационные ресурсы.

Для студентов Института радиоэлектроники и информационных технологий — РТФ и Физико-технологического института УрФУ. Может быть использовано и для других направлений подготовки и специальностей инженерных институтов.

УДК 519.1(042.4)

ББК 519.1(042.4)

Изображение на обложке:
идея и дизайн Л. А. Крохина, выполнено
с помощью нейросети Stable Diffusion
на веб-сервисе dreamstudio.ai

ISBN 978-5-7996-3577-0

© Уральский федеральный университет, 2022

Оглавление

Предисловие	6
Обозначения и сокращения	9
Глава 1. Случайные события	12
1.1. Стохастичность окружающего мира	12
1.2. Пространство элементарных событий	17
1.3. Примеры построения пространства элементарных событий	21
1.4. Операции над событиями. Алгебра событий	25
Глава 2. Вероятность. Различные подходы	27
2.1. Статистический подход к вероятности	27
2.2. Классическая вероятность	30
2.3. Комбинаторные вычисления	32
2.4. Достоинства и недостатки классической модели	42
2.5. Геометрическая вероятность	43
Глава 3. Аксиоматический подход к вероятности. Основные формулы теории вероятностей	50
3.1. Идеология аксиоматического подхода	50
3.2. Простейшие вероятностные модели	52
3.3. Формула сложения	54
3.4. Условная вероятность. Теорема умножения	57
3.5. Обсуждение термина <i>независимые события</i>	60
3.6. Композиция испытаний	61
Глава 4. Основные формулы (продолжение)	65
4.1. Схема гипотез. Формулы полной вероятности и Байеса	65
4.2. Последовательности испытаний	68
4.3. Схема Бернулли	69
4.4. Формулы Лапласа и Пуассона	70

Глава 5. Случайные величины	76
5.1. Случайная величина: определение понятия	76
5.2. Дискретные случайные величины	78
5.3. Числовые характеристики случайных величин	79
5.4. Биномиальное распределение	81
5.5. Случайные величины непрерывного типа	83
5.6. Функция распределения случайных величин	84
Глава 6. Некоторые замечательные распределения	87
6.1. Распределение Пуассона	87
6.2. Равномерное распределение	88
6.3. Показательный (экспоненциальный) закон распределения	89
6.4. Нормальный закон распределения	92
Глава 7. Случайные векторы (системы случайных величин)	95
7.1. Закон распределения случайного вектора с дискретными составляющими	96
7.2. Закон распределения случайного вектора с непрерывными составляющими	100
7.3. Функция распределения случайного вектора	101
Глава 8. Зависимые и независимые случайные величины	104
8.1. Свойства математического ожидания и дисперсии	105
8.2. Корреляционный момент. Коэффициент корреляции	106
Глава 9. Неслучайные функции случайного аргумента	110
9.1. Функция случайной величины	110
9.2. Функция непрерывного случайного аргумента	112
9.3. Числовые характеристики функций	116
9.4. Закон распределения функции случайного вектора	119
Глава 10. Распределение суммы, произведения и отношения двух случайных величин	121
10.1. Задача композиции	123
10.2. Композиционная устойчивость нормального распределения	127
10.3. Распределения Стьюдента и χ^2	128
10.4. Распределение Рэлея	131
Глава 11. Предельные теоремы теории вероятностей	134
11.1. Сходимость по вероятности. Неравенство Чебышёва	134

11.2. Закон больших чисел в форме Бернулли и в форме Чебышёва	136
11.3. Формулы Лапласа. Понятие о центральной предельной теореме	138
Глава 12. Случайные функции	141
12.1. Основные понятия	141
12.2. Случайные функции. Корреляционная теория	142
12.3. Дифференцирование и интегрирование случайных функций	144
12.4. Стационарные случайные процессы	148
12.5. Спектральное разложение стационарного случайного процесса	149
Глава 13. Математическая статистика	152
13.1. Основные понятия	152
13.2. Описательная статистика. Визуализация	153
13.3. Оценки параметров распределения случайной величины	158
13.3.1. Точечная оценка параметров распределения	158
13.3.2. Интервальные оценки параметров распределения	159
13.4. Оценка параметров нормального распределения	161
13.5. Оценка параметра p распределения Бернулли	164
13.6. Проверка статистических гипотез	168
13.7. Критерий Пирсона χ^2	170
Глава 14. Элементы дисперсионного и корреляционного анализа	176
14.1. Дисперсионный анализ	176
14.1.1. Описание методики	176
14.1.2. Постоянное число наблюдений	178
14.2. Регрессионный анализ в теории вероятностей	180
14.3. Тема и содержание учебной лабораторной работы	184
Список библиографических ссылок	186

Предисловие

Одной из целей математического образования инженера является знакомство с математическими моделями, используемыми в специальных инженерных дисциплинах. Особенность радиотехнических задач состоит в важности учета случайности, связанной как с физической природой электромагнитных процессов в веществе, так и с большой сложностью изучаемых систем. В теории вероятностей основной объект — случайная величина, т. е. числовая функция, заданная на пространстве элементарных событий. Это удобная математическая модель характеристики системы, подверженной случайному воздействию. Однако методы теории вероятностей применимы лишь для достаточно идеализированной статической постановки задачи. Если же нас интересует эволюция системы во времени, а на нее непрерывно оказывают влияние случайные факторы, то аппарата теории вероятностей становится недостаточно. Для этих целей используется теория случайных функций. Поэтому знакомство с ее элементами совершенно необходимо каждому студенту радиотехнической специальности.

В учебном пособии сначала даются основные понятия теории вероятностей: случайное событие и вероятность. Затем выводятся основные формулы ТВ и предлагается их наглядная интерпретация для решения практических задач. Еще одно из важнейших понятий — случайная величина определяется с формальной точки зрения, как функция особого вида, отображающая пространство случайных событий на числовое множество. Такой подход принят в университетах и нашей страны, и зарубежных для специальностей с повышенной математической подготовкой, однако в данном пособии он несколько «облегчен» с учетом целевой аудитории. Специфический способ задания такой функции — закон распределения случайных величин — определяется и обсуждается отдельно для дискретных и непрерывных случайных величин, случайных векторов и функций случайного аргумента. Затем вводится по-

нятие случайной функции, изложены существующие подходы к описанию и классификации случайных функций. Корреляционная теория, как наиболее практически важная, рассматривается применительно к стационарным случайным функциям.

В последних главах даются основные понятия математической статистики, обсуждается ее роль и место в статистических исследованиях. Рассмотрены задачи первичной обработки данных, оценки параметров распределений, проверки статистических гипотез и понятия о корреляционном и регрессионном анализе.

Многие иллюстрации, справочные сведения и задачи доступны посредством технологии добавленной реальности через QR-метки. Для этого читателю предлагается сканировать посредством смартфона или иным способом QR-метку (рис. 1), пройти по появившейся ссылке и затем следовать указаниям.

Пособие поможет читателю-студенту лучше ориентироваться в учебной литературе по теории вероятностей.

Можно порекомендовать посмотреть, хотя бы бегло, основополагающие работы А. Н. Колмогорова [1, 2], классический учебник Е. Н. Вентцель [3]. Основательная книга В. С. Пугачева [4] позволит любознательному студенту лучше разобраться в некоторых темах (например, свойства функций случайных величин), традиционно трудно воспринимаемых студентами инженерных специальностей. В решении задач поможет книга, прямо ориентированная на инженерные приложения [5]. Первую часть знаменитого двухтомника В. Феллера [6] осилит далеко не каждый студент, но более упорные и подготовленные найдут там много интересного и полезного. Многие годы в ведущих университетах нашей страны используются учебники В. П. Чистякова [7], Ю. А. Розанова [8] и др. У Б. В. Гнеденко [9] можно подробно ознакомиться с историей теории вероятностей [9]. Более детальное изложение случайных процессов можно посмотреть в работе С. М. Рытова [10], а также популярном в МАИ базовом курсе теории вероятностей [11]. Последний учебник по уровню сложности приближается к математическим курсам.

Доступное и довольно обстоятельное изложение комбинаторики — без которой задачи теории вероятностей не решить — читатель может найти в известной книге [12].



Рис. 1. Метка для доступа к сетевым дополнениям

В настоящем учебном пособии широко используются визуализации различного вида, причем представленные не только как традиционные рисунки — структурные элементы текста с номером и подписью. Иконические и символические нелитеральные знаки непосредственно внедряются в текст как его неотъемлемая составляющая, поэтому подобные «квазирисунки» не нумеруются.

В последние годы такой процесс наблюдается в самых разных текстах — и полиграфических, и сетевых. Лингвисты называют их поликодовыми [13], отмечая смешение различных семиотических систем, при этом в физико-математических и химических текстах оно ни в коем случае не является синкретическим. «Например, в физико-математических, химических текстах, в текстах, посвященных техническим разработкам, формулы, символические изображения, графики, таблицы, технические рисунки, геометрические фигуры и другие изобразительные элементы являются смысловыми компонентами текста, передающими основное содержание текста» [14]. Как уже было отмечено выше, многие компоненты поликодового текста вынесены за пределы собственно страниц книги. Это касается как иллюстраций, особенно анимационного характера, так и гипертекстовых, и мультимедийных фрагментов. Такая организация дидактических материалов реализует специфическую форму технологии дополненной реальности¹. Содержание страниц книги дополняется изображением на экране смартфона [15].

Следуя политике УрФУ, направленной на интернационализацию образовательного процесса, в настоящем пособии упоминаются основные особенности англоязычной терминологии зарубежных курсов ТВ и МС, а также даются соответствующие синонимы наименований основных понятий, принятых в отечественной традиции.

Набор и верстка данного пособия были выполнены автором в замечательной и популярной во всем мире системе верстки научных текстов $\text{LaTeX}2\epsilon$ [16], созданной самоотверженным многолетним трудом Дональда Кнута и безвозмездно переданной Миру.

Все иллюстрации были сделаны в графических расширениях $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}2\epsilon$, созданных последователями Кнута и совершенствуемых постоянным упорным трудом многих энтузиастов высочайшей квалификации, объединенных в глобальное сообщество $\text{T}_{\text{E}}\text{X}$ USER GROUP [17].

¹ Данный термин является почти дословным переводом оригинального: *augmented reality*.

Обозначения и сокращения

В учебном пособии автор придерживался той системы терминологии и обозначений, которая была использована в основополагающей работе А. Н. Колмогорова [1] (на русском языке см. [2]).

ЭСИ — эксперимент со случайным исходом.

Ω — пространство элементарных событий (ПЭС).

ω — элементарное событие или исход.

$\{\omega\}$ — одноэлементное случайное событие (не смешивать с исходом).

$\{\omega_1, \omega_2, \dots, \omega_n, \}$ — [сложное] случайное событие, состоящее из нескольких исходов; детализированная запись содержит составляющие его элементарные события.

$A, B, C \dots$ — события, краткая запись (используются заглавные латинские буквы первой половины алфавита).

$X, Y, Z \dots$ — числовые множества действительных чисел (используются заглавные латинские буквы конца алфавита).

$A \cup B$ ($A + B$), $A \cap B$ ($A \cdot B$), $A \setminus B$ ($A - B$), \bar{A} — обозначения соответственно суммы, произведения, разности двух событий и противоположного события (в скобках — сокращенный вариант записи).

ξ, ξ_i, η, ζ — случайные величины (СВ).

F_ξ — функция распределения случайной величины ξ .

f_ξ — плотность [распределения] вероятности случайной величины ξ .

СКО — среднее квадратичное отклонение.

МО — математическое ожидание.

QR(Quick Responce) — графический примитив, кодирующий небольшой фрагмент текста; удобен как машиночитаемое представление URL — адреса документа в сети Интернет.

■ — конец решения.

□ — конец доказательства.

ТВ и МС — общепринятое сокращение названий математических дисциплин *теория вероятностей* и *математическая статистика*.

ИТ — информационные технологии.

ООФ — область определения функции.

МЗФ — множество значений функции.

ЗБЧ — закон больших чисел.

Случайный. 1. Возникший, появившийся
непредвиденно. 2. Бывающий лишь изредка...

С. И. Ожегов

*Вероятность есть число, вокруг которого имеют
тенденцию группироваться частоты, причем при
возрастании численности этих серий эта
тенденция проявляется со все большей
отчетливостью и точностью...*

А. Н. Колмогоров

Глава 1

Случайные события

1.1. Стохастичность окружающего мира

Окружающий нас мир многообразен и сложен. Человечество занимается его изучением столько времени, сколько оно существует. Основой наших знаний всегда было наблюдение, выделение каких-то повторяющихся при одних и тех же условиях событий.

Поэтому одной из древнейших наук является астрономия. Сначала люди замечали и описывали то, что они наблюдали – смену дня и ночи, времен года, а затем пытались дать этому объяснение. Не всегда то, что видит наблюдатель, соответствует действительному положению вещей. Так, видимое суточное движение Солнца по небосводу далеко не сразу связали с осевым вращением Земли. Заметим, что пресловутый вопрос ВЦИОМ «Что вращается: Земля вокруг Солнца или Солнце вокруг Земли?» совершенно некорректен, поскольку механическое движение по определению рассматривается относительно выбранной системы отсчета.

Классическая механика впервые приобрела статус настоящей науки благодаря гению Ньютона, сформулировавшему свои знаменитые три закона, причем отнюдь не очевидные. Жизненный опыт, скажем, говорит о том, что двигать, перемещать предметы можно только прикладывая определенные усилия. И лишь анализ многочисленных наблюдений за движением реальных тел позволил сформулировать первый закон механики, по которому движение с постоянной скоростью не требует никакого воздействия. Современная механика — это наука, позволяющая не только понимать закономерности движения физических тел, но и предсказывать на многие годы и с высокой точностью траектории планет и искусственных космических аппаратов. Ярким

примером является проект Voyager [18], с содержанием которого рекомендуем ознакомиться, используя данный QR.

В упомянутом проекте главными были выбор времени старта и траектории, использующей «гравитационный маневр» (ГМ) [19]. Начальный импульс был выбран таким, чтобы при приближении аппарата к Юпитеру существенно изменить его траекторию, используя гравитационное поле гигантской планеты. Тонкость расчета состояла в том, чтобы разогнать аппарат, но избежать падения на Юпитер. Рассчитанная таким образом траектория позволяла «посетить» Сатурн, а потом и Уран с Нептуном. Так были получены прекрасные фотографии этих планет и их спутников. Время на прохождение всех четырех исследованных планет сократилось до 12 лет (по сравнению с 30 годами без использования ГМ — по гомановской траектории). Причем космический аппарат сберег свой ограниченный запас топлива и других ресурсов.

Подробное объяснение ГМ как интересного примера применения классической механики дается, например, в статье [19]. Заметим, что ГМ применялся и в других проектах, а впервые ГМ совершила советская «Луна-3»¹ в 1959 г.

Время запуска «Вояджеров» было приурочено к так называемому «параду планет» — все четыре планеты (Юпитер, Сатурн, Уран и Нептун) оказываются на одной прямой линии и могут быть исследованы в одном полете. Эта конфигурация длится непродолжительное время и повторяется только через 176 лет. Для планируемого проекта стало большой удачей, что ближайшее «окно» для начала экспедиции ожидалось в интервале с 1975 по 1977 г. Именно в этом промежутке времени и рассчитывались моменты старта обоих «Вояджеров» — 20 августа 1977 г. и 5 сентября 1977 г.

14 февраля 1990 г. был сделан снимок нашей голубой планеты с самого большого расстояния от Земли. Знаменитый снимок «Pale Blue Dot» был сделан по окончании основной миссии «Вояджера-1», когда аппарат находился примерно в 6 миллиардах (!) километров от нас. И наша земная наука позволила дистанционно навести фотообъектив в нужном направлении! Это реально серьезная задача — в космосе нет ни верха, ни низа. Как определить направление на Землю? Расположение звезд вокруг космического аппарата совершенно непохоже на ту картину, которую мы наблюдаем с поверхности нашей планеты. Снимок можно посмотреть в wiki и, кстати, сравнить с дру-




¹Советская автоматическая межпланетная станция для изучения Луны и космического пространства.

гим знаменитым снимком нашей планеты — «The Blue Marble». «Вояджеры» в космосе уже 44 года, за это время они преодолели порядка 20 млрд км. Это настоящая эпопея, богатая драматическими событиями, естественно, технического характера. Аппаратура, стоящая на борту, очень старая — ламповые передатчики, комплектующие и программное обеспечение 1970-х гг. Инженерам приходилось настраивать частоты и другие характеристики сигналов связи, математики и программисты модернизировали операционную систему, протоколы обмена. И все это дистанционно, на удалении сотен миллионов километров. Один сеанс обмена данными длился многие часы.

Научный подход к изучению некоего природного явления заключается в выделении наиболее существенных его свойств и пренебрежении менее важными. Практически все современные исследования природы проводятся в виде *экспериментов*, т. е. изучаемый объект или явление наблюдается в специально организованных условиях. Закономерность будет считаться имеющей место, если при одних и тех же условиях будет регистрироваться один и тот же результат. Познание человеком окружающего его мира началось с обнаружения эмпирических закономерностей — если мы наблюдаем некоторое событие *A*, то через некоторое время произойдет событие *B*. В дальнейшем на основании подобных наблюдений сформировались такие естественные науки, как астрономия, физика, химия. В свою очередь, развитие этих наук представило математике различные абстрактные понятия и модели, а также и основной исследовательский метод — логическое рассуждение. В классической логике применяется четкая цепочка: **причина** \Rightarrow **следствие**. Чем лучше ученые понимали и постигали закономерности окружающего мира, тем большее удивление вызывали такие ситуации, в которых при идентичных начальных условиях результат предсказать было совершенно невозможно.



Археологические исследования показывают, что предмет , который мы теперь называем *игральной костью*, был известен многим народам тысячи лет назад. Людей чрезвычайно впечатляла полная невозможность предсказать результат его падения. Появилась популярная игра «в кости», которая, кстати, привлекла внимание математиков того времени, чьи наблюдения за азартными играми положили начало новой математической теории.

С точки зрения механики все дело в том, что небольших размеров кубик из материала малой плотности и со скругленными ребрами очень неустойчив. Процедура выбрасывания из стаканчика нескольких таких кубиков приводит к полной невозпроизводимости начальных условий — векторов скоростей поступательного и вращательного движения. К тому же небольшие изменения

ориентации кубика в момент отскока от стола еще сильнее запутывают ситуацию. Так что «многократное повторение опыта в идентичных условиях» в контексте игры в кости означает лишь традиционные действия игрока. В этом отношении школьная задача о движении тела, брошенного под углом к горизонту, относится к очень мало чувствительному к начальным условиям движению. Хотя и здесь имеется неопределенность начальных условий. Поэтому в реальности снаряды, выпущенные из одной и той же пушки при постоянной наводке, не попадают в одну точку. Правда, разброс этих точек на много порядков меньше расстояния до самой пушки.

Некоторые снайперы из специальных винтовок, используя особые патроны (очень точная дозировка массы пороха и высокая точность изготовления пуль) и компьютерные системы наведения, могут многократно поразить цель на дистанциях более километра [20] (сканируйте QR).

А вот результат падения специальным образом подброшенной монетки непредсказуем (именно поэтому таким способом осуществляется в футболе распределение ворот между командами). Процедура выбрасывания монетки (неконтролируемые начальные скорости поступательного и вращательного движения) не позволяет достаточно точно повторить начальные условия. Однако в этом случае можно ожидать, что зоркий и чрезвычайно ловкий человек может «накрыть» падающую монету в нужной ему фазе вращения.



Карточные игры также «завязаны» на ограниченность органолептической чувствительности обычного человека. Поэтому со стороны рубашки карты одинаковы на вид и на ощупь неразличимы. Однако для хитроумных шулеров в раскладке карт никакой случайности нет!

Кроме хорошо известных азартных игр, внимание ученых привлекло *броуновское движение*. Это очень интересное явление (обнаружено Р. Броуном в 1827 г.), причем до некоторого времени не только непонятное с точки зрения классической физики, но, пожалуй, даже загадочное. Маленькие частички краски в воде вели себя как живые — беспорядочно дергались, постепенно смещаясь в поле зрения микроскопа. Предложенное Эйнштейном в 1905 г. объяснение и теоретический вывод основной закономерности опирались на совершенно новые идеи: а) молекулярного строения вещества и б) статистического описания его свойств [21]. Мы такими сложными задачами заниматься не будем. Однако качественное разъяснение дать можно. Частицы краски были достаточно велики (порядка 10^{-6} м), чтобы их можно было наблюдать в оптический микроскоп. Но площадь поверхности и мас-

са каждой частички все же были таковы, что удары молекул жидкости с разных направлений не компенсировались. Это приводило к ее постоянному беспорядочному смещению то в одну, то в другую стороны. Более детальное рассмотрение явления см. по ссылке [22] (сканируйте QR).

Другие многочисленные примеры «случайности» связаны с отсутствием знания, информации. Сел человек в лифт, а где он живет, к кому пришел мы не знаем, он может выйти на любом этаже. В одной американской телевизионной игре (Let's Make a Deal) ведущий предлагал игроку угадать, за какой из трех дверей находится приз. После того как игрок указывал выбранную дверь, открывалась одна из оставшихся и все видели, что там приза нет. Ведущий (он знает, где приз) предлагал игроку возможность повторного выбора. Фактически мы получаем довольно хитроумную задачу теории вероятностей, известную как Monty Hall problem. Решение ее бурно обсуждалось на форумах многие годы.



Диагностика и лечение болезней — еще одна область для применения теоретико-вероятностных и статистических методов. Устройство человеческого организма настолько сложно и индивидуализировано, что с определенностью, детерминированно предсказать реакцию конкретного человека на некое лекарство невозможно. Здесь возможен только статистический подход. К сожалению, тема актуальная и больная.

Отдельно надо сказать о квантовой механике. Наука исходит из особых свойств микрочастиц, изучение которых чрезвычайно опосредовано и основано на логической интерпретации косвенных измерений.

Можно упомянуть опыты Штерна и Герлаха [23], которые подтвердили принципиальную неразличимость микрочастиц и дискретность их состояний, наличие спина [21, с. 713]. Поэтому классическая механика совершенно неприменима для описания свойств и движения микрочастиц. Состояние микрочастицы описывается *волновой функцией*, имеющей прямой теоретико-вероятностный смысл и являющейся решением уравнения Шредингера.



В настоящем пособии мы ограничимся рассмотрением таких явлений, в которых случайность определяется их природой либо ограниченностью наших знаний.

Одно из основных понятий теории вероятностей (ТВ) — *случайное событие*. Во избежание недоразумений будем в дальнейшем различать употребление этого словосочетания в *широком смысле*, т. е. применительно к *реаль-*

ным событиями, событиям, которые можно наблюдать, которые происходят с реальными объектами окружающего нас мира, и в *узком смысле* — как абстрактного понятия математической теории.

Случайность реального события подразумевает некоторую неопределенность результата наблюдения при одних и тех же начальных условиях. Как иногда говорят, случайное событие может произойти или не произойти. Разберемся с этим понятием подробнее, а затем перейдем к математической терминологии.

1.2. Пространство элементарных событий

Начала ТВ связаны с изучением закономерностей азартных игр: кости, карточные игры, рулетка, различные лотереи. У всех азартных игр имеется общее: первый этап (выбрасывание костей из стаканчика, тасование и сдача карт, запуск колеса рулетки. . .), который единообразно повторяется много раз, и второй этап — подсчет выпавших на костях очков, состав полученного при сдаче набора карт, место на рулеточном колесе, на котором остановился шарик... Весь интерес как раз в том, что предсказать заранее, что получится на втором этапе, невозможно. Как говорится, результат случаен.

Введем абстрактную модель, объединяющую основные характерные особенности азартных игр — *эксперимент со случайным исходом (ЭСИ)*. На рис. 1.1 представлена схема этой модели. Первая часть — *испытание*, она отображает условия, правила проведения игры, вторая — содержит все возможные *исходы*, состав ее также отчасти определяется правилами игры. Если игрок хочет разобраться с правилами подробнее, попытаться выявить какие-то закономерности, то ему придется уточнить, что именно лучше считать исходом.

Эксперимент может проводить сам наблюдатель, выбрасывая, например, игральный кубик. Может быть и гораздо более сложный механизм организации эксперимента, такой как «Спортлото» или другая лотерея. Наконец, это может быть и некий естественный процесс, происходящий безо всякого вмешательства человека. Отметим сразу элемент идеализации, имеющий место уже на начальном этапе изучения реального случайного явления. Мы не пытаемся отслеживать временную эволюцию объекта, т. е. изменение его состояния со временем, считаем, что оно, это изменение, происходит мгновенно.

Время лишь выделяет два этапа эксперимента, естественным образом упорядочивая их. Это *испытание*, включающее в себя исходное состояние объекта и воздействие, и результат, называемый *исходом* (см. рис. 1.1).

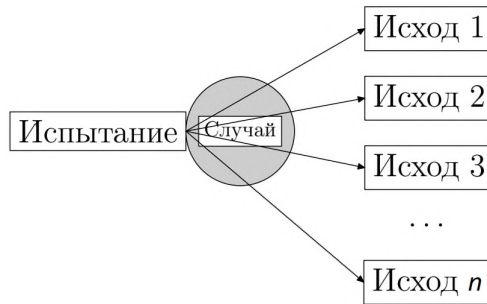


Рис. 1.1. Схема эксперимента со случайным исходом

Возможность *многократного*, хотя бы в принципе, повторения условий нашего эксперимента очень существенна и будет подробно обсуждаться ниже. Такое реальное событие называют **массовым**. Примеров случайных событий можно привести очень много. Случайным является результат выбрасывания монетки, игрального кубика, розыгрыша лотереи. Хотя результат футбольного матча также невозможно предсказать заранее, но для конкретных команд данное событие никак нельзя считать массовым и применять к нему все последующие результаты не имеет смысла.

Перейдем теперь к математическому пониманию термина *случайное событие*. Предполагаем, что можно заранее перечислить, описать все возможные результаты опыта, или *исходы*. Что именно подразумевать под *исходом* из описания нашего эксперимента не следует и должно быть специально оговорено, чтобы избежать недоразумений. Особенно надо помнить об этом при решении задач.

Вместо «исход» часто говорят *элементарное событие*, имея в виду, что более *детальная* фиксация результата эксперимента не производится. Естественно, что реально произойти может только одно событие из этого «списка» (говорят, что такие события *несовместны*). Форма монетки такова, что на твердую поверхность она может упасть гербом (Г) или цифрой (Ц) вверх.

Эмпирический факт — невозможно предсказать, какой стороной выпадет монетка. Под этим понимают невозможность каким-либо способом уверенно предсказывать результат выбрасывания. Это значит, что повторяя свое предсказание несколько раз, вы будете угадывать примерно столько же раз, сколько и ошибаться. Точно так же невозможно уверенно определить характеристики игральной карты, которая вынимается из тщательно перетасованной колоды. Теория вероятностей ничем не поможет вам в подобных

ситуациях: ни определить сторону монетки, ни масть и достоинство игральной карты. Если не интересоваться координатами места падения монетки (ее часто просто ловят рукой), то получается естественно два возможных исхода.

Для игры в кости кубики изготавливают из материала, плотность которого небольшая (кость!), этот кубик чрезвычайно верткий, он поворачивается при малейшем воздействии. Бросающему нет никакой возможности добиться того, чтобы кубик упал определенным заранее образом. Место падения на столе, ориентация во внимание не принимается — так уж повелось, правила игры такие. Игральный кубик может выпасть любой из своих шести граней вверх, значит — шесть исходов.

Именно поэтому считаем возможными ровно шесть исходов по числу граней для каждого из бросаемых кубиков — мы еще вернемся к этой игре.

В других примерах число возможных исходов будет иным. Бывают эксперименты, в которых затруднительно ограничить количество исходов, и его удобно считать бесконечным. Скажем, по условию монетка выбрасывается до первого появления герба. Нет никакой разумной причины ограничить максимальное число бросаний.

Итак, наличие более одного возможного исхода — вот в чем заключается «случайность» реального явления. Множество возможных исходов полностью содержит всю «индивидуальность» случайного явления и может рассматриваться как основа для его изучения. Результатом испытания будет наступление одного и только одного из исходов. Примеры показывают, что никаких разумных ограничений на свойства этого множества предложить нельзя, оно может быть совершенно произвольным. Конкретная природа составляющих его элементов для математики несущественна, они первичны и элементарны, подобно точкам в геометрии. Приведем теперь формальные определения.

Определение. Произвольное множество Ω назовем **пространством элементарных событий** (ПЭС). Элементы ω этого множества будем называть *элементарными событиями*, или *исходами*.

Надо отметить, что описание того, что именно считать исходом, определяет сам исследователь с учетом особенностей данного процесса. Чуть ниже мы подробнее поговорим об этом.

Можно сформулировать условия, налагаемые на результат эксперимента, так, что им будут удовлетворять несколько исходов. Скажем, в эксперименте с выбрасыванием кубика можно наблюдать событие «число выпавших очков

четное». Такие события называют *сложными* или просто событиями, без прилагательного, а математической моделью для них будет любое подмножество $A \subseteq \Omega$ элементов из Ω .

Принципы построения ПЭС

1. **Детализация.** При решении прикладных задач очень важно выделить то, что *существенно* для понимания природы изучаемого явления, и не принимать во внимание несущественное, только затрудняющее решение задачи.
2. **Полнота.** В то же время нельзя пренебрегать важными свойствами объекта.
3. **Несовместность.** Формулировка того, что считать исходом данного ЭСИ, должна исключать фиксацию одновременного наступления двух и более исходов. Результатом может быть *только один* исход.

Определение. *Случайным событием* называют произвольное подмножество ПЭС Ω данного ЭСИ.

Ω содержит все возможные исходы, такое событие называют *достоверным событием*. Это математическая «тень», модель реального события, которое обязательно произойдет в результате данного ЭСИ, мы будем его обозначать через **U** или **D**.

Невозможное событие — это событие, обозначаемое **V** или **H**, не содержащее ни одного исхода.

Это название также достаточно удачное, поскольку соответствует реальному событию, которое не может произойти в условиях данного ЭСИ.

Нельзя, однако, забывать, что в ТВ мы имеем дело с моделью, а не реальным явлением. Составляя ПЭС, мы не учитывали несущественные, по нашему мнению, результаты эксперимента. А в реальности монетка может закатиться в щель, игральный кубик упадет со столика...

В своей знаменитой книге В. Феллер отмечал: «... природа пространства элементарных событий не играет роли для нашей теории. Для нас это пространство (вместе с заданным на нем распределением вероятностей) определяется идеализированным экспериментом» [6, с. 20].

1.3. Примеры построения пространства элементарных событий

Решение задач по теории вероятностей (а они содержат, как правило, описание некоторых, если так можно выразиться, житейских ситуаций) начинается с моделирования, формализации этих ситуаций. Иначе говоря, необходимо выполнить построение *пространства элементарных событий*.

Для этого, во-первых, мы определяем, что именно будем считать элементарным событием, исходом данного опыта, а во-вторых, описываем множество этих исходов. В рассмотренных ранее примерах с монеткой и кубиком все очевидно (вернее, существует традиция). В других задачах построение ПЭС может быть сделано несколькими способами и заслуживает специального обсуждения.

Рассмотрим выбрасывание двух одинаковых монеток. Какое выбрать множество исходов? Трехэлементное $\{ГГ, ГЦ, ЦЦ\}$ или четырехэлементное $\{ГГ2, ГЦ2, ЦГ2, ЦЦ2\}$? С логической точки зрения оба варианта возможны. Мы же предпочитаем второй, поскольку он лучше отражает *природу* данного эксперимента, позволяет объяснить, почему комбинация $\{ГЦ\}$ выпадает чаще, чем $\{ГГ\}$ или $\{ЦЦ\}$.

Дело в том, что монетки, пусть даже совершенно новые, только из-под пресса, в принципе различимы. Можно пометить монетки очень маленькой меткой или радиоактивным изотопом, и это никак не повлияет на их движение во время полета. Далее, отметку можно проверить, а можно и не проверять. Это также на характер движения повлиять не может. Монетки «не знают», будем мы их различать или нет. Значит, удобнее и при выборе модели фиксировать порядок символов Г и Ц, т. е. в качестве ПЭС выбрать $\{ГГ2, ГЦ2, ЦГ2, ЦЦ2\}$.

Если еще добавить вполне естественное предположение о *равновозможности* исходов (последнее означает отсутствие предпочтительности одного или нескольких из них перед остальными), то мы получаем очень даже подходящую модель. Комбинация $\{ГЦ\}$ при бросании будет выпадать чаще, потому что это сложное событие, состоящее из двух элементарных — $\{ГГЦ2\}$ и $\{ЦГ2\}$.

А вот электроны неразличимы в принципе. И если придумать эксперимент, аналогичный выбрасыванию монеток, то электроны, имеющие направление спина вверх или вниз, будут иметь либо два варианта параллельной ориентации спинов, либо вариант антипараллельной. Более подходящей будет первая модель — $\{ГГ, ГЦ, ЦЦ\}$.

В литературе по ТВ можно встретить следующую легенду, в которой фигурирует эксперимент по выбрасыванию трех одинаковых игральных костей. Некий игрок в кости обратился за разъяснениями к знаменитому Паскалю (в других книгах к Галилею или Гюйгенсу — в общем, к большому математическому «авторитету»). По мнению этого игрока, сумма очков на трех костях должна одинаково часто принимать значения 11 и 12. Ведь каждое из этих чисел можно представить в виде суммы трех натуральных слагаемых шестью способами. Действительно,

$$11 = 1 + 5 + 5 = 1 + 4 + 6 = 2 + 3 + 6 = 2 + 4 + 5 = 3 + 3 + 5 = 3 + 4 + 4,$$

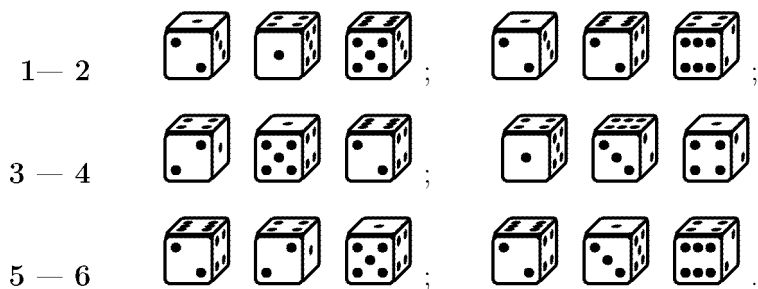
$$12 = 1 + 5 + 6 = 2 + 4 + 6 = 2 + 5 + 5 = 3 + 4 + 5 = 3 + 3 + 6 = 4 + 4 + 4.$$

На практике же 11 выпадает чаще, чем 12. Парадокс легко разрешим, если принять во внимание принципиальную различимость игральных костей. Ведь суть этого парадокса в несоответствии результата, предсказанного в рамках модели, и реально наблюдаемого. Игральные кости двигаются, согласно законам классической механики, как различные макроскопические тела. Поэтому исходом надо считать *упорядоченный* набор трех чисел из $\{1, 2, 3, 4, 5, 6\}$. Тогда подсчет количества исходов, «благоприятствующих» событиям «сумма 11» или «сумма 12», надо проводить иначе. Комбинация $\{1 + 5 + 5\}$ — не элементарное событие и может реализоваться тремя способами:

$$1 + 5 + 5 = 5 + 1 + 5 = 5 + 5 + 1,$$

так как единичка может выпасть на любой из трех костей.

Комбинация $\{1 + 4 + 6\}$ реализуется шестью способами:



А вот комбинация $\{4 + 4 + 4\}$ есть исход. Попробуйте теперь подсчитать сами — и вы получите для 11 очков 27 исходов, а для 12 очков — 25. Полное число возможных исходов (число элементов ПЭС) будет $6 \cdot 6 \cdot 6 = 216$.

Точно так же поступаем и при рассмотрении эксперимента с выбрасыванием двух кубиков (см. рис. 1.2). Модель, адекватно учитывающая свойства













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

Рис. 1.2. Наглядное представление пространства элементарных событий для эксперимента с выбрасыванием двух кубиков

реального эксперимента, содержит 36 исходов. В этом случае также различным значениям суммы очков будет соответствовать разное число исходов. Мы еще вернемся к этому эксперименту на с. 76.

Задача 1. Двое уговорились встретиться на следующих условиях. Каждый появляется на месте встречи в произвольный момент времени, скажем, от 13 до 14 часов, ожидает другого не более 10 минут, затем уходит. Построить пространство элементарных событий (ПЭС).

Решение

В качестве исхода будем фиксировать моменты прихода участников на место встречи (с бесконечной точностью — это уже часть нашей идеализации реальной задачи). Получаем множество упорядоченных пар чисел $\{t_1, t_2\}$, каждое из которых принимает любое значение из $[13; 14]$. Для простоты можно отсчитывать время в минутах от 13 часов. Геометрическим образом ПЭС будет квадрат (рис. 1.3). Теперь построим событие $A = \{\text{двое встретились}\}$. Условие встречи

$$|t_1 - t_2| \leq 10.$$

Нетрудно сообразить, что это будет своеобразная шестиугольная «полоска» вдоль диагонали квадрата. ■

Под описанное условие может подходить протокол встречи резидента

и связника. Случайность момента прихода обуславливается соображениями конспирации — каждый должен убедиться, что за ним не следят. Этой задаче можно придать не только детективно-шпионский оттенок. Можно говорить о приемном устройстве, которое обрабатывает полученный сигнал в течение некоторого промежутка времени. Если второй сигнал поступит, пока устройство занято, то он не будет принят. Таким устройством является и телефонный узел, и банковский компьютер, принимающий переводы, и счетчик Гейгера, и многое другое. Аналогичная ситуация возникает в порту, если к причалу одновременно приходят два парохода. Разгрузка подошедшего первым занимает некоторое время, в течение которого второй будет ждать.

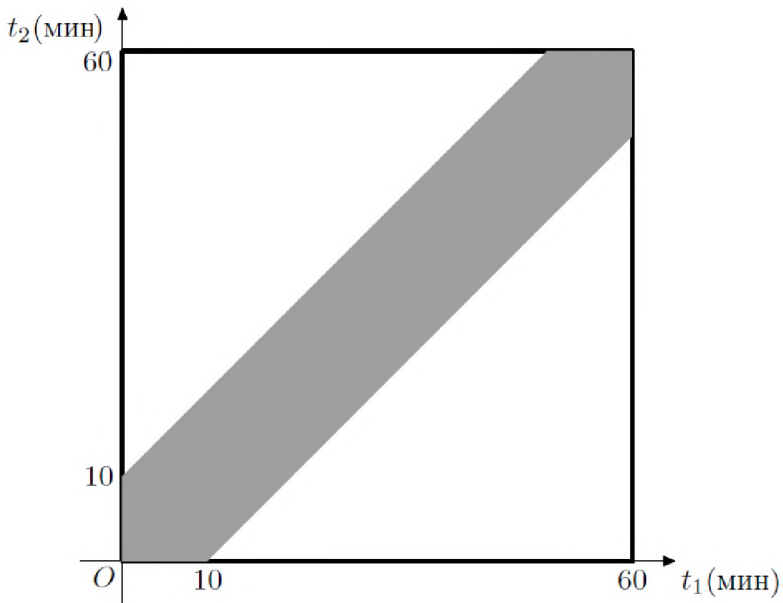


Рис. 1.3. ПЭС для задачи о встрече

1.4. Операции над событиями. Алгебра событий

На множестве случайных событий, связанных с данным ЭСИ, можно определить формальные операции.

События A_1 и A_2 называются *равными*, если они состоят из одних и тех же исходов, т. е. при теоретико-множественной интерпретации равенство событий — это равенство множеств.

$$A_1 = A_2 \Leftrightarrow (\omega \in A_1 \Leftrightarrow \omega \in A_2).$$

Установить равенство — значит убедиться, что все исходы события A_1 входят в A_2 , а исходы A_2 — в A_1 . Заметим, что в реальной жизни установить, что событие, являющееся страховым случаем по договору страхования, и случившаяся с вами неприятность совпадают (равны), может потребовать значительных усилий.

Определение. Суммой событий A и B называется событие C , состоящее из исходов, принадлежащих хотя бы одному из суммируемых событий. В стандартных обозначениях теории множеств имеем $C = A \cup B$.

Краткое обозначение: $A + B = C$. Сумма событий может быть интерпретирована как реальное событие, состоящее в наступлении **хотя бы одного** из суммируемых событий (рис. 1.4).

Определение. Произведением событий A и B называется событие C , состоящее из исходов, принадлежащих **обоим** событиям-сомножителям. В стандартных обозначениях теории множеств имеем $C = A \cap B$.

Краткое обозначение: $A \cdot B = C$. Произведение можно интерпретировать как реальное событие, состоящее в одновременном наступлении **обоих** событий-сомножителей. При проектировании и программировании систем управления, при составлении финансовых договоров и контрактов, заключении соглашений очень важно правильно сопоставлять реальные события и их математические модели (рис. 1.4).

Определение. Разностью событий A и B называется событие C , состоящее из исходов события A , не принадлежащих событию B . В стандартных обозначениях теории множеств имеем $C = A \setminus B$.

Определение. Событием, противоположным [по отношению] к событию A , называется событие \bar{A} , состоящее из всех таких исходов $\omega \in \Omega$, которые не входят в данное событие (так называемое **дополнение** множества A до множества Ω всех элементарных исходов). $\bar{A} = \{\omega | \omega \in \Omega \setminus A\}$.

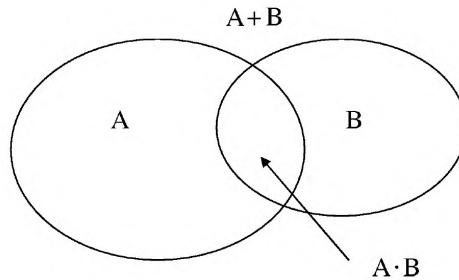


Рис. 1.4. Диаграмма Вена, иллюстрирующая сумму и произведение двух событий

Операции можно выполнять многократно, при этом имеют место следующие свойства.

Свойства операций:

1. $(A + B) + C = A + (B + C)$;
2. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$;
3. $(A + B) \cdot C = A \cdot C + B \cdot C$;
4. $\overline{A + B} = \overline{A} \cdot \overline{B}$;
5. $\overline{A \cdot B} = \overline{A} + \overline{B}$.

Доказательства опустим, они не очень сложные, познакомиться с ними можно, например, в работах [6] или [11].

Определение. События A и B называются **несовместными**, если $A \cdot B = \mathbf{V}$, т. е. произведение событий A и B есть невозможное событие. События A, B, \dots (список может быть бесконечным) называются **попарно несовместными**, если любые два различных события из этого списка несовместны.

Множество событий \mathcal{A} называют *алгеброй событий*, если выполняются три условия: а) $\Omega \in \mathcal{A}$; б) если $A \in \mathcal{A}$, то $\overline{A} \in \mathcal{A}$; в) если $A, B \in \mathcal{A}$, то и $A + B \in \mathcal{A}$.

Более сложная конструкция, чем ПЭС, понадобится в дальнейшем для строгого определения вероятности (для конечных ПЭС достаточно работать со множеством всех подмножеств Ω).

Глава 2

Вероятность. Различные подходы

Вероятность — одно из базовых понятий теории вероятностей (да и статистики тоже). Заметим, что используется это слово в самых различных контекстах и в разном смысле. В Толковом словаре русского языка Ожегова читаем: «Вероятность... Возможность исполнимости, осуществимости чего-либо» [24]. В обыденной речи о степени вероятности говорят на основании личного опыта, интуиции.

В естественных науках и в технике термин *вероятность* имеет более точную формулировку и означает *объективно существующую количественную меру возможности наступления случайного события* [21].

Обоснованием того, что случайному событию действительно можно сопоставить некоторое число, являются опыт, наблюдения за реальными явлениями, что, собственно, и отражает слово *объективно*.

2.1. Статистический подход к вероятности

Как уже отмечалось, понятие *вероятность* возникло из человеческого опыта как отражение объективно присущей случайным явлениям закономерности, называемой **устойчивостью относительной частоты**. Рассмотрим некоторое массовое случайное явление. **Массовость** подразумевает принципиальную возможность наблюдать это явление многократно в идентичных условиях. **Случайность** же состоит в том, что результатом может быть элементарное событие из некоторого множества. Будем проводить серию из n испытаний и фиксировать некоторое, вообще говоря, сложное событие A . Пусть оно наступило в m_A опытах.

Величину m_A называют *частотой* (в старых учебниках писали *ча́стость*) события A , а отношение

$$N_A(n) = \frac{m_A}{n}$$

относительной частотой события A . Многочисленные наблюдения показали, что относительная частота при больших n демонстрирует *устойчивость*. Иначе говоря, ее значения мало отличаются от некоторой величины N_A , одной и той же в различных сериях испытаний. Характер приближения относительной частоты к своему устойчивому значению будет более подробно рассмотрен позже, при знакомстве с законом больших чисел. Здесь же мы в качестве иллюстративного примера привели лишь график зависимости относительной частоты события от количества опытов в эксперименте, который проводился студентами на лекциях известного математика Б. В. Гнеденко и описан в его популярном учебнике [9]. Эксперимент состоял в следующем. Студенты принесли несколько колод карт, а затем сто раз провели разделение колод на две равные части. Наблюдалось событие A : {в каждой половинке одинаковое количество карт красной и черной масти}. Естественно, фиксировалось это событие не при каждом испытании (разделении колоды), поэтому относительная частота изменялась при увеличении количества испытаний немонотонно (см. рис. 2.1).

Во всех экспериментах такого рода, а их в истории известно множество, при увеличении числа испытаний до больших значений (сотни и тысячи) наблюдались небольшие нерегулярные колебания графика относительно некоторого «устойчивого» значения. Полученное значение можно принять за **вероятность** (статистическую) события A .

$$P(A) = N_A.$$

В учебниках здесь обычно ставится знак приближенного равенства, поскольку при повторении опыта значение частоты продолжает изменяться и отличается от точного N_A . Но это совершенно обычная ситуация при проведении измерений, наблюдаемая при выполнении любой лабораторной работы. И при измерении массы, и при измерении длины, и при измерении силы тока и напряжения цифры, стоящие в младших разрядах результата, не повторяются. Говорят, что результат любого измерения имеет определенную погрешность, причем часто она является принципиально неустранимой. Вот в этом смысле и надо понимать записанное равенство. Иначе говоря, полученная величина имеет некоторую погрешность.

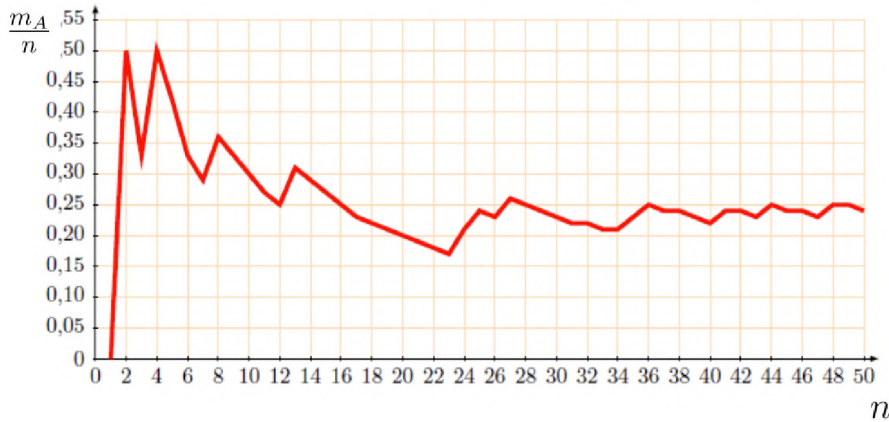


Рис. 2.1. Изменение относительной частоты случайного события $\frac{m_A}{n}$ при увеличении числа испытаний n

По следующей метке QR [25] вы можете посмотреть компьютерное моделирование известного эксперимента Бюффона с иглой.



