

КЛАССИЧЕСКАЯ УЧЕБНАЯ ЛИТЕРАТУРА
ПО МАТЕМАТИКЕ

А. А. СВЕШНИКОВ

ПРИКЛАДНЫЕ МЕТОДЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Под редакцией кандидата физико-математических наук
О. И. ЗАЙЦА

РЕКОМЕНДОВАНО
Учебно-методическим объединением
по университетскому политехническому образованию
в качестве учебника для студентов высших учебных заведений,
обучающихся по направлению подготовки магистров
«Системный анализ и управление»



Санкт-Петербург • Москва • Краснодар
2012

ББК 22.171я73

С 24

Свешников А. А.

С 24 Прикладные методы теории вероятностей: Учебник / Под ред. О. И. Зайца. — СПб.: Издательство «Лань», 2012. — 480 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-1219-8

Учебник известного российского ученого содержит систематическое изложение базового курса теории вероятностей. За основу взята аксиоматика А. Н. Колмогорова, однако использование сложного математического аппарата (теории меры, функционального анализа) сведено к минимуму. Это делает книгу доступной для лиц, владеющих стандартным вузовским курсом высшей математики. Учебник сочетает достаточную строгость изложения с прикладной направленностью. Отразив, в известной мере, современную аксиоматику теории вероятностей, в остальном автор не выходит за рамки того уровня строгости и общности изложения, который характерен для технических вузов. Материал иллюстрируется большим числом примеров и задач.

Учебник адресован, прежде всего, студентам и аспирантам технических вузов с повышенным уровнем математической подготовки, специализирующимся по направлениям «Системный анализ и управление», «Прикладная математика и информатика», может быть также использован научными и инженерно-техническими работниками.

ББК 22.171я73

Рецензенты:

М. П. ГАНИН — доктор технических наук, профессор Военно-морской академии им. Н. Г. Кузнецова; *В. Е. КЛАВДИЕВ* — кандидат технических наук, доцент, зав. кафедрой «Прикладная математика» Санкт-Петербургского государственного политехнического университета.

*Работа по подготовке книги проводилась в рамках программы
«Национальный исследовательский университет»*

Обложка
Н. А. ГОНЧАРОВА

© Издательство «Лань», 2012

© А. А. Свешников, наследники, 2012

© Издательство «Лань»,

художественное оформление, 2012

ПРЕДИСЛОВИЕ РЕДАКТОРА

Предлагаемая читателю книга является одной из последних работ известного российского ученого, заслуженного деятеля науки и техники РСФСР, доктора технических наук, профессора Арама Арутюновича Свешникова (1911–1979). Профессор Свешников является одним из ведущих отечественных специалистов в области прикладных методов теории вероятностей и автором ряда широко известных учебных пособий, неоднократно переиздававшихся как в нашей стране, так и за рубежом¹.

Недавно издательство «Лань» выпустило в свет новое издание хорошо известного задачника [21], ранее выдержавшего два издания на русском языке и переведенного на английский, немецкий, польский и чешский языки. В том же издательстве вышел и новый, ранее не издававшийся курс теории марковских процессов [23]. Обе эти книги с интересом были встречены современным читателем, а их тираж быстро разошелся. В настоящее время издательство «Лань» готовит к выпуску наиболее известную книгу ученого «Прикладные методы теории случайных функций» [22], которая, как и задачник [21], дважды переиздавалась в нашей стране и была переиздана за рубежом.

Все перечисленные выше издания были задуманы автором, как составные части большого курса прикладных вероятностных методов, работа над которым продолжалась много лет. К сожалению, массовый читатель пока лишен возможности познакомиться со всем этим курсом целиком. Базовый для всего изложения курс теории вероятностей [27], [28], [29] был

¹ Краткий обзор научно-педагогической деятельности А. А. Свешникова с соответствующими библиографическими ссылками можно найти в предисловии к книге [23].

издан давно (около тридцати лет назад) небольшим тиражом и в сокращенном виде.

Само издание не было идеальным. Во-первых, объем каждой части был жестко фиксирован. Это привело к сокращениям, деформации текста, нарушению в ряде случаев авторского замысла (например, все примеры набирались петитом, чего сам Свешников никогда не делал). Ряд вопросов (например, аксиоматика теории вероятностей) был урезан до минимума.

Во-вторых, автор не успел просмотреть корректуру книги — техническим редактированием занимались его ученики, которым пришлось согласиться с многочисленными сокращениями. Наконец, ряд вопросов вообще не вошел в пособие, так как планировался к выпуску в следующих, так и не написанных частях (например, предельные теоремы).

По нашему мнению, пришло время переиздать весь курс теории вероятностей [27], [28], [29], устранив отмеченные технические недостатки первого издания и максимально приблизив книгу к первоначальному авторскому замыслу. Мы полагаем, что такая книга будет полезна современному читателю и органично впишется в серию изданий [21], [23], а также планируемое переиздание книги [22].

В предлагаемое издание вошел весь материал пособий [27], [28], [29] (частично с более полных сохранившихся черновиков), часть текста была восстановлена с помощью конспекта лекций.

На наш взгляд, книга представляет хорошо продуманный и тщательно проработанный курс, прошедший длительную апробацию в технических вузах (в частности, в Военно-морской академии и Санкт-Петербургском государственном политехническом университете). По своему содержанию, методике изложения и целевой направленности книга относится к той категории вероятностных курсов, которые в наше время являются дефицитными на книжном рынке (это курсы для технических вузов с повышенным уровнем математической подготовки).

Хотя книги А. А. Свешникова были написаны более тридцати лет назад, они и по сей день остаются востребованными студентами вузов, инженерными и научно-техническими ра-

ботниками. Предлагаемая книга, как мы полагаем, является ценным и необходимым дополнением уже известного научно-педагогического наследия ученого, потому что представляет исходную, базовую часть как для задачника [21], так и для последующих книг серии [22], [23], посвященных вопросам теории случайных процессов.

Мы надеемся, что данное издание окажется столь же полезным широкому читателю, как и преподавателям, студентам кафедры прикладной математики СПбГПУ, где эта книга более тридцати лет используется в учебном процессе.

О. И. Заяц

ВВЕДЕНИЕ

Возникновение теории вероятностей обычно относят к середине XVII века, когда впервые в работах Б. Паскаля², П. Ферма³ и Х. Гюйгенса⁴ были сформулированы правила подсчета вероятностей различных исходов азартных игр. Несмотря на то, что самые первые работы по теории вероятностей были посвящены только азартным играм, авторы этих работ уже тогда указывали, что на примере азартных игр удастся проследить закономерности, являющиеся основой очень интересной и глубокой науки. Таким образом, можно сказать, что первым объектом изучения в теории вероятностей были *случайные события*, причем особого вида, механизм возникновения которых в чем-то аналогичен механизму получения различных результатов в азартных играх.

Для подобных событий характерны три особенности: во-первых, всегда имеется четко сформулированный комплекс контролируемых условий (правила игры), при котором наблюдается исход испытания; во-вторых, имеется еще один, скрытый от игроков, комплекс условий, влияющий на исход испытания, не контролируемый явно в данном опыте, но обладающий известной стабильностью от опыта к опыту; наконец, в-третьих, можно выделить ряд промежуточных исходов испытания, реализацию каждого из которых (вследствие симметрии схемы возникновения самого случайного события) допустимо считать равновозможной. Например, при бросании игрального кубика («кости») падение кубика на любую грань следует признать равновозможным, если кубик не имеет де-

² Паскаль Блез (1623–1662) — французский математик, физик, философ.

³ Ферма Пьер (1601–1665) — французский юрист и математик.

⁴ Гюйгенс Христиан (1629–1695) — нидерландский механик, физик, математик и астроном.

фектов. Аналогичная картина наблюдается и в карточных играх, где появление любой карты из колоды также следует признать равновозможным (если, конечно, колода хорошо перемешана, что предусматривается правилами игры).

Именно для событий подобного типа в середине XVII века в работах Паскаля, Ферма, Гюйгенса и были впервые сформулированы основные правила вычисления особой количественной меры возможности возникновения случайных событий, названной *вероятностью* случайного события.

Дальнейшее развитие теории шло как по линии усложнения моделей исследуемых случайных событий, так и по линии усовершенствования методов расчета их вероятностей. Следующим естественным шагом явилось введение операций над случайными событиями и установление общих правил исчисления вероятностей при этих операциях. Известную логическую завершенность разработанная теория приобрела к началу XVIII века благодаря работам Я. Бернулли⁵. Он ввел и детально изучил схему повторных независимых испытаний, впоследствии получившую его имя, и доказал в этой схеме первую предельную теорему теории вероятностей. Теорема проясняет связь между вероятностью события и частотой его появления при многократном повторении опыта. Бернулли доказал принципиально важный факт: при неограниченном возрастании числа испытаний вероятность любого сколь угодно малого отклонения частоты от соответствующей вероятности бесконечно мала.

Первоначально такое содержание новой науки вполне удовлетворяло потребностям естествознания и практики. Однако постепенно наука и техника поставили ряд совершенно новых задач, выходящих за рамки привычных моделей. Такие задачи возникли, в частности, в геодезии, астрономии, теории стрельбы. Во всех этих приложениях исследователи столкнулись со случайными экспериментами более сложной природы, чем уже известная им модель случайного события. Результат эксперимента вовсе не обязан сводиться к одному из двух возможных исходов (произошло некое событие или нет). Гораздо

⁵ Бернулли Якоб (1654–1705) — швейцарский математик.

чаще целью является получение какой-либо физической величины. Многочисленные исследования показали, что эта величина в целом ряде задач неоднозначно определяется условиями опыта, несмотря на самый тщательный контроль над правильностью его проведения.

Так возник новый объект вероятностного исследования — *случайные величины*. Случайная величина в результате опыта может принять заранее неизвестное значение. Однако его появление не лишено полностью какой-либо математической закономерности. Эта закономерность является более сложной, опосредованной и скрытой от глаз экспериментатора. Теория вероятностей имеет дело только с такими величинами, для которых может быть указана некоторая объективная количественная мера. Существование вероятности в этом случае пришлось постулировать уже независимо от первоначальной схемы, использованной при формулировке основ теории применительно к случайным событиям типа тех, которые происходят в азартных играх.

Понятие случайной величины является в известном смысле обобщением понятия случайного события. Каждому случайному событию можно сопоставить случайную величину, например, равную единице, если событие произошло, и нулю в противном случае. Далее все результаты, которые можно получить в схеме случайных событий, в принципе можно вывести из свойств так построенной случайной величины. Вместе с тем, понятие случайной величины, разумеется, гораздо более содержательно, чем понятие случайного события, и охватывает гораздо более широкий класс задач.

При изучении случайных величин разработанная ранее методология и общие принципы не изменились, но появился новый математический аппарат, необходимый для решения более широкого круга задач. Если первоначальная теория случайных событий обходилась в основном алгеброй и комбинаторикой, то теперь пришлось шире использовать разработанный к тому времени аппарат математического анализа (дифференциальное и интегральное исчисления, линейную алгебру, методы оптимизации, теорию рядов и т. п.). Теория вероятностей из небольшого по объему и изолированного предмета на границе

физики и математики выросла в разветвленную дисциплину, тесно связанную с другими направлениями математики.

Формирование основ теории исследования случайных величин продолжалось весь XVIII век и захватило начало XIX века. Важность этого направления для развития всей математики в целом была осознана многими крупными учеными. В этот период были введены основные числовые и функциональные характеристики случайных величин, получены базовые формулы и теоремы, исследованы важнейшие типовые распределения вероятностей, решено множество актуальных прикладных задач. Свой вклад в развитие теории внесли многие крупнейшие математики того времени, среди которых необходимо выделить П. Лапласа⁶, К. Гаусса⁷, С. Пуассона⁸. В частности, Лапласом и Пуассоном были доказаны первые предельные теоремы в схеме случайных величин, имеющие важное прикладное и методологическое значение.

В современном виде основные классические разделы курса теории вероятностей, касающиеся случайных величин, сформировались в конце XIX — начале XX века. Значительный вклад на этом этапе внесли представители русской математической школы, в частности, В. Я. Буняковский⁹, П. Л. Чебышев¹⁰, А. А. Марков¹¹, А. М. Ляпунов¹², которым удалось поставить и решить ряд принципиально важных для развития общей теории проблем. Это касалось как классических разделов теории (например, предельных теорем), так и новых разделов, впоследствии получивших интенсивное развитие, скажем, теории марковских цепей.

⁶ Лаплас Пьер Симон (1749–1827) — французский математик, физик и астроном.

⁷ Гаусс Карл Фридрих (1777–1855) — немецкий математик, астроном, физик и геодезист.

⁸ Пуассон Симеон Дени (1781–1840) — французский механик, физик и математик.

⁹ Буняковский Виктор Яковлевич (1804–1889) — российский математик.

¹⁰ Чебышев Пафнутий Львович (1821–1894) — российский математик и механик.

¹¹ Марков Андрей Андреевич (1856–1922) — российский математик.

¹² Ляпунов Александр Михайлович (1857–1918) — российский математик и механик.

Дальнейшее расширение объекта исследования теории вероятностей произошло в начале XX века, когда начал разрабатываться еще один новый раздел данной науки — теория *случайных функций*. Необходимость исследования случайных функций впервые возникла у физиков при изучении явлений типа броуновского движения.

Броуновским движением называют непрерывное хаотическое движение мелких частиц, взвешенных в жидкости или газе. Видимые лишь под микроскопом частицы описывают сложные хаотические траектории. Движение не ослабевает со временем, не зависит от химических свойств среды, но зависит от температуры и вязкости среды, а также от размера частиц. Броуновское движение объясняется бесчисленным числом соударений частицы с беспорядочно движущимися молекулами окружающей среды. Каждое отдельное столкновение оказывает на частицу пренебрежимо малое воздействие, но вместе они производят тот эффект, который и обнаруживается при наблюдениях под микроскопом.

Описанное явление было открыто Р. Броуном¹³ в 1827 году. Вероятностная постановка задач естествознания тогда уже была обычным делом, однако математический аппарат исследования подобных явлений в теории вероятностей того времени отсутствовал. Поэтому первые теоретические работы, изучающие случайные функции типа броуновского движения, были выполнены значительно позже, лишь в начале XX века, и принадлежали физикам (в частности, А. Эйнштейну¹⁴, М. Планку¹⁵ и др.). Похожие процессы исследовал и Л. Башелье¹⁶, который использовал их для моделирования курсовой стоимости акций на бирже.

Основы математической теории случайных функций как раздела теории вероятностей были заложены в тридцатых годах XX века. Отечественная наука продолжала занимать здесь значительное, а по ряду направлений и ведущее положение. В этой связи необходимо упомянуть труды А. Н. Кол-

¹³ Броун Роберт (1773–1858) — английский ботаник.

¹⁴ Эйнштейн Альберт (1879–1955) — немецкий физик.

¹⁵ Планк Макс (1858–1947) — немецкий физик.

¹⁶ Башелье Людовик (1870–1946) — французский математик и экономист.

могорова¹⁷, А. Я. Хинчина¹⁸, а также Е. Е. Слуцкого¹⁹. В настоящее время теория случайных функций является наиболее интенсивно развивающимся разделом теории вероятностей, а вся эта теория в целом оказывает мощное влияние на многие другие разделы математики.

Параллельно с усовершенствованием математических методов теории вероятностей увеличивалось число применений вероятностных методов к разнообразным теоретическим и прикладным задачам, связанным с исследованием случайных явлений. При этом, как правило, механизм возникновения случайных явлений отличался от схемы, принятой при формулировке первоначальных, исходных интуитивно ясных положений теории вероятностей (схема равновозможных исходов испытания). Несмотря на это, существование вероятности в рассматриваемых задачах по-прежнему постулировалось, и использовались правила вычисления вероятностей, установленные ранее для более частного вида случайных событий.

В ряде задач (например, в теории ошибок, в молекулярной физике, в теории страхования и т. п.) это применение оказалось успешным и позволило получить ценные теоретические и прикладные результаты. При этом интересно отметить, что, хотя теория вероятностей оперирует с вероятностями, принимающими значения в интервале $[0, 1]$, наибольшую практическую ценность обычно имело выявление событий, для которых вероятность близка к нулю или к единице. В первом случае тем самым устанавливается, что данное событие практически невозможное, во втором — практически достоверное.

Наряду с бесспорными успехами, иногда применение вероятностных методов оказывалось бесплодным или давало результаты, явно противоречащие практике. Например, в XVIII веке и в первой половине XIX века были попытки применения теории вероятностей к определению необходимого числа свидетелей для вынесения правильного приговора в суде и к другим подобным задачам. Несмотря на то, что

¹⁷ Колмогоров Андрей Николаевич (1903–1987) — российский математик.

¹⁸ Хинчин Александр Яковлевич (1894–1959) — российский математик.

¹⁹ Слуцкий Евгений Евгеньевич (1880–1948) — российский математик.

подобными задачами занимались весьма крупные математики (например, Лаплас), никаких практически полезных результатов в этом направлении получить не удалось.

Как очевидно в настоящее время, неуспех применения вероятностных методов в этих случаях объясняется тем, что в подобных задачах само понятие вероятности неприемлемо. Здесь нет количественной меры объективной возможности возникновения случайного события, которую в других задачах определяют как вероятность. Поэтому применение теории вероятностей к подобным, «не вероятностным» задачам не могло дать положительных результатов и в какой-то мере дискредитировало вероятностные методы вообще.

Таким образом, развитие теории вероятностей отличалось от развития других математических дисциплин прежде всего тем, что основные определения и методы теории, сформулированные для частного вида случайных событий, стали затем применяться к более сложным случайным объектам без уточнения условий, при выполнении которых это применение допустимо. Преодоление этой исторически сложившейся математической непоследовательности построения теории вероятностей относится уже к XX веку, когда был принят обычный для математики метод: на основании изучения свойств случайных явлений и рассмотрения простейших схем их возникновения были сформулированы основные аксиомы, из которых вся теория вероятностей выводится путем дедукции без добавочных допущений. Впервые подобный подход к построению теории вероятностей достаточно полно был изложен С. Н. Бернштейном²⁰ в 1917 году и получил свое завершение в работе А. Н. Колмогорова 1933 года, в которой за основу при формулировке аксиом была принята теория множеств.

Формулировка основных аксиом теории вероятностей, как и аксиом в любой математической науке, может быть осуществлена различными способами, если при этом выполняются общие для всякой аксиоматики требования полноты и непротиворечивости. Аксиоматика Колмогорова в настоящее время является общепринятой, поскольку она наиболее удоб-

²⁰ Бернштейн Сергей Натанович (1880–1968) — российский математик.

на для построения теории и связывает теорию вероятностей с другой математической дисциплиной — теорией множеств.

После установления основных аксиом построение теории вероятностей полностью аналогично построению любой математической дисциплины. Однако при использовании вероятностных методов для решения конкретной задачи по-прежнему возникает вопрос, насколько в данном случае подобное применение допустимо, то есть насколько принятые аксиомы отвечают свойствам тех случайных объектов, которые мы хотим исследовать. Этот вопрос является уже не математическим, а физическим, поскольку требует учета свойств различных физических объектов и анализа механизма возникновения их случайности.

Таким образом, имеет место следующая ситуация: теория вероятностей может быть строго построена на основании нескольких принятых аксиом, однако эти аксиомы не обладают универсальностью в том смысле, что имеются случайные явления, для которых они не выполняются, и следовательно, в каждом конкретном случае требуется проверка справедливости этих аксиом. Одним из способов подобной проверки является применение предельных теорем теории вероятностей, которые устанавливают определенные свойства результатов многократного наблюдения. Если результаты опыта противоречат предельным теоремам, то применимость теории вероятностей к данному случаю должна быть поставлена под сомнение.

Первые предельные теоремы были доказаны еще на начальном этапе развития теории вероятностей (теорема Я. Бернулли в 1705 г., теорема Лапласа в 1812 г.), причем и Бернулли, и Лаплас придавали большое значение достигнутым ими результатам именно потому, что эти теоремы позволяли установить связь между экспериментальными данными и теорией.

Очевидно, что каждый раз проверять соответствие результатов наблюдения предельным теоремам для доказательства применимости теории вероятностей практически невозможно, так как в этом случае станет практически невозможным сколько-нибудь широкое применение вероятностных методов. К счастью, необходимость в такой проверке не возникает, поскольку можно указать целые группы явлений, для которых

справедливость основных аксиом теории вероятностей может считаться установленной. К явлениям подобного типа относятся все так называемые массовые явления, в которых речь идет о большом числе примерно одинаковых объектов или о большом числе случайных факторов примерно одного типа: молекулярные явления, ошибки измерения, демография, ошибки систем автоматического управления, шумы в линиях радиопередач и т. п. Тем не менее, всегда при обращении к методам теории вероятностей в новой области исследователь должен критически проанализировать условия задачи для того, чтобы быть уверенным в правомерности подобного применения.

Кроме классического и аксиоматического построения теории вероятностей, Р. Мизесом²¹ был предложен так называемый частотный подход к определению вероятности. Согласно этому подходу вероятность появления случайного события определялась как предел отношения числа появлений событий к числу произошедших опытов n , когда $n \rightarrow \infty$. Несмотря на то, что первое время концепция Мизеса была довольно популярна среди экспериментаторов, она не выдерживает критики с ряда точек зрения и должна быть признана несостоятельной.

В данном курсе принято аксиоматическое построение теории вероятностей, однако для того, чтобы введение аксиом сделать более естественным и понятным читателю, вначале излагается «классическое» построение теории. Аксиомы же вводятся только после предварительного обсуждения свойств частостей событий и являются логическим обобщением экспериментально наблюдаемых свойств этих частостей на абстрактное понятие вероятности.

Курс предназначен для студентов технических вузов с повышенным уровнем математической подготовки, а также специалистов-прикладников, желающих углубить свои знания в области теории вероятностей. Поэтому обозначения и терминология, принятые в настоящее время в публикациях по теории вероятностей, адресованных профессиональным математикам, несколько упрощены. Книга содержит достаточно строгое и полное изложение предмета. Учитывая целевую на-

²¹ Мизес Рихард Эдлер (1883–1958) — немецкий математик и механик.

правленность пособия и предполагаемый контингент читателей, автор старался свести к минимуму использование математического аппарата, выходящего за рамки программы математики для технических вузов (прежде всего, функционального анализа и теории меры).

Книга возникла на основе курса лекций, читавшегося на протяжении ряда лет автором студентам физико-механического факультета Ленинградского политехнического института (ЛПИ). Она была задумана как первая часть трехсеместрового курса прикладных вероятностных методов. В силу ряда причин первоначально была опубликована вторая часть [22], которая выдержала ряд переизданий у нас в стране и за рубежом. Последнюю, третью часть составил прикладной курс теории непрерывных (диффузионных) марковских процессов²².

В настоящей книге при изложении материала сохранена историческая последовательность развития теории вероятностей. Вначале на элементарном уровне рассматривается исчисление вероятностей случайных событий, затем достаточно подробно разбираются свойства случайных величин и, наконец, более сложные модели систем случайных величин. Читатель, желающий продолжить знакомство с предметом и освоить также теорию случайных функций, может обратиться к курсу автора [22]. По целевой направленности, стилю изложения, методике подачи материала (включая обозначения и определения) указанный курс не отличается от настоящего и образует с ним одну серию²³.

²² Эта книга недавно выпущена издательством «Лань» (см. [23]). — *Прим. ред.*

²³ Заключительную часть серии составляет книга [23]. — *Прим. ред.*

ЭЛЕМЕНТАРНАЯ ТЕОРИЯ СЛУЧАЙНЫХ СОБЫТИЙ

§ 1. СООТНОШЕНИЯ МЕЖДУ СЛУЧАЙНЫМИ СОБЫТИЯМИ

Предположим, что имеется определенный комплекс условий, который может быть воспроизведен любое число раз. Осуществление этого комплекса будем называть *испытанием*, а явления, происходящие в результате испытания, — *событиями*.

В качестве испытания можно рассматривать бросание игральной кости или монеты, извлечение карты из колоды, включение в цепь электронного прибора, проверку качества продукции в ОТК и т. п. Если в результате испытания некоторое событие происходит неизбежно, то такое событие называется *достоверным*. Наоборот, если в результате испытания событие не может произойти никогда, то такое событие называется *невозможным*. В данной главе будем обозначать достоверное событие через U , а невозможное — через V .

Пусть, например, испытанием является бросание игральной кости. Достоверное событие при этом — получение какого-либо числа от 1 до 6 на верхней грани, невозможное — получение числа, большего 6.

Если комплекс условий, принятый для данного испытания, определяет исход испытания неоднозначно, то событие, могущее произойти в результате испытания, называется *случайным*. В примере с игральной костью таким событием является получение определенного целого числа из интервала $[1, 6]$.

При включении прибора в цепь случайным событием является отказ прибора или отсутствие отказа.

Для обозначений случайных событий будем использовать большие буквы латинского алфавита: $A, B, C, \dots$, снабжая их иногда индексами.

Говорят, что событие A влечет за собой событие B (или из A следует B), если при наступлении A неизбежно происходит событие B . Это утверждение обозначается $A \subset B$, что означает также, что событие A включено в событие B , является его составной частью.

Пусть, например, событие A означает получение четного числа очков игральной кости, а событие B — получение числа 6. Ясно, что в данном случае $B \subset A$, так как при получении числа 6 условие четности выполнено автоматически.

Допустим, что событие A влечет за собой B , но, в свою очередь, и B всегда влечет A . Ясно, что такие события фактически неразличимы и на деле представляют одно и то же событие. Их называют *равными* (либо *тождественными*, *эквивалентными* или *совпадающими*) и пишут $A = B$. Введение этого понятия позволяет записывать равенство между событиями, получаемыми с помощью алгебраических операций над ними. Таких операций несколько.

Произведением двух событий A и B называется событие C , состоящее в одновременном наступлении обоих событий A и B . Для обозначения произведения событий используется или алгебраический знак умножения, или знак $\cap$, употребляемый в теории множеств для обозначения пересечения множеств, то есть пишут

$$C = A \cdot B = AB = A \cap B. \quad (1.1)$$

Произведение нескольких событий $A_1, A_2, \dots, A_n$ называется событие C , состоящее в одновременном наступлении всех указанных событий. В соответствии с принятыми обозначениями

$$C = \prod_{k=1}^n A_k = A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{n-1} \cap A_n = \bigcap_{k=1}^n A_k. \quad (1.2)$$

Суммой двух событий A и B называется событие C , состоящее в наступлении хотя бы одного из событий A и B . Для обозначения суммы используется или алгебраический знак плюс, или знак $\cup$, применяемый в теории множеств для обозначения объединения множеств

$$C = A + B = A \cup B. \quad (1.3)$$

Для суммы n событий $A_1, A_2, \dots, A_n$, означающей наступление хотя бы одного из совокупности событий $A_k (k = \overline{1, n})$, имеем аналогично

$$C = \sum_{k=1}^n A_k = \bigcup_{k=1}^n A_k. \quad (1.4)$$

Поясним геометрическую интерпретацию введенных операций с помощью так называемых *диаграмм Венна*¹, представленных на рисунке 1.1. Допустим, что испытание заключается в выборе по случайному закону одной единственной точки из заданной области Ω (на рис. 1.1 эта область представляет собой прямоугольник).

События A и B сводятся к попаданию случайной точки в части, отмеченные соответствующей буквой. Штриховка показывает результирующее событие при той или иной операции над заданными событиями. Очевидно, достоверному событию U отвечает весь прямоугольник (рис. 1.1а), а невозможному — полное отсутствие какой-либо штриховки (рис. 1.1б). Сумма событий $A + B$ эквивалентна попаданию случайной точки в область, получаемую наложением областей A и B (рис. 1.1в), а их произведение AB — в общую часть областей A и B (рис. 1.1г). Если событие A влечет за собой событие B , то область A целиком лежит внутри области B (рис. 1.1д).

Рис. 1.1е иллюстрирует понятие *несовместных событий*. События A и B несовместны, если они не могут произойти оба одновременно, то есть выполняется равенство

$$AB = V. \quad (1.5)$$

¹ Венн Джон (1834–1923) — английский логик и математик.

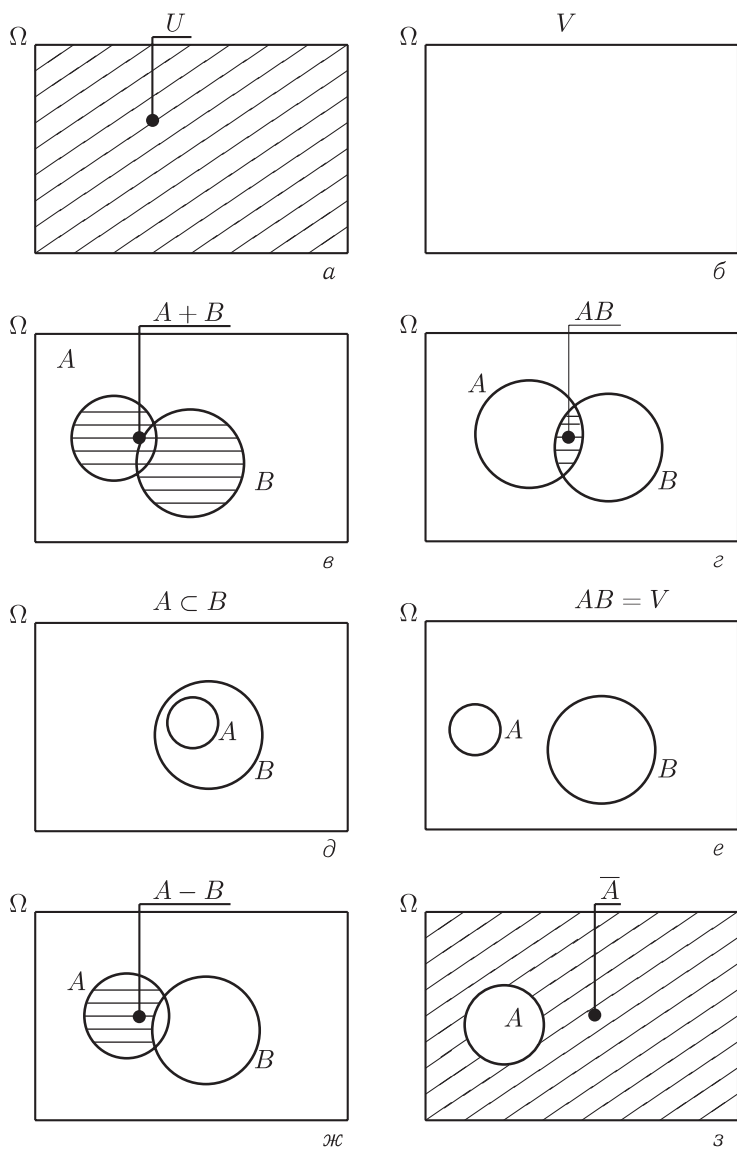


Рис. 1.1

Геометрическая интерпретация операций над событиями

Несовместность событий в известном смысле слова является свойством, противоположным их включению друг в друга. Событие, являющееся частью другого, всегда происходит вместе с ним, но с несовместным событием такое не может произойти никогда.

Пусть, например, A означает получение не менее трех тузов из пяти сданных подряд карт, B — получение двух тузов в следующих пяти картах, сданных из той же колоды, а C — получение при второй сдаче менее двух тузов. Очевидно, события A и B несовместны, и событие A влечет C , хотя и не совпадает с последним (C может наступить и при невыполнении A).

Часто вводят также *разность* событий. Применительно к паре событий A и B этим термином называют событие C , состоящее в том, что A произойдет, а B — нет. Для обозначения разности используют либо знак минус, либо косую черту, принятую в теории множеств для обозначения дополнения множеств (см. рис. 1.1ж),

$$C = A - B = A \setminus B. \quad (1.6)$$

Операция вычисления разности обратна по отношению к суммированию событий. Из выполнения (1.6) следует, что

$$A = B + C = B \cup C. \quad (1.7)$$

Если событие A влечет за собой событие B , то есть $A \subset B$, то тогда

$$AB = A, \quad A + B = B. \quad (1.8)$$

и, в частности, при $B = A$ имеем

$$AA = A, \quad A + A = A. \quad (1.9)$$

Аналогичные равенства выполняются при перемножении или суммировании любого числа совпадающих событий.

Приведенные равенства (1.8) и (1.9) показывают, что свойства операций умножения и сложения применительно к случайным событиям в некоторых случаях могут существенно отличаться от привычных свойств этих операций в алгебре

чисел. Вместе с тем непосредственно из определения суммы и произведения событий следует, что для сложения и умножения событий выполняются обычные *переместительный* и *сочетательный* законы

$$A + B = B + A, \quad AB = BA, \quad (1.10)$$

$$(A + B) + C = A + (B + C), \quad (AB)C = A(BC), \quad (1.11)$$

а также *распределительный* закон умножения по отношению к сложению

$$(A + B)C = AC + BC. \quad (1.12)$$

Операции с участием достоверного и невозможного событий приводят к следующему результату

$$A + U = U, \quad AU = A, \quad A + V = A, \quad AV = V. \quad (1.13)$$

Невозможное событие V при обеих операциях ведет себя как нуль в области чисел. Достоверное событие U при умножении аналогично числовой единице, а при сложении ведет себя особым образом.

Последний из рисунков 1.1 поясняет понятие *противоположного* события. События называются противоположными, если появление одного из них означает непоявление другого. Противоположное событие обозначается одной и той же буквой, что и исходное, но с чертой сверху. Например, $\bar{A}$ — событие, противоположное A . При этом

$$A\bar{A} = V, \quad A + \bar{A} = U. \quad (1.14)$$

Очевидно, противоположные события A и $\bar{A}$ несовместны. Противоположное событие может быть выражено как дополнение исходного события до достоверного

$$\bar{A} = U \setminus A = U - A. \quad (1.15)$$

События $A_1, A_2, \dots, A_n$ образуют *полную группу*, если при каждом испытании обязательно должно произойти хотя бы одно из них, то есть

$$\sum_{k=1}^n A_k = U, \quad (1.16)$$

где U — достоверное событие. Если, например, в каждом опыте может появиться событие A , а A_k — появление события A ровно k раз в серии из n опытов, то события $A_0, A_1, \dots, A_n$ образуют полную группу событий. Полную группу образуют также противоположные события A и $\bar{A}$.