Рассмотрим основные свойства полученной статистической вероятности.

Во-первых, ее возможные значения $0 \leq \mathbf{P}(A) \leq 1$. Равенства соответствуют невозможному событию, $\mathbf{P}(A) = 0$, и достоверному событию, $\mathbf{P}(A) = 1$. Чем ближе вероятность к нулю, тем реже наблюдается данное событие. Если же вероятность близка к единице, то такое событие будет наблюдаться часто. Подчеркнем, что сейчас речь идет об экспериментально наблюдаемом свойстве относительной частоты!

Рассмотрим теперь статистические вероятности двух *несовместных* событий A и B . Событие $A + B$ произойдет $m_A + m_B$ раз. Значит,

$$N_{A+B} = \frac{m_A + m_B}{n} = \frac{m_A}{n} + \frac{m_B}{n} = N_A + N_B.$$

И это реально наблюдается (с учетом уже упомянутой погрешности).

Для события \bar{A} , противоположного событию A , имеем

$$N_{\bar{A}} = \frac{n - m_A}{n} = 1 - \frac{m_A}{n} = 1 - N_A.$$

Другие полезные свойства статистической вероятности будут рассмотрены ниже при знакомстве с основными формулами ТВ.

Статистический подход к вероятности дает эмпирическую основу для построения математической теории. Свойства вероятности нашли отражение, как мы увидим далее, в аксиомах абстрактной вероятности. Статистические методы являются в некоторых случаях источником получения исходных вероятностей одних событий и входят как начальные данные в расчет вероятностей других событий. Кроме того, интерпретация результатов вероятностных расчетов также возможна лишь в рамках статистического подхода.

Конечно, данный подход сам по себе ограничен, поскольку статистическая вероятность не может быть получена иначе, чем в многочисленных экспериментах. Нередко их проведение практически невозможно, а в некоторых случаях хотя и возможно, но необоснованно дорого и сложно.

Следующий в нашем рассмотрении подход позволяет вычислять вероятность априори, без опытов, на основе некоторых предположений о свойствах изучаемого явления. Это *классическая формула вероятности*.

2.2. Классическая вероятность

Это, пожалуй, исторически первая возможность вычисления вероятностей. Пусть все события из множества $\{E_1, E_2, \dots, E_n\}$ элементарных исходов **равновозможны**, т. е. у нас нет никаких оснований предполагать, что какой-то исход будет встречаться чаще остальных. Это свойство исходов *не является математическим понятием*, оно предполагается интуитивно понятным и отражает свойства реального эксперимента, модель которого мы изучаем. По существу, предположение о равновозможности лишь ограничивает применимость полученных с его помощью результатов.

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

Функция **P**, определенная таким образом, называется **классической вероятностью**. Хрестоматийный вариант «озвучивания» этой формулы следующий: «Вероятность события *A* есть отношение числа исходов, *благоприятствующих* событию *A*, к *общему* числу возможных исходов».

Введение классической формулы свело теоретико-вероятностные задачи к комбинаторным. Нам надо подсчитать количество подмножеств конечного множества, удовлетворяющих определенным условиям. Надо сказать, что при этом приходится проводить довольно-таки хитроумные подсчеты [12].

Пример 1. В партии из 1000 изделий имеются 10 дефектных. Найти вероятность того, что среди 50 изделий, взятых наудачу из этой партии, ровно 3 окажутся дефектными.

Решение

Условие этой задачи типовое, использованный при решении подход может применяться и к другим подобным задачам.

Наречие «наудачу» не имеет четко определенного смысла и лишь оправдывает применение классической формулы. В практических методиках контроля качества реализация этого условия представляет серьезную проблему.

Общее число возможных исходов данного ЭСИ является *числом сочетаний* и равно: $n = C_{1000}^{50}$ (обозначение в отечественной нотации).

Вычислим m . Возьмем исход, входящий в искомое событие A . Он состоит в том, что три изделия из 10 — дефектные, а остальные 47 — годные. Два разных исхода, входящих в A , могут отличаться только тем, какие именно 3 дефектных изделия выбрали из 10 (таких возможностей C_{10}^3), и тем, какие 47 изделий выбрали из 990 недефектных изделий (таких возможностей C_{990}^{47}) (общее число выбранных изделий по условию 50). Не меняя дефектные детали, будем перебирать возможные наборы по 47 штук из годных. Количество возможных выборов стандартных изделий равно C_{990}^{47} . Изменим состав трех дефектных изделий и снова будем перебирать все составы годных. Следовательно, общее число благоприятных исходов равно:

$$\underbrace{C_{990}^{47} + C_{990}^{47} + \dots + C_{990}^{47}}_{C_{10}^3 \text{ слагаемых}} = C_{10}^3 \cdot C_{990}^{47}.$$

Итак, общее число благоприятных исходов равно: $m = C_{10}^3 C_{990}^{47}$. В комбинаторике приведенные выше очевидные соображения называют правилом произведения. Используя формулу $C_n^m = \frac{n!}{m!(n-m)!}$, получаем, что искомая вероятность равна:

$$\begin{aligned} p &= \frac{C_{10}^3 C_{990}^{47}}{C_{1000}^{50}} = \frac{10! \cdot 990!}{7!3! \cdot 47!943!} = \\ &= \frac{10 \cdot 9 \cdot 8 \cdot 50 \cdot 49 \cdot 48 \cdot 950 \cdot 949 \cdot 948 \cdot 947 \cdot 946 \cdot 945 \cdot 944}{2 \cdot 3 \cdot 1000 \cdot 999 \cdot 998 \cdot 997 \cdot 996 \cdot 995 \cdot 994 \cdot 993 \cdot 992 \cdot 991} \approx 0.01. \end{aligned}$$

■

Следствием определения классической вероятности являются несколько простых свойств:

1. $P(A) = \frac{m}{n} \geq 0$ для каждого события A .
2. Для достоверного события $P(\Omega) = 1$.
3. Пусть события A и B несовместны, причем A состоит из m_1 исходов, а B — из m_2 исходов. Тогда

$$P(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

4. Вероятность противоположного события равна:

$$P(\bar{A}) = \frac{n - m_A}{n} = 1 - \frac{m_A}{n} = 1 - P(A).$$

5. **Свойство монотонности.** Если событие A влечет событие B (это значит, что все его исходы входят в B , $A \subseteq B$), то $P(A) \leq P(B)$. Это свойство легко доказать, если учесть $A = A \cdot B + \bar{A} \cdot B$ и применить предыдущее свойство.

Заметим, что полученные свойства повторяют замеченные на опыте свойства относительной частоты.

2.3. Комбинаторные вычисления

Вычисление вероятностей в классической схеме связано с комбинаторными подсчетами, т. е. вычислением количества элементов множества, определенного некоторыми условиями.

Основная идея этих подсчетов — использование *эквивалентного множества* (рис. 2.2), поскольку множества, связанные этим отношением, имеют одинаковое число элементов.

Этим термином в математике называют два множества A и B , связанные взаимно-однозначным соответствием $A \sim B$. Каждому $a \in A$ соответствует единственный элемент $b \in B$, и обратно.

Принцип равновеликости эквивалентных множеств позволяет получить одну полезную формулу.

Пусть имеется n -элементное множество. Сколько подмножеств можно составить?

Для ответа на этот вопрос перенумеруем элементы и запишем n -разрядное двоичное число по следующему правилу.

Поставим в i -м разряде 1, если i -й элемент входит в подмножество
 ($\underbrace{0 \dots 1}_{i\text{-й элемент}} \dots$), и 0 — если не входит ($\underbrace{0 \dots 0}_{i\text{-й элемент}} \dots$).

Пустому подмножеству соответствует число, во всех разрядах которого стоят нули. Второму тривиальному подмножеству (это само множество) соответствует число, во всех разрядах которого стоят единички. Одноэлементным подмножествам будут соответствовать n чисел с единичками в каком-то одном разряде. Слева направо или справа налево нумеровать разряды — несущественно. Например:

110000010101.

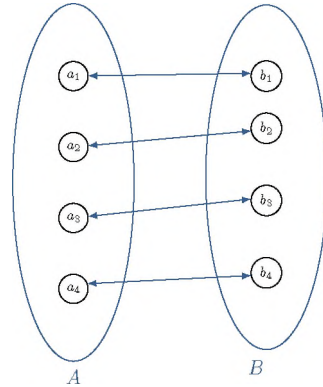


Рис. 2.2. Визуальная метафора эквивалентности множеств (биекция)

Здесь «закодировано» 5-элементное подмножество 12-элементного множества. А вот другое 5-элементное подмножество:

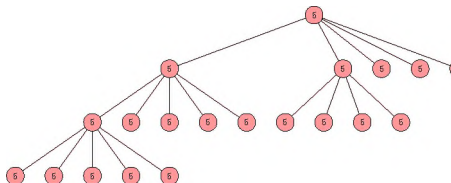
111110000000.

Очевидно, что множество всех подмножеств n -элементного множества¹ эквивалентно множеству всех n -разрядных двоичных чисел.

Итак, различных подмножеств (включая пустое — изображающее невозможное событие) n -элементного множества столько, сколько n -разрядных двоичных чисел. А это 2^n .

Подобное рассуждение можно провести и для другой категории задач. Например, вы подбираете код для кодового замка: на замке 4 рукоятки, на каждой 5 цифр. Полное число комбинаций — количество 4-значных чисел в 5-ричной системе счисления. Это 5^4 . Наглядное представление для подобного подсчета дает *дерево вариантов*. Мы приведем только фрагмент, но для решения задачи этого достаточно.

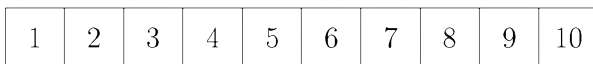
¹Его называют *фактормножеством*.



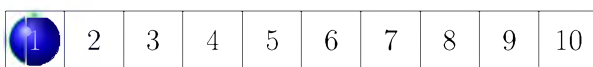
Одна из наиболее известных *комбинаций* для n -элементного множества — **перестановка** его элементов. Предполагается, что элементы множества упорядочены (перенумерованы) и мы изменяем порядок. Более точное определение: подстановка (перестановка) — биективное отображение n -элементного множества на себя. Подсчитаем полное число перестановок n -элементного множества.

Соответствующие подсчеты наиболее наглядны в модели «шариков» и «коробочек». Перенумерованные шарики (n штук) раскладываем по n перенумерованным коробочкам — один шарик в одну коробочку. Сколько таких раскладок получится?

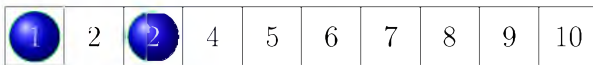
Будем рассуждать следующим образом. Пусть сначала все коробочки пустые.



Первый шарик можно положить в любую из n коробочек, т. е. n способами.



Второй шарик можно положить только $n - 1$ способом — одна коробочка занята.



Итак, первые два шарика можно положить $n(n - 1)$ способами (*правило умножения*). Причем каждый из них побывает в каждой коробочке.



Третий шарик можно положить только $n - 2$ способами — две коробочки заняты.



А всего для трех шариков существует $n(n - 1)(n - 2)$ способов укладки. Понятно, что предпоследний можно положить двумя способами, а последний — единственным.

Полное число перестановок n предметов: $n! = n(n - 1)(n - 2) \dots 2 \cdot 1$.

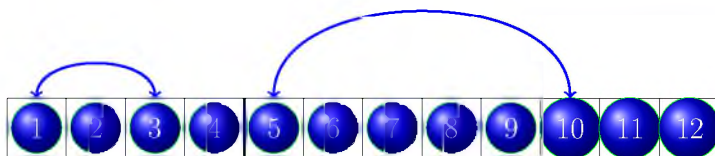
Приведенное рассуждение является *индукцией*, но индукцией интуитивной. В математике используется строгий метод доказательства утверждений — «математическая индукция», который, однако, имеет дело с готовым утверждением или формулой. Предложенный подход позволяет проследить, как получается, возникает общая формула. Многие известные математики называют его математическим наблюдением, в частности, Дж. Пойа (Georg Polya) в своей знаменитой «Mathematics and plausible reasoning».

Вторая важная комбинация — **сочетание**. Так называют m -элементное подмножество n -элементного множества. Напомню, что множество определяется только составом элементов!

Для подсчета числа сочетаний опять воспользуемся шариками и коробочками.

Жирной чертой отделим левые m (в примере 4) коробочек. В них лежит выбранное подмножество шариков. Исходный состав: $A_1 = \{1, 2, \dots, m\}$.

Произведем все возможные перестановки, их $n!$. А теперь внимание! Если не переносить шарики через границу, то состав сочетания (шарики слева от границы) не изменяется.



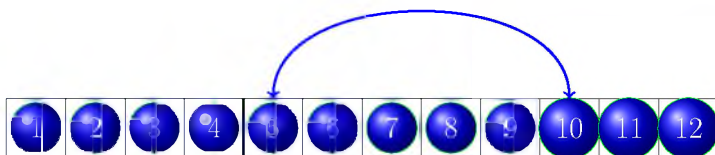
Сколько же перестановок отвечает сочетанию A_1 ? Каждая из $m!$ перестановок слева от черты вместе с каждой из $(n - m)!$ перестановок справа от черты дает $m!(n - m)!$ перестановок. Переставим шарики 1 и 5 — это даст нам новое сочетание A_2 . Состав шариков слева от черты изменился. Опять проведем все перестановки без переноса через черту, их снова $m!(n - m)!$.

Итак, каждому сочетанию соответствует ровно $m!(n - m)!$ перестановок. Значит, полное число сочетаний будет таким:

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

Это число еще называют *биномиальным коэффициентом*, поскольку именно такие коэффициенты при произведениях $a^{n-m}b^m$, $m \in \{0, 1, \dots, n\}$, получаются при разложении *бинома Ньютона* $(a + b)^n$.

Размещением называют упорядоченное подмножество. Если вернуться к предыдущему примеру, то перестановки слева от черты также дают новые размещения, а справа — не меняют ни состав, ни порядок элементов слева.



Поэтому полное число размещений равно:

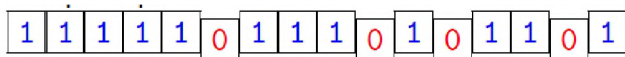
$$A_n^m = \frac{n!}{(n - m)!}.$$

Пример 2. Пусть имеется 5 сортов некоторого продукта, причем каждый сорт в достаточном количестве. Сколькими способами можно составить комплекты из 12 экземпляров этого продукта?

Решение

Модель: каждому комплекту сопоставим двоичное число — код. Единица изображает один экземпляр, разные сорта разделяются ноликом.

Например,



соответствует набору из 5 экземпляров первого сорта, 3 — второго, 1 — третьего, 2 — четвертого и 1 — пятого.

А вот в этом комплекте все экземпляры первого сорта.



Ясно, что искомое количество комплектов в точности равно количеству 16-значных двоичных чисел с четырьмя нулями.

Выполним подсчет. Полное число перестановок 16 цифр — это $16!$. Однако перестановка между собой нулей не меняет значения числа, как и перестановка между собой единиц.

Нули можно переставить $4!$ способами, а единички — $12!$. Значит, одно и то же число дают $4! \cdot 12!$ перестановок. Делим полное число перестановок на количество перестановок на один комплект и получаем искомое количество комплектов:

$$\frac{16!}{4!12!}.$$

В общем случае n экземпляров из m сортов можно составить $\frac{(m+n-1)!}{(m-1)!n!}$ способами.

Это число можно представить в виде

$$\frac{(m+n-1)!}{(m-1)!n!} = C_{m+n-1}^n. \quad (2.1)$$

В комбинаторике данную формулу называют числом сочетаний с повторением. Мы видим, что после некоторых размышлений можно обойтись без знания этой формулы. ■

Хочу заметить, что все использованные графические иллюстрации — это не просто красивые картинки. Graphic Aids, как их называют, помогают понять и пояснить ход логических рассуждений.

В докомпьютерные времена полиграфические возможности, имевшиеся в распоряжении авторов математических текстов, были достаточно бедными. Криволинейные интегралы, трехэтажные формулы представляли большую проблему для наборщиков.

Тем не менее метод stars and bars — так его назвал Вильям Феллер в своей знаменитой книге по теории вероятностей [6] (у других авторов «sticks and stones», «balls and bars») — описан в wiki [26] (можно посмотреть по QR-ссылке) и широко используется в научной и учебной литературе.

Рекомендую им пользоваться при решении учебных задач.

Кстати, как было отмечено в предисловии, такие *полнокодированные* тексты широко распространены и стали предметом изучения лингвистов (см., например, [13]).



В примере 2 комбинацию stars and bars можно изобразить так:

*****|***|*|**|* — для первой картинке и

*****||| — для второй. Такой простейшей визуализации вполне достаточно для понимания условия и проведения решения, хотя она и не так наглядно, как те, что получаются при использовании современных цифровых технологий.

Рассмотрим несколько простых примеров.

Сколькими способами можно разместить два шарика в три ячейки? Введем графический код: первый разряд содержит номер ячейки первого шарика, второй — второго: $\underbrace{1}_a \underbrace{3}_b$. Ясно, что кодирование взаимно-однозначно, т. е. количество размещений равно количеству таких кодов, а их — 9 ($3 \cdot 3$). Если ячеек 4, то размещений $4^2 = 16$.

Задачу о пирожных обычно относят к модели раскладки неразличимых шаров по различным ячейкам, в нашем случае — 12 шаров по пяти ячейкам. Шары (пирожные) мы не различаем, ведь по условию они имеют одну цену, а вот ячейки (сорта) различаем.

К этой же модели можно свести задачу о разложении натурального числа n на k слагаемых с учетом порядка.

Например, если $n = 10$ и $k = 4$, то нам надо найти количество решений:

$$x_1 + x_2 + x_3 + x_4 = 10, x_i \in \mathbb{N}.$$

Каждое решение есть упорядоченная четверка натуральных чисел (x_1, x_2, x_3, x_4) — точка соответствующего пространства. Если допустить $x_i = 0$,¹ то задача будет полностью эквивалентна задаче о пирожных при $n = 10$ и $m = 4$ и по формуле (2.1) получим $C_{13}^3 = 286$. Если же ограничиться натуральными значениями, т. е. $x_i \neq 0$, то нам приходится учитывать, что в комплекте должны быть пирожные каждого сорта.

Как же решать задачу? Возьмем по одному пирожному каждого сорта. Осталось выбрать только еще шесть четырех сортов. Диаграмма Graphic Aids будет содержать $10 - 4 = 6$ stars и 3 разделительных bars (рисуем самостоятельно). Получаем

$$\frac{(6 + 3)!}{3! \cdot 6!} = 168. \quad (2.2)$$

¹Напомню, что в отечественной традиции число 0 не входит во множество \mathbb{N} , т. е. мы расширяем множество допустимых значений.

Пример 3. На полке стоят 12 книг. Сколькими способами можно взять 5 не стоящих рядом?

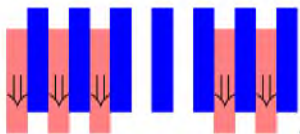
Решение

Воспользуемся простенькой картинкой: * | * | * | * | * | * | * | * . Символ | (bar) представляет оставшиеся после выбора книги — их будет 7. Символ * (star) — это возможные места, откуда можно взять книги. Подумаем, как можно взять книги согласно условию задачи. Стрелочка показывает, откуда книги взяты. Можно взять книжки так: |↑|↑|↑|↑|↑|↑|↑. Если их перенумеровать слева направо, то это визуализация варианта: взяты книжки с номерами {2, 4, 6, 8, 12}. Другой вариант: ||↑|↑|↑|↑|↑|. Здесь набор номеров {3, 5, 7, 9, 11}.

Очевидно, что каждая картинка соответствует допустимому выбору: стрелки перемежаются оставленными книгами.

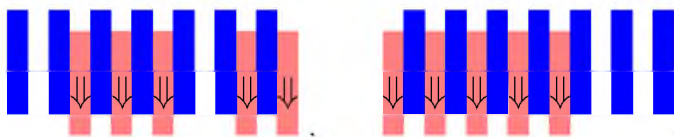
Есть восемь мест (звездочек), откуда можно взять 5 книг. Это сочетания, число которых $C_8^5 = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{5 \cdot 4 \cdot 3 \cdot 2} = 56$. Именно столько способов выполнить условие задачи. ■

Современная цифровая полиграфия дает новые возможности визуализации. Согласитесь, что вот такая иллюстрация гораздо более наглядная:



Взяты книги с номерами 1, 3, 5, 9 и 11.

А вот следующие две комбинации понятны всем:



Задача 2. За знаменитым столом сидят король Артур и 11 рыцарей. Каждый враждует со своими соседями слева и справа. Сколькими способами можно отобрать возможную команду из пяти рыцарей так, чтобы среди них не было враждующих? Король на то и король, что может войти в команду, а может и нет.

Решение

Заметим, что эту задачу можно свести к предыдущей.

Действительно, пусть Артур идет в поход. Два его соседа в этом случае участвовать не могут. Остается 9 рыцарей, для наглядности можно считать, что круг разорван и эти 9 рыцарей стоят в ряд (как книги на полке). Нам надо выбрать четырех, не стоящих рядом. Можете сами проделать все рассуждения из задачи о книгах и получить:

$$C_{9-4+1}^4 = C_6^4 = \frac{6!}{4! \cdot 2!} = 15.$$

В варианте, когда Артур в поход не идет, выбираем пятерых из одиннадцати. Круг опять разомкнулся, и мы имеем задачу с 11 книгами, из которых выбираем 5 несмежных. Повторяем предыдущие рассуждения и получаем

$$C_{11-5+1}^5 = C_7^5 = \frac{7!}{5! \cdot 2!} = 21.$$

Осталось сложить эти результаты, окончательный ответ: 36 вариантов.

$$\text{Заметим, что } C_{12-5-1}^{5-1} + C_{12-5}^5 = \frac{12}{12-5} \frac{(12-5)! \cdot 6 \cdot 6 \cdot 5!}{5!} = 36. \quad \blacksquare$$

Итак, в порядке краткого знакомства с комбинаторикой мы получили несколько полезных формул (количество подмножеств, перестановок, сочетаний и размещений), фактически решая простые задачи общего вида. Другие задачи (про книги, про рыцарей) мы решали для частных случаев, т. е. при заданных значениях параметров.

Для математики характерна самая общая постановка задач. Вот несколько примеров такого подхода.

Задача 3. Докажем два простых тождества:

$$C_n^k = C_{n-1}^k + C_{n-1}^{k-1} \quad \text{и} \quad C_{n-k-1}^{k-1} + C_{n-k}^k = \frac{n}{n-k} C_{n-k}^k.$$

Решение

Проведя тождественные преобразования, убедимся в справедливости записанных равенств для допустимых значений параметров.

$$C_{n-1}^k + C_{n-1}^{k-1} = \frac{(n-1)!}{k! \cdot (n-k-1)!} + \frac{(n-1)!}{(k-1)! \cdot (n-k)!}.$$

Приведем дроби к общему знаменателю, домножив числитель и знаменатель каждой дроби соответственно на $(n-k)$ и на k .

Получим

$$\frac{(n-k) \cdot (n-1)! + k \cdot (n-1)!}{k! \cdot (n-k)!} = \frac{n!}{k! \cdot (n-k)!}.$$

■

Предлагаем доказать второе тождество самостоятельно. Подсказка: постарайтесь вынести в левой части общий множитель C_{n-k}^k . В частном случае равенство было получено при решении задачи 2 о рыцарях на с. 39.

Рассмотрим два подхода к доказательству справедливости комбинаторных формул (тождеств).

Задача 4. Проверить справедливость формулы для суммы сочетаний:

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n. \quad (2.3)$$

Решение

Вариант 1. Используем бином Ньютона:

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

Запишем с его помощью:

$$2^n = (1+1)^n = C_n^0 \cdot 1^n + C_n^1 \cdot 1^{n-1} \cdot 1^1 + C_n^2 \cdot 1^{n-2} \cdot 1^2 + \dots + C_n^{n-2} \cdot 1^2 \cdot 1^{n-2} + C_n^{n-1} \cdot 1^{n-1} \cdot 1 + C_n^n \cdot 1.$$

Отсюда с очевидностью следует, что

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^{n-2} + C_n^{n-1} + C_n^n = 2^n. \quad (2.4)$$

Вариант 2. Рассуждаем так: число сочетаний C_n^k — по определению количество k -элементных подмножеств n -элементного множества. Далее подсчитываем полное число всевозможных подмножеств — это как раз и будет левая часть рассматриваемого тождества.

Само значение мы уже вычисляли ранее (с. 33) — 2^n . ■

Итак, два независимых способа доказательства продемонстрированы. Можно было применить и метод математической индукции.

Более подробно с комбинаторными моделями можно познакомиться в работе [12] и в других рекомендованных выше учебниках.

2.4. Достоинства и недостатки классической модели

Рассмотренная ранее классическая модель известна и использовалась с очень давних времен. Математикам VIII — XIX вв. формула представлялась очевидной. Заметим, что свойства классической вероятности совпадают со свойствами статистической. Это не случайно, и несколько позже мы рассмотрим данную модель с более современной точки зрения.

Классический подход представляется предпочтительным, если у нас есть основания считать, что исследуемое явление обладает необходимыми свойствами симметрии (равновозможность!). Надо заметить, что наличие данного свойства определяется условием конкретной прикладной задачи, фактически это элемент идеализации при переходе к математической модели. К сожалению, многие студенты не понимают важность этого этапа решения. Приведем высказывание основателя теории вероятностей А. Н. Колмогорова: «На вопрос, какие случаи можно считать равновозможными, математика не даст ответа. При бросании костей условия выпадения любой из шести граней представляются нам одинаковыми. Кроме того, представляется естественным считать, что различные комбинации верхних граней двух костей тоже одинаково правдоподобны. Разделение всех возможных исходов на исключаяющие друг друга равновозможные случаи достаточно *деликатно*» [27, с. 12].

Этому требованию удовлетворяет большое количество интересных и практически важных задач. В случае, когда классический результат примерно совпадает с наблюдаемой частотой, можно заключить, что мы правильно понимаем суть изучаемого явления. В противном случае наше предположение о равновозможности не соответствует реально наблюдаемому, например, необходимых свойств симметрии нет (гнутая монетка, кубик со смещенным центром масс, крапленые карты и пр.).

В романе Дж. Лондона о золотоискателях «Смок Беллью» (Smoke Bellew) приводится история о том, как главный герой — бывший студент — заметил несимметричность рулеточного стола в салуне. После некоторых наблюдений ему удалось построить выигрышную стратегию для этого рулеточного стола. Несомненно, автор был неплохо знаком со статистикой!

Более адекватное построение модели может быть в этом случае выполнено и теоретически, и статистически. Первый подход более характерен для теоретической физики, и примером является сравнительное рассмотрение свойств классических частиц, бозонов и фермионов. Второй подход применяют в технике, экономике и социологии, где изучаемые системы слишком сложны и законы их функционирования неизвестны.

С практической точки зрения недостатком классического подхода является и то, что он применим лишь к конечным ПЭС.

2.5. Геометрическая вероятность

Известно обобщение классического подхода на бесконечные ПЭС. Это так называемый *геометрический* подход, который на интуитивном уровне использовался еще на заре развития теории. Мы приведем слегка осовремененный вариант.

Пусть M — множество из следующего списка:

- а) отрезок прямой $[a, b]$;
- б) G — ограниченная часть плоскости (геометрическая фигура);
- в) W — ограниченная часть пространства (геометрическое тело).

Назовем **мерой** этого множества числовую функцию μ , значение которой для подмножества множества M равно соответственно:

- а) длине ($\mu([a, b])$);
- б) площади ($\mu(G)$);
- в) объему ($\mu(W)$).

Если ПЭС и интересующему нас событию A как множествам можно сопоставить такую меру, то такие события принято называть измеримыми. Вспомним задачу 1 о встрече двоих (с. 24). Достоверному событию, ПЭС можно сопоставить площадь квадрата, а событию {двое встретились} — площадь шестиугольника.

За вероятность события A принимают

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}. \quad (2.5)$$

$P(A)$ называют **геометрической вероятностью**. Таким образом, геометрическая вероятность представляет собой **долю** площади фигуры A в фигуре Ω .

Рассмотрим задачу, известную в истории математики как «задача Бюффона об игле». Мы уже упоминали опыты Бюффона ранее (см. с. 29) и давали там же QR-ссылку на компьютерную анимацию эксперимента с иглой.

Задача 5. На плоскость, расчерченную параллельными прямыми, отстоящими друг от друга на расстояние a , наудачу брошена игла длиной l ($l < a$). Найти вероятность того, что игла пересечет линию.

Решение

В качестве исхода эксперимента будем фиксировать расстояние x от центра иглы до ближайшей линии и угол между линией и иглой, θ .

Построим вспомогательный чертеж (рис. 2.3) и получим

$$0 \leq \theta \leq \pi; \quad 0 \leq x \leq \frac{a}{2}.$$

В качестве ПЭС можно взять прямоугольник со сторонами π и $\frac{a}{2}$. Событие $A = \{\text{игла пересекает линию}\}$ наступает, как нетрудно сообразить, при условии $x \leq \frac{l}{2} \sin \theta$ (рис. 2.3).

Вероятность $\mathbf{P}(A)$ считаем как отношение площадей криволинейной трапеции, ограниченной сверху линией $x = l \sin \theta$, и прямоугольника.

$$\begin{cases} S_A = \frac{l}{2} \int_0^{\pi} \sin \theta \, d\theta = l \\ S_{\Omega} = \pi \frac{a}{2} \end{cases} \Rightarrow \mathbf{P}(A) = \frac{S_A}{S_{\Omega}} = \frac{2l}{\pi a}. \quad (2.6)$$

■

В своих работах, посвященных геометрическим вероятностям, Бюффон ставил целью показать, что «геометрия может быть использована в качестве аналитического инструмента в области теории вероятностей». Дело в том, что для вычисления вероятности по классической формуле используется только арифметика. Действительно, рассмотренные нами ранее задачи фактически

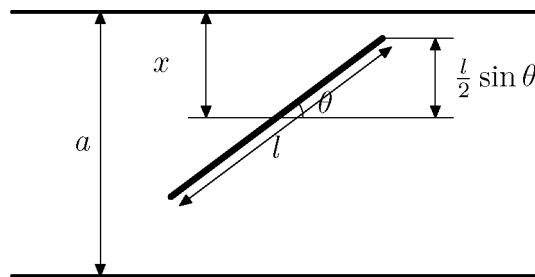


Рис. 2.3. Характеристики исхода бросания иглы

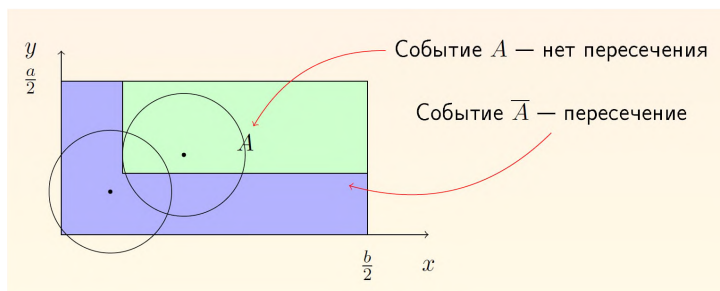
сводились к комбинаторике, т. е., по сути, к арифметическим действиям с действительными числами.

Эксперименты с иглой со времен этих публикаций проводились много раз, определялась статистическая вероятность и по формуле (2.6) подсчитывалось значение числа π , обеспечивающее равенство теоретической и статистической вероятностей. У последователей Бюффона получались очень приличные результаты в диапазоне $3.15 \div 3.17$. В компьютерных экспериментах автора при $n = 1000$ бросаний доля «успехов» была такова, что вычисляемое по формуле (2.6) значение для π оказалось 3.15. Кстати, в упомянутой выше анимации можно увидеть на экране изменение этого значения при увеличении числа бросков. Надо отметить, что полученные результаты ни в коей мере не являются «измерением» π по причинам, к сожалению, отнюдь не очевидным. Во-первых, теория вероятностей вовсе не требует равенства частоты и вероятности. Во-вторых, измерения углов, расстояний, да и самого факта пересечения приближенные.

Задача 6. Игра франк-карро (также исследованная Бюффеном в его работе «Мемуар об игре франк-карро») состоит в следующем: пол разграфлен на одинаковые фигуры. На пол бросается монета, ее диаметр ($2r$) меньше каждой из сторон, и монета целиком укладывается внутрь фигуры. Чему равна вероятность того, что брошенная наудачу монета пересечет одну или две стороны фигуры?

Решение

Для определенности рассмотрим покрытие плоскости прямоугольниками со сторонами a и b , $b > a > 2r$. Исход эксперимента определяется парой чисел (x, y) — расстояниями от центра монетки до ближайших вертикальной и горизонтальной нарисованных фигур соответственно. ПЭС для данного ЭСИ имеет очень простой вид:



Замечание. На рисунке абстрактный прямоугольник (ПЭС), а не кусочек пола. Условие пересечения сторон совершенно очевидно: понятно, что если

точка (x, y) (центр монеты) попадет внутрь малого прямоугольника (событие A), то она в точности окажется внутри одной из нарисованных фигур.

Значит, вероятность того, что монета пересечет по меньшей мере одну из сторон основного прямоугольника, вычисляемая как отношение соответствующих площадей, равна

$$2r \frac{a + b - 2r}{ab}$$

Автором было проведено компьютерное моделирование этой задачи (рис. 2.4). Полученное хорошее согласие экспериментального значения вероятности с теоретическим позволяет предложить метод оценки размера квадратной сетки. В этом случае параметр один — « a ». Его значение можно оценить, зная размер монетки и устойчивое значение относительной частоты. Именно такой метод (в идейном смысле) был использован в исторических опытах Резерфорда по рассеянию α -частиц для оценки размеров ядра атома.

После Бюффона задачи на геометрические вероятности стали систематически включаться в трактаты и учебники по теории вероятностей. Так, в знаменитой книге Лапласа «Аналитическая теория вероятностей» были подробно рассмотрены все задачи Бюффона.

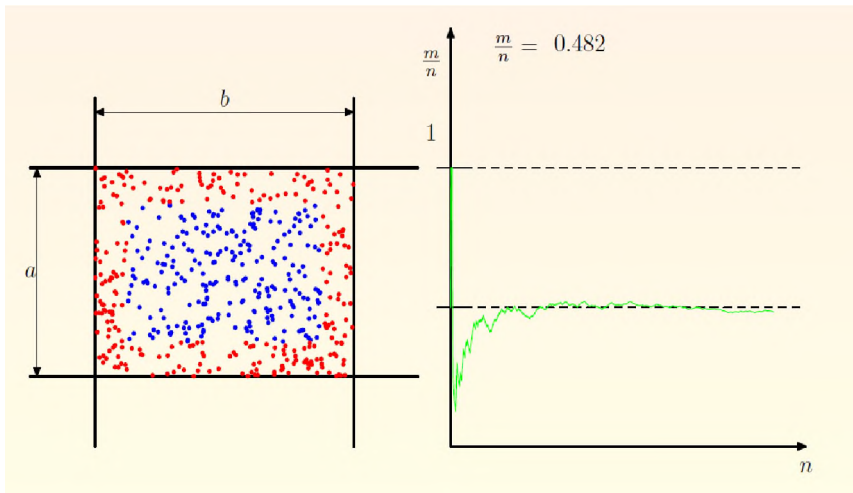


Рис. 2.4. Кадр компьютерной анимации автора, для наглядности отображались только центры монет. График демонстрирует устойчивость частоты ($\frac{m}{n} = 0.482$) и хорошее согласие с теоретическим результатом 0.5

Вторая задаче об игле, сформулированная и решенная Бюффеном, отличалась от первой тем, что плоскость была разграфлена на квадраты. В более

общей постановке Лапласа две системы параллельных прямых представляли не что иное, как систему координатных линий на плоскости. Расстояние между линиями первой системы равно a , второй — b . Эту задачу Лаплас решил и нашел, что вероятность пересечения одной из линий брошенной иглой

равна $4r \frac{a+b-r}{\pi ab}$. Кстати, тем самым он обнаружил ошибку в решении

Бюффона.

Давайте рассмотрим подробно решение этой задачи.

Задача 7. Плоскость разграфлена двумя системами параллельных прямых, представляющих ортогональную систему координатных линий на плоскости. Расстояние между линиями первой системы равно a , второй — b . На плоскость бросается игла длиной $2r$ ($2r < a < b$). Чему равна вероятность того, что игла пересечет хотя бы одну линию?

Решение

Введем переменные, описывающие исход соответствующего ЭСИ: (x, y, θ) (показаны на рис. 2.5). События: $A = \{\text{пересекает вертикальную линию}\}$, $B = \{\text{пересекает горизонтальную линию}\}$. События A и B совместны, для их

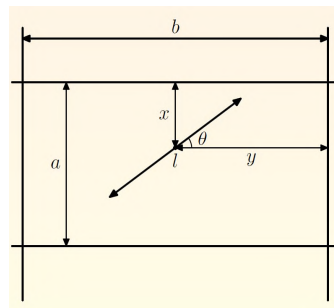


Рис. 2.5. Схема ко второй задаче Бюффона

суммы (хотя бы одно пересечение) применяем формулу сложения:

$$P(A + B) = P(A) + P(B) - P(A \cdot B). \quad (2.7)$$

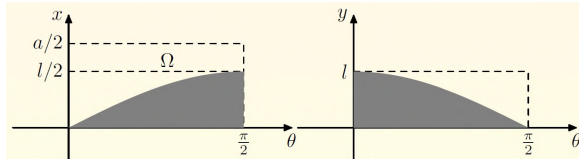
Вероятности событий A и B фактически известны из обычной задачи Бюффона:

$$P(A) = \frac{4r}{\pi a}, \quad P(B) = \frac{4r}{\pi b}. \quad (2.8)$$

Условия пересечений линий разного направления в данном случае выглядят так: $x \leq r \cdot \sin(\theta)$ и $y \leq r \cdot \cos(\theta)$ (см. рис. 2.5). Поэтому $P(B)$ вычисляется иначе, чем в задаче с одной системой линий, но результаты совпадают:

$$\int_0^{\pi/2} r \cdot \cos(\alpha) d\alpha = r \Rightarrow P(B) = 4 \frac{r}{\pi b}. \quad (2.9)$$

В геометрической модели, а именно она используется для решения, вероятность равна отношению объемов. Вероятность произведения в (2.7) связана с одновременным выполнением обоих условий, т. е. $x \leq r \cdot \sin(\theta)$, $y \leq r \cdot \cos(\theta)$. ПЭС представляет собой прямоугольный параллелепипед. События A и B представляют собой цилиндры с образующими вдоль осей OY и OX соответственно, а проекции на соответствующие координатные плоскости (Θ, X) и (Θ, Y) — криволинейные трапеции:



Тело события $A \cdot B$ — пересечение этих цилиндров (рис. 2.6). Построив его, находим пределы для вычисления тройного интеграла, дающего объем этого события:

$$V_{AB} = \iiint_{AB} dV = \int_0^{\pi/2} \int_0^{r \sin(\theta)} \int_0^{r \cos(\theta)} 1 \, dy \, dx \, d\alpha = \frac{1}{2} r^2.$$

$$P(AB) := \frac{V_{AB}}{V_{\Omega}} = 8 \cdot \frac{\frac{1}{2} r^2}{\pi a b} = 4 \cdot \frac{r^2}{\pi a b}.$$

Подставляя полученные вероятности в (2.7), получаем окончательно

$$P(A + B) = \frac{4r}{\pi a} + \frac{4r}{\pi b} - \frac{4r^2}{\pi a b} = 4r \frac{a + b - r}{\pi a b}.$$

■

Заметим, что полученный результат совпадает с результатом Лапласа (приведен на с. 47).

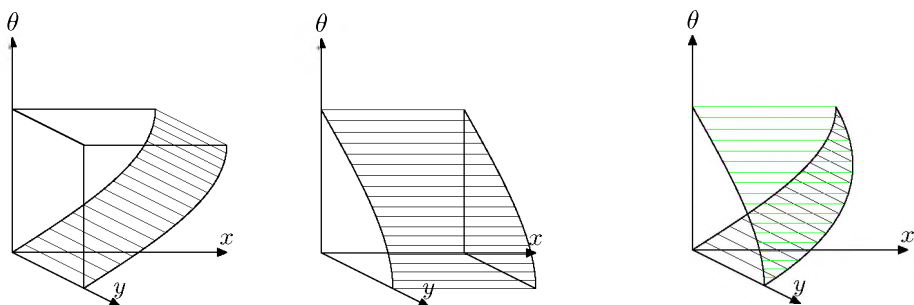


Рис. 2.6. Иллюстрация построения произведения событий $A \cdot B$ в решаемой задаче: слева направо событие A , событие B , их пересечение — событие $A \cdot B$

Мы сделали своеобразный исторический экскурс в процесс становления теории вероятностей. Во многом подходы основателей науки были интуитивными, но это была гениальная интуиция. С современной точки зрения геометрический вероятность означает введение *равномерно распределенных* случайных величин (будет подробное рассмотрение). Таковыми являются все характеристики экспериментов Бюффона. Однако случайные величины позволяют описывать данные эксперименты более адекватно их природе. Игрок бросает монетку прицельно, а отклонения от цели тем менее вероятны, чем эти отклонения больше. Поэтому распределение вероятностей оказывается существенно неравномерным.

Глава 3

Аксиоматический подход к вероятности. Основные формулы теории вероятностей

3.1. Идеология аксиоматического подхода

Современное определение вероятности было предложено в 30-е гг. советским математиком А. Н. Колмогоровым и принято практически всеми математиками. За основу были взяты основные свойства статистической вероятности. Из них выделен минимум свойств, из которых логически следуют остальные. Ведь с точки зрения математики вероятность — это числовая функция. Областью ее определения является множество всех мыслимых событий данного эксперимента, а множеством значений — отрезок $[0, 1]$. Возможность определить вероятности для конкретного эксперимента, т. е. какие именно значения из этого отрезка принимает вероятность некоего события, не является предметом ТВ.

Понятно, что вероятность должна иметь определенное значение для любого события, наблюдение которого в эксперименте возможно. Поскольку множество событий должно быть замкнуто относительно операций сложения и умножения, то основной математической конструкцией будет *алгебра событий*.

Определение. Совокупность \mathcal{A} подмножеств Ω называется алгеброй при выполнении следующих условий:

- $\Omega \in \mathcal{A}, \emptyset \in \mathcal{A}$;
- Если $A \in \mathcal{A}$, то $\bar{A} \in \mathcal{A}$;
- Если $A, B \in \mathcal{A}$, то $A \cup B \in \mathcal{A}$ и $A \cap B \in \mathcal{A}$.

Заметим, что данное определение несколько упрощено, интересующимся можно порекомендовать работу [29]. Если последнее условие расширить на бесконечное (счетное) множество операндов, то используют термин σ -алгебра.

Приведем несколько примеров:

1. Пусть Ω — некоторое множество, тогда $\mathcal{A}_* = \{\Omega, \emptyset\}$ является тривиальной алгеброй.
2. Если $A \subset \mathcal{A}$, $\mathcal{A} = \{\Omega, \emptyset, A, \bar{A}\}$.
3. Совокупность всех подмножеств (фактормножество) конечного $\Omega(\mathcal{A}^*)$.
4. Возьмем $\Omega = R$ и рассмотрим множества $(-\infty, a)$, $[b, c]$, $[d, -\infty)$. Конечные объединения этих множеств являются алгеброй.
5. $\mathcal{A} = \{[0, 1], \emptyset, [0, \frac{1}{3}], [\frac{1}{3}, 1]\}$ алгеброй **не является**.

Для конечного ПЭС никаких проблем нет — любое подмножество (а их будет конечное число) входит в алгебру. Таким образом, алгеброй в этом случае будет множество всех подмножеств. Для бесконечных же ПЭС возникают определенные логические проблемы, которые в нашем учебном пособии обсуждаться не будут. Некоторый намек будет дан только при чуть более строгом рассмотрении геометрической модели.

Определение. Пусть проводится некоторый эксперимент, которому соответствует ПЭС Ω . Рассмотрим возникающую алгебру событий \mathcal{A} . Определим на множестве событий \mathcal{A} числовую функцию \mathbf{P} , обладающая следующими свойствами:

1. $\mathbf{P}(A) \geq 0$;
2. $\mathbf{P}(\Omega) = 1$;
3. Если события A и B несовместны, то $\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B)$.

Такая функция \mathbf{P} называется **вероятностью**. Для события A значение $\mathbf{P}(A)$ называется **вероятностью события A** .

Последняя аксиома позволяет вычислять вероятность для суммы любого конечного числа попарно несовместных событий.

Из данного определения можно вывести еще много свойств вероятности, важнейшие из которых будут получены позднее. $A + \bar{A} = \Omega$, $A \cdot \bar{A} = V$.

Поэтому, применяя вторую и третью аксиомы, получаем $P(A) + P(\bar{A}) = 1$.
Еще одно: $P(A) = 1 - P(\bar{A}) \leq 1$.

В строгой математической ТВ основой является *вероятностное пространство*, или тройка $\langle \Omega, \mathcal{A}, P \rangle$. Дело в том, что для данного ПЭС алгебру, как показано в предыдущем примере, можно определить по-разному. Ну а функцию P тем более, лишь бы она удовлетворяла аксиомам. В прикладных задачах выбор алгебры и вероятности определяется содержанием, свойствами изучаемой системы. Во многих случаях функцию P можно определить классической или геометрической формулами. В других многочисленных задачах применяются более сложные методики. Часть из них будет рассмотрена в следующей главе, посвященной *основным формулам* теории вероятностей.

Задача 8. Правильная монета подбрасывается до тех пор, пока герб не выпадет 2 раза подряд. Построить ПЭС и алгебру. Найти вероятность того, что число подбрасываний не превысит 5.

Решение

Рассмотрим множество конечных цепочек символов Γ и Π , длина которых не меньше 2, в которых пара $\Gamma\Gamma$ стоит только в конце цепочки, а также множество бесконечных цепочек из Γ и Π , не содержащих сочетания $\Gamma\Gamma$. Все это — исходы нашего эксперимента. В качестве σ -алгебры (множество событий бесконечно!) возьмем множество всех подмножеств.

Вероятность на σ -алгебре определим так: каждой цепочке длины n сопоставим значение вероятности $1/2^n$, а каждой бесконечной цепочке — 0.

Событие {число бросков не превысит 5} состоит из исходов $\{\Gamma\Gamma, \Pi\Gamma\Gamma, \Gamma\Pi\Gamma\Gamma, \Gamma\Pi\Pi\Gamma\Gamma, \Pi\Pi\Gamma\Gamma, \Pi\Pi\Pi\Gamma\Gamma\}$.

$$P(\{\text{число бросков не превысит 5}\}) = 1/4 + 1/8 + 2/16 + 3/32 = 19/32.$$

■

3.2. Простейшие вероятностные модели

Упомянутые ранее классическая и геометрическая формулы естественным образом получаются в рамках двух вероятностных пространств. Мы также конкретизируем и наполним строго определенным смыслом термин «равновозможность».