В приведенных примерах полную группу образуют несовместные события. В общем случае события в полной группе могут быть и совместными. Например, события A , B и $\overline{A+B}$ образуют полную группу, хотя A и B могут быть и совместны.

При решении задач часто возникает потребность в упрощении формул, выражающих одни события через другие. Часто эти упрощения могут быть выполнены путем применения к суммам и произведениям событий правил алгебры и простых логических рассуждений.

Пример 1.1. Упростить выражение

$$A = (B + C) \cdot (B + \bar{C}) \cdot (\bar{B} + C).$$

Решение. Так как при любых событиях B и C справедливо равенство $B + BC = B$, то получаем

$$(B + C) \cdot (B + \bar{C}) = B + BC + B\bar{C} + V = B,$$

а тогда

$$A = BC.$$

Подобно разобранным примеру преобразуется и произведение

$$A = \prod_{k=1}^n (B + C_k). \quad (1.17)$$

После перемножения и аналогичных упрощений получим в результате следующее выражение

$$A = B + \prod_{k=1}^n C_k. \quad (1.18)$$

Пример 1.2. Проводится стрельба по мишени, состоящей из десяти кругов, ограниченных концентрическими окружностями с радиусами r_k , ($k = \overline{1, 10}$), занумерованными так, что

$r_1 < r_2 < \dots < r_{10}$. Событие A_k означает попадание при одном выстреле в круг радиуса r_k . Упростить выражение для следующих сложных событий

$$B = \prod_{k=6}^{10} A_k, \quad C = \sum_{k=1}^5 A_k.$$

Решение. Так как появление события A_6 влечет за собой появление всех событий, соответствующих попаданиям в области больших размеров, то $B = A_6$. По аналогичным соображениям событие C отвечает попаданию в круг наибольшего радиуса r_k из всех k , представленных во второй сумме, то есть $C = A_5$.

Используя операции сложения и умножения событий, можно представить сложное событие в виде различных комбинаций более простых событий. Выбор конкретной формы представления зависит от характера решаемой задачи и удобства последующих вычислений.

Пример 1.3. Проводится шахматный турнир. Пусть A_i обозначает выигрыш определенным участником турнира своей партии в i -ом туре, а A — выигрыш хотя бы одной из двух первых партий. Выразить событие A через A_1 и A_2 тремя различными способами: 1) в виде суммы совместных слагаемых; 2) в виде суммы несовместных слагаемых; 3) вычисляя $\bar{A}$ и переходя к противоположному событию.

Решение. Напрашивающийся тривиальный вариант дает $A = A_1 + A_2$. Теперь попытаемся выразить то же событие через сумму несовместных слагаемых. Выиграть хотя бы одну из двух партий можно либо выиграв первую и проиграв вторую ($A_1\bar{A}_2$), либо выиграв вторую и проиграв первую ($\bar{A}_1A_2$), либо добившись успеха в обеих партиях (A_1A_2). Следовательно, $A = A_1\bar{A}_2 + \bar{A}_1A_2 + A_1A_2$. Третий способ исходит из того, что искомое событие A противоположно проигрышу обеих партий, то есть $\bar{A} = \bar{A}_1\bar{A}_2$. Поэтому $A = \overline{\bar{A}_1 \cdot \bar{A}_2}$. Итак, в результате имеем три эквивалентные формы представления интересующего нас события

$$A = A_1 + A_2 = A_1\bar{A}_2 + \bar{A}_1A_2 + A_1A_2 = \overline{\bar{A}_1 \cdot \bar{A}_2}.$$

Далее рассмотрим пример, требующий более сложных логических рассуждений.

Пример 1.4. Определить случайное событие X из уравнения

$$X + A\overline{X} = B,$$

при условии, что $A \subset B$.

Решение. Прежде всего, перепишем исходное равенство в эквивалентном виде

$$X + A = B.$$

Далее заметим, что X должно включать в себя разность $B \setminus A$. Дело в том, что в левую часть последнего равенства входит событие A , которое несовместно с $B \setminus A$, а в правую — само событие B , которое, естественно, включает $B \setminus A$. Равенство события $X + A$ событию B без включения $B \setminus A$ в X невозможно.

Представим поэтому X в виде

$$X = (B \setminus A) + A_0,$$

причем будем подбирать A_0 несовместным с $B \setminus A$. Поскольку A_0 несовместно с $B \setminus A$, то оно может быть включено лишь в сумму $\overline{B} + A$. Включение A_0 в $\overline{B}$ невозможно, так как иначе сумма $X + A$ не равнялась бы B . Следовательно, A_0 должно быть включено в A . Итак, окончательно искомое событие X представляется в виде

$$X = (B \setminus A) + A_0, \quad A_0 \subset A,$$

причем A_0 — произвольное событие, удовлетворяющее указанному включению.

Приведенные примеры показывают, что преобразование сложных событий к наиболее удобному виду иногда требует достаточно длинных и утомительных логических рассуждений. Для сокращения и стандартизации рассуждений используют специальные приемы. Часто оказывается полезным следующее тождество, справедливое для любых событий A и B :

$$\overline{A \cdot B} = \overline{A} + \overline{B}. \quad (1.19)$$

Формулу (1.19) в теории случайных событий называют *формулой инверсии*. В математической логике и теории множеств аналогичную формулу связывают с именем де Моргана².

Справедливость (1.19) доказывается элементарно. Действительно, событие $\overline{AB}$ означает, что A и B не появились совместно, но то же самое означает и сумма $\overline{A} + \overline{B}$, стоящая в правой части равенства. Переходя к противоположным событиям, приходим к эквивалентной форме записи правила инверсии

$$\overline{A + B} = \overline{A} \cdot \overline{B}. \quad (1.20)$$

Обе эти формулы легко обобщаются на произвольное число событий.

Докажем, например, формулу, обобщающую исходную запись теоремы инверсии для двух событий (1.19)

$$\overline{\prod_{k=1}^n A_k} = \sum_{k=1}^n \overline{A_k}, \quad (n \geq 2). \quad (1.21)$$

Для доказательства применим метод математической индукции. База индукции имеется, так как при $n = 2$ формула (1.21) уже доказана. Допустим, что она справедлива при некотором произвольном $n \geq 2$. Рассмотрим произведение, аналогичное (1.21), но содержащее $(n + 1)$ сомножитель A_k . Представим это произведение в виде двух событий, одним из которых является A_{n+1}

$$\prod_{k=1}^{n+1} A_k = A_{n+1} \prod_{k=1}^n A_k, \quad (1.22)$$

вычислим противоположные события от обеих частей и применим в правой части формулу инверсии. В итоге получим

$$\overline{\prod_{k=1}^{n+1} A_k} = \overline{A_{n+1}} + \overline{\prod_{k=1}^n A_k}. \quad (1.23)$$

Заменяя здесь второе слагаемое согласно индуктивному предположению (1.21) приходим к уравнению, получаемому

² Де Морган Августус (1806–1871) — шотландский математик и логик.

из (1.21) заменой n на $(n + 1)$. Следовательно, доказываемое правило инверсии справедливо для любого $n \geq 2$.

Совершенно аналогично обобщается и вторая формула этого правила (1.20), которая приобретает вид

$$\overline{\sum_{k=1}^n A_k} = \prod_{k=1}^n \bar{A}_k, \quad (n \geq 2). \quad (1.24)$$

Оба варианта правила инверсии совершенно равноценны. При преобразовании сложных комбинаций событий их иногда приходится применять последовательно по несколько раз.

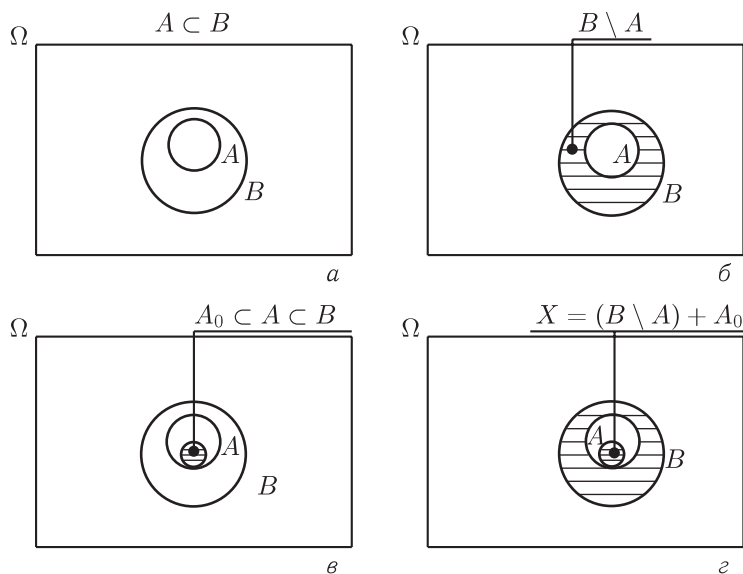


Рис. 1.2
К примерам 1.4 и 1.6

Пример 1.5. Найти событие X из уравнения

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

где A, B — произвольные заданные случайные события.

Решение. Обозначим левую часть последнего равенства через Y и найдем $\overline{Y}$. Применяя формулу инверсии вначале в виде (1.19), а затем в виде (1.20), имеем

$$\begin{aligned}\overline{Y} &= \overline{X + A} + \overline{X + \overline{A}} = \overline{X} \cdot \overline{A} + \overline{X} \cdot A = \\ &= \overline{X} \cdot (\overline{A} + A) = \overline{X} \cdot U = \overline{X}.\end{aligned}$$

По условию $\overline{Y} = \overline{B}$, откуда $X = B$.

Далее рассмотрим продолжение примера 1.4 и докажем прямой подстановкой, что найденное в этом примере решение действительно удовлетворяет исходному уравнению. Конечно, такую проверку правильности решения можно было бы провести непосредственно на основе определений, данных в начале этого параграфа, однако правило инверсии существенно сокращает и упрощает рассуждения, сводя их к стандартной математической схеме.

Пример 1.6. Пусть A, B — заданные случайные события, удовлетворяющие условию $A \subset B$. Доказать, что любое событие вида

$$X = (B \setminus A) + A_0,$$

где $A_0 \subset A$, при подстановке в выражение

$$Y = X + A\overline{X}$$

дает результат $Y = B$.

Решение. Вычисляя по правилу инверсии Y , находим

$$\overline{Y} = \overline{X} \cdot (\overline{A} + X) = \overline{X} \cdot \overline{A} + \overline{X} \cdot X = \overline{X} \cdot \overline{A} + V = \overline{X} \cdot \overline{A}.$$

Теперь вычислим по правилу инверсии $\overline{X} = \overline{(B \setminus A) \cdot \overline{A}_0}$. При соблюдении включения $A \subset B$ событие, противоположное разности $B \setminus A$, есть $\overline{B} + A$. Подставляя это в выражение для $\overline{X}$, а само $\overline{X}$ — в выражение для $\overline{Y}$, получаем

$$\begin{aligned}\overline{Y} &= (\overline{B} + A) \cdot \overline{A}_0 \cdot \overline{A} = \overline{B} \cdot \overline{A}_0 \cdot \overline{A} + A \cdot \overline{A}_0 \cdot \overline{A} = \\ &= \overline{B} \cdot \overline{A}_0 \cdot \overline{A} + V = \overline{B} \cdot \overline{A} \cdot \overline{A}_0.\end{aligned}$$

Вновь производя инверсию, имеем

$$Y = B + A + A_0,$$

причем оба слагаемых A и A_0 включены в B . Применение второй формулы (1.8) дает $Y = B$, каким бы ни было событие A_0 и событие A , что и требовалось доказать.

В заключении заметим, что неизвестное случайное событие может определяться на основе уравнений единственным образом (как в примере 1.5), уравнение может иметь несколько решений (как в примере 1.4) или совсем не иметь решений (если, например, в примере 1.4 включение $A \subset B$ заменить на обратное $B \subset A$).

§ 2. КЛАССИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Рассмотрим случайные события частного вида, для которых понятие вероятности может быть введено элементарно. Такое определение обычно называют *классическим*, поскольку этот способ был использован при первоначальном построении теории вероятностей Ферма, Гюйгенсом, Паскалем. Предположим, что в результате испытания появляется одно из n несовместных и равновозможных событий $E_1, E_2, \dots, E_n$, которые будем называть элементарными событиями или возможными исходами. Эти события образуют полную группу, то есть

$$\sum_{k=1}^n E_k = U, \quad (2.1)$$

а их несовместность означает, что при любых $k \neq j$

$$E_k E_j = V \quad (k = \overline{1, n}; \quad j = \overline{1, n}; \quad k \neq j). \quad (2.2)$$

Равновозможность элементарных событий понимается в том смысле, что, например, вследствие симметрии появление любого из них не имеет объективного преимущества перед появлением любого другого.

Обозначим через m число исходов, при осуществлении каждого из которых происходит событие A . Эти исходы назы-

ваются *благоприятствующими* событию A . Тогда вероятностью $\mathbf{P}(A)$ появления события A при одном испытании будем называть отношение числа m возможных исходов, благоприятствующих событию A , к общему числу n всех возможных исходов, то есть

$$\mathbf{P}(A) = \frac{m}{n}. \quad (2.3)$$

Формула (2.3) соответствует интуитивному представлению о вероятности как о мере объективной возможности появления случайного события A при одном испытании. Когда $m = 0$, событие A невозможно, поскольку оно не может произойти ни при каком испытании. В этом случае вероятность по определению равна нулю

$$\mathbf{P}(V) = 0. \quad (2.4)$$

С увеличением числа m исходов, благоприятствующих событию A , вероятность $\mathbf{P}(A)$ увеличивается, что соответствует повышению объективной возможности появления события A . Если $m = n$, то событие всегда происходит при любом испытании, то есть является достоверным. Вероятность такого события равна единице

$$\mathbf{P}(U) = 1. \quad (2.5)$$

Вероятность $\mathbf{P}(A)$ любого случайного события A больше вероятности невозможного события V и меньше вероятности достоверного события U . Поэтому из приведенного определения вероятности (2.3) следует, что всегда

$$0 \leq \mathbf{P}(A) \leq 1. \quad (2.6)$$

Если $m \ll n$, то $\mathbf{P}(A) = \frac{m}{n} \ll 1$, а событие A при одном испытании может появиться только в редких случаях. Такое событие называется *практически невозможным*. Наоборот, если $n - m \ll n$, то событие A будет появляться практически при каждом испытании. В этом случае $\mathbf{P}(A) = \frac{m}{n} \approx 1$, а событие A называется *практически достоверным*.

Что считать практически невозможным, а что — практически достоверным событием, то есть какими отклонениями вероятности $\mathbf{P}(A)$ от нуля и от единицы можно пренебречь, определяется физическим содержанием рассматриваемой задачи. Например, если событие A — отказ какого-либо прибора, который может быть устранен путем его повторного включения, то вероятность $\mathbf{P}(A) = 0,10$ часто может быть признана малой, и ею можно пренебречь. Если реализация события A влечет за собой катастрофические последствия (взрыв, аварию и т. п.), то даже меньшая на порядок вероятность $\mathbf{P}(A) = 0,01$ должна уже быть признана большой.

Как мы увидим дальше, основной интерес в теории вероятностей представляет определение условий, при которых вероятность события или близка к единице, или близка к нулю, так как в этих случаях можно делать уверенный прогноз ожидаемого результата опыта, несмотря на случайный характер интересующего нас явления.

При расчете вероятности $\mathbf{P}(A)$ случайного события A по формуле (2.3) выбор возможных исходов и числа m благоприятствующих событию A исходов часто представляет достаточно сложную задачу, которую иногда удается упростить, выбрав систему элементарных событий наиболее удачным способом.

Во многих задачах при вычислении вероятности $\mathbf{P}(A)$ по формуле (2.3) приходится пользоваться формулами комбинаторики для различных соединений (сочетаний, размещений, перестановок и т. п.).

Пусть, например, в партии из k изделий доброкачественных l штук. Требуется определить вероятность того, что при выборе наудачу s изделий среди них окажется доброкачественных r штук.

Число способов, которыми можно отобрать из партии в k изделий s любых изделий, равно числу сочетаний C_k^s из k элементов по s . Так как все эти способы объективно равновозможны, то общее число возможных исходов $n = C_k^s = \frac{k!}{s!(k-s)!}$. Из l доброкачественных изделий взять r штук

можно C_l^r способами, а из $k - l$ бракованных изделий взять $s - r$ штук можно C_{k-l}^{s-r} способами. Общее число m благоприятствующих исходов равно произведению числа способов, которыми можно отобрать r доброкачественных изделий, на число способов, которыми можно взять $s - r$ бракованных изделий. Следовательно, $m = C_l^r C_{k-l}^{s-r}$. Искомая вероятность в данном случае будет

$$\mathbf{P}(A) = \frac{m}{n} = \frac{C_l^r C_{k-l}^{s-r}}{C_k^s}. \quad (2.7)$$

О величине вероятности $\mathbf{P}(A)$ случайного события A в какой-то степени можно судить по тому, как часто появляется данное событие в серии независимых испытаний. Предположим, что при одних и тех же условиях произведено N независимых испытаний, в результате которых событие A имело место M раз. Назовем *частотой* $\tilde{\mathbf{P}}(A)$ появления события A в данной серии испытаний отношение M к N , то есть положим³

$$\tilde{\mathbf{P}}(A) = \frac{M}{N}. \quad (2.8)$$

В отличие от вероятности, являющейся постоянным числом для данных условий опыта, частота в разных сериях будет принимать различные случайные значения. Тем не менее между вероятностью, являющейся объективной мерой возможности появления случайного события при каждом испытании, и частотой, являющейся реализацией этой возможности в данной серии испытаний, существует определенное соответствие, отражающее тот факт, что более вероятные события должны появляться, как правило, чаще. Это соответствие будет подробно исследовано при рассмотрении предельных теорем теории вероятностей. Пока отметим только, что для невозможного V и достоверного U события значение частоты совпадает со значением вероятности, так как

$$\tilde{\mathbf{P}}(V) = 0, \quad \tilde{\mathbf{P}}(U) = 1, \quad (2.9)$$

³В прикладной литературе вместо термина «частота» обычно употребляют термин «*частость*».

поскольку при классическом определении вероятности невозможное событие не может произойти ни в одном опыте, а достоверное событие имеет место в каждом опыте.

Существенно отметить, что если для некоторого события A получено, что его частота $\hat{\mathbf{P}}(A) = 0$, это еще не означает, что A — невозможное событие, поскольку нулевое значение частоты может быть получено случайно и при $\mathbf{P}(A) \neq 0$. Аналогично равенство частоты единице еще не доказывает, что мы имеем дело с достоверным событием. Например, если из десяти деталей, взятых для контроля из большого количества однотипных деталей, ни одна не оказалась бракованной, это еще не означает, что вся партия сплошь состоит из доброкачественных деталей. Точно так же, если все выбранные детали оказались бракованными, это еще не значит, что в партии совсем нет доброкачественных деталей.

Рассмотрим примеры на непосредственный подсчет вероятностей.

Пример 2.1. Для моделирования случайного события A , вероятность которого $\mathbf{P}(A) = 0,72$, заготовлено n карточек, на m из которых написана буква A . Какие минимальные числа m и n следует взять, чтобы при случайном извлечении карточки вероятность появления карточки с буквой A равнялась $\mathbf{P}(A)$?

Решение. Пронумеруем мысленно все карточки и будем понимать под элементарным событием E_k появление карточки, помеченной числом k . Полная группа элементарных событий состоит из n событий, из которых только m благоприятствуют появлению карточки с буквой A . Следовательно, вероятность этого события равняется $\frac{m}{n}$. По условию $\frac{m}{n} = \frac{72}{100} = \frac{18}{25}$, то есть минимальное значение $n = 25$, а соответствующее значение $m = 18$.

Пример 2.2. Из партии s деталей, среди которых l доброкачественных, наудачу взято r штук. Какова вероятность $\mathbf{P}(A)$, что все взятые детали окажутся доброкачественными ($r < l$)?

Решение. Из s деталей взять r штук можно различными способами. Пронумеруем эти способы в произвольном порядке и будем понимать под элементарным событием E_k реализацию

k -го способа. Число элементарных событий $n = C_s^r$. Из общего числа элементарных событий благоприятствуют A только те, при которых все r деталей доброкачественные. Число таких событий $m = C_l^r$. Таким образом,

$$P(A) = \frac{C_l^r}{C_s^r} = \frac{l(l-1)\dots(l-r+1)}{s(s-1)\dots(s-r+1)}.$$

Пример 2.3. Куб, все грани которого окрашены, распилен на тысячу кубиков одинакового размера. Полученные мелкие кубики тщательно перемешаны. Определить вероятность p_k того, что наудачу извлеченный кубик будет иметь k окрашенных граней $k = (0, 3)$.

Решение. Всего кубиков $n = 1000$. Мысленно представим себе, как они располагались в исходном кубе до перемешивания. Полностью неокрашены были все внутренние кубики, не выходящие на поверхность. Они образуют меньший куб, вдоль ребра которого укладывается 8 мелких кубиков. Поэтому общее число неокрашенных кубиков $m_0 = 8^3 = 512$.

Вероятность извлечь такой кубик $p_0 = \frac{m_0}{n} = 0,512$. Кубики, окрашенные с одной стороны, примыкают к поверхности исходного куба, но не касаются его ребер. На каждой грани куба их $8^2 = 64$, а общее число для всех шести граней составит $m_1 = 64 \cdot 6 = 384$. Значит, $p_1 \frac{m_1}{n} = 0,384$. Куб имеет 12 ребер, на каждом из которых имеется по 8 кубиков, раскрашенных лишь с двух сторон, так что $m_2 = 8 \cdot 12 = 96$, а

тогда $p_2 = \frac{m_2}{n} = 0,096$. Наконец, раскрашенные с трех сторон кубики лежат только в вершинах исходного куба. Их восемь,

и $m_3 = 8$, $p_3 = \frac{m_3}{n} = 0,008$.

Пример 2.4. Имеется двузначное случайное число N_2 , каждая из цифр которого равновозможна от 0 до 9. Определить вероятность того, что после возведения этого числа в куб обе последние цифры в результате окажутся равными единице.

Решение. Представим N_2 в виде десятичного разложения $N_2 = a + 10b$, причем каждая из цифр числа a и b случайным образом пробегает значения от 0 до 9, и каждое из этих значений равновозможно. Общее число возможных вариантов получающихся чисел $n_2 = 100$, и все они также равновозможны. После возведения в куб имеем

$$N_2^3 = a^3 + 30a^2b + 300ab^2 + 1000b^3.$$

Легко видеть, что на последнюю цифру N_2^3 влияет только значение a . Чтобы N_2^3 оканчивалось единицей, необходимо, чтобы $a = 1$, так как только это значение благоприятствует заданному событию. Для получения второй цифры рассмотрим число

$$\frac{N_2^3 - 1}{10} = 3b + 10(3b^2) + 100b^3,$$

последняя цифра которого будет единицей только при $b = 7$. Таким образом, благоприятствующая нашему событию комбинация цифр a и b оказалась единственной ($a = 1$, $b = 7$, так что $N_2 = 71$, $N_2^3 = 357911$). Следовательно, $m_2 = 1$, а тогда

$$\text{искомая вероятность } p_2 = \frac{m_2}{n_2} = 0,01.$$

Отметим, что полученный результат сохраняет силу для любого k -значного числа N_k , ($k \geq 2$). Действительно, обозначим через n_k общее число вариантов задания таких чисел, а через m_k — число благоприятствующих вариантов. Очевидно, факт наступления интересующего нас события зависит только от двух последних цифр числа. Оставшиеся $k - 2$ первые цифры могут быть произвольными. Поэтому при любом $k \geq 2$ имеем $m_k = m_2 \cdot 10^{k-2}$, $n_k = n_2 \cdot 10^{k-2}$, где общий множитель 10^{k-2} дает число всевозможных способов задания первых $k - 2$

цифр. Следовательно, $p_k = p_2 = \frac{m_2}{n_2} = 0,01$ при любом $k \geq 2$.

Пример 2.5. В очереди перед открытием билетной кассы стоит $2n$ человек, из которых у n человек имеются деньги десятирублевого достоинства и у n человек — деньги пятирублевого достоинства. Определить вероятность $\mathbf{P}(A)$ того, что ни одному человеку не придется ждать сдачи, если первоначально в кассе денег не было и каждый человек покупает один

билет, который стоит пять рублей, а расположение людей в очереди случайно и не зависит от имеющихся у них денег.

Решение. Пронумеруем мысленно всех лиц, стоящих в очереди, присвоив им номера от 1 до $2n$. Так как порядок в очереди случаен, то всего будет $(2n)!$ возможных чередований номеров. Из этого числа нужно отобрать такие числовые расстановки, при которых перед каждым человеком, имеющим десять рублей, число людей, расплатившихся деньгами пятирублевого достоинства, было бы хотя бы на единицу больше числа людей, имевших деньги десятирублевого достоинства. Легко видеть, что определение этого числа расстановок сопряжено с большими трудностями. Однако задача сразу решается, если перейти к другой системе элементарных событий. Заметим, что из каждой расстановки, при которой ни одному покупателю не придется ждать сдачи, можно получить расстановку, при которой лишь один покупатель вынужден будет ждать сдачи. Наоборот, если при некоторой расстановке в очереди один покупатель вынужден ждать сдачи, то, поменяв местами двух человек в очереди, можно получить расстановку, при которой ни одному человеку не придется ждать. Таким образом, между расстановками указанных двух типов имеется взаимно однозначное соответствие и, следовательно, число таких расстановок одинаково. Аналогично можно показать, что число расстановок, при которых k человек вынуждены ждать сдачу, равно числу расстановок, при которых сдачу ожидают $(k - 1)$ человек. Следовательно, если назвать расстановку, при которой ожидать сдачу будут k человек ($k = \overline{0, n}$), расстановкой k -го типа, то все эти расстановки равновозможны и могут быть приняты за элементарные события E_k . Из этих событий только E_0 благоприятствует событию A , то есть $m = 1$, а так как всего таких событий $(n + 1)$, то

$$\mathbf{P}(A) = \frac{1}{n + 1}.$$

Существует и другой способ решения поставленной задачи, основанный на ином выборе элементарных событий E_k . Будем характеризовать состояние очереди с помощью соответствующей перестановки из n пятерок и n десятков. Общее

число вариантов расстановки, очевидно, равняется C_{2n}^n . Как показано в [9], число благоприятствующих расстановок равняется $C_{2n}^n - C_{2n}^{n+1}$. Этот результат получен из геометрических соображений путем сопоставления каждой расстановке некоторой ломаной, заданной в интервале $[0, 2n]$, и анализа свойств таких ломаных. В итоге приходим к тому же результату, что и в первом методе

$$\mathbf{P}(A) = \frac{C_{2n}^n - C_{2n}^{n+1}}{C_{2n}^n} = \frac{1}{n+1}.$$

Приведенный пример показывает, что полная система равно-возможных исходов E_k не является единственной. Рациональный ее выбор способен существенно облегчить решение задачи, а менее удачный — сделать его более трудоемким.

§ 3. ГЕОМЕТРИЧЕСКИЕ ВЕРОЯТНОСТИ

При классическом определении вероятности появление события A предполагается, что число n возможных исходов и число m благоприятствующих событию A исходов конечно. Это определение естественным образом можно обобщить, считая, что число возможных исходов E_k перестает быть конечным, но остается счетным. В этом случае под вероятностью следует понимать предел отношения числа благоприятствующих событию A исходов к общему числу исходов n при $n \rightarrow \infty$. Переход к бесконечному счетному числу исходов позволяет получить обобщение классического определения вероятности (2.3) на ряд случаев, при которых в качестве исхода рассматривается положение точки в некоторой области пространства, размерность которого не играет принципиальной роли. Так как положение точки в пространстве определяется непрерывно изменяющимися значениями ее координат, то в данном случае число возможных исходов не только будет бесконечно, но даже не будет являться счетным. Тем не менее, если предположить, что попадание точки в любые два равновеликие объема заданной области, выбранной в рассматриваемом пространстве, равновозможно, то классическое определение вероятности может быть применено и к этому случаю.

Определяемые таким способом вероятности называют *геометрическими вероятностями*.

Рассмотрим этот вопрос подробнее на примере двумерного пространства. Выберем некоторую плоскую область Ω (рис. 3.1), попадание случайной точки в любые равновеликие части которой равновозможно. Выделим в области Ω подобласть A и будем считать, что попадание случайной точки в эту подобласть эквивалентно реализации события A . Для того чтобы воспользоваться классическим определением вероятности, разобьем область Ω на n равновеликих элементов площадью s_0 каждый. Если площади областей A и Ω соизмеримы, можно так выбрать s_0 , чтобы в области A расположилось целое число m элементов. Если площади не соизмеримы, то этого же можно добиться с любой точностью, неограниченно уменьшая s_0 .

Будем понимать под элементарным исходом E_k попадание случайной точки в элемент с номером k . Тогда, сохраняя классическое определение вероятности для бесконечного (но счетного) числа исходов, получаем в пределе

$$\mathbf{P}(A) = \lim_{s_0 \rightarrow 0} \frac{m}{n} = \lim_{s_0 \rightarrow 0} \frac{ms_0}{ns_0} = \frac{S_A}{S}, \quad (3.1)$$

где через S обозначена площадь всей области Ω , а через S_A — площадь области A .

Приведенные рассуждения не зависят от числа измерений области Ω . Поэтому, если попадание точки в любые равновеликие части выбранной области Ω произвольного числа

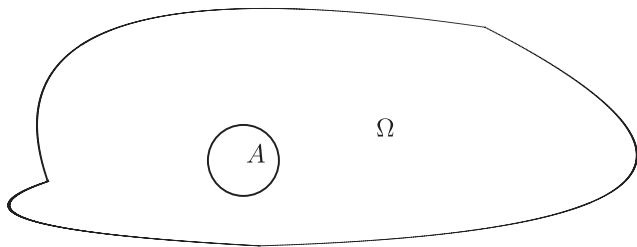


Рис. 3.1

К геометрическому определению вероятности

измерений равновозможно, то вероятность события A будет равна отношению соответствующей меры (то есть длины, площади, объема и т. п.) части области, попадание в которую благоприятствует событию A , к мере всей области Ω .

Геометрический подход к определению вероятности может быть использован и в том случае, когда задача по своей природе не является геометрической, но имеется один или несколько параметров, значения которых равновозможны⁴ для соответствующих интервалов в указанном выше смысле слова. Такие параметры можно рассматривать как координаты случайной точки и свести рассматриваемую задачу к определению вероятности попадания случайной точки внутрь некоторой области, то есть к определению геометрической вероятности. При формулировке задач подобного типа всегда должно быть четко указано, значения каких параметров можно считать равновозможными. Например, при формулировке задач на геометрические вероятности часто говорится о произвольно проводимых прямых, произвольно поставленных точках (см., например, [11]). Без четкого пояснения смысла слова «произвольно» задаче можно придать самое различное математическое содержание и, следовательно, получить совершенно различные ответы. В этом отношении весьма поучительны задачи, известные под названием *парадоксов Бертрана*⁵. Рассмотрим в качестве примера две такие задачи.

Задача 1. Определить вероятность того, что взятая произвольно хорда окружности будет больше стороны вписанного в окружность равностороннего треугольника.

Условие задачи не полно, так как слова «взятая произвольно хорда» допускают различное трактование. Рассмотрим, например, следующие три способа выбора «наудачу» взятой хорды.

Способ 1. Проводим перпендикулярно хорде BB' диаметр MM' (рис. 3.2а) и построим вписанный равносторонний тре-

⁴ В дальнейшем, в том случае, когда значения некоторого параметра X одинаково возможны для любых равновеликих интервалов, выбранных в интервале $[a, b]$, мы будем говорить для краткости, что «значения X равновозможны в интервале $[a, b]$ ».

⁵ Бертран Жозеф Луи Франсуа (1822–1900) — французский математик.

угольник с вершиной в точке M . Точку пересечения диаметра с хордой обозначим через C , а с основанием треугольника — через A . Отразив точку A симметрично относительно центра окружности, получаем точку A' . Далее будем считать произвольной такую хорду, пересечение которой с диаметром MM' равновозможно в любой его точке. Тогда искомая вероятность P_1 сводится к вероятности попадания точки C в промежуток AA' . Следовательно,

$$P_1 = \frac{\overline{AA'}}{\overline{MM'}} = \frac{1}{2}. \quad (3.2)$$

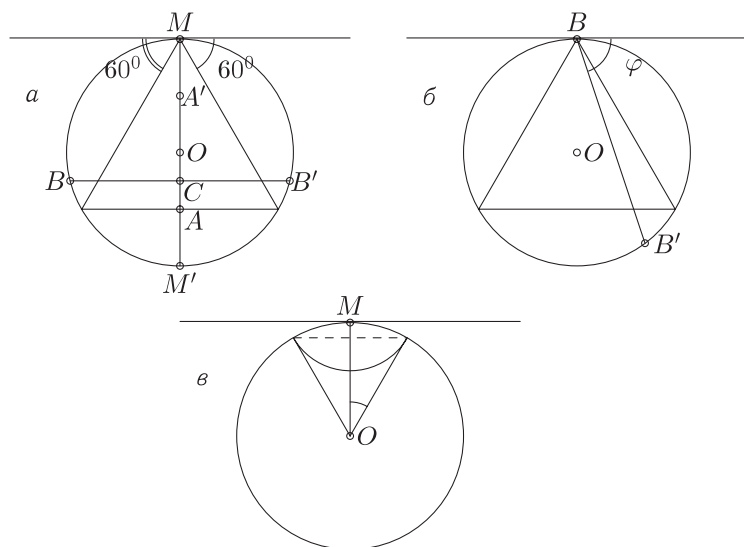


Рис. 3.2
К парадоксам Бертрана

Способ 2. По-прежнему обозначая концы хорды через B и B' , проведем в точке B касательную к окружности и построим равносторонний вписанный треугольник (рис. 3.2б). Будем считать произвольной хорду, образующую с касательной произвольный угол φ , значения которого равновозможны от 0°

до 180^0 . Очевидно, интересующему нас событию благоприятствует попадание φ в интервал от 60^0 до 120^0 . Поэтому искомая вероятность

$$P_2 = \frac{60}{180} = \frac{1}{3}. \quad (3.3)$$

Способ 3. Условимся характеризовать положение хорды, задавая точку A ее пересечения с перпендикуляром OA , опущенным из центра окружности. Произвольной будем считать такую хорду, для которой положение точки A равновозможно внутри окружности (см. рис. 3.2а). Легко видеть, что искомому событию благоприятствует попадание точки A внутрь окружности вдвое меньшего радиуса, чем исходная. Обозначая радиус внешней окружности через r_1 , а внутренней через r_2 , по формуле (3.1) находим

$$P_3 = \frac{\pi r_2^2}{\pi r_1^2} = \left(\frac{r_2}{r_1}\right)^2 = \frac{1}{4}. \quad (3.4)$$

Подчеркнем, что возможны и другие ответы в зависимости от того, как именно понимается произвольность поведения хорды. Мы ограничились приведенными классическими вариантами, принадлежащими Берtrandу.

Задача 2. Определить вероятность того, что угловая величина меньшей дуги большего круга между выбранными наудачу на поверхности шара двумя точками M и M_1 будет меньше α .

Вследствие симметрии задачи можно фиксировать положение одной из точек, например, точки M (рис. 3.2б). Тогда условию задачи будут удовлетворять точки M_1 , попадающие внутрь телесного угла, ограниченного по азимуту величиной α

$$\omega = \int_0^{2\pi} \int_0^\alpha \sin \vartheta d\vartheta d\varphi = 2\pi(1 - \cos \alpha), \quad (3.5)$$

а искомая вероятность

$$P_1 = \frac{\omega}{4\pi} = \frac{1}{2}(1 - \cos \alpha) = \sin^2 \frac{\alpha}{2}. \quad (3.6)$$

Возможно и иное толкование условия задачи: будем считать выбранными наудачу точками M и M_1 две точки, случайно расположенные на дуге произвольно выбранного большого круга. В этом случае искомая вероятность $P_2 = \frac{2\alpha}{2\pi} = \frac{\alpha}{\pi}$. Полученные значения вероятности резко отличаются друг от друга, что особенно заметно при малых α , так как в этом случае асимптотически

$$\frac{P_1}{P_2} = \frac{\sin^2 \frac{\alpha}{2}}{\frac{1}{\pi} \alpha} = \frac{\pi}{4} \alpha [1 + o(\alpha)], \quad (3.7)$$

то есть P_1 и P_2 имеют даже разный порядок малости, относительно α .

Парадоксам Бертрана впервые дал строгое объяснение А. Пуанкаре⁶. Он показал, что предлагаемые в них варианты вычисления вероятности относятся не к одной и той же задаче, а к совершенно разным задачам. Схожесть этих задач чисто внешняя, а исследуемые случайные события не эквивалентны друг другу. Источник парадокса скрыт нечеткой формулировкой исходных допущений. Обратимся, скажем, к задаче 1. Положение хорды в круге определяется заданием двух параметров, но выбрать их можно по-разному. В первом случае задавалось направление нормали к хорде и расстояние до нее от центра; во втором — один из концов хорды и угол ее поворота относительно него; в третьем — декартовы координаты основания перпендикуляра, опущенного из центра к хорде. Между любыми двумя парами параметров положения существует взаимно однозначная связь. Если считать, что какая-либо одна пара параметров равновозможна в соответствующей области, то это однозначно определяет случайный разброс всех остальных параметров положения. Для других способов выбора параметров положения равновозможность нарушается, и пользоваться геометрическим определением вероятности (3.1) будет уже нельзя.

⁶ Пуанкаре Анри (1854–1912) — французский математик, физик, астроном и философ.

Рассмотрим теперь ряд примеров, при решении которых используется геометрическое определение вероятности. Начнем с классической задачи, рассмотренной еще в начале XVIII века Ж. Бюффеном⁷.

Пример 3.1 (Задача Бюффона). На плоскость, покрытую сеткой параллельных прямых, бросается «игла» (отрезок прямой) длиной l . Определить вероятность того, что игла пересечет хотя бы одну прямую, если расстояние между соседними прямыми равно a (рис. 3.3).

Решение. В соответствии со смыслом задачи, расстояние x от центра иглы O до ближайшей прямой, расположенной снизу, и угол φ , образованный иглой с направлением прямых, равновозможны в интервалах $[0, a]$ и $[0, \pi/2]$ соответственно. Следовательно, их можно рассматривать как прямоугольные координаты случайной точки, поставленной внутри прямоугольника со сторонами a и $\pi/2$ (рис. 3.3), и для решения задачи применять метод геометрической вероятности. Для этого определим область A изменения координат x и φ случайной точки, соответствующей пересечению иглой хотя бы одной прямой. Очевидно, что пересечения не будет, если выполняются неравенства

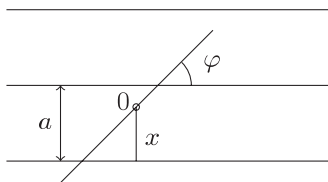


Рис. 3.3
К задаче Бюффона

$$x + \frac{1}{2} \cdot l \sin \varphi \leq a, \quad x - \frac{1}{2} \cdot l \sin \varphi \geq 0,$$

то есть границы указанной области определяются условиями:

$$\frac{1}{2} \cdot l \sin \varphi \leq x \leq a - \frac{1}{2} \cdot l \sin \varphi.$$

Вид области A будет различен при $l \leq a$ и $l > a$. В первом случае область A будет состоять из двух частей, как это

⁷ Бюффон Жорж Луи Леклерк (1707–1788) — французский естествоиспытатель.

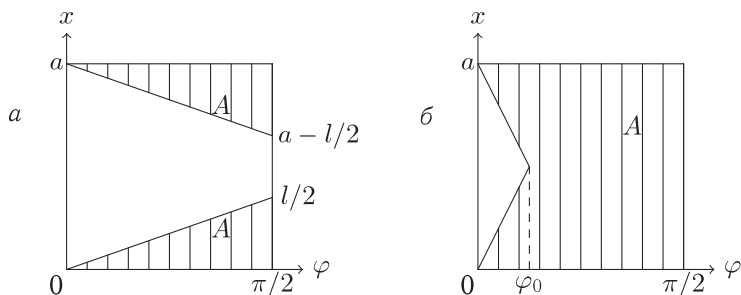


Рис. 3.4
Область допустимых значений параметров
в задаче Бюффона

изображено на рис. 3.4а, а для площади области A получим

$$S_A = l \int_0^{\pi/2} \sin \varphi d\varphi = l. \quad (3.8)$$

Для второго случая область A имеет вид, изображенный на рис. 3.4б и площадь S_A определится равенством ($\sin \varphi_0 = a/l$):

$$S_A = l \int_0^{\varphi_0} \sin \varphi d\varphi + \left(\frac{\pi}{2} - \varphi_0\right) a = \left(\frac{\pi}{2} - \varphi_0\right) a + l(1 - \cos \varphi_0).$$

Таким образом, для искомой вероятности получим

$$\mathbf{P}(A) = \frac{S_A}{\frac{\pi a}{2}} = \begin{cases} \frac{2l}{\pi a}, & l \leq a, \\ \left(1 - \frac{2\varphi_0}{\pi}\right) + \frac{2l(1 - \cos \varphi_0)}{\pi a}, & \sin \varphi_0 = \frac{a}{l}, \quad l \geq a. \end{cases}$$

Особенно простой ответ будет при $l = a$, когда

$$\mathbf{P}(A) = \frac{2}{\pi}.$$

Пример 3.2 (Задача о встрече). Два человека условились встретиться в 12 часов дня в определенном месте. Пришедший первым ждет второго не более 20 минут, после чего уходит. Определить вероятность встречи (событие A), если приход каждого равновозможен в интервале 12–13 часов.

Решение. Обозначим через t_1 время прихода первого, а через t_2 — время прихода второго. Тогда пару чисел (t_1, t_2) можно рассматривать как прямоугольные координаты случайной точки, попадание которой равновозможно в любое место квадрата со стороной, равной единице (1 час). Следовательно, можно применить метод геометрических вероятностей. Будем отсчитывать t_1 и t_2 от предполагаемого

момента встречи (12 часов), построим квадрат со стороной, равной единице, и выделим в этом квадрате область A , соответствующую тем значениям t_1 и t_2 , при которых встреча происходит. Так как для осуществления встречи должно выполняться неравенство $|t_1 - t_2| \leq 1/3$, то указанной областью будет шестиугольник, заштрихованный на рис. 3.5. Площадь этого шестиугольника $S_A = 1 - (2/3)^2$. Следовательно, искомая вероятность

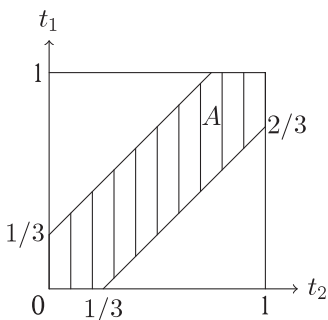


Рис. 3.5

Область допустимых значений параметров в задаче о встрече

$$P(A) = \frac{1 - (2/3)^2}{1^2} = \frac{5}{9}.$$

Задачи, подобные рассмотренной задаче о встрече, могут иметь различное физическое содержание. Например, к задаче подобного типа сводится задача обнаружения передатчика, работающего по неизвестной частоте, и ряд других задач.

Пример 3.3 (Задача о составлении треугольника). От каждого из трех отрезков одинаковой длины l «отламываются» отрезки произвольной длины. Какова вероятность, что из

полученных таким образом трех отрезков можно будет составить треугольник?

Решение. Согласно условию задачи координаты x, y, z полученных отрезков независимы и равновозможны в интервале $[0, l]$. Следовательно, их можно принять в качестве координат случайной точки, положение которой равновозможно внутри куба, ребро которого равно l . Область A , благоприятствующая указанному в условии примера событию, определяется неравенствами:

$$x + y \geq z, \quad y + z \geq x, \quad z + x \geq y.$$

Каждое из неравенств определяет плоскость, отсекающую от куба пирамиду, высотой которой служит отрезок координатной оси длиной l , а основанием — прямоугольный треугольник с катетами длины l (рис. 3.6). Объем каждой из этих пирамид равен $\frac{1}{6} l^3$. Следовательно, объем области, благоприятствующей событию A , равен

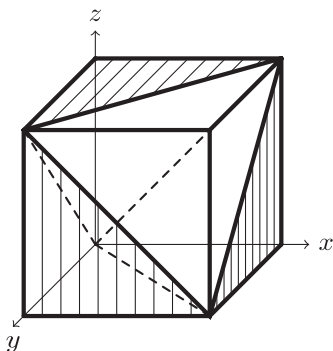


Рис. 3.6

Область допустимых значений параметров в задаче о составлении треугольника

$$l^3 - 3 \cdot \frac{1}{6} l^3 = \frac{1}{2} l^3,$$

а искомая вероятность

$$\mathbf{P}(A) = \frac{\frac{1}{2} l^3}{l^3} = \frac{1}{2}.$$

Пример 3.4. В круг вписан правильный n -угольник ($n \geq 3$). В тот же круг наудачу бросается точка. Определить вероятность p_n того, что эта точка окажется вне многоугольника.

Решение. По условию координаты точки равновозможны в круге. Обозначим его радиус через r , площадь через S , а

площадь вписанного многоугольника через S_n . Искомая вероятность

$$\mathbf{P}(A) = \frac{S - S_n}{S} = 1 - \frac{S_n}{S},$$

так что задача сводится к вычислению площади S_n .

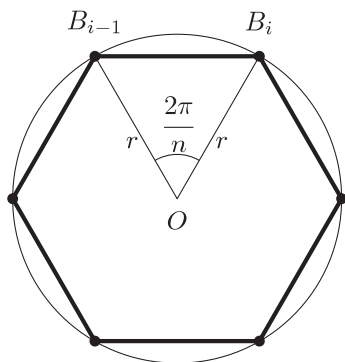


Рис. 3.7

К вычислению площади
правильного n -угольника

Разобьем многоугольник на n одинаковых равнобедренных треугольников с общей вершиной в центре круга (рис. 3.7). Каждый такой треугольник имеет угол $\frac{2\pi}{n}$ между боковыми сторонами длиной r . Следовательно, площадь каждого треугольника

$$S_0 = \frac{r^2}{2} \sin \frac{2\pi}{n},$$

а общая площадь многоугольника $S_n = nS_0$. Поскольку площадь круга $S = \pi r^2$, найдем отсюда

$$p_n = 1 - \frac{1}{\pi r^2} \cdot \frac{n \cdot r^2}{2} \sin \frac{2\pi}{n} = 1 - \frac{n}{2\pi} \sin \frac{2\pi}{n}.$$

Для конкретных числовых значений n имеем

$$p_3 = 1 - \frac{3\sqrt{3}}{4\pi} = 0,587, \quad p_4 = 1 - \frac{2}{\pi} = 0,363,$$

$$p_6 = 1 - \frac{3\sqrt{3}}{2\pi} = 0,173.$$

При больших n эта вероятность убывает обратно пропорционально n^2 .

Пример 3.5. Генератор стандартных случайных чисел выдает числовую последовательность, все элементы которой равновозможны в промежутке $[0, 1]$, причем каждое новое число генерируется независимо от всех предыдущих. Какова вероятность того, что произведение двух таких чисел будет

не меньше, чем θ , но при этом их сумма не превзойдет $1 + \theta$, ($0 \leq \theta \leq 1$).

Решение. Пусть x и y — выданные генератором числа. По условию двукратный запуск генератора эквивалентен случайному эксперименту, в котором случайная точка с координатами (x, y) наудачу выбирается из единичного квадрата

$$\Omega = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

имеющего единичную площадь. Благоприятствующая событию A область значений (x, y) определяется неравенствами

$$xy \geq \theta, \quad x + y \leq 1 + \theta$$

и показана на рис. 3.8. Сверху эта область ограничена прямой $y = 1 + \theta - x$, а снизу — гиперболой $y = \theta/x$. Площадь сегмента, ограниченного указанными границами вычисляется в виде следующего интеграла по промежутку $[\theta, 1]$

$$S_A = \int_{\theta}^1 \left[(1 + \theta - x) - \frac{\theta}{x} \right] dx = (1 + \theta)(1 - \theta) - \frac{1}{2}(1 - \theta^2) + \theta \ln \theta.$$

С учетом того, что $S = 1$, после приведения подобных членов находим

$$p = \mathbf{P}(A) = \frac{1}{2}(1 - \theta^2) + \theta \ln \theta.$$

Функция $p(\theta)$ монотонно убывает с ростом θ от значения $p(0) = 1/2$ до нуля при $\theta = 1$. Например, при $\theta = 1/2$ имеем

$$p(1/2) = 3/8 - 1/2 \ln 2 = 0,0284.$$

Таким образом, за исключением тривиального случая $\theta = 0$ искомая вероятность всегда меньше 50%.

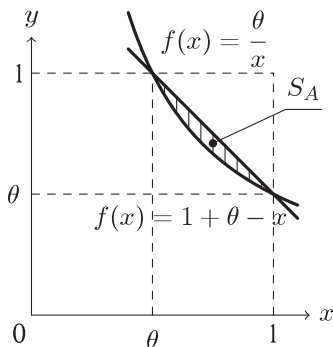


Рис. 3.8
К вычислению площади сегмента

Пример 3.6. Генератор из примера 3.5. последовательно запускается n раз. Найти вероятность того, что полученная последовательность случайных чисел окажется неубывающей.

Решение. Обозначим i -е случайное число через x_i и сформируем из них вектор $\vec{x} = \{x_1, \dots, x_n\}$. Значения $\vec{x}$ равновозможны в n -мерном единичном кубе:

$$\Omega = \{\vec{x} : 0 \leq x_i \leq 1, \quad i = \overline{1, n}\}.$$

Объем этого куба $S = 1$.

Интересующее нас событие A эквивалентно попаданию точки $\vec{x}$ в область

$$\Omega_A = \{\vec{x} : 0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1\}.$$

Объем данной области дается интегралом

$$S_A = \int_0^1 \int_0^{x_n} \int_0^{x_{n-1}} \dots \int_0^{x_2} dx_1 \dots dx_{n-2} dx_{n-1} dx_n,$$

а искомая вероятность

$$\mathbf{P}(A) = \frac{S_A}{S} = S_A$$

совпадает со значением выписанного повторного интеграла.

Для его вычисления введем последовательность функций

$$\varphi_0(x) = 1, \quad \varphi_n(x) = \int_0^x \varphi_{n-1}(y) dy, \quad (n \geq 1).$$

Воспользовавшись методом математической индукции, нетрудно показать, что

$$\varphi_n(x) = \frac{x^n}{n!}, \quad (n = \overline{0, \infty}).$$

Легко видеть, что

$$S_A = \varphi_n(1),$$

следовательно, вероятность неубывания построенной последовательности случайных чисел

$$\mathbf{P}(A) = \frac{1}{n!}$$

и быстро уменьшается с ростом n .

§ 4. УСЛОВНАЯ ВЕРОЯТНОСТЬ. ТЕОРЕМА УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ. НЕЗАВИСИМОСТЬ

Продолжим рассмотрение основных положений теории вероятностей, исходя пока из классического определения вероятности (или эквивалентного ему геометрического определения). Введем новое понятие: будем называть *условной вероятностью* $\mathbf{P}(A|B)$ события A вероятность этого события, вычисленную в предположении, что событие B уже произошло. Естественно, понятие условной вероятности имеет смысл только в том случае, если $\mathbf{P}(B) \neq 0$, то есть событие B не является невозможным. В том случае, когда $\mathbf{P}(A|B) = \mathbf{P}(A)$, то есть вероятность события A не зависит от того, произошло или не произошло событие B , говорят, что событие A *не зависит* от события B . Если же $\mathbf{P}(A|B) \neq \mathbf{P}(A)$, то события A и B называют *зависимыми*.

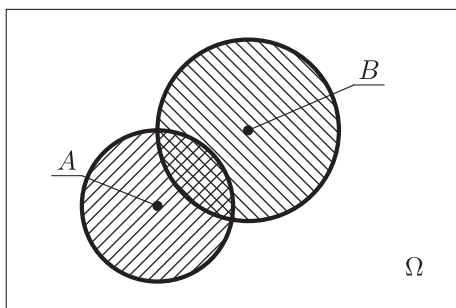


Рис. 4.1

К вычислению условных вероятностей

Установим соотношения между вероятностью $\mathbf{P}(AB)$ произведения двух событий, условными вероятностями $\mathbf{P}(A|B)$, $\mathbf{P}(B|A)$ и ранее введенными вероятностями $\mathbf{P}(A)$, $\mathbf{P}(B)$, которые для их отличия от условных вероятностей иногда называют *безусловными*. Воспользуемся геометрическим определением вероятности.

Предположим, что испытание состоит в выборе наудачу точки внутри некоторой области Ω , геометрическая мера которой равна S , понимая под словом «наудачу», что попадание точки в любую равновеликую часть области Ω равновозможно. Будем считать для наглядности область двумерной (см. рис. 4.1), хотя все последующие рассуждения не зависят от числа измерений. Выделим в области Ω две области — A и B и будем считать попадание случайной точки в область A событием A , а попадание в область B — событием B . Обозначим геометрическую меру (в данном случае — площадь) областей A , B и области AB , являющейся пересечением областей A и B (заштрихована на рисунке), через S_A , S_B и S_{AB} соответственно. Тогда вероятности событий A , B и AB определяются формулами:

$$\mathbf{P}(A) = \frac{S_A}{S}, \quad \mathbf{P}(B) = \frac{S_B}{S}, \quad \mathbf{P}(AB) = \frac{S_{AB}}{S}. \quad (4.1)$$

Если событие B произошло, то наудачу выбранная точка лежит в области B . При этом вероятность события A равно отношению площади S_{AB} , благоприятствующей появлению события AB , к площади S_B , благоприятствующей появлению события B (считаем $S_B \neq 0$). Поэтому условная вероятность события A определяется формулой

$$\mathbf{P}(A|B) = \frac{S_{AB}}{S_B}. \quad (4.2)$$

Аналогично для условной вероятности $\mathbf{P}(B|A)$ события B получим выражение ($S_A \neq 0$)

$$\mathbf{P}(B|A) = \frac{S_{AB}}{S_A}. \quad (4.3)$$

Событие A не зависит от B , если $\mathbf{P}(A|B) = \mathbf{P}(A)$, то есть при $S_{AB}/S_B = S_A/S$. В противном случае событие A зависит от события B . Для независимости события B от A должно выполняться условие $S_{AB}/S_A = S_B/S$, совпадающее с условием независимости события A от B . Следовательно, если событие A не зависит от события B , то и событие B не зависит от A . Иначе говоря, два события могут быть только взаимно зависимыми или взаимно независимыми.

Связь между вероятностью произведения событий, условными и безусловными вероятностями этих событий при классическом построении теории вероятностей устанавливается *теоремой умножения вероятностей*, согласно которой вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое из этих событий произошло.

Для вывода этой теоремы запишем вероятность $\mathbf{P}(AB)$ произведения событий A и B в виде ($S_A \neq 0$, $S_B \neq 0$)

$$\mathbf{P}(AB) = \frac{S_{AB}}{S} = \frac{S_A}{S} \cdot \frac{S_{AB}}{S_A} = \frac{S_B}{S} \cdot \frac{S_{AB}}{S_B}.$$

Учитывая равенства (4.1)–(4.3), для вероятности $\mathbf{P}(AB)$ получаем

$$\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B|A) = \mathbf{P}(B)\mathbf{P}(A|B), \quad (4.4)$$

что и требовалось доказать.

Для взаимно независимых событий одновременно $\mathbf{P}(A|B) = \mathbf{P}(A)$ и при этом $\mathbf{P}(B|A) = \mathbf{P}(B)$. В этом случае теорема умножения вероятностей (4.4) принимает наиболее простой вид

$$\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B), \quad (4.5)$$

то есть вероятность произведения двух независимых событий равняется произведению вероятностей этих событий.

Формулу (4.4) можно, разумеется, получить, исходя не только из геометрического, но и из классического определения вероятности. Пусть из общего числа n равновозможных, единственно возможных и несовместных исходов $E_k (k = \overline{1, n})$

событию A благоприятствуют m исходов, а событию AB — l исходов. В этом случае

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

Используя тождество $\frac{l}{n} = \frac{m}{n} \cdot \frac{l}{m}$, получаем $\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B|A)$, то есть прежнюю формулу (4.4).

Формула (4.4) может быть обобщена на любое число событий. Действительно, например, для трех событий A , B и C имеем

$$\mathbf{P}(ABC) = \mathbf{P}(AB)\mathbf{P}(C|AB) = \mathbf{P}(A)\mathbf{P}(B|A)\mathbf{P}(C|AB). \quad (4.6)$$

Рассуждая аналогично (4.6), получаем, что для вероятности произведения любого числа n событий $A_1, \dots, A_n$ справедлива формула, обобщающая (4.5)

$$\mathbf{P}\left(\prod_{k=1}^n A_k\right) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1A_2)\dots \dots \mathbf{P}(A_n \left| \prod_{k=1}^{n-1} A_k \right.). \quad (4.7)$$

Совокупность произвольного числа $n \geq 2$ случайных событий, которые относятся к одному и тому же опыту (эксперименту), называют *системой случайных событий*. Формула (4.7) выражает теорему умножения вероятностей для произвольной такой системы. Вероятность произведения произвольного числа событий равняется произведению вероятности одного из них на условные вероятности остальных событий. При этом условная вероятность i -го по счету сомножителя в произведении вычисляется при условии, что все события с меньшими номерами произошли.

Как мы видели, для двух событий теорема умножения вероятностей существенно упрощается, если события независимы. Для систем с большим числом событий определение независимости нескольких событий нуждается в уточнении, а сама эта независимость может пониматься в различном смысле слова.

Обратимся к системе n событий $A_1, \dots, A_n$. Эта система включает $\frac{n(n+1)}{2}$ подсистем вида A_i, A_j , состоящих из двух событий ($i = \overline{1, n-1}, j = \overline{i, n}$). Вопрос о наличии зависимости между каждой парой событий можно решить с помощью данного ранее определения. Если любые два события системы независимы, то события $A_1, \dots, A_n$ называют *попарно независимыми*.

Рассмотрим условную вероятность $\mathbf{P}(A_i | \prod_{s=1}^k A_i)$. Если она совпадает с соответствующей безусловной вероятностью, то есть

$$\mathbf{P}(A_i | \prod_{s=1}^k A_{i_s}) = \mathbf{P}(A_i), \quad (4.8)$$

то будем говорить, что *событие A_i не зависит от группы событий $A_{i_1}, \dots, A_{i_k}$* . Теперь предположим, что в системе $\{A_i\}_{i=1}^n$ событие A_2 не зависит от A_1 , A_3 не зависит от A_1 и A_2 и вообще при любом $2 \leq i \leq n$ событие A_i не зависит от $\{A_j\}_{j=1}^{i-1}$. Тогда, заменяя в формуле (4.7) все условные вероятности согласно (4.8) на безусловные, приводим (4.7) к виду, аналогичному (4.5)

$$\mathbf{P}(\prod_{k=1}^n A_k) = \prod_{k=1}^n \mathbf{P}(A_k). \quad (4.9)$$

Если для системы событий выполняется равенство (4.9), то события $A_1, \dots, A_n$ называют *взаимно независимыми* (иногда говорят просто «*независимыми*»). Для системы независимых событий вероятность их произведения равна произведению вероятностей перемножаемых событий для любого $n \geq 2$. Отметим, что для взаимной независимости n событий при $n > 2$ уже не достаточно их попарной независимости, как это имело место при $n = 2$ (см. далее пример 4.4).

Наиболее употребительное достаточное условие взаимной независимости трех и более событий таково. Будем называть события $A_1, \dots, A_n$ *независимыми в совокупности*, если любая условная вероятность вида (4.8) при произвольном

$1 \leq i \leq n$, произвольном $1 \leq k \leq n - 1$ и произвольных комбинациях индексов $1 \leq i_1 < i_2 < \dots < i_k \leq n$ равняется безусловной вероятности $\mathbf{P}(A_i)$. Очевидно, для событий, независимых в совокупности, (4.9) всегда выполняется. Для того чтобы убедиться в этом, достаточно выразить условные вероятности, входящие в (4.7) с помощью (4.8), положив $k = i - 1$, $i_s = s$.

Понятие независимости в совокупности иногда определяют несколько иначе. Говорят, что этот вид независимости имеет место, если

$$\mathbf{P}\left(\prod_{s=1}^k A_{i_s}\right) = \prod_{s=1}^k \mathbf{P}(A_{i_s}) \quad (4.10)$$

для любого $2 \leq k \leq n$ и любых комбинаций индексов $1 \leq i_1 < \dots < i_k \leq n$. Очевидно, оба приведенных определения независимости событий в совокупности эквивалентны друг другу. Доказательство этого факта несложно и для экономии места здесь не приводится.

Независимость в совокупности сильнее других видов независимости ограничивает свойства системы событий. Из нее автоматически следуют и попарная, и взаимная независимость (первая соответствует $k = 2$ в (4.10), а вторая — $k = n$). Попарная и взаимная независимость могут наблюдаться в любых комбинациях: обе присутствовать (например, при наличии независимости в совокупности), может присутствовать только первая (см. пример 4.4), только вторая (см. пример 4.5) или отсутствовать обе сразу.

Далее рассмотрим примеры применения полученных формул.

Пример 4.1. Группа из n лиц с помощью жребия выбирает одного своего представителя. Для этого в урну поместили один выигрышный билет и $(n - 1)$ билет без выигрыша. Жребий тянут по очереди, пока не будет извлечен выигрышный билет. Какова вероятность того, что он достанется человеку, тянущему жребий j -м по счету?

Решение. Обозначим через A_j событие, состоящее в том, что жребий вытянул подошедший к урне j -м по порядку. Поскольку событие A_j может осуществиться только при усло-

вии, что до этого жребий никто не вытянул, то при любом $2 \leq j \leq n$ имеем

$$\begin{aligned} \mathbf{P}(A_j) &= \mathbf{P}(\overline{A_1} \overline{A_2} \dots \overline{A_{j-1}}) \mathbf{P}(A_j | \overline{A_1} \overline{A_2} \dots \overline{A_{j-1}}) = \\ &= \mathbf{P}(\overline{A_1}) \mathbf{P}(\overline{A_2} | \overline{A_1}) \dots \mathbf{P}(\overline{A_{j-1}} | \overline{A_1} \overline{A_2} \dots \overline{A_{j-2}}) \mathbf{P}(A_j | \overline{A_1} \overline{A_2} \dots \overline{A_{j-1}}), \\ \mathbf{P}(\overline{A_1}) &= \frac{n-1}{n}, \quad \mathbf{P}(\overline{A_2} | \overline{A_1}) = \frac{n-2}{n-1}, \dots \\ \dots, \mathbf{P}(\overline{A_{j-1}} | \overline{A_1} \overline{A_2} \dots \overline{A_{j-2}}) &= \frac{n-(j-1)}{n-(j-2)}, \\ \mathbf{P}(A_j | \overline{A_1} \overline{A_2} \dots \overline{A_{j-1}}) &= \frac{1}{n-(j-1)}, \\ \mathbf{P}(A_j) &= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \dots \cdot \frac{n-(j-1)}{n-(j-2)} \cdot \frac{1}{n-(j-1)} = \frac{1}{n}, \end{aligned}$$

то есть вероятность вытянуть жребий для всех участников одинакова и не зависит от очередности их подхода к урне.

Пример 4.2. С помощью 6 карточек, на которых написано по одной букве, составлено слово «карета». Карточки перемешиваются, а затем наугад извлекаются по одной. На первой извлеченной карточке оказалась буква «р». Какова вероятность, что в порядке поступления букв образуется слово «ракета»?

Решение. Обозначая событие, состоящее в появлении данной буквы большой буквой того же наименования, для искомой вероятности P получаем

$$\begin{aligned} p &= \mathbf{P}(AKETA|P) = \mathbf{P}(A|P) \cdot \mathbf{P}(K|PA) \cdot \mathbf{P}(E|PAK) \times \\ &\times \mathbf{P}(T|PAKE) \cdot \mathbf{P}(A|PAKET) = \frac{2}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{60}. \end{aligned}$$

Пример 4.3 Имеется партия из m изделий, среди которых l бракованных. Наудачу выбирается s изделий (причем $s \leq m - l$) и проводится их последовательная проверка. Партия принимается, если не окажется ни одного бракованного изделия. Определить вероятность того, что партия будет принята.

Решение. Обозначим через A_k случайное событие, состоящее в том, что k из проверяемых изделий доброкачественное ($k = \overline{1, s}$). Тогда искомая вероятность

$$p = \mathbf{P}(A_1 A_2 \dots A_s) = \mathbf{P}(A_1) \mathbf{P}(A_2 | A_1) \dots \mathbf{P}(A_s | \prod_{k=1}^{s-1} A_k).$$

Событию A_1 благоприятствует $m-l$ равновозможных исходов (число небракованных изделий) из общего числа m . Поэтому $\mathbf{P}(A_1) = (m-l)/m$. Если произошло событие A_1 , то число оставшихся изделий $(m-l)$, из них небракованных $(m-l-1)$. Поэтому $\mathbf{P}(A_2 | A_1) = (m-l-1)/(m-1)$. Аналогично получаем $\mathbf{P}(A_3 | A_1 A_2) = \frac{m-l-2}{m-2}$. В общем случае по индукции получаем

$$\mathbf{P}(A_k | A_1 A_2 \dots A_{k-1}) = \frac{m-l-(k-1)}{m-(k-1)}, \quad (k = \overline{1, s}).$$

Искомая вероятность

$$p = \frac{m-l}{m} \cdot \frac{m-l-1}{m-1} \cdot \frac{m-l-2}{m-2} \cdot \dots \cdot \frac{m-l-s+1}{m-s+1} = \frac{(m-l)!(m-s)!}{m!(m-l-s)!}$$

или после алгебраических преобразований и сокращения факториалов

$$p = \left(1 - \frac{l}{m}\right) \cdot \left(1 - \frac{l}{m-1}\right) \times \dots \times \left(1 - \frac{l}{m-s+1}\right).$$

Данный пример можно решить и без применения теоремы умножения вероятностей. Действительно, искомая вероятность равна вероятности того, что среди взятых s изделий нет ни одного бракованного. Общее число возможных способов взять s изделий из m равно $n = C_m^s$. Из них благоприятствующих исходов $C_m^0 C_{m-l}^s$.

Пример 4.4. На станцию обслуживания поступили четыре автомашины: у одной неисправен мотор, у другой помят

кузов, у третьей разбиты фары, у четвертой имеются все три указанные неисправности. Пусть A, B, C — события, заключающиеся в том, что у первой наудачу взятой для ремонта машины неисправен мотор (A), помят кузов (B), разбиты фары (C). Являются ли данные события независимыми попарно и в совокупности?

Решение. Каждому из трех введенных событий благоприятствуют по два исхода из четырех возможных. Поэтому

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

Если взята автомашина, у которой помят кузов (произошло событие B), то вероятность того, что у этой автомашины неисправен мотор, $\mathbf{P}(A|B) = \frac{1}{2}$. Аналогично находим, что

$$\mathbf{P}(A|C) = \mathbf{P}(B|A) = \mathbf{P}(B|C) = \mathbf{P}(C|B) = \frac{1}{2}.$$

Таким образом,

$$\mathbf{P}(A|B) = \mathbf{P}(A|C) = \mathbf{P}(A) = \frac{1}{2},$$

$$\mathbf{P}(B|A) = \mathbf{P}(B|C) = \mathbf{P}(B) = \frac{1}{2},$$

$$\mathbf{P}(C|A) = \mathbf{P}(C|B) = \mathbf{P}(C) = \frac{1}{2},$$

а это означает, что события *попарно независимые*.