Вероятностное пространство — $\langle \Omega, \mathcal{A}, \mathcal{P} \rangle$, ПЭС Ω — дискретное, поэтому в алгебру входят все подмножества. Гипотезу «равновозможности» определим так: все одноэлементные события имеют одну и ту же вероятность

$P(\{\omega_i\}) = c$. Используя аксиомы вероятности, получим

$$P(\Omega) = P\left(\bigcup_{i=1}^n \{\omega_i\}\right) = \sum_{i=1}^n P(\{\omega_i\}) = |\Omega| \cdot c = 1. \quad (3.1)$$

Отсюда следует уже знакомая классическая формула

$$P(A) = P\left(\bigcup_{\omega_i \in A} \{\omega_i\}\right) = \sum_{\omega_i \in A} P(\{\omega_i\}) = \frac{|A|}{|\Omega|}. \quad (3.2)$$

Вычисления по этой формуле выполняем с использованием комбинаторных методов.

В случае геометрической модели ситуация сложнее. Для определенности будем рассматривать ПЭС Ω в виде плоской геометрической фигуры Φ , имеющей некоторую площадь S_Ω . В этом случае не любое подмножество Ω будет событием, поскольку подмножество-событие должно иметь площадь. Этот пример показывает, почему в математической теории вероятностей в качестве базового понятия вводят алгебру событий.

Площадь плоской геометрической фигуры Φ — подходящая мера для определения вероятности. Действительно, $S \geq 0$, обладает свойством монотонности, т. е. площадь фигуры не меньше площади ее части, $\Phi_1 \subseteq \Phi$. Наконец, разбив фигуру на непересекающиеся части $\Phi = \Phi_1 \cup \Phi_2$, получим $S = S_1 + S_2$. Осталось получить условие нормировки $P(\Omega) = 1$.

Итак, событием будет любая часть фигуры Ω , имеющая сколь угодно малую площадь.

Разобьем ПЭС на произвольное число равновеликих событий. Гипотеза равновозможности в данной модели подразумевает, что все равновеликие события имеют одинаковую вероятность.

Таким образом, вероятность события пропорциональна его площади.

$$P(A) = c \cdot S_A, P(\Omega) = c \cdot S_\Omega = 1, \Rightarrow c = \frac{1}{S_\Omega}. \quad (3.3)$$

$$P(A) = \frac{S_A}{S_\Omega}. \quad (3.4)$$

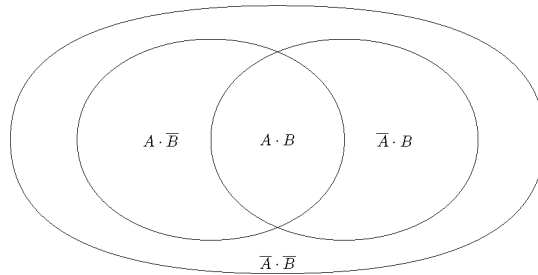
Аналогичные соображения для одномерного и трехмерного случаев ограничивают подмножества такими, которые имеют длину или объем соответственно. Общая формула может быть записана так:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}, \quad (3.5)$$

где μ — мера множества (события), т. е. длина, площадь или объем соответственно.

3.3. Формула сложения

Аксиоматически заданы, естественно, не все свойства вероятности. Остальные свойства непосредственно следуют из определения. Ранее мы вычисляли вероятность либо по классической формуле, либо по формуле геометрической вероятности, сейчас мы получим несколько важных формул для вычисления вероятностей. Они позволяют решать более сложные задачи, находить вероятности одних событий через известные вероятности других. Рассмотрим также несколько довольно широко применимых моделей или, как их традиционно называют, схем: схему гипотез, схему Бернулли, полиномиальную схему. Первая полезная формула составляет содержательную часть теоремы сложения.



Теорема 1. Для любых событий A и B имеет место равенство

$$P(A + B) = P(A) + P(B) - P(A \cdot B).$$

Доказательство

Идея доказательства довольно прозрачная. Представим событие $A + B$ в виде суммы несовместных событий:

$$A + B = A + B\bar{A}, \quad \text{проверяем: } A \cdot (B \cdot \bar{A}) = \emptyset, \text{ верно.}$$

По третьей аксиоме

$$P(A + B) = P(A) + P(B\bar{A}). \quad (3.6)$$

Также и для события B :

$$B = A \cdot B + \bar{A} \cdot B \Rightarrow P(B) = P(AB) + P(\bar{A}B). \quad (3.7)$$

Исключив из (3.6) и (3.7) $P(\bar{A}B)$, получаем искомое.

$$P(A + B) = P(A) + P(B) - P(AB). \quad (3.8)$$

□

Последовательно применяя формулу (3.8), получаем для трех событий

$$P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC). \quad (3.9)$$

В качестве примера рассмотрим следующий эксперимент.

Задача 9. Цифры 1, 2, 3 пишутся в произвольном порядке. Какова вероятность того, что по крайней мере одна из цифр окажется на надлежащем (в смысле нумерации) месте?

Решение

ПЭС Ω эксперимент состоит из множества перестановок трех символов. $|\Omega| = 3! = 6$. Пусть A есть множество перестановок, в которых цифра 1 стоит на первом месте, B — множество перестановок, в которых цифра 2 стоит на втором месте, C — множество перестановок, в которых цифра 3 стоит на третьем месте. Нам нужно вычислить $P(A + B + C)$. Ясно, что

$$P(A) = P(B) = P(C) = \frac{2}{6} = \frac{1}{3}.$$

Действительно, если какая-нибудь цифра стоит на надлежащем месте, то имеются две возможности переставить остальные две цифры из общего числа $3! = 6$ возможных перестановок трех цифр. Далее

$$P(AB) = P(AC) = P(BC) = \frac{1}{6}, \quad P(ABC) = \frac{1}{6},$$

так как в каждом из этих случаев возникает только одна возможность — все три цифры стоят на своих местах. И тогда формула (3.9) дает нам

$$P(A + B + C) = 3 \cdot \frac{1}{3} - 3 \cdot \frac{1}{6} + \frac{1}{6} = 1 - \frac{1}{2} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}. \quad \blacksquare$$

Вывести самостоятельно формулу для $P(A + B + C + D)$ и применить ее к эксперименту, аналогичному предыдущему, в котором будут участвовать 4 цифры. Ответ: соответствующая вероятность равна $\frac{5}{8} = 0,6250$.

Общая формула для суммы n событий имеет вид:

$$P(A_1 + A_2 + \dots + A_n) = \sum_i P(A_i) - \sum_{i \neq j} P(A_i \cdot A_j) + \sum_{i \neq j \neq k} P(A_i \cdot A_j \cdot A_k) - \dots \pm P(A_1 \dots A_n), \quad (3.10)$$

где символы \sum обозначают суммирование по всем возможным произведениям, содержащим два, три, ..., $(n - 1)$ событий из слагаемых, стоящих в левой части, т. е. A_1, A_2, \dots, A_n .

Вид формулы может быть установлен из «математических наблюдений» (Феллер [6]) за тем, как из формулы (3.8) была выведена формула (3.9), а затем и формула для четырех слагаемых в примере на с. 55. Доказательство для n слагаемых поводится методом математической индукции.

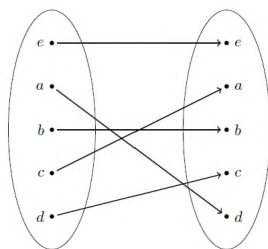
С помощью (3.10) решается большое количество задач из курса ТВ. В частности, можно заключить, что если n цифр $1, 2, 3, \dots, n$ написаны в произвольном порядке, то вероятность того, что по крайней мере одна из цифр окажется на надлежащем месте, равна:

$$P_n = 1 - \frac{1}{2!} + \frac{1}{3!} - \dots \pm \frac{1}{n!}, \quad (3.11)$$

причем перед последним членом ставится знак «+» при нечетном n или «-» при четном. В частности, при $n = 5$ эта вероятность равна:

$$P_5 = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} = \frac{19}{30}. \quad (3.12)$$

Очень важно, решая задачу «нематематического вида», отбросить несущественные детали и выделить главное, построить математическую модель. И тогда вы с удивлением поймете, что хрестоматийные задачи «о рассеянной секретарше», «о новогодних подарках» (она же Christmas gifts task) и многие другие сводятся к формуле (3.10).



Все эти задачи можно рассматривать в модели подстановка взаимно-однозначное отображение n элементного множества в себя (биекция). ПЭС будет состоять из всех возможных подстановок, а их $n!$. Обратите внимание на такой тип отображения, при котором $k < n$ элементов отображаются в себя — остаются *неподвижными*. Если же неподвижных элементов нет, то используется термин *беспорядок* из n элементов, D_n . Приведем справочную рекуррентную формулу для числа беспорядков:

$$D_n = n \cdot D_{n-1} + (-1)^n,$$

очевидно, что $D_1 = 0$, $D_2 = 1$. Теперь вы можете рассмотреть и случай, когда секретарша из пяти написанных писем два все же отправила по адресу.

3.4. Условная вероятность. Теорема умножения

Определим новую числовую функцию событий.

Определение. Вероятностью события A при условии наступления события B (условная вероятность) называется функция, определенная соотношением

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cdot B)}{\mathbf{P}(B)}, \mathbf{P}(B) > 0. \quad (3.13)$$

Она позволяет вычислить вероятность наступления события A в ситуации, когда известно, что событие B произошло. С формальной точки зрения мы определили новую функцию, причем это функция двух переменных — событий. Зафиксируем событие B и покажем, что определенная формулой функция от A удовлетворяет аксиомам вероятности.

1. $\mathbf{P}(A|B) \geq 0$. Это следствие определения и первой аксиомы вероятности.

2. $\mathbf{P}(\Omega|B) = \frac{\mathbf{P}(\Omega \cdot B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B)}{\mathbf{P}(B)} = 1$, так как $\Omega \cdot B = B$.

3. $\mathbf{P}(\bar{A}|B) = 1 - \mathbf{P}(A|B)$. Действительно, построим цепочку импликаций: $A + \bar{A} = \Omega \Rightarrow AB + \bar{A}B = B \Rightarrow P(AB) + P(\bar{A}B) = P(B)$. Делим обе части равенства на $P(B)$ и получаем искомое свойство.

4. Для несовместных A_1 и A_2 имеем: $\mathbf{P}(A_1 + A_2|B) = \mathbf{P}(A_1|B) + \mathbf{P}(A_2|B)$. Опять строим цепочку импликаций: $(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B \xrightarrow{A_1 \cdot A_2 = \emptyset} \xrightarrow{A_1 \cdot A_2 = \emptyset} P((A_1 + A_2) \cdot B) = P(A_1 \cdot B) + P(A_2 \cdot B) \xrightarrow{def} \frac{P((A_1 + A_2) \cdot B)}{P(B)} = \frac{P(A_1 \cdot B)}{P(B)} + \frac{P(A_2 \cdot B)}{P(B)}$. Последнее равенство и есть искомое.

Итак, все это означает, что функция, определенная формулой (3.13) и называемая условной вероятностью, действительно является вероятностью при фиксированном B ($P(B) > 0$).

Формула умножения вероятностей непосредственно следует из определения условной вероятности:

$$P(A \cdot B) = P(A|B)P(B). \quad (3.14)$$

Последовательно применяя (3.14), получим: $P(A \cdot B \cdot C) = P(A|B\dot{C}) \cdot P(B|C) \cdot P(C)$ и т. п.

Определение. События A и B называются **статистически независимыми** (для краткости — **независимыми**), если $\mathbf{P}(A) = \mathbf{P}(A|B)$.

Тогда $\mathbf{P}(B) = \mathbf{P}(B|A)$, поскольку

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(AB)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A|B)\mathbf{P}(B)}{\mathbf{P}(A)} = \mathbf{P}(B). \quad (3.15)$$

Из определения парной независимости и (3.15) следует: если события A и B независимы, то $\mathbf{P}(A \cdot B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$. Доказательство данного утверждения предоставляется читателям в качестве простого упражнения. Сама формула будет использована в дальнейшем.

Иначе говоря, парное отношение независимости обладает симметричностью, события A и B взаимно независимы. Можно показать, что в этом случае независимы и пары (\bar{A}, B) ; (\bar{A}, \bar{B}) ; (A, \bar{B}) .



В качестве примера применения основных формул по данному QR [30] можете ознакомиться с «funny story» — задачей, занимавшей в 90-е гг. XX в. многих любителей интернета.

Определение. События $\{A_1, A_2, \dots, A_n\}$ называются **независимыми в совокупности**, если $\mathbf{P}(A_1 \cdot A_2 \dots A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2) \dots \mathbf{P}(A_n)$.

Для событий $\{A_1, A_2, \dots, A_n\}$ надо различать попарную независимость и независимость в совокупности, что иллюстрирует следующая задача («пример Бернштейна»).

Задача 10. Три грани тетраэдра окрашены соответственно в красный (К), зеленый (З) и синий (С) цвета, а четвертая — частично в каждый из этих трех цветов (КЗС). Изучить вопрос о независимости событий К, З и С. Здесь буква обозначает исход, состоящий в выпадении вниз грани, на которой присутствует соответствующий цвет.

Решение

В данном эксперименте выбрасывается однородное тело правильной геометрической формы. Поэтому все четыре исхода можно считать равновероятными, значит, вероятность сложного события {в основании будет красный цвет} равна $1/2$. Условные вероятности типа $\mathbf{P}(К|З)$ также равны по $1/2$, так как красный цвет присутствует на двух гранях из четырех. Аналогично для остальных цветов. Значит, попарная независимость имеет место, а вот произведение событий К, З и С будет исходом с вероятностью $1/4$.

Отсюда следует, что условие независимости в совокупности —
 $P(K) \cdot P(Z) \cdot P(C) = P(K \cdot Z \cdot C)$ — не выполняется: $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \neq \frac{1}{4}$. ■

Задача 11. Хорошая задача. В семьях Иванова и Петрова по два ребенка. Ответим на два вопроса:

У Ивановых один из детей — точно мальчик (знаем откуда-то). Какова вероятность того, что оба ребенка мальчики?

У Петровых старший ребенок — мальчик. Какова вероятность того, что оба маленьких Петрова мальчики?

Решение

Для разумного ответа на эти вопросы построим модель эксперимента со случайным исходом. Фамилия главы семьи в нашей модели несущественна, абстрактная семья с двумя детьми. ПЭС состоит из четырех элементарных: (М, Д), (М, М), (Д, М) и (Д, Д). Условное обозначение: (младший, старший). Исходя из известной статистики рождений, предполагается, что события рождение мальчика и рождение девочки независимые и равновероятные события.

Построим схему эксперимента со случайным исходом и соответствующее ему ПЭС.

№ исхода	(младший, старший)	Вероятность
1	(М, М)	1/4
2	(М, Д)	1/4
3	(Д, М)	1/4
4	(Д, Д)	1/4

Применяем теорию вероятностей. Условие задачи означает, что нам нужно вычислить значения двух условных вероятностей в данном ПЭС. Начнем с первой, которую для краткости так и обозначим — P_1 . Получим, составив отношение вероятностей произведения событий (М, М) и (М, М)+(М, Д)+(Д, М), а это (М, М), и события (М, М)+(М, Д)+(Д, М) (3.13):

$$P_1 = \frac{P(\text{М, М})}{P((\text{М, М})+(\text{М, Д})+(\text{Д, М}))} = \frac{1/4}{3/4} = 1/3. \quad (3.16)$$

Вычисление P_2 аналогично, только надо учесть, что ((М, М)+(Д, М))·(М, М)=(М, М). Поэтому $P_2 = \frac{P(\text{М, М})}{P((\text{М, М})+(\text{Д, М}))} = \frac{1/4}{1/2} = 1/2$. ■

Рассмотренная выше задача широко известна с 1959 г. из статьи известного популяризатора науки Гарднера и считалась одним из парадоксов теории

вероятностей [31]. Советуем внимательно разобраться с приведенным решением, а также ознакомиться и с другими вариантами условия. Кажущаяся парадоксальность объясняется лишь нечеткой формулировкой условия некоторыми авторами, допускающей двусмысленность формальной интерпретации. Надо также отметить, что семейный антураж, по сути, несущественен, можно дважды бросать монетку и фиксировать выпадение герба или цифры — полная аналогия.

3.5. Обсуждение термина *независимые события*

Независимость событий — это еще один термин ТВ, который используется как в узком (см. выше), так и в широком смысле. В первой главе этого учебного пособия обсуждалось понятие *случайность* и некоторые причины ее появления в окружающем мире. Естественные науки изучают неслучайные, детерминированные связи между физическими характеристиками объектов и явлений, выражают их в форме функциональных зависимостей. Именно в виде функциональных зависимостей сформулированы основные физические законы: Ньютона, Ома, газовые законы и др. Случайные явления просто по определению так описываться не могут. Ведь при идентичных начальных условиях гарантированно предсказать результат невозможно. Тем не менее обнаружены и изучаются другие виды зависимостей, не являющиеся функциональными. Их называют статистическими, в курсе ТВ и МС именно знакомство с такого вида зависимостями является основной целью обучения.

Статистические зависимости активно применяются в очень многих областях человеческой деятельности, в частности, в медицине. Влияние различных жизненных факторов и особенностей конкретных людей на их предрасположенность к некоторому заболеванию и восприимчивость к применяемым лекарствам изучают статистическими методами. Итак, зависимости между характеристиками объектов и явлений могут быть различной силы. Крайними в этом спектре будут функциональная зависимость и статистическая независимость.

Пусть в серии из n опытов наблюдаются события A , B и $A \cdot B$.

$$N_A = \frac{m_A}{n}, N_B = \frac{m_B}{n}, N_{A \cdot B} = \frac{m_{A \cdot B}}{n}.$$

Сравним две величины: $\frac{m_{A \cdot B}}{m_B}$ и $\frac{m_A}{n}$. Первая есть доля исходов, связанных с одновременным появлением событий A и B среди исходов с B , вторая — доля исходов с наблюдением A среди всех опытов. Скажем, A — {смерть больного

от рака легких}, $B = \{\text{большой курил}\}$. Многочисленные наблюдения показывают, что $\frac{m_{A \cdot B}}{m_B} \gg \frac{m_A}{n}$. Это дает основание медикам утверждать, что курение способствует данному заболеванию (наш Минздрав предупреждает, а американский — бьет долларом!).

Если для каких-то других событий A и B обнаруживается, что

$$\frac{m_{A \cdot B}}{m_B} = \frac{m_A}{n},$$

то делается вывод о их *статистической независимости*.

Заметим, что

$$\frac{m_{A \cdot B}}{m_B} = \frac{m_{A \cdot B}}{n} \frac{n}{m_B} = \frac{N_{A \cdot B}}{N_B}.$$

Это отношение было бы естественно принять за статистическую условную вероятность, т. е. вероятность наступления события A при условии, что B произошло. Таким образом, теоретическая «независимость» абстрактных случайных событий есть отражение статистической независимости соответствующих реальных событий — прообразов.

3.6. Композиция испытаний

Сейчас мы рассмотрим еще одну конструкцию, используемую при моделировании вероятностных явлений. Пусть проводятся сразу два испытания. Это может быть стрельба по мишени двух стрелков. Пусть вероятность попадания для первого — p_1 , а для второго — p_2 .

Обозначим исходы первого испытания $\{A_1, \bar{A}_1\}$, а второго испытания $\{A_2, \bar{A}_2\}$. Можно рассмотреть эти два испытания как некоторое *составное* испытание, исходами которого являются всевозможные комбинации исходов первого и второго: $\{(A_1, A_2), (A_1, \bar{A}_2), (\bar{A}_1, A_2), (\bar{A}_1, \bar{A}_2)\}$.

Найдем вероятности исходов. Очевидно, что

$$p((A_1; (A_2 + \bar{A}_2))) = p_1, p((A_1 + \bar{A}_1), A_2) = p_2.$$

В более общем случае имеем два ПЭС: $\Omega_1 = \{\omega_1^1, \dots, \omega_n^1\}$ и $\{\Omega_2 = \{\omega_1^2, \dots, \omega_n^2\}\}$. Будем считать, что вероятности $P_1(\{\omega_j^1\})$ и $P_2(\{\omega_j^2\})$ известны.

Построим новое ПЭС, $\Omega = \{\omega_{ij} = (\omega_i^1, \omega_j^2) | \omega_i^1 \in \Omega_1, \omega_j^2 \in \Omega_2\}$. Его называют *прямым произведением* Ω_1 и Ω_2 или *композицией* и обозначают $\Omega_1 \times \Omega_2$.

Подмножества Ω назовем событиями, причем они будут в некотором смысле «сложными» по отношению к событиям первых множеств. Дело в том, что

ω_{ij} можно интерпретировать как одновременное наблюдение двух элементарных событий: одного в первом испытании, а второго — во втором.

На новой алгебре событий вероятности определим следующим образом: $P(\{\omega_{ij}\}) = P_1(\{\omega_i^1\}) \cdot P_2(\{\omega_j^2\})$, а для события A

$$P(A) = \sum_{\omega_{ij} \in A} P(\{\omega_{ij}\}). \quad (3.17)$$

Рассмотрим теперь два события A и B из новой алгебры. A состоит из таких пар (ω_i^1, ω_j^2) , что $\omega_i^1 \in A_1, \omega_j^2 \in \Omega_2$, т. е. исход второго эксперимента — любой. Используя геометрическую аналогию, такие события называют *цилиндрическими*. Вероятность этого события

$$\begin{aligned} P(A) &= \sum_{\omega_i^1 \in A_1, \omega_j^2 \in \Omega_2} P_1(\{\omega_i^1\}) P_2(\{\omega_j^2\}) = \\ &= \sum_{\omega_i^1 \in A_1} P_1(\{\omega_i^1\}) \sum_{\omega_j^2 \in \Omega_2} P_2(\{\omega_j^2\}) = P_1(A_1). \end{aligned} \quad (3.18)$$

Событие же B состоит из таких пар (ω_i^1, ω_j^2) , что $\omega_i^1 \in \Omega_1, \omega_j^2 \in A_2$. Для вероятности аналогично получается

$$P(B) = \sum_{\omega_i^1 \in \Omega_1, \omega_j^2 \in A_2} P_1(\{\omega_i^1\}) P_2(\{\omega_j^2\}) = P_2(A_2). \quad (3.19)$$

Произведение событий A и B состоит из таких пар (ω_i^1, ω_j^2) , что $\omega_i^1 \in A_1, \omega_j^2 \in A_2$. Значит, вероятность, $A \cap B$ будет равна:

$$\begin{aligned} P(A \cap B) &= \sum_{\omega_i^1 \in A_1, \omega_j^2 \in A_2} P_1(\{\omega_i^1\}) P_2(\{\omega_j^2\}) = \\ &= \sum_{\omega_i^1 \in A_1} P_1(\{\omega_i^1\}) \sum_{\omega_j^2 \in A_2} P_2(\{\omega_j^2\}) = P_1(A_1) \cdot P_2(A_2) = P(A) \cdot P(B). \end{aligned} \quad (3.20)$$

Таким образом, абстрактный *критерий независимости* событий в данном случае приобретает очень ясный смысл — он отражает отсутствие какой-либо связи между исходами двух различных экспериментов. Конечно, формально полученный результат есть следствие задания вероятности на новой алгебре, $P(\{\omega_{ij}\}) = P_1(\{\omega_i^1\}) \cdot P_2(\{\omega_j^2\})$, при другом определении вероятности события A и B могут и не быть независимыми.

Итак, «совместное», «одновременное» наблюдение событий может моделироваться двояко. Во-первых, как произведение событий из одного ПЭС: $A \cdot B | A \subset \Omega; B \subset \Omega$. Например, эксперимент — выбрасывание кубика, событие A — {выпало четное число очков}, событие B — {выпало число очков, кратное трем}. Произведение A и B — {выпало 6 очков}. Во-вторых, как элемент композиции соответствующих ПЭС двух испытаний, как в задаче об одновременной стрельбе двоих.

Пример 4. Выбрасывается два игральных кубика. Построить композицию событий $A_1 = \{1, 2, 3\}$ на первом кубике, $A_2 = \{2, 3\}$ на втором кубике.

Решение

ПЭС для двух экспериментов соответственно Ω_1 и Ω_2 , в каждом по шесть исходов, а в $\Omega_1 \times \Omega_2$ — 36. Ведь по определению прямого (декартова) произведения множеств это множество упорядоченных пар элементов этих множеств. В алгебре композиции рассматриваем события $A_1^c = (\{1, 2, 3\}, \Omega_2)$, т. е. в первом эксперименте наблюдаем A_1 на первом кубике, а на втором — что угодно и $A_2^c = (\Omega_1, \{2, 3\})$ (на первом кубике что угодно, на втором — 2 или 3).

Искомую композицию обозначим

$$(A_1, A_2) = \{(1, 2); (1, 3); (2, 2); (2, 3); (3, 2); (3, 3)\}.$$

Можно сказать, что событие (A_1, A_2) моделирует «одновременное» наблюдение двух событий A_1 и A_2 , т. е. оно наблюдается при выпадении любой из перечисленных комбинаций исходов первого и второго ПЭС. Вероятности:

$$P(A_1^c) = P(A_1) = \frac{3}{6} = \frac{1}{2}; \quad P(A_2^c) = \frac{1}{3}; \quad P((A_1, A_2)) = \frac{6}{36} = \frac{1}{2} \cdot \frac{1}{3}. \quad (3.21)$$

События A_1^c и A_2^c независимы! ■

Пример 5. Выбрасывается две монетки. Построить композицию.

Решение

Вариант 1. $\Omega = \Omega_1 \times \Omega_2 = \{(h, h), (h, t), (t, t), (t, h)\}$. Обозначим вероятности элементарных исходов обоих экспериментов положительными a и b , пусть $a + b = 1$. Для суперсобытий (двойных серий)

$$P(hh) = a^2, \quad P(ht) = ab, \quad P(th) = ab, \quad P(tt) = b^2.$$

Такое определение вероятностей на новой алгебре согласуется с аксиомами

$$a^2 + 2ab + b^2 = (a + b)^2 = 1.$$

Определяем события новой алгебры: первой выпала решка H_1 , второй — решка H_2 . Вероятности событий:

$$P(H_1) = P(hh) + P(ht) = a^2 + ab = a, \quad P(H_2) = P(th) + P(hh) = ab + a^2 = a.$$

Проверяем условие независимости.

$$H_1 \cdot H_2 = P(hh) = a^2 = P(H_1) \cdot P(H_2).$$

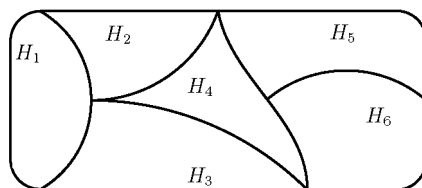
Вариант 2. Этот же эксперимент можно описать иначе. Выпадение решки имеет одинаковую вероятность при каждом броске. События независимы по определению. ■

Глава 4

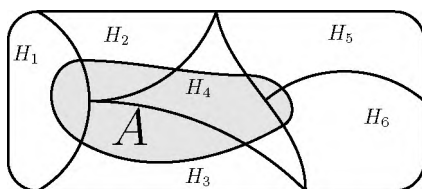
Основные формулы (продолжение)

4.1. Схема гипотез. Формулы полной вероятности и Байеса

Многие задачи теории вероятностей решаются в рамках одной теоретической модели — *схемы гипотез*.



Пусть H_1, H_2, \dots, H_n — множество попарно несовместных событий, $H_1 + \dots + H_n = \Omega$. Событие A может произойти только совместно с некоторыми из них.



Итак, множество событий H_1, H_2, \dots, H_n , определяющих разбиение ПЭС, называют *гипотезами*. Используется еще термин «полная группа событий».

Любое событие A может произойти совместно со всеми или частью событий из полной группы: $A \cdot (H_1 + \dots + H_n) = A$.

Можно предложить следующую трактовку схемы гипотез, поясняющую и происхождение названия, и характерные задачи, в которых она применяется.

Допустим, что предлагается провести опыт, условия которого заранее полностью неизвестны, о них можно сделать лишь n взаимоисключающих предположений (гипотез). Вероятности гипотез известны, известны и условные вероятности интересующего нас события A в рамках каждой гипотезы. Оказывается, что по этим данным можно вычислить вероятность $\mathbf{P}(A)$.

Теорема 2. Пусть H_1, H_2, \dots, H_n — полная группа событий (гипотез). Тогда

$$\mathbf{P}(A) = \mathbf{P}(H_1) \cdot \mathbf{P}(A|H_1) + \dots + \mathbf{P}(H_n) \cdot \mathbf{P}(A|H_n).$$

Доказательство

Будем исходить из определения гипотез:

$$H_1 + H_2 + \dots + H_n = \Omega, \quad H_i \cdot H_j = V, \quad i \neq j.$$

Умножим обе части равенства на событие A , учитывая, что $A \cap \Omega = A$, получаем

$$(A \cap H_1) \cup (A \cap H_2) \cup \dots \cup (A \cap H_n) = A. \quad (\text{Это события!}) \quad (4.1)$$

В схеме гипотез событие A представляется в виде суммы несовместных событий $(A \cap H_i)$. Значит, применяя формулу сложения, получаем

$$P(A) = P(H_1 \cap A) + P(H_2 \cap A) + \dots + P(H_n \cap A). \quad (\text{Это вероятности!}) \quad (4.2)$$

В формулах (4.1) и (4.2) добавлены важные (!) замечания. Первое равенство — это равенство **событий** с соответствующими операциями над **событиями**, второе — **числовое** равенство, операции — арифметические. Применим теорему умножения для каждого аргумента слагаемых в (4.2) и получим искомую формулу:

$$P(A) = \sum_{i=1}^n P(A|H_i)P(H_i). \quad (4.3)$$

Эту формулу называют **формулой полной вероятности**. □

Следует обратить внимание на то, что символы операций сложения событий и вероятностей в (4.1) и (4.2) взяты разные, чтобы подчеркнуть — операнды разной природы.

Пример 6. Имеются три урны одинакового вида. Известно, что в одной 2 белых и 3 черных шара, во второй — 4 белых и 1 черный шар, в третьей — 3

белых шара. Вам предложена одна из этих урн. Какова вероятность достать из нее белый шар?

Решение

Вы не знаете, какая именно урна предложена, поэтому применяем схему гипотез. Обозначим гипотезы: H_i — предложена урна с номером i . Вероятность каждой из гипотез в нашем случае равна $1/3$ (равновозможность). Условные вероятности $P(A|H_i)$ легко вычисляются. Используя формулу полной вероятности (4.3), получим

$$P(A) = \frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{4}{5} + \frac{1}{3} \cdot 1 = \frac{11}{15}.$$

■

Эксперименты, моделируемые схемой гипотез, допускают получение новых сведений о свойствах объекта или явления из результатов проведенных экспериментов. Это позволяет сделать *формула Байеса*.

Теорема 3. В условиях теоремы о полной вероятности (т. е. в схеме гипотез) имеет место равенство

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)} = \frac{P(H_i)P(A|H_i)}{P(H_1)P(A|H_1) + \dots + P(H_n)P(A|H_n)}. \quad (4.4)$$

Доказательство

В самом деле, с одной стороны,

$$P(A \cdot H_i) = P(A)P(H_i|A), \quad (4.5)$$

с другой стороны,

$$P(A \cdot H_i) = P(H_i)P(A|H_i). \quad (4.6)$$

Приравнивая правые части равенств (4.5) и (4.6), получаем с помощью формулы полной вероятности требуемое заключение. □

Полученная формула (4.4) для условных вероятностей гипотез называется формулой Байеса (часто говорят *формулы Байеса*, так как в обеих частях стоит свободная переменная $i \in \{1, 2, \dots, n\}$). Эта формула позволяет «переоценить» вероятности гипотез по результатам эксперимента. Вероятности $P(H_i|A)$ называют **апостериорными**, поскольку они учитывают результаты опыта, в отличие от **априорных** вероятностей гипотез.

Пример 7. Из 18 стрелков 5 попадают в мишень с вероятностью 0,8; 7 — с вероятностью 0,7; 4 — с вероятностью 0,6 и 2 — с вероятностью 0,5. Стрелок,

выбранный по жребию из этих 18 стрелков, произвел выстрел и попал в мишень. Какова вероятность того, что стрелок принадлежит к первой группе?

Решение

Условия задачи укладываются в схему гипотез. Так как результат эксперимента уже известен, то следует применить формулу Байеса. В качестве гипотез возьмем следующие события: H_1 — стрелок принадлежит к первой группе, H_2 — ко второй, H_3 — к третьей, H_4 — к четвертой, A — стрелок попал в мишень. Тогда

$$\begin{aligned} \mathbf{P}(A) &= \mathbf{P}(H_1)\mathbf{P}(A|H_1) + \mathbf{P}(H_2)\mathbf{P}(A|H_2) + \mathbf{P}(H_3)\mathbf{P}(A|H_3) + \\ &\quad + \mathbf{P}(H_4)\mathbf{P}(A|H_4) = \\ &= \frac{5}{18} \cdot 0.8 + \frac{7}{18} \cdot 0.7 + \frac{4}{18} \cdot 0.6 + \frac{2}{18} \cdot 0.5 = 41/60; \\ \mathbf{P}(H_1|A) &= \frac{\frac{5}{18} \cdot 0.8}{\frac{41}{60}} = \frac{40}{123}. \end{aligned}$$

■

Заметим, что априорная вероятность того, что выбранный по жребию стрелок принадлежит к первой группе, равна $\frac{5}{18} = 0.2777\dots$. Если же известно, что этот стрелок попал, то вероятность немного увеличилась. Это и понятно — в первой группе собрались самые меткие стрелки и известно, что выбранный стрелок попал.

4.2. Последовательности испытаний

Ранее (в главе 3) мы ввели в рассмотрение композицию двух ЭСИ и получили несколько интересных результатов. Заметим, что актуальность этого вопроса подтверждает и наличие его в хрестоматийной работе Феллера [6]. Сейчас определим и обсудим еще несколько теоретических моделей, применяющихся при изучении случайных явлений. По традиции их называют схемами. Рассмотрим совокупность нескольких ЭСИ. Составим композицию связанных с ними ПЭС $\Omega_1, \dots, \Omega_n$. Исходами нового испытания, «суперисходами», являются упорядоченные наборы исходов $(\omega_i^1, \omega_j^2, \dots, \omega_k^n)$, $\omega_i^l \in \Omega_l$. В литературе по ТВ «суперисходы» принято называть *последовательностями испытаний*, причем никакого «временного» упорядочения не подразумевается. Не имеет смысла говорить, что одно испытание «предшествует» другому или

«произошло раньше», чем другое. Данная модель применяется и при одновременном осуществлении всей серии, важна лишь формальная упорядоченность, т. е. у каждого испытания есть номер или другой идентификационный признак.

Серьезно разработана модель, в которой задаются а) вероятности исходов первого испытания; б) для остальных считается, что вероятность i -го исхода n -го испытания зависит только от исхода предыдущего испытания. Эти вероятности составляют *матрицу вероятностей* перехода. А сама модель называется *цепь Маркова*.

Отсутствие какой-либо связи между испытаниями, входящими в серию (их независимость), моделируется знакомым нам способом:

$$\mathbf{P}((\omega_i^1, \omega_j^2, \dots, \omega_k^n)) = \mathbf{P}_1(\omega_i^1) \cdot \mathbf{P}_2(\omega_j^2) \dots \mathbf{P}_n(\omega_k^n).$$

Простейшая из таких моделей — схема Бернулли.

4.3. Схема Бернулли

Определение. Схемой Бернулли называют последовательность n испытаний, в каждом из которых имеется два исхода (A, \bar{A}), вероятность наступления события A в каждом из них равна p , а вероятность противоположного $\bar{A} - q = 1 - p$.

Следуя Феллеру [6], мы опускаем в определении прилагательное «независимых», используемое во многих отечественных учебниках, поскольку соответствующее свойство явно присутствует в точной формулировке. В современной терминологии мы имеем дело со специфическим ПЭС, состоящим из всевозможных упорядоченных n -элементных цепочек вида $(A, \bar{A}, \bar{A}, A, \dots, A)$. В каждой позиции может стоять A или \bar{A} , т. е. всего 2^n исходов.

Интерес представляет «число успехов» события A , т. е. сколько раз событие произошло.

Теорема 4. Пусть вероятность наступления события A в каждом из n испытаний, проходящих по схеме Бернулли, равна p . Тогда вероятность того, что в серии из n испытаний событие A наступит ровно m раз, равна:

$$p_n(m) = C_n^m p^m q^{n-m}, \quad (4.7)$$

где $q = 1 - p$.

Доказательство

Заметим, что «суперсобытие» {событие A произошло k раз} есть сумма всех таких серий:

$$\underbrace{(A, A \dots A \quad \bar{A}, \bar{A} \dots \bar{A})}_{k \text{ — компонент}},$$

в которых A произошло в k каких-то испытаниях из n . Вероятность каждой такой серии в силу постулированной на алгебре независимости одна и та же: $p^m q^{n-m}$. Самих серий столько, сколько можно составить сочетаний k «успешных» испытаний из всех n испытаний, т. е. C_n^m .

Итак, получаем $P_n(k) = C_n^k p^k q^{n-k}$. \square

Заметим, что правая часть есть k -й член разложения бинома $(p + q)^n$, так что полученное соотношение называют еще биномиальным законом распределения.

Обобщением схемы Бернулли (биномиальной схемы) является схема проведения эксперимента, которая отличается от схемы Бернулли тем, что при каждом испытании происходит не одно из событий A или \bar{A} , а одно из m несовместных событий. Иначе говоря, ПЭС содержит более двух исходов. Получается *полиномиальная схема*.

Утверждение 1. Пусть в каждом из n независимых испытаний может произойти m исходов с вероятностями p_1, p_2, \dots, p_m . Тогда вероятность того, что каждое из событий A_i произойдет k_i раз, $i = 1, \dots, m$, будет выражаться соответствующим членом полиномиального разложения $(p_1 + p_2 + \dots + p_m)^n$.

$$P_n(k_1, k_2, \dots, k_m) = \frac{n!}{k_1! k_2! \dots k_m!} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}.$$

Это утверждение принимается без доказательства.

4.4. Формулы Лапласа и Пуассона

Вычисления непосредственно по формулам Бернулли проводить достаточно сложно, даже имея в распоряжении компьютер. (Что уж говорить о докомпьютерной эре!) Однако вполне приемлемую точность можно получить, используя приближенные формулы для соответствующих значений параметров. Для больших значений n и вероятностей p , далеких от 0 и 1, известны

формулы Лапласа. При малых значениях p (или q) хорошо работает *формула Пуассона.* Познакомимся с этими формулами и попробуем применить их к решению задач.

Пример 8. Вероятность брака на данном станке составляет 0,005. Найти вероятность того, что из 10 000 изделий окажутся бракованными а) ровно 40, б) не более 70.

Решение

Решение легко записывается по формуле Бернулли:

$$\text{а) } P_{10000}(40) = C_{10000}^{40} 0.995^{9960} 0.005^{40};$$

$$\text{б) } p(k \leq 70) = \sum_{m=0}^{70} C_{10000}^m 0.995^{10000-m} 0.005^m.$$

Вот только вычислять непосредственно затруднительно, скажем,

$$C_{10000}^m = \frac{\overbrace{10000 \cdot 9999 \cdot \dots \cdot (10000 - m + 1)}^{(m \text{ множителей})}}{m!}. \quad (4.8)$$

Здесь в числителе 40 множителей и столько же в знаменателе: ни записать явно (без условной символики), ни в строку калькулятора ввести. А что было делать математикам в XVIII в.? Воспользуемся формулами Лапласа. Не будем проводить строгий вывод, ограничимся качественным описанием результата (plausible reasoning — правдоподобные рассуждения по Поа [32]).

Точки с координатами $(k, P_n(k))$ при больших значениях n прекрасно ложатся на кривую Гаусса с параметрами $a = np$ — матожидание и $\sigma = \sqrt{npq}$ — СКО (рис. 4.1). Отсюда следует, что вычисление по локальной формуле сводится к вычислению ординаты точки графика с $x = k$. Интегральная формула — фактически интегральная сумма интеграла от самой функции $\frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-\frac{x^2}{2}}$, $x = \frac{k - np}{\sqrt{npq}}$ при $\Delta x = 1$. Для $n \sim 100$ погрешность будет достаточно малой.

Локальная формула Лапласа (аргумент x приведен к стандартному табличному виду):

$$\sqrt{npq} \cdot p_n(k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x = \frac{k - np}{\sqrt{npq}}.$$

Прежде чем открывать таблицы, проведем некоторые предварительные вычисления вручную:

$$\sqrt{npq} = \sqrt{10000 \cdot 0.005 \cdot 0.995} = \sqrt{49.75} \simeq 7,$$

$$p_n(k) \simeq \frac{1}{7\sqrt{2\pi}} e^{-\frac{1.42^2}{2}} = \frac{1.1456}{7} = 0.0026.$$

Значение функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ найдено по таблице (она табулирована как в бумажных справочниках, так и в электронных). Есть таблицы и другой функции – **функции Лапласа**:

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

Интегральная формула Лапласа:

$$P(k_1 \leq k \leq k_2) = \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right).$$

Применяя формулу Лапласа, по таблицам находим ответ на второй вопрос задачи:

$$p(k \leq 70) = \Phi\left(\frac{20}{7}\right) - \Phi\left(\frac{-50}{7}\right) = \Phi(2.84) + 0.5 = 0.9975.$$

■

Формула Пуассона применяется в схеме Бернулли.

Пусть $p \ll 1$ или $q \ll 1$, в этом случае говорят, что явления **редкие**. Формально получить результат для вероятности в этом случае можно, перейдя к пределу по $n \rightarrow \infty$ при условии $n \cdot p = a = \text{const}$.

$$p \rightarrow 0, \quad n \cdot p = a = \text{const}.$$

Исключим p , положив $p = \frac{a}{n}$. Преобразуем формулу Бернулли.

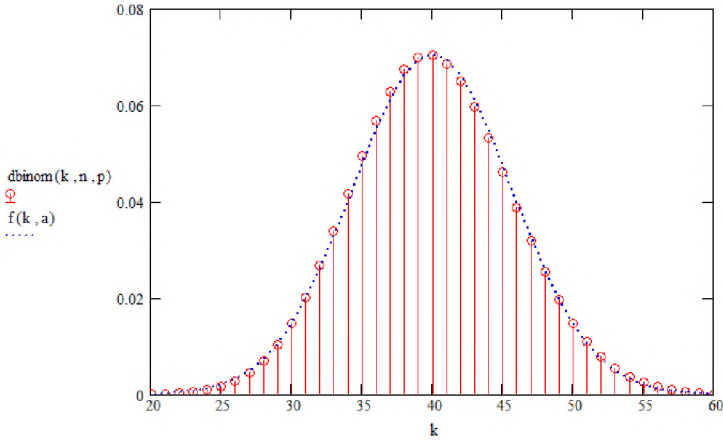


Рис. 4.1. Результат сравнения значений $P_n(k)$ и ординат точек графика функции φ , соответствующих абсциссам $x = k$ (получено автором в программе MathCad)

$$\begin{aligned}
 p_n(k) &= \frac{n!}{k!(n-k)!} \left(\frac{a}{n}\right)^k \left(1 - \frac{a}{n}\right)^{n-k} = \\
 &= \frac{n(n-1)\dots(n-k+1)}{k!} \left(\frac{a}{n}\right)^k \left(1 - \frac{a}{n}\right)^{n-k} = \\
 &= \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{a^k}{k!} \left(1 - \frac{a}{n}\right)^{n-k} \rightarrow \frac{a^k}{k!} e^{-a}. \quad (4.9)
 \end{aligned}$$

$$P(k) = \frac{a^k}{k!} e^{-a}.$$

Погрешность и условия применимости формулы можно получить из разложения

$$\ln \left(1 - \frac{a}{n}\right)^n = n \left(-\frac{a}{n} - \frac{1}{2} \frac{a^2}{n^2} - \dots\right).$$

При $p < 0.1$ формула Пуассона считается прекрасным приближением.

Пример 9. В начале Великой Отечественной войны при подавляющем превосходстве немецкой авиации применялась стрельба по низколетящему самолету всем батальоном или полком. Попасты из винтовки в летящий на

бредущем полете самолет очень маловероятно, $p = 0.001$. Подсчитаем вероятность попадания в цель не менее чем двумя пулями при числе выстрелов $n = 5000$.

$$a = np = 0.001 \cdot 5000 = 5,$$

$$p(k \geq 2) = 1 - p(0) - p(1) = 1 - e^{-5} - 5e^{-5} = 0.9596.$$

$$p(4) = p(5) = 0.1751.$$

Если стрелявших меньше, то и вероятность будет ниже.

Рассмотрим хрестоматийный так называемый «кавалерийский пример», под эти наименованием он известен в истории теории вероятностей. Непосредственно использовалась замечательная книга Габора Секей «Парадоксы в теории вероятностей и математической статистике» [33].

Пример 10. «За 20 лет между 1875 и 1894 годами в 14 различных кавалерийских корпусах германской армии была собрана статистика... когда солдат был убит ударом копыта... 196 солдат погибли» [33, с. 34]. Полезно на данном примере подробно описать применение формулы Пуассона, так как в источнике процедура описана довольно лаконично, есть только ссылка на оригинальную работу Ладислава Борткевича [34]¹.

Фактически была применена схема Бернулли, длина последовательности вычислялась как произведение числа корпусов на количество лет, т. е. получается $14 \times 20 = 280$ «испытаний» (наблюдений). Именно в виде таблицы размерности 14 на 20 статистические данные были представлены у Борткевича [Там же, с. 24]. Среднее число трагических исходов будет $\frac{196}{280} = 0.70 = np$. Это значение (достаточно малая величина) было взято в качестве параметра формулы Пуассона $\lambda = 0.7$, который равен математическому ожиданию (среднему значению числа k).

Итак, во-первых, имеются статистические данные о распределении количества погибших по 280 наблюдениям. Во-вторых, мы можем вычислить «теоретическое» значение вероятности числа «успехов» (k) в схеме Бернулли и тем самым оценить, во скольких наблюдениях будет количество смертей, равное k . Иначе говоря, вычислим значение $n \cdot P_n(k) = 280 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda}$. Представим результаты в виде таблицы, статистические данные приведены только

¹Л. Борткевич (1868–1931) — выдающийся российский математик, внесший значительный вклад в прикладную статистику. В его цитированной работе рассматриваемая задача проанализирована значительно полнее и сделан вывод о том, что «Die Kongruenz der Theorie mit erfahrung läßt sowohl nichts wünschen übrig». Русский вариант приведен в тексте.

те, которые имеются в источнике (внимательный читатель наверняка заметит этот факт: $32 \times 2 + 91 = 155$, а не 196).

k (число смертей)	0	1	2
Фактически наблюдаемое	140	91	32
Оценка по Пуассону	139	97	34

У Г. Секей приводится такая оценка сравнения этих данных: «... практика и теория оказались в столь хорошем согласии, что вряд ли можно было ожидать большего». Справедливости ради заметим, что сам Л. Борткевич в своей книге гораздо более подробно анализировал статистические данные и привел уточненную таблицу для 200 наблюдений. Именно в этой таблице и оказалось настолько хорошее согласие теории с практикой, что он сам сделал данную высокую оценку. Его оригинальная фраза на немецком языке приведена в подстрочном примечании на с. 74.

Примеры величин, подчиняющихся распределению Пуассона: количество определенных товаров, проданных в течение дня; количество опечаток на странице текста; число телефонных соединений с определенным номером в течение дня; число молекул гемоглобина, видимое на предметном стекле микроскопа. Даже распределение количества забитых голов в играх чемпионата по футболу (редкие явления!) хорошо укладывается в рамки данного распределения.