Если у взятой автомашины помят кузов и разбиты фары (произошло событие BC), то это значит, что взята четвертая автомашина. Вероятность того, что у этой автомашины неисправен мотор, равна единице. Следовательно, $\mathbf{P}(A|BC) = 1$. Аналогично получаем, что и $\mathbf{P}(B|AC) = \mathbf{P}(C|AB) = 1$. Так как величины этих условных вероятностей получились отличными от $\mathbf{P}(A)$, $\mathbf{P}(B)$ и $\mathbf{P}(C)$, то события A, B и C не являются *независимыми в совокупности*.

Пример 4.5. На плоскость наудачу бросается игральная кость в форме правильного тетраэдра. Две грани его окрашены в зеленый, а две другие в красный цвет. Кость

подбрасывается три раза. Пусть события A и B состоят в том, что, соответственно, в первом и третьем опыте выпал зеленый цвет, а событие C означает, что в серии трех опытов ни разу не было перехода с красного на зеленое. Проверить события A , B , C на попарную, взаимную независимость и независимость в совокупности.

Решение. Правильный тетраэдр падает на каждую из своих четырех граней с вероятностью $1/4$. Согласно правилам раскраски граней выпадение любого из двух имеющихся цветов в одном опыте происходит с вероятностью $1/2$. Каждое из событий A и B не зависит ни от каких других событий и происходит с вероятностью

$$\mathbf{P}(A) = \mathbf{P}(B) = 1/2.$$

Вероятность события C найдем, используя классическое определение вероятности. В серии из трех бросков равновозможны и единственно возможны $n = 2^3 = 8$ исходов (вариантов чередования цветов). Из них событию C благоприятствует $m = 4$ варианта (333, 3ЗК, ЗКК, ККК). Следовательно,

$$\mathbf{P}(C) = \frac{m}{n} = \frac{1}{2}.$$

Подсчитаем вероятность произведения всех трех событий. По теореме умножения вероятностей

$$\mathbf{P}(ABC) = \mathbf{P}(C|AB)\mathbf{P}(AB) = \mathbf{P}(C|AB)\mathbf{P}(A)\mathbf{P}(B).$$

Ясно, что $\mathbf{P}(C|AB) = \frac{1}{2}$. Поэтому

$$\mathbf{P}(ABC) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \frac{1}{8},$$

то есть рассматриваемые три события *взаимно независимы* в смысле определения (4.9).

Теперь вычислим условные вероятности $\mathbf{P}(A|C)$ и $\mathbf{P}(B|C)$. Как мы видели, событию C благоприятствуют четыре исхода опыта. При этом событие A происходит лишь в трех из них (333, 3ЗК, ЗКК), а событие B — лишь в одном (333). Отсюда следует

$$\mathbf{P}(A|C) = \frac{3}{4}, \quad \mathbf{P}(B|C) = \frac{1}{4}.$$

Поскольку $\mathbf{P}(A|C) \neq \mathbf{P}(A)$ и $\mathbf{P}(B|C) \neq \mathbf{P}(B)$, то *попарная независимость* отсутствует, а вместе с ней отсутствует и *независимость в совокупности*.

§ 5. ТЕОРЕМА СЛОЖЕНИЯ ВЕРОЯТНОСТЕЙ. НЕСОВМЕСТНЫЕ СОБЫТИЯ

Теорема сложения вероятностей позволяет выразить вероятность $\mathbf{P}(A + B)$ суммы двух случайных событий A и B через вероятности $\mathbf{P}(A)$ и $\mathbf{P}(B)$ этих событий и вероятность $\mathbf{P}(AB)$ их произведения. При выводе этой теоремы воспользуемся сначала геометрической интерпретацией. Вероятности событий A , B и AB в этом случае определяются равенствами (4.1). Случайному событию $A + B$ из общей области Ω с геометрической мерой S благоприятствует область $(A + B)$, являющаяся объединением областей A и B (см. рис. 4.1). Геометрическая мера этой области $S_{A+B} = S_A + S_B - S_{AB}$, а потому

$$\frac{S_{A+B}}{S} = \frac{S_A}{S} + \frac{S_B}{S} - \frac{S_{AB}}{S}. \quad (5.1)$$

Так как $\mathbf{P}(A + B) = S_{A+B}/S$, то с учетом равенств (4.1) из (5.1) получаем

$$\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(AB). \quad (5.2)$$

Таким образом, вероятность суммы двух случайных событий равна сумме вероятностей этих событий за вычетом вероятности их произведения.

События A и B , как мы помним (см. § 1), называют *несовместными*, если они одновременно не могут произойти в одном опыте. Иногда используют также термин *непересекающиеся* события. Систему из произвольного числа $n \geq 2$ событий $A_1, \dots, A_n$ называют *несовместной* в данном опыте, если в результате этого опыта никакие два из них не могут появиться вместе друг с другом. Подчеркнем, что понятие несовместности относится к данным конкретным условиям случайного

эксперимента. Два события несовместные в одном опыте, могут оказаться совместными в другом. Например, выпадение орла и решки при одном бросании монеты несовместны, однако они становятся совместными, если совершаются $k \geq 2$ бросаний.

Математически несовместность событий A и B записывается в виде равенства $AB = V$, где V — невозможное событие. Поскольку при этом $\mathbf{P}(AB) = 0$, то для несовместных событий теорема сложения (5.2) приобретает наиболее простой вид

$$\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B). \quad (5.3)$$

Таким образом, вероятность суммы двух несовместных событий равна сумме вероятностей этих событий. В частном случае, когда $B = \bar{A}$, а потому $A + B = A + \bar{A} = U$, где U — достоверное событие, из (5.3) следует, что

$$\mathbf{P}(A) + \mathbf{P}(\bar{A}) = 1, \quad (5.4)$$

что эквивалентно

$$\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A), \quad (5.5)$$

то есть вероятность противоположного события дополняет вероятность исходного события до единицы.

Теорему сложения вероятностей можно доказать и с помощью классического определения вероятности. Пусть случайному событию A благоприятствует m элементарных событий E_k из их общего числа n , событию B благоприятствует l исходов, а событию AB — s исходов. Тогда $\mathbf{P}(A + B) = (m + l - s)/n$. Так как $\mathbf{P}(A) = m/n$, $\mathbf{P}(B) = l/n$, $\mathbf{P}(AB) = s/n$, то получаем $\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(AB)$, то есть прежнюю формулу (5.2).

Формула (5.2) может быть легко обобщена на любое число слагаемых. Например, для трех событий A , B и C находим

$$\mathbf{P}(A + B + C) = \mathbf{P}(A + B) + \mathbf{P}(C) - \mathbf{P}[(A + B)C].$$

Так как

$$\mathbf{P}[(A + B)C] = \mathbf{P}(AC) + \mathbf{P}(BC) - \mathbf{P}(ABC),$$

то с учетом (5.2) получим

$$\begin{aligned} \mathbf{P}(A + B + C) &= \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \\ &\quad - \mathbf{P}(AB) - \mathbf{P}(AC) - \mathbf{P}(BC) + \mathbf{P}(ABC). \end{aligned} \quad (5.6)$$

Для произвольного числа n событий $A_1, A_2, \dots, A_n$ методом индукции может быть доказана следующая формула:

$$\begin{aligned} \mathbf{P}\left(\sum_{k=1}^n A_k\right) &= \sum_{k=1}^n \mathbf{P}(A_k) - \\ &\quad - \sum_{k=1}^{n-1} \sum_{j=k+1}^n \mathbf{P}(A_k A_j) + \sum_{k=1}^{n-2} \sum_{j=k+1}^{n-1} \sum_{i=j+1}^n \mathbf{P}(A_k A_j A_i) + \dots \\ &\quad \dots + (-1)^{n-1} \mathbf{P}\left(\prod_{k=1}^n A_k\right). \end{aligned} \quad (5.7)$$

Если система событий $A_1, A_2, \dots, A_n$ несовместна, то из (5.7) получим равенство, аналогичное (5.3)

$$\mathbf{P}\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbf{P}(A_k), \quad (5.8)$$

то есть вероятность суммы любого числа несовместных событий равна сумме вероятностей этих событий.

Если вероятности произведения событий A_k не зависят от номеров событий, входящих в произведения, и определяются только числом сомножителей, то формула (5.7) упрощается и принимает вид

$$\begin{aligned} \mathbf{P}\left(\sum_{k=1}^n A_k\right) &= C_n^1 \mathbf{P}(A_1) - C_n^2 \mathbf{P}(A_1 A_2) + \\ &\quad + C_n^3 \mathbf{P}(A_1 A_2 A_3) - \dots + (-1)^{n-1} \mathbf{P}\left(\prod_{k=1}^n A_k\right), \end{aligned}$$

или

$$\mathbf{P}\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} C_n^k \mathbf{P}\left(\prod_{j=1}^k A_j\right). \quad (5.9)$$

Используя общую формулу (5.7), можно получить выражение для вероятности $\mathbf{P}\left(\prod_{k=1}^n A_k\right)$ произведения n событий через вероятности сумм различных комбинаций сомножителей. Для вывода соответствующей формулы запишем равенство (5.7) применительно к противоположным событиям. Тогда

$$\begin{aligned} \mathbf{P}\left(\sum_{k=1}^n \overline{A_k}\right) &= \sum_{k=1}^n \mathbf{P}(\overline{A_k}) - \\ &- \sum_{k=1}^{n-1} \sum_{j=k+1}^n \mathbf{P}(\overline{A_k A_j}) + \sum_{k=1}^{n-2} \sum_{j=k+1}^{n-1} \sum_{i=j+1}^n \mathbf{P}(\overline{A_k A_j A_i}) - \dots \\ &\dots + (-1)^{n-1} \mathbf{P}\left(\prod_{k=1}^n \overline{A_k}\right). \end{aligned} \quad (5.10)$$

С помощью (1.21), (1.24) и (5.5) получим

$$\mathbf{P}\left(\prod_{k=1}^n \overline{A_k}\right) = \mathbf{P}\left(\overline{\prod_{k=1}^n A_k}\right) = 1 - \mathbf{P}\left(\prod_{k=1}^n A_k\right), \quad (5.11)$$

$$\mathbf{P}\left(\prod_{k=1}^r \overline{A_k}\right) = \mathbf{P}\left(\overline{\sum_{k=1}^r A_k}\right) = 1 - \mathbf{P}\left(\sum_{k=1}^r A_k\right). \quad (5.12)$$

Подставляя эти вероятности в (5.10) и учитывая, что

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

приходим к следующей формуле для вероятности произведения n событий:

$$\begin{aligned} \mathbf{P}\left(\prod_{k=1}^n A_k\right) &= \sum_{k=1}^n \mathbf{P}(A_k) - \sum_{k=1}^{n-1} \sum_{j=k+1}^n \mathbf{P}(A_k + A_j) + \\ &+ \sum_{k=1}^{n-2} \sum_{j=k+1}^{n-1} \sum_{i=j+1}^n \mathbf{P}(A_k + A_j + A_i) - \dots \\ &\dots + (-1)^{n-1} \mathbf{P}\left(\sum_{k=1}^n A_k\right). \end{aligned} \quad (5.13)$$

Формулы (5.7) и (5.13) близки по своей структуре. Их часто объединяют в одну группу, называя *формулами включения-исключения*, или *формулами Пуанкаре*. Они особенно эффективны для систем событий, независимых в совокупности, а также для симметричных систем событий. Поясним понятие симметрии подробнее. Говорят, что система событий $A_1, \dots, A_n$ *симметрична*, если при любом $1 \leq k \leq n$ вероятности $\mathbf{P}\left(\sum_{s=1}^k A_{i_s}\right)$ и $\mathbf{P}\left(\prod_{s=1}^k A_{i_s}\right)$ зависят только от числа k , но не зависят от конкретных значений индексов $i_1, \dots, i_k$.

Как мы видели, для симметричной системы событий первая формула включения-исключения (5.7) принимает вид (5.9). Аналогичному преобразованию можно подвергнуть при наличии симметрии и вторую формулу (5.13)

$$\mathbf{P}\left(\prod_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} C_n^k \mathbf{P}\left(\sum_{j=1}^k A_j\right). \quad (5.14)$$

Для симметричной системы независимых событий сумма (5.9) еще более упрощается и может быть выражена в конечном виде

$$\mathbf{P}\left(\sum_{k=1}^n A_k\right) = 1 - [1 - \mathbf{P}(A_1)]^n, \quad (5.15)$$

причем вместо A_1 здесь может стоять любое другое событие системы A_j , ($j = \overline{1, n}$), так как в силу симметрии вероятности всех этих событий одинаковы.

Пример 5.1. Два стрелка стреляют в тире по одной мишени. Вероятность поразить мишень первым стрелком p_1 , вторым стрелком — p_2 . Определить вероятность поражения цели.

Решение. Пусть событие A_1 — поражение мишени первым стрелком, событие A_2 — вторым стрелком. По условию $\mathbf{P}(A_1) = p_1$, $\mathbf{P}(A_2) = p_2$. События A_1 и A_2 независимы. Искомая вероятность

$$\mathbf{P}(A_1 + A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1)\mathbf{P}(A_2) = p_1 + p_2 - p_1p_2.$$

Тот же результат можно получить и несколько иначе. Если ввести в рассмотрение события A_1 и A_2 , соответствующие промахам первого и соответственно второго стрелка, то

$$\begin{aligned}\mathbf{P}(A_1 + A_2) &= 1 - \overline{\mathbf{P}(A_1 + A_2)} = 1 - \overline{\mathbf{P}(\overline{A_1} \cdot \overline{A_2})} = \\ &= 1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1p_2.\end{aligned}$$

Пример 5.2. Минное заграждение состоит из n мин, поставленных в одну линию перпендикулярно возможному курсу корабля. Корабль может подорваться только на одной мине, причем при взрыве одной мины остальные не повреждаются. Минное заграждение пересекается последовательно и независимо друг от друга кораблями, причем пересечение линии мин для каждого корабля равновозможно в любом месте. Определить вероятность P подрыва m -го корабля, если вероятность подрыва корабля на k -й мине ($k = \overline{1, n}$) для корабля, первым пересекающего заграждение, равна p ($p < 1/n$).

Решение. Обозначим через A_k — случайное событие, состоящее в том, что на k -й мине $m - 1$ кораблей не подорвутся, а m -й подорвется ($k = \overline{1, n}$). Вероятность такого события $\mathbf{P}(A_k) = q^{m-1}p$, где $q = 1 - p$. Искомая вероятность $P = \mathbf{P}(A_1 + A_2 + \dots + A_n)$.

Так как m -й корабль может подорваться только на одной мине, то события $A_1, \dots, A_n$ несовместны. Воспользовавшись формулой (5.8), получим

$$P = \sum_{k=1}^n \mathbf{P}(A_k) = n\mathbf{P}(A_1) = npq^{m-1}.$$

Пример 5.3 (Задача о совпадениях). Некто, желая отправить n писем n лицам, заготовил n конвертов с адресами этих лиц, а затем распределил письма по конвертам случайным образом. Определить вероятность того, что m писем будут иметь правильные адреса ($m \leq n$).

Решение. Обозначим: событие A — адреса написаны правильно для m писем; событие B — остальные $(n - m)$ писем имеют неправильные адреса. Тогда искомая вероятность $p = \mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B|A)$.

Случайное событие A можно представить в виде суммы несовместных случайных событий, каждое из которых соответствует определенной последовательности писем с правильными и неправильными адресами.

Используя теорему сложения вероятностей для случая несовместных событий с одинаковыми вероятностями, получаем

$$\mathbf{P}(A) = C_n^m \mathbf{P}(A_1 A_2 \dots A_m),$$

где случайное событие A_k означает, что в k -м по порядку письме оказался правильный адрес. Так как письма распределены по конвертам случайным образом, то согласно классическому определению вероятности

$$\mathbf{P}(A_1) = \frac{1}{n}, \quad \mathbf{P}(A_2|A_1) = \frac{1}{n-1}, \quad \mathbf{P}(A_3|A_1 A_2) = \frac{1}{n-2}.$$

В общем случае

$$\mathbf{P}(A_s|A_1 A_2 \dots A_{s-1}) = \frac{1}{n-s+1}, \quad (s = \overline{1, m}).$$

Поэтому вероятность события A выразится в виде

$$\mathbf{P}(A) = C_n^m \frac{1}{n} \cdot \frac{1}{n-1} \cdot \dots \cdot \frac{1}{n-m+1} = \frac{n!}{m!(n-m)!} \cdot \frac{(n-m)!}{n!} = \frac{1}{m!}.$$

Условную вероятность $\mathbf{P}(B|A)$ представим в виде

$$\mathbf{P}(B|A) = 1 - \mathbf{P}(C),$$

где случайное событие C означает, что хотя бы у одного из оставшихся писем написан правильный адрес. Пусть событие $C_k (k = \overline{1, n-m})$ означает, что на k -м из оставшихся конвертов написан правильный адрес. Очевидно, что $C = C_1 + C_2 + \dots + C_{n-m}$. Тогда имеем

$$\mathbf{P}(C_k) = \frac{1}{n-m} = \frac{1}{1!C_{n-m}^1},$$

$$(k = \overline{1, n-m}),$$

$$\mathbf{P}(C_k C_j) = \frac{1}{(n-m)(n-m-1)} = \frac{1}{2!C_{n-m}^2},$$

$$(k \neq j = \overline{1, n-m}).$$

Аналогично предыдущему получим, что вероятности произведений при равных количествах событий совпадают, причем

$$\mathbf{P}(C_1 C_2 \dots C_s) = \frac{(n-m-s)!}{(n-m)!} = \frac{1}{s!C_{n-m}^s}, \quad (s = \overline{1, n-m}).$$

Используя формулу включения-исключения для суммы событий (5.7), находим

$$\begin{aligned} \mathbf{P}(C) &= C_{n-m}^1 \frac{1}{1!C_{n-m}^1} - C_{n-m}^2 \frac{1}{2!C_{n-m}^2} + \\ &+ C_{n-m}^3 \frac{1}{3!C_{n-m}^3} - \dots + (-1)^{n-m-1} \frac{1}{(n-m)!} = \\ &= \sum_{k=1}^{n-m} \frac{(-1)^{k-1}}{k!} = 1 - \sum_{k=0}^{n-m} \frac{(-1)^k}{k!}. \end{aligned}$$

Вспоминая определение события C , получаем, что

$$\mathbf{P}(B|A) = \sum_{k=0}^{n-m} \frac{(-1)^k}{k!},$$

а потому искомая вероятность

$$p = \frac{1}{m!} \sum_{k=0}^{n-m} \frac{(-1)^k}{k!}.$$

При большом значении разности $(n - m)$ асимптотически

$$\sum_{k=0}^{n-m} \frac{(-1)^k}{k!} \approx \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1},$$

следовательно, в этом случае получаем приближенное значение искомой вероятности в виде

$$p^* \approx \frac{1}{m!} \cdot e^{-1}.$$

Последней формулой практически можно пользоваться при большом n и любых m , так как при m малом имеем $(n - m) \gg \gg 1$, а при m , сравнимом с n , получаем $p \approx 0$ вследствие большого значения факториала, стоящего в знаменателе.

О точности последней приближенной формулы можно судить по данным таблицы 5.1. В этой таблице приводятся точные значения искомой вероятности p , а также относительная погрешность

$$\delta = \frac{p - p^*}{p}.$$

Отметим, что погрешность δ зависит только от разности $m - n$, поэтому данные последней строки таблицы можно использовать при произвольном задании n , взяв соответствующее m .

Таблица 5.1

Сравнение точного и приближенного решения при $n = 10$

m	0	2	4
p	0,368	0,184	$0,153 \cdot 10^{-1}$
δ	$0,650 \cdot 10^{-7}$	$0,680 \cdot 10^{-5}$	$0,478 \cdot 10^{-3}$
m	6	8	10
p	$0,521 \cdot 10^{-3}$	$0,124 \cdot 10^{-4}$	$0,275 \cdot 10^{-6}$
δ	$0,190 \cdot 10^{-1}$	0,264	0,632

Как показывает анализ числовых данных, приближенная формула всегда дает оценку искомой вероятности снизу, причем при $n - m \geq 5$ ее относительная погрешность не превосходит 1% и быстро убывает с ростом $n - m$. Если m близко к n , но при этом оба этих числа достаточно велики, то δ может достигать порядка 60%, однако оба значения p и p^* настолько малы, что интересующее нас событие (совпадение адресов и писем) практически является невозможным.

Пример 5.4. В n одинаковых урнах по 1 черному и l белых шаров. Испытание состоит в том, что из выбранной наудачу урны извлекается шар. Определить вероятность того, что после m испытаний ($l + 1 \geq m \geq n$) в урнах не останется черных шаров, если вероятность извлечь черный шар из выбранной урны равна p . Рассчитать эту вероятность при $n = 3$, $m = 4$, $l = 4$.

Решение. Обозначим через A извлечение черного шара из k -й урны ($k = \overline{1, n}$). Искомая вероятность $P = \mathbf{P}\left(\prod_{k=1}^n A_k\right)$ может быть определена по формуле (5.13). Так как вероятности различных сумм событий A_k ($k = \overline{1, n}$) не зависят от номеров этих событий, а определяются только числом слагаемых, то искомая вероятность

$$\begin{aligned} P &= \mathbf{P}\left(\prod_{k=1}^n A_k\right) = C_{n^1} \mathbf{P}(A_1) - C_{n^2} \mathbf{P}(A_1 + A_2) + \\ &+ C_{n^3} \mathbf{P}(A_1 + A_2 + A_3) - \dots + (-1)^{n-1} \mathbf{P}\left(\sum_{k=1}^n A_k\right) = \\ &= \sum_{k=1}^n (-1)^{k-1} C_n^k \mathbf{P}\left(\sum_{j=1}^k A_j\right). \end{aligned}$$

При одном опыте вероятность того, что шар будет извлечен из какой-либо из k урн, равна $\frac{k}{n} \cdot p$, где $p = \frac{1}{l+1}$. Вероятность того, что при m опытах не будут извлечены шары ни из

одной из выбранных k урн, равна $\left(1 - \frac{k}{n} \cdot p\right)^m$, а вероятность извлечь шар хотя бы из одной из этих k урн будет следующей:

$$\mathbf{P}\left(\sum_{j=1}^k A_j\right) = 1 - \left(1 - \frac{k}{n} \cdot p\right)^m.$$

Искомая вероятность находится по формуле (5.9)

$$P = \sum_{k=1}^n (-1)^{k-1} C_n^k \left[1 - \left(1 - \frac{k}{n} \cdot p\right)^m\right].$$

Для указанных в условии задачи числовых данных находим

$$\begin{aligned} \mathbf{P}(A_1 A_2 A_3) &= 3 \left[1 - \left(1 - \frac{1}{15}\right)^4\right] - 3 \left[1 - \left(1 - \frac{2}{15}\right)^4\right] + \\ &+ \left[1 - \left(1 - \frac{3}{15}\right)^4\right] = \frac{4}{625}. \end{aligned}$$

Пример 5.5 (Задача о жюри). Жюри состоит из трех экспертов, каждый из которых индивидуально принимает правильное решение с вероятностью p_i , ($i = \overline{1, 3}$), причем независимо от своих коллег. Жюри в целом выносит решение по большинству голосов. Известно, что k членов жюри являются профессионалами своего дела и принимают верное решение с одинаковой вероятностью $p_i = p = \text{const}(i)$. Оставшиеся $3 - k$ человек, будучи дилетантами, выносят вердикт с вероятностью $p_i = 1/2$. Определить для произвольного $0 \leq k \leq 3$ вероятность принятия правильного решения при таком составе жюри.

Решение. Пусть событие A_i обозначает правильное голосование i -го члена жюри. По условию система случайных событий (A_1, A_2, A_3) независима в совокупности. Введем случайные события B_j — верное решение приняли j экспертов ($j = \overline{0, 3}$). Система событий B_0, B_1, B_2, B_3 несовместна,

а составляющие ее события выражаются через A_i следующим образом:

$$\begin{aligned} B_0 &= \overline{A_1} \overline{A_2} \overline{A_3}, & B_1 &= A_1 \overline{A_2} \overline{A_3} + \overline{A_1} A_2 \overline{A_3} + \overline{A_1} \overline{A_2} A_3, \\ B_2 &= A_1 A_2 \overline{A_3} + A_1 \overline{A_2} A_3 + \overline{A_1} A_2 A_3, & B_3 &= A_1 A_2 A_3. \end{aligned}$$

Все слагаемые, входящие в выражение для B_1 и B_2 , несовместны, поэтому по теореме сложения вероятностей и теореме их умножения

$$\begin{aligned} \mathbf{P}(B_0) &= (1 - p_1)(1 - p_2)(1 - p_3), \\ \mathbf{P}(B_1) &= p_1(1 - p_2)(1 - p_3) + p_2(1 - p_1)(1 - p_3) + \\ &\quad + p_3(1 - p_1)(1 - p_2), \\ \mathbf{P}(B_2) &= p_1p_2(1 - p_3) + p_1p_3(1 - p_2) + p_2p_3(1 - p_1), \\ \mathbf{P}(B_3) &= p_1p_2p_3. \end{aligned}$$

Обозначим через C вынесение правильного вердикта жюри в целом простым большинством голосов. Ясно, что $C = B_2 + B_3$, поэтому

$$\mathbf{P}(C) = \mathbf{P}(B_2) + \mathbf{P}(B_3),$$

поскольку все B_j несовместны.

Осталось учесть соотношение профессионалов и дилетантов в коллегии судей. Обозначим через $\varphi_k(p)$ вероятность справедливого решения при наличии k профессионалов, компетентность которых характеризуется вероятностью p . Для большей наглядности кроме результирующего значения φ_k выпишем также и вероятности $\mathbf{P}(B_2)$ и $\mathbf{P}(B_3)$.

Опуская промежуточные выкладки, получаем

$$\begin{aligned} \varphi_0 &= \frac{1}{2}, & \mathbf{P}(B_2) &= \frac{3}{8}, & \mathbf{P}(B_3) &= \frac{1}{8}, \\ \varphi_1 &= \frac{1}{4} + \frac{p}{2}, & \mathbf{P}(B_2) &= \frac{1}{4} + \frac{p}{4}, & \mathbf{P}(B_3) &= \frac{1}{4} \cdot p, \\ \varphi_2 &= p, & \mathbf{P}(B_2) &= p - \frac{p^2}{2}, & \mathbf{P}(B_3) &= \frac{p^2}{2}, \\ \varphi_3 &= 3p^2 - 2p^3, & \mathbf{P}(B_2) &= 3p^2(1 - p), & \mathbf{P}(B_3) &= p^3. \end{aligned}$$

Полученные результаты графически представлены на рис. 5.1.

Если в жюри входят только дилетанты ($k = 0$), то обоснованно выбрать один из двух вариантов решения не представляется возможным, и поэтому $\varphi_0 \equiv 1/2$. Увеличение числа профессионалов повышает вероятность правильного решения, однако при $k = 1$ эта вероятность не превосходит 75%, даже если профессиональный член жюри обладает высокой компетентностью ($p = 1$). Получить φ_k , близкое к единице, можно только

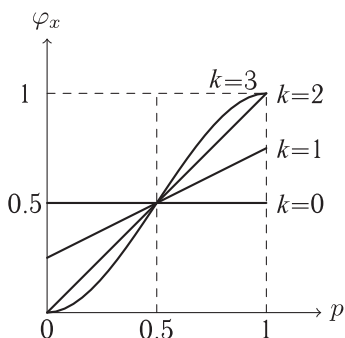


Рис. 5.1

Зависимость вероятности φ_k от параметра p

при достаточно профессиональном жюри: во-первых, специалистов должно быть больше, чем дилетантов и, во-вторых, компетентность специалистов должна быть высокой (p близко к единице). Интересно заметить, что при одном легкомысленном члене жюри ($k = 2$) вероятность принятия верного решения $\varphi_2(p) = p$, то есть жюри по эффективности работы равноценно одному отдельно взятому специалисту.

Оценим, насколько жюри, составленное из трех специалистов, работает эффективнее одного эксперта-специалиста. Для этого рассмотрим разность

$$\Delta(p) = \varphi_3(p) - \varphi_2(p)$$

и найдем максимум Δ в промежутке $\left(\frac{1}{2}, 1\right)$. Производная $\Delta'(p)$ обращается в нуль при $p = p_*$, где

$$p_* = \frac{1}{2} + \frac{1}{2\sqrt{3}} = 0,789.$$

В этой точке отклонение φ_3 от φ_2 максимально и равняется

$$\Delta_{\max} = \Delta(p_*) = \frac{1}{6\sqrt{3}} = 0,0962.$$

Следовательно, замена одного эксперта на жюри из трех профессионалов той же квалификации позволяет повысить надежность принимаемых решений примерно на 9,6%, причем при условии, что p близко к p_* . Если же к экспертизе привлекаются малокомпетентные специалисты (p близко к 0,5), либо специалисты высокого класса (p близко к 1), то разница между коллегией экспертов и одним экспертом практически исчезает.

Пример 5.6 (Продолжение задачи о жюри). Допустим, что сохраняются все условия задачи 5.5 за исключением требования независимости работы экспертов. Жюри сформировано из специалистов, $p_i = p$ ($i = \overline{1,3}$). Независимо от остальных работает лишь один (скажем, первый) эксперт, а два других принимают одинаковые решения с вероятностью α , удовлетворяющей условию $p^2 \leq \alpha \leq p$. Требуется найти вероятность φ вынесения правильного решения.

Замечание. Поясним последнее ограничение на α . Оно означает, что $\mathbf{P}(A_2 A_3) = \alpha$. Ясно, что при заданном значении вероятности p значение α не может быть совершенно произвольным. Например, в случае независимости A_2 от A_3 имеем $\alpha = p^2$. Если же второй и третий эксперты голосуют всегда абсолютно синхронно, то есть события A_2 и A_3 эквивалентны, то тогда $\alpha = p$. В задаче требуется изучить вариант «мягкого» согласования позиции членов жюри, когда $p^2 \leq \alpha \leq p$.

Разумеется, данная постановка задачи не является единственно возможной. Так, возможен вариант конфликтной ситуации, когда эксперты голосуют не согласованно, а в оппозиции друг другу (крайний случай дается равенством $A_2 = \overline{A_3}$). Тогда оказывается, что $0 \leq \alpha \leq p^2$. Этот случай в данном примере мы оставляем в стороне.

Решение. Ранее полученные выражения для событий B_j сохраняют силу, однако вычисление вероятностей $\mathbf{P}(B_j)$ усложняется, так как система событий (A_1, A_2, A_3) уже не будет независимой. Конечно, A_1 не зависит от A_2 и A_3 , но два последних события связаны друг с другом.

Используя независимость A_1 от остальных событий, получаем

$$\mathbf{P}(B_0) = \mathbf{P}(\overline{A_1}) \cdot \mathbf{P}(\overline{A_2} \overline{A_3}),$$

$$\begin{aligned}\mathbf{P}(B_1) &= \mathbf{P}(A_1)\mathbf{P}(\overline{A_2} \overline{A_3}) + \mathbf{P}(\overline{A_1})\mathbf{P}(A_2\overline{A_3} + A_3\overline{A_2}), \\ \mathbf{P}(B_2) &= \mathbf{P}(A_1)\mathbf{P}(A_2\overline{A_3} + A_3\overline{A_2}) + \mathbf{P}(\overline{A_1})\mathbf{P}(A_2A_3), \\ \mathbf{P}(B_3) &= \mathbf{P}(A_1)\mathbf{P}(A_2A_3).\end{aligned}$$

В эти выражения сразу же подставляем $\mathbf{P}(A_2A_3) = \alpha$.

Для вычисления вероятности $\mathbf{P}(\overline{A_2} \overline{A_3})$ воспользуемся формулой инверсии (1.20), что дает

$$\begin{aligned}\mathbf{P}(\overline{A_2} \overline{A_3}) &= \mathbf{P}(\overline{A_2 + A_3}) = 1 - \mathbf{P}(A_2 + A_3) = \\ &= 1 - \mathbf{P}(A_2) - \mathbf{P}(A_3) + \mathbf{P}(A_2A_3) = 1 - 2p + \alpha.\end{aligned}$$

Далее обратимся к вычислению вероятности события

$$A_2\overline{A_3} + A_3\overline{A_2} = (A_2 \setminus A_3) + (A_3 \setminus A_2),$$

представляющего сумму разностей событий A_2 минус A_3 и A_3 минус A_2 . Такую сумму часто называют *симметрической разностью* A_2 и A_3 . Воспользовавшись диаграммами Венна (см. § 1) нетрудно показать, что

$$\mathbf{P}(A_2\overline{A_3} + A_3\overline{A_2}) = \mathbf{P}(A_2 + A_3) - \mathbf{P}(A_2A_3) = 2(p - \alpha).$$

Опуская несложные промежуточные выкладки, находим

$$\begin{aligned}\mathbf{P}(B_0) &= (1 - p)(1 - 2p + \alpha), \\ \mathbf{P}(B_1) &= p(1 - 2p + \alpha) + 2(1 - p)(p - \alpha), \\ \mathbf{P}(B_2) &= 2p(p - \alpha) + \alpha(1 - p), \\ \mathbf{P}(B_3) &= \alpha p, \\ \varphi = \mathbf{P}(C) &= \mathbf{P}(B_2) + \mathbf{P}(B_3) = 2p^2 - 2p\alpha + \alpha.\end{aligned}$$

По условию $p^2 \leq \alpha \leq p$, поэтому при таких α

$$p \leq \varphi \leq 3p^2 - 2p^3,$$

причем левое неравенство реализуется при $A_2 = A_3$, когда эксперты голосуют одинаково, а правое — в случае независимости A_2 от A_3 , то есть при полной независимости решений, принимаемых различными экспертами. Как видим, какое-либо

согласование позиций членов жюри ($\alpha > p^2$) приводит к снижению эффективности его работы.

К такому же результату приводит и наличие антагонизма между экспертами, когда некоторые из них стараются голосовать в противовес другим. Математически этот случай можно рассмотреть аналогично разобранному выше, если задать вероятность $\beta = \mathbf{P}(A_2 \overline{A_3})$, наложив на нее ограничение вида $p(1-p) \leq \beta \leq p$. При независимом голосовании $\beta = p(1-p)$, а при антагонистическом голосовании ($A_2 = \overline{A_3}$) имеем $\beta = p$. При антагонистическом голосовании решение задачи очевидно: все решает первый эксперт, занимающий правильную позицию с вероятностью p . Следовательно, аналогично случаю $A_2 = A_3$ имеем $\varphi = p$. Можно показать, что при промежуточных значениях $\beta \in [p(1-p), p]$ искомая вероятность φ находится в пределах от p до $3p^2 - 2p^3$, как и в случае согласования решений экспертами. Точно также, как и в этом случае, самым справедливым является жюри из независимых экспертов, обеспечивающее максимум φ .

§ 6. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ

Пусть случайное событие A может произойти только вместе с одним из n несовместных событий $H_1, H_2, \dots, H_n$, образующих полную группу. В дальнейшем будем называть события H_k ($k = \overline{1, n}$) *гипотезами*. Предположим, что для каждой гипотезы известны условные вероятности $\mathbf{P}(A|H_k)$ появления события A .

Формула полной вероятности определяет связь вероятности $\mathbf{P}(A)$ события A с вероятностями $\mathbf{P}(H_k)$ гипотез H_k и условными вероятностями $\mathbf{P}(A|H_k)$ появления события A при всех этих гипотезах.

Чтобы выразить вероятность $\mathbf{P}(A)$ через совокупность вероятностей $\mathbf{P}(H_k)$ и $\mathbf{P}(A|H_k)$, заметим прежде всего, что так как гипотезы образуют полную группу событий, то их сумма дает достоверное событие U , то есть

$$\sum_{k=1}^n H_k = U, \quad \sum_{k=1}^n \mathbf{P}(H_k) = 1.$$

Следовательно, справедливы тождества

$$A = UA = \left(\sum_{k=1}^n H_k \right) A = \sum_{k=1}^n H_k A. \quad (6.1)$$

При $k = j$ события H_k и H_j несовместны, так как $H_k H_j = V$. Поэтому несовместными являются и события $(H_k A)$ и $(H_j A)$

$$(H_k A)(H_j A) = V.$$

Воспользовавшись теоремой сложения для n несовместных событий, из (6.1) получим

$$\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(H_k A). \quad (6.2)$$

По теореме умножения вероятностей

$$\mathbf{P}(H_k A) = \mathbf{P}(H_k) \mathbf{P}(A|H_k),$$

поэтому формулу полной вероятности (6.2) для вероятности события A можно записать в окончательном виде

$$\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(H_k) \mathbf{P}(A|H_k). \quad (6.3)$$

Приведенные рассуждения сохраняют свою силу и в случае бесконечного числа гипотез, если только система гипотез $\{H_j\}_{j=1}^{\infty}$ остается полной, то есть выполняются условия

$$\sum_{j=1}^{\infty} H_j = U; \quad H_k H_j = V, \quad (j = \overline{1, \infty}, k = \overline{1, \infty}, j \neq k). \quad (6.4)$$

При этом

$$\mathbf{P}(A) = \sum_{k=1}^{\infty} \mathbf{P}(H_k) \mathbf{P}(A|H_k). \quad (6.5)$$

Несмотря на кажущуюся простоту и очевидность формула полной вероятности широко используется в теории вероятностей и ее приложениях. Она является основным средством

вычисления вероятностей сложных событий. Дело в том, что условные вероятности часто легко вычисляются или даже задаются по условию задачи. Кроме того, формула полной вероятности может быть полезна для уяснения механизма рассматриваемого случайного явления. Имеются целые крупные разделы теории вероятностей, например, теория цепей Маркова, где формула полной вероятности играет определяющую роль.

Рассмотрим простейший частный случай формулы (6.3), когда $n = 2$, а в качестве системы гипотез выбраны $H_1 = B$, $H_2 = \bar{B}$, где B — некоторое произвольное событие, отличное от достоверного и невозможного события ($0 < \mathbf{P}(B) < 1$). Тогда

$$\mathbf{P}(A) = \mathbf{P}(A|B)\mathbf{P}(B) + \mathbf{P}(A|\bar{B})\mathbf{P}(\bar{B}). \quad (6.6)$$

За счет рационального выбора события B часто удается существенно упростить расчет искомой вероятности A .

Далее рассмотрим примеры применения формулы полной вероятности.

Пример 6.1. На складе имеется десять приборов, из которых девять исправны, а один дефектный. Вероятность того, что исправный прибор во время опыта даст отказ, равна 0,001; вероятность отказа для дефектного прибора равна 0,5. Какова вероятность, что взятый наудачу прибор даст отказ?

Решение. Обозначим отказ взятого прибора событием A . Введем в рассмотрение две гипотезы: H_1 — взятый наудачу прибор исправный, H_2 — взятый прибор дефектный. В соответствии с условиями задачи

$$\mathbf{P}(H_1) = \frac{9}{10}, \quad \mathbf{P}(H_2) = \frac{1}{10}, \quad \mathbf{P}(A|H_1) = 0,01, \quad \mathbf{P}(A|H_2) = 0,85.$$

Применяя теорему о полной вероятности, получаем

$$\mathbf{P}(A) = \sum_{k=1}^2 \mathbf{P}(H_k)\mathbf{P}(A|H_k) = 0,9 \cdot 0,01 + 0,1 \cdot 0,85 = 0,094.$$

Пример 6.2. Двое поочередно бросают монету. Выигрывает тот, кто первый получит «орел». Каковы вероятности выиграть первому, второму?

Решение. Обозначим выигрыш первого A_1 , выигрыш второго — A_2 . Обозначим через H_k событие, состоящее в том, что первому игроку придется бросить монету k раз. События H_k можно рассматривать как гипотезы, причем их число бесконечно. Для того чтобы имела место гипотеза H_k , необходимо, чтобы и у первого, и у второго игрока $k-1$ раз подряд выпадала цифра («решка»). Следовательно,

$$\mathbf{P}(H_k) = \frac{1}{2^{k-1}} \cdot \frac{1}{2^{k-1}} = \frac{1}{2^{2k-2}}.$$

Вероятность $\mathbf{P}(A_1|H_k) = 1/2$, независимо от номера гипотезы. Следовательно,

$$\mathbf{P}(A) = \sum_{k=1}^{\infty} \mathbf{P}(H_k) \mathbf{P}(A_1|H_k) = \frac{1}{2} \cdot \sum_{k=1}^{\infty} \frac{1}{2^{2(k-1)}} = \frac{1}{2} \cdot \sum_{j=0}^{\infty} \frac{1}{4^j} = \frac{2}{3}.$$

Аналогично для второго игрока

$$\mathbf{P}(A_2|H_k) = \frac{1}{4}, \quad \mathbf{P}(A_2) = \frac{1}{4} \cdot \sum_{k=1}^{\infty} \frac{1}{2^{2(k-1)}} = \frac{1}{3}.$$

Легко убедиться, что $\mathbf{P}(A_1) + \mathbf{P}(A_2) = 1$, что является проверкой правильности решения задачи, поскольку события A_1 и A_2 образуют полную группу.

Пример 6.3. Прибор вышел из строя вследствие выхода из строя одного из двух блоков. Известно, что первый блок выходит из строя с вероятностью P_1 , второй — с вероятностью P_2 . При однократной проверке неисправного блока причина неисправности обнаруживается с вероятностью P . Произведено n_1 проверок первого блока и n_2 проверок второго блока. Как распорядиться числами n_1 и n_2 , чтобы при фиксированном общем числе проверок $n = n_1 + n_2$ вероятность обнаружения неисправности была бы наибольшей?

Решение. Пусть гипотеза H_1 — прибор вышел из строя вследствие отказа первого блока, гипотеза H_2 — вследствие отказа второго блока. В соответствии с формулой условной вероятности

$$\mathbf{P}(H_1) = \frac{P_1}{P_1 + P_2}, \quad \mathbf{P}(H_2) = \frac{P_2}{P_1 + P_2}.$$

Если A — обнаружение неисправности, то

$$\mathbf{P}(A|H_1) = 1 - (1 - P)^{n_1}, \quad \mathbf{P}(A|H_2) = 1 - (1 - P)^{n_2}.$$

Следовательно, безусловная вероятность события A

$$\mathbf{P}(A) = \frac{P_1}{P_1 + P_2} [1 - (1 - P)^{n_1}] + \frac{P_2}{P_1 + P_2} [1 - (1 - P)^{n_2}].$$

Положив в этой формуле $n_2 = n - n_1$, найдем максимальное значение $\mathbf{P}(A)$, рассматривая n_1 как непрерывную переменную, то есть определим n_1 так, чтобы $\frac{d\mathbf{P}(A)}{dn_1} = 0$. Выполнив очевидные преобразования, получим

$$n_1 = \frac{1}{2} \left[n + \frac{\ln \frac{P_2}{P_1}}{\ln(1 - P)} \right].$$

Найденное значение n_1 будет, вообще говоря, нецелым, поэтому в качестве окончательного значения n_1 нужно взять ближайшее целое число (приближенное решение) или найти вероятность $\mathbf{P}(A)$ для двух ближайших к n_1 целых чисел n'_1 и n^*_1 , ($n'_1 \leq n_1 \leq n^*_1$) и выбрать то из них, для которого вероятность $\mathbf{P}(A)$ будет иметь наибольшее значение.

Пример 6.4 (Продолжение примера 6.3). Первый блок из примера 6.3 был подвергнут двум проверкам, и никаких неисправностей в нем обнаружено не было. Какова вероятность того, что и в случае третьей проверки в нем также не будет обнаружено никаких дефектов?

Решение. Введем гипотезы H_1 — первый блок поврежден, H_2 — первый блок исправен. Аналогично предыдущему примеру

$$\mathbf{P}(H_1) = \frac{P_1}{P_1 + P_2}, \quad \mathbf{P}(H_2) = \frac{P_2}{P_1 + P_2}.$$

Введем события: A — первый блок выдержал две проверки, B — он выдержал третью проверку. В задаче нужно найти условную вероятность $\mathbf{P}(B|A)$.

Рассмотрим произведение событий A и B . Оно означает, что блок выдержит без отказов все три проверки. По формуле полной вероятности

$$\mathbf{P}(AB) = \mathbf{P}(AB|H_1)\mathbf{P}(H_1) + \mathbf{P}(AB|H_2)\mathbf{P}(H_2).$$

Легко видеть, что

$$\mathbf{P}(AB|H_1) = (1 - P)^3, \quad \mathbf{P}(AB|H_2) = 1,$$

поэтому

$$\mathbf{P}(AB) = \frac{(1 - P)^3 P_1 + P_2}{P_1 + P_2}.$$

Аналогичным образом ищем вероятность события A :

$$\begin{aligned} & \mathbf{P}(A)\mathbf{P}(A|H_1)\mathbf{P}(H_1) + \mathbf{P}(A|H_2)\mathbf{P}(H_2), \\ & \mathbf{P}(A|H_1) = (1 - P)^2, \quad \mathbf{P}(A|H_2) = 1, \\ & \mathbf{P}(A) = \frac{(1 - P)^2 P_1 + P_2}{P_1 + P_2}. \end{aligned}$$

Далее по теореме умножения вероятностей будем иметь

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(AB)}{\mathbf{P}(A)} = \frac{(1 - P)^3 P_1 + P_2}{(1 - P)^2 P_1 + P_2}.$$

Пример 6.5. Два игрока играют в следующую игру. Каждый выбирает произвольное целое число от 1 до n . Вероятность выбрать i очков для первого игрока равняется p_i , а для второго — q_i . Вектора $\vec{p} = \{p_i\}_{i=1}^n$ и $\vec{q} = \{q_i\}_{i=1}^n$ называются стратегиями игроков. Далее играющие проводят партию игры, исход которой заранее не определен. Если первый игрок выбирает i очков, а второй — j очков, то вероятность выигрыша первого игрока составляет α_{ij} . Матрица $\alpha = \|\alpha_{i,j}\|_{i,j=1}^n$ называется матрицей игры. Определить вероятность выигрыша первого игрока. При заданной матрице игры и заданной стратегии второго игрока найти оптимальную стратегию для первого из них.

Решение. Введем систему гипотез H_{ij} — первый игрок выбросил i , а второй — j очков. По формуле полной вероятности, обозначая за A выигрыш первого игрока, получим

$$\mathbf{P}(A) = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} p_i q_j = \vec{p}^T \alpha \vec{q},$$

где T обозначает символ транспонирования.

Введем вектор

$$\vec{\beta} = \alpha \vec{q}$$

и оптимизируем при заданном $\vec{\beta}$ стратегию первого игрока $\vec{p}$. Требуется решить следующую задачу на условный экстремум

$$\begin{aligned} \vec{p}^T \vec{\beta} &\rightarrow \max, \\ p_i &\geq 0, \quad \sum_{i=1}^n p_i = 1. \end{aligned}$$

Поскольку и целевая функция и ограничения линейны, поставленная задача относится к задачам линейного программирования [5], для решения которых разработаны специальные методы, в частности, симплекс-метод, однако, в данном случае поставленная задача линейного программирования может быть легко решена элементарными средствами.

Действительно, обозначим через k номер наибольшей компоненты вектора $\vec{\beta}$, то есть будем считать, что

$$\beta_{\max} = \beta_k > \beta_i, \quad i = \overline{1, n}; \quad i \neq k.$$

Представим целевую функцию в виде

$$\mathbf{P}(A) = \sum_{i=1}^n p_i \beta_i = \beta_{\max} \sum_{i=1}^n p_i + \sum_{i=1}^k p_i (\beta_i - \beta_{\max}).$$

Вследствие поставленных ограничений на p_i

$$\mathbf{P}(A) = \beta_{\max} + \sum_{i=1, i \neq k}^k p_i (\beta_i - \beta_{\max}).$$

В последнем выражении второе слагаемое не может быть положительным и обращается в нуль, только если $p_i \equiv 0$ при всех $i \neq k$. Следовательно, оптимальная стратегия первого игрока заключается в том, чтобы перед началом партии выбрать k очков

$$p_i^{(opt)} = \begin{cases} 1, & i = k, \\ 0, & i \neq k, \end{cases}$$

при этом вероятность выигрыша

$$\mathbf{P}(A)_{opt} = \beta_{\max}.$$

Стратегия, при которой одна из компонент $\vec{p}$ равна единице, а остальные обращаются в нуль, называется *чистой* стратегией. В данной задаче возможны и более сложные оптимальные стратегии игры, которые уже будут *смешанными*.

Допустим, что наибольшая компонента $\vec{\beta}$ не единственна, а имеются, скажем, s компонент с номерами $k_1, \dots, k_s$, превосходящих остальные компоненты и равных друг другу

$$\beta_{\max} = \beta_{k_1} = \dots = \beta_{k_s} > \beta_i, \quad (i = \overline{1, n}; \quad i \neq k_j; \quad j = \overline{1, s}).$$

Тогда, повторяя предыдущие рассуждения, получим

$$p_i^{(opt)} = 0, \quad (i = \overline{1, n}; \quad i \neq k_j, \quad j = \overline{1, s}).$$

Что касается остальных компонент $\vec{p}$, то они могут выбираться произвольно, лишь бы соблюдались поставленные в условии ограничения на p_i . Таким образом, первый игрок должен выбирать с произвольными вероятностями одно из чисел $k_1, \dots, k_s$. На соответствующие вероятности накладывается единственное ограничение — условие нормировки

$$\sum_{j=1}^s p_{k_j} = 1.$$

Всем этим стратегиям соответствует одно и то же значение вероятности выигрыша $\beta_{\max}$.

Пример 6.6. Для моделирования случайных чисел, представленных в двоичной системе в виде

$$x = \sum_{i=0}^n c_i 2^i,$$

строится последовательность $\{c_i\}_{i=0}^n$, в которой случайным образом чередуются цифры 0 и 1. Значения c_i вычисляются рекуррентно с помощью следующего случайного эксперимента. Вначале принимается $c_0 = 0$. Пусть известна i -я цифра числа c_i . Тогда с вероятностью $0 < \alpha < 1$ следующая $(i+1)$ -я цифра $c_{i+1} = c_i$, а с вероятностью $(1 - \alpha)$ имеем $c_{i+1} \neq c_i$. Требуется определить вероятность $p_j(n)$ того, что в старшем разряде формируемого случайного числа будет получена цифра j , $j = \overline{0, 1}$.

Решение. Сформируем вектор $\vec{p}(k) = \{p_0(k), p_1(k)\}$, состоящий из вероятностей $p_j(k)$. По условию $\vec{p}(0) = \{1, 0\}$. Допустим, что нам известен результат моделирования k -й цифры. Введем систему гипотез $H_j^{(k)} - c_k = j$, $(k = \overline{1, n-1}; j = \overline{0, 1})$. Ясно, что вероятности таких гипотез $\mathbf{P}(H_i^{(k)}) = p_i(k)$.

Рассмотрим формирование $(k+1)$ -й цифры числа. Система гипотез $H_0^{(k)}, H_1^{(k)}$ является полной. В соответствии со сформулированными в условии правилами моделирования, по формуле полной вероятности находим

$$\begin{cases} p_0(k+1) = \alpha p_0(k) + (1 - \alpha) p_1(k), \\ p_1(k+1) = (1 - \alpha) p_0(k) + \alpha p_1(k) \end{cases}$$

или в векторной форме

$$\vec{p}(k+1) = A \vec{p}(k),$$

где матрица A имеет вид

$$A = \begin{vmatrix} \alpha & (1 - \alpha) \\ (1 - \alpha) & \alpha \end{vmatrix}.$$

Полученное рекуррентное соотношение для $\vec{p}(k)$ позволяет записать решение в форме

$$\vec{p}(k) = A^k \vec{p}(0),$$

и задача сведется к вычислению степеней матрицы A .

С помощью метода математической индукции нетрудно показать, что

$$A^k = \frac{1}{2} \begin{vmatrix} 1 + (2\alpha - 1)^k & 1 - (2\alpha - 1)^k \\ 1 - (2\alpha - 1)^k & 1 + (2\alpha - 1)^k \end{vmatrix}.$$

При $k = 1$ это равенство очевидно, а при произвольном $k > 1$ оно следует из матричного тождества

$$\begin{aligned} & \begin{vmatrix} 1 + (2\alpha - 1)^k & 1 - (2\alpha - 1)^k \\ 1 - (2\alpha - 1)^k & 1 + (2\alpha - 1)^k \end{vmatrix} \cdot \begin{vmatrix} \alpha & (1 - \alpha) \\ (1 - \alpha) & \alpha \end{vmatrix} = \\ &= \begin{vmatrix} 1 + (2\alpha - 1)^{k+1} & 1 - (2\alpha - 1)^{k+1} \\ 1 - (2\alpha - 1)^{k+1} & 1 + (2\alpha - 1)^{k+1} \end{vmatrix}. \end{aligned}$$

Вспоминая выражение для $\vec{p}(0)$, получаем

$$\vec{p}(n) = \frac{1}{2} \{1 + (2\alpha - 1)^n, 1 - (2\alpha - 1)^n\}.$$

По условию $0 < \alpha < 1$, то есть исключаются случаи $\alpha = 0$ и $\alpha = 1$, когда последовательность $\{c_i\}$ сплошь состоит из нулей или единиц. Поэтому $|2\alpha - 1| < 1$ и асимптотически

$$\lim_{n \rightarrow \infty} p_0(n) = \lim_{n \rightarrow \infty} p_1(n) = \frac{1}{2}.$$

Это означает, что для достаточно длинных чисел появление в старшем разряде нуля и единицы равновероятно.

Рассмотренное в данном примере сложное событие (появление заданной цифры в старшем разряде числа) определенным образом зависит от исходов других, ранее произошедших событий (появления тех или иных цифр в младших разрядах). Такие случайные эксперименты довольно часто встречаются в прикладных задачах, и их в общем виде можно описать следующим образом.

Допустим, что проводится серия из n испытаний. В каждом испытании возможны m исходов $A_1, \dots, A_m$, образующих полную группу событий. Если вероятности исходов в каждом очередном испытании зависят от результатов ранее проведенных испытаний, то говорят, что имеет место *цепь случайных событий*.

Цепи событий подробно изучал в начале XX в. А. А. Марков. Он выделил важный класс таких цепей, для которого вероятности исходов в любом испытании зависят от результата только одного непосредственно предшествующего ему испытания и не зависят от того, как закончились все ранее проведенные испытания. Это свойство, называемое *марковским свойством*, означает отсутствие у системы памяти. Будущее поведение такой системы зависит только от ее нынешнего состояния и не зависит от того, как это состояние было достигнуто. Цепь, обладающую указанным свойством, называют *цепью Маркова*.

В разобранным примере мы имеем конечную ($m = 2 < \infty$) и однородную ($A = \text{const}(k)$) цепь Маркова. Цепи Маркова имеют важное прикладное значение [31], однако их подробное изучение выходит за рамки данного курса. Отметим лишь, что формула полной вероятности является одним из основных инструментов изучения таких цепей.

§ 7. ТЕОРЕМА ГИПОТЕЗ (ФОРМУЛА БАЙЕСА)

Будем исходить из тех же предположений, что и при выводе формулы полной вероятности, то есть из предположения, что имеется полная группа гипотез H_k с вероятностями $\mathbf{P}(H_k)$ и для каждой гипотезы известна условная вероятность $\mathbf{P}(A|H_k)$ появления события A .

Кроме того, предположим, что проведен опыт, в результате которого наблюдалось событие A . Теорема гипотез дает ответ на вопрос, как изменятся вероятности гипотез после окончания опыта. Вероятности $\mathbf{P}(H_k|A)$ гипотез после испытания называют *апостериорными*, в отличие от вероятностей $\mathbf{P}(H_k)$, которые в этом случае называют *априорными*.

Для доказательства теоремы гипотез (формулы Т. Байеса⁸) применим теорему умножения вероятностей к произведению событий A и H_k , записав ее в двух видах:

$$\begin{aligned}\mathbf{P}(AH_k) &= \mathbf{P}(A)\mathbf{P}(H_k|A), \\ \mathbf{P}(AH_k) &= \mathbf{P}(H_k)\mathbf{P}(A|H_k).\end{aligned}\tag{7.1}$$

Отсюда и следует искомая формула Байеса

$$\mathbf{P}(H_k|A) = \frac{\mathbf{P}(H_k)\mathbf{P}(A|H_k)}{\mathbf{P}(A)}, \quad (k = \overline{1, n}). \tag{7.2}$$

Вероятность $\mathbf{P}(A)$ события A здесь определяется по формуле полной вероятности (6.3), а именно

$$\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(H_k)\mathbf{P}(A|H_k). \tag{7.3}$$

Просуммировав равенства (7.2) по всем значениям k , получим

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

откуда следует

$$\sum_{k=1}^n \mathbf{P}(H_k|A) = 1. \tag{7.4}$$

Таким образом, после проведения испытания, как и до его проведения, гипотезы H_k по-прежнему составляют полную группу событий.

Рассмотрим примеры применения формулы Байеса.

Пример 7.1 (см. пример 6.1). Из десяти приборов наудачу выбран один, который дал отказ. Какова вероятность того, что взятый наудачу прибор был исправным, если исправный прибор дает отказ с вероятностью 0,01, дефектный прибор — с вероятностью 0,85, а среди всех десяти приборов лишь один был дефектным?

⁸ Байес Томас (1702–1761) — английский математик.

Решение. Гипотеза H_1 — взятый наудачу прибор исправный, H_2 — взятый прибор дефектный. Априорные вероятности гипотез $\mathbf{P}(H_1) = 9/10$, $\mathbf{P}(H_2) = 1/10$. Условные вероятности получения отказа $\mathbf{P}(A|H_1) = 0,01$, $\mathbf{P}(A|H_2) = 0,85$. Апостериорные вероятности

$$\begin{aligned}\mathbf{P}(H_1|A) &= \frac{\mathbf{P}(H_1)\mathbf{P}(A|H_1)}{\sum_{j=1}^2 \mathbf{P}(H_j)\mathbf{P}(A|H_j)} = \\ &= \frac{0,9 \cdot 0,01}{0,9 \cdot 0,01 + 0,1 \cdot 0,85} = \frac{9}{94} \approx 0,096,\end{aligned}$$

$$\begin{aligned}\mathbf{P}(H_2|A) &= \frac{\mathbf{P}(H_2)\mathbf{P}(A|H_2)}{\sum_{j=1}^2 \mathbf{P}(H_j)\mathbf{P}(A|H_j)} = \\ &= \frac{0,1 \cdot 0,85}{0,9 \cdot 0,01 + 0,1 \cdot 0,85} = \frac{85}{94} \approx 0,904.\end{aligned}$$

Искомая вероятность $\mathbf{P}(H_1|A)$ после испытания стала примерно в 10 раз меньше своего априорного значения $\mathbf{P}(H_1)$. Вторая апостериорная вероятность $\mathbf{P}(H_2|A)$ может быть использована для проверки правильности численных расчетов, поскольку согласно (7.4) должно быть

$$\mathbf{P}(H_1|A) + \mathbf{P}(H_2|A) = 1.$$

Пример 7.2. Получена партия из восьми изделий. При проверке половины партии три оказались технически исправными, а одно — бракованным. Какова вероятность, что при проверке трех последующих изделий одно из них окажется исправным, а два — бракованными, если любое количество бракованных изделий в данной партии равновероятно?

Решение. Пусть случайное событие A означает, что из четырех взятых изделий три исправных, а одно бракованное. В качестве гипотезы H_k возьмем событие, означающее, что среди восьми изделий имеется ровно k исправных. До первой

проверки изделий все введенные гипотезы были равновозможны, то есть

$$\mathbf{P}(H_k) = \frac{1}{9}, \quad (k = \overline{0, 8}).$$

Если в партии из восьми изделий исправных k , а бракованных $(8 - k)$, то вероятность взять три исправных изделия и одно бракованное определяется формулой

$$\mathbf{P}(A|H_k) = \frac{C_k^3 C_{8-k}^1}{C_8^4}, \quad (k = \overline{0, 8}).$$

Отсюда получаем, что

$$\begin{aligned} \mathbf{P}(A|H_0) = 0, \quad \mathbf{P}(A|H_1) = 0, \quad \mathbf{P}(A|H_2) = 0, \quad \mathbf{P}(A|H_3) = \frac{1}{14}, \\ \mathbf{P}(A|H_4) = \frac{8}{35}, \quad \mathbf{P}(A|H_5) = \frac{3}{7}, \quad \mathbf{P}(A|H_6) = \frac{4}{7}, \\ \mathbf{P}(A|H_7) = \frac{1}{2}, \quad \mathbf{P}(A|H_8) = 0. \end{aligned}$$

Используя формулу полной вероятности, находим

$$\mathbf{P}(A) = \frac{1}{9} \cdot \left(\frac{1}{14} + \frac{8}{35} + \frac{3}{7} + \frac{4}{7} + \frac{1}{2} \right) = 0,2.$$

После проверки половины партии, в результате произошло событие A , согласно формуле Байеса для вероятностей гипотез получаем следующие новые значения:

$$\begin{aligned} \mathbf{P}(H_0|A) = \mathbf{P}(H_1|A) = \mathbf{P}(H_8|A) = 0, \\ \mathbf{P}(H_3|A) = \frac{5}{126}, \quad \mathbf{P}(H_4|A) = \frac{8}{63}, \quad \mathbf{P}(H_5|A) = \frac{5}{21}, \\ \mathbf{P}(H_6|A) = \frac{20}{63}, \quad \mathbf{P}(H_7|A) = \frac{5}{18}. \end{aligned}$$

Обозначим через B случайное событие, состоящее в том, что при проверке трех последующих изделий одно из них окажется исправным, а два — бракованными. Событие B может

произойти только совместно с гипотезами H_4 и H_5 , так как только в этом случае общая партия содержит не менее четырех исправных и трех бракованных изделий. Вероятности указанных гипотез после первой проверки равны соответственно $8/63$ и $5/21$. Вероятности события B при этих гипотезах будут такими

$$\mathbf{P}(B|H_4) = \frac{C_1^1 C_3^2}{C_4^3} = \frac{3}{4}, \quad \mathbf{P}(B|H_5) = \frac{C_2^1 C_2^2}{C_4^3} = \frac{1}{2}.$$

Воспользовавшись формулой полной вероятности, получаем

$$\mathbf{P}(B) = \frac{8}{63} \cdot \frac{3}{4} + \frac{5}{21} \cdot \frac{1}{2} = \frac{3}{14}.$$

Пример 7.3. Прибор вышел из строя вследствие неисправности одного из двух блоков. Согласно паспорту прибора при данном сроке эксплуатации вероятность выхода из строя первого блока P_1 в три раза превосходит вероятность выхода из строя второго блока P_2 . Вышедший из строя блок можно обнаружить путем тестовых испытаний. При однократной проверке дефектного блока неисправность выявляется с вероятностью $P = 75\%$. Была проведена серия из $n = 5$ независимых проверок первого блока, n_1 выбиралось так, чтобы вероятность обнаружения неисправности была максимальна (см. пример 6.3). Испытания не позволили обнаружить неисправность. Какова вероятность наличия дефектов в i -м блоке с учетом проведенных испытаний ($i = \overline{1, 2}$)?

Решение. Гипотеза H_i — неисправен i -й блок ($i = \overline{1, 2}$). Априорные вероятности гипотез

$$\mathbf{P}(H_1) = \frac{P_1}{P_1 + P_2} = \frac{3}{4}, \quad \mathbf{P}(H_2) = \frac{P_2}{P_1 + P_2} = \frac{1}{4}.$$

Пусть A — отсутствие отказов в обоих блоках по результатам испытаний. Условные вероятности:

$$\mathbf{P}(A|H_1) = (1 - P)^{n_1}, \quad \mathbf{P}(A|H_2) = (1 - P)^{n_2}.$$

Вероятность события A по формуле полной вероятности

$$\mathbf{P}(A) = \frac{P_1}{P_1 + P_2}(1 - P)^{n_1} + \frac{P_2}{P_1 + P_2}(1 - P)^{n - n_1}, \quad (n = \overline{1, 4}).$$

Числовые значения вероятности $\mathbf{P}(A)$ приводятся в таблице 7.1.

Таблица 7.1
Вероятность отсутствия отказов
при испытаниях

n_1	0	1	2	3	4	5
$\mathbf{P}(A)$	0,750	0,189	0,0508	0,0273	0,0654	0,251

Легко видеть, что наименьшая вероятность безрезультатного окончания испытаний имеет место при $n = 3$ и составляет менее 3%. Интересно заметить, что приближенная формула для n_1 в примере 6.3, в данном случае дает

$$n_1 = \frac{1}{2} \left[n + \frac{\ln \frac{P_1}{P_2}}{\ln(1 - P)} \right] = \frac{1}{2} \left[5 + \frac{\ln 3}{\ln 4} \right] = 2,896,$$

что хорошо согласуется с точным оптимальным значением $n_1^{(opt)} = 3$.

При $n_1 = 3$ получаем по формуле Байеса

$$\begin{aligned} \mathbf{P}(A) &= \frac{3}{4} \left(\frac{1}{4} \right)^3 + \frac{1}{4} \left(\frac{1}{4} \right)^2 = \frac{7}{256} \approx 0,273, \\ \mathbf{P}(H_1|A) &= \frac{\mathbf{P}(H_1)\mathbf{P}(A|H_1)}{\mathbf{P}(A)} = \frac{\frac{3}{4} \cdot \left(\frac{1}{4} \right)^3}{(7/256)} = \frac{3}{7} \approx 0,429, \\ \mathbf{P}(H_2|A) &= \frac{\mathbf{P}(H_2)\mathbf{P}(A|H_2)}{\mathbf{P}(A)} = \frac{\frac{1}{4} \cdot \left(\frac{1}{4} \right)^2}{(7/256)} = \frac{4}{7} \approx 0,571. \end{aligned}$$

Поскольку в результате испытаний произошло весьма маловероятное в случае неисправности первого блока событие A , то

соотношение вероятностей гипотез изменилось. Более вероятным стало предположение, что все-таки неисправен априори надежный второй блок.

Пример 7.4 (Продолжение примера 7.3). Первый блок из примера 7.3 проверяли k раз, но никаких неисправностей в нем не было обнаружено. Какова вероятность того, что данный блок полностью технически исправен?

Решение. Введем гипотезы: H_1 — проверяемый блок неисправен, H_2 — блок полностью исправен, а также событие A — после k испытаний дефектов не выявлено.

$$\mathbf{P}(A|H_1) = (1 - P)^k, \quad \mathbf{P}(A|H_2) = 1,$$

$$\mathbf{P}(H_1) = \frac{P_1}{P_1 + P_2}, \quad \mathbf{P}(H_2) = \frac{P_2}{P_1 + P_2}.$$

По формуле полной вероятности

$$\mathbf{P}(A) = \sum_{i=1}^2 \mathbf{P}(A|H_i) \mathbf{P}(H_i) = \frac{P_1(1 - P)^k + P_2}{P_1 + P_2}$$

искомая априорная вероятность

$$\mathbf{P}(H_2|A) = \frac{\mathbf{P}(A|H_2) \mathbf{P}(H_2)}{\mathbf{P}(A)} = \frac{P_2}{P_1(1 - P)^k + P_2}.$$

При неограниченном увеличении числа проверок, которые не выявляют каких-либо дефектов

$$\lim_{k \rightarrow \infty} \mathbf{P}(H_2|A) = 1.$$

Физический смысл этого равенства очевиден: если при тщательной проверке блока никаких неисправностей не обнаруживается, то естественно считать данный блок полностью исправным. В связи с этим возникает вопрос: сколько проверок k нужно провести, чтобы можно было считать факт исправности практически установленным. Рассмотрим этот вопрос подробнее.