Глава 5

Случайные величины

5.1. Случайная величина: определение понятия

Рассмотрим знакомый нам эксперимент с выбрасыванием двух игральных кубиков (см. с. 23). Соответствующее ПЭС, содержащее 36 исходов, изобразим наглядно (рис. 5.1). Теперь каждый исход отметим точкой и соединим стрелочкой с числом — суммарным количеством выпавших очков (рис. 5.2). Получается визуализация отображения (функции, которую называют случайной величиной) с областью определения Ω (множество исходов ЭСИ — ПЭС) и множеством значений X . В данном примере $X = \{2, 3, \dots, 12\}$.

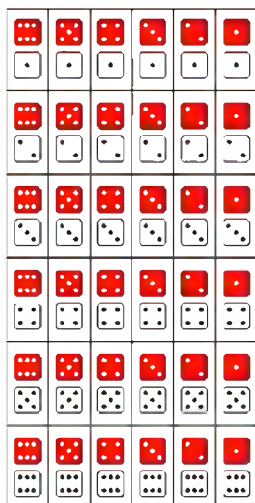


Рис. 5.1. Схема ПЭС эксперимента с двумя кубиками

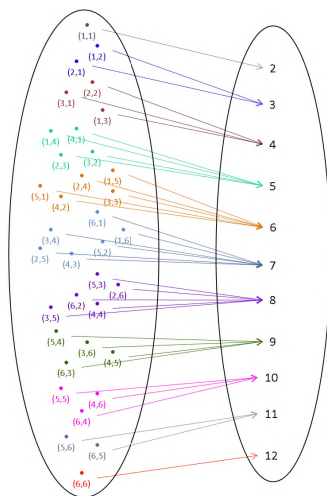
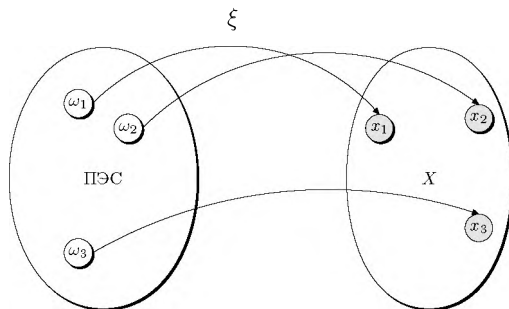


Рис. 5.2. Отображение ПЭС на множество значений суммы очков, выпавших на двух игральных кубиках

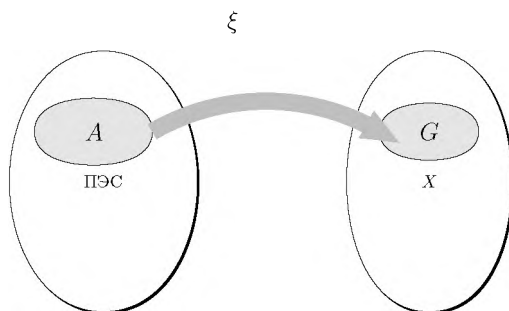
Дадим общее определение. Пусть заданы (Ω, \mathcal{A}, P) , т. е. ПЭС, множество событий, замкнутое относительно сложения, умножения и дополнения, и функция вероятности.



Определение. Если задано отображение $\xi : \Omega \rightarrow X$ из ПЭС в числовое множество, то говорят, что задана *случайная величина* (СВ) ξ .

Это фактически означает, что каждому исходу $\omega \in \Omega$ по некоторому правилу сопоставляется определенное число $x \in X$. Таким образом, можно говорить о вероятностях того, что в результате ЭСИ будут получены определенные значения СВ, $P(\omega : \xi(\omega) \mapsto x)$. В теории вероятностей функциональная природа СВ не является объектом рассмотрения. Мы обращаем внимание только на правую часть схемы. А именно нас будет интересовать вероятность того, что значение СВ попадает в некоторое подмножество $G \subset X$. $P(A) = P(\omega \in A : \xi(\omega) \in G) = P(\xi \in G)$.

СВ полностью описывается *законом распределения*. Этот удобный описательный термин [35] мы будем подробно обсуждать ниже, заметим лишь, что закон распределения позволяет определить для СВ ξ вероятность принимать любое подмножество возможных значений $P(\xi \in G)$.



If (Ω, \mathcal{A}, P) is a probability space, then a **random variable** on Ω is a measurable function $\xi : (\Omega, \mathcal{A}) \rightarrow S$ to a measurable **space** S (frequently taken

to be the real numbers with the standard measure). Приведенный абзац взят должен продемонстрировать заявленное в предисловии ограничение уровня строгости изложения материала, но с сохранением принципиально важных свойств основных понятий.

Для инженерных приложений используется выделение **непрерывных** (НСВ) и **дискретных** (ДСВ) случайных величин со специфическими формами законов распределения.

5.2. Дискретные случайные величины

Множество значений ДСВ X конечное или счетное. В рассмотренном выше примере сумма выпавших очков на двух игральных кубиках — ДСВ.

Закон распределения для ДСВ задается *рядом распределения*: для каждого элемента $x_i \in X$ задается вероятность того, что $\xi = x_i$, точнее, вероятность события $\{\omega_i : \omega_i \rightarrow x_i\}$. Мы используем общепринятые упрощенные обозначения, оговаривая это обстоятельство.

В учебных задачах ряд распределения будет представлен таблицей вида

x_i	x_1	x_2	\dots	x_n
p_i	p_1	p_2	\dots	p_n

При решении учебных задач нам будут даны как ЭСИ, так и описание правила, задающего СВ. Для СВ, определенной как сумма очков на двух игральных кубиках, очевидно (см. рис. 5.2) получим следующий ряд распределения:

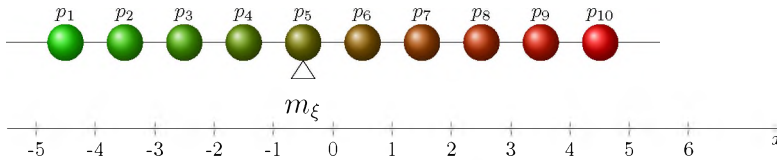
x_i	2	3	\dots	7	\dots	11	12
p_i	1/36	2/36	\dots	6/36	\dots	2/36	1/36

Надо заметить следующее: приведенная выше таблица фактически определяет функцию, в англоязычной традиции ее называют probability mass function (pmf), и именно она конкретизирует понятие «закон распределения» для ДСВ. Строго говоря, любая таблица такого вида, во второй строке которой стоят неотрицательные числа $p_i \geq 0$, $\sum p_i = 1$, определяет закон распределения ДСВ.

5.3. Числовые характеристики случайных величин

Кроме полного исчерпывающего описания СВ законом распределения, используется еще и сокращенное описание — числовыми характеристиками. Такое описание более наглядно для практических приложений, поскольку числовые характеристики имеют прямой теоретико-вероятностный смысл. Кроме того, при статистическом изучении СВ закон распределения определить значительно труднее.

Можно использовать механическую аналогию, ДСВ моделировать набором материальных точек, расположенных вдоль прямой с координатами x_i и имеющих массы $m_i = p_i$ (сравните с упомянутой выше pmf).



Определение. Математическое ожидание (МО) ДСВ задается определяющим соотношением:

$$M[\xi] = \sum_{k=1}^n x_k p_k. \quad (5.1)$$

Это *операторная* запись, она представляет правило, т. е. данную формулу, которая сопоставляет СВ как функции специфического вида действительное число.

Для обозначения самого значения оператора математического ожидания мы будем использовать более короткую запись — m_ξ (МО : $\xi \mapsto m_\xi$).

В механической модели СВ аналог МО — центр масс (тяжести) системы материальных точек на прямой.

В описании свойств числовых характеристик мы будем использовать некоторые обозначения и термины, не определенные ранее. Переменная C — «неслучайная» СВ, она принимает одинаковые значения для всех исходов и является аналогом постоянной функции в матанализе. Точное определение суммы $\xi_1 + \xi_2$, произведения $\xi_1 \cdot \xi_2$ и независимости СВ мы дадим позднее. Тогда же будут приведены и доказательства соответствующих свойств.

Свойства математического ожидания

1. $M[C] = C$;
2. $M[C\xi] = C \cdot m_\xi$;
3. $M[\xi_1 + \xi_2] = M[\xi_1] + M[\xi_2]$;
4. Для **НЕЗАВИСИМЫХ** $M[\xi_1 \cdot \xi_2] = M[\xi_1] \cdot M[\xi_2]$.

Вторая важная числовая характеристика СВ — **дисперсия**.

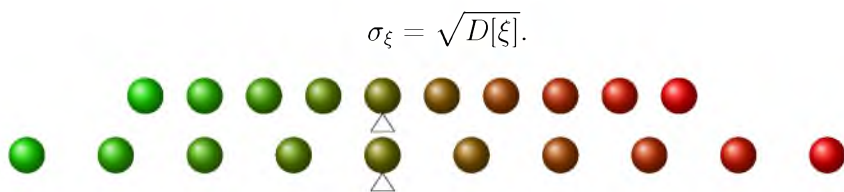
Определение. Дисперсия (только для ДСВ¹):

$$D[\xi] = M[(\xi - m_\xi)^2] = \sum_{k=1}^n (x_k - m_\xi)^2 \cdot p_k. \quad (5.2)$$

Часто приводят и словесный вариант: «Дисперсия — математическое ожидание квадрата отклонения СВ от своего МО».

Дисперсия характеризует разброс значений СВ относительно математического ожидания. Механический аналог — момент инерции системы материальных точек относительно оси вращения.

Дисперсия имеет размерность квадрата самой СВ, действительно, сравните: (метры → квадратные метры, рубли → квадратные рубли!). Поэтому в приложениях удобнее использовать другую характеристику — среднеквадратичное отклонение.



Чем больше σ_ξ , тем далее значения СВ отклоняются от МО — больше **разброс** значений.

¹Формулы для ДСВ и НСВ в начальных курсах ТВ отличаются — это частные случаи строгого математического определения. Замечание относится к обеим числовым характеристикам (МО и дисперсии).

Свойства дисперсии

1. $D[C] = 0$;
2. $D[C\xi] = C^2\sigma_\xi^2$;
3. Для **НЕЗАВИСИМЫХ** $D[\xi_1 \pm \xi_2] = D[\xi_1] + D[\xi_2]$.

В приведенных выше определениях слово *независимых* было намеренно выделено, поскольку никакого математического определения этого термина пока не было дано. Мы еще вернемся к более строгому рассмотрению этих свойств (матожидания и дисперсии) и даже проведем доказательства их справедливости.

5.4. Биномиальное распределение

Биномиальным называют распределение знакомой нам ДСВ — число успехов в схеме Бернулли. Можно дать и формальное определение, задав множество значений СВ и вероятности каждого значения. Фактически мы имеем дело с целым классом распределений, поскольку в формулу Бернулли входят два параметра: $0 < p < 1$ и $n \in \mathbb{N}$.

$$X = \{0, 1, 2, \dots, n\}; P_n(k) = C_n^k p^k (1-p)^{n-k}, k \in X. \quad (5.3)$$

Для распределения Бернулли можно продемонстрировать своеобразный «график» зависимости $P_n(k)$ (рис. 5.3). Кавычки поставлены, поскольку согласно определению графика функции² на рисунке должны стоять отдельные точки с координатами $(k, P_n(k))$. Для большей наглядности прибегают к использованному в данном случае изобразительному приему, существуют и другие приемы визуализации pmf (график именно этой функции и получился). Чаще всего это stem-diagram, пример которой мы уже показывали на с. 73 (рис. 4.1).

Рассмотрим числовые характеристики биномиального распределения. Матожидание вычисляем по определению:

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \Rightarrow M[\xi] = \sum k \cdot C_n^k p^k (1-p)^{n-k}.$$

²Графиком функции f называется множество точек плоскости с координатами $(x, f(x))$, $x \in X$.

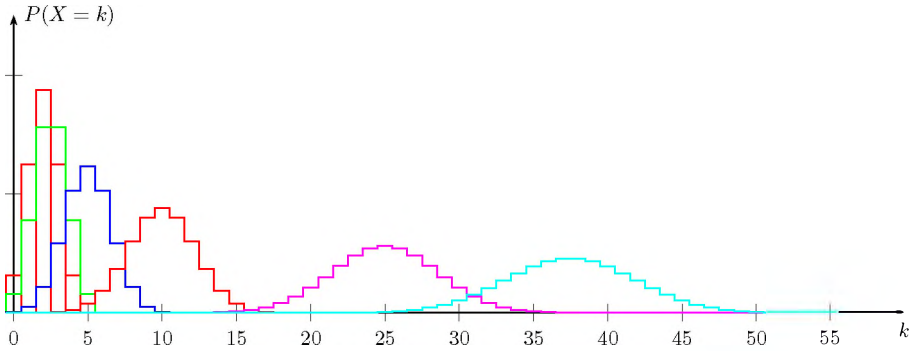


Рис. 5.3. Графики $P_n(k)$ для возрастающих значений параметра n ($p = 0.5$): точки соответствуют серединам горизонтальных отрезков. Увеличение n приводит к сдвигу максимума вправо

Непосредственно сделать это очень сложно, удобно ввести вспомогательную *индикаторную* СВ ν (индикатор). Если событие A произошло, то значение ν равно 1, в противном случае — 0. Получается ряд распределения:

x_i	0	1
p_i	$1 - p$	p

Вычисляем по определению: $M[\nu] = p$, $D[\nu] = p - p^2 = pq$. Число успехов k можно рассматривать как сумму значений индикаторов отдельных испытаний. Тогда для математического ожидания и дисперсии СВ k , с учетом независимости испытаний, получаем: $M[k] = np$, $D[k] = npq$.

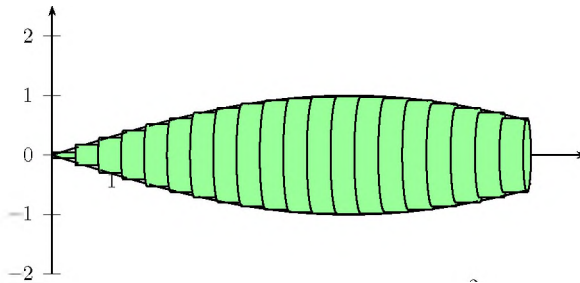
Заметим, что для числовых характеристик относительной частоты успехов получаются примечательные значения: $M\left[\frac{k}{n}\right] = p$, $D\left[\frac{k}{n}\right] = \frac{pq}{n} \xrightarrow{n \rightarrow \infty} 0$.

Это очень важные результаты! Вспомним свойство устойчивости относительной частоты, продемонстрированное на с. 29 (рис. 2.1), а также при компьютерном моделировании экспериментов Бюффона, ссылки на которые приведены на с. 29 и с. 46 (рис. 2.4). Асимптотическое поведение относительной частоты совершенно непохоже на асимптотику таких функций, как $\exp(-x)$ или $\arctg x$. Полученная асимптотика математического ожидания и дисперсии биномиального распределения является отражением этого экспериментально наблюдаемого свойства в теории и подтверждением ее непротиворечивости. Дальнейшее изучение этого вопроса будет проведено далее при рассмотрении *закона больших чисел*.

5.5. Случайные величины непрерывного типа

Второй тип случайной величины называют *непрерывными случайными величинами* (НСВ). Строгое определение пока давать не будем, ограничимся лишь замечанием: в отличие от ДСВ, принимающей конечное множество значений (или счетное), НСВ может принимать любые значения из промежутка (a, b) , объединения таких промежутков или на всем множестве \mathbb{R} .

Закон распределения НСВ не может быть задан в виде ряда распределения, как для ДСВ. Воспользуемся еще раз механической моделью. Для ДСВ такой моделью может служить система грузиков различной массы, расположенных вдоль невесомого стержня, а для НСВ — сплошной прямолинейный стержень, масса которого принимается за единицу.



В механике распределение массы стержня характеризуется линейной плотностью. Выделяем участок стержня от x до $x + \Delta x$ с массой Δm , составляем отношение $\frac{\Delta m}{\Delta x}$ (средняя линейная плотность), затем делаем предельный переход $\Delta x \rightarrow 0$.

Аналогичным образом можно задать и *распределение вероятности*. Выделим интервал от x до $x + \Delta x$, в который СВ попадает с вероятностью Δp , составим отношение и перейдем к пределу:

$$f_{\xi}(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta p}{\Delta x}. \quad (5.4)$$

Полученная функция f_{ξ} называется *плотностью [распределения] вероятности* (англоязычный аналог pdf — **p**robability **d**istribution **f**unction). Для краткости слово в квадратных скобках будем опускать, заметим, что преподаватели зарубежных университетов также на лекциях и в учебной литературе используют сокращения типа pmf, pdf, rv (random variable). Лаконичность в обозначениях — общематематическая традиция, правда, несколько нарушаемая в учебных пособиях начального уровня.

Свойства плотности вероятности

1. Определена на всей числовой оси (за исключением, быть может, некоторых изолированных точек).

2. Неотрицательность: $f_{\xi}(x) \geq 0$, $x \in (-\infty, +\infty)$.

3. Нормировка: $\int_{-\infty}^{+\infty} f_{\xi}(x) dx = 1$.

4. Вероятность попадания в интервал:

$$P(x_1 < \xi < x_2) = \int_{x_1}^{x_2} f_{\xi}(x) dx.$$

С использованием плотности вероятности мы можем вычислить числовые характеристики НСВ. Расчетные формулы:

$$M[\xi] = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx; \quad (5.5)$$

$$D[\xi] = \int_{-\infty}^{+\infty} (x - m_{\xi})^2 f_{\xi}(x) dx. \quad (5.6)$$

Ниже будут представлены плотности вероятности наиболее интересных в прикладном плане НСВ.

5.6. Функция распределения случайных величин

Неудобство различных форм закона распределения для ДСВ и НСВ очевидно, к тому же этими типами СВ далеко не исчерпываются.

Будем использовать короткое обозначение $\{\xi < x\}$ для события $\{\omega : \xi(\omega) < x\}$.

Определение. Функция $F_{\xi} : \mathbb{R} \mapsto [0, 1]$, $F_{\xi}(x) = P(\xi < x)$, $-\infty < x < +\infty$, называется *функцией распределения* СВ ξ .

Так как

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

то по аксиоме вероятности $P(\xi < x_2) = P(x_1 \leq \xi < x_2) + P(\xi < x_1)$

или

$$P(x_1 \leq \xi < x_2) = F_{\xi}(x_2) - F_{\xi}(x_1).$$

Далее, $P(\xi \geq x) = 1 - F_\xi(x)$. Значит, величина скачка в точке разрыва x_0 равна вероятности $P(\xi = x_0)$:

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

Свойства функции распределения

1. Определена и не убывает на всей числовой оси. Для $x_2 > x_1$
 $\{\xi < x_2\} = \{\xi < x_1\} \cup \{x_1 \leq \xi < x_2\} \Rightarrow$
 $\Rightarrow P(x_1 \leq \xi < x_2) = F_\xi(x_2) - F_\xi(x_1) \geq 0.$
2. Ограничена в области определения: $0 \leq F_\xi(x) \leq 1.$
3. Непрерывна слева в каждой точке: $F_\xi(x) = F_\xi(x - 0).$
4. Может иметь точки разрыва 1-го рода: $F_\xi(x + 0) = P(\xi \leq x).$
5. $P(x_1 \leq \xi \leq x_2) = F_\xi(x_2 + 0) - F_\xi(x_1), P(\xi = x) = F_\xi(x + 0) - F_\xi(x).$
6. Предельное поведение:

$$\lim_{x \rightarrow -\infty} F_\xi(x) = 0, \quad \lim_{x \rightarrow +\infty} F_\xi(x) = 1.$$

Таким образом, получаем, что для НСВ вероятность принимать определенное значение равна нулю. Для НСВ функция распределения непрерывна, а в точках дифференцируемости

$$f_\xi(x) = \lim_{\Delta x \rightarrow 0} \frac{P(\xi \in (x, x + \Delta x))}{\Delta x} = \frac{dF_\xi(x)}{dx}. \quad (5.7)$$

Есть еще **смешанные** СВ. В этом случае на некоторых промежутках распределение непрерывное, т. е. СВ может принимать любые значения. Кроме того, есть отдельные значения, которые СВ принимает с конечными вероятностями. Функция распределения такой СВ имеет точки разрыва первого рода, соответствующие отдельным значениям, и промежутки монотонного возрастания, где она непрерывна.

Функция распределения F_ξ , таким образом, является универсальной формой закона распределения СВ. В англоязычной традиции эта функция имеет

очень удачное название — **Cumulative distribution function (CDF)**, выражающее как определяющее соотношение, так и основное свойство — неубывание. Плотность вероятности, в свою очередь, применяется в физике и технических науках как более наглядное представление распределения вероятности, имеющее физическую аналогию (см. с. 83).

Закон распределения НСВ может быть задан двумя способами: функцией распределения F_ξ или плотностью вероятности f_ξ . Принятые для этих функций обозначения очень удобны, поскольку именно так в математическом анализе обозначают связанную пару функций — первообразную и производную. Действительно, согласно определениям и формуле (5.7),

$$f_\xi(x) = \frac{dF_\xi(x)}{dx}, \quad (5.8)$$

т. е. плотность вероятности есть производная функции распределения. Значит, функция распределения есть первообразная плотности вероятности:

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt. \quad (5.9)$$

Надо обязательно заметить, что в математически более строгом изложении теории вероятностей, начиная с основополагающих работ А. Н. Колмогорова, именно F_ξ рассматривается как основная форма закона распределения случайной величины ξ [2, с. 39]. Определение для функции f_ξ фактически является и определением непрерывных случайных величин.

Если для данной СВ ξ существует интегрируемая неотрицательная функция f_ξ , для которой справедливо тождество

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt,$$

ее называют *абсолютно непрерывной* или просто непрерывной случайной величиной (НСВ). Функция f_ξ называется плотностью вероятности.

Практически дословные определения даются и в большинстве серьезных учебников, см., например: [11, с. 56] или [7, с. 81].

Глава 6

Некоторые замечательные распределения

6.1. Распределение Пуассона

Это распределение числа успехов для редких явлений определяется знаковой формулой (5.3), ранее приведенной на с. 81. Однако можно не вспоминать ее происхождение и определить дискретную случайную величину (ДСВ) (рис. 6.1) со следующим законом распределения:

$$X = \{0, 1, 2, \dots\}; P(k) = \frac{\lambda^k}{k!} e^{-\lambda}, k \in \mathbb{N},$$

где X — множество значений (бесконечное); $P(k)$ — аналитическое представление ряда распределения.

Теорема 5. Для закона Пуассона имеем $\mathbf{M}[k] = \lambda$, $\mathbf{D}[k] = \lambda$.

Доказательство

$$\mathbf{M}[\xi] = \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = e^{-\lambda} \lambda e^{\lambda} = \lambda.$$

$$\begin{aligned} \mathbf{D}[\xi] &= \mathbf{M}[\xi^2] - \mathbf{M}[\xi]^2 = \sum_{m=0}^{\infty} m^2 \frac{\lambda^m}{m!} e^{-\lambda} - \lambda^2 = \\ &= e^{-\lambda} \lambda \sum_{m=1}^{\infty} \lambda^{m-1} \frac{m}{(m-1)!} - \lambda^2. \end{aligned}$$

Заметим, что для $m \geq 2$ имеем

$$\frac{m}{(m-1)!} = \frac{(m-1) + 1}{(m-1)!} = \frac{1}{(m-2)!} + \frac{1}{(m-1)!},$$

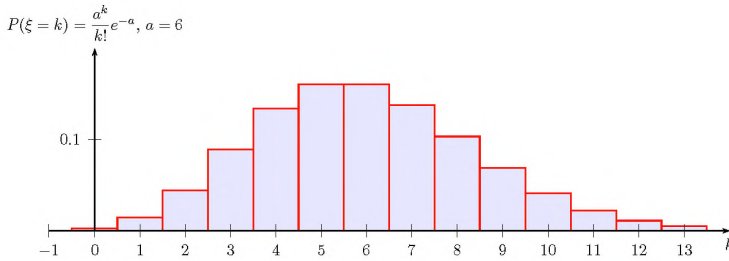


Рис. 6.1. График распределения Пуассона в форме столбчатой диаграммы (гистограммы)

ПОЭТОМУ

$$\begin{aligned} \sum_{m=1}^{\infty} m \frac{\lambda^{m-1}}{(m-1)!} &= 1 + \sum_{m=2}^{\infty} \frac{\lambda^{m-1}}{(m-2)!} + \sum_{m=2}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \\ &= 1 + \lambda \sum_{m=2}^{\infty} \frac{\lambda^{m-2}}{(m-2)!} + \sum_{m=2}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = 1 + \lambda e^{\lambda} + e^{\lambda} - 1. \end{aligned}$$

Таким образом,

$$D[\xi] = e^{-\lambda} \lambda \sum_{m=1}^{\infty} m \frac{\lambda^{m-1}}{(m-1)!} - \lambda^2 = e^{-\lambda} \lambda (\lambda e^{\lambda} + e^{\lambda}) - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Теорема доказана. □

6.2. Равномерное распределение

Определение. Равномерный закон — это закон распределения СВ ξ , плотность вероятности которой задается формулой

$$f_{\xi}(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a < x < b, \\ 0, & x > b, \end{cases} \quad (6.1)$$

где a, b — некоторые числа, причем $a < b$. Используется краткое обозначение $\xi \sim U(a, b)$.

Числовые характеристики случайной величины, распределенной равномерно (РРСВ), вычисляются очень легко:

$$\mathbf{M}[\xi] = \frac{a + b}{2}, \quad \mathbf{D}[\xi] = \frac{(b - a)^2}{12}.$$

РРСВ можно считать время ожидания транспортного средства, прибывающего на остановку через равные промежутки времени (время ожидания поезда метро). При считывании показаний на стрелочном приборе результат берется по ближайшему делению, ошибка рассматривается как РРСВ на отрезке: $[-\frac{T}{2}, \frac{T}{2}]$, где T — цена деления шкалы.

Одно из важных приложений равномерного распределения — описание погрешностей аналого-цифрового преобразования. В ходе этой операции непрерывный сигнал, который в данный момент времени может иметь любое значение, преобразуется в двоичное число с определенным количеством разрядов. Поскольку с помощью ограниченного количества разрядов можно представить только дискретный набор величин, появляется погрешность преобразования, представляющая собой разность между фактическим значением сигнала и ближайшей дискретной величиной.

Если гармонический сигнал передается из одного места в другое на некотором расстоянии от первого, то при условии, что длина пути распространения во много раз превышает длину волны сигнала, его фаза в точке приема может с полным основанием считаться случайной. Поскольку трудно указать какую-либо физическую причину, в соответствии с которой нужно было бы одно из значений фазового угла предпочесть другим, обычно предполагают, что фаза распределена равномерно в пределах угла 2π .

6.3. Показательный (экспоненциальный) закон распределения

Поскольку *показательное распределение* имеют СВ непрерывного типа, его задают плотностью вероятности.

Определение. Показательным называют закон распределения НСВ ξ , определяемый плотностью вероятности вида (рис. 6.2):

$$f_{\xi}(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases} \quad (6.2)$$

Используется также и название *экспоненциальный*, которое прямо обозначает используемую элементарную функцию. Здесь $\lambda > 0$ — параметр. Очевидно, что функция распределения (первообразная для f_ξ) будет иметь вид:

$$F_\xi(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases} \quad (6.3)$$

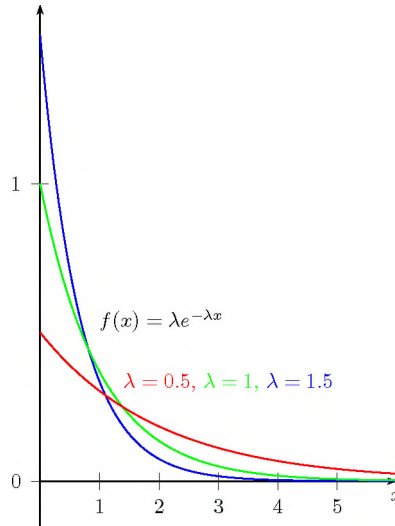


Рис. 6.2. Графики показательного распределения при различных λ

Для показательного закона характерным является равенство

$$\mathbf{M}[\xi] = \sigma_\xi = \sqrt{\mathbf{D}[\xi]} = \frac{1}{\lambda}.$$

Действительно, интегрируя по частям, получаем

$$\mathbf{M}[\xi] = \int_0^{+\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \cdot (-\lambda x \cdot e^{-\lambda x} - e^{-\lambda x}) \Big|_0^{+\infty} = \frac{1}{\lambda}.$$

Аналогично, дважды интегрируя по частям, получаем

$$\begin{aligned} \mathbf{D}[\xi] &= \mathbf{M}[\xi^2] - \mathbf{M}[\xi]^2 = \int_0^{+\infty} x^2 \cdot \lambda \cdot e^{-\lambda x} dx - \frac{1}{\lambda^2} = \\ &= \left(-x^2 \cdot e^{-\lambda x} - \frac{2x}{\lambda} \cdot e^{-\lambda x} - \frac{2}{\lambda^2} \cdot e^{-\lambda x} \right) \Big|_0^{+\infty} - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \end{aligned}$$

Во многих приложениях, например в теории надежности, СВ ξ имеет смысл времени безотказной работы. Тогда математическое ожидание $m_\xi = T$ интерпретируется как среднее время безотказной работы, а в атомной физике — как период полураспада. В этих случаях параметр распределения переопределяют так: $\lambda = \frac{1}{T}$.

Еще одно важное свойство показательного распределения — отсутствие *последствия*. В строгой формулировке оно звучит следующим образом:

Теорема 6. *Для случайной величины, распределенной по показательному закону, условная вероятность того, что ξ примет значение в пределах $0 \leq \alpha < \xi < \alpha + \beta$, при условии $\alpha \geq \xi$, равна:*

$$p(\alpha < \xi < \alpha + \beta | \alpha < \xi) = F(\beta).$$

В частности, эта условная вероятность не зависит от α .

Доказательство

$$\begin{aligned} p(\alpha < \xi < \alpha + \beta | \alpha < \xi) &= \frac{p(\alpha < \xi < \alpha + \beta)}{p(\xi \geq \alpha)} = \\ &= \frac{F(\alpha + \beta) - F(\alpha)}{1 - F(\alpha)} = \frac{e^{-\lambda\alpha} - e^{-\lambda(\alpha+\beta)}}{1 - (1 - e^{-\lambda\alpha})} = 1 - e^{-\lambda\beta} = F(\beta). \end{aligned}$$

□

Например, если время безаварийной работы устройства распределено по показательному закону (что иногда наблюдается у устройств, не имеющих изнашивающихся частей), то вероятность того, что оно не выйдет из строя, начиная от данного момента до истечения заданного промежутка времени, не зависит от того, сколько оно проработало до данного момента времени.

Предположим, что событие может произойти в любой момент времени. Причем вероятность того, что оно произойдет в промежутке $(t, t + \tau)$, пропорциональна τ и не зависит от t . В этом случае момент времени наступления события является СВ, имеющей показательное распределение.

6.4. Нормальный закон распределения

Определение. Нормальный закон распределения, или закон Гаусса, задается следующей плотностью вероятности:

$$f_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (6.4)$$

где a и σ — параметры распределения. Условное обозначение: $\xi \sim \mathbf{N}(a, \sigma)$.

Указанные параметры распределения случайной величины ξ имеют прямой теоретико-вероятностный смысл: $a = \mathbf{M}[\xi]$, σ — среднеквадратичное отклонение, поскольку $\mathbf{D}[\xi] = \sigma^2$.

Для получения этого результата воспользуемся одним известным интегралом (интеграл Пуассона):

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1.$$

$$M[\xi] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dt, \quad (6.5)$$

Сделаем в (6.5) замену $\frac{x-a}{\sigma} = t$, это соответствует последовательному выполнению преобразований *центрирования* и *нормирования* СВ.

$$M[\xi] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (t\sigma + a)e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ae^{-\frac{t^2}{2}} dt = a \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = a.$$

Представляем интеграл в виде суммы двух интегралов, первый равен нулю как интеграл от нечетной функции в симметричных пределах. В оставшемся, вынося постоянную a , получаем желаемый результат. Аналогичная выкладка для дисперсии (σ^2) гораздо более громоздкая, тем не менее ее можно сделать и самостоятельно. Как значение параметра σ^2 влияет на форму графика функции f_{ξ} показано на рис. 6.3.

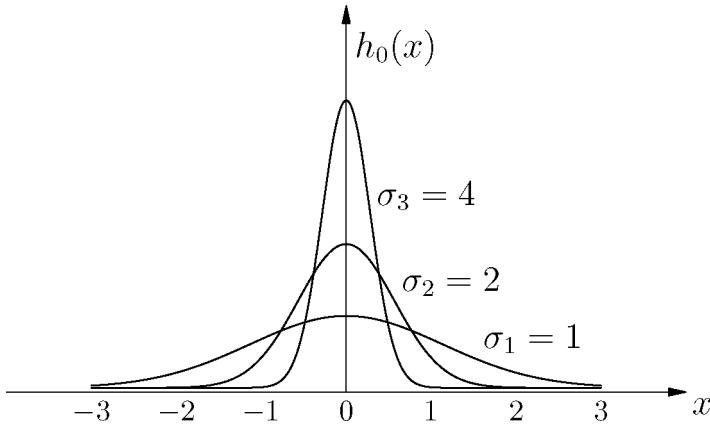


Рис. 6.3. Графики плотности вероятности нормального распределения (кривая Гаусса) при $a = 0$ и разных σ

Функция распределения выражается через функцию Лапласа Φ , которая табулирована и встроена в любую математическую систему (MathCad, StatGraph etc).

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right),$$

здесь использовано следующее определение функции Лапласа:

$$\Phi(x) = \frac{1}{2\pi} \int_{-\infty}^x e^{-t^2/2} dt.$$

Полезная формула:

$$P(\alpha < \xi < \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \quad (6.6)$$

Нормальное распределение применяется чрезвычайно широко по причинам, которые будут обсуждаться позднее. Оно прекрасно описывает распределение людей одной национальности и пола по росту, по массе тела, размеру обуви и многие другие величины. На нем основана и система допусков и посадок — основа конструкторских расчетов в машиностроении.

Полезно знать интересное свойство этого распределения, фактически это геометрическое свойство кривой Гаусса — *правило «3σ»*. Площадь криволинейной трапеции, ограниченной сверху этой кривой, а также вертикальными

прямыми $x = \pm 3\sigma$, с очень высокой точностью равна единице. Именно поэтому распределение можно применять и к таким величинам, которые по своей физической природе принимают ограниченное множество значений (рис. 6.4).

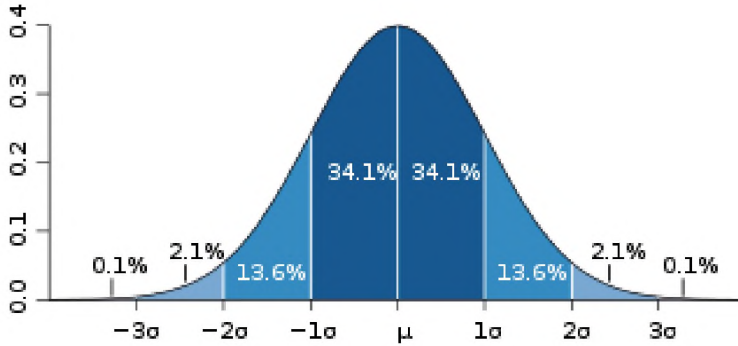


Рис. 6.4. График плотности вероятности нормального распределения. Отмечены характерные значения вероятностей для интервалов

В связи с удивительной вездесущностью нормального распределения квалифицированный специалист в любой области (наука, производство, торговля, социология и пр.) должен знать эти свойства кривой Гаусса.

Глава 7

Случайные векторы (системы случайных величин)

Рассмотрим эксперимент со случайным исходом, которому соответствует некоторое ПЭС Ω . Пусть теперь каждому исходу по некоторому правилу сопоставлена упорядоченная пара действительных чисел (x, y) . При этом $x \in X$, а $y \in Y$. Получаем отображение ПЭС: $\Omega \mapsto X \times Y$ (пример такого отображения представлен на рис. 7.1). Фактически мы имеем две СВ ξ и η , определенные на одном множестве исходов Ω .

Однако *совместное* рассмотрение и описание двух СВ как составляющих одного *случайного вектора* (говорят еще о *системе* случайных величин) гораздо более информативно, чем раздельное. Как мы убедимся далее, это дает нам возможность корректно рассматривать вопрос о *зависимости* или *независимости* СВ.

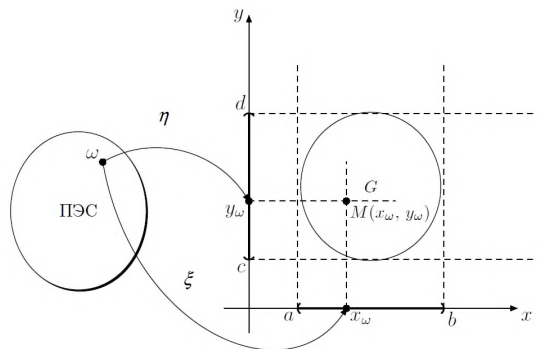


Рис. 7.1. Схема системы двух случайных величин

Небольшое обсуждение можно провести прямо сейчас. Пусть составляющие случайного вектора $\xi: \Omega \mapsto (a, b)$ и $\eta: \Omega \mapsto (c, d)$, а множество его

значений — G (они изображены на рис. 7.2). Возьмем две точки: $(x_1, y_1) \in G$ и $(x_1, y_2) \notin G$. Понятно, что пара (x_1, y_2) не является значением случайного вектора, ей не соответствует никакой исход и $P((\xi, \eta) = (x, y)) = 0$. Однако $y_2 \in (c, d)$ и как $P(\eta = y_2) \neq 0$, так и $P(\xi = x_1) \neq 0$. Поэтому принимает величина ξ значение x_1 или не принимает, зависит от значения второй составляющей η . Пара (x_1, y_1) может наблюдаться, а пара (x_1, y_2) — нет. В этом и состоит *зависимость* между СВ ξ и η .

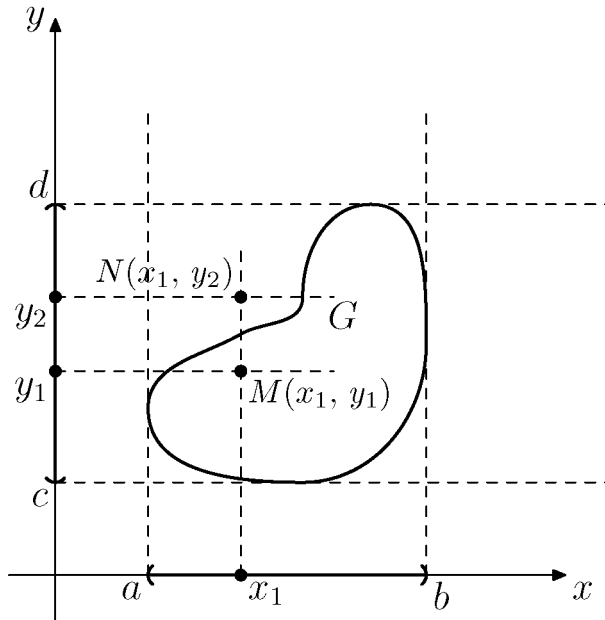


Рис. 7.2. Значения случайного вектора при фиксированном значении одной составляющей

7.1. Закон распределения случайного вектора с дискретными составляющими

Мы уже знаем, что случайные величины задаются *законом распределения*. Для ДСВ — это ряд распределения, а для НСВ — функция плотности вероятности f_ξ . Наиболее общая форма закона распределения СВ произвольного типа — функция распределения F_ξ .

Случайный вектор с двумя дискретными составляющими задается специальной таблицей. Каждая строка соответствует одному значению первой со-

ставляющей ξ , а каждый столбец — значению второй составляющей η . На пересечении строки i и столбца j записана вероятность $p_{ij} = P(\xi = x_i, \eta = y_j)$.

$$P((\xi, \eta) \in D) = \sum_{(x_i, y_j) \in D} p_{ij}.$$

$\xi \backslash \eta$	y_1	\dots	y_j	\dots	y_n	
x_1	p_{11}	\dots	p_{1j}	\dots	p_{1n}	$p^{(1)}$
x_2	p_{21}	\dots	p_{2j}	\dots	p_{2n}	$p^{(2)}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_i	p_{i1}	\dots	p_{ij}	\dots	p_{in}	$p^{(i)}$
\dots	\dots	\dots	\dots	\dots	\dots	\dots
x_m	p_{m1}	\dots	p_{mj}	\dots	p_{mn}	$p^{(m)}$
	p_1	\dots	p_j	\dots	p_n	

В англоязычной традиции используется обозначение pmf (Probability Mass Functin) и символ $f(i, j)$. Название отсылает нас к механической модели случайных величин,

Это полное описание двумерного случайного вектора. Используя данную таблицу, можно найти вероятность того, что случайный вектор принимает значения в произвольном множестве $D \subset R^2$. Для этого нужно просто сложить вероятности тех комбинаций (пар (x_i, y_j)) значений составляющих, для которых $(x_i, y_j) \in D$:

$$P((\xi, \eta) \in D) = \sum_{(x_i, y_j) \in D} p_{ij}.$$

Кроме того, по этой таблице мы легко находим законы распределения каждой из составляющих случайного вектора (их называют для определенности *безусловными*). Англоязычный аналог — «marginal» distribution, в отличие опять же от «conditional». Первый и последний столбцы содержат *безусловный* закон распределения СВ ξ , первая и последняя строки — для η .

$$p_j = \sum_{i=1}^m p_{ij}, p^{(i)} = \sum_{j=1}^n p_{ij}.$$

Математические ожидания находятся как обычно:

$$m_\xi = M[\xi] = \sum_{i=1}^m x_i p^{(i)}, m_\eta = M[\eta] = \sum_{j=1}^n y_j p_j.$$

Пара чисел (m_ξ, m_η) называется *центром распределения*. Можно предложить механическую модель случайного вектора с дискретными составляющими в виде расположенных на плоскости материальных точек с координатами (x_i, y_j) и массами p_{ij} . В этой модели (m_ξ, m_η) — положение центра тяжести системы.

Совместное описание двух СВ *более информативно*, чем отдельное. Действительно, восстановить всю таблицу по отдельным рядам распределения составляющих, вообще говоря, невозможно.

Есть более «дружелюбное» для исследователя (хотя и менее информативное) описание взаимозависимости двух составляющих. Это *условные законы* распределения:

$$P(\xi = x_i | \eta = y_j) = \frac{p_{ij}}{p_j}, \quad P(\eta = y_j | \xi = x_i) = \frac{p_{ij}}{p^{(i)}}.$$

В этом случае зависимость одной составляющей от другой проявляется в том, что ряд распределения ξ при $\eta = y_j$ отличается от такового при $\eta = y_{j_2}$.

Для каждого условного закона распределения находим *условное* матожидание. Для каждой составляющей оно вычисляется как обычно, но с использованием условных вероятностей, соответствующих фиксированному значению другой составляющей. Например, $M[\xi | \eta = y_j] = \sum x_i p^{(i)}$.

Функциональная зависимость условного матожидания одной составляющей от значения другой составляющей называется *регрессией*.

Условные законы и их числовые характеристики давно и эффективно применяются при статистических исследованиях. Допустим, проведен некий опрос группы покупателей о том, купили ли бы они определенный товар. Полученные результаты позволяют обнаружить гендерную зависимость формы ответа.

	Да	Нет	Не знаю	Всего
Мужчины	110	85	25	220
Женщины	185	75	20	280
Все	295	160	45	500

Оформление результатов опроса как измерения значения случайного вектора (двумерная выборка) позволяет судить о предпочтительности данного товара и среди всех покупателей, и отдельно среди женщин и мужчин. Для получения аналогов вероятностей абсолютные значения, стоящие в подобных таблицах, обычно пересчитывают в относительные (проценты или доли) —

это более наглядно. В данном исследовании сравнение безусловной доли ответивших «Да» (среди всех): $\frac{295}{500} = 0.59$, и условной (среди женщин): $\frac{185}{280} = 0.66$, позволяет сделать вывод о гендерной зависимости привлекательности товара.

Любопытно происхождение термина *регрессия*, особенно с учетом его обычного понимания как ухудшения или уменьшения. В Математическом энциклопедическом словаре приводится следующая информация по этому поводу: «Первоначально термин "регрессия"(математич.) был употреблен английским статистиком Ф. Гальтоном (1886) в теории наследственности в следующем специальном смысле: "возвратом к среднему состоянию"(*regression to mediocrity*) было названо явление, состоящее в том, что дети тех родителей, рост которых превышает среднее значение на a единиц, имеют в среднем рост, превышающий среднее значение меньше, чем на a единиц» [35, с. 523]. Можно посмотреть и соответствующую визуализацию (рис. 7.3), взятую из открытых интернет-источников.

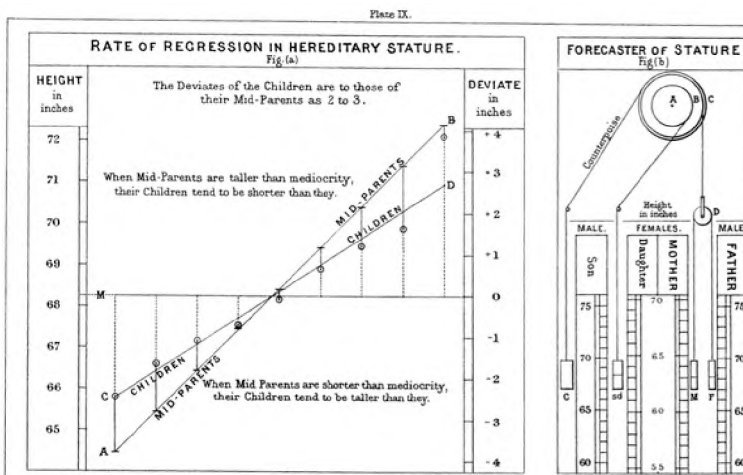


Рис. 7.3. Иллюстрация из оригинальной работы Ф. Гальтона

7.2. Закон распределения случайного вектора с непрерывными составляющими

Двумерный вектор с непрерывными составляющими задается неотрицательной функцией *плотность вероятности*, зависящей от двух переменных (рис. 7.4):

$$f_{\xi\eta}(x, y), f_{\xi\eta} : \mathbb{R}^2 \mapsto \mathbb{R}_+. \quad (7.1)$$

Функцию $f_{\xi\eta}$ можно определить аналогично одномерному случаю как предел средней плотности вероятности:

$$f_{\xi\eta}(x, y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(x < \xi < x + \Delta x, y < \eta < y + \Delta y)}{\Delta x \Delta y}. \quad (7.2)$$

Вероятность того, что вектор примет значения во множестве $A \subseteq \mathbb{R}^2$, равна

$$\iint_A f_{\xi\eta}(x, y) dx dy.$$

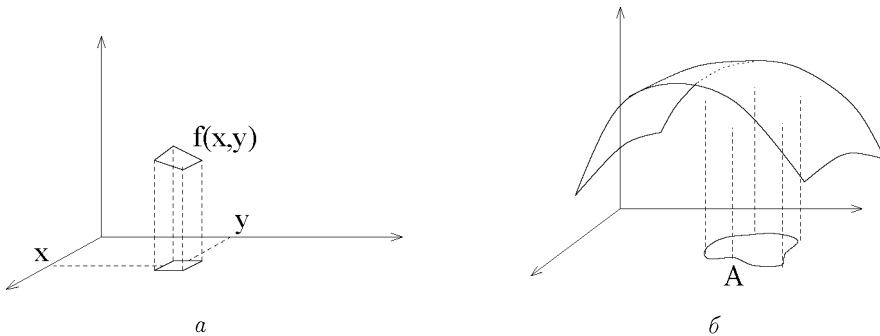


Рис. 7.4. Геометрический смысл плотности вероятности: *a* — вероятность попадания в $\Delta x \times \Delta y$ равна $f(x, y)\Delta x\Delta y$; *б* — вероятность попадания вектора во множество A

Представление закона распределения в форме $f_{\xi\eta}$ удобно при использовании в приложениях ТВ к специальным дисциплинам. Явное аналитическое выражение для функции в этих случаях может быть записано из модельных соображений или получено в результате вычислений.