Пример 7.5 (Продолжение примера 7.4). Определить, при каком числе проверок k апостериорная вероятность исправности блока из предыдущего примера будет превосходить $1 - \varepsilon$, где ε — заданное малое число ($\varepsilon \ll 1$).

Решение. Рассмотрим неравенство

$$\mathbf{P}(H_2|A) > 1 - \varepsilon.$$

Подставляя в него выражение апостериорной вероятности, полученное в предыдущем примере, получаем неравенство для определения искомого ограничения на число проверок k

$$\frac{P_2}{P_1(1-P)^k + P_2} > 1 - \varepsilon.$$

Разрешая полученное неравенство относительно $(1-P)^k$, находим

$$(1-P)^k < \frac{\varepsilon P_2}{(1-\varepsilon)P_1}.$$

При любом, сколь угодно малом ε всегда можно подобрать такое число испытаний $k_{\min}$, что при $k \geq k_{\min}$ это неравенство будет выполняться. Минимальное число требуемых для этого испытаний определяется равенством

$$k_{\min} = \left\lceil \frac{\ln \frac{\varepsilon P_2}{(1-\varepsilon)P_1}}{\ln(1-P)} \right\rceil + 1,$$

где квадратные скобки обозначают символ целой части числа.

Для числовых данных, приведенных в примере 7.3, критическое число испытаний $k_{\min}$ представлено в таблице 7.2.

Таблица 7.2

Критическое число проверок

ε	0,1	0,01	0,001	0,0001
$k_{\min}$	3	5	6	8

Как видим, требуемое число проверок сравнительно невелико и позволяет с высокой надежностью выявить исправный блок.

ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТИ ПОЯВЛЕНИЯ СОБЫТИЯ ПРИ ПОВТОРНЫХ ИСПЫТАНИЯХ

§ 8. ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ ПОЯВЛЕНИЯ СОБЫТИЯ В СЕРИИ НЕЗАВИСИМЫХ ИСПЫТАНИЙ С ДВУМЯ ВОЗМОЖНЫМИ ИСХОДАМИ

Предположим, что проводится серия из n *независимых испытаний*, в результате каждого из которых происходит событие A , или противоположное ему событие $\bar{A}$. Вероятность $P(A)$ появления события A при различных испытаниях может либо оставаться неизменной, либо меняться от испытания к испытанию. Независимость испытаний означает, что вероятность появления события A при любом очередном испытании не зависит от результатов всех предыдущих испытаний. Подобная схема впервые была рассмотрена Я. Бернулли и называется *схемой Бернулли*.

Примерами независимых испытаний являются: подбрасывание монеты или игральной кости, испытание различных приборов на надежность, регистрация числа телефонных вызовов за определенный интервал времени и тому подобное. Если условия опыта остаются неизменными, то вероятности различных исходов испытания остаются постоянными. В противном случае эти вероятности могут меняться от опыта к опыту, хотя испытания по-прежнему остаются независимыми. Например, число телефонных вызовов, поступающих за один час на некоторый коммуникатор, существенно зависит от времени суток, хотя появление отдельного вызова можно считать

случайным событием, не зависящим от тех вызовов, которые уже имели место до этого момента времени.

При решении многих прикладных задач возникает необходимость определения вероятности $P_m^{(n)}$ того, что в серии из n независимых испытаний событие A произойдет ровно m раз, ($m = \overline{0, n}$). Такие задачи возникают при контроле качества продукции, в телефонии, в различного рода информационных, вычислительных системах и так далее.

Найдем указанную вероятность сначала в простейшем случае, когда вероятность появления случайного события A при каждом испытании одинакова и равна p , а потому вероятность q появления противоположного события $\overline{A}$ при каждом испытании равна $q = 1 - p$.

Обозначим через A_k случайное событие, означающее появление события A при k -м испытании ($k = \overline{1, n}$). Введем также случайное событие B_m , состоящее в том, что при n испытаниях событие A произойдет ровно m раз. Данное событие может быть представлено в виде суммы некоторого числа несовместных событий. Каждое из них является произведением, составленным из m событий типа A_k и $n - m$ событий типа $\overline{A_k}$. Любое слагаемое соответствует одному из возможных способов появления события A ровно m раз при n независимых испытаниях. Таким образом, можно записать

$$\begin{aligned} B_m = & A_1 A_2 \dots A_m \overline{A_{m+1}} \overline{A_{m+2}} \dots \overline{A_n} + \\ & + A_1 A_2 \dots A_{m-1} \overline{A_m} \overline{A_{m+1}} \dots \overline{A_n} + \dots + \\ & + \overline{A_1} \overline{A_2} \dots \overline{A_{n-m}} A_{n-m+1} \dots A_n. \end{aligned} \quad (8.1)$$

Первое слагаемое в правой части равенства (8.1) соответствует случаю, когда событие A происходит уже при первых m испытаниях, а затем ни разу не происходит. Последнее слагаемое соответствует другому крайнему случаю, когда при первых $(n - m)$ испытаниях событие A не имело места, а затем произошло m раз подряд.

Вероятность того, что событие A появится m раз и не появится $(n - m)$ раз при определенном варианте чередования появлений и непоявлений этого события, не зависит от характера чередований событий A , $\overline{A}$ и равна $p^m q^{n-m}$. Так как

искомая вероятность $P_m^{(n)} = \mathbf{P}(B_m)$, а B_m , согласно (8.1), является суммой C_n^m равновероятных несовместных событий, то для вероятности $P_m^{(n)}$ появления события A ровно m раз при n независимых испытаниях, проводимых в одинаковых условиях, справедлива следующая формула, полученная впервые Я. Бернулли:

$$P_m^{(n)} = C_n^m p^m q^{n-m}. \quad (8.2)$$

Вероятность $P_m^{(n)}$ равна коэффициенту при $p^m q^{n-m}$ в разложении бинома $(p+q)^n$ по степеням p . Поэтому совокупность вероятностей $P_m^{(n)}$ называют *биномиальным законом распределения*. Сумма вероятностей $P_m^{(n)}$, очевидно, равна единице, следовательно, события B_m образуют полную группу. Действительно, учитывая (8.2), имеем

$$\sum_{m=0}^n P_m^{(n)} = \sum_{m=0}^n C_n^m p^m q^{n-m} = (p+q)^n = 1. \quad (8.3)$$

Через вероятности $P_m^{(n)}$ появления события A при n испытаниях m раз можно выразить и ряд других важных вероятностей, представляющих интерес в приложениях. Например, вероятность $R_m^{(n)}$ того, что при n независимых испытаниях событие A произойдет не менее m раз, дается равенством

$$R_m^{(n)} = \sum_{k=m}^n P_k^{(n)}. \quad (8.4)$$

Воспользовавшись тождеством $\sum_{k=0}^n P_k^{(n)} = 1$, эту формулу можно также записать в виде

$$R_m^{(n)} = 1 - \sum_{k=0}^{m-1} P_k^{(n)}. \quad (8.5)$$

При небольших m формула (8.5) удобнее для вычисления, чем формула (8.4).

Вероятность $P_0^{(n)}$ того, что при n испытаниях событие A не произойдет ни разу, равна $P_0^{(n)} = q^n$. Тогда вероятность

$R_1^{(n)}$ того, что при n независимых испытаниях событие A произойдет хотя бы раз, согласно (8.5), определится формулой

$$R_1^{(n)} = 1 - q^n. \quad (8.6)$$

Формула (8.6) позволяет найти необходимое число испытаний n , при которых событие A произойдет хотя бы один раз с вероятностью не менее заданного значения P . Замечаем, что наименьшее вещественное значение n , при котором выполняется указанное условие, равно $\frac{\ln(1-P)}{\ln(1-p)}$. Вероятность $R_1^{(n)}$ поэтому будет не менее заданного значения P , если число опытов

$$n \geq \left[\frac{\ln(1-P)}{\ln(1-p)} \right] + 1, \quad (8.7)$$

где квадратные скобки обозначают целую часть числа.

Рассмотрим зависимость вероятности $P_m^{(n)}$ от числа m при неизменном числе испытаний n . Имеем согласно (8.2)

$$\frac{P_{m+1}^{(n)}}{P_m^{(n)}} = \frac{(n-m)p}{(m+1)q}.$$

Обозначим через m_0 число появлений события A , при котором $P_m^{(n)}$ достигает наибольшего значения, и будем называть это число *вероятнейшим числом* появления события A в серии из n опытов. Если m_0 — вероятнейшее значение, то справедливы неравенства

$$P_{m_0-1}^{(n)} \leq P_{m_0}^{(n)}, \quad P_{m_0}^{(n)} \geq P_{m_0+1}^{(n)},$$

что с учетом предыдущей формулы дает

$$np - q \leq m_0 \leq np + p. \quad (8.8)$$

Поскольку m_0 — число целое, то при $(np + p) < 1$ вероятнейшим значением будет $m_0 = 0$, то есть наиболее вероятным исходом рассматриваемой серии будет непоявление события A ни разу. Если же $(np + p) > 1$, и не является целым, то вероятнейшее число появлений события A больше нуля и равно

тому целому числу, которое попадает в интервал $[np - q, np + p]$ (поскольку длина этого интервала равна единице, такое число всегда существует). Следовательно, вероятности $P_m^{(n)}$ с ростом m вначале при $m < m_0$ возрастают, а затем при $m > m_0$ начинают убывать. Наконец, если $(np + p)$ — целое число, то целым будет и левая граница неравенств (8.8). В этом случае вероятнейших значений два:

$$m'_0 = (n + 1)p, \quad m''_0 = np - q = m'_0 - 1.$$

Рассмотрим теперь общий случай независимых испытаний, когда вероятность появления события A может меняться от испытания к испытанию. Такие испытания иногда называют *неоднородными испытаниями Бернулли*. В этом случае искомая вероятность $P_m^{(n)}$ появления события A ровно m раз при n независимых испытаниях также совпадает с вероятностью $\mathbf{P}(B_m)$ события B_m , определяемого формулой (8.1). Все C_n^m слагаемых данного выражения по-прежнему остаются несовместными событиями, однако, вероятности этих событий теперь будут различны, так как они зависят от того, в каких именно испытаниях появилось m раз событие A . Учитывая это отличие, для искомой вероятности $P_m^{(n)}$ получаем

$$\begin{aligned} P_m^{(n)} = & p_1 p_2 \dots p_m q_{m+1} q_{m+2} \dots q_n + \\ & + p_1 p_2 \dots p_{m-1} q_m q_{m+1} \dots q_{n-1} p_n + \\ & + \dots + q_1 q_2 \dots q_{n-m} p_{n-m+1} p_{n-m+2} \dots p_n, \end{aligned} \quad (8.9)$$

где в правой части равенства должны быть выписаны все C_n^m возможных комбинаций m множителей типа p_k и $n - m$ множителей типа q_j , причем номера индексов k и j должны равняться порядковому номеру сомножителя. В частных случаях формула (8.9) принимает более простой вид. Например, для вероятности $P_0^{(n)}$ того, что событие A не произойдет ни одного раза, и для вероятности $R_1^{(n)}$ того, что событие A произойдет хотя бы один раз, получим

$$P_0^{(n)} = \prod_{q=1}^n q_k, \quad R_1^{(n)} = 1 - \prod_{k=1}^n q_k. \quad (8.10)$$

Особый интерес представляет рассмотрение схемы Бернулли в том случае, когда вероятность p появления события A в одном опыте мала, но число опытов n настолько велико, что произведение np не является малым. Практическое значение этого случая связано с тем, что произведение np , как это будет видно из дальнейшего, позволяет судить о среднем числе появлений события A в серии испытаний, которую можно рассматривать как последовательность очень большого числа испытаний (в пределе — бесконечно большого), в каждом из которых событие A может появиться с весьма малой (в пределе нулевой) вероятностью. К таким событиям относятся, в частности, радиоактивный распад, при котором вероятность распада каждого ядра исчезающе мала, однако, число ядер настолько велико, что общее число ядер, распавшихся в некоторый интервал времени, уже не является малым. Аналогичная картина наблюдается в телефонии, где вероятность вызовов, поступающих на АТС (например, в течение часа) от определенного абонента, является величиной малой, но вследствие многочисленности абонентов общее число вызовов уже нельзя считать пренебрежимо малым. Аналогичная картина наблюдается и в ряде других прикладных задач теории массового обслуживания, теории надежности, в задачах экономического содержания и тому подобных.

Итак, рассмотрим вероятность $P_m^{(n)}$, определяемую формулой (8.2), в предельном случае, когда $n \rightarrow \infty$, $p \rightarrow 0$, но при этом $np = a = \text{const}$, то есть найдем предел

$$P_m = \lim_{\substack{n \rightarrow \infty \\ np=a}} C_n^m p^m (1-p)^{n-m}.$$

Положив под знаком предела $p = a/n$ и выразив число сочетаний C_n^m через факториалы, после сокращений и очевидных преобразований получим

$$P_m = \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-m+2)(n-m+1)}{n^m} \cdot \frac{a^m}{m!} \cdot \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^m}.$$

Дробь $\frac{n(n-1)\dots(n-m+1)}{n^m}$ распадается на конечное число m множителей вида $(n-j)/n$, каждый из которых стремится к единице при $n \rightarrow \infty$. Поэтому

$$\lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-m+1)}{n^m} = 1.$$

Воспользовавшись хорошо известным из анализа пределом

$$\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}$$

и учитывая, что при фиксированном показателе степени m

$$\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^m = 1,$$

окончательно получим

$$P_m = \frac{a^m}{m!} e^{-a}. \quad (8.11)$$

Совокупность вероятностей P_m , ($m = \overline{0, \infty}$) называется *законом Пуассона*, который иногда также называют *законом редких событий*.

Пример 8.1. Определить наиболее вероятное число m_0 появлений события A и соответствующую ему вероятность $P_{\max}$ при $n = 10$ независимых испытаниях, если вероятность появления события A при одном испытании $\mathbf{P}(A) = 0,21$ и не зависит от номера испытания.

Решение. Так как $(n+1)p = 2,31$ число дробное, то $m_0 = 2$,

$$P_{\max} = C_{10}^2 (0,21)^2 (0,79)^8 \approx 0,18.$$

Пример 8.2. Радиотехническое устройство имеет в своем составе следующие элементы: $n_1 = 1000$ электронных ламп, $n_2 = 2000$ транзисторов и $n_3 = 4000$ контактных соединений. Определить вероятность отказа устройства в течение заданного времени эксплуатации, если отказ наступает при выходе

из строя хотя бы одного из указанных элементов, выходы из строя каждого элемента — события независимые, а за указанное время каждая лампа может выйти из строя с вероятностью $p_1 = 0,001$, транзистор — с вероятностью $p_2 = 0,0004$, контактное соединение — с вероятностью $p_3 = 0,0005$.

Решение. Задача соответствует схеме Бернулли с общим числом испытаний $n = n_1 + n_2 + n_3$ и переменной вероятностью появления события A (отказ элемента).

Применяя формулу (8.10), получаем

$$\begin{aligned} P(A) &= R_1^{(n)} = 1 - (1 - p_1)^{n_1} (1 - p_2)^{n_2} (1 - p_3)^{n_3} = \\ &= 1 - \exp\{n_1 \ln(1 - p_1) + n_2 \ln(1 - p_2) + n_3 \ln(1 - p_3)\}. \end{aligned}$$

Так как вероятности $p_k (k = \overline{1, 3})$ малы, то, воспользовавшись степенным разложением для логарифма

$$\ln(1 - p_k) = -p_k - \frac{1}{2} p_k^2 - \frac{1}{3} p_k^3 - \dots,$$

получим

$$\begin{aligned} P(A) &\approx 1 - \exp\{-n_1 p_1 - n_2 p_2 - n_3 p_3\} = \\ &= 1 - \exp\{-1 - 0,8 - 2\} = 1 - e^{-3,8} = 0,9776, \end{aligned}$$

причем удержание одного члена при разложении $\ln(1 - p_k)$ обеспечивает в данном случае правильность всех десятичных знаков, указанных в ответе.

Пример 8.3. Число появлений события A подчиняется закону Пуассона с параметром $a = 2,5$. Определить вероятнейшее число появлений события A и соответствующую ему вероятность.

Решение. Переходя в формуле (8.8) к пределу $p \rightarrow 0$, ($np = a$), получаем

$$1,5 \leq m_0 \leq 2,5, \quad m_0 = 2.$$

По формуле (8.11) находим

$$P'_{m_0} = \frac{2,5^2}{2!} e^{-2,5} = 0,2564.$$

Пример 8.4. Проводятся испытания Бернулли, характеризуемые параметрами n и p . При каких значениях этих параметров вероятность получить четное число благоприятных исходов P_* превосходит вероятность получить нечетное их число P_{**} ?

Решение. Представим вероятность получения четного числа благоприятных исходов в виде

$$\begin{aligned} P_* &= \sum_{k=0}^{[n/2]} P_{2k}^{(n)} = \sum_{k=0}^{[n/2]} C_n^{2k} p^{2k} (1-p)^{n-2k} = \\ &= (1-p)^n \sum_{k=0}^{n/2} C_n^{2k} \left(\frac{p}{1-p} \right)^{2k}, \end{aligned}$$

где квадратные скобки обозначают целую часть числа.

По формуле бинома Ньютона¹

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

Легко видеть, что сумма одних лишь четных степеней λ может быть найдена отсюда путем сложения $(1+\lambda)^n$ и $(1-\lambda)^n$, в результате чего все нечетные степени λ пропадут, что дает

$$\sum_{k=0}^{[n/2]} C_n^{2k} \lambda^{2k} = \frac{1}{2} ((1+\lambda)^n + (1-\lambda)^n).$$

Применяя последнюю формулу к вычислению P_* , находим

$$\begin{aligned} P_* &= (1-p)^n \cdot \frac{1}{2} \left(\left(1 + \frac{p}{1-p} \right)^n + \right. \\ &\quad \left. + \left(1 - \frac{p}{1-p} \right)^n \right) = \frac{1}{2} (1 + (1-2p)^n). \end{aligned}$$

¹ Ньютон Исаак (1643–1727) — английский физик, механик, астроном и математик.

Подобным же образом вычисляя P_{**} , имеем

$$P_{**} = \frac{1}{2}(1 - (1 - 2p)^n).$$

Интересующее нас условие $P_* > P_{**}$ приводит к неравенству

$$(1 - 2p)^n > 0.$$

Отсюда следует, что при четном n должно выполняться условие $p \neq 1/2$, а при нечетном — условие $p < 1/2$. Иначе говоря, при четном числе опытов их исходы не должны быть равновероятны, а при нечетном числе опытов благоприятный исход каждого опыта должен быть менее вероятен, чем неблагоприятный.

Пример 8.5 (Блуждания Бернулли). Частица совершает случайные блуждания вдоль числовой оси, стартуя из начала координат. Она последовательно перемещается за один шаг на величину Δ , причем оба направления движения на каждом шаге равновозможны. Какова вероятность того, что за $2m$ шагов частица удалится от исходной точки на расстояние, не большее чем $2k \cdot \Delta$?

Решение. Поставленной задаче можно сопоставить схему Бернулли, состоящую из $2m$ опытов, где в роли события A_k выступает, например, движение частицы на данном шаге в положительном направлении, а в роли противоположного события $\overline{A_k}$ — в отрицательном направлении. По условию все перемещения, а значит, и все события A_k не зависят друг от друга.

Допустим, что из зарезервированных $2m$ шагов ровно N были сделаны в положительном направлении. Тогда $2m - N$ шагов были совершены в противоположном направлении. По условию имеем

$$|N - (2m - N)| \leq 2k.$$

Решая это неравенство относительно N , получим $m - k \leq N \leq m + k$, а тогда искомая вероятность $P_{k,m}$ выражается в виде

$$P_{k,m} = \sum_{j=m-k}^{m+k} P_j^{(2m)}.$$

По условию перемещения вправо и влево на каждом шаге равновероятны, следовательно, рассматриваемая схема Бернулли характеризуется параметром $p = q = \frac{1}{2}$, поэтому

$$P_j^{(2m)} = C_{c,m}^j \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{2m-j} = C_{2m}^j \left(\frac{1}{2}\right)^{2m}.$$

Искомая вероятность выражается в виде

$$P_{k,m} = \left(\frac{1}{2}\right)^{2m} \cdot \sum_{j=m-k}^{m+k} C_{2m}^j.$$

В частности, вероятность вернуться после $2m$ скачков в исходное положение дается равенством

$$P_{o,m} = \left(\frac{1}{2}\right)^{2m} \cdot C_{2m}^m.$$

Эту формулу можно пояснить следующим образом. Общее число всевозможных способов совершить $2m$ скачков вправо либо влево составляет 2^{2m} , причем благодаря равновозможности направлений скачка все эти 2^{2m} вариантов траектории движения частицы будут равновозможными. Интересующему нас событию благоприятствуют те варианты, в которых m скачков совершаются в прямом, и столько же — в обратном направлении. Число таких вариантов равняется числу способов разместить m положительных скачков в серии из $2m$ скачков, то есть числу сочетаний C_{2m}^m . В соответствии с классическим определением вероятности отсюда следует выражение для $P_{o,m}$, приведенное выше.

Представляет интерес асимптотика вероятности $P_{o,m}$ при больших m . Выражая число сочетаний C_{2m}^m через факториалы, находим

$$P_{o,m} = \left(\frac{1}{2}\right)^{2m} \frac{2m!}{(m!)^2}.$$

Далее воспользуемся формулой Стирлинга² [9]

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{\Theta}{12n}}, \quad (0 < \Theta < 1),$$

которая дает приближенное выражение для факториала при больших n

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

По формуле Стирлинга при $m \gg 1$

$$P_{o,m} \approx \frac{1}{\sqrt{\pi m}},$$

то есть, чем больше интервал наблюдения за частицей, тем менее вероятно ее возвращение в начальную точку в конце этого интервала.

Пример 8.6 (Обобщение примера 5.5). Жюри состоит из $(2k + 1)$ эксперта, каждый из которых независимо от других принимает правильное решение с вероятностью p . Итоговое решение жюри определяется большинством голосов. Какова вероятность того, что это итоговое решение будет правильным.

Решение. Поскольку решение принимается экспертами независимо друг от друга, и вероятность p для всех них одинакова, работу жюри можно описать с помощью испытаний Бернулли с параметрами $(2k + 1)$ и p . Вероятность вынесения справедливого вердикта таким составом жюри

$$q_{2k+1} = \sum_{j=k+1}^{2k+1} P_j^{(2k+1)} = \sum_{j=k+1}^{2k+1} C_{2k+1}^j p^j (1-p)^{2k+1-j}.$$

В частном случае $k = 1$ имеем

$$q_3 = C_3^2 p^2 (1-p) + C_3^3 p^3 = 3p^2 - 2p^3,$$

что совпадает с решением для этого частного случая, ранее полученным в примере 5.5.

² Стирлинг Джеймс (1692–1770) — шотландский математик.

Пример 8.7. Готовые изделия перед отправкой заказчикам проходят два этапа проверки. На первом этапе проверяются техническая исправность и работоспособность изделия. Образцы, не прошедшие эту проверку, изымаются и направляются на разборку и ремонт. Технически исправные изделия подвергаются углубленной проверке, в зависимости от результатов которой их относят к первому или второму сорту. Было изготовлено n изделий. Вероятность пройти первый этап испытаний для каждого из них равняется α , а вероятность того, что технически исправное изделие относится к первому сорту, равняется β . Какова вероятность того, что в данной партии окажется ровно k первосортных изделий?

Решение. В данном примере мы имеем дело с так называемыми *рекуррентными испытаниями Бернулли*. В этой модели последовательно проводится несколько серий обычных испытаний Бернулли, причем в каждой последующей серии испытаний участвуют только объекты, успешно прошедшие предыдущую серию. Общее число таких серий испытаний l может быть произвольным. Если на каком-то этапе испытаний все опыты закончились безрезультатно, повторные испытания заканчиваются, даже если не проведены все l серий. В нашем случае рассматривается простейшая двухэтапная схема проверки ($l = 2$).

Основная сложность таких задач состоит в том, что мы заранее знаем число опытов лишь на первом этапе испытаний. На всех последующих этапах это число может оказаться произвольным целым от 0 до n . Эту трудность можно обойти двумя способами.

Способ 1. Введем систему гипотез. H_j — первый этап проверки прошли j изделий ($j = \overline{0, n}$). По формуле полной вероятности искомая вероятность

$$q_k = \sum_j j = k^n P_j^{(k)}(\beta) P(H_j), \quad (k = \overline{0, n}),$$

где $P_j^{(k)}(\beta)$ обозначает вероятность вида (8.2), соответствующую $p = \beta$. В этих обозначениях

$$\mathbf{P}(H_j) = C_n^j \alpha^j (1 - \alpha)^{n-j} = P_n^{(j)}(\alpha).$$

В результате получим

$$q_k = \sum_{j=k}^n C_j^k \beta^k (1-\beta)^{j-k} C_n^j \alpha^j (1-\alpha)^{n-j}.$$

Заменяя в этой сумме индекс суммирования j на $i = j - k$, имеем

$$q_k = \sum_{i=0}^{n-k} C_n^{i+k} C_{i+k}^k \alpha^{i+k} (1-\alpha)^{n-k-i} \beta^k (1-\beta)^i.$$

Далее выразим оба числа сочетаний через факториалы

$$C_n^{i+k} = \frac{n!}{(i+k)!(n-i-k)!}, \quad C_{i+k}^k = \frac{(i+k)!}{k!i!}.$$

После очевидных алгебраических преобразований

$$q_k = (\alpha\beta)^k \frac{n!}{k!(n-k)!} \sum_{i=0}^{n-k} C_{n-k}^i (\alpha(1-\beta))^i (1-\alpha)^{n-k-i}.$$

Если в последующем равенстве ввести число сочетаний

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

и вычислить сумму по формуле бинома

$$\begin{aligned} \sum_{i=0}^{n-k} C_{n-k}^i (\alpha(1-\beta))^i (1-\alpha)^{n-k-i} &= \\ &= (\alpha(1-\beta) + (1-\alpha))^{n-k} = (1-\alpha\beta)^{n-k}, \end{aligned}$$

окончательно получим

$$q_k = C_n^k (\alpha\beta)^k (1-\alpha\beta)^{n-k}.$$

Легко видеть, что полученные вероятности соответствуют схеме Бернулли с параметром $p = \alpha\beta$. Это наводит на мысль,

что существует более простой и короткий способ решения. Он заключается в агрегировании, то есть укрупнении испытаний, переходе к сложным двухэтапным испытаниям каждого отдельно взятого изделия.

Способ 2. Введем случайные события A_{js} — изделие с номером j успешно прошло s -й этап испытаний ($j = \overline{1, n}$; $s = 1, 2$), а также события $A_j = A_{j1}A_{j2}$ — изделие j оказалось первосортным. По условию

$$\mathbf{P}(A_{j1}) = \alpha, \mathbf{P}(A_{j2}|A_{j1}) = \beta, (j = \overline{1, n}).$$

По теореме умножения вероятностей

$$\mathbf{P}(A_j) = \mathbf{P}(A_{j1}A_{j2}) = \mathbf{P}(A_{j1})\mathbf{P}(A_{j2}|A_{j1}) = \alpha\beta.$$

При такой трактовке испытаний

$$q_k = P_k^{(n)}(\alpha\beta) = C_n^k(\alpha\beta)^k(1 - \alpha\beta)^{n-k},$$

что совпадает с ранее полученным другим способом выражением.

§ 9. ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ ПОЯВЛЕНИЯ СОБЫТИЙ В СЕРИИ ИСПЫТАНИЙ С ЧИСЛОМ ВОЗМОЖНЫХ ИСХОДОВ, БОЛЬШИМ ДВУХ

Усложним схему Бернулли, предположив, что возможным результатом каждого из независимых испытаний является не одно из пары событий A и $\overline{A}$, а любое из несовместных событий $A_1, A_2, \dots, A_k$ ($k \geq 2$), которые составляют полную группу, то есть $\sum_{j=1}^k A_j = U$. Будем считать, что вероятности $p_1, p_2, \dots, p_k$ появления этих событий не изменяются от испытания к испытанию. Так как события A_j ($j = \overline{1, k}$) несовместные, то $\sum_{j=1}^k p_j = 1$.

Определим вероятность $P_{m_1, m_2, \dots, m_k}^{(n)}$ того, что события $A_1, \dots, A_k$ произойдут соответственно $m_1, m_2, \dots, m_k$ раз, где,

разумеется, $\sum_{j=1}^k m_j = n$. При $k = 2$ вероятность $P_{m_1, m_2}^{(n)}$ уже нам известна и определяется формулой (8.2), в которой p надо заменить на p_1 , q — на p_2 , m — на m_1 и $(n - m)$ — на m_2 . Следовательно, при $k = 2$

$$P_{m_1, m_2}^{(n)} = \frac{n!}{m_1! m_2!} p_1^{m_1} p_2^{m_2}. \quad (9.1)$$

Покажем, что структура формулы сохранится при любом k , то есть, что при $k \geq 2$

$$\begin{cases} P_{m_1, \dots, m_k}^{(n)} = \frac{n!}{m_1! m_2! \dots m_k!} p_1^{m_1} \dots p_k^{m_k}, \\ \sum_{j=1}^k m_j = n, \quad \sum_{j=1}^k p_j = 1. \end{cases} \quad (9.2)$$

Доказательство проведем по индукции.

Для этого достаточно положить, что событие A_k является суммой двух несовместных событий A'_k и A'_{k+1} , имеющих вероятности появления в одном испытании p'_k и p'_{k+1} соответственно, каждое из которых появляется m'_k и m'_{k+1} раз в рассматриваемой серии n испытаний (очевидно, что $p'_k + p'_{k+1} = p_k$, $m'_k + m'_{k+1} = m_k$). В этом случае вместо множителя $p_k^{m_k}$, дающего в (9.2) вероятность появления события A_k ровно m_k раз, нужно подставить вероятность вида (9.1)

$$P_{m'_k m'_{k+1}} = \frac{m_k!}{m'_k! m'_{k+1}!} (p'_k)^{m'_k} (p'_{k+1})^{m'_{k+1}}. \quad (9.3)$$

Выполнив указанную подстановку и произведя сокращение на $m_k!$, получим

$$p_{m_1, \dots, m'_{k+1}}^{(n)} = \frac{n!}{m_1! m_2! \dots m'_k! m'_{k+1}!} p_1^{m_1} \dots (p'_k)^{m'_k} (p'_{k+1})^{m'_{k+1}},$$

то есть формулу того же типа, что и (9.2), чем и доказывается справедливость этой формулы при любом $k \geq 2$.

Совокупность вероятностей $P_{m_1, \dots, m_k}^{(n)}$ называется *полиномиальным законом распределения*. Иногда этот закон называют еще *мультиномиальным*. Испытания описанного выше типа, когда число исходов $k > 2$, называют *сложными испытаниями Бернулли*.

Пример 9.1. Вероятность того, что изделие не имеет брака, равна 0,90, что имеет устранимый брак — 0,09, имеет неустранимый брак — 0,01. Определить вероятность того, что среди взятых наудачу трех изделий хотя бы одно небракованное и хотя бы одно с устранимым браком.

Решение. Пусть событие A_1 означает, что взятое изделие небракованное, A_2 — с устранимым браком, а A_3 — с неустранимым браком. Вероятности этих событий $\mathbf{P}(A_1) = p_1 = 0,90$; $\mathbf{P}(A_2) = p_2 = 0,09$; $\mathbf{P}(A_3) = p_3 = 0,01$. Взяты три изделия, поэтому число испытаний $n = 3$. Искомая вероятность равна вероятности того, что среди взятых наугад изделий хотя бы одно небракованное ($m_1 \geq 1$) и хотя бы одно с устранимым браком ($m_2 \geq 1$), поэтому искомая вероятность

$$\begin{aligned} p &= P_{1,1,1}^{(3)} + P_{2,1,0}^{(3)} + P_{1,2,0}^{(3)} = \\ &= \frac{3!}{1! 1! 1!} \cdot 0,9 \cdot 0,9 \cdot 0,01 + \frac{3!}{1! 2! 0!} \cdot 0,9^2 \cdot 0,09 + \\ &\quad + \frac{3!}{1! 2! 0!} \cdot 0,9 \cdot 0,09^2 = 0,245. \end{aligned}$$

Пример 9.2. В квадрате, разделенном на четыре равновеликие части, случайным образом выбирается шестнадцать точек. Какова вероятность, что в каждой четверти квадрата окажется по четыре точки?

Решение. Вероятность попадания случайной точки в любую четверть постоянна и равна $1/4$. Следовательно, по формуле (9.2) надо положить $p_1 = p_2 = p_3 = p_4 = 1/4$, $m_1 = m_2 = m_3 = m_4 = 4$, $n = 16$. Искомая вероятность

$$p = P_{4,4,4,4}^{(16)} = \frac{16!}{4! 4! 4! 4!} \cdot \left(\frac{1}{4}\right)^{16}.$$

Вычисление по этой формуле можно упростить, если воспользоваться формулой Стирлинга, дающей приближенное значение факториала (см. пример 8.5). В данном случае запишем ее в уточненном виде [9]

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \exp \left\{ \frac{1}{12n} - \frac{1}{360n^3} + o\left(\frac{1}{n^5}\right) \right\}. \quad (9.4)$$

Если в показателе экспоненты сохранить только первое слагаемое, то уже при $n \geq 4$ формула обеспечивает две верные значащие цифры. Используя эту формулу, для искомой вероятности получаем

$$p \approx \frac{\sqrt{2\pi \cdot 16} \cdot 16^{16} e^{-16}}{(\sqrt{2\pi \cdot 4})^4 \cdot 4^{16} e^{-16} \cdot 4^{16}} = \frac{\sqrt{2\pi \cdot 16}}{4\pi^2 \cdot 16} \approx 0,014.$$

Если же в формуле Стирлинга (9.4) в показателе экспоненты учесть следующее слагаемое, то получим $p = 0,013$, причем теперь все выписанные десятичные знаки будут точными.

Пример 9.3. Доказать, что сумма всех вероятностей (9.2) полиномиального закона распределения вероятностей

$$S = \sum_{m_1 + \dots + m_k = n} \frac{n!}{m_1! m_2! \dots m_k!} p_1^{m_1} \dots p_k^{m_k}$$

удовлетворяет условию нормировки $S \equiv 1$.

Решение. Преобразуем выражение (9.2) для полиномиальных вероятностей тождественно к виду

$$\begin{aligned} P_{m_1, \dots, m_k}^{(n)} &= \frac{n!}{m_1! (n - m_1)!} \cdot \frac{(n - m_1)!}{m_2! (n - m_1 - m_2)!} \dots \\ &\dots \frac{(n - m_1 - \dots - m_{k-2})!}{m_{k-1}! (n - m_1 - \dots - m_{k-1})!} \times \\ &\times \frac{(n - m_1 - \dots - m_{k-1})!}{m_k!} \cdot p_1^{m_1} \dots p_k^{(m_k)}. \end{aligned}$$

Для дальнейшего удобно ввести суммы

$$\nu_j = \sum_{i=1}^j m_i, \quad (j = \overline{0, k}).$$

Ясно, что $\nu_0 = 0$ и $\nu_k = n$.

Из первого дополнительного условия (9.2), налагаемого на индексы m_i , следует, что

$$m_k = n - \sum_{i=0}^{k-1} m_i,$$

поэтому последний факториальный сомножитель в выражении для $P_{m_1 \dots m_k}^{(n)}$ обращается в единицу

$$\frac{(n - m_1 - \dots - m_{k-1})!}{m_k!} = \frac{m_k!}{m_k!} \equiv 1,$$

а сама эта вероятность представима в виде

$$P_{m_1 \dots m_k}^{(n)} = C_{n-\nu_0}^{m_1} C_{n-\nu_1}^{m_2} \dots C_{n-\nu_{k-2}}^{m_{k-1}} p_1^{m_1} \dots p_{k-1}^{m_{k-1}} p_k^{n-\nu_{k-2}-m_{k-1}}.$$

Поскольку на индексы m_i наложено одно линейное ограничение в форме равенства их суммы n , то сумма S фактически является $(k-1)$ -кратной. Ее можно представить в виде следующей $(k-1)$ -кратной повторной суммы

$$S_i = \sum_{m_i=0}^{n-\nu_{i-1}} \dots \sum_{m_{k-1}=0}^{n-\nu_{k-2}} C_{n-\nu_{i-1}}^{m_i} \dots C_{n-\nu_{k-2}}^{m_{k-1}} p_i^{m_i} \dots p_k^{n-\nu_{k-2}-m_{k-1}},$$

$$(i = \overline{1, (k-1)}).$$

Ясно, что $S = S_1$. Докажем по индукции, что

$$S_i = (p_i + \dots + p_k)^{n-\nu_{i-1}}, \quad (i = \overline{1, k-1}).$$

При $i = k-1$ имеем по формуле бинома

$$S_{k-1} = \sum_{m_{k-1}=0}^{n-\nu_{k-2}} C_{n-\nu_{k-2}}^{m_{k-1}} p_{k-1}^{m_{k-1}} p_k^{n-\nu_{k-2}-m_{k-1}} =$$

$$= (p_{k-1} + p_k)^{n-\nu_{k-2}},$$

что соответствует индуктивному предположению при $i = k-1$. Допустим, что это предположение справедливо при произвольном i , тогда

$$\begin{aligned}
S_{i-1} &= \sum_{m_{i-1}=0}^{n-\nu_{i-2}} C_{n-\nu_{i-2}}^{m_{i-1}} p_{i-1}^{m_{i-1}} S_i = \\
&= \sum_{m_{i-1}=0}^{n-\nu_{i-2}} C_{n-\nu_{i-2}}^{m_{i-1}} p_{i-1}^{m_{i-1}} (p_i + \dots + p_k)^{n-\nu_{i-1}} = \\
&= \sum_{m_{i-1}=0}^{n-\nu_{i-2}} C_{n-\nu_{i-2}}^{m_{i-1}} p_{i-1}^{m_{i-1}} (p_i + \dots + p_k)^{n-\nu_{i-2}-m_{i-1}} = \\
&= (p_{i-1} + \dots + p_k)^{n-\nu_{i-1}},
\end{aligned}$$

то есть индуктивное предположение верно при уменьшении индекса на единицу, а значит, и при всех $1 \leq i \leq k-1$.

При $i = 1$ находим отсюда окончательно

$$S = S_1 = (p_1 + \dots + p_k)^{n-\nu_0} = (p_1 + \dots + p_k)^n \equiv 1,$$

что следует из второго ограничения (9.2), накладываемого на вероятности p_j . Таким образом, условие нормировки выполняется при любых вероятностях исходов тождественно.

Пример 9.4 (Продолжение примера 8.6). Жюри состоит из $2k + 1$ специалистов. Все они выносят решения независимо друг от друга. Компетентность членов жюри примерно одинакова. Каждый из них выносит правильное решение с вероятностью p , противоположное решение — с вероятностью q , а с вероятностью $r = 1 - p - q$ воздерживается от голосования по данному вопросу. Решение принимается большинством голосов. Если голоса членов жюри, принявших участие в голосовании, разделились поровну, то решение не принимается, и назначаются новые слушания. Найти вероятность принятия решения при такой технологии голосования.

Решение. В данном случае имеем сложные испытания Бернулли с тремя исходами в каждом испытании и $(2k + 1)$ испытанием: A_1 — член жюри проголосовал «за», A_2 — проголосовал «против», A_3 — воздержался. По условию $p_1 = p$, $p_2 = q$, $p_3 = 1 - p - q$.

Вероятность того, что m_1 экспертов поддержат правильное решение, m_2 его отвергнут, а оставшиеся $2k + 1 - m_1 - m_2$ воздержатся, дается равенством

$$\begin{aligned} P_{m_1, m_2, 2k+1-m_1-m_2}^{2k+1} &= \\ &= \frac{(2k+1)!}{m_1! m_2! (2k+1-m_1-m_2)!} \times \\ &\quad \times p^{m_1} q^{m_2} (1-p-q)^{2k+1-m_1-m_2}. \end{aligned}$$

Обозначая искомую вероятность точно так же, как в примерах 5.5 и 8.6, через q_{2k+1} , имеем

$$\begin{aligned} q_{2k+1} &= \sum_{i=0}^{2k+1} \sum_{j=0}^{\min(2k+1-i, i-1)} P_{i, j, 2k+1-i-j}^{(2k+1)} = \\ &= \sum_{i=0}^{2k+1} \sum_{j=0}^{\min(2k+1-i, i-1)} \frac{(2k+1)!}{i! j! (2k+1-i-j)!} \times \\ &\quad \times p^i q^j (1-p-q)^{2k+1-i-j}. \end{aligned}$$

Во внутренней сумме j должно быть меньше, чем i , так как суммируются лишь вероятности, отвечающие принятию верного решения. При этом одновременно $j \leq 2k + 1 - i$, поскольку общее число членов жюри равняется $2k + 1$. Разность $(2k + 1 - i)$ минус j дает число членов жюри, воздержавшихся при голосовании.

Рассмотрим предельный случай $r \rightarrow 0$, когда голосующие не воздерживаются. В этом случае в сумме остаются лишь слагаемые, характеризуемые $2k + 1 - i - j = 0$, то есть $j = 2k + 1 - i$, а тогда внутренняя сумма по j пропадает. При этом по-прежнему должно выполняться условие $j \leq i - 1$, которое сводится к неравенству

$$2k + 1 - i \leq i - 1,$$

что эквивалентно ограничению на пределы суммирования во внешней сумме

$$i \geq k + 1.$$

В результате получим

$$\begin{aligned}
 q_{2k+1} &= \sum_{i=k+1}^{2k+1} \frac{(2k+1)!}{i!(2k+1-i)!} p^i (1-p)^{2k+1-i} = \\
 &= \sum_{i=k+1}^{2k+1} C_{2k+1}^i p^i (1-p)^{2k+1-i}.
 \end{aligned}$$

Этот результат в точности совпадает с ранее полученным в задаче 8.6.

Пример 9.5. Имеется партия из n изделий. Каждое изделие с вероятностью α является качественным, а с вероятностью $(1-\alpha)$ — бракованным. Проводится выборочный контроль качества. Изделие, поступившее на стол контроля, с вероятностью β может быть проверено, а с вероятностью $(1-\beta)$ выпускается без проверки. Какова вероятность того, что число проверенных годных изделий будет не меньше, чем число непроверенных бракованных изделий?

Решение. Имеем сложные испытания Бернулли с четырьмя исходами: A_1 — проверено качественное изделие, A_2 — качественное изделие прошло стол контроля без проверки, A_3 — выявлен случай брака, A_4 — бракованное изделие выпущено без проверки. Поскольку наличие брака и выбор изделия на контроль считаются независимыми событиями, то все эти исходы независимы. Вероятности исходов

$$p_1 = \alpha\beta, \quad p_2 = \alpha(1-\beta), \quad p_3 = (1-\alpha)\beta, \quad p_4 = (1-\alpha)(1-\beta).$$

Вероятность того, что будет проверено m_1 качественных изделий, m_2 таких изделий пройдут без проверки и будет выявлено в ходе проверок m_3 бракованных изделий

$$\begin{aligned}
 P_{m_1, m_2, (n-m_1-m_2-m_3)}^{(n)} &= \frac{n!}{m_1! m_2! m_3! (n-m_1-m_2-m_3)!} \times \\
 &\times (\alpha\beta)^{m_1} (\alpha(1-\beta))^{m_2} ((1-\alpha)\beta)^{m_3} ((1-\alpha)(1-\beta))^{n-m_1-m_2-m_3}.
 \end{aligned}$$

Для получения ответа следует просуммировать все такие вероятности для всевозможных значений m_i , для которых

$m_1 \geq m_4 = n - m_1 - m_2 - m_3$, то есть для значений m_i , удовлетворяющих неравенству

$$m_1 \geq n - m_1 - m_2 - m_3.$$

Учитывая, что сумма всех m_i не должна превосходить n , получаем искомую вероятность в виде

$$q = \sum_{m_1=0}^n \sum_{m_2=0}^{n-m_1} \sum_{m_3=n-2m_1-m_2}^{n-m_1-m_2} \frac{n!}{m_1! m_2! m_3! (n - m_1 - m_2 - m_3)!} \times \\ \times \alpha^{m_1+m_2} \beta^{m_1+m_3} (1 - \alpha)^{n-m_1-m_3} (1 - \beta)^{n-m_1-m_2}.$$

§ 10. ПРИМЕНЕНИЕ ПРОИЗВОДЯЩИХ ФУНКЦИЙ И РЕКУРРЕНТНЫХ УРАВНЕНИЙ ДЛЯ ВЫЧИСЛЕНИЯ ВЕРОЯТНОСТЕЙ СЛОЖНЫХ СОБЫТИЙ

Рассматриваемые выше методы вычисления вероятностей сложных событий основаны на непосредственном применении основных теорем теории вероятностей. Однако, во многих задачах получение окончательного результата таким способом затруднительно, а в некоторых случаях и практически невыполнимо, так как требует сложных комбинаторных расчетов, суммирования вероятностей различных исходов испытания и других трудоемких и утомительных вычислений. Например, простое суммирование всех полиномиальных вероятностей (9.2) представляет достаточно громоздкую и трудоемкую задачу (см. пример 9.3).

Для того чтобы избежать подобных трудоемких вычислений или по крайней мере существенно их сократить, разработаны специальные искусственные приемы, наиболее эффективными из которых являются *метод производящих функций* и *метод рекуррентных соотношений*. Оба этих метода являются универсальными и находят применение во многих вероятностных задачах, однако здесь они будут рассмотрены на примере вычисления вероятностей сложных событий в схеме повторных независимых испытаний.

Начнем с метода производящих функций. Допустим, что имеется некоторая конечная ($n < \infty$) или бесконечная

($n = \infty$) система случайных событий $\{E_m\}_{m=0}^n$, которые для удобства занумерованы, начиная с $m = 0$. Вероятности событий обозначим через $p_m = P(E_m)$, ($m = \overline{0, n}$). Предположим, что система $\{E_m\}$ — полная, тогда должно выполняться условие нормировки

$$\sum_{m=0}^n p_m = 1, \quad (10.1)$$

причем в случае $n = \infty$ слева стоит бесконечный ряд.

Назовем производящей функцией вероятностей p_m функцию $G(u)$, определяемую в виде суммы степеней u^m , коэффициентами при которых являются вероятности p_m :

$$G(u) = \sum_{m=0}^n p_m u^m. \quad (10.2)$$

Аргумент u , вообще говоря, является комплексным, однако, во многих задачах можно ограничиться вещественными значениями u . Сама по себе переменная u не имеет какого-либо глубокого физического смысла. Она является всего лишь удобным средством, чтобы перейти от набора чисел $p_0, p_1, \dots, p_n$ к функции $G(u)$, через которую, как будет показано далее, легко выражаются все эти числа.

Если $n < \infty$, то производящая функция является полиномом степени n , следовательно, она определена при произвольных значениях своего аргумента. При $n = \infty$ имеем бесконечный степенной ряд, являющийся рядом Тейлора³. Независимо от конкретного вида вероятностей p_m ряд всегда сходится при $|u| \leq 1$, поскольку при таких значениях u

$$|G(u)| \leq \sum_{m=0}^{\infty} p_m |u|^m \leq \sum_{m=0}^{\infty} p_m = 1. \quad (10.3)$$

Таким образом, при любом n , как при конечном, так и при бесконечном, производящая функция (10.2) всегда определена относительно комплексного аргумента u в круге единичного радиуса $|u| \leq 1$.

³ Тейлор Брук (1685–1731) — английский математик и философ.

Если функция G известна, то вероятности p_m проще всего могут быть определены путем дифференцирования

$$p_m = \frac{1}{m!} \left. \frac{d^m G(u)}{du^m} \right|_{u=0}, \quad (m = \overline{0, n}). \quad (10.4)$$

Действительно, дифференцируя при $n < \infty$ сумму (10.2) и подставляя $u = 0$, имеем

$$\frac{d^k}{du^k} [p_m u^m] |_{u=0} = \begin{cases} m! p_m, & k = m, \\ 0, & k \neq m, \end{cases} \quad (10.5)$$

поскольку при $k < m$ получаем слева тождественный нуль, а при $k > m$ — некоторую положительную степень u , которая обращается в нуль при подстановке $u = 0$.

В случае $n = \infty$ все эти рассуждения сохраняют силу, так как ряд (10.2) внутри единичного круга $|u| \leq 1$ сходится равномерно относительно u , и его можно дифференцировать почленно.

Выделить коэффициент p_m в разложении (10.2) можно и иным способом. Разделим обе части этого равенства на u^{m+1} и, рассматривая u как комплексную переменную, проинтегрируем полученные выражения по произвольному гладкому замкнутому контуру Γ , который целиком лежит в единичном круге $|u| \leq 1$ и содержит точку $u = 0$ внутри себя. В качестве такого контура можно, например, взять окружность $|u| = \rho$, центр которой совпадает с началом координат, а радиус $\rho < 1$.

Из теории функций комплексного переменного хорошо известно, что при произвольном целом j , справедливо равенство [24]

$$\frac{1}{2\pi i} \int_{|z|=\rho} \frac{dz}{z^j} = \begin{cases} 1, & j = 1, \\ 0, & j \neq 1, \end{cases} \quad (10.6)$$

где i обозначает мнимую единицу. Интегрируя отношение $\frac{G(u)}{u^{m+1}}$ по окружности $|u| = \rho$ и учитывая (10.6), обнаруживаем, что из всей суммы (10.2) остается лишь одно слагаемое $2\pi i p_m$ и, следовательно, для определения вероятности p_m

можно воспользоваться формулой

$$p_m = \frac{1}{2\pi i} \int_{|u|=\rho} \frac{G(u)}{u^{m+1}} du, \quad (m = \overline{0, n}), \quad (10.7)$$

эквивалентной (10.4).

Выражения (10.4) и (10.7) являются универсальными и имеют силу для любых вероятностей p_m и любой полной системы событий E_m . Применим их к изучению повторных независимых испытаний в схеме Бернулли. Пусть $E_m = B_m$, где B_m определяется (8.1) и означает, что в серии из n опытов m из них закончились успешно. Как мы уже установили, при этом $p_m = P_m^{(n)}$, где $P_m^{(n)}$ определяется биномиальным распределением (8.2). Согласно (10.2) производящая функция биномиальных вероятностей по формуле бинома выразится в виде

$$\begin{aligned} G(u) &= \sum_{m=0}^n P_m^{(n)} u^m = \sum_{m=0}^n C_n^m p^m q^{n-m} u^m = \\ &= \sum_{m=0}^n C_n^m (pu)^m q^{n-m} = (pu + q)^n. \end{aligned} \quad (10.8)$$

Подобным же образом подсчитывается производящая функция закона Пуассона (8.11). Допустим, что число испытаний $n \rightarrow \infty$, но при этом $p = \frac{a}{n} \rightarrow 0$. Тогда имеем $n = \infty$, $p_m = P_m$, где P_m определяется (8.11), откуда

$$\begin{aligned} G(u) &= \sum_{m=0}^{\infty} p_m u^m = \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} u^m = \\ &= e^{-a} \sum_{m=0}^{\infty} \frac{(au)^m}{m!} = e^{-a} e^{au} = e^{a(u-1)}. \end{aligned} \quad (10.9)$$

Легко видеть, что обе производящие функции (10.8) и (10.9) удовлетворяют условию

$$G(1) = 1, \quad (10.10)$$

которое является прямым следствием условия нормировки (10.1) и должно выполняться для любой полной системы событий $\{E_m\}_{m=0}^n$.

Помимо свойства (10.10) производящие функции обладают целым рядом других важных и полезных свойств, используя которые часто можно определить $G(u)$, минуя предварительное вычисление вероятностей p_m . Тогда метод производящих функций дает альтернативный способ вычисления указанных вероятностей, основанный на применении общих формул (10.4) и (10.7).

Покажем в качестве примера, как получить выражения (10.8) и (10.9), не прибегая к рассуждениям из параграфа 8 и не вычисляя вначале вероятности $P_m^{(n)}$ и P_m . При этом воспользуемся следующим свойством производящих функций.

Свойство 1. Пусть проводятся две независимых серии испытаний объемом, соответственно, n_1 и n_2 опытов. Обозначим через $n = n_1 + n_2$, общее число опытов, а через G_{n_1} , G_{n_2} , G_n — производящие функции для вероятностей числа успешных опытов в первой, второй и в обеих сериях. Тогда

$$G_n(u) = G_{n_1}(u)G_{n_2}(u). \quad (10.11)$$

Для доказательства этого свойства обозначим, как и раньше, через $P_m^{(n)}$ вероятность получения m «успехов» в серии из n опытов. Для каждого из сомножителей в правой части (10.11)

$$G_{n_i}(u) = \sum_{s=0}^{n_i} P_s^{(n_i)} u^s, \quad (i = \overline{1, 2}). \quad (10.12)$$

Вычислим произведение функций (10.12) при $i = 1$ и $i = 2$, которое обозначим через G_* :

$$G_*(u) = G_{n_1}(u)G_{n_2}(u). \quad (10.13)$$

Коэффициенты разложения G_* по степеням u по аналогии с (10.2) обозначим через p_m^* . Ясно, что G_* представляет собой некоторый полином степени $n = n_1 + n_2$.

Легко видеть, что коэффициент p_m^* , представляет собой сумму произведений $P_{s_1}^{(n_1)}$, $P_{s_2}^{(n_2)}$, у которых $s_1 + s_2 = m$. Эту

сумму можно записать в виде

$$p_m^* = \sum_{s_1+s_2=m} P_{s_1}^{(n_1)} P_{s_2}^{(n_2)} = \sum_{s=0}^m P_s^{(n_1)} P_{m-s}^{(n_2)}. \quad (10.14)$$

Теперь обратимся к производящей функции $G_n(u)$ общей серии испытаний. Введем наряду с событием B_m , определяемым (8.1), аналогичные события $B_m^{(1)}$ и $B_m^{(2)}$ для каждой из коротких серий опытов. Если известно, что всего событие A произошло m раз, то в первой серии оно могло появиться любое число $0 \leq s \leq m$ раз, а затем $m - s$ раз во второй серии

$$B_m = \sum_{s=0}^m B_s^{(1)} B_{m-s}^{(2)}. \quad (10.15)$$

Все слагаемые в сумме (10.15) несовместны, а сомножители в каждом из них независимы, поэтому

$$P_m^{(n)} = \sum_{s=0}^m P_s^{(n_1)} P_{m-s}^{(n_2)}. \quad (10.16)$$

Сравнивая правые части (10.14) и (10.16), заключаем

$$P_m^* = P_m^{(n)}, \quad (m = \overline{0, n}), \quad (10.17)$$

а это означает, что

$$G_*(u) = G_n(u), \quad (10.18)$$

что и доказывает сформированное свойство (10.1).

Применим полученный результат для нахождения производящей функции, соответствующей схеме Бернулли. Если вероятность события A в одном испытании равна p , а вероятность противоположного события $\bar{A}$ есть $q = (1 - p)$, то производящая функция для одного испытания ($n = 1$) в соответствии с общим определением (10.2) будет равняться

$$G_1(u) = pu + q. \quad (10.19)$$

По индукции элементарно доказывается, что для серии из n независимых испытаний

$$G_n(u) = (pu + q)^n. \quad (10.20)$$

Действительно, если это выражение верно при некотором n , то при добавлении к серии еще одного независимого испытания по свойству 1

$$\begin{aligned} G_{n+1}(u) &= G_n(u)G_1(u) = \\ &= (pu + q)^n(pu + q) = (pu + q)^{n+1}, \end{aligned} \quad (10.21)$$

откуда следует справедливость (10.20) при любом $n \geq 1$. Это выражение, разумеется, совпадает с (10.8).

Применяя формулу (10.4), получаем искомые вероятности $P_m^{(n)}$:

$$\begin{aligned} P_m^{(n)} &= \frac{1}{m!} \frac{d^m}{du^m} [(pu + q)^n] |_{u=0} = \\ &= \frac{n(n-1)\dots(n-m+1)}{m!} \cdot (pu + q)^{n-m} p^m |_{u=0} = \\ &= \frac{n!}{m!(n-m)!} p^m q^{n-m}, \end{aligned} \quad (10.22)$$

что совпадает с формулой (8.2), полученной ранее путем непосредственного применения теорем умножения и сложения вероятностей.

Подобным же способом легко получить распределение Пуассона (8.11). Для этого нужно положить в (10.20)

$q = 1 - p$, $p = \frac{a}{n}$, а затем устремить n к бесконечности, считая a фиксированным:

$$G(u) = \lim_{n \rightarrow \infty} \left[1 + \frac{a}{n}(u-1) \right]^n = e^{a(u-1)}. \quad (10.23)$$

Далее по формуле (10.4) находим

$$\begin{aligned} P_m &= \frac{1}{m!} \frac{d^m}{du^m} [e^{a(u-1)}]_{u=0} = \\ &= \frac{1}{m!} e^{-a} e^{au} a^m |_{u=0} = e^{-a} \frac{a^m}{m!}, \end{aligned} \quad (10.24)$$

что совпадает с ранее полученным другим способом выражением (8.11).

Обратимся теперь к неоднородным испытаниям Бернулли, когда в серии независимых испытаний вероятность появления события A зависит от номера испытания k и равна p_k , так что $q_k = 1 - p_k$. Нетрудно показать, что свойство 1 сохраняет силу и для таких испытаний. Последовательно применяя формулу (10.11) n раз, получим

$$G_n(u) = \prod_{k=1}^n (p_k u + q_k). \quad (10.25)$$

Это выражение позволяет определить вероятности $P_m^{(n)}$ более простым способом, чем при непосредственном применении основных теорем теории вероятностей (см. §8). Например, для вероятности однократного появления события она дает:

$$P_1^{(n)} = \frac{d}{du} \left[\prod_{k=1}^n (p_k u + q_k) \right]_{u=0} = \prod_{k=1}^n q_k \sum_{i=1}^n \frac{p_i}{q_i}. \quad (10.26)$$

Метод производящих функций особенно эффективен в случае сложных испытаний Бернулли с числом возможных исходов, большим двух. Предположим, что каждое отдельно взятое испытание завершается появлением одного из k несовместных событий $A_1, A_2, \dots, A_k$, вероятности которых $p_1, p_2, \dots, p_k$ одинаковы для всех опытов и в сумме дают единицу.

Определим по аналогии с (10.2) производящую функцию в виде суммы произведений вероятностей $P_{m_1 \dots m_k}^{(n)}$ на произведение соответствующих степеней $u_1, \dots, u_k$

$$G_n(u_1, \dots, u_k) = \sum_{m_1 + \dots + m_k = n} P_{m_1 \dots m_k}^{(n)} u_1^{m_1} \dots u_k^{m_k}. \quad (10.27)$$

Если производящая функция вероятностей $P_{m_1 \dots m_k}^{(n)}$ известна, то сами эти вероятности легко определяются путем ее дифференцирования по формуле

$$P_{m_1 \dots m_k}^{(n)} = \frac{1}{m_1! \dots m_k!} \frac{\partial^n}{\partial u_1^{m_1} \dots \partial u_k^{m_k}} \times \\ \times G_n(u_1, \dots, u_k) \Big|_{u_1 = \dots = u_k = 0}, \quad (10.28)$$

которая выводится аналогично формуле (10.4), отвечающей случаю $k = 2$. Таким образом, между производящей функцией и связанными с ней вероятностями по-прежнему имеет место взаимно однозначное соответствие.

Как и в случае обычной схемы Бернулли, приведем два способа вычисления G_n . Вначале воспользуемся уже известными нам формулами биномиального распределения (9.2). С их помощью выражение (10.27) преобразуется к виду

$$G_n(u_1, \dots, u_k) = \\ = \sum_{m_1 + \dots + m_k = n} \frac{n!}{m_1! \dots m_k!} (p_1 u_1)^{m_1} \dots (p_k u_k)^{m_k}, \quad (10.29)$$

с той разницей, что все p_j заменены на произведения $p_j u_j$.

Сумма S уже вычислялась в примере 9.3. После достаточно длинных и громоздких выкладок было получено, что

$$S = (p_1 + \dots + p_k)^n. \quad (10.30)$$

Заменяя в этом выражении p_j на $p_j u_j$, находим

$$G_n(u_1, \dots, u_k) = (p_1 u_1 + \dots + p_k u_k)^n. \quad (10.31)$$

То же самое выражение легко получается методом производящих функций без каких-либо сложных вычислений аналогично выводу формулы (10.20). Очевидно, при $n = 1$ имеем

$$G_1(u_1, \dots, u_k) = (p_1 u_1 + \dots + p_k u_k). \quad (10.32)$$

Легко проверить, что свойство 1 сохраняет силу и в случае сложных испытаний Бернулли. Поэтому индукцией по n приходим к формуле (10.32), из которой с учетом (10.20) следует,

что

$$P_{m_1 \dots m_k}^{(n)} = \frac{1}{m_1! \dots m_k!} \cdot n! p_1^{m_1} \dots p_k^{m_k}, \quad (10.33)$$

то есть прежняя формула (9.2) для полиномиального распределения, полученная ранее более сложным способом.

Как уже отмечалось, для полиномиального закона распределения сохраняет силу свойство 1 производящих функций, установленное ранее для биномиального закона. Наряду с этим производящая функция полиномиального закона обладает рядом дополнительных полезных свойств, использование которых позволяет существенно упростить решение прикладных задач. Доказательство их несложно, и его рекомендуется провести читателю самостоятельно в качестве упражнения.

Свойство 2. Предположим, что проводятся сложные испытания Бернулли с k исходами каждого испытания $A_1, \dots, A_k$. Приравняем к единице какие-либо $k - l$ аргументов производящей функции, скажем, $u_{l+1}, \dots, u_k$. Тогда функция

$$G_n^*(u_1, \dots, u_l) = G_n(u_1, \dots, u_k)|_{u_{l+1}=\dots=u_k=1} \quad (10.34)$$

даст производящую функцию для вероятностей числа появления оставшихся исходов $A_1, \dots, A_l$.

Свойство 3. Приравняем друг к другу какие-либо l аргументов производящей функции (10.32), например, положим $u_1 = \dots = u_l = u$, а оставшиеся $k - l$ аргументов приравняем к единице, то есть $u_{l+1} = \dots = u_k = 1$. Полученная функция

$$G_n^*(u) = G_n(u_1, \dots, u_k)|_{u_1=\dots=u_l=u, u_{l+1}=\dots=u_k=1} \quad (10.35)$$

есть производящая функция появления группы исходов $A + \dots + A_1 = \dots = A_l$.

Свойство 4. Допустим, что проводятся сложные испытания Бернулли с k исходами $A_1, \dots, A_k$, причем каждому из исходов A_j присвоен некоторый целый ранг $r_j \geq 0$. Обозначим через R суммарный ранг всех исходов, реализовавшихся в серии из n испытаний. Очевидно, что R является некоторым неотрицательным целым случайным числом. Тогда функция

$$G_n^*(u) = G_n(u^{r_1}, \dots, u^{r_k}) \quad (10.36)$$

является производящей для вероятностей появления различных значений суммарного ранга R в данной серии испытаний.

К последнему свойству необходимо сделать ряд замечаний. Испытания, проводимые по описанной в нем схеме, некоторые авторы называют *ранжированными испытаниями Бернулли*. Подобная схема является достаточно общей и включает в себя в качестве частного случая ряд более простых моделей, разобранных ранее. Например, при $k = 2$, $A_1 = A$, $A_2 = \bar{A}$, $r_1 = 1$, $r_2 = 0$ имеем обычную классическую схему Бернулли, и R обозначает число появлений события A в серии. Если же при произвольном k выбрать некоторое $l < k$, принять $r_1 = \dots = r_l = 1$, $r_{l+1} = \dots = r_k = 0$, то свойство 4 сводится к свойству 3, а R обозначает число появлений какого-либо события из группы $A_1, \dots, A_l$.

Ранжированные испытания довольно часто встречаются в приложениях. В частности, они актуальны в задачах предсказания результатов спортивных состязаний. Пусть, например, два спортсмена проводят серию из n игр, в каждой из которых возможны k исходов состязания. Типичными вариантами исходов игры могут быть ничья, выигрыш или проигрыш первого игрока. При реализации каждого из этих исходов первому игроку начисляется определенное количество очков (баллов). Задача заключается в том, чтобы найти вероятность получения первым игроком заданного числа очков по окончании всего матча.

Отметим, что при формулировке свойства 4 предполагалось, что ранги исходов r_j *неотрицательны*. Иногда полезно отказаться от этого предположения и допустить отрицательные значения r_j . Например, можно поставить задачу определения вероятности того, что *разность* числа появления каких-либо двух исходов равняется заданному числу.

Можно показать, что свойство 4 сохраняет свою силу при любых целочисленных рангах r_j , как положительных, так и отрицательных. При этом, однако, в функции (10.37) могут появиться отрицательные степени u , что противоречит данному вначале определению производящей функции (10.2).

Обобщим это определение. Будем считать, что события E_m и соответствующие им вероятности $p_m = \mathbf{P}(E_m)$ занумерованы

ны как положительными, так и отрицательными индексами $m = \overline{-n_1, n_2}$, причем допускается $n_1 = \infty$ и(или) $n_2 = \infty$.

При таком определении производящая функция будет включать в себя как положительные, так и отрицательные степени аргумента

$$G(u) = \sum_{m=-n_1}^{n_2} p_m u^m. \quad (10.37)$$

Если оба числа n_j конечны, то функция G определена при любых $u \neq 0$. Пусть $n_2 = \infty$, но n_1 конечно. Тогда ряд (10.38) сходится в единичном круге $|u| \leq 1$ за исключением точки $u = 0$. В случае $n_1 = \infty$ ряд (10.38) представляет собой ряд Лорана⁴ [24]. Если при этом все-таки $n_2 < \infty$, то его область сходимости дается неравенством $|u| \geq u_0$, то есть является внешней частью круга радиуса u_0 . Наиболее сложен случай, когда оба n_j бесконечны. Область сходимости такого ряда Лорана имеет вид кругового кольца $u_0 \leq |u| \leq 1$.

Независимо от конкретных значений n_j точка $u = 0$ является изолированной особой точкой функции $G(u)$ и не входит в область ее определения. Поэтому для получения вероятности p_m формулы (10.4) оказываются непригодными, однако, сохраняют свою силу формулы (10.7), которыми можно пользоваться при любых целых m .

Если допустить к рассмотрению производящие функции вида (10.38), то свойство 4 будет выполняться и для отрицательных значений рангов r_j . Рассмотрим один его важный частный случай, который оформим в виде еще одного нового отдельного свойства.

Свойство 5. Положим в производящей функции (10.32) $u_j = u$, $u_l = \frac{1}{u}$, $u_s = 1$ при всех $s \neq j$, $s \neq l$.

Полученная функция

$$G_n^*(u) = G_n(u_1, \dots, u_k)|_{u_i = \varphi_i(u)}, \quad (10.38)$$

⁴ Лоран Пьер Альфонс (1813–1854) — французский математик, по профессии военный инженер.

где

$$\varphi_i(u) = \begin{cases} u, & i = j, \\ u^{-1}, & i = l, \\ 1, & i \neq l, j, \end{cases} \quad (10.39)$$

является производящей для вероятностей различных значений разности чисел появления в серии испытаний исходов A_j и A_l .