Как и для дискретного случайного вектора, зная (7.2), мы можем получить безусловные законы распределения составляющих.

$$f_{\xi}(x) = \int_{-\infty}^{+\infty} f_{\xi\eta}(x, y) dy, \quad f_{\eta}(y) = \int_{-\infty}^{+\infty} f_{\xi\eta}(x, y) dx.$$

Например, найдем законы распределения составляющих вектора с плотностью $f_{\xi\eta}(x, y) = \frac{1}{2\pi} \exp -(x^2 + y^2)/2$. Совершенно ясно, они также будут распределены нормально.

Определяют и *условные* плотности распределения составляющих:

$$\varphi(x|y) = \frac{f_{\xi\eta}(x, y)}{f_{\eta}(y)}, f_{\eta}(y) > 0.$$

Набор графиков такой функции при различных значениях параметра y описывает зависимость между СВ.

Пример 11. Совместная плотность вероятности двух СВ

$$f_{\xi\eta} = \begin{cases} \frac{12}{5}x(2-x-y), & 0 \leq x \leq 1, \quad 0 \leq y \leq 1; \\ 0 & x \notin [0, 1], \quad y \notin [0, 1]. \end{cases}$$

Найти условную плотность составляющей ξ .

Решение

Ясно, что нетривиальный вид этой функции будет только в первой четверти. Поскольку

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dx = \int_0^1 \frac{12}{5}x(2-x-y) dx = \frac{2}{3} - \frac{y}{2},$$

получаем

$$\varphi(x|y) = \frac{f_{\xi\eta}(x, y)}{f_{\eta}(y)} = \frac{6x(2-x-y)}{4-3y}.$$

Несколько графиков этой функции представлены на рис. 7.5 (с. 103). ■

7.3. Функция распределения случайного вектора

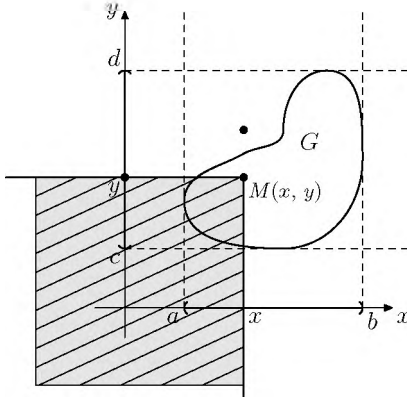
Двумерная функция распределения — универсальная форма закона распределения для произвольного случайного вектора.

$$F_{\xi\eta}(x, y) = P(\xi < x, \eta < y) = P(\{\omega \in \Omega | \xi(\omega) < x, \eta(\omega) < y\}). \quad (7.3)$$

Свойства функции (7.3) (наглядная интерпретация на рис. 7.6) приведем без подробного обоснования:

1. $F_{\xi\eta} : \mathbb{R}^2 \mapsto [0, 1]$.

2. Не убывает ни по одному из аргументов.
3. Безусловные законы распределения составляющих.



Интуитивно понятно, что сдвигая одну из сторон угла, мы получим маргинальную функцию распределения только для одной СВ: если горизонтальную вверх, то мы получаем условие $\xi < x$, если вертикальную вправо, то $\eta < y$. Значит,

$$\lim_{x \rightarrow +\infty} F_{\xi\eta}(x, y) = F_{\eta}(y),$$

$$\lim_{y \rightarrow +\infty} F_{\xi\eta}(x, y) = F_{\xi}(x).$$

4. Из тех же наглядно-геометрических соображений получаем еще два предельных свойства. Первый предельный переход дает одномерную функцию распределения, а второй — применяется к одномерной (ее свойства нам уже известны).

$$\lim_{\substack{x \rightarrow +\infty, \\ y \rightarrow +\infty}} F_{\xi\eta}(x, y) = 1,$$

$$\lim_{x \rightarrow -\infty, y \rightarrow -\infty} F_{\xi\eta}(x, y) = 0.$$

5. Вероятность попадания в прямоугольную область. Получается на основе определения (7.3) и формулы сложения для вероятностей:

$$P(x_1 \leq \xi < x_2, y_1 \leq \eta < y_2) = F_{\xi\eta}(x_2, y_2) - F_{\xi\eta}(x_1, y_2) - F_{\xi\eta}(x_2, y_1) + F_{\xi\eta}(x_1, y_1).$$

Подставим правую часть последнего равенства в (7.2), получим важное соотношение

$$f_{\xi\eta}(x, y) = \frac{\partial^2 F_{\xi\eta}(x, y)}{\partial x \partial y}.$$

Можно и наоборот — найти функцию распределения по плотности вероятности:

$$F_{\xi\eta}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{\xi\eta}(x', y') dx' dy'.$$

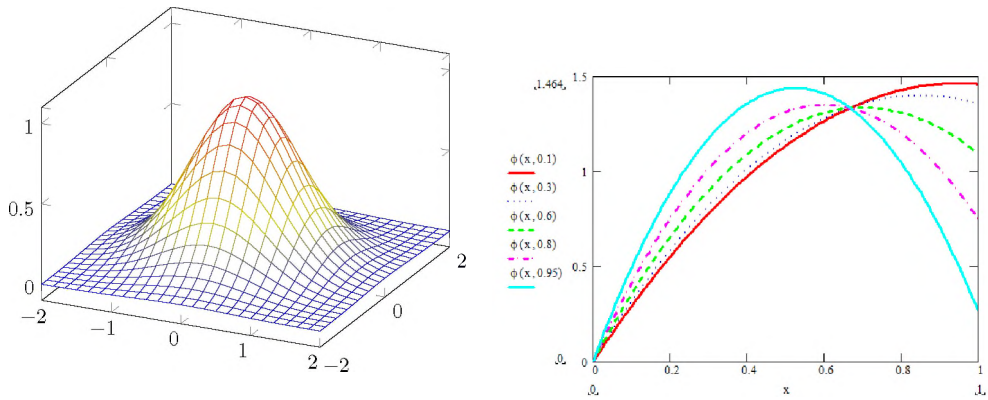
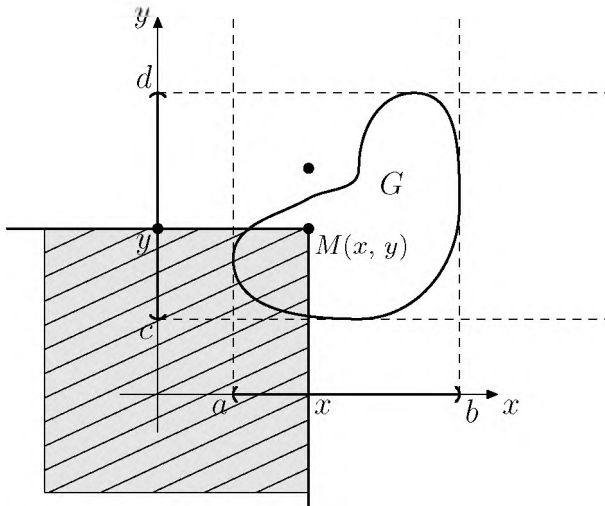


Рис. 7.5. Графики условных вероятностей

Рис. 7.6. Геометрический смысл значения функции распределения случайного вектора $F_{\xi\eta}(x, y)$ — вероятность попадания в бесконечный угол с вершиной в точке (x, y)

Глава 8

Зависимые и независимые случайные величины

Концепция статистической *независимости* обсуждалась нами ранее (раздел 3.5.). Альтернатива независимости для случайных величин также требует аккуратного и детального рассмотрения. Прежде всего сама постановка вопроса о зависимости/независимости случайных(!) величин предполагает, что эти величины заданы на одном и том же ПЭС, связанном с некоторым стохастическим явлением. Это важное обстоятельство было отмечено в главе 7. Совместное рассмотрение двух (и более) СВ позволяет изучать и описывать связь между ними (зависимость одной СВ от другой).

Пусть определено *вероятностное пространство* $\langle \Omega, \mathcal{A}, \mathcal{P} \rangle$ и две СВ. В теории вероятностей *независимость* случайных событий характеризуется критерием

$$P(A \cdot B) = P(A) \cdot P(B). \quad (8.1)$$

Опираясь на это определение, введем понятие независимости для СВ ξ и η . Во-первых, они должны рассматриваться совместно, т. е. быть составляющими случайного вектора, определенного на ПЭС некоторого эксперимента со случайным исходом и принимающего значения в $X \times Y \subseteq \mathbb{R}^2$. Кроме того, в соответствии с (8.1), для любых $G_x \subseteq X$, $G_y \subseteq Y$ должно иметь место

$$P(\{\xi \in G_x, \eta \in G_y\}) = P(\{\xi \in G_x\}) \cdot P(\{\eta \in G_y\}). \quad (8.2)$$

Определение. Назовем СВ $\xi: \Omega \mapsto X$ и $\eta: \Omega \mapsto Y$ независимыми, если справедливо (8.2).

Теорема 7. Для независимости СВ $\xi: \Omega \mapsto X$ и $\eta: \Omega \mapsto Y$ необходимо и достаточно, чтобы $F_{\xi\eta}(x, y) = F_{\xi}(x) \cdot F_{\eta}(y)$, $\forall (x, y) \in \mathbb{R}^2$.

В частных случаях для ДСВ

$$p_{ij} = p_i \cdot p^{(j)} \quad (i = 1, \dots, m, j = 1, \dots, n), \quad (8.3)$$

а для НСВ

$$f_{\xi\eta}(x, y) = f_\xi(x) \cdot f_\eta(y), \quad (\forall (x, y) \in \mathbb{R}^2). \quad (8.4)$$

8.1. Свойства математического ожидания и дисперсии

Ранее на с. 80 были перечислены основные свойства числовых характеристик СВ. Сейчас мы проведем вывод этих свойств.

Математическое ожидание суммы двух ДСВ:

$$\begin{aligned} M[\xi_1 + \xi_2] &= M[\eta] = \sum_{i,j} (x_i + x_j) p_{ij} = \sum_i x_i \sum_j p_{ij} + \\ &+ \sum_j x_j \sum_i p_{ij} = \sum_i x_i p_i + \sum_j x_j p_j = M[\xi_1] + M[\xi_2]. \end{aligned} \quad (8.5)$$

Результат справедлив и при наличии зависимости между слагаемыми.

Математическое ожидание произведения двух независимых СВ:

$$\begin{aligned} M[\xi_1 \cdot \xi_2] &= M[\eta] = \sum_{i,j} x_i \cdot x_j p_{ij} = \sum_{i,j} x_i \cdot x_j p_i p_j = \\ &= \sum_i x_i p_i \cdot \left(\sum_j p_j x_j \right) = M[\xi_1] \cdot M[\xi_2]. \end{aligned} \quad (8.6)$$

Выкладки, аналогичные проведенным в (8.5) и (8.6), можно провести и для НСВ. Например, для произведения независимых СВ:

$$M[\xi_1 \cdot \xi_2] = \iint_{\mathbb{R}^2} xy f_{\xi\eta} dx dy = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_\xi(x) f_\eta(y) = M[\xi_1] \cdot M[\xi_2]. \quad (8.7)$$

Запишем дисперсию суммы двух СВ, выполним перегруппировку аргумента оператора D , применим свойства матожидания:

$$\begin{aligned} D[\xi_1 \pm \xi_2] &= M\left[\left((\xi_1 \pm \xi_2) - (m_{\xi_1} \pm m_{\xi_2}) \right)^2 \right] = M[(\xi_1 \pm \xi_2)^2 - 2(\xi_1 \pm \xi_2) \cdot \\ &\cdot (m_{\xi_1} \pm m_{\xi_2}) + (m_{\xi_1} \pm m_{\xi_2})^2] = M[(\xi_1 - m_{\xi_1})^2] + M[(\xi_2 - m_{\xi_2})^2] \mp \\ &\mp 2M[\xi_1 \xi_2] - 2m_{\xi_1} m_{\xi_2} = D[\xi_1] + D[\xi_2] \pm 2K_{\xi_1, \xi_2}, \end{aligned} \quad (8.8)$$

где $K_{\xi_1\xi_2} = M[\xi_1\xi_2] - m_{\xi_1}m_{\xi_2}$, так называемый **корреляционный момент**. Для независимых СВ он в силу свойства матожидания произведения равен нулю, поэтому для независимых СВ

$$D[\xi_1 \pm \xi_2] = D[\xi_1] + D[\xi_2].$$

Результаты легко обобщаются и на системы более двух СВ.

8.2. Корреляционный момент. Коэффициент корреляции

Рассмотрим случайный вектор с двумя составляющими (ξ, η) . Числовые характеристики

$$M[\xi], M[\eta], D[\xi] = M[(\xi - m_\xi)^2], D[\eta] = M[(\eta - m_\eta)^2]$$

вычисляются по безусловным (маргинальным) законам и ничего о зависимости составляющих не говорят. Новым по сравнению с одномерным случаем будет *смешанный* момент второго порядка

$$k_{\xi\eta} = M[(\xi - m_\xi)(\eta - m_\eta)], \quad (8.9)$$

называемый *корреляционным* моментом, при наличии более двух составляющих получится матрица соответствующей размерности. Для вычисления $k_{\xi\eta}$ требуется совместный закон распределения составляющих случайного вектора. Если составляющие будут ДСВ, то совместный закон распределения задан таблицей p_{ij} . Тогда

$$k_{\xi\eta} = \sum_{i=1}^m \sum_{j=1}^n (x_i - m_\xi)(y_j - m_\eta)p_{ij}.$$

Рассмотрим подробнее теоретико-вероятностный смысл корреляционного момента. Для этого преобразуем определяющее соотношение (8.9) и используем свойства математического ожидания:

$$\begin{aligned} k_{\xi\eta} &= M[(\xi - m_\xi)(\eta - m_\eta)] = M[\xi \cdot \eta + m_\xi m_\eta - \xi m_\eta - m_\xi \eta] = \\ &= M[\xi \cdot \eta] - M[\xi]M[\eta]. \end{aligned}$$

Для независимых СВ $M[\xi \cdot \eta] = M[\xi]M[\eta]$ (это было доказано в (8.6)). Поэтому при $k_{\xi\eta} \neq 0$ имеет место *зависимость* между СВ. Следующий пример показывает, что обратное не имеет места, т. е. корреляционный момент может равняться нулю и при наличии зависимости между СВ.

Пример 12. Рассмотрим случайный вектор с непрерывными составляющими. Пусть

$$f_{\xi\eta}(x, y) = \begin{cases} \frac{1}{\pi R^2}, & x^2 + y^2 \leq R^2; \\ 0 & x^2 + y^2 > R^2. \end{cases}$$

Доказать, что составляющие не являются независимыми, но $k_{\xi\eta} = 0$.

Решение

Плотность вероятности не факторизуется: $f_{\xi\eta}(x, y) \neq f_{\xi}(x) \cdot f_{\eta}(y)$. Действительно, найдем маргинальную плотность вероятности составляющей ξ . Заметим, что интегрирование по y нетривиально только при $-R \leq x \leq R$.

$$f_{\xi}(x) = \int_{-\infty}^{+\infty} f_{\xi\eta}(x, y) dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2} dy = 2 \frac{\sqrt{R^2-x^2}}{\pi R^2}.$$

Аналогично и для $f_{\eta}(y)$ (с заменой $x \rightleftharpoons y$). Итак, критерий независимости не выполняется.

$$4 \frac{\sqrt{R^2-x^2}}{\pi R^2} \cdot \frac{\sqrt{R^2-y^2}}{\pi R^2} \neq \frac{1}{\pi R^2}. \quad (8.10)$$

Вычислим $k_{\xi\eta}$. Тот факт, что $m_{\xi} = m_{\eta} = 0$, ясен из соображений симметрии и механической интерпретации МО (центр тяжести диска с плотностью, заданной в условии, в его геометрическом центре). Можете, конечно, проверить вычислением.

Получаем для корреляционного момента

$$k_{\xi\eta} = \iint_{\mathbb{R}^2} xyf(x, y) dx dy = \frac{1}{\pi R^2} \iint_D xy dx dy = \frac{1}{\pi R^2} \int_{-R}^R x dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y dy = 0.$$

Здесь D — круг радиуса R с центром в начале координат. ■

Итак, равенство нулю корреляционного момента *необходимое*, но не *достаточное* условие независимости СВ (рис. 8.1). При $k_{\xi\eta} \neq 0$ имеет место особый вид зависимости, называемой *корреляционной*. Для СВ, связанных корреляционной зависимостью, дисперсия их суммы не равна сумме дисперсий.

$$\begin{aligned} D[\xi \pm \eta] &= M[(\xi \pm \eta) - (m_{\xi} \pm m_{\eta})]^2 = \\ &= M[\xi^2 + \eta^2 \pm \xi \cdot \eta] = D[\xi] + D[\eta] \pm 2M[\xi \cdot \eta]. \end{aligned} \quad (8.11)$$

$f(x,y)$		y			$f_X(x)$
		-1	0	1	
x	-1	0.20	0	0.20	0.4
	0	0	0.20	0	0.2
	1	0.20	0	0.20	0.4
$f_Y(y)$		0.4	0.2	0.4	1

Рис. 8.1. Пример закона распределения дискретного двумерного вектора (составляющие зависимы, но некоррелированы)

Корреляционный момент — величина размерная, его значение изменяется при изменении единиц измерения. Невозможно сравнивать степень корреляционной зависимости между парами величин различной размерности.

Определим операции *центрирования* $\overset{\circ}{\xi}$ и *нормирования* $\hat{\xi}$ СВ:

$$\overset{\circ}{\xi} = \xi - m_\xi, \quad \overset{\circ}{\eta} = \eta - m_\eta, \quad \hat{\xi} = \frac{\overset{\circ}{\xi} - m_\xi}{\sigma_\xi}, \quad \hat{\eta} = \frac{\overset{\circ}{\eta} - m_\eta}{\sigma_\eta}.$$

Нормированная СВ имеет нулевое матожидание и единичную дисперсию. Кроме того, она безразмерная.

Определение. Коэффициент корреляции (КК) СВ ξ и η — корреляционный момент нормированных СВ $\hat{\xi}$ и $\hat{\eta}$.

$$r_{\xi\eta} = \frac{k_{\xi\eta}}{\sigma_\xi\sigma_\eta}, \quad (8.12)$$

$$\frac{k_{\xi\eta}}{\sigma_\xi\sigma_\eta} = \frac{M[\overset{\circ}{\xi} \cdot \overset{\circ}{\eta}]}{\sigma_\xi\sigma_\eta} = M\left[\frac{\overset{\circ}{\xi}}{\sigma_\xi} \frac{\overset{\circ}{\eta}}{\sigma_\eta}\right] = M[\hat{\xi}\hat{\eta}] = K_{\hat{\xi}\hat{\eta}}.$$

Именно эта величина используется для количественной характеристики *тесноты* корреляционной связи.

Свойства $r_{\xi\eta}$:

1. КК — величина безразмерная, ее значения ограничены условием $|r_{\xi\eta}| \leq 1$. Это следует из (8.11) и (8.12):

$$D[\overset{\circ}{\xi} \pm \overset{\circ}{\eta}] = 2 \pm 2r_{\xi\eta} \geq 0.$$

2. Масштабное преобразование не меняет $|r_{\xi\eta}|$:

$$|r_{\alpha\xi, \beta\eta}| = \left| \frac{K_{\alpha\xi, \beta\eta}}{\sigma_{\alpha\xi} \sigma_{\beta\eta}} \right| = \left| \frac{\alpha\beta K_{\xi, \eta}}{|\alpha||\beta| \sigma_{\xi} \sigma_{\eta}} \right| = \left| \frac{K_{\xi, \eta}}{\sigma_{\xi} \sigma_{\eta}} \right| = |r_{\xi\eta}|.$$

3. Изменение начала отсчета переменных не меняет значения $r_{\xi\eta}$ (как и $K_{\xi\eta}$), поскольку не меняются разности $\xi - m_{\xi}$.
4. Равенство $|r_{\xi\eta}| = 1$ есть Н и Д условие линейной зависимости между ξ и η .

Последнее свойство рассмотрим подробнее.

Доказательство

Необходимость непосредственно следует из свойств коэффициента корреляции. Если $\eta = a\xi + b$, то $r_{\xi, a\xi+b} = r_{\xi\xi} = \frac{D[\xi]}{\sigma_{\xi}^2} = 1$.

Для доказательства достаточности воспользуемся формулами (8.11) и (8.12).

Пусть $r_{\xi\eta} = 1$, тогда $D[\overset{\circ}{\xi} - \overset{\circ}{\eta}] = 0$. Отсюда следует неслучайность, т. е. постоянство аргумента дисперсии. А если учесть, что его МО равно нулю, то получаем $\overset{\circ}{\xi} = \overset{\circ}{\eta}$.

Это фактически и означает наличие линейной зависимости. □

Глава 9

Неслучайные функции случайного аргумента

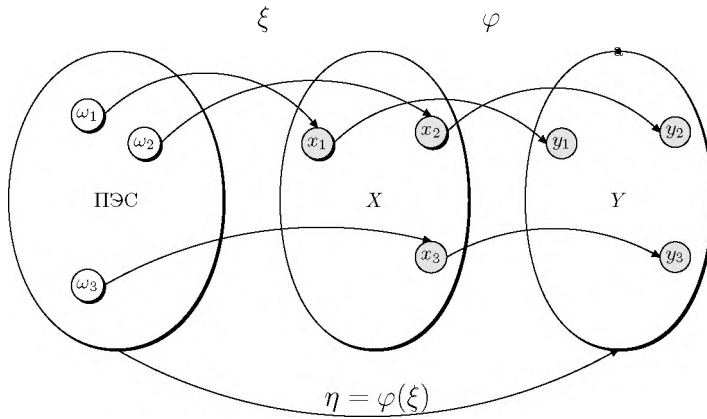
9.1. Функция случайной величины

Понятие *функция* хорошо вам известно из курса математического анализа. В приложениях математики в естественных науках *функция* используется как математическая модель некоторой объективно существующей закономерности, связывающей характеристики изучаемого объекта.

Определение. Если каждому элементу множества $X \subseteq \mathbb{R}$ по некоторому правилу сопоставляется единственный элемент множества $Y \subseteq \mathbb{R}$, то говорят, что задана функция с областью определения X и множеством значений Y , $f : X \mapsto Y$.

Пусть теперь X (область определения f) — множество значений некоторой СВ ξ . Тогда можно рассматривать и Y как множество значений СВ η . Действительно, если результатом ЭСИ явился исход $\omega_i \in \Omega$, то ξ приняла определенное (это важно!) значение $x_i \in X$, которому соответствует *определенное* $y_i = \varphi(x_i)$. Ни в первом отображении $\omega_i \xrightarrow{\xi} x_i$, ни во втором — $x_i \xrightarrow{\varphi} y_i$ никакой случайности нет! Вся случайность связана с появлением ω_i !

Событие $\{\xi = x_i\}$ совпадает с событием $\{\eta = y_i : y_i = \varphi(x_i)\}$, оба события имеют одно и ту же вероятность $P(\{\omega_i\})$. Таким образом, задавая на множестве значений случайной величины ξ некоторую функцию $\varphi : X \mapsto Y$, мы получаем новую случайную величину η . Используем обозначение $\eta = \varphi(\xi)$. В некоторых простейших задачах каких-то новых методов для изучения распределения вероятности η не требуется.



Определение. Функция СВ. Пусть на множестве X значений СВ $\xi : \Omega \mapsto X$ определена функция $\varphi : X \mapsto Y$. Тогда суперпозиция отображений $\Omega \xrightarrow{\xi} X \xrightarrow{\varphi} Y$ задает СВ η , называемую **функцией СВ** ξ , $\eta = \varphi(\xi)$.

Для примера рассмотрим закон распределения и числовые характеристики функции ДСВ.

Задача 12. СВ задана рядом распределения. Найти закон распределения СВ $\eta = \xi^2$ и ее числовые характеристики.

ξ	-2	-1	0	1	2
p_i	0.1	0.05	0.25	0.4	0.2

Решение

Множество значений СВ η :

$$Y = \{0, 1, 4\}.$$

Учитывая равенство событий, находим

$$\begin{aligned} \{\xi = 0\} &= \{\eta = 0\} \Rightarrow P(\xi = 0) = P(\eta = 0) = 0.25, \\ \{\eta = 1\} &= \{\xi = -1\} + \{\xi = 1\} \Rightarrow P(\eta = 1) = P(\xi = -1) + P(\xi = 1) = 0.45, \\ \{\eta = 4\} &= \{\xi = -2\} + \{\xi = 2\} \Rightarrow P(\eta = 4) = P(\xi = -2) + P(\xi = 2) = 0.3. \end{aligned}$$

Итак, закон распределения СВ η :

η	0	1	4
p_i	0.25	0.45	0.3

Математическое ожидание:

$$M[\eta] = 0 \cdot 0.25 + 1 \cdot 0.45 + 4 \cdot 0.3 = (-2)^2 \cdot 0.1 + (-1)^2 \cdot 0.05 + 0^2 \cdot 0.25 + 1^2 \cdot 0.4 + 2^2 \cdot 0.2 = 1.65 = M[\xi^2].$$

Получается, что для вычисления математического ожидания не требуется закон распределения самой СВ η , достаточно знать распределение только аргумента, СВ ξ . Это важно для практических целей обстоятельство мы рассмотрим ниже. ■

9.2. Функция непрерывного случайного аргумента

Случайные величины непрерывного типа при всей их актуальности в прикладном плане значительно сложнее дискретных в плане математическом. Это обстоятельство отмечал В. Феллер в обстоятельном предисловии еще к первому изданию своей замечательной книги по теории вероятностей [6], объясняя выделение НСВ в отдельный второй том. Работая со студентами технических специальностей, преподавателю приходится снижать уровень строгости при рассмотрении НСВ и решении задач, включающих этот тип СВ. Некоторая компенсация достигается большей опорой на интуитивный подход с активным включением математических визуализаций. В современной лингвистике даже появился термин «поликодовый текст»¹, хотя сама практика совмещения различных семантических систем, распределения информационного потока по визуальным и аудиоканалам хорошо изучена с различных точек зрения и прекрасно поддерживается новыми цифровыми технологиями [36].

Итак, нашей целью является нахождение закона распределения НСВ η , являющейся функцией другой НСВ ξ с известным законом распределения. Среди возможных методов решения этой задачи мы выберем *метод функции распределения* (МФР или cdf method), для применения которого не требуются знания, выходящие за пределы обычного для технических вузов курса высшей математики.

Основная идея метода МФР состоит в следующем. Во-первых, закон распределения каждой из СВ будем представлять функцией распределения. Это, кстати, позволяет работать с широким классом отображений $\mapsto^{\mathcal{L}}$. Во-вторых, будем исходить из определяющего соотношения: $F_{\xi}(x) = P(\xi < x)$. Надо найти функцию $F_{\eta}(y) = P(\eta < y)$, выразив ее значение через $P(\xi \in G_y)$, где $G_y = \{x : \varphi(x) < y\}$. В простейшем случае линейного отображения мы просто разрешаем параметрическое неравенство $\varphi(x) < y$. Итак, собы-

¹Есть и другие варианты: креолизованный текст, мультимодальный текст и др.

тие $\{a \cdot \xi + b < y\}$ очевидно совпадает с $\{\xi < \frac{y-b}{a}\}$ при $a > 0$ либо с $\{\xi > \frac{y-b}{a}\}$ при $a < 0$. Значит, $F_\eta(y) = F_\xi(\frac{y-b}{a})$ в первом случае и $F_\eta(y) = \left(1 - F_\xi(\frac{y-b}{a})\right)$ — во втором.

Для функций φ более сложного вида решение неравенства $\{x : \varphi(x) < y\}$ в общем виде сложно и практически бесполезно. Учитывая принятый нами уровень строгости, будем активно использовать геометрические свойства графиков φ , пренебрегая детальным аналитическим исследованием. Поскольку значения СВ ξ и η связаны функциональной зависимостью φ , то точки с координатами (ξ, η) можно рассматривать как результат измерения составляющих случайного вектора. Они будут ложиться на кривую $y = \varphi(x)$, но необязательно составлять ее всю(!).

Применим МФР для φ , строго возрастающей на $X_0 \subset X$ (ООФ φ), при этом используем наглядные соображения (рис. 9.1). Возрастание на всем X не обязательно. Ищем $F_\eta(y) = P(\eta < y)$, $y \in \mathbb{R}$. Очевидно, что событие $\{\eta < y_0\}$ равно событию $\{\xi < x_0 | y_0 = \varphi(x_0)\}$ (рис. 9.1).

$$\begin{aligned} P(\eta < y_0) &= P(\xi \in G_x) = P(\xi < \varphi_0^{-1}(y_0)) = F_\xi(\varphi_0^{-1}(y_0)), \\ G_x &= \{x | y_0 < \varphi_0(x) < y_0\} = (x, \varphi_0^{-1}(y)). \end{aligned} \quad (9.1)$$

Индекс 0 напоминает, что это «укороченная» функция, определенная на X_0 .

$$F_\eta(y) = P(\eta < y) = P(\xi < \varphi_0^{-1}(y)) = F_\xi(\varphi_0^{-1}(y)).$$

Здесь используется обозначение функции, обратной по отношению к φ_0 (для монотонной функции закономерно введение обратной). Положим $X_0 = (x_1, x_2) \Rightarrow Y = (y_1, y_2)$, $y_1 = \varphi(x_1)$, $y_2 = \varphi(x_2)$. Тогда

$$F_\xi(x) = \begin{cases} 0, & x < x_1, \\ P_x, & x_1 < x \leq x_2 \text{ (заданная функция)}, \\ 1, & x > x_2, \end{cases}$$

поэтому

$$F_\eta(y) = \begin{cases} 0, & y \leq y_1, \\ P(x < \varphi_0^{-1}(y)), & y_1 < y \leq y_2, \\ 1, & y > y_2. \end{cases}$$

Для НСВ закон распределения удобнее использовать в виде плотности вероятности $f_\eta(y)$.

$$\begin{aligned} f_\eta(y) &= \frac{dF_\eta(y)}{dy} = \frac{d}{dy} F_\xi(\varphi^{-1}(y)) = \\ &= \frac{dF_\xi(x)}{dx} \Big|_{x=\varphi^{-1}(y)} \cdot \frac{d\varphi^{-1}(y)}{dy} = f_\xi(\varphi^{-1}(y)) \cdot \frac{d\varphi^{-1}(y)}{dy}. \\ f_\eta(y) &= f_\xi(\varphi^{-1}(y)) \cdot \frac{d\varphi^{-1}(y)}{dy}. \end{aligned} \quad (9.2)$$

Рассмотрим другой частный случай – $\varphi(x)$ монотонно убывает на множестве значений СВ ξ (рис. 9.2).

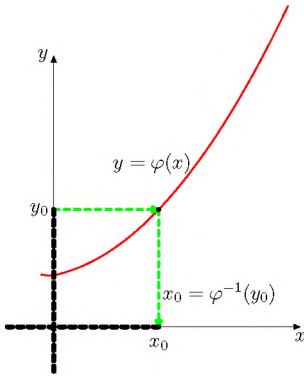


Рис. 9.1. Эскиз монотонно возрастающей функции

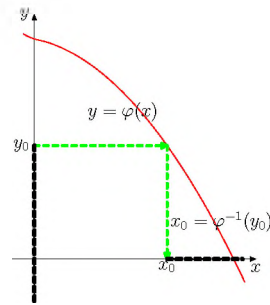


Рис. 9.2. Эскиз монотонно убывающей функции

Имеем

$$\{\eta < y_0\} = \{\xi > \varphi_0^{-1}(y_0)\}.$$

Воспользуемся известным соотношением для вероятностей взаимно противоположных событий A и \bar{A} :

$$F_\eta(y) = P(\eta < y) = P(\xi > \varphi_0^{-1}(y)) = 1 - P(\xi < \varphi_0^{-1}(y)) = 1 - F_\xi(\varphi_0^{-1}(y)).$$

Для плотности вероятности получим

$$f_\eta(y) = -f_\xi(\varphi^{-1}(y)) \cdot \frac{d\varphi^{-1}}{dy}. \quad (9.3)$$

Заметим, что $\frac{d\varphi^{-1}}{dy} < 0$. Для монотонной функции обе формулы можно объединить и применять с учетом свойств данной φ :

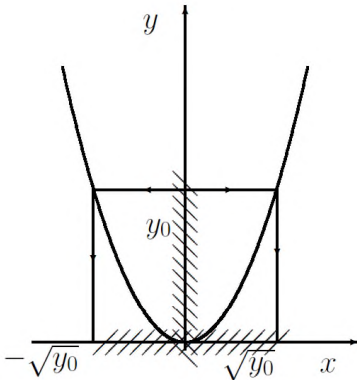
$$f_{\eta}(y) = f_{\xi}(\varphi^{-1}(y)) \cdot \left| \frac{d\varphi^{-1}}{dy} \right|. \quad (9.4)$$

Если φ — немонотонная функция, то можно разбить ООФ на промежутки монотонности и рассматривать отдельно каждый промежуток. Однако это связано с определенными математическими сложностями, и мы рекомендуем применять МФР с использованием геометрических свойств графика функции φ . Заметим, что похожий метод предлагают и другие авторы ([4, 7]).

Задача 13. Пусть НСВ ξ задана своей плотностью вероятности f_{ξ} , $\varphi(x) = x^2$. (Эта функциональная зависимость СВ будет использоваться далее.) Найти закон распределения СВ $\eta = \varphi(\xi)$.

Решение

Воспользуемся МФР. Построим график функции, на котором обозначим штриховкой события $\{0 < \eta < y_0\}$ и $\{0 < \eta < \sqrt{y_0}\}$.

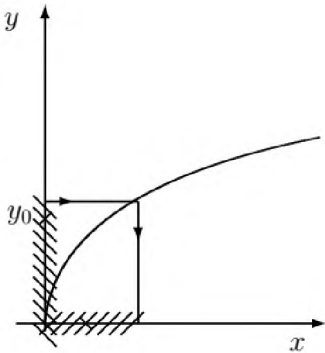


Очевидно, что отрицательных значений СВ η принимать не может. При $y_0 > 0$ прообразом $(0, y_0)$ будет $(-\sqrt{y_0}, \sqrt{y_0})$.

$$\begin{aligned} F_{\eta}(y) = f_{\eta}(y) &= 0 \text{ при } y < 0, \\ \text{а при } y \geq 0 \quad F_{\eta}(y) &= P(\eta < y) = \\ &= P(-\sqrt{y} < \xi < \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_{\xi}(x) dx. \end{aligned}$$

Итак,

$$f_{\eta}(y) = \begin{cases} \frac{f_{\xi}(\sqrt{y}) + f_{\xi}(-\sqrt{y})}{2\sqrt{y}}, & y > 0, \\ 0. & \end{cases} \quad (9.5)$$



Отсюда (9.5), кстати, легко выводится закон распределения и обратной функции, $\xi = \psi(\eta) = \sqrt{\eta}$:

$$f_{\xi}(x) = \begin{cases} 2 \cdot x \cdot f_{\eta}(x^2), & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (9.6)$$

9.3. Числовые характеристики функций

Вычисление моментов СВ $\eta = \varphi(\xi)$ и $\eta = \varphi(\xi_1, \xi_2, \dots, \xi_n)$ — матожидания и дисперсии можно проводить, не имея закона распределения η , а зная лишь распределение аргумента. Это чрезвычайно важное обстоятельство широко используется в приложениях теории вероятностей. Сейчас мы проведем упрощенный вывод соответствующей формулы.

Будем считать, что функция φ монотонно возрастает (по крайней мере на множестве возможных значений случайного аргумента ξ). Запишем определяющее соотношение для матожидания непрерывной СВ:

$$M[\eta] = \int_{-\infty}^{+\infty} y f_{\eta}(y) dy. \quad (9.7)$$

В рассматриваемом случае имеем

$$M[\eta] = \int_{-\infty}^{+\infty} y f_{\xi}(\varphi^{-1}(y)) \cdot \frac{d\varphi^{-1}(y)}{dy} dy.$$

Сделаем формальную замену переменных интегрирования $y = \varphi(x)$, т. е. перейдем к интегрированию по x . Как обычно, делаем замену в подынтегральной функции, в пределах интегрирования, и в дифференциале. Получим очень простое и удобное выражение для матожидания функции φ СВ ξ :

$$M[\varphi(\xi)] = \int_{-\infty}^{+\infty} \varphi(x) f_{\xi}(x) dx. \quad (9.8)$$

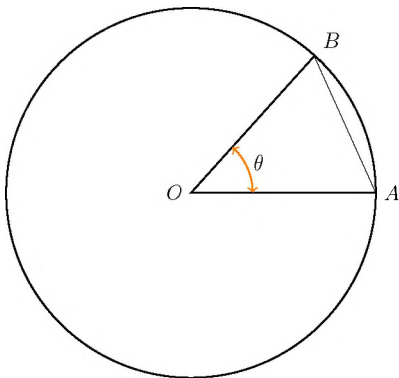
Использовать эту формулу можно и для немонотонной функциональной зависимости. Аналогичную формулу можно записать и для функции случайного вектора:

$$M[\varphi(\{\xi_1, \xi_2\})] = \iint_{-\infty}^{+\infty} \varphi(x, y) f_{\xi_1, \xi_2}(x, y) dS. \quad (9.9)$$

Рассмотрим несколько достаточно нетривиальных примеров нахождения числовых характеристик СВ.

Пример 13. На окружность радиуса r наудачу ставятся две точки, которые затем соединяются между собой и с центром окружности. Найти математическое ожидание и дисперсию площади полученного треугольника.

Решение



Учитывая симметрию задачи, можно считать, что первая точка имеет фиксированные полярные координаты $(r, 0)$. Положение же второй точки определяется проведенным из центра окружности лучом, образующим с первой стороной треугольника угол θ . Архаичный термин «наудачу» в современной терминологии подразумевает равномерное распределение значения угла в $[0, 2\pi]$.

Функциональное выражение для случайной величины S — площади треугольника получается в виде

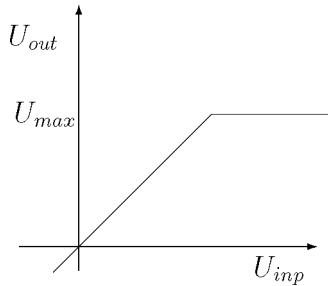
$$S = \frac{r^2}{2} |\sin \theta|.$$

По формуле (9.8) находим математическое ожидание площади:

$$M[S] = \frac{1}{2} r^2 \int_0^{2\pi} |\sin x| \frac{1}{2\pi} dx = \frac{r^2}{\pi}.$$

■

Пример 14. Напряжение, подаваемое на вход ограничителя, распределено по нормальному закону с параметрами m, σ . Найти числовые характеристики напряжения на выходе.

**Решение**

Вольтамперная характеристика ограничителя изображена на рисунке и имеет параметр U_{\max} . Моделировать зависимости между напряжениями можно следующей функцией:

$$U_{\text{out}} = \min\{U_{\text{inp}}, U_{\max}\}.$$

Отсюда матожидание напряжения на выходе равно:

$$M[U_{\text{out}}] = \int_{-\infty}^{+\infty} f_{\xi}(x) \min\{x, U_{\max}\} dx. \quad (9.10)$$

Здесь $f_{\xi}(x)$ — плотность вероятности напряжения на входе. Надо заметить, что подход к нахождению числовых характеристик по формулам (9.8) и (9.9) в данной задаче позволяет избежать нахождения плотности вероятности напряжения на выходе, являющейся обобщенной функцией. Нетрудно, хотя и громоздко, выполнить интегрирование в (9.10):

$$\begin{aligned} \int_{-\infty}^{+\infty} f_{\xi}(x) \min\{x, U_{\max}\} dx &= \int_{-\infty}^{U_{\max}} f_{\xi}(x)x dx + U_{\max} \cdot \int_{U_{\max}}^{+\infty} f_{\xi}(x) dx = \\ &= \int_{-\infty}^{U_{\max}} f_{\xi}(x)x dx + U_{\max} \cdot (1 - F_{\xi}(U_{\max})). \end{aligned}$$

Теперь учтем явный вид закона распределения входного напряжения:

$$M[U_{\text{out}}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{U_{\max}} x \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) dx + U_{\max} \cdot \frac{1}{\sigma\sqrt{2\pi}} \int_{U_{\max}}^{+\infty} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) dx.$$

Вычислить в обычном смысле такой интеграл невозможно, через элементарные функции он не выражается. Можно записать короче, поскольку фактически в правой части стоит функция F_{ξ} нормального распределения:

$$M[U_{\text{out}}] = U_{\max} + F_{\xi}(U_{\max}) \cdot (1 - U_{\max}). \quad (9.11)$$

Удобнее окончательный результат записать через табулированные функцию Лапласа $\Phi(x)$ и функцию φ :

$$M[U_{\text{out}}] = U_{\text{max}} - \sigma \cdot \left(\tau \left(\Phi(\tau) + 0.5 \right) + \varphi(\tau) \right), \quad \tau = \frac{U_{\text{max}} - m}{\sigma}.$$

Теперь найдем дисперсию, воспользовавшись известной формулой для ее вычисления: $D[U_{\text{out}}] = M[U_{\text{out}}^2] - M[U_{\text{out}}]^2$.

$$\begin{aligned} M[U_{\text{out}}^2] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{U_{\text{max}}} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx + U^2 \frac{1}{\sigma\sqrt{2\pi}} \int_{U_{\text{max}}}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \\ &= (m^2 + \sigma^2) (\Phi(\tau) + 0.5) + U_{\text{max}}^2 (0.5 - \Phi(\tau)) - \frac{2\sigma m + \sigma^2 \tau}{\sqrt{2\pi}} \exp\left(-\frac{\tau^2}{2}\right). \end{aligned}$$

$$\begin{aligned} D[U_{\text{out}}] &= \sigma^2 \left((1 + \tau^2) (\Phi(\tau) + 0.5) + \frac{\tau}{\sqrt{2\pi}} \exp\left(-\frac{\tau^2}{2}\right) - (\tau(\Phi(\tau) + 0.5) \right. \\ &\quad \left. + \frac{1}{2\pi} \exp\left(-\frac{\tau^2}{2}\right))^2 \right). \end{aligned}$$

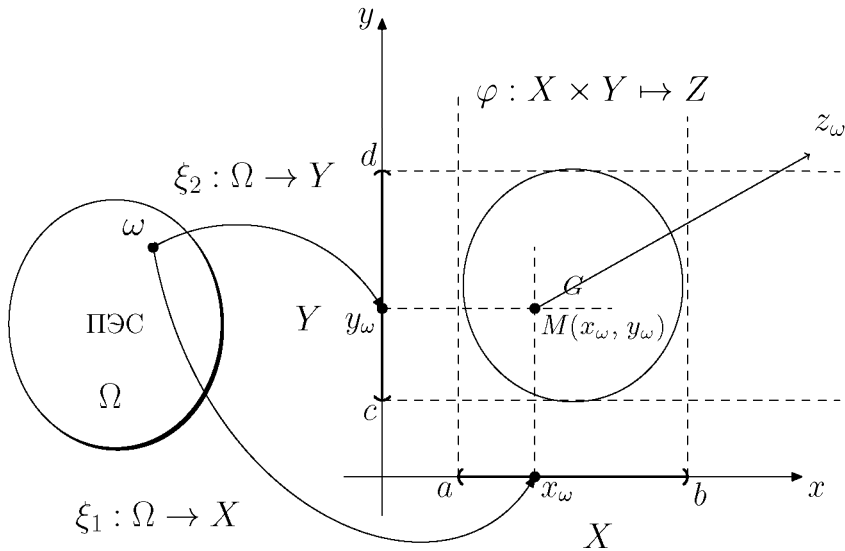
■

9.4. Закон распределения функции случайного вектора

Пусть на ПЭС Ω некоторого ЭСИ определены две СВ: $\xi_1 : \Omega \mapsto X = (a, b)$ и $\xi_2 : \Omega \mapsto Y = (a, b)$. Случайный вектор (ξ_1, ξ_2) имеет своими значениями множество упорядоченных пар действительных чисел $G \subset X \times Y \subset \mathbb{R}^2$. Определим функцию двух переменных $\varphi : X \times Y \rightarrow Z$. Суперпозиция

$$\Omega \xrightarrow{\xi_1, \xi_2} G \xrightarrow{\varphi} Z_0 \subset Z \quad (9.12)$$

дает новую случайную величину $\eta : \Omega \mapsto Z_0$. Принято записывать $\eta = \varphi(\xi_1, \xi_2)$.



Такая форма записи лишь задает правило, по которому вычисляются значения СВ η . Манипулировать явной записью формулы для φ можно лишь в очень ограниченном смысле.

Рассмотрим здесь только функцию двумерного случайного вектора, это не слишком громоздко, но позволяет обсудить все принципиальные отличия от функции одномерной случайной величины. И, конечно, не будем углубляться в теоретические тонкости, т. е. условия существования и свойства новой СВ, ограничимся примерами.

Поставим следующую общую задачу. Задан закон распределения случайного вектора с непрерывными составляющими (ξ, η) посредством двумерной плотности вероятности $f_{\xi, \eta}(x, y)$. Требуется найти закон распределения в виде плотности вероятности СВ $\zeta = \varphi(\xi, \eta)$, $f_{\zeta}(\cdot)$.

Как и в одномерном случае, начинаем с того, что определим функцию распределения для новой СВ.

$$F_{\zeta}(z) = P(\zeta < z) = \iint_{(G_z)} f_{\xi, \eta}(x, y) dx dy. \quad (9.13)$$

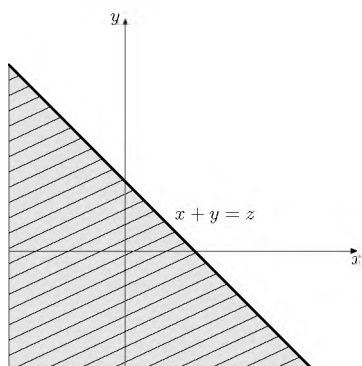
Здесь $G_z = \{(x, y) | \varphi(x, y) < z\}$. Конечно, $G_z \subseteq G$, т. е. подмножество множества значений вектора-аргумента, параметрически зависящее от z . Производная функции распределения и дает искомую плотность вероятности.

Глава 10

Распределение суммы, произведения и отношения двух случайных величин

Рассматривать проблему нахождения закона распределения $\zeta = \varphi(\xi, \eta)$ в общем виде бесполезно, разберем подробно несколько наиболее часто встречающихся на практике видов зависимостей φ .

1. Пусть $\varphi(x, y) = x + y$.



$F_{\zeta}(z) = \iint_{G_z} f_{\xi, \eta}(x, y) ds$. Область G_z определяется неравенством $x + y < z$, используя его геометрический смысл, нетрудно записать пределы повторного интеграла. Получаем

$$F_{\zeta}(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} dy f_{\xi, \eta}(x, y).$$

Продифференцировав F_{ζ} по z (предполагая выполнения достаточных для этого условий), получим плотность вероятности

$$f_{\zeta}(z) = \frac{dF_{\zeta}(z)}{dz} = \int_{-\infty}^{+\infty} dx f_{\xi, \eta}(x, z - x).$$

Представляет интерес случай суммы двух *независимых* СВ. Как известно, независимость составляющих случайного вектора приводит к факторизации двумерной плотности вероятности:

$$f_{\xi,\eta}(x, y) = f_{\xi} \cdot f_{\eta},$$

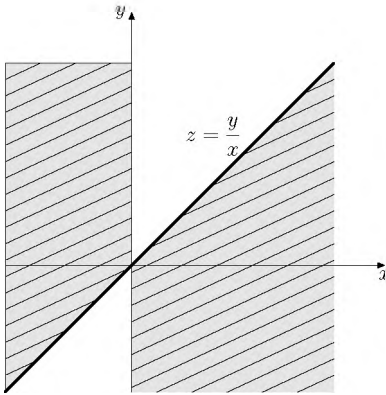
поэтому

$$f_{\zeta}(z) = \int_{-\infty}^{+\infty} f_{\xi}(x) \cdot f_{\eta}(z-x) dx. \quad (10.1)$$

Выражение, стоящее в правой части (10.1), называют **сверткой** функций $f_{\xi}(\cdot)$ и $f_{\eta}(\cdot)$, имеется специальное обозначение для свертки — $f_1 * f_2$. Свертка довольно часто встречается в различных математических приложениях, например, в теории сигналов и цепей. В данном случае свертка двух плотностей вероятности $f_{\xi}(\cdot)$ и $f_{\eta}(\cdot)$ дает плотность вероятности суммы независимых СВ ξ и η .

Нахождение распределения суммы независимых СВ называют **задачей композиции**, и ниже мы рассмотрим ее решение подробнее.

2. Следующая полезная функция $\varphi(x, y) = \frac{y}{x}$, $\zeta = \varphi(\xi, \eta)$. Область G_z определяется неравенством $\frac{y}{x} < z$.



Разберемся подробнее, как она выглядит. Пусть $z > 0$, тогда $y = z \cdot x$, $x \neq 0$ — граница области G . Это прямая линия. В правой полуплоскости неравенство $\frac{y}{x} < z$, при $z > 0$ выполняется ниже границы, в левой полуплоскости — выше границы. Если уменьшать z , то линия будет поворачиваться вокруг начала координат по часовой стрелке, при этом расположение точек области относительно границы останется прежним: в правой части — ниже линии, в левой — выше.

Расставляем пределы по правилу «от точки до точки, от линии до линии». Находим функцию распределения:

$$F_{\zeta}(z) = \int_{-\infty}^0 dx \int_{zx}^{+\infty} dy f_{\xi,\eta}(x, y) + \int_0^{+\infty} dx \int_{-\infty}^{zx} dy f_{\xi,\eta}(x, y). \quad (10.2)$$

Плотность вероятности получим, дифференцируя (10.2) по правилу Лейбница,

$$f_{\zeta}(z) = \int_{-\infty}^{+\infty} dx |x| f_{\xi, \eta}(x, zx). \quad (10.3)$$

Данное соотношение можно использовать при выводе формул некоторых полезных для матстатистики распределений. Надо заметить, что вид формулы (10.3) в разных источниках различный и зависит от того, какая СВ стоит в числителе отношения, а какая — в знаменателе.

3. Последняя из полезных для дальнейшего рассмотрения функций: $\varphi(x, y) = x \cdot y$, которая определчет СВ $\zeta = \varphi(\xi, \eta)$. Область G легко построить, используя вышеприведенные геометрические соображения и выкладки, аналогичные предыдущим. Результат:

$$f_{\zeta}(z) = \int_{-\infty}^{+\infty} f_{\xi}(x) \cdot f_{\eta}(z/x) \frac{1}{|x|} dx.$$

Другой вариант (и не единственный) см. в работе [11].

10.1. Задача композиции

Решим задачу композиции двух равномерных распределений, т. е. найдем распределение суммы независимых СВ $\xi_1 + \xi_2$. Пусть плотность вероятности СВ ξ_1 и ξ_2 одна и та же:

$$f_{\xi_i}(x) = \begin{cases} 1, & x \in [1, 2]; \\ 0 & x \notin [1, 2]. \end{cases}$$

Наша задача сводится к вычислению интеграла типа свертки:

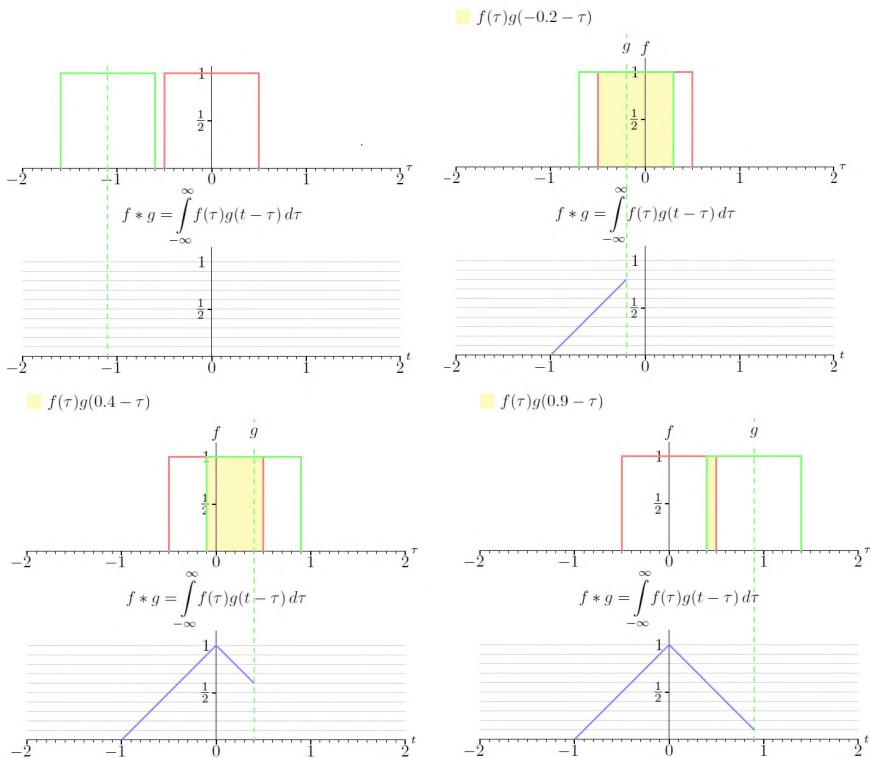
$$f_{\xi_1 + \xi_2}(y) = \int_{-\infty}^{+\infty} f_{\xi_1}(x) \cdot f_{\xi_2}(y - x) dx,$$

однако технически это сделать не так просто. Дело в том, что интегрируется произведение «склеенных» функций (англоязычный вариант похожий — *pieceswise*). Более строгое название — функции с ограниченным носителем. Последний термин применяют к области, в которой функция отлична от нуля.

Поэтому в зависимости от значения y подынтегральные выражения будут различные в том смысле, что только на пересечении носителей сомножителей подынтегральное выражение отлично от нуля.

Облегчим себе задачу следующим приемом, который можно назвать *мультипликацией*. Нарисуем графики функций $f_{\xi_1}(x)$ и $f_{\xi_2}(y-x)$. Первый будем называть «неподвижным» (относительно оси абсцисс), а второй, положение которого относительно оси абсцисс зависит от значения параметра y , — «подвижным». Этот факт известен еще из школьной математики: график функции $f_{\xi_2}(y-x) = f_{\xi_2}(-(x-y))$ отличается от графика $f_{\xi_2}(x)$ зеркальным отражением относительно оси ординат ($f(x) \rightarrow f(-x)$) и трансляцией вдоль ОХ ($f(-x) \rightarrow f(-(x-y))$). Последнее преобразование представляет собой параллельный перенос на y вправо, если $y > 0$, при $y < 0$ — влево на $|y|$. Изменение параметра y от $-\infty$ до $+\infty$ можно интерпретировать как движение этого графика слева направо.

Несколько характерных кадров этой «мультипликации» изображены ниже.



Теперь нетрудно сообразить, что в зависимости от значения параметра y свертка принимает вид

$$f_{\xi_1+\xi_2}(y) = \begin{cases} 0, & y \notin [2, 4]; \\ \int_{y-1}^y dx, & 2 \leq y \leq 3; \\ \int_{y-2}^{\frac{1}{2}} dx, & 3 < y \leq 4. \end{cases}$$

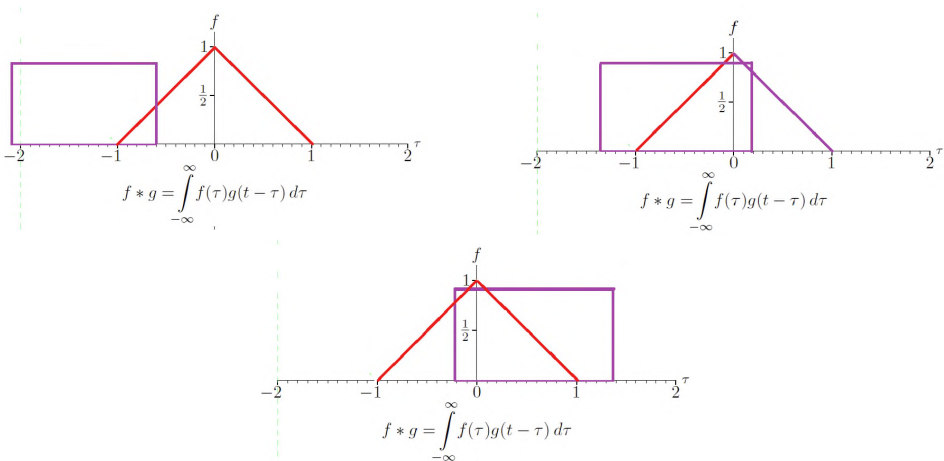
Окончательно получаем для плотности вероятности суммы двух равномерно распределенных СВ:

$$f_{\xi_1+\xi_2}(y) = \begin{cases} 0, & y \notin [2, 4]; \\ y - 2, & 2 \leq y \leq 3; \\ 4 - y, & 3 < y \leq 4. \end{cases}$$

Такое распределение называют *треугольным* или распределением Симпсона. Вы можете посмотреть полный видеоролик анимации этого подхода к вычислению свертки показательного и треугольного распределений по ссылке на QR [37].



Интерес представляет сумма трех и более независимых равномерно распределенных СВ. Для трех задача сводится к свертке исходной плотности вероятности и той функции, что получена только что (для двух слагаемых). «Неподвижным» можно считать график плотности вероятности распределения Симпсона. Рисуем опять кадры «мультипликации» и расставляем пределы интегрирования.



$$f_{\xi_1+\xi_2+\xi_3}(y) = \begin{cases} 0, & y \notin [3, 6]; \\ \int_2^{y-1} (x-2)dx, & 3 \leq y \leq 4; \\ \int_{y-2}^3 (x-2)dx + \int_3^{y-1} (4-x)dx, & 4 < y \leq 5; \\ \int_{y-2}^4 (4-x)dx, & 5 \leq y \leq 6. \end{cases}$$

Форма графика функции плотности состояний претерпела еще одно изменение, теперь линия гладко «склеена» из кусочков парабол.

$$f_{\xi_1+\xi_2+\xi_3}(x) = \begin{cases} 0, & y \notin [3, 6]; \\ \frac{1}{2}x^2 - 3x + \frac{9}{2}, & 3 \leq x \leq 4; \\ 9x - \frac{39}{2} - x^2, & 4 < x \leq 5; \\ 18 - 6x + \frac{1}{2}x^2, & 5 \leq x \leq 6. \end{cases}$$

Считать вручную, конечно, утомительно, но результат того стоит. Увеличим количество слагаемых, только считать сами не будем, доведем количество слагаемых до 10 с помощью математической программы Maple. Ни текст программы, ни сами выражения для полученных функций-сверток приводить не будем, покажем только графики. Итак, слева направо в соответствии с командой

```
> plot({f[2](x), f[4](x), f[6](x), f[10](x)}
> }, x=0..20);
```

приводятся графики плотностей вероятности для 2, 4, 6 и 10 слагаемых (рис. 10.1).

Заметим, что крайний правый график практически неотличим от кривой нормального распределения с соответствующими параметрами. Этот результат прекрасно объясняется в теории вероятностей центральной предельной теоремой, согласно которой распределение суммы большого числа независимых слагаемых мало отличается от нормального. С практической точки зрения полученный результат применяется при генерировании нормально распределенных псевдослучайных чисел. В качестве таковых можно брать сумму десяти последовательных равномерно распределенных чисел.

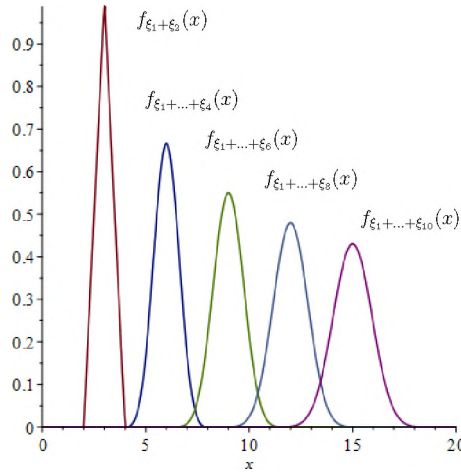


Рис. 10.1. Результат вычислений плотности вероятности суммы независимых СВ для различного числа слагаемых

10.2. Композиционная устойчивость нормального распределения

Сейчас будут рассмотрены важные для математической статистики законы распределения. При выводе соответствующих формул мы будем использовать результаты, полученные ранее (см. с. 92).

Нормальное распределение занимает особое место среди других распределений. В частности, нормальное распределение **композиционно устойчиво**. Иначе говоря, сумма двух независимых нормально распределенных слагаемых также имеет нормальное распределение. Можно добавить третье слагаемое, четвертое и т. д. Это очень важное свойство нормального распределения, которым мы воспользуемся ниже.

Пример 15. СВ ξ и η имеют нормальное распределение $N(a_1, \sigma_1)$ и $N(a_2, \sigma_2)$ соответственно. Решить задачу композиции.

Воспользуемся формулой

$$f_{\xi+\eta}(z) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} dx \exp\left(-\frac{(x-a_1)^2}{2\sigma_1^2}\right) \cdot \exp\left(-\frac{(z-x-a_2)^2}{2\sigma_2^2}\right).$$

Преобразуем выражение, стоящее в показателе экспоненты (естественно, предварительно обе экспоненты объединили):

$$x^2\sigma_2^2 - 2xa_1\sigma_2^2 - \dots + \sigma_1^2x^2 = (\sigma_1^2 + \sigma_2^2) \cdot \left(x - \frac{\sigma_1^2(z - a_2) + \sigma_2^2a_1^2}{\sigma_1^2 + \sigma_2^2} \right)^2 - \\ - \frac{(\sigma_1^2(z - a_2) + \sigma_2^2a_1^2)^2}{\sigma_1^2 + \sigma_2^2} + \sigma_2^2a_1^2 + (z - a_2)^2\sigma_1^2.$$

После замены переменной интегрирования

$$x \rightarrow x - \frac{\sigma_1^2(z - a_2) + \sigma_2^2a_1^2}{\sigma_1^2 + \sigma_2^2} \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2\sigma_2^2}}$$

получаем интеграл Пуассона, значение которого известно (см. с. 92):

$$\frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(a_1^2\sigma_2^2 + (z - a_2)^2\sigma_1^2 - \frac{\sigma_1^2(z - a_2) + \sigma_2^2a_1^2}{\sigma_1^2 + \sigma_2^2})^2}{2\sigma_1^2\sigma_2^2}\right) \cdot \\ \cdot \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(a_1 - z + a_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right).$$

Итак, плотность вероятности суммы двух нормально распределенных независимых слагаемых имеет вид:

$$f_{\xi+\eta}(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z - a)^2}{2\sigma^2}\right).$$

Получился нормальный закон распределения с параметрами $a = a_1 + a_2$ и $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$. Кстати, именно такими должны быть матожидание и дисперсия суммы в соответствии с их общими свойствами.

10.3. Распределения Стьюдента и χ^2

Сначала мы рассмотрим закон распределения, имеющий необычное обозначение χ^2 (рис. 10.2) (читается по-русски «хи-квадрат», а по-английски, как ни странно, [kaɪ skweɪ]¹). Это распределение широко используется в математической статистике. Соответствующую СВ обозначают χ_n^2 , где n — число степеней свободы (так называют этот параметр).

¹В 1990-е гг. был популярен редактор Chi-Writer, который попал в нашу страну неофициально. У нас его называли *Чирайтер*, это слово даже писалось(!) на обложках книг-пособий. А через некоторое время в Москву заехал автор, имя которого было Chi. Оказалось, звучит это имя как Кай!

Пример 16. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — система попарно независимых нормально распределенных СВ с параметрами $a = 0$ и $\sigma = 1$. Найти закон распределения СВ $\chi_n^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2$.

Решение

Закон распределения НСФ удобно задавать функцией $f_{\chi_n^2}$ (вы вполне можете сделать это самостоятельно!). Положим $n = 1$. Воспользуемся уже полученным результатом для распределения квадрата СВ (см. формулу (9.5) на с. 115).

$$F_{\chi_1^2}(x) = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{x}} \exp\left(-\frac{t^2}{2}\right) dt, x \geq 0;$$

$$f_{\chi_1^2}(x) = \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right), x \geq 0.$$

Положим $n = 2$, тогда

$$f_{\chi_2^2}(z) = \int_{-\infty}^{+\infty} dx f_{\chi_1^2}(x) f_{\chi_1^2}(z-x) = C \cdot z^{\frac{2}{2}-1} \cdot e^{-\frac{z}{2}}.$$

Есть основания предположить², что для произвольного n закон распределения будет иметь вид:

$$f_{\chi_n^2}(z) = C_n \cdot z^{\frac{n}{2}-1} \cdot e^{-\frac{z}{2}}, z \geq 0.$$

Конечно же, при $z < 0$ $f_{\chi_n^2}(z) = 0$.

Докажем по индукции:

$$\begin{aligned} f_{\chi_{n+1}^2} &= \int_0^x f_{\chi_n^2}(t) f_{\chi_1^2}(x-t) = \frac{C_n}{\sqrt{2\pi}} \int_0^x t^{\frac{n}{2}-1} e^{-\frac{t}{2}} (x-t)^{-\frac{1}{2}} e^{-\frac{x-t}{2}} dt = \\ &= \frac{C_n}{\sqrt{2\pi}} e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1} \left(\int_0^1 u^{\frac{n}{2}-1} (1-u)^{-1/2} du \right), (x \geq 0). \end{aligned} \quad (10.4)$$

²Для этого надо бы еще один частный случай рассмотреть, $n = 3$.

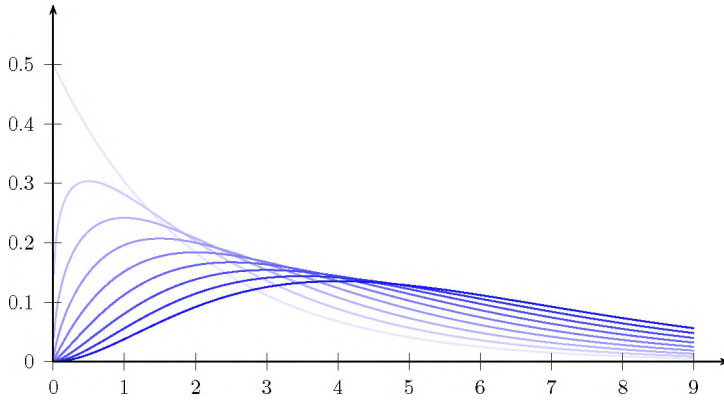


Рис. 10.2. Графики распределения χ^2 при $n = 2 \div 10$

Найдем, наконец, константу C_n :

$$\int_0^{\infty} C_n x^{-\frac{n}{2}-1} e^{-\frac{x}{2}} dx = 1 \Rightarrow C_n = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})}.$$

Итак, искомая плотность вероятности

$$f_{\chi_n^2}(x) = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x \geq 0.$$

Справедливости ради отметим, что эта функция табулирована, в любом математическом или статистическом программном пакете она присутствует.

Еще одно хорошо известное распределение — **распределение Стьюдента** (рис. 10.3). Под таким оригинальным псевдонимом публиковал свои статьи математик Уильям Д. Госсет, работавший на пивоваренном(!) заводе Гиннеса. Это распределение СВ $\zeta_n = \frac{\xi}{\sqrt{\chi_n^2}}$, где ξ имеет распределение $N[0,1]$, а χ_n^2 — распределение хи-квадрат.

Сначала найдем распределение СВ ζ_n^2 , используя уже полученные результаты для распределения хи-квадрат и распределения отношения двух СВ:

$$f_{\zeta_n^2} = \int_0^{\infty} z f_{\chi_n^2}(z) f_{\xi}(zx) dx = \frac{C_n}{\sqrt{2\pi}} \int_0^{\infty} t^{\frac{n}{2}} e^{-\frac{t}{2}} e^{-\frac{z^2 t}{2}} \frac{dt}{\sqrt{zt}}.$$

Получим

$$f_{\zeta_n^2}(z) = K_n \frac{(z+1)^{\frac{n+2}{2}}}{\sqrt{z}}.$$

А теперь можно заняться и искомым законом распределения ζ_n :

$$f_{\zeta_n}(x) = |x| f_{\zeta_n^2}(x^2) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi} \Gamma(\frac{n}{2})} (x^2 + 1)^{-\frac{n+1}{2}}, x \in \mathbb{R}.$$

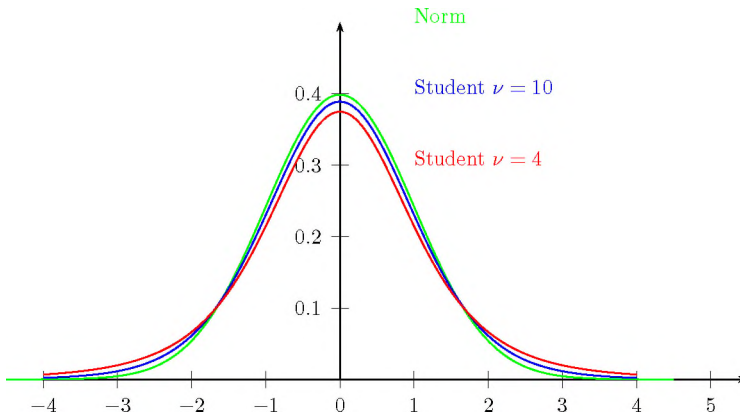


Рис. 10.3. Графики распределения Стьюдента

10.4. Распределение Рэлея

Первоначально это распределение ввел лорд Рэлей в 1880 г. при рассмотрении огибающей суммы ряда гармонических колебаний разной частоты. Оно также встречается при пристрелке пушек, ракет и другого огнестрельного оружия, если разбросы (отклонения от цели) в каждом из двух взаимно перпендикулярных направлений независимы и распределены по нормальному закону.

Таким образом, если начало прямоугольной системы координат считать целью, а разброс по осям прямоугольной системы координат обозначить через X и Y , то промах (отклонение от цели) будет выглядеть как СВ

$R = (X^2 + Y^2)^{1/2}$. График f_R для различных значений дисперсии σ^2 показан на рис. 10.4.

Получим явный аналитический вид этой функции. Если X и Y — независимые гауссовские случайные величины (нормальное распределение) с нулевыми математическими ожиданиями и одинаковыми дисперсиями σ^2 , то плотность вероятностей для R будет записываться в виде

$$f_R(r) = \begin{cases} r/\sigma^2 \exp(-r^2/2\sigma^2), & r \geq 0; \\ 0, & r < 0. \end{cases} \quad (10.5)$$

Это и есть *рэлеевская* плотность распределения вероятностей. Выкладку при желании можно провести и самостоятельно, все необходимые промежуточные вычисления уже рассматривались — и сумма квадратов (см. с. 129), и квадратный корень (см. с. 116). Обратите внимание на то, что максимум этой функции соответствует стандартному отклонению (СКО σ) и что она несимметрична относительно этого значения. Последнее означает, что малые отклонения от цели вероятнее больших. Математическое ожидание случайной величины, распределенной по закону Рэлея, легко определяется и равно:

$$M[R] = \int_0^{+\infty} r f_R(r) dr = \int_0^{+\infty} \left(\frac{r^2}{\sigma^2} \right) \exp\left(-\frac{r^2}{2\sigma^2}\right) dr = \sqrt{\pi/2}\sigma, \quad (10.6)$$

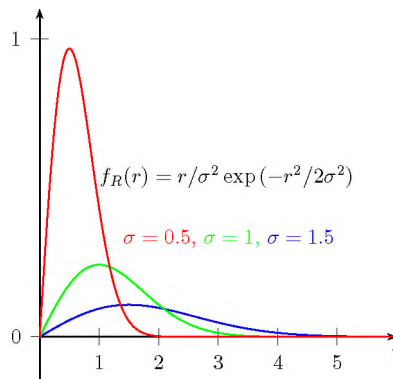


Рис. 10.4. Плотности вероятности распределения Рэлея

а средний квадрат имеет вид:

$$M[R^2] = \int_0^{+\infty} r^2 f_R(r) dr = \int_0^{+\infty} (r^2/\sigma^2) \exp(-r^2/2\sigma^2) r dr = 2\sigma^2. \quad (10.7)$$

При этом дисперсия случайной величины R равна:

$$\sigma_R^2 = M[R^2] - m_R^2 = \left(2 - \frac{\pi}{2}\right) \sigma^2.$$

Обратите внимание, что полученное значение дисперсии отличается от дисперсии σ^2 гауссовских случайных величин, из которых «составлена» рассматриваемая рэлеевская величина. В отличие от гауссовских случайных величин, для случайной величины, распределенной по закону Рэлея, и математическое ожидание и дисперсия зависят от одного и того же параметра σ^2 , в результате чего они не могут изменяться независимо друг от друга.

Функция распределения вероятностей для рэлеевской величины находится непосредственно из соответствующей плотности вероятностей, которая легко интегрируется:

$$F_R(r) = \begin{cases} 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right), & r \geq 0; \\ 0, & r < 0. \end{cases} \quad (10.8)$$

Глава 11

Предельные теоремы теории вероятностей

11.1. Сходимость по вероятности. Неравенство Чебышёва

В теории пределов числовых последовательностей ключевым было понятие *сходимости*. Строгое определение содержало неравенство $|x_k - A| < \varepsilon$. Если рассматривать систему (последовательность) СВ, то аналогичное неравенство $|\xi_n - A| < \varepsilon$ становится случайным событием, которое может не произойти при некоторых комбинациях значений n и ε . Последнее зависит от вида законов распределения СВ. Математики предложили и используют в ТВ специфические виды сходимости: *по вероятности*, «*в среднем*», «*почти наверное*». Мы ограничимся первым из пересчисленных.

Сходимость *по вероятности* последовательности СВ означает, что

$$\lim_{n \rightarrow \infty} P(|\xi_n - \xi| < \varepsilon) = 1.$$

Можно предложить следующую интерпретацию сходимости по вероятности. При увеличении числа испытаний большие отклонения от предельного значения становятся относительно реже.

Мы изложим здесь замечательное утверждение, принадлежащее одному из крупнейших математиков прошлого века П. Л. Чебышёву (1821—1894). Оно интересно не только широтой формулировки, но и исключительно простой идеей доказательства. В основе последующих рассуждений лежит неравенство, обнаруженное П. Л. Чебышёвым и позднее нашедшее многочисленные применения как в теории вероятностей, так и в других математических дисциплинах.

Теорема 8. Если случайная величина ξ имеет конечную дисперсию, то при любом положительном ε имеет место неравенство

$$P\left(|\xi - M[\xi]| \geq \varepsilon\right) \leq \frac{D[\xi]}{\varepsilon^2}.$$

Доказательство

Пусть x_i — возможные значения величины ξ и p_i — их вероятности. По определению

$$D[\xi] = \sum_i (x_i - m_\xi)^2 p_i.$$

Если из этой суммы мы выбросим все слагаемые, для которых $|x_i - m_\xi| < \varepsilon$, т. е. сохраним лишь те, для которых $|x_i - m_\xi| \geq \varepsilon$, то от этого сумма может только уменьшиться. Следовательно,

$$D[\xi] \geq \sum_{|x_i - m_\xi| \geq \varepsilon} (x_i - m_\xi)^2 p_i.$$

Сумму, стоящую справа, можно уменьшить еще более, если все множители $(x_i - m_\xi)^2$ заменить допустимым для них минимумом ε^2 :

$$D[\xi] \geq \varepsilon^2 \sum_{|x_i - m_\xi| \geq \varepsilon} p_i = \varepsilon^2 P\left(|\xi - M[\xi]| \geq \varepsilon\right).$$

Отсюда и следует искомое неравенство Чебышёва:

$$P(|\xi - m_\xi| > \varepsilon) \leq \frac{D[\xi]}{\varepsilon^2}.$$

□

Доказательство неравенства можно провести и для НСВ с плотностью вероятности f_ξ (предполагая существование как m_ξ , так и $D[\xi]$). Последнее накладывает на функцию f_ξ требование достаточно быстрого убывания при $|x| \rightarrow \infty$.

Пусть ξ — СВ непрерывного типа, имеющая ограниченную дисперсию. Это, в частности, означает, что плотность вероятности $f_\xi(x)$ убывает на бесконечности достаточно быстро для обеспечения сходимости не только интеграла $\int_{-\infty}^{\infty} f_\xi(x) dx$, но и $\int_{-\infty}^{\infty} (x - m_\xi)^2 f_\xi(x) dx$.

Выделим окрестность математического ожидания:

$$S_\varepsilon(m_\xi) = (m_\xi - \varepsilon, m_\xi + \varepsilon).$$

Используем определение дисперсии и несколько очевидных неравенств, аналогичных приведенным выше для ДСВ.

$$\begin{aligned} D[\xi] &= \int_{-\infty}^{+\infty} (x - m_\xi)^2 f_\xi(x) dx \geq \int_{x \notin S_\varepsilon} (x - m_\xi)^2 f_\xi(x) dx \geq \\ &\geq \varepsilon^2 \int_{x \notin S_\varepsilon} f_\xi(x) dx = \varepsilon^2 P(|\xi - m_\xi| > \varepsilon), P(|\xi - m_\xi| > \varepsilon) \leq \frac{D[\xi]}{\varepsilon^2}. \end{aligned}$$

11.2. Закон больших чисел в форме Бернулли и в форме Чебышёва

Законом больших чисел (ЗБЧ) в теории вероятностей называют ряд утверждений относительно предельного поведения числовых характеристик суммы большого числа СВ. Сформулируем и докажем два утверждения, известных как ЗБЧ в форме Бернулли и ЗБЧ в форме Чебышёва. Наиболее просто доказывается и имеет наглядную интерпретацию теорема Бернулли о законе больших чисел. Докажем ее.

Теорема 9. *Относительная частота события в схеме Бернулли сходится по вероятности к значению p .*

Доказательство

Ранее мы уже вычисляли матожидание и дисперсию распределения Бернулли: $m = np$, $\sigma^2 = np(1 - p)$. Для относительной частоты будет соответственно

$$M\left[\frac{m}{n}\right] = p, D\left[\frac{m}{n}\right] = \frac{pq}{n}.$$

Воспользуемся неравенством Чебышёва:

$$\begin{aligned} P\left(\left|\frac{m}{n} - p\right| \geq \varepsilon\right) &\leq \frac{D[m/n]}{\varepsilon^2} = \frac{pq}{n\varepsilon^2}, \\ 1 \geq P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) &< 1 - \frac{pq}{n\varepsilon^2}. \end{aligned}$$

Перейдем к пределу и получим искомое:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1.$$

□

Теорема Чебышёва о ЗБЧ имеет более строгую формулировку и доказательство ее сложнее.

Теорема 10. Пусть имеется последовательность попарно независимых случайных величин $\{\xi_1, \xi_2, \dots, \xi_n\}$ с математическими ожиданиями $M[\xi_k]$ и ограниченными в совокупности дисперсиями $D[\xi_k] : D[\xi_k] \leq C, \forall k$.

Тогда среднее арифметическое системы СВ сходится по вероятности к среднему арифметическому матожиданий.

Доказательство

По свойствам матожидания

$$M\left[\frac{1}{n} \sum_i \xi_k\right] = \frac{1}{n} \sum_i M[\xi_k].$$

Воспользуемся неравенством Чебышёва и условием попарной независимости СВ:

$$D\left[\frac{1}{n} \sum_i \xi_k\right] = \frac{1}{n^2} \sum_i D[\xi_k] \leq \frac{C \cdot n}{n^2} = \frac{C}{n},$$

$$P\left(\left|\frac{1}{n} \sum_i \xi_k - \frac{1}{n} \sum_i M[\xi_k]\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} D\left[\frac{1}{n} \sum_i \xi_k\right] \leq \frac{C}{n^2 \varepsilon^2} \rightarrow 0.$$

Для противоположного события знак неравенства в вероятности меняем на противоположный и учитываем $P(A) + P(\bar{A}) = 1$.

Получаем

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_i \xi_k - \frac{1}{n} \sum_i M[\xi_k]\right| < \varepsilon\right) = 1.$$

А это и означает справедливость доказываемого утверждения. В случае же равенства всех матожиданий величине m будем иметь

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_i \xi_k - m\right| < \varepsilon\right) = 1.$$

□

Следствие. Если в условиях теоремы СВ имеют равные матожидания (что представляет интерес для статистики), то среднее арифметическое системы СВ сходится по вероятности к матожиданию.

ЗБЧ в форме Бернулли имеет простую житейскую интерпретацию: если какое-то событие может произойти ($p > 0$), то оно рано или поздно произойдет.

Полученные выше неравенства и оценки используются довольно широко. В частности, при оценке степени опасности некоторых воздействий на человеческий организм учитывается ограниченность продолжительности жизни. Если время, которое необходимо для патологического изменения в организме при данной концентрации вещества в окружающей среде, существенно превосходит продолжительность жизни, то такая концентрация считается допустимой.

Пример 17. Покажем, что вероятность того, что при 300 бросках частота выпадения шестерки отклонится от p не более чем на 0.01.

Решение

Воспользуемся полученным неравенством Чебышёва:

$$P\left(\left|\frac{m}{n} - \frac{1}{6}\right| \geq 0.01\right) \geq 1 - \frac{5}{6 \cdot 6 \cdot 300 \cdot 10^{-4}}.$$

Результат тривиальный. ■

11.3. Формулы Лапласа. Понятие о центральной предельной теореме

Важность и вездесущность нормального распределения уже упоминались неоднократно ранее. Мы использовали это распределение для нахождения вероятности $P_n(k)$ в схеме Бернулли по формулам Лапласа (см. с. 72). Качественным обоснованием было сравнение изображений на одном графике (см. рис. 4.1) кривой Гаусса и точек $(k, P_n(k))$. Пожалуй, наиболее впечатляющим был результат компьютерного эксперимента по сложению десяти равномерно распределенных СВ (см. рис. 10.1, с. 127). Было обнаружено, что график плотности вероятности практически неотличим от кривой Гаусса.

Это обстоятельство не случайно. В истории ТВ известно большое количество теорем, в которых при различных ограничениях доказывалось, что **суммы большого числа СВ имеют распределение, близкое к нормальному**. Формулировки этих теорем достаточно далеко выходят за пределы программы нашего учебного курса (например, [7, с. 152]). Отметим значительный вклад в изучение этого вопроса нашего соотечественника великого математика А. М. Ляпунова, который и ввел в обиход название «центральная предельная теорема» (ЦПТ). Программы для технических специальностей предусматривают знакомство только с *понятием о ЦПТ*. Основной

теоретический факт состоит в том, что сумма большого количества случайных величин имеет приближенно нормальное распределение при условиях: а) независимости или слабой зависимости между ними, б) пренебрежимости каждой из них в совокупности. Последнее условие означает малость вклада в результирующее значение суммы.

С проявлением закономерности случайных явлений, отраженных в ЦПТ, люди встречаются постоянно: рост и вес представителей одной народности имеют нормальное распределение; министерство обороны заказывает военную форму и обувь в соответствии с нормальным распределением количества призывников с одинаковыми размерами одежды и обуви. Параметры любого изделия отличаются от проектных, это принципиально неустранимо. На отклонение диаметра вала от расчетного в процессе изготовления влияет и неуравновешенность заготовки, положение ее в зажиме патрона станка, изношенность резца и многое другое. Каждый фактор дает небольшой независимый от других вклад, но суммарная погрешность будет иметь нормальное распределение. В первой главе мы указывали на разброс снарядов при стрельбе, здесь также работает масса факторов, которые в сумме дают случайное нормально распределенное отклонение точки попадания от цели. В разделах по матстатистике мы будем широко использовать следствия из ЦПТ.

Полезным проявлением свойств нормального распределения являются *локальная* и *интегральная* формулы Лапласа.

Приведем их без строгого доказательства, используя правдоподобные рассуждения по Пойа (plausible reasoning [32]). Обе формулы имеют прямое отношение к распределению Бернулли. Дело в том, что число успехов в схеме Бернулли можно рассматривать как сумму значений индикаторных СВ, относящихся к каждому испытанию серии. Согласно ЦПТ, при больших n эту сумму можно с хорошей точностью интерпретировать нормально распределенной СВ ξ . Параметры этого распределения легко найти, ведь это МО и СКО ($\sigma = \sqrt{D[\xi]}$).

Поскольку числовые характеристики каждого индикатора-слагаемого нам известны (см. с. 82), а операторы МО и дисперсии в данном случае линейны, то мы получаем

$$m_{\xi} = \sum p = np, \quad D[\xi] = \sum pq = npq.$$

Все слагаемые каждой суммы одинаковые, поэтому суммирование сводится просто к умножению на n — количество слагаемых.

Итак, вероятности числа успехов в схеме Бернулли при больших n ¹ можно довольно точно вычислять через табулированную функцию

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

Для этого применяем **локальную** формулу Лапласа:

$$\sqrt{npq} \cdot p_n(k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x = \frac{k - np}{\sqrt{npq}}.$$

Если же необходимо вычислить $P(k_1 \leq k \leq k_2)$, т. е. просуммировать $P_n(k)$, то прибегаем к интегрированию, как это и делается для НСВ. Как мы уже отмечали на с. 118, фактическое интегрирование не проводится — используется **функция Лапласа**:

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

Интегральная формула Лапласа:

$$P(k_1 \leq k \leq k_2) = \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right).$$

¹Конечно, строго говоря, нужна оценка погрешности, но в учебных задачах мы ограничимся условием $n \sim 100$ и более.

Глава 12

Случайные функции

12.1. Основные понятия

Пусть Ω — пространство случайных событий и $T \subset \mathbb{R}$.

Определение. Если по некоторому правилу каждой паре (ω, t) , $\omega \in \Omega$ и $t \in T$, сопоставляется некоторое число $x \in \mathbb{R}$, то говорят, что задана *случайная функция* (СФ).

В приложениях чаще всего роль параметра t играет время, видимо, поэтому вместо термина «случайная функция» используется термин «случайный процесс».

Для каждого фиксированного значения параметра t имеем СВ $\xi(t)$, называемую *сечением* СФ. Если же зафиксировать элемент пространства событий ω , то получим обычную функцию переменной t , называемую *реализацией* СФ. Иначе ее называют выборочной функцией, а для случайных процессов — также и траекторией.

Еще в школе вы познакомились со способами задания «обычных» функций: аналитическим, графическим и табличным. А как задать случайную функцию? В теории существует два подхода.

Можно рассматривать реализации как элементы некоторого функционального пространства L . СФ — это отображение $\Omega \rightarrow L$. Получается некоторая аналогия со случайными величинами, только теперь каждому $\omega \in \Omega$ сопоставляется не **число**, а **функция** из L , и нужно каким-то образом задать распределение вероятности в L . Такой подход очень сложен и не представляет для нас практического интереса.

Перейдем к другому способу описания СФ. Наша задача — найти вероятность того, что СФ принимает значения вблизи $(x_1^0, x_2^0, x_3^0, \dots, x_n^0)$ в моменты «времени» $t_1, t_2, \dots, t_n, n \in \mathbb{N}$.

Пусть для определенности каждое сечение СФ $\xi(t)$ с данным значением параметра t — это непрерывная СВ с плотностью вероятности $f_1(x, t)$. Таким образом, задана вероятность того, что СВ принимает в момент «времени» t значения из некоторой малой окрестности точки x_0 , $\Delta p = f_1(x_0, t) \cdot \Delta x$.

Упорядоченная пара сечений $(\xi(t_1), \xi(t_2))$, $t_1, t_2 \in T$ может рассматриваться как двумерный случайный вектор с законом распределения в форме плотности вероятности $f_2(x_1, x_2; t_1, t_2)$. Можно найти вероятность того, что этот случайный вектор принимает значение из двумерной окрестности точки (x_1^0, x_2^0) , $\Delta p_2 = f_2(x_1, x_2; t_1, t_2) \Delta x_1 \Delta x_2$ (функция принимает значения вблизи (x_1^0, x_2^0) в моменты «времени» t_1 и t_2). Можно рассматривать любую упорядоченную совокупность сечений $(\xi_1(t_1), \xi_2(t_2), \dots, \xi_n(t_n))$ при любых $t_1, t_2, \dots, t_n \in T$ как случайный вектор, заданный n -мерной плотностью вероятности $f_n(\{x_n\}; \{t_n\})$.

В случае конечного множества T мы придем к многомерному случайному вектору. Если же множество значений параметра t бесконечное, то полное описание случайной функции $\xi(t)$ требует и бесконечной иерархии функций вида $f_n(\{x_n\}; \{t_n\})$. Употребление слова *иерархия* подчеркивает тот факт, что все «нижестоящие» функции порядка $k < n$ могут быть получены из функции порядка n последовательным интегрированием по «иксовым» переменным

$$\begin{aligned} f_{n-1}(x_1, x_2, \dots, x_{n-1}; t_1, t_2, \dots, t_{n-1}) &= \\ &= \int_{-\infty}^{+\infty} f_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_n. \end{aligned}$$

В инженерных приложениях теории исследователи используют те или иные приближенные подходы: либо многомерные функции распределения «расщепляются» ими на произведения двух- или трехмерных, либо авторы ограничиваются моментами низких порядков. В некоторых случаях функциональный вид реализаций определяется природой изучаемого объекта, а случайность сводится к параметрической зависимости от одной или нескольких СВ.

12.2. Случайные функции. Корреляционная теория

Рассмотрим один из подходов к теории СФ, наиболее важный для инженерных приложений, — *корреляционную теорию*.

Будем описывать СФ несколькими основными числовыми характеристиками, имеющими вполне ясный и наглядный смысл.

Математическое ожидание определяется для каждого сечения обычным образом и является функцией параметра t .

$$M[\xi(t)] = m_\xi(t) = \int_{-\infty}^{+\infty} x \cdot f_1(x, t) dt. \quad (12.1)$$

Эту характеристику можно рассматривать как некоторую среднюю реализацию СФ. Разброс относительно математического ожидания, как обычно, характеризуется *дисперсией* СФ:

$$D[\xi(t)] = M[(\xi(t) - m_\xi(t))^2]. \quad (12.2)$$

Третья числовая характеристика — *корреляционная функция*, $K_\xi(t_1, t_2)$,

$$K_\xi(t_1, t_2) = M[\overset{\circ}{\xi}(t_1), \overset{\circ}{\xi}(t_2)], \quad \overset{\circ}{\xi}(t) = \xi(t) - m_\xi(t). \quad (12.3)$$

Она описывает зависимость между сечениями СФ, соответствующими различным значениям t_1 и t_2 .

Приведем основные свойства математического ожидания, дисперсии и корреляционной функции:

- $D[\xi(t)] = K_\xi(t, t)$ — это очевидное следствие определений (12.2) и (12.3);
- прибавим к СФ $\xi(t)$ неслучайную функцию $\varphi(t)$. Это просто означает, что к каждому сечению прибавили неслучайное слагаемое, зависящее от t :

$$M[\eta(t)] = M[\xi(t) + \varphi(t)] = m_\xi(t) + \varphi(t).$$

Использовано известное свойство математического ожидания: $M[\varphi(t)] = \varphi(t)$.

Посмотрим, как изменится корреляционная функция:

$$K_\eta(t_1, t_2) = M[\overset{\circ}{\eta}(t_1) \cdot \overset{\circ}{\eta}(t_2)] = K_\xi(t_1, t_2), \quad (12.4)$$

$$\overset{\circ}{\eta}(t) = \eta(t) - m_\eta(t) = \xi(t) + \varphi(t) - m_\xi(t) - \varphi(t) = \overset{\circ}{\xi}(t). \quad (12.5)$$

Это означает, что прибавление к СФ неслучайной функции не изменяет корреляционной функции, а следовательно, и дисперсии $D[\xi(t)]$:

$$D[\xi(t) + \varphi(t)] = D[\xi(t)];$$

- умножим СФ на неслучайную функцию $\varphi(t)$:

$$\begin{aligned}\eta(t) &= \varphi(t) \cdot \xi(t), \\ M[\eta(t)] &= \varphi(t) \cdot M[\xi(t)] = \varphi(t) \cdot m_\xi(t); \\ K_\eta(t_1, t_2) &= M[(\eta(t_1) - m_\eta(t_1)) \cdot (\eta(t_2) - m_\eta(t_2))] = \\ &= \varphi(t_1)\varphi(t_2)M[(\xi(t_1) - m_\xi(t_1)) (\xi(t_2) - m_\xi(t_2))] = \varphi(t_1)\varphi(t_2)K_\xi(t_1, t_2);\end{aligned}$$

- свойство симметрии: $K_\xi(t_2, t_1) = K(t_1, t_2)$;
- ограниченность корреляционной функции:

$$|K_\xi(t_1, t_2)| \leq D[\xi(t_1)]D[\xi(t_2)].$$

Заметим, что перечисленные числовые характеристики наиболее важны для практики, поскольку могут быть измерены. А коль скоро различие свойств двух случайных процессов не сказывается на корреляционной функции и матожидании, то они практически оказываются неразличимы.

12.3. Дифференцирование и интегрирование случайных функций

В фундаменте классического математического анализа лежит понятие сходимости числовой последовательности. Как вы, наверное, помните, исходя из него можно определить предел функции, производную, определенный интеграл. Для СВ обычное понятие сходимости неприменимо, так как нельзя гарантировать выполнение неравенства $|\xi_n - A| < \epsilon$, поскольку в левой части стоит случайная величина. В теории вероятностей мы встречали понятие *сходимость по вероятности*. Напомним, что последовательность СВ $\{\xi_n\}_{n=1}^\infty$ сходится по вероятности к A , если

$$\lim_{n \rightarrow \infty} P(|\xi_n - A| < \epsilon) = 0, \quad \forall \epsilon > 0, \quad \text{обозначение: } \xi_n \xrightarrow{P} A.$$

Можно разъяснить это свойство так: по мере увеличения числа испытаний n большие отклонения СВ ξ_n от предела A становятся относительно реже.

Определим еще одно понятие сходимости для СВ. (Да, не удивляйтесь, для СВ не существует какого-то одного «естественного» понятия сходимости!)

Определение. Последовательность СВ $\{\xi_n\}_{n=1}^\infty$ сходится в среднеквадратичном к СВ ξ , если

$$\lim_{n \rightarrow \infty} M[(\xi_n - \xi)^2] = 0.$$

Общепринятое обозначение

$$l.i.m.\xi_n = \xi \quad (\text{limit in the mean square}).$$

Можно доказать, что из сходимости в среднеквадратичном следует сходимость по вероятности.

Аналогично определяется и сходимость (предел в точке) для функции.

Определение. Говорят, что случайная функция $\xi(t)$ сходится в среднеквадратичном (далее в среднем) при $t \rightarrow t_0$ к случайной величине ξ_0 , если

$$\lim_{t \rightarrow t_0} M[\xi(t) - \xi_0]^2 = 0.$$

Краткое обозначение:

$$l.i.m.\xi(t) = \xi_0.$$

С формальной точки зрения ни одно из определений сходимости не является предпочтительным. В корреляционной теории применяется понятие сходимости в среднем, что объясняется физическим смыслом (подробнее см. ниже).

На основании понятия **сходимость в среднем** можно определить и непрерывность СФ (в среднем!), если

$$l.i.m.\Delta t \rightarrow 0 (\xi(t + \Delta t) - \xi(t)) = 0,$$

и производную СФ:

$$\xi'(t) = l.i.m.\Delta t \rightarrow 0 \frac{\xi(t + \Delta t) - \xi(t)}{\Delta t},$$

которая сама является СФ.

В рамках корреляционной теории нас интересуют матожидание и корреляционная функция. Приведем без доказательства следующие утверждения:

1) для непрерывности в среднем СФ $\xi(t)$, $t \in (a, b)$, необходимо и достаточно, чтобы $m_\xi(t)$ была непрерывна на (a, b) , а $K_\xi(t_1, t_2)$ непрерывна при $t_1 = t_2 = t \in (a, b)$;

2) для существования производной в среднем СФ $\xi(t)$ на интервале (a, b) необходимо и достаточно, чтобы $m_\xi(t)$ была дифференцируема на (a, b) и существовала смешанная производная $\frac{\partial^2 K_\xi}{\partial t_1 \partial t_2}$, $t_1 = t_2 = t \in (a, b)$;

3) матожидание производной в среднем СФ $\xi(t)$ равно производной от матожидания, т. е. $M[\xi'(t)] = \frac{d}{dt}m_\xi(t)$;

4) корреляционная функция производной в среднем СФ равна смешанной производной от корреляционной функции K_ξ , т. е. $K_{\xi'}(t_1, t_2) = \frac{\partial^2 K_\xi(t_1, t_2)}{\partial t_1 \partial t_2}$.

Введем теперь понятие стохастического интеграла. Пусть на (a, b) задана СФ $\xi(t)$. Составим для некоторого разбиения $a = t_0 < t_1 \cdots < t_n = b$ интегральную сумму

$$\sum_{i=1}^n \varphi(t_i^*) \xi(t_i^*) \Delta t_i, \quad t_i^* \in [t_{i-1}, t_i].$$

Здесь $\varphi(t)$ — неслучайная функция, непрерывная на (a, b) , также СВ. Если интегральная сумма сходится в среднем при $\max\{\Delta t_i\} \rightarrow 0$ к некоторой СВ ζ , последнюю называют стохастическим интегралом функции $\xi(t)$ на (a, b) . Обозначим

$$\zeta = \int_a^b \varphi(t) \xi(t) dt.$$

Если верхний предел будет переменной $\tau \in (a, b)$, то стохастический интеграл с переменным верхним пределом — СФ $\zeta(t)$.

Приведем еще два утверждения:

1) матожидание стохастического интеграла равно интегралу от матожидания:

$$M[\zeta(t)] = M\left[\int_a^t \varphi(s) \xi(s) ds\right] = \int_a^t \varphi(s) m_\xi(s) ds;$$

2) корреляционная функция стохастического интеграла равна:

$$K_\zeta(t_1, t_2) = \int_a^{t_1} d\tau_1 \int_a^{t_2} d\tau_2 \varphi(\tau_1) \varphi(\tau_2) K_\xi(\tau_1, \tau_2).$$

Как видите, в рамках корреляционной теории интегрирование и дифференцирование СФ сводится к таковым операциям в *обычном* смысле для матожидания и корреляционной функции. Это обстоятельство тесным образом связано с выбранным типом сходимости.

Пример 18. Для СФ $\xi(t)$ задано $m_\xi(t) = t^2 + 4$. Найти матожидание СФ $\eta(t) = t \cdot \xi'(t) + t^2$.

Решение

$$M[t \cdot \xi'(t) + t^2] = t \cdot M[\xi'(t)] + t^2 = t \cdot \frac{dm_\xi(t)}{dt} + t^2 = 3t^2.$$

■

Пример 19. Математическая модель некоторого устройства

$$y(t) = \int_0^t x(s) ds,$$

где $x(t)$ — сигнал на входе, а $y(t)$ — сигнал на выходе.

Пусть на вход подается случайный сигнал, описываемый СФ $\xi(t)$. В рамках корреляционного подхода случайная функция задается матожиданием $m_\xi = 1 + \sin^2 \omega t$ и корреляционной функцией $K_\xi(t_1, t_2) = 4 \cos \omega t_1 \cos \omega t_2$. Найти $m_\eta(t)$ и $K(t_1, t_2)$, где η — сигнал на выходе.

Решение

$$\begin{aligned} m_\eta(t) &= \int_0^t m_\xi(s) ds = \int_0^t (1 + \sin^2(\omega s)) ds = \frac{3}{2}t - \frac{1}{\omega t} \sin(2\omega t), \\ K_\eta(t_1, t_2) &= \int_0^{t_1} ds_1 \int_0^{t_2} ds_2 K_\xi(s_1, s_2) = \int_0^{t_1} ds_1 \int_0^{t_2} ds_2 4 \cos(\omega s_1) \cos(\omega s_2) = \\ &= \frac{4}{\omega^2} \sin(\omega t_1) \sin(\omega t_2), \\ D(\eta(t)) &= K_\eta(t, t) = \frac{4}{\omega^2} \sin^2(\omega t). \end{aligned}$$

■

Как уже отмечалось, функциональный вид реализации случайного процесса может быть известен из физических соображений. Скажем, гармоническое колебание может иметь случайный сдвиг по фазе или/и случайную амплитуду. В этом случае имеет смысл своеобразное «аналитическое» задание СФ:

$$\xi(t) = \Psi \cdot \sin(3t + \varphi).$$

Для описания же нам будет достаточно аппарата теории вероятностей, в частности, методов получения числовых характеристик и распределений функций СВ.

12.4. Стационарные случайные процессы

Представим себе, что мы изучаем некую реальную систему, находящуюся длительное время в постоянных макроскопических условиях. В качестве примера можно взять броуновское движение частичек краски в воде или проводник с током в очень слабом электрическом поле. Примеры выбраны так, чтобы для описания характеристик системы потребовалась случайная функция.

В случае броуновского движения такой функцией будет каждая координата частицы краски в произвольный момент времени. Случайной будет и сила тока, отличающаяся от величины, даваемой законом Ома, это явление называют тепловым шумом. Основное свойство рассматриваемых реальных процессов — стационарность, т. е. неизменность макроскопических характеристик с течением времени. Для описания таких реальных процессов применяются СФ, удовлетворяющие определенным свойствам, так называемые *стационарные случайные функции*. Чаще их называют *стационарными случайными процессами*.

Дадим определения.

Определение. Стационарность СФ в узком смысле означает, что все конечномерные плотности вероятности инвариантны относительно произвольной трансляции по времени.

Это значит, что выбор начала отсчета времени t не влияет на свойства случайного процесса (СП). В частности, должны иметь место такие свойства:

$$m_{\xi}(t) = m_{\xi}(t + \tau) = m_{\xi}; \quad (12.6)$$

$$K_{\xi}(t_1, t_2) = K_{\xi}(t_1 + \tau, t_2 + \tau). \quad (12.7)$$

Последнее возможно лишь при условии

$$K_{\xi}(t_1, t_2) = K_{\xi}(|t_1 - t_2|). \quad (12.8)$$

В рамках корреляционной теории мы учитываем только первые два момента СП. Поэтому условие стационарности ослабляется.

Определение. СП называют **стационарным в широком смысле**, если выполняются только (12.6) и (12.7).

Из (12.8) и общих свойств корреляционных функций следует, что

$$K_\xi(\tau) = K_\xi(-\tau) \text{ и } |K_\xi(\tau)|^2 \leq K_\xi(0) = D[\xi].$$

Пример 20. Пусть $\xi(t) = \cos(t + \alpha)$, где α — равномерно распределенная на $[0, 2\pi]$ СВ. Доказать, что $\xi(t)$ стационарный СП.

Решение

Найдем матожидание СП $\xi(t)$:

$$\begin{aligned} M[\xi(t)] &= M[\cos(t + \alpha)] = M[\cos t \cos \alpha - \sin t \sin \alpha] = \\ &= \cos t \cdot M[\cos \alpha] - \sin t \cdot M[\sin \alpha]. \end{aligned}$$

Учитывая, что

$$M[\sin \alpha] = \frac{1}{2\pi} \int_0^{2\pi} \sin x dx = 0 = M[\cos \alpha],$$

получаем $M[\xi(t)] = 0$.

Найдем корреляционную функцию $K_\xi(t_1, t_2)$:

$$\overset{\circ}{\xi}(t) = \xi(t) = \cos(t + \alpha),$$

поэтому

$$\begin{aligned} K_\xi(t_1, t_2) &= M[\cos(t_1 + \alpha) \cos(t_2 + \alpha)] = \\ &= \frac{1}{2} M[\cos(t_2 - t_1) + \cos(t_2 + t_1 + 2\alpha)]. \end{aligned}$$

Для второго слагаемого матожидание равно нулю. Итак,

$$K_\xi(t_1, t_2) = \frac{1}{2} \cos(t_2 - t_1),$$

изучаемый СП является стационарным в широком смысле. ■

12.5. Спектральное разложение стационарного случайного процесса

Рассмотрим СП специального вида:

$$\xi(t) = m_\xi + \sum_{k=1}^n U_k \cos(\omega_k t) + V_k \sin(\omega_k t), \quad (12.9)$$

где m_ξ , ω_k — постоянные, а U_k и V_k — центрированные некоррелированные СВ с дисперсиями D_k :

$$M[U_k] = M[V_k] = 0; \quad M[V_i \cdot V_j] = M[U_i \cdot U_j] = D_i \delta_{ij}; \quad M[U_i V_j] = 0. \quad (12.10)$$

Найдем для (12.9) корреляционную функцию. С учетом (12.10) это сделать очень легко:

$$\begin{aligned} K_\xi(t_1, t_2) &= M[\overset{\circ}{\xi}(t_1) \cdot \overset{\circ}{\xi}(t_2)] = \\ &= M \left[\sum_{k,l=1}^n (U_k \cos(\omega_k t_1) + V_k \sin(\omega_k t_1)) (U_l \cos(\omega_l t_2) + V_l \sin(\omega_l t_2)) \right] = \\ &= \sum_{k=1}^n (M[U_k^2] \cos(\omega_k t_1) \cos(\omega_k t_2) + M[V_k^2] \sin(\omega_k t_1) \sin(\omega_k t_2)) = \\ &= \sum_{k=1}^n D_k \cos \omega_k(t_2 - t_1). \quad (12.11) \end{aligned}$$

Следовательно, СП (12.9) есть стационарный в широком смысле, его называют *процессом с дискретным спектром*, а выражение (12.9) — *спектральным представлением*. Дадим наглядную интерпретацию этим терминам. Каждое слагаемое суммы (12.9) можно записать в виде

$$\xi(t) = U_k \cos(\omega_k t) + V_k \sin \omega_k t = A + k \sin(\omega_k t + \alpha_k)$$

и рассматривать как математическую модель гармонического колебания с частотой ω_k , случайной амплитудой $A_k \equiv \sqrt{U_k^2 + V_k^2}$ и случайной фазой $\omega_k t + \alpha_k$, $\alpha_k \equiv \arctg \frac{U_k}{V_k}$. СП (12.9) есть сумма таких гармонических колебаний. Заметим, что дисперсия

$$D[\xi(t)] = K_\xi(t, t) = \sum_{k=1}^n D_k.$$

Энергия гармонического колебания пропорциональна квадрату амплитуды, значит, D_k пропорциональна матожиданию энергии случайного колебания с частотой ω_k . Стационарный СП при $t \in [-\tau, \tau]$ всегда можно представить в виде (12.9). Если же $t \in (-\infty, +\infty)$ и K_ξ непериодическая, но абсолютно интегрируемая, то выполняются условия разложимости $K_\xi(\tau)$ в интеграл Фурье.

$$K_\xi(\tau) = \int_0^\infty S_\xi(\omega) \cos(\omega\tau) d\omega; \quad S_\xi(\omega) = \frac{2}{\pi} \int_0^\infty K_\xi(\tau) \cos(\omega\tau) d\tau. \quad (12.12)$$

СП называют процессом с *непрерывным спектром*, а $S_\xi(\omega)$ — *спектральной плотностью* СП. Поскольку

$$D[\xi(t)] = K_\xi(0) = \int_0^\infty S_\xi(\omega) d\omega,$$

а дисперсия пропорциональна энергии колебания, то $S_\xi(\omega)$ также называют *энергетическим спектром*.

Пример 21. Найти энергетический спектр СП $\xi(t)$, имеющего нулевое матожидание и корреляционную функцию

$$K_\xi(\tau) = \sigma^2 \exp(-\alpha|\tau|) \cos \omega_0 \tau, \quad \alpha > 0.$$

Решение

$$\begin{aligned} S_\xi(\omega) &= \frac{2}{\pi} \int_0^\infty \sigma^2 \exp(-\alpha|\tau|) \cos(\omega_0 \tau) \cos(\omega \tau) d\tau = \\ &= \frac{\alpha \sigma^2}{\pi} \left(\frac{1}{\alpha^2 + (\omega - \omega_0)^2} + \frac{1}{\alpha^2 + (\omega + \omega_0)^2} \right). \end{aligned} \quad (12.13)$$

■

Изложенные выше сведения по теории СП являются первоначальным знакомством с понятиями и спецификой этого раздела математики. Уровень и строгость материала ориентированы на инженерные приложения, в частности, теории сигналов и цепей. Для дальнейшего изучения можно порекомендовать работы [8, 10].

Глава 13

Математическая статистика

13.1. Основные понятия

В названии нашего учебного пособия объединены две математические дисциплины — теория вероятностей и математическая статистика (МС). ТВ является дедуктивной наукой, она основана на системе аксиом, из которых логическим путем выводятся свойства объектов. Нашей целью при изучении ТВ было научиться находить вероятности сложных событий через вероятности более простых.

Как отмечалось в предисловии, пособие ориентировано на приложения математических методов, поэтому кроме знакомства с **математической** статистикой мы будем рассматривать еще и *прикладную* статистику. Если МС — математическая дисциплина, с присущей математике строгостью формулировок и доказательств, то прикладная — является по сути научно-прикладной деятельностью государственных и частных институтов.

В 1954 г. в своем известном учебнике академик АН УССР Б. В. Гнеденко дал следующее определение:

«Статистика состоит из трех разделов:

- 1) сбор статистических сведений, то есть сведений, характеризующих отдельные единицы каких-либо массовых совокупностей;
- 2) статистическое исследование полученных данных, заключающееся в выяснении тех закономерностей, которые могут быть установлены на основе данных массового наблюдения;
- 3) разработка приемов статистического наблюдения и анализа статистических данных. Последний раздел, собственно, и составляет содержание *математической статистики*» [9].

Прикладная статистика основана на индуктивном методе — по эмпирическим данным об отдельных объектах делаются заключения общего характе-

ра, касающиеся всего множества. МС использует язык и понятия ТВ и вместе с ней обеспечивает прикладную статистику приемами и методами, которые позволяют изучать закономерности и свойства случайных явлений и сложных систем.

Цель МС — разработка методов сбора и обработки статистических (опытных) данных, на основании которых составляется цельное представление о некотором случайном явлении. В нашем курсе мы познакомимся как с основными понятиями и задачами МС, так и приемами и методами прикладной статистики.

Объект изучения МС — случайная величина, ее закон распределения, его параметры и характеристики. В прикладной же статистике объект — некоторое множество, элементы которого обладают свойством X (оно может выражаться числом, набором чисел, а может быть и нечисловым). Это множество называют *генеральной совокупностью* (ГС). С формальной точки зрения ГС — множество значений некоторой СВ. Непосредственное изучение всей ГС невозможно, поэтому изучается лишь подмножество — *выборка*. Этот термин в МС применяют как к набору значений признака X элементов ГС, полученному в измерении, так и к многомерной СВ, составляющие которой считаются независимыми и имеющими распределение изучаемой СВ (см. таблицу).

Понятие	Математическая статистика	Прикладная статистика
Генеральная совокупность (ГС)	Множество значений СВ ξ	Множество реальных объектов
Выборка	Вектор $(\xi_1, \xi_2, \dots, \xi_k)$ с независимыми составляющими	Подмножество ГС
Варианта	Значение СВ, полученное при измерении	Значение признака объекта ГС

13.2. Описательная статистика. Визуализация

В зарубежной литературе по статистике фигурируют два важных термина: *Discriptive statistic* — описательная статистика и *Exploratory statistic* — исследовательская статистика. Оба эти этапа статистического изучения систем мы и рассмотрим.

Пусть элементы ГС имеют признак X (его еще называют **вариантой**). Проведено n измерений, результаты которых составляют большой массив

чисел — **простой статистический ряд**; n называют **объемом выборки**. В исходном состоянии этот набор чисел не позволяет исследователю получить какие-либо свойства исследуемой варианты (СВ).

Первичная обработка заключается в группировке. Если выборка содержит лишь небольшое количество различных чисел, т. е. значения повторяются, то группировка приводит к выборочному ряду распределения.

x_1	x_2	\dots	x_k
m_1	m_2	\dots	m_k

Здесь в первой строке стоят различные значения в порядке возрастания, во второй — *частоты* этих значений, т. е. m_i указывает, сколько раз значение x_i повторялось в выборке. Естественно,

$$\sum_{k=1}^n m_k = n.$$

Размах вариации (выборки) — разность максимального и минимального значений варианты. Графическое представление выборки — полигон частот. Это ломаная линия, полученная последовательным соединением точек (x_i, m_i) отрезками прямых.

Если же измерение проводилось с высокой точностью (нет повторов) или объем выборки очень большой, то производится интервальная группировка. Интервал, занимаемый значениями варианты, делится на небольшое (20–30) количество равных интервалов, частота определяется для каждого интервала. Это позволяет лучше выделить закономерные различия значений варианты и исключить несущественные.

(x_1, x_2)	(x_2, x_3)	\dots	(x_k, x_{k+1})
m_1	m_2	\dots	m_k

Визуализация интервального ряда — *гистограмма* (рис. 13.2). Это ступенчатая фигура, составленная из прямоугольников. Основаниями являются интервалы варианты, а высота пропорциональна частоте. Коэффициент пропорциональности выбирается так, чтобы полная площадь всех прямоугольников равнялась единице.

Есть еще один довольно остроумный способ первичной обработки выборки, автор — замечательный математик и статистик Джон Тьюки (John Wilder Tukey)¹. Он назвал этот метод «Стебель и лист» (*Stem_and_Leaf*). Вы просто переписываете выборку особым способом. Старшие разряды значений варианты записываются в столбик, а младшие — в строчку за соответствующими старшими. Выглядит это так:

Stem	Leaf
1	1 1 2 3 3 4 4
1	5 6 6 8
2	0 3
2	7 8
3	
3	5 7 8 8
4	0 0 0 1 2 4 4 4
4	5 5 6 7 7 7 8 8 9

Первое достоинство — легкость выполнения, при этом наибольшее и наименьшее значения получаются естественно. Кроме того, распределение визуализируется в своеобразную *квазигистограмму*.

Числовые характеристики выборки — выборочное среднее:

$$\bar{x}_B = \frac{1}{n} \sum m_i x_i,$$

и выборочная дисперсия:

$$D_B = \frac{1}{n} \sum (\bar{x}_B - x_i)^2 m_i.$$

Есть еще ряд характеристик.

Квантиль — значение переменной, отделяющей заданную часть ранжированного ряда данных. Чаще всего ранжированный ряд делят на две, четыре, десять и сто частей. При этом квантили называют соответственно медианой (меньше этого числа ровно половина выборки), квартилями (меньше первого из них 25 % выборки), децилями (меньше первого — 10 %) и процентилями (1 %). Значение квантили находят по таблицам.

Эти характеристики очень наглядно представляются оригинальной визуализацией Тьюки — *box-with-whiskers* (также *Box and Whisker*, по-разному)

¹Один из выдающихся ученых прошлого века (1915–2000), достаточно упомянуть FFT (быстрое преобразование Фурье), многочисленные оригинальные методы статистической визуализации и многое другое.

(рис. 13.1). У нас отечественного общепринятого термина нет, говорят *боксовая диаграмма*. Здесь наглядно отмечаются Q_1 , Q_3 — соответственно первый и третий квартили (они же 25 % и 75 % процентиля), межквартильные расстояния IQR , выбросы и многое другое.

Заметим, что выборочные среднее и медиана существенно различаются. В механической аналогии первое будет центром тяжести стержня, а вторая — просто делит его на две половинки одинаковой массы. Надеюсь, всем понятно, что масса может распределяться по-разному и центр тяжести соответственно не обязательно делит тело на две равнотяжелые части.

Выборочный коэффициент асимметрии служит для характеристики смещения вариационного ряда относительно среднего значения по величине и направлению. Если *полигон* асимметричен, то одна из его ветвей, начиная с вершины, имеет более пологий «спуск», чем другая. В случае отрицательного коэффициента асимметрии более пологий «спуск» полигона наблюдается слева, в противном случае — справа. В первом случае асимметрию называют левосторонней, а во втором — правосторонней. Асимметрия менее 0,5 считается малой.

Важной характеристикой является также мода — наиболее часто встречающееся значение.

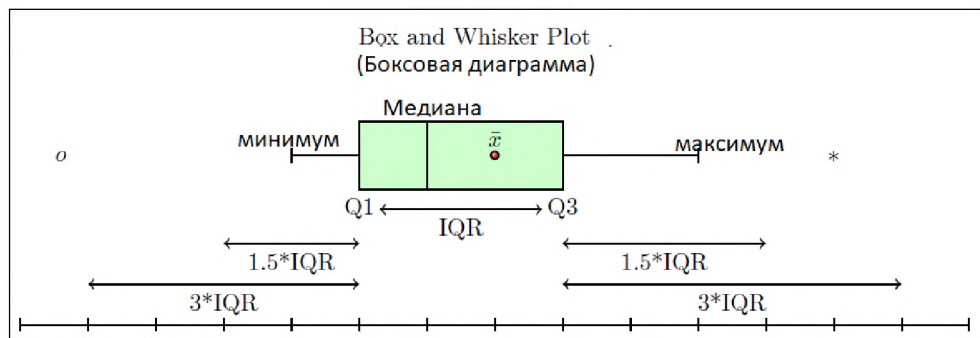


Рис. 13.1. Вот такую богатую информацию отображает «коробка с усами». Более подробное описание метода и обозначений см. в [38]

Пример 22. Рассмотрим учебную выборку:

20.5	20.7	20.8	21.0	21.0	21.4	21.5	22.0	22.1	22.5
22.6	22.6	22.7	22.7	22.9	22.9	23.1	23.3	23.4	23.5
23.6	23.6	23.6	23.9	24.1	24.3	24.5	24.5	24.8	24.8
24.9	24.9	25.1	25.1	25.2	25.6	25.8	25.9	26.1	26.7

Первичная обработка — определение экстремальных значений и разбиение интервала на группы (классы). Технологии контроля качества предполагают сбор информации на рабочих местах, т. е. вручную. Работник поочередно зачеркивает значения в выборке и ставит условный значок в таблице. Было уже отмечено четыре значения в этом интервале ||||, добавил еще одно, стало ||||| — так обозначают пять. Встретилось еще одно значение в этом интервале, стало шесть |||||. Последующие числа обозначают аналогично — семь: |||||, восемь: |||||, девять: |||||, десять: |||||.

Это очень удобно и не требует особого внимания. В дальнейшем исходные данные переводятся в обычные числа и идет последующая обработка:

(20.45, 21.35)	20.5–21.3		5	5/36	20.9
(21.35, 22.25)	21.4–22.2		4	4/36	21.8
(22.25, 23.15)	22.3–23.1		8	8/36	22.7
(23.15, 24.05)	23.2–24.0		7	7/36	23.6
(24.05, 24.95)	24.1–24.9		8	8/36	24.5
(24.95, 25.85)	25.0–25.8		5	5/36	25.4
(25.85, 26.75)	25.9–26.7		3	3/36	26.3

Затем полученное выборочное распределение визуализируется. Наиболее распространены гистограммы, их форма отвечает психофизиологическим особенностям зрительного восприятия человека. Соотношение длин столбиков анализируется на предаттентивном уровне — практически мгновенно (рис. 13.2).

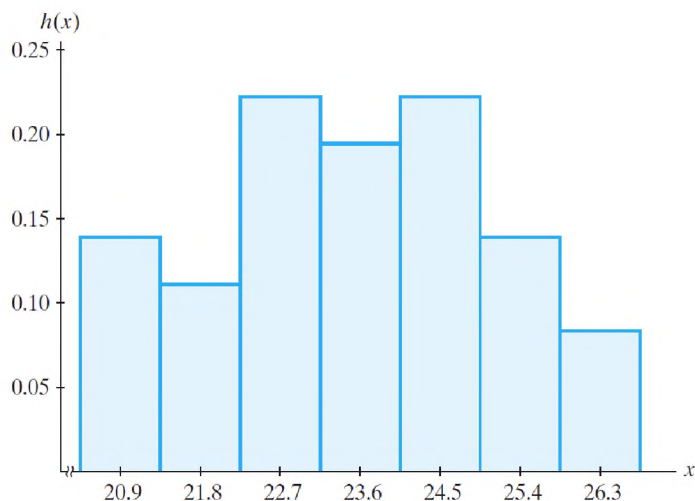


Рис. 13.2. Визуализация учебной выборки

13.3. Оценки параметров распределения случайной величины

13.3.1. Точечная оценка параметров распределения

Одна из важнейших задач математической статистики — оценка параметров неизвестного закона распределения СВ ξ . Эта задача возникает, если имеются четкие модельные соображения относительно того, что закон распределения относится к некоторому известному классу, например, нормальному или равномерному. Возникает необходимость определить параметры, которые имеются для каждого класса распределений: λ — у показательного, (m, σ) — у нормального и т. д.

Оценка параметра одним числом называется *точечной*. Фактически это СВ θ^* , являющаяся функцией выборки $(\xi_1, \xi_2, \dots, \xi_k)$. Здесь ξ_i — независимые СВ, имеющие распределение СВ ξ . Поэтому для всех ξ_i матожидания и дисперсии одинаковы.

К точечной оценке предъявляются требования *несмещенности, эффективности и состоятельности*.

1. Оценка называется **несмещенной**, если ее МО равно оцениваемому параметру $M[\theta^*] = \theta$. В качестве примера можно указать выборочное среднее —

несмещенную оценку МО изучаемой СВ:

$$M[\bar{x}_B] = \frac{1}{n} M[\sum m_i \xi_i] = \frac{1}{n} \sum m_i m_\xi = m_\xi.$$

Заметим, что выборочная дисперсия таким свойством не обладает. Несмещенной оценкой дисперсии является *исправленная* выборочная дисперсия:

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x}_B)^2. \quad (13.1)$$

2. **Состоятельность** оценки означает, что она сходится по вероятности к оцениваемому параметру. Согласно ЗБЧ, выборочное среднее этому условию удовлетворяет.

3. **Эффективность** предполагает малую дисперсию данной оценки в сравнении с другими. Найдем дисперсию выборочного среднего:

$$D[\bar{x}_B] = \frac{1}{n^2} D[\sum m_i \xi_i] = \frac{1}{n^2} \sum m_i D[\xi] = \sigma^2/n.$$

Учтено, что отдельные измерения независимы и дисперсия постоянная.

Итак, в дальнейшем будем использовать \bar{x}_B в качестве точечной оценки параметра МО. Для точечной оценки параметра σ^2 или дисперсии применяют *несмещенную дисперсию*.

13.3.2. Интервальные оценки параметров распределения

Приближенные вычисления всегда сопровождаются указанием погрешности, т. е. указывается интервал $(x^* - \Delta, x^* + \Delta)$, в котором лежит неизвестное точное значение x^* . Чем уже этот интервал, тем точнее известно значение величины. Но ни одна физическая величина не может быть измерена «совершенно точно» — $\Delta \neq 0$.

Поставим задачу составить подобный интервал для параметра распределения Θ , используя выборку значений СВ:

$$\Theta_1^* < \Theta < \Theta_2^*. \quad (13.2)$$

Здесь Θ_i — две СВ, являющиеся функциями выборки.

Определение. Интервальной называется оценка параметра Θ двумя числами — концами *доверительного* интервала (13.2).

Выполнение неравенства (13.2) — это случайное событие, его вероятность называют *доверительной вероятностью* или *надежностью* γ . Интересно отметить, что оцениваемая величина хоть и неизвестна, но имеет некоторое неслучайное значение, а вот концы доверительного интервала меняются от выборки к выборке. Поэтому говорят, что доверительный интервал *накрывает* оцениваемую величину с вероятностью γ .

На рис. 13.3 изображены доверительные интервалы для нескольких выборок, соответствующих z -статистике (см. далее). Как видно, в большинстве случаев интервал накрывает оцениваемый параметр — матожидание СВ $\xi \sim N(m, \sigma)$. В данном случае $\gamma = 0.9$ (90-процентный доверительный интервал).

Надежность задается специалистом в предметной области, в учебных задачах мы будем полагать: $\gamma = 0.9, 0.95, 0.99$.

Задача — при заданной надежности установить доверительный интервал для параметра распределения.

Основное — построить вспомогательную СВ, называемую *статистикой*, с известным законом распределения, который не зависит от оцениваемого параметра. Это очень сложная задача, поэтому многие известные методики интервального оценивания и статистики носят фамилии своих создателей. В нашем курсе будут рассмотрены интервальные оценки параметров нормального распределения.

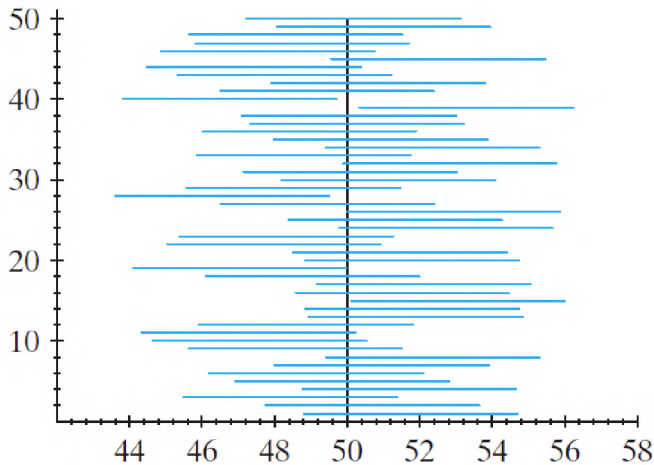


Рис. 13.3. Интерпретация доверительного интервала

13.4. Оценка параметров нормального распределения

Как известно, нормальное распределение (Гаусса) задается двумя параметрами. Это МО m и дисперсия σ^2 . Для каждого из перечисленных случаев строится своя статистика:

- σ известно, оцениваем m ;
- σ неизвестно, оцениваем m ;
- m известно, оцениваем σ ;
- m неизвестно, оцениваем σ .

1. Пусть СВ ξ имеет нормальное распределение $N(m, \sigma)$ с известным значением σ . \bar{x}_B как сумма независимых СВ с распределениями $N(m, \sigma)$ имеет также нормальное распределение — мы показывали, что данное распределение композиционно устойчиво. Его параметры находим с помощью свойств МО. Получим $M[\bar{x}_B] = m$, $D[\bar{x}_B] = \sigma^2/n$. Выполним переход к новой СВ:

$$\frac{\bar{x}_B - m}{\sigma/\sqrt{n}}. \quad (13.3)$$

Эта статистика имеет нормальное распределение $N(0, 1)$, не зависящее от оцениваемого параметра m . Общепринятое название: z -статистика.

Составим доверительный интервал с заданной надежностью γ , т. е. новая статистика (13.3) должна попадать в него с вероятностью γ .

$$\left| \frac{\bar{x}_B - m}{\sigma/\sqrt{n}} \right| < z_\gamma,$$

z_γ находится по таблице плотности распределения $N(0, 1)$, или по таблице квантилей. Можно использовать и подходящее программное обеспечение. Разрешаем неравенство относительно оцениваемого параметра и получаем доверительный интервал.

Здесь есть маленькая хитрость. Интервал двусторонний, поэтому вероятность непопадания в него что слева, что справа будет $\frac{1-\gamma}{2}$. Поэтому для отыскания границы z_γ требуется выполнение равенства $F_\xi(z_\gamma) = \gamma + \frac{1-\gamma}{2} = \frac{1+\gamma}{2}$. Значит, при $\gamma = 0.9$ функция распределения $F_\xi(z_\gamma) = 0.95$. Соответственно в таблице квантилей (а это обратная по отношению к функции распределения функция) аргумент будет 0,95. Функция Лапласа опре-

деляется по-разному: в одних таблицах она совпадает с F_ξ , в других — $F_\xi(x) = \Phi(x) + 0.5$.

Далее находится сам доверительный интервал (рис. 13.4):

$$\bar{x}_B - \frac{z_\gamma \sigma}{\sqrt{n}} < m < \bar{x}_B + \frac{z_\gamma \sigma}{\sqrt{n}}.$$

Пример 23. Агрегат наполняет баночки майонезом. Масса нетто 250 г. Реально имеется случайный разброс по массе, который моделируется нормально распределенной СВ ξ с $m = 250$ г, $\sigma = 2.5$ г.

Для контроля взвешиваем $n = 25$ баночек. Получаем массы ξ_1, \dots, ξ_{25} , это выборка для ξ .

$$\mu = \frac{1}{n} \sum_{i=1}^{25} \xi_i, \Rightarrow \bar{x} = \frac{1}{25} \sum_{i=1}^{25} x_i = 250.2 \text{ г.}$$

Если масса баночки есть нормально распределенная величина с $\sigma = 2.5$ г, то и выборочное среднее также имеет нормальное распределение с $\sigma/\sqrt{n} = 0.5$ г.

Находим доверительный интервал для $\gamma = 0.95 \Rightarrow t_\gamma = 1.96$:

$$250.2 - 1.96 \cdot 0.5 < m < 250.2 + 1.96 \cdot 0.5 \Rightarrow 249.22 < m < 251.18.$$

2. Пусть теперь СВ ξ имеет нормальное распределение $N(m, \sigma)$ с неизвестными как m , так и σ .

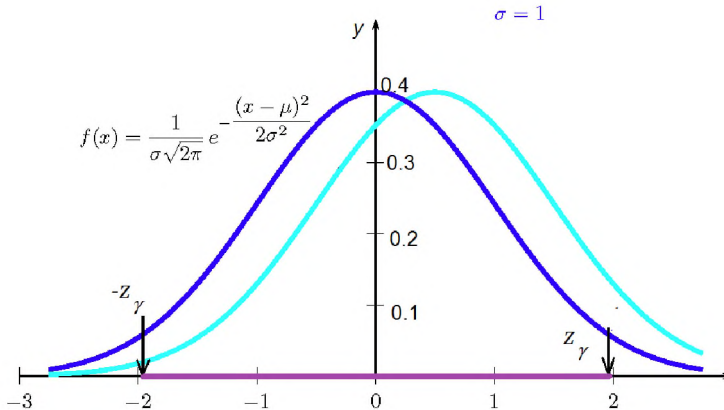


Рис. 13.4. Доверительный интервал $(-z_\gamma, z_\gamma)$

В этом случае выборочные характеристики \bar{x}_B и s^2 независимы. Первая имеет распределение $N(m, \sigma/\sqrt{n})$, а величина $\frac{ns^2}{\sigma^2}$ — имеет распределение χ^2 с $n - 1$ степенью свободы. Составим отношение

$$\frac{\frac{\bar{x}_B - m}{\sigma/\sqrt{n}}}{\sqrt{\frac{1}{n-1} \frac{ns^2}{\sigma^2}}} = \frac{\bar{x}_B - m}{s/\sqrt{n-1}}.$$

Это отношение имеет рассматривавшееся ранее распределение Стьюдента:

$$\frac{\xi}{\sqrt{\chi_n^2/n}},$$

с $n - 1$ степенью свободы.

Итак, для построения доверительного интервала для МО в данном случае используется t -статистика — $\frac{\bar{x}_B - m}{s/\sqrt{n-1}}$. Ее распределение известно и не зависит от оцениваемого параметра. Построение доверительного интервала аналогично предыдущему, только использовать надо таблицы распределения Стьюдента (рис. 13.5).

Оценка параметра σ проводится с использованием статистик (рис. 13.6):

- $\frac{nD_B}{\sigma^2} = \chi_n^2$, имеющей распределение χ^2 с n степенями свободы (при известном m);
- $\frac{(n-1)}{\sigma^2}s^2 = \chi_{n-1}^2$, имеющей распределение χ^2 с $n - 1$ степенями свободы (при неизвестном m).

Доверительный интервал будет ограничен двумя величинами. Находим их следующим образом. Поскольку вероятность попадания в искомый интервал равна γ , то вероятность непопадания будет $1 - \gamma$. Приравниваем вероятности «непопадания влево» и «непопадания вправо». Основанием является равновозможность этих ситуаций. Левая граница доверительного интервала χ_1^2 теперь будет квантилью порядка $(1 - \gamma)/2$, а правая — χ_2^2 будет квантилью порядка $(1 + \gamma)/2$. Искомый доверительный интервал получается такой:

$$\frac{ns^2}{\chi_2^2} < \sigma^2 < \frac{ns^2}{\chi_1^2}. \quad (13.4)$$

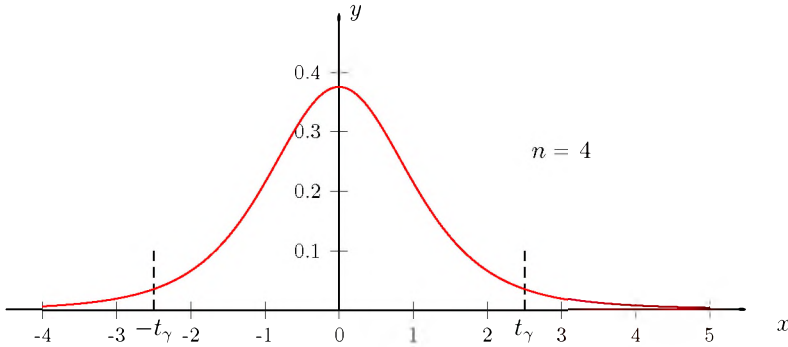


Рис. 13.5. Доверительный интервал $(-t_\gamma, t_\gamma)$ t -распределения Стьюдента

13.5. Оценка параметра p распределения Бернулли

Одна из основных теоретических моделей теории вероятностей – схема Бернулли: последовательность независимых испытаний, в каждом из которых событие может произойти с вероятностью p или не произойти с вероятностью $q = 1 - p$. В статистической интерпретации за вероятность можно взять отношение N_A/N , где N_A – количество объектов с признаком A , а N – общее количество объектов. По терминологии статистики, последняя величина – объем генеральной совокупности, а само отношение называют *генеральной долей*.

Определение генеральной доли по выборке сводится к задаче нахождения интервальной оценки параметра распределения Бернулли – p . При больших объемах выборки n распределение Бернулли можно описывать нормальным распределением. Для этого переходим к новой переменной $\frac{k - np}{\sqrt{npq}}$. Здесь k – число «успехов» в схеме Бернулли, т. е. $\frac{k}{n}$ и есть выборочная доля признака.

Обозначим $p = \frac{N_A}{N}$ и $p^* = \frac{k}{n}$. Распределение $p^* \sim N\left(p, \sqrt{p(1-p)/n}\right)$.

Для биномиального распределения воспользуемся нормальной аппроксимацией ($pn \gg 10$).

$$P\left(\left|\frac{k}{n} - p\right| < \delta\right) = 2\Phi\left(\frac{\delta}{\sigma_B}\right) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sqrt{pq}}\right),$$

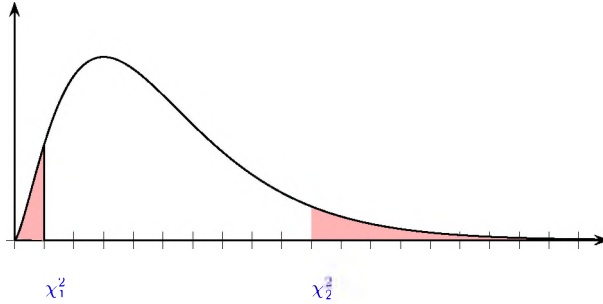


Рис. 13.6. Доверительный интервал (χ_1^2, χ_2^2) распределения χ^2

$$\delta = \frac{t\sqrt{pq}}{\sqrt{n}} \Rightarrow \left| \frac{k}{n} - p \right| < \frac{t\sqrt{pq}}{\sqrt{n}} \Rightarrow \left(\frac{k}{n} - p \right)^2 < \frac{t^2 pq}{n}.$$

Если доля опрошенных (выборочная) p^* , то генеральная доля оценивается интервалом. При заданной надежности γ доверительный интервал для генеральной доли получим из неравенства

$$|p - p^*| < t_\gamma \sqrt{\frac{pq}{n}} \text{ или } |p - p^*| < t_\gamma \sqrt{\frac{p(1-p)}{n}}.$$

Это неравенство можно решить как квадратное неравенство и получить границы интервала:

$$p_{1,2} = \frac{1}{2} \frac{2np^* + t_\gamma^2 \pm \sqrt{4np^*t_\gamma^2 + t_\gamma^4 - 4t_\gamma^2np^{*2}}}{n + t_\gamma^2}.$$

Иначе

$$p_{1,2} = \frac{1}{n + t_\gamma^2} \left(np^* + \frac{t_\gamma^2}{2} \pm t_\gamma \sqrt{\frac{np^*(n - np^*)}{n} + \frac{t_\gamma^2}{4}} \right).$$

Обычно правую часть раскладывают до первого порядка по $1/n$ (большие n порядка сотен):

$$p_{1,2} = p^* \pm Z_{1-\alpha/2} \sqrt{\frac{p^*(1-p^*)}{n}}.$$

Мы получили интересный результат, величину $\sigma(p^*) = \sqrt{\frac{p^*(1-p^*)}{n}}$ называют среднеквадратичной ошибкой выборки.

Возникают три задачи:

- Дано n и $\alpha = 1 - \gamma$, найти доверительный интервал.
- Даны доверительный интервал и n , оценить надежность.
- Оценить n для заданной надежности и доверительного интервала.

Последняя задача (оценить объем выборки, позволяющий при заданной надежности обеспечить определенный доверительный интервал) непосредственно по приведенной формуле решена быть не может, поскольку содержит неизвестную величину p^* . В этом случае подкоренное выражение неизвестно и его оценивают «по максимуму». $p^*(1-p^*) = 0.25$ – это школьная задачка на производную. Получим для объема выборки значение $n = t_\gamma^2 \frac{0.25}{\Delta^2}$. При $\gamma = 0.9$, $t_\gamma = 1.96$ и $\Delta = 0.04$ это будет $n = 600$.

Доверительный интервал можно интерпретировать так. Пусть при опросе 40 % опрошенных объявили, что будут голосовать за N . 90 % доверительный интервал означает, что среди всех избирателей процент будет от 38 до 42 %. А вот 95 % доверительный интервал будет от 36 до 44 %. Очень важным фактором является количество опрошенных.

Пример 24. Рассмотрим частотное распределение выборочной совокупности 812 школьников по количеству имевшихся у них наличных денег (1987 г.).

Сумма наличных, руб.	Частота, m_i	$x_i \cdot m_i$
До 3	25	75
3–10	48	312
10–25	230	4 025
25–50	280	10 500
50–100	160	12 000
100–150	52	6 500
150–200	12	2 100
Больше 200	5	1 000
Всего	812	36 512

Решение

Найдем выборочное среднее $\bar{x}_в = 45$ руб. Далее оценим СКО, считая распределение нормальным:

$$\sigma = 36.85105155 \approx 36.9.$$

Доверительный интервал с центром в m будет иметь полуширину

$t_\gamma \sigma / \sqrt{n} = 3 \cdot 1.3 = 3.9$ руб. Это при $\gamma = 0.997$. Уменьшив доверительную вероятность до 0.683, мы сузим доверительный интервал до (43.7; 46.3). Это нужно понимать так: повторяя проверку наличия в разных классах, мы примерно в 68 % случаев получим среднее значение для класса в этом интервале. ■

Преобразовав формулу для полуширины интервала, получим возможность оценивать объем выборки:

$$\Delta = t \sqrt{\frac{p^*(1-p^*)}{n}} \Rightarrow n = t^2 \frac{p^*(1-p^*)}{\Delta^2}.$$

Неизвестное подкоренное выражение оценим по максимуму.

$$p(1-p) \rightarrow 0.25 - \max.$$

Для обеспечения относительной погрешности 2 % получим, что $\gamma = 0.997$, т. е. $t = 3$, $n = 3^2 \frac{0.25}{0.02^2} = 5625$.

Приведем оценки объемов выборки для генеральных совокупностей, обеспечивающих с надежностью $p = 0.954$ ошибку 5 %.

Объем ГС	Объем выборки
500	222
1000	286
2000	333
4000	360
5000	370
10 000	385
100 000	398
	400

13.6. Проверка статистических гипотез

Статистическая гипотеза представляет собой предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки.

Примерами статистических гипотез являются предположения: генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу. В первой из них высказано предположение о виде закона распределения, а во второй — о параметрах двух распределений.

Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют *непараметрическими*, в противном случае — *параметрическими*.

Различают *простые* и *сложные* гипотезы. Гипотезу называют простой, если она исчерпывающе описывает закон распределения случайной величины. Например, СВ имеет экспоненциальное распределение с определенным параметром λ .

Постановка задачи требует формулирования двух, одна называется нулевой или проверяемой H_0 , другая H_1 — альтернативной или конкурирующей. Принятие нулевой означает отклонение альтернативной, и наоборот. Надо отметить, что гипотезы можно переставлять, а при данной нулевой формулировать различные альтернативные. Вводится вспомогательная СВ — *статистика критерия* (критерий согласия). Обозначим эту величину через z , ее значение является функцией от выборки $z = z(x_1, x_2, \dots, x_n)$.

Все множество значений критерия делится на два непересекающихся подмножества V_0 и V_1 . Если значение критерия z попадает в область V_0 , то гипотеза H_0 принимается, как еще говорят, «нет оснований для отклонения гипотезы». Если же — в область V_1 , то проверяемая гипотеза отклоняется и принимается альтернативная H_1 . Множество V_0 называется *областью принятия гипотезы* или областью допустимых значений, а множество V_1 — *критической областью* или областью отклонения гипотезы.

Критическая область строится так, чтобы при истинности нулевой гипотезы попадание критерия в нее было *практически невозможным* событием. Вероятность «практически невозможного события» принимается достаточно малой и называется *уровнем значимости* α . Выбор одной области однозначно определяет и другую область.

Проверка гипотезы может привести к ошибочному результату, ведь значение критерия случайно — оно определяется по выборке. Возможные ситуации представлены в таблице.

H_0	Действие	Вероятность	Результат
Истинная	Принимается	$1 - \alpha$	Нет ошибки
Истинная	Отвергается	α	Ошибка 1-го рода
Ложная	Принимается	β	Ошибка 2-го рода
Ложная	Отвергается	$1 - \beta$	Верно

Принято считать, что ошибка 1-го рода приводит к более тяжелым последствиям. Исходя из этого и строится схема проверки — какую именно гипотезу берем за нулевую. Вероятность этой ошибки довольно прозрачно называют *уровнем значимости*. Вероятность события, противоположного ошибке 2-го рода, называют *мощностью* критерия.

При кажущейся абстрактности рассматриваемая задача имеет большую практическую применимость. Ее приходится решать и при контроле качества продукции, и при приеме товаров у поставщика, и при анализе эффективности новых препаратов и методик лечения.

Пример 25. Большая партия изделий может содержать некоторую долю дефектных. Поставщик утверждает, что эта доля составляет 5%; покупатель предполагает, что доля дефектных изделий равна 10%. Условия поставки: из партии случайным образом отбирается и проверяется 10 изделий; партия принимается на условиях поставщика, если при проверке обнаружено не более одного дефектного изделия; в противном случае партия принимается на условиях покупателя. Сформулировать эту задачу в терминах теории проверки статистических гипотез и ответить на следующие вопросы: а) Каковы статистика критерия, область ее значений, критическая область? б) Какое распределение имеет статистика критерия? в) В чем состоят проверяемая и альтернативная гипотезы? г) В чем состоят ошибки 1-го и 2-го рода и каковы их вероятности?

Решение

Ясно, что результат проверки будет определяться количеством бракованных изделий в выборке.

Возможные значения критерия $V = \{0, 1, 2, \dots, 10\}$, критическая область $V_1 = \{2, 3, \dots, 10\}$.

Доля бракованных изделий определяет статистическую вероятность того, что наугад взятое изделие будет бракованное. Таким образом, процедура

нахождения выборочного значения статистики критерия представляет собой схему Бернулли, а распределение — биномиальное $B(10, p)$. Параметр p различен: для нулевой гипотезы H_0 $p = 0.05$ (это утверждение поставщика), для альтернативной гипотезы H_1 $p = 0.1$ (так считает покупатель).

Из условий задачи следует именно такое расположение гипотез, а оно выгодно именно продавцу. Ошибка 1-го рода, которая приводит к более тяжелым последствиям, состоит в установлении цены продукции с учетом 10 % брака — мы отбрасываем истинную гипотезу о 5 % брака. Вероятность ее подсчитывается по биномиальной формуле

$$\alpha = 1 - C_{10}^0 0.05^0 0.95^{10} - C_{10}^1 0.05^1 0.95^9 = 0.086.$$

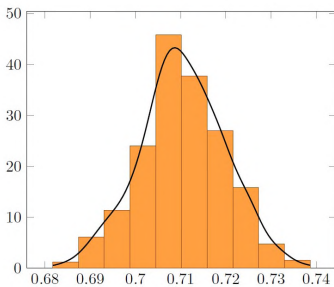
Вероятность ошибки 2-го рода — вероятность того, что СВ, распределенная по закону $B(10, 0.1)$, примет значения из области принятия нулевой гипотезы $V_0 = \{0, 1\}$.

$$\beta = C_{10}^0 0.1^0 0.9^{10} + C_{10}^1 0.1^1 0.9^9 = 0.736.$$

■

13.7. Критерий Пирсона χ^2

Познакомимся с одной популярной методикой проверки гипотезы о законе распределения СВ. Используемая в ней статистика критерия носит фамилию выдающегося математика Карла Пирсона.



Пусть первичная обработка выборки $\{x_1, x_2, \dots, x_n\}$ привела к построению гистограммы такой формы, что можно предположить нормальность изучаемой СВ. Параметры распределения найдем по выборке наблюдений, используя оценки неизвестных параметров предполагаемого закона распределения случайной величины ξ .

Далее область возможных значений случайной величины разбивается на r множеств $\Delta_1, \Delta_2, \dots, \Delta_r$. Это может быть r интервалов в случае, когда ξ — непрерывная случайная величина, или r групп, состоящих из отдельных значений, для ДСВ.

Пусть p_i есть вероятность для величины ξ при данном распределении $F_\xi(x)$ принять значение, принадлежащее i -му множеству Δ_i (интервалу или

группе), m_i — число значений варианты из общего их числа n в выборке $\{x_1, x_2, \dots, x_n\}$, попавших в Δ_i . При этом, очевидно, должны выполняться условия

$$\sum_{i=1}^n m_i = r, \quad \sum_{i=1}^r p_i = 1.$$

В качестве количественной меры расхождения данных выборки $\{m_1, m_2, \dots, m_i\}$ с «теоретическими» данными $\{m_1^*, m_2^*, \dots, m_i^*\}$, $m_i^* = p_i n$ рассмотрим величину

$$\chi_n^2 = \sum_{k=1}^r \frac{(m_k - m_k^*)^2}{m_k^*}. \quad (13.5)$$

Ее аналитический вид напоминает встречающуюся ранее, при нахождении интервальных оценок нормального распределения, статистику

$$\frac{ns^2}{\sigma^2} = \sum_{k=1}^n \frac{(x_k - \bar{x}_B)^2}{\sigma^2},$$

имеющую распределение χ^2 с уменьшенным на единицу числом степеней свободы (за счет линейной связи слагаемых — вспомните определение \bar{x}_B .) Величина

$$\frac{m_k - p_k n}{\sqrt{p_k n}}$$

только *асимптотически* нормальна, поэтому в группах выборки должно выполняться условие $np_k \geq 5$. При невыполнении следует объединить соседние группы в одну.

Отличие величины (13.5) в том, что нам неизвестны параметры распределения, они заменены оценками $m = \bar{x}_B$, $\sigma = s$. Это обстоятельство снижает число степеней свободы еще на две.

Итак, считаем, что статистика критерия (13.5) (критерий Пирсона) при проверке гипотезы о нормальном распределении имеет распределение хи-квадрат с числом степеней свобод $r - 1 - l$. Здесь l — число неизвестных параметров, в случае нормального распределения $l = 2$.

Критическая область односторонняя, т. е. имеет вид $(\chi_{1-\alpha}^2, \infty)$. Мы допускаем случайное различие эмпирических и теоретических частот, но считаем практически невозможным то, что $\chi_{\text{набл}}^2$ превзойдет $\chi_{1-\alpha}^2$. Как обычно, невозможность характеризуется уровнем значимости α . Индекс $1 - \alpha$ у критического значения означает, что это квантиль порядка $1 - \alpha$; α — вероятность

превзойти $\chi_{1-\alpha}^2$, а $1-\alpha$ — принять значение меньше. В таблицах обычно стоят именно квантили.

Из рис. 13.7 с очевидностью следует, что различия эмпирических и теоретических частот обязательно будут. Слишком близкие значения маловероятны, в некоторых источниках их называют подозрительными. А вот большие отклонения критерия Пирсона от нуля также маловероятны при истинности нулевой гипотезы.

Схема применения критерия Пирсона

Указанная задача возникает после первичной обработки выборки, если имеются признаки нормального распределения: это прежде всего вид гистограммы, а также характерное распределение частот.

Постановка задачи. Требуется, используя критерий Пирсона, проверить гипотезу о том, что генеральная совокупность X распределена нормально.

1. Вычислить выборочную среднюю \bar{x}_B и выборочное среднее квадратическое отклонение s , принимаемые за точечные оценки параметров теоретического распределения.
2. Вычислить теоретические частоты. Это можно сделать так:

$$m_i^* = \frac{nh}{s} \cdot \varphi(u_i), \quad u_i = \frac{x_i - \bar{x}_B}{s}, \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2),$$

где n — объем выборки (сумма всех частот); h — шаг (разность между двумя соседними вариантами). Можно и по-другому, используя функцию Лапласа,

$$p_i = \Phi\left(\frac{b_i - \bar{x}_B}{s}\right) - \Phi\left(\frac{a_i - \bar{x}_B}{s}\right), \quad m_i^* = np_i, \quad [a_i, b_i] = \Delta_i.$$

3. Сравнить эмпирические и теоретические частоты с помощью критерия Пирсона. Для этого: а) составляют расчетную таблицу, по которой находят наблюдаемое значение критерия (13.5); б) по таблице критических точек распределения χ^2 , по заданному уровню значимости α и числу степеней свободы $k = r - 3$ (r — число групп выборки) находят критическую точку $\chi_{кр}^2$ правосторонней критической области.

Если $\chi_{набл}^2 < \chi_{кр}^2$, то нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности. Другими словами, эмпирические и теоретические частоты различаются незначимо (случайно).

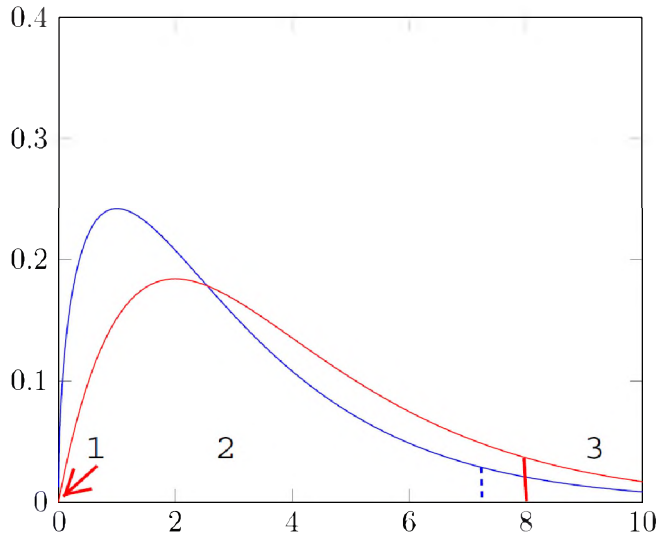


Рис. 13.7. Области значения статистики χ^2 : 1) подозрительно малые различия частот — маловероятно; 2) высокая вероятность выборочных значений критерия χ^2 , такие расхождения частот вполне возможны; 3) **критическая область** — значимо малая вероятность таких значений критерия, считаем гипотезу и наблюдаемые значения несогласующимися

Если $\chi_{\text{набл}}^2 > \chi_{\text{кр}}^2$, то гипотезу отвергают. Другими словами, эмпирические и теоретические частоты различаются значимо.

Замечание. Малочисленные частоты ($m_i < 5$) следует объединить; в этом случае и соответствующие им теоретические частоты также надо сложить. Если производилось объединение частот, то при определении числа степеней свободы по формуле следует в качестве r принять число групп выборки, оставшихся после объединения частот.

Пример 26. Выборка: 20,3; 15,4; 17,2; 19,2; 23,3; 18,1; 21,9; 15,3; 16,8; 13,2; 20,4; 16,5; 19,7; 20,5; 14,3; 20,1; 16,8; 14,7; 20,8; 19,5; 15,3; 19,3; 17,8; 16,2; 15,7; 22,8; 21,9; 12,5; 10,1; 21,1; 18,3; 14,7; 14,5; 18,1; 18,4; 13,9; 19,1; 18,5; 20,2; 23,8; 16,7; 20,4; 19,5; 17,2; 19,6; 17,8; 21,3; 17,5; 19,4; 17,8; 13,5; 17,8; 11,8; 18,6; 19,1. Провести первичную обработку: выполнить интервальную группировку (7 интервалов), найти числовые характеристики, построить гистограмму [28].

Решение

Объем выборки $n = 55$. Размах выборки $w = 23,8 - 10,1 = 13,7$. Длина интервала группировки $b = 13,7/7 \equiv 2$. В качестве первого интервала удоб-

но взять интервал 10–12. $\bar{x}_B = 17.84$, $s^2 = 8.53$. Результаты группировки сведены в таблицу.

Номер интервала	Границы интервала	Середина интервала	Частота	Относительная частота
1	10–12	11	2	0.0364
2	12–14	13	4	0.0727
3	14–16	15	8	0.1455
4	16–18	17	12	0.2182
5	18–20	19	16	0.2909
6	20–22	21	10	0.1818
7	22–24	23	3	0.0545

Гистограмма для данной выборки представлена на рис. 13.8.

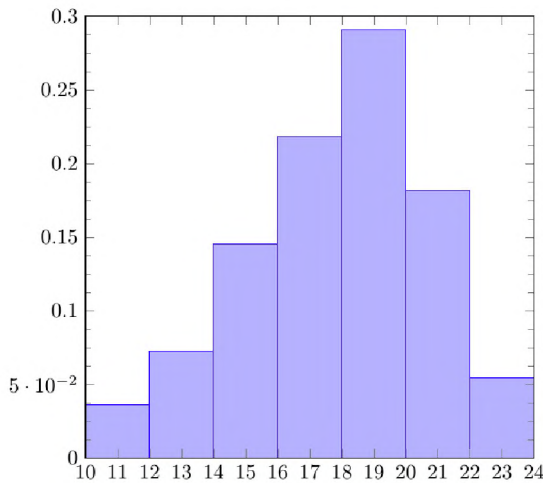


Рис. 13.8. Гистограмма для выборки (пример 26)

Пример 27. В примере 26 была построена гистограмма выборочного распределения. Проверить гипотезу о нормальности этого распределения при уровне значимости $\alpha = 0.1$. ■

Решение

В примере 26 приведена таблица группировок выборки. Первый и последний интервал присоединяем к соседним — они слишком малочисленны.

Номер интервала	Границы интервала	Частота m_k	np_k	$m_k - np_k$	$\frac{(m_k - np_k)^2}{np_k}$
1	$(-\infty, 14)$	6	5.274	0.727	0.010
2	14—16	8	9.273	-1.273	0.175
3	16—18	12	14.168	-2.168	0.332
4	18—20	16	13.662	-2.338	0.400
5	$(20, \infty)$	13	12.633	0.366	0.011

Получаем выборочное значение $\chi^2_{\text{набл}} = 0.928$. В таблице квантилей распределения χ^2 или в таблице критических значений для односторонней области находим $\chi^2_{\text{кр}} = 4.61$, взяв число степеней свободы равным $5 - 3 = 2$ (ведь число группировок уменьшилось!).

Поскольку $\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$, нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности (рис. 13.9). ■

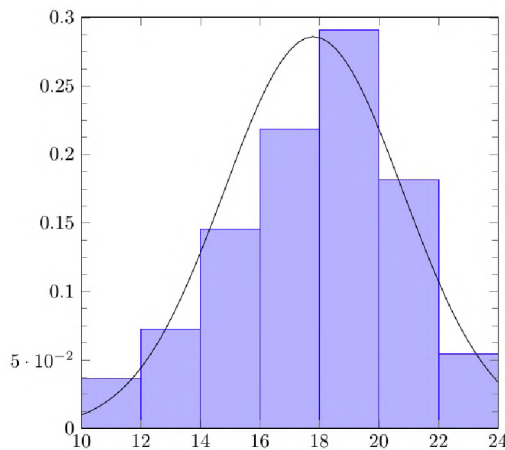


Рис. 13.9. Сравнение гистограммы и теоретической кривой

Рассмотренную методику можно применять и для других предполагаемых распределений. Довольно актуальной является проверка гипотезы о равномерном распределении.

Глава 14

Элементы дисперсионного и корреляционного анализа

14.1. Дисперсионный анализ

14.1.1. Описание методики

Рассмотрим кратко очень перспективную методику изучения свойств сложных систем. Она интенсивно используется в медицине при сравнении методик лечения заболеваний, определении эффективности новых лекарственных препаратов, выявлении причин возникновения болезней. Речь идет о факторном анализе, простейший вариант которого состоит в следующем.

Данная методика позволяет изучать зависимость двух случайных величин, одна из которых — *нечисловая*. Это значит, что ее значения нельзя складывать или перемножать. Например, значениями ее может быть *цвет, тип* применяемого лекарства. Нас может интересовать сравнительное качество обучения вождению автомобиля разными инструкторами.

Такую СВ принято называть фактором, а ее значения — уровнями фактора.

Пусть на случайную величину X действует фактор F , имеющий p постоянных уровней. Например, X — это суммарное количество посетителей магазина в течение дня, а фактор F — день недели. Предполагается, что дисперсия не зависит от уровня фактора F . Требуется проверить, согласуются ли экспериментальные данные с гипотезой об однородности случайной величины X относительно фактора F , т. е. о независимости $\mathbf{M}[X]$ от фактора F .

Пусть фактор F имеет p постоянных уровней и на каждом уровне F_i проведено q_i измерений.

Будем рассматривать всю выборку объема $\sum_{i=1}^p q_i$, x_{ij} — наблюдаемое значение X , причем первый индекс — это уровень фактора, а второй — номер измерения.

$Q = \sum_{i=1}^p \sum_{j=1}^{q_i} (x_{ij} - \bar{x})^2$ — *общая* сумма квадратов отклонений наблюдаемых значений от общей средней.

Изучим влияние фактора на результат измерения. Для этого просуммируем по всем измерениям для одного уровня:

$$\sum_j x_{ij} = q_i \cdot \bar{x}_i. \quad (14.1)$$

Можно назвать \bar{x}_i внутригрупповым средним.

Преобразуем и выражение для дисперсии:

$$\begin{aligned} \sum_i \sum_j (x_{ij} - \bar{x})^2 &= \sum_i \sum_j (x_{ij} - \bar{x}_i + \bar{x}_i - \bar{x})^2 = \\ &= \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum_i \sum_j (\bar{x}_i - \bar{x})^2 + 2 \sum_i \sum_j (x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x}) = \\ &= \sum_{i=1}^p \sum_{j=1}^{q_i} (x_{ij} - \bar{x}_i)^2 + \sum_{i=1}^p \sum_{j=1}^{q_i} (\bar{x}_i - \bar{x})^2 + 2 \sum_{i=1}^p \left((\bar{x}_i - \bar{x}) \sum_{j=1}^{q_i} (x_{ij} - \bar{x}_i) \right). \end{aligned}$$

Заметим, что

$$\sum_j (x_{ij} - \bar{x}_i) = \sum_j x_{ij} - \sum_j \bar{x}_i = q_i \cdot \bar{x}_i - q_i \cdot \bar{x}_i = 0.$$

Остается только

$Q_1 = \sum_{i=1}^p \sum_{j=1}^{q_i} (\bar{x}_i - \bar{x})^2$ — *факторная* сумма квадратов отклонений групповых средних от общей средней, которая характеризует рассеяние «между группами»;

$Q_2 = \sum_{i=1}^p \sum_{j=1}^{q_i} (x_{ij} - \bar{x}_i)^2$ — *остаточная* сумма квадратов отклонений наблюдаемых значений группы от своей групповой средней, которая характеризует рассеяние внутри групп.

$$Q = Q_1 + Q_2.$$

Очевидно, что $\bar{x} = \frac{1}{q_1 + q_2 + \dots + q_p} \sum_{i=1}^p \sum_{j=1}^{q_i} x_{ij}$, при этом $q_1 + q_2 + \dots + q_p = n$ — объем выборки. Нетрудно проверить, что

$$Q_1 = \sum_{i=1}^p \sum_{j=1}^{q_i} x_{ij}^2 - n\bar{x}^2; \quad (14.2)$$

$$Q_2 = \sum_{i=1}^p \left(\frac{1}{q_i} \left(\sum_{j=1}^{q_i} x_{ij} \right)^2 \right) - n\bar{x}^2. \quad (14.3)$$

В дальнейшем будем анализировать и сравнивать несмещенные оценки дисперсий (13.1). Для этого разделим величины Q_1 и Q_2 на соответствующее число степеней свободы: n — объем выборки, p — количество уровней фактора, столько же и групповых средних, связанных одним уравнением (14.1).

В Q_2 входят все групповые средние — p связей. Получим: $s^2 = \frac{Q}{n-1}$ — **общая** дисперсия; $s_1^2 = \frac{Q_1}{p-1}$ — **факторная** дисперсия; $s_2^2 = \frac{Q_2}{n-p}$ — **остаточная** дисперсия.

Если гипотеза о равенстве групповых средних верна, то факторная и остаточная дисперсии являются несмещенными оценками неизвестной генеральной дисперсии и поэтому различаются незначимо. Тогда можно применить статистику критерия, связанную с распределением Фишера (F -распределение) [7, с. 184]. Кроме того, если факторная дисперсия окажется меньше остаточной, то получаем, что экспериментальные данные согласуются с рассматриваемой гипотезой об однородности случайной величины относительно фактора F .

Поясним более наглядно. Выборочные значения варианты X имеют разброс относительно МО, характеризуемый дисперсией. Может оказаться, что все значения разбиваются на группы, соответствующие различным уровням фактора и локализованные вблизи групповых средних. Если сами групповые средние существенно (значимо) отличаются между собой по сравнению с отличием от МО (общее среднее), то можно выдвинуть гипотезу о влиянии уровня фактора на значение X . Полученная F -статистика позволяет выполнить количественный анализ значимости этого влияния в технологии проверки статистических гипотез.

14.1.2. Постоянное число наблюдений

Будем считать, что число наблюдений на каждом из уровней постоянно и равно q (т. е. объем выборки для каждого уровня равен q). Тогда общий объем выборки равен $n = pq$.

Номер испытания	Уровни фактора F_i			
	F_1	F_2	...	F_p
1	x_{11}	x_{21}	...	x_{p1}
2	x_{12}	x_{22}	...	x_{p2}
...
q	x_{1q}	x_{2q}	...	x_{pq}
Групповая средняя	\bar{x}_1	\bar{x}_2	...	\bar{x}_p

Тогда из формул для общего случая получаем:

- $Q = \sum_{i=1}^p \sum_{j=1}^q (x_{ij} - \bar{x})^2$ — *общая* сумма квадратов отклонений наблюдаемых значений от общей средней;
- $Q_1 = q \sum_{i=1}^p (\bar{x}_i - \bar{x})^2$ — *факторная* сумма квадратов отклонений групповых средних от общей средней, которая характеризует рассеяние «между группами»;
- $Q_2 = \sum_{i=1}^p \sum_{j=1}^q (x_{ij} - \bar{x}_i)^2$ — *остаточная* сумма квадратов отклонений наблюдаемых значений группы от своей групповой средней, которая характеризует рассеяние внутри групп.

Как и в общем случае, справедливы следующие формулы:

$$Q = \left(\sum_{i=1}^p \sum_{j=1}^q x_{ij}^2 \right) - pq \bar{x}^2, \quad Q_1 = \left(\sum_{i=1}^p \frac{1}{q} \left(\sum_{j=1}^q x_{ij} \right)^2 \right) - pq \bar{x}^2.$$

В этом случае имеем: $s^2 = \frac{Q}{pq-1}$ — общая дисперсия; $s_1^2 = \frac{Q_1}{p-1}$ — факторная дисперсия; $s_2^2 = \frac{Q_2}{p(q-1)}$ — остаточная дисперсия. Для ответа на вопрос о согласованности гипотезы о независимости дисперсий от уровня фактора F с экспериментальными данными вновь пользуемся F -распределением Фишера.

В нашем случае гипотеза отвергается, если фактическое отношение $F = s_1^2/s_2^2$ больше критического $F(\alpha, n-1, pq-p)$.

Например, $F = \frac{1660}{454.4} = 3,65$, а критическое $F_k(0.05, 3, 16) = 3.24$. В этом случае гипотеза отвергается.

14.2. Регрессионный анализ в теории вероятностей

Знакомясь с системами случайных величин (случайными векторами), вы узнали, что совместное их описание, наблюдение и изучение дают ключ к пониманию того, в чем же собственно состоит зависимость между СВ. Количественно описывать статистическую зависимость можно набором условных законов распределения и регрессией, мерой особого вида зависимости является коэффициент корреляции.

Регрессионным анализом называют раздел математической статистики, объединяющий практические методы исследования регрессионной зависимости между СВ по выборочным данным. Целью регрессионного анализа является установление общего вида уравнения регрессии, а также проверка соответствующих статистических гипотез.

Рассмотрим двумерную случайную величину (ξ, η) . Мы уже отмечали, что возможен один из следующих вариантов:

- между величинами ξ и η нет никакой зависимости (независимые СВ), когда для элемента ω_0 генеральной совокупности по информации о значении случайной величины $\xi(\omega_0)$ нельзя сделать никакого вывода о значении $\eta(\omega_0)$;
- зависимость между ξ и η является функциональной, когда по значению $x_i(\omega_0)$ можно *однозначно* вычислить значение $\eta(\omega_0) = \varphi(\xi(\omega_0))$;
- зависимость имеет *статистический характер*, когда по значению $\xi(\omega_0)$ вычислить точное значение $\eta(\omega_0)$, вообще говоря, нельзя, но можно сделать некоторые выводы о законе распределения случайной величины η при условии, что ξ принимает данное определенное значение, определить условное МО. Именно этот случай нас в данный момент интересует.

Изучение зависимости между двумя СВ начинается с проведения совместного измерения их значений. Важно, что измерение проводится сразу для всей пары. В результате мы получим двумерную выборку объема n — множество пар чисел (x_i, y_i) . Визуализация этих данных заключается в построении так называемой **диаграммы рассеивания** (другое название — **поле корреляций**).

Уже сама картинка взаимного расположения точек много дает опытному исследователю (рис. 14.1).

Более аккуратное описание зависимости можно сделать, используя **регрессию** и **коэффициент корреляции**.

При знакомстве с законами распределения составляющих случайного вектора мы ввели понятие *теоретическая регрессия*. Так называют зависимость условного математического ожидания одной составляющей от значения дру-

гой составляющей (для системы двух СВ). В статистике используется *опытная*, или эмпирическая, регрессия — это зависимость выборочного среднего одной варианты \bar{x}_{y_j} при фиксированном значении y_j второй от самих этих значений.

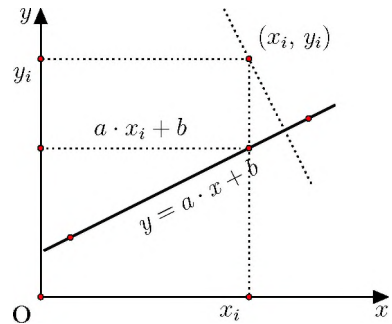
Естественно, такую зависимость имеет смысл строить и отображать графически, если фиксированному значению y_j в выборке соответствует несколько различных значений x_{ij} . Графиком будет ломаная, отрезки которой соединяют последовательно точки (\bar{x}_{y_j}, y_j) .

Однако чаще всего под регрессией понимают функцию, которая находится по выборке в результате определенной специальной процедуры.

Мы рассмотрим только линейную регрессию и получим соответствующее уравнение прямой линии — график линейной регрессии. Кстати, именно эту прямую линию мы видим в центре рис. 14.1.

Для решения задачи используется метод наименьших квадратов (МНК). Это значит, что прямая выбирается таким образом, чтобы сумма квадратов расстояний до нее для всех опытных точек была минимальна.

Решение задачи упрощается, если использовать геометрические соображения. Опуская произвольную точку (x_i, y_i) выборки перпендикуляры на координатные оси и саму линию регрессии, мы получим прямоугольный треугольник. Ясно, что данное построение можно проделать для любой опытной точки, гипотенузы этих треугольников параллельны (между собой), а длина одного из катетов каждого равна расстоянию от соответствующей точки до прямой регрессии.



Очевидно, что все такие треугольники для всех опытных точек подобны, и для каждого длина гипотенузы $|y_i - a \cdot x_i - b|$ пропорциональна длине катета с одним и тем же коэффициентом. Это означает, что минимум суммы квадратов расстояний будет при таком же расположении прямой регрессии на диаграмме рассеивания, что и минимум значения величины

$$F(a, b) = \sum (y_i - a \cdot x_i - b)^2.$$

Таким образом, мы можем искать коэффициенты a и b уравнения $y = ax + b$ прямой линии регрессии из условия минимизации функции, заданной формулой

$$F(a, b) = \sum (y_i - a \cdot x_i - b)^2. \quad (14.4)$$

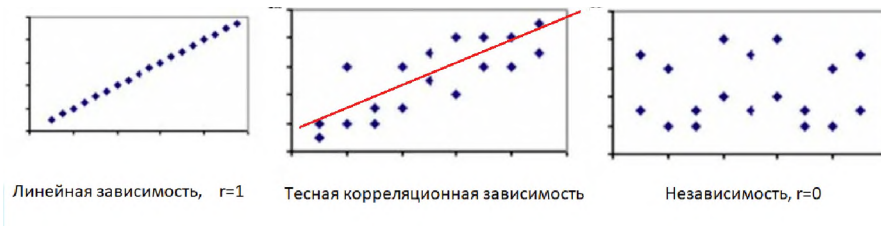


Рис. 14.1. Расположение опытных точек при различной тесноте связи между составляющими случайного вектора

В точке экстремума частные производные обращаются в ноль:

$$\frac{\partial}{\partial a} F(a, b) = 0, \quad \frac{\partial}{\partial b} F(a, b) = 0.$$

Получаем

$$\begin{aligned} & \begin{cases} \sum_{i=1}^n \cdot 2(ax_i + b - y_i) \cdot \frac{\partial(ax_i + b - y_i)}{\partial a} = 0; \\ \sum_{i=1}^n \cdot 2(ax_i + b - y_i) \cdot \frac{\partial(ax_i + b - y_i)}{\partial b} = 0, \end{cases} \Leftrightarrow \\ & \Leftrightarrow \begin{cases} \sum_{i=1}^n \cdot 2(ax_i + b - y_i) \cdot x_i = 0; \\ \sum_{i=1}^n \cdot 2(ax_i + b - y_i) = 0, \end{cases} \Leftrightarrow \\ & \Leftrightarrow \begin{cases} a \left(\sum_{i=1}^n x_i^2 \right) + b \left(\sum_{i=1}^n x_i \right) - \sum_{i=1}^n y_i x_i = 0; \\ a \left(\sum_{i=1}^n x_i \right) + b - \left(\sum_{i=1}^n y_i \right) = 0. \end{cases} \end{aligned}$$

Заметим, что простым преобразованием переменных можно лианеризовать и другие виды функциональной зависимости (степенной или показательной).

Введем обозначения:

$$\begin{aligned} \bar{x}_B &= \frac{1}{n} \left(\sum_{i=1}^n x_i \right), \quad \bar{y}_B = \frac{1}{n} \left(\sum_{i=1}^n y_i \right), \\ \bar{x}_B^2 &= \frac{1}{n} \left(\sum_{i=1}^n x_i^2 \right), \quad \bar{x}_B \bar{y}_B = \frac{1}{n} \left(\sum_{i=1}^n y_i \cdot x_i \right). \end{aligned} \quad (14.5)$$

Система принимает вид

$$\begin{cases} a\bar{x}_B^2 + b\bar{x}_B = \overline{xy}; \\ a\bar{x}_B + b = \bar{y}_B. \end{cases}$$

Решая полученную систему из двух уравнений с двумя неизвестными, получаем

$$a = \frac{\overline{x_i \cdot y_i} - \bar{x}_B \cdot \bar{y}_B}{\bar{x}_B^2 - (\bar{x}_B)^2}, \quad (14.6)$$

$$b = \bar{y}_B - a \cdot \bar{x}_B. \quad (14.7)$$

Итак, знаменатель (14.6) — выборочная дисперсия варианты X , а числитель — выборочный корреляционный момент (ср. с корреляционным моментом):

$$k[X, Y] = \overline{x \cdot y} - \bar{x}_B \cdot \bar{y}_B. \quad (14.8)$$

Корреляционный момент имеет физическую размерность, равную произведению размерностей случайных величин X и Y , т. е. если X измеряется в метрах, а Y — в метрах в секунду, то $k[X, Y]$ измеряется в квадратных метрах в секунду. Естественно, иметь в качестве статистической характеристики размерную величину неудобно (это мы уже отмечали при обсуждении теоретического корреляционного момента).

В отличие от корреляционного момента, коэффициент корреляции является величиной безразмерной. Это относится и к **выборочному коэффициенту корреляции** (ВКК) величин X и Y , который определим соотношением

$$r_{X,Y} = \frac{k[X, Y]}{\sigma_X \cdot \sigma_Y}. \quad (14.9)$$

С использованием этих обозначений получаем уравнение искомой прямой в виде

$$y = r_{XY} \cdot \frac{\sigma_Y}{\sigma_X} \cdot (x - \bar{x}_B) + \bar{y}_B. \quad (14.10)$$

Начальной ординатой будет \bar{y}_B , точка пересечения с осью абсцисс — \bar{x}_B . Угловой коэффициент анализировать сложнее. Однако при близких значениях СКО обеих вариант угловой коэффициент пропорционален ВКК. При нулевом его значении прямая горизонтальна, при положительном/отрицательном соответственно наклон положительный/отрицательный.

Разброс точек относительно прямой регрессии определяется также ВКК и СКО. В простейшем случае исследователь просто визуально рассматривает

расположение облака точек относительно прямой. ВКК достаточно информативен только при значениях, близких по модулю единице (тесная корреляционная зависимость). А вот малые (по модулю) его значения могут быть связаны как со слабой зависимостью, так и с другими особенностями закона распределения изучаемых СВ. Например, нелинейная функциональная зависимость может сопровождаться даже нулевым значением ВКК.

Специализированные статистические пакеты позволяют легко и быстро пересчитать значения выборки. Скажем, перейти к логарифмам вариант — одной или обеих, вместо x_i подставить x_i^2 . Если новый ВКК будет близок единице (по модулю), то тем самым устанавливается другая функциональная форма регрессии — показательная, степенная или иная.

Известны выборочные характеристики, более информативные, чем ВКК, но это уже объект рассмотрения курсов статистики.

14.3. Тема и содержание учебной лабораторной работы

Тема лабораторной работы: «Изучение зависимости между двумя случайными величинами».

Основные задачи:

1. Сбор статистического материала. Необходимо получить выборку двух СВ, связанных с одним ЭСС. Например, температура воздуха и время, необходимое для поездки от дома до университета, или вес и рост студента вашей группы. Объем выборки должен составлять не менее 20 пар. Статистический материал выбирается САМОСТОЯТЕЛЬНО! Рекомендуется подобрать такую пару величин, чтобы а) они были случайными; б) связаны с одним испытанием; в) можно было предполагать некоторую связь между ними.
2. Выполнение первичной обработки выборки: группировка по необходимости, нахождение числовых характеристик, построение диаграммы рассеивания.
3. Вывод уравнения прямой линии регрессии.
4. Вычисление параметров линии регрессии и оценка коэффициента корреляции. Построение линии регрессии на диаграмме рассеивания.
5. Анализ полученных результатов (положение линии регрессии на диаграмме рассеивания, значение коэффициента корреляции).

6. Подготовка отчета по работе (оформление протокола работы, графической части и выводов).

Пример 28. По данной выборке значений случайного вектора [28] изучить возможную зависимость между переменными.

x	y	x	y	x	y	x	y	x	y
8,35	3,50	10,50	6,00	11,35	9,50	12,15	6,00	12,85	9,50
8,74	1,49	10,75	2,50	11,50	6,00	12,25	8,05	13,15	9,02
9,25	6,40	10,76	5,74	11,50	9,00	12,35	5,01	13,25	6,49
9,50	4,50	11,00	8,50	11,62	8,50	12,50	7,03	13,26	10,50
9,75	5,00	11,00	5,26	11,75	10,00	12,76	7,53	13,40	7,51
10,24	7,00	11,25	8,00	12,00	9,00	12,85	6,01	13,50	10,00
13,65	9,50	14,50	10,00	13,75	8,51	14,75	12,00	14,00	11,00
15,25	12,50	14,23	8,40	16,00	11,50	14,26	10,00	16,00	13,00
14,51	9,50	16,25	12,00						

Решение

Найдем оценки параметров линии регрессии, коэффициента корреляции, построим график. Все необходимые для этого формулы были получены: для выборочных средних (14.5), для параметров линии регрессии (14.6) и (14.7), для коэффициента корреляции (ВКК) (14.9). Однако вычисления прямо по этим формулам не очень рациональны, особенно при ручном счете, результирующие формулы для уравнения линии регрессии (14.10) и ВКК (14.9) можно упростить. Вот как это делается:

$$a(x - \bar{x}_B) + \bar{y}_B = \frac{kXY}{D[X]}(x - \bar{x}_B) + \bar{y}_B = \frac{Q_{xy}}{Q_x}(x - \bar{x}_B) + \bar{y}_B; r_{[X,Y]} = \frac{k[XY]}{\sigma_X \sigma_Y} = \frac{Q_{xy}}{\sqrt{Q_x Q_y}},$$

здесь $Q_x = n \cdot D[X]$, $Q_y = n \cdot D[Y]$, $Q_{xy} = n \cdot k[XY]$. Сокращаются общие множители в отношениях и эффективность вычислительного процесса повышается.

Объем выборки $n = 42$. Предварительно вычислим

$$\sum x_i = 522, 23, \quad \sum y_i = 336, 41, \quad \sum x_i^2 = 6652, 25, \quad (14.11)$$

$$\sum y_i^2 = 2987, 80, \quad \sum x_i y_i = 4358, 626. \quad (14.12)$$

Выборочные средние находятся по формулам

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i; \quad \bar{y}_B = \frac{1}{n} \sum_{i=1}^n y_i \Rightarrow \bar{x}_B = 12, 34, \quad \bar{y}_B = 8, 011. \quad (14.13)$$

Затем вычисляем вспомогательные величины:

$$\begin{aligned} Q_x &= \sum (x_i - \bar{x}_B)^2 = \sum x_i^2 - \frac{1}{n} \left(\sum x_i \right)^2; \\ Q_y &= \sum (y_i - \bar{y}_B)^2 = \sum y_i^2 - \frac{1}{n} \left(\sum y_i \right)^2; \\ Q_{xy} &= \sum (x_i - \bar{x}_B)(y_i - \bar{y}_B) = \sum x_i y_i - \frac{1}{n} \sum x_i \sum y_i. \end{aligned} \quad (14.14)$$

По формулам (14.14) находим

$$Q_x = 158,3182, Q_y = 292,5958, Q_{xy} = 175,19. \quad (14.15)$$

Наносим точки на диаграмму рассеивания и строим по уравнению прямую регрессии (рис. 14.2). Для полного анализа кроме визуального рассмотрения взаимного расположения экспериментальных точек и линии регрессии — что тоже важно — необходимо получить выборочный коэффициент корреляции. Окончательно из соотношений (14.14) получаем $r = 0.813$. ■

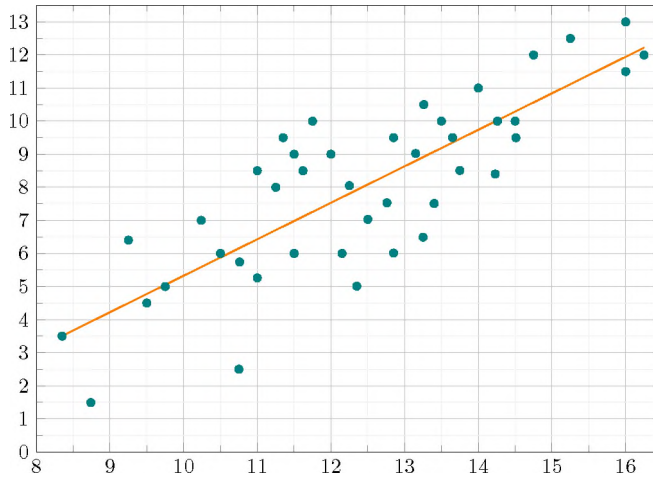


Рис. 14.2. Диаграмма рассеивания и прямая линия регрессии

Можно сделать вывод о довольно тесной корреляционной зависимости между величинами. Предполагая, что в нашем примере именно ξ влияет на η , мы можем, используя уравнение регрессии, количественно оценить это влияние.

Список библиографических ссылок

1. *Kolmogorov A. N.* Grundbegriffe der Wahrscheinlichkeitsrechnung, in Ergebnisse der Mathematik. Berlin, 1933.
2. *Колмогоров А. Н.* Основные понятия теории вероятностей. 2-е изд. М. : Наука, 1974. 120 с.
3. *Вентцель Е. С.* Теория вероятностей. М. : Наука, 1969. 576 с.
4. *Пугачев В. С.* Теория вероятностей и математическая статистика. М. : Наука, 1979. 496 с.
5. *Вентцель Е. С., Обчаров Л. А.* Теория вероятностей и ее инженерные приложения. М. : Наука, 1988. 480 с.
6. *Феллер В.* Введение в теорию вероятностей и ее приложения : в 2 т. : пер. с англ. М. : Изд-во иностр. лит., 1984. Т. 1. 528 с. : ил.
7. *Чистяков В. П.* Курс теории вероятностей. М. : Наука, 1982. 256 с.
8. *Розанов Ю. А.* Теория вероятностей, случайные процессы и математическая статистика : учебник для вузов. М. : Наука, 1989. 320 с.
9. *Гнеденко Б. В.* Курс теории вероятностей. М. : Наука, 1988. 448 с.
10. *Рытов С. М.* Введение в статистическую радиофизику. М. : Наука, 1966. 404 с.
11. Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами : учеб. пособие для студентов высш. техн. учеб. заведений / А. И. Кибзун, Е. Р. Горяинова, А. В. Наумов ; под ред. А. И. Кибзуна. 3-е изд., перераб. и доп. М. : Физматлит, 2007. 231 с. : ил.
12. *Виленкин Н. Я.* Комбинаторика. М. : Наука, 1969.
13. *Попова Т. И., Колесова Д. В.* Поликодовый vs вербальный текст в академическом лекционном дискурсе // Общество. Коммуникация. Образование. 2020. Т. 11, № 3. С. 78–87.
14. Поликодовость как проявление интердискурсивности в испанском научном дискурсе // Портал научных журналов Российского университета дружбы народов. URL: https://journals.rudn.ru/polylinguality/article/view/1776/1249/ru_RU (дата обращения: 03.04.2022).

15. *Krochin A. L., Spitsin S. E., Krokhin L. A.* VISUAL METAPHORS OF MATHEMATICAL OBJECTS AND AUGMENTED REALITY TECHNOLOGY IN HIGH EDUCATION // ФИЗИКА. ТЕХНОЛОГИИ. ИННОВАЦИИ. IX Международная молодежная научная конференция ФТИ-2022 URL: http://fizteh.urfu.ru/fileadmin/user_upload/site_19855/Conference/2022/Tesisy_FTF-2022_Wo1IBN_1.0.pdf, p. 926 (дата обращения: 03.11.2022).
16. *Гуссенс М., Муттельбах Ф., Самарин А.* Путеводитель по пакету \LaTeX и его расширению $\text{\LaTeX}2\epsilon$: пер. с англ. М. : Мир, 1999. 606 с.
17. TeX Users Group : [site]. URL: <https://tug.org/> (date of access: 03.02.2021).
18. How Gravity Assist Work // Space.com: [site]. URL: <https://www.space.com/38226-how-gravity-assists-work-osiris-rex.html> (date of access: 03.02.2021).
19. Voyager : Jet Propulsion Laboratory : [site]. URL: <https://voyager.jpl.nasa.gov/> (date of access: 03.02.2021).
20. Как снайперы поражают цели на дистанции в 2 километра // Яплакал : [сайт]. URL: <https://www.yaplakal.com/forum2/topic2176029.html> (дата обращения: 03.02.2021).
21. Физический энциклопедический словарь / гл. ред. А. М. Прохоров. М. : Большая рос. энцикл., 1995. 928 с.
22. Броуновское движение // Крохин Александр Леонидович : [сайт]. URL: <https://tutor4you/matstat3> (дата обращения: 03.02.2021).
23. Большой энциклопедический словарь : в 2 т. / гл. ред. А. М. Прохоров. М. : Сов. энцикл., 1991. 1456 с.
24. *Ожегов С. И., Шведова Н. Ю.* Толковый словарь русского языка / Рос. акад. наук, Ин-т. рус. яз. им. В. В. Виноградова. 4-е изд., доп. М. : Азбуковник, 1997. 944 с.
25. Simulation of Buffons Needle // Yihui Xie : [site]. URL: <https://yihui.org/animation/example/buffon-needle> (date of access: 03.02.2021).
26. Stars and bars (combinatorics) // Википедия : свободная энциклопедия. URL: [https://en.wikipedia.org/wiki/Stars_and_bars_\(combinatorics\)](https://en.wikipedia.org/wiki/Stars_and_bars_(combinatorics)) (дата обращения: 03.02.2022).
27. *Колмогоров А. Н., Журбенко И. Г., Прохоров А. В.* Введение в теорию вероятностей. М. : Наука, 1982. 160 с.

28. Сборник задач по математике для вузов. Теория вероятностей и математическая статистика : учеб. пособие для вузов/под ред. Э. А. Вуколова, А. В. Ефимова и др. М. : Наука. 1990. 432 с.
29. Алгебра событий // Новосибирский государственный университет : [сайт]. URL: <https://tvims.nsu.ru/chernova/tv/lec/node9.html> (дата обращения: 03.02.2021).
30. Ask Good Questions : [site]. URL: <https://askgoodquestions.blog/2020/06/15/50-which-tire/> (date of access: 03.04.2022).
31. Парадокс мальчика и девочки : Википедия : свободная энциклопедия. https://ru.wikipedia.org/wiki/%D0%9F%D0%B0%D1%80%D0%B0%D0%B4%D0%BE%D0%BA%D1%81_%D0%BC%D0%B0%D0%BB%D1%8C%D1%87%D0%B8%D0%BA%D0%B0_%D0%B8_%D0%B4%D0%B5%D0%B2%D0%BE%D1%87%D0%BA%D0%B8 (дата обращения: 03.04.2022).
32. *По́я Д.* Математика и правдоподобные рассуждения. : пер. с англ. И. А. Вайнштейна ; под ред. С. А. Яновской. 2-е изд., испр. М. : Наука, 1975. 464 с.
33. *Секей Г.* Парадоксы теории вероятностей и математической статистики : пер. с англ. М. : Мир 240 с.: ил.
34. *Bortkiewicz L. von* Das Gesetz der Kleinen Zahlen. Leipzig : B. G. Teubner, 1898. 52 p.
35. Математический энциклопедический словарь/ гл. ред. Ю. В. Прохоров. М. : Сов. энцикл., 1988. 845,[2] с. : ил.
36. *Крохин А. Л.* Принципы и технология математической визуализации : учеб. пособие / А. Л. Крохин. Екатеринбург : Изд-во Урал. ун-та, 2014. 137, [3] с.
37. Яндекс Диск. URL: <https://disk.yandex.ru/i/H6VWFvqtHYAjKA> (дата обращения: 03.04.2022).
38. Компактность представления информации : Википедия : свободная энциклопедия. URL: [http://ru.wikipedia.org/wiki/Ящик с усами](http://ru.wikipedia.org/wiki/Ящик_с_усами) (дата обращения: 03.04.2022).

Учебное издание

Крохин Александр Леонидович

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

С ЭЛЕМЕНТАМИ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ

Учебное пособие

Заведующий редакцией *М. А. Овечкина*

Редактор *Н. В. Чапаева*

Корректор *Н. В. Чапаева*

Компьютерная верстка в \LaTeX *А. Л. Крохина*

Подписано в печать 12.12.2022. Формат 70×100 1/16.
Бумага офсетная. Цифровая печать. Усл. печ. л. 16,12.
Уч.-изд. л. 11,2. Тираж 30 экз. Заказ № 200.

Издательство Уральского университета
Редакционно-издательский отдел ИПЦ УрФУ
620083, Екатеринбург, ул. Тургенева, 4
E-mail: rio@urfu.ru
Тел. +7 (343) 389-94-79, 350-43-28

Отпечатано в Издательско-полиграфическом центре УрФУ
620083, Екатеринбург, ул. Тургенева, 4
Тел. +7 (343) 358-93-06, 350-58-20, 350-90-13
Факс +7 (343) 358-93-06
<http://print.urfu.ru>

Для заметок

Для заметок



КРОХИН АЛЕКСАНДР ЛЕОНИДОВИЧ

Кандидат физико-математических наук, доцент, преподаватель департамента информационных технологий и автоматике ИРИТ — РТФ Уральского федерального университета. Более двадцати лет читает математические курсы в Физико-технологическом институте и Институте радиоэлектроники и информационных технологий УрФУ. Область научных интересов: различные аспекты визуализации и поликодовые тексты. Занимается разработкой мультимедийных учебных курсов и лекционных презентаций с применением математических визуализаций.