В данной задаче становится понятным физический смысл отрицательных индексов при вероятностях p_m в (10.38). Дело в том, что разность числа появления в эксперименте каких-либо двух исходов может быть как положительной, так и отрицательной, а число m указывает на ее возможное значение.

Свойство 5 описывает особый случай ранжированных испытаний Бернулли, когда $r_j = 1$, $r_l = -1$, а ранги всех остальных исходов испытания приравняются к нулю. При этом суммированный ранг R всей серии дает разность числа появления A_j и A_l . Легко видеть, что $-n \leq R \leq n$, так что в производящей функции (10.38) $n_1 = n_2 = n$. Применение свойства 5 поясним на простом примере.

Пусть, например, два игрока играют $2n$ партий, в каждой из которых первый игрок выигрывает с вероятностью p , проигрывает с вероятностью q , делает ничью с вероятностью $r = 1 - p - q$. Требуется определить вероятность P того, что первый игрок выиграет на две партии больше, чем второй.

Так как здесь возможны три исхода каждого испытания, то производящая функция будет функцией трех аргументов

$$G(u_1, u_2, u_3) = (pu_1 + qu_2 + ru_3)^n.$$

Положим $u_1 = u$, $u_2 = 1/u$, $u_3 = 1$ и обозначим полученную таким образом функцию через $R(u)$. Тогда при разложении функции $R(u)$ в ряд по степеням u коэффициент при u^m даст вероятность того, что первый игрок выиграл на m партий больше, чем второй. Таким образом, для нахождения искомой вероятности достаточно в разложении функции $R(u)$ определить коэффициент при u^2 . Имеем

$$R(u) = \left(pu + \frac{q}{u} + r \right)^n = \sum_{j=0}^n \sum_{l=0}^j C_n^j C_{n-j}^l r^{n-j-l} q^l p^j u^{j-l}.$$

В полученной двойной сумме нужно сохранить только слагаемые, у которых показатель степени при u равен двум, то есть слагаемые, у которых $j - l = 2$. Положив затем в полученном выражении $u = 1$, получим искомую вероятность. В результате имеем

$$\begin{aligned} P &= \sum_{j=2}^n C_n^j C_{n-j}^{j-2} r^{n-2j+2} q^{j-2} p^j = \\ &= \sum_{j=2}^n \frac{n!}{j! ((j-2)! (n-2j+2)!)} r^{n-2j+2} q^{j-2} p^j. \end{aligned}$$

В частном случае, при $r = 0$, $n = 2n_1$,

$$P = \frac{(2n_1)!}{(n+1)! (n-1)!} p^{n_1+1} q^{n_1-1}.$$

Дальнейшие примеры применения метода производящих функций разобраны в конце параграфа, а сейчас перейдем к изложению второго метода.

В методе *рекуррентных уравнений* вначале составляется система уравнений, определяющих некоторую совокупность вероятностей, в число которых входит и вероятность искомого события. Эти уравнения могут быть алгебраическими, дифференциальными или какой-либо более сложной природы, а число уравнений может быть как конечным, так и бесконечным.

При определении искомых вероятностей, пользуясь составленной системой рекуррентных уравнений, в некоторых задачах бывает целесообразно предварительно найти производящую функцию всей совокупности вероятностей, в других случаях — непосредственно решать эти уравнения для получения искомой вероятности.

Рассмотрим типичный пример применения рекуррентных уравнений, ограничившись простейшим случаем серии независимых испытаний. Уравнения в данном случае будут алгебраическими, а при их составлении рассматривается, например, постепенное увеличение числа испытаний в серии на единицу,

и для каждого такого шага устанавливается соотношение между вероятностями различных чисел появления событий при n и $(n + 1)$ -м испытаниях. Сущность этого метода можно пояснить на решении задачи, исторически называемой *задачей о разорении игрока*, хотя подобные же задачи часто возникают в самых различных приложениях. Ниже приведем эту задачу в ее классической формулировке.

Задача о разорении игрока. Два игрока A и B играют серию партий, каждую из которых игрок A может выиграть с вероятностью p , проиграть — с вероятностью q и сделать ничью — с вероятностью $r = 1 - p - q$. Каковы вероятности P_A и P_B , что игра закончится разорением игрока A или разорением игрока B , если выигравший партию получает от проигравшего один рубль, причем в начале игры игрок A имел a рублей, а игрок B — b рублей?

Так как $P_A + P_B = 1$ (считается, что игра ведется без ограничения числа партий до полного разорения одного из игроков), то достаточно определить P_A . Введем кроме интересующей нас вероятности P_A вероятность $\mathbf{P}(c)$ разорения игрока A в том случае, когда у него в ходе игры осталось c рублей. Очевидно, что $\mathbf{P}(a) = P_A$, причем $0 \leq c \leq a + b$. Если перед очередной партией игрок A имеет капитал в c рублей, то он может разориться, выиграв очередную партию и увеличив этим свой капитал на единицу, проиграв и уменьшив свой капитал также на единицу, и, наконец, сделав в очередной партии ничью (капитал не изменится). Вероятности указанных событий равны соответственно $p\mathbf{P}(c + 1)$, $q\mathbf{P}(c - 1)$, $r\mathbf{P}(c)$. Следовательно, применяя теорему сложения вероятностей для несовместных событий, получим искомую систему уравнений

$$\mathbf{P}(c) = p\mathbf{P}(c + 1) + q\mathbf{P}(c - 1) + r\mathbf{P}(c), \quad 0 < c < (a + b), \quad (10.40)$$

при решении которой надо учесть, что $\mathbf{P}(0) = 1$ (игрок A разорился), а $\mathbf{P}(a + b) = 0$ (разорился игрок B).

Решение рекуррентных уравнений часто упрощается, если им удастся придать вид соотношений между приращениями вероятностей. Применим этот способ и в данном случае.

Заменив $r = 1 - p - q$, после очевидных преобразований вместо (10.40), получим

$$p[\mathbf{P}(c+1) - \mathbf{P}(c)] = q[\mathbf{P}(c) - \mathbf{P}(c-1)], \quad 0 < c < (a+b). \quad (10.41)$$

Написав набор равенств (10.41) для всех значений капитала первого игрока от единицы до c и перемножив левые и правые части этих равенств, после сокращения на одинаковые множители, получим

$$\mathbf{P}(c+1) = \mathbf{P}(c) - \left(\frac{q}{p}\right)^c \cdot [1 - \mathbf{P}(1)]. \quad (10.42)$$

Снова выписав все полученные равенства (10.42) для значений параметра от единицы до $c-1$ и произведя на этот раз суммирование, после приведения подобных членов, получим

$$\mathbf{P}(c) = \mathbf{P}(1) - \frac{\left[\left(\frac{q}{p}\right)^c - \frac{q}{p}\right] [1 - \mathbf{P}(1)]}{\frac{q}{p} - 1}, \quad (10.43)$$

$$0 < c \leq (a+b).$$

Для определения неизвестной вероятности $\mathbf{P}(1)$, подставив в последнее равенство $c = a + b$, получим

$$1 - \mathbf{P}(1) = \frac{\frac{q}{p} - 1}{\left(\frac{q}{p}\right)^{a+b} - 1}, \quad (10.44)$$

что после подстановки в (10.42) для $c = a$ дает искомую вероятность

$$P_A = \mathbf{P}(a) = \frac{q^a(q^b - p^b)}{q^{a+b} - p^{a+b}}. \quad (10.45)$$

Вероятность разорения второго игрока

$$P_B = 1 - P_A = \frac{p^b(q^a - p^a)}{q^{a+b} - p^{a+b}}. \quad (10.46)$$

В частном случае, когда $p = q = 1/2$, то есть игроки равны по силе, для P_A и P_B получим неопределенные выражения, раскрывая которые, будем иметь

$$P_A = \frac{b}{a+b}, \quad P_B = \frac{a}{a+b}. \quad (10.47)$$

Следовательно, для равных по силе игроков вероятность разорения пропорциональна капиталу партнера.

Рассмотрим примеры на применение метода производящих функций и метода рекуррентных уравнений.

Пример 10.1. Для обеспечения надежной работы прибора поставлены параллельно три элемента, причем прибор может работать только в том случае, когда исправен хотя бы один из этих элементов. Вероятности безотказной работы элементов за некоторое фиксированное время T соответственно равны 0,9, 0,8 и 0,7. Определить вероятность безотказной работы прибора за время T и вероятности отказов одного, двух и трех элементов, если отказ любого элемента не изменяет вероятности безотказной работы двух других.

Решение. Для определения искомых вероятностей составим производящую функцию (10.25) при $n = 3$, $p_1 = 0,9$, $q_1 = 0,1$, $p_2 = 0,8$, $q_2 = 0,2$, $p_3 = 0,7$, $q_3 = 0,3$. В этом случае

$$\begin{aligned} G(u) &= (0,1 + 0,9u)(0,2 + 0,8u)(0,3 + 0,7u) = \\ &= 0,006 + 0,092u + 0,398u^2 + 0,504u^3. \end{aligned}$$

Так как

$$G(u) = \sum_{k=0}^3 P_k^{(3)} u^k,$$

где $P_k^{(3)}$ — вероятность того, что из трех элементов безотказно за время работы T работают k элементов ($k = 0, 1, 2, 3$), то искомые вероятности будут равняться

$$P_0^{(3)} = 0,006, \quad P_1^{(3)} = 0,092, \quad P_2^{(3)} = 0,398, \quad P_3^{(3)} = 0,504.$$

Вероятность безотказной работы прибора в течение времени T равна вероятности того, что будет исправен хотя бы один

из трех элементов. Поэтому указанная вероятность выражается в виде

$$R_1^{(3)} = 1 - P_0^{(3)} = 1 - 0,006 = 0,994.$$

Пример 10.2. Датчик случайных чисел с равной вероятностью выдает любое целое число от нуля до k . Определить вероятность того, что а) сумма n случайных чисел будет равна заданному числу l , где $0 \leq l \leq nk$; б) сумма n случайных чисел не меньше заданного числа s , ($s \leq nk$); в) две суммы по n случайных чисел одинаковые.

Решение. а) Будем считать, что испытание состоит в выборе наудачу одного числа. Тогда выбор n чисел эквивалентен проведению n испытаний. При каждом испытании имеется $k+1$ равновозможных исходов $A_0, \dots, A_k$, причем A_j соответствует получению с датчика числа j , ($j = \overline{0, k}$).

Вероятность любого события $\mathbf{P}(A_j) = p_j = \frac{1}{k+1}$, не зависит от j , поэтому для производящей функции (10.30) имеем

$$G(u_0, \dots, u_k) = \frac{1}{(k+1)^n} (u_0 + \dots + u_k)^n.$$

По свойству 4 для получения производящей функции различных значений суммы чисел следует положить $u_j = u^j$, ($j = \overline{0, k}$). Обозначим производящую функцию для вероятностей появления различных значений этой суммы через $S(u)$. Тогда в разложении по степеням u функции

$$\begin{aligned} S(u) &= G(1, u, u^2, \dots, u^k) = \\ &= \frac{1}{(k+1)^n} \cdot (1 + u + u^2 + \dots + u^k) = \\ &= \frac{1}{(k+1)^n} \cdot \left(\frac{1 - u^{k+1}}{1 - u} \right)^n \end{aligned}$$

коэффициент при u^l будет равен вероятности того, что сумма n случайных чисел равна заданному числу l .

Используя формулу бинома, имеем

$$(1 - u^{k+1})^n = 1 - C_n^1 u^{k+1} + C_n^2 u^{2(k+1)} - \\ - C_n^3 u^{3(k+1)} + \dots + (-1)^n u^{n(k+1)},$$

а с помощью биномиального ряда получаем

$$(1 - u)^{-n} = 1 + C_n^{n-1} u + C_{n+1}^{n-1} u^2 + C_{n+2}^{n-1} u^3 + \dots$$

Искомая вероятность находится путем деления этих двух разложений друг на друга и выделения коэффициента при u^l :

$$P_l = \frac{1}{(k+1)^n} \cdot \left[C_{l+n-1}^{n-1} - C_n^1 C_{l+n-1-(k+1)} + \right. \\ \left. + C_n^2 C_{l+n-2(k+1)}^{n-1} - \dots + (-1)^n C_n^r C_{l+n-1-(k+1)}^{n-1} \right],$$

где r — наибольшее целое число, при котором

$$l + n - 1 - r(k+1) \geq n - 1,$$

то есть r — целая часть числа $l/(k+1)$. Если, например, $k = 9$, $l = 21$, а $n = 6$, то

$$P_{21} = 10^{-6} (C_{26}^5 - C_6^1 C_{16}^5 + C_6^2 C_6^5) = 0,03966.$$

б) Искомая вероятность R_s может быть определена с помощью равенства

$$R_s = 1 - \sum_{l=0}^{s-1} P_l.$$

Справедливо следующее тождество:

$$\sum_{l=\alpha(k+1)}^{s-1} C_{l+n-1-\alpha(k+1)}^{n-1} = \\ = \sum_{j=0}^{s-1-\alpha(k+1)} C_{j+n-1}^{n-1} = C_{s+n-1-\alpha(k+1)}^n,$$

поэтому искомая вероятность выражается в виде

$$R_s = 1 - \frac{1}{(k+1)^n} \cdot \left[C_{s+n-1}^n - C_n^1 C_{s+n-1-(k+1)}^n + \right. \\ \left. + C_n^2 C_{s+n-1-2(k+1)}^n - \dots + (-1)^n C_n^r C_{s+n-1-r(k+1)}^n \right],$$

где r — целая часть числа $\frac{s-1}{k+1}$.

Если, например, $k = 9$, $n = 3$, $s = 25$, то

$$R_{25} = 1 - 10^{-3}(C_{27}^3 - C_3^1 C_{17}^3 + C_3^2 C_7^3) = 0,01.$$

в) Имеем аналогично пункту а):

$$S(u) = \frac{1}{(k+1)^n} \left(\frac{1-u^{k+1}}{1-u} \right)^n = \sum_{l=0}^{nk} P_l u^l,$$

где P_l — вероятность того, что сумма из n чисел равна l .

Для другой серии из n случайных чисел в функции

$$S_1(v) = \frac{1}{(k+1)^n} \left(\frac{1-v^{k+1}}{1-v} \right)^n = \sum_{j=0}^{nk} P_j v^j$$

коэффициент при v^j равен вероятности того, что сумма из n чисел равна j . Примем $v = 1/u$, тогда при разложении функции $S(u)S_1(1/u)$ в ряд по степеням u коэффициент при u^0 даст искомую вероятность $P = \sum_{j=0}^{nk} P_j^2$ равенства двух сумм из n случайных чисел. Имеем

$$S(u)S_1\left(\frac{1}{u}\right) = \frac{1}{(k+1)^{2n}} \cdot \frac{1}{u^{kn}} \cdot \left(\frac{1-u^{k+1}}{1-u} \right)^{2n}.$$

Используя записанные выше выражения, получим

$$P = \frac{1}{(k+1)^{2n}} \left[C_{kn+2n-1}^{2n-1} - \right. \\ \left. - C_{2n}^1 C_{kn+2n-1-(k+1)}^{2n-1} + C_{2n}^{2n} C_{kn+2n-1-2(k+1)}^{2n-1} - \dots \right. \\ \left. \dots + (-1)^r C_{2n}^r C_{kn+2n-1-r(k+1)}^{2n-1} \right],$$

где r — целая часть числа $kn/(k+1)$.

Пример 10.3 (Задача о «счастливых» билетах). Определить вероятность получения трамвайного билета, у которого суммы первых трех и последних трех цифр номера равны, считая, что номер шестизначный и все его цифры равновероятны и независимы.

Решение. Рассмотрим случайный эксперимент, в котором последовательно разыгрываются все шесть цифр номера. Этот эксперимент можно трактовать как две серии из $n_1 = n_2 = 3$ сложных испытаний Бернулли, в каждом из которых возможны десять различных исходов (появление той или иной цифры), причем все эти исходы равновероятны. Данные испытания являются ранжированными: если выпадает цифра j , ($j = \overline{0,9}$), то к общей сумме добавляется ровно j очков.

Особенность задачи заключается в наличии двух серий испытаний, по каждой из которых берется своя сумма очков. Дело сводится к нахождению вероятности того, что обе эти суммы равны.

Сформулированная задача легко решается, если использовать общие свойства производящих функций. Вначале нужно записать производящую функцию для вероятностей получения определенного набора цифр в первой и второй сериях. Далее по свойству 4 находится производящая функция вероятностей суммы первых, а затем — вторых трех цифр номера. Далее, по свойству 1, строится совместная производящая функция для вероятностей обеих указанных сумм. Наконец, свойство 5 позволяет определить производящую функцию для вероятностей разности первой и второй суммы очков. Остается лишь выделить в ней коэффициент при нулевой степени аргумента.

Подробно расписывать все эти вычисления нет необходимости, так как они были в общем виде приведены при рассмотрении пункта в) в примере 10.2. Ясно, что наша новая задача соответствует $k = 9$ (числа записываются в десятичной системе исчисления), $n = 3$ (сумма цифр берется по первым трем и последним трем разрядам). Используя последнюю формулу примера 10.2, находим

$$P = 10^{-6}(C_{32}^5 - C_6^1 C_{22}^5 + C_6^2 C_{12}^5) = 0,055252.$$

Таким образом, доля «счастливых» билетов среди всех составляет более 5%. Приблизительно каждый восемнадцатый билет является «счастливым», и возможность вытащить его не является столь уж редким событием.

Пример 10.4. Игра k лиц состоит в том, что каждый поочередно бросает монету; выигрывает первый, получивший орла. Определить вероятность выигрыша для каждого из игроков.

Решение. Обозначим вероятность выигрыша j -го игрока через P_j . Этот игрок будет бросать монету только в том случае, если все предыдущие $(j-1)$ игроков не выиграли. Вероятность этого события $(1/2)^{j-1}$. Бросив монету, игрок выигрывает или сразу с вероятностью $1/2$, если выпадает орел, или (если выпадает решка с вероятностью $1/2$) после того, как все следующие $(k-j)$ игроков не выиграли (вероятность этого события $(1/2)^{k-j}$), после чего игрок может выиграть с вероятностью P_j .

Применяя теоремы умножения и сложения вероятностей, получаем уравнение для определения P_j :

$$P_j = \frac{1}{2^{j-1}} \left[\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2^{k-j}} P_j \right],$$

откуда находим

$$P_j = \frac{0,5^j}{1 - 0,5^k}.$$

В частном случае $k = 2$, соответствующем игре двух лиц,

$$P_1 = \frac{2}{3}, \quad P_2 = \frac{1}{3}.$$

Пример 10.5. Радиосигнал принят, если приняты подряд m импульсов, испускаемых передатчиком в определенные моменты времени. Определить вероятность того, что сигнал будет принят после прослушивания приемником n последовательных импульсов, если вследствие помех импульс принимается с вероятностью p и не принимается с вероятностью $q = 1 - p$, независимо от того, приняты предыдущие импульсы или нет.

Решение. Обозначим через p_k вероятность того, что имеется одна серия из m импульсов, которая закончится на k -м импульсе. Через q_k обозначим вероятность того, что среди первых k импульсов не будет ни одной серии из m принятых последовательно импульсов. Очевидно, что при $k < m$ значение $p_k = 0$, а $q_k = 1$. Вероятность появления хотя бы одной серии из m принятых импульсов за первые k переданных импульсов равна $1 - q_k$. Такая серия впервые может произойти при s -м испытании с вероятностью p_s ($s = \overline{m, k}$). События, имеющие указанные вероятности, несовместные, поэтому имеет место равенство

$$1 - q_k = \sum_{s=m}^k p_s,$$

из которого следует, что при любом k

$$q_k - q_{k-1} = -p_k, \quad (k = \overline{1, \infty}).$$

Умножая обе части этого выражения на u^k и суммируя по k от единицы до бесконечности, получаем

$$\sum_{k=1}^{\infty} q_k u^k - \sum_{k=1}^{\infty} q_{k-1} u^k = - \sum_{k=1}^{\infty} p_k u^k.$$

Введем производящие функции для вероятностей p_k и q_k

$$G(u) = \sum_{k=0}^{\infty} p_k u^k, \quad R(u) = \sum_{k=0}^{\infty} q_k u^k,$$

тогда полученное выше равенство можно записать в терминах введенных производящих функций в виде

$$R(u) - q_0 - uR(u) = p_0 - G(u).$$

Так как $p_0 + q_0 = 1$, то $G(u) = 1 - (1 - u)R(u)$. При $k \geq m$ серия из m последовательных принятых импульсов может не возникнуть только одним из следующих m способов: сначала имеет место серия из $s - 1$ принятых импульсов, затем

импульс не принят, и наконец, среди оставшихся $k - s$ переданных импульсов не будет серии из m принятых импульсов ($s = \overline{1, m}$).

Вероятности указанных независимых событий соответственно равны p^{s-1} , q_{k-s} и $q = 1 - p$, поэтому справедливы равенства

$$q_k = q \sum_{s=1}^m p^{s-1} q_{k-s}, \quad (k = \overline{m, \infty}).$$

Умножив последнее равенство на u^k и суммируя по k , получим

$$\sum_{k=m}^{\infty} q_k u^k = q \sum_{s=1}^m p^{s-1} u^s \sum_{k=m}^{\infty} q_{k-s} u^{k-s}.$$

Имеем

$$\begin{aligned} \sum_{k=m}^{\infty} q_k u^k &= R(u) - \sum_{k=0}^{m-1} u^k = R(u) - \frac{1 - u^m}{1 - u}, \\ \sum_{k=m}^{\infty} q_{k-s} u^{k-s} &= R(u) - \frac{1 - u^{m-s}}{1 - u}, \end{aligned}$$

поэтому справедливо равенство

$$\begin{aligned} R(u) - \frac{1 - u^m}{1 - u} &= R(u) q u \frac{1 - (pu)^m}{1 - pu} - \\ &\quad - \frac{qu}{1 - u} \sum_{s=1}^m (up)^{s-1} (1 - u^{m-s}), \end{aligned}$$

которое можно записать еще так:

$$\begin{aligned} R(u) \left[1 - qu \frac{1 - (pu)^m}{1 - pu} \right] &= \frac{1}{1 - u} \left[1 - u^m - \right. \\ &\quad \left. - qu \frac{1 - (pu)^m}{1 - pu} + u^m (1 - p^m) \right]. \end{aligned}$$

Из этого выражения находим

$$R(u) = \frac{1 - u - (pu)^m + p^m u^{m+1}}{(1 - u)(1 - u + qp^m u^{m+1})}.$$

Поэтому производящая функция

$$G(u) = \frac{(pu)^m(1 - pu)}{1 - u + qp^m u^{m+1}}.$$

Искомая вероятность

$$P_n = \frac{1}{n!} \left. \frac{d^n G(u)}{du^n} \right|_{u=0}.$$

Аналогично может быть решена задача, если для приема сигнала достаточно, чтобы приемник из любой серии m импульсов принял l импульсов.

Пример 10.6. Проводятся стандартные испытания Бернулли с параметрами n и p . С какой вероятностью событие произойдет в них четное число раз? Решение получить методом рекуррентных соотношений.

Решение. Эта задача ранее была решена в примере 8.4 путем формальных алгебраических вычислений на основе формулы бинома. Теперь решим ее другим способом.

Обозначим через p_k вероятность того, что после k опытов число появления события оказалось четным. Перед проведением k -го опыта удобно ввести две гипотезы H_{k1} и H_{k2} . Первая означает, что до k -го опыта событие произошло четное число раз, а вторая — нечетное число раз. Априорные вероятности гипотез

$$P(H_{k1}) = p_{k-1}, \quad P(H_{k2}) = 1 - p_{k-1}.$$

По формуле полной вероятности

$$p_k = p_{k-1}(1 - p) + (1 - p_{k-1})p,$$

откуда следует

$$p_k = p + p_{k-1}(1 - 2p).$$

Последнее равенство представим в виде

$$p_k - \frac{1}{2} = (1 - 2p) \left(p_{k-1} - \frac{1}{2} \right), \quad (k = \overline{1, n}).$$

По индукции нетрудно показать, что

$$p_k - \frac{1}{2} = (1 - 2p)^k \left(p_0 - \frac{1}{2} \right), \quad (k = \overline{1, n}),$$

поскольку $p_0 = 1$, то получаем $p_n = \frac{1}{2}[1 + (1 - 2p)^n]$, что совпадает с примером 8.4.

§ 11. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Пределные теоремы теории вероятностей устанавливают количественные соответствия между наблюдаемыми в эксперименте результатами испытаний и вероятностными характеристиками изучаемых случайных явлений. Пределными эти теоремы называются потому, что устанавливаемые ими соотношения выполняются с тем большей вероятностью, чем больше число проведенных испытаний (в пределе — с вероятностью единица).

Устанавливая соответствие между вероятностной мерой случайных явлений и результатами опыта, предельные теоремы имеют непосредственное прикладное значение. В частности, если окажется, что результаты опыта противоречат предельным теоремам, то это будет означать, что принятая вероятностная схема в данном случае требует корректировки или вообще должна быть отброшена.

Прикладное значение предельных теорем было ясно уже на начальном этапе развития теории вероятностей таким математикам, как Я. Бернулли и П. Лаплас, которые и доказали две первые предельные теоремы, относящиеся к последовательности n независимых испытаний, в каждом из которых событие A может произойти с фиксированной вероятностью $\mathbf{P}(A) = p$ («схема Бернулли»). Обозначим число появлений события A

в серии через m , а его отношение к числу испытаний n (*частота*, или *частость*, появления события A) через $\bar{p}$:

$$\bar{p} = \frac{m}{n}. \quad (11.1)$$

При $p = 0$ ($A = V$ — невозможное событие) имеем $\bar{p} = 0$, при $p = 1$ ($A = U$ — достоверное событие), получаем $\bar{p} = 1$. Если же $0 < p < 1$, то предсказать до опыта значение $\bar{p}$ невозможно, поскольку оно зависит от случая, однако интуитивно представляется правдоподобным, что при большом n частота $\bar{p}$ появления события A не должна сильно отличаться от вероятности p его появления. Теорема Бернулли объясняет это интуитивное убеждение.

Теорема Бернулли

При неограниченном увеличении числа n независимых испытаний частота $\bar{p}$ появления события A сходится по вероятности к вероятности p появления этого события при одном испытании, то есть

$$\lim_{n \rightarrow \infty} \mathbf{P}\{|\bar{p} - p| < \varepsilon\} = 1, \quad (11.2)$$

где ε — любое сколь угодно малое положительное число.

Таким образом, при достаточно большом n с вероятностью, сколь угодно близкой к единице, можно быть уверенным, что ошибка при замене вероятности p частотой $\bar{p}$ не превзойдет по модулю сколь угодно малого заданного значения ε . Следовательно, частоту $\bar{p}$ можно рассматривать как экспериментальное значение вероятности p .

Частота $\bar{p}$ является средним значением числа появлений события A в серии испытаний, приходящимся на одно проведенное испытание. Следовательно, теорема Бернулли устанавливает связь между средним и вероятностью при увеличении числа испытаний. Такие предельные теоремы в теории вероятности объединяют под общим названием «законов больших чисел».

Математический смысл теоремы Бернулли сводится к следующему. Представим вероятность, стоящую под знаком пре-

дела в (11.2), в виде

$$\begin{aligned} Q_n &= \mathbf{P}\{|\bar{p} - p| < \varepsilon\} = \\ &= \mathbf{P}\{(p - \varepsilon)n < m < (p + \varepsilon)n\} = \\ &= \mathbf{P}\{m_1 \leq m \leq m_2\}, \end{aligned}$$

где

$$m_1 = [(p - \varepsilon)n] + 1, \quad m_2 = [(p + \varepsilon)n],$$

причем квадратные скобки обозначают целую часть числа, а оба граничных значения m_1 и m_2 зависят от n и неограниченно возрастают при $n \rightarrow \infty$.

Вероятность Q_n представляется суммой

$$Q_n = \sum_{m=m_1}^{m_2} P_m^{(n)},$$

в которую входят биномиальные вероятности (8.2). Ранее (см. § 8) было установлено, что наибольшие из этих вероятностей имеют номера m , близкие к np . Все они входят в сумму Q_n , так как из определения m_1 и m_2 следует, что $m_1 \leq np \leq m_2$. Исключенными из суммы Q_n оказываются вероятности $P_m^{(n)}$ с номерами, близкими к нулю, а также близкими к n , которые достаточно быстро стремятся к нулю при $n \rightarrow \infty$.

Приведенные рассуждения можно было бы подтвердить и прямым вычислением, доказав теорему Бернулли аналогично тому, как далее в этом же параграфе будет доказана теорема Лапласа, однако в этом нет необходимости. Дело в том, что теорема Бернулли является частным случаем теоремы Чебышева, которая будет доказана далее в специальной главе, посвященной предельным теоремам. Доказательство теоремы Чебышева несложно, однако требует применения нового математического аппарата (теории случайных величин). Поэтому не будем приводить отдельно доказательство теоремы Бернулли, а перейдем к рассмотрению второй предельной теоремы — теоремы Лапласа.

Теорема Лапласа также относится к схеме Бернулли, однако является уже не разновидностью закона больших чисел,

а принадлежит к другой группе предельных теорем, объединяемых под общим названием «центральных предельных теорем». Они касаются определения вероятностей для сумм большого числа случайных величин.

Теорема Лапласа

Если m — число появлений события A в серии n независимых испытаний, $p = \mathbf{P}(A)$ — вероятность появления события A в одном испытании, $q = 1 - p$, причем $0 < p < 1$, то

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ z_1 \leq \frac{m - np}{\sqrt{npq}} \leq z_2 \right\} = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-t^2/2} dt. \quad (11.3)$$

Для доказательства прежде всего заметим, что

$$z_1 \leq \frac{m - np}{\sqrt{npq}} \leq z_2$$

эквивалентно неравенству

$$np + z_1 \sqrt{npq} \leq m \leq np + z_2 \sqrt{npq}.$$

Обозначая эти границы изменения m через m_1 и m_2 , формуле (11.3) можно придать другой вид, позволяющий найти приближенное значение вероятности попадания числа m в заданный интервал при большом n :

$$\mathbf{P} \{ m_1 \leq m \leq m_2 \} \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{m_1 - np}{\sqrt{npq}}}^{\frac{m_2 - np}{\sqrt{npq}}} e^{-t^2/2} dt. \quad (11.4)$$

Интеграл, стоящий справа в (11.3) и (11.4), не выражается через элементарные функции. Поскольку он часто встречается в вероятностных расчетах, то он табулирован [17]. Для этого вводят специальную функцию, называемую интегральной

функцией Лапласа и определяемую формулой

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt. \quad (11.5)$$

Функцию (11.5) часто называют *интегралом вероятностей*, *функцией ошибок*, используют и другие наименования. Отметим также, что существует несколько табулированных функций, через которые выражается интеграл (11.3), отличающиеся коэффициентом перед интегралом, видом подынтегральной функции и пределами интегрирования. Например, интегральными функциями Лапласа иногда называют функции

$$\Phi_1(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt, \quad \Phi_2(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt,$$

а также другие подобные им функции. Во избежание недоразумений в нашем курсе в дальнейшем под функцией Лапласа всегда будет пониматься функция вида (11.5).

Используя обозначение (11.5), формуле (11.3) можно придать вид

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P} \left\{ z_1 \leq \frac{m - np}{\sqrt{npq}} \leq z_2 \right\} = \\ = \frac{1}{2} \left[\Phi \left(\frac{m_2 - np}{\sqrt{npq}} \right) - \Phi \left(\frac{m_1 - np}{\sqrt{npq}} \right) \right], \end{aligned} \quad (11.6)$$

а приближенное равенство (11.4) тогда переписать так:

$$\begin{aligned} \mathbf{P} \{ m_1 \leq m \leq m_2 \} \approx \\ \approx \frac{1}{2} \left[\Phi \left(\frac{m_2 - np}{\sqrt{npq}} \right) - \Phi \left(\frac{m_1 - np}{\sqrt{npq}} \right) \right]. \end{aligned} \quad (11.7)$$

Перейдем непосредственно к доказательству теоремы. Для этого представим вероятность $\mathbf{P} \{ m_1 \leq m \leq m_2 \}$ как сумму

вероятностей $P_m^{(n)}$ различных чисел появления события A . Учитывая формулу (8.2) для биномиального закона распределения, имеем

$$\mathbf{P}\{m_1 \leq m \leq m_2\} = \sum_{m=m_1}^{m_2} \frac{n!}{m!(n-m)!} p^m q^{n-m}. \quad (11.8)$$

Далее рассчитаем более подробно различные слагаемые, входящие в формулу (11.8).

Как уже отмечалось, наибольшее значение вероятности $P_m^{(n)}$ достигается при m , близком к np . Это число можно считать большим, поскольку при доказательстве теоремы Лапласа принимается, что $n \rightarrow \infty$. При малых значениях m , а также значениях, близких к n , вероятности $P_m^{(n)}$ малы, так как в этом случае коэффициент C_n^m не компенсирует малости произведения $p^m q^{n-m}$. Поэтому в формуле (11.8) можно считать, что m велико, а именно, имеет порядок величины n , и, следовательно, все факториалы в выражении для вероятности $P_m^{(n)}$ можно заменить, пользуясь формулой Стирлинга [17]

$$n! = \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{\frac{\Theta}{12n}}, \quad (0 < \Theta < 1). \quad (11.9)$$

Выполнив эту замену, получим

$$P_m^{(n)} = \frac{n^{n+1/2} \cdot p^m q^{n-m} \cdot e^{\frac{\Theta}{12n}}}{\sqrt{2\pi} m^{m+1/2} (n-m)^{n-m+1/2}}. \quad (11.10)$$

Введя новую переменную Z_m , положив

$$Z_m = \frac{m - np}{\sqrt{npq}}, \quad (11.11)$$

будем иметь

$$\begin{cases} m = np \left(1 + \sqrt{\frac{q}{np}} Z_m\right), \\ n - m = nq \left(1 - \sqrt{\frac{p}{nq}} Z_m\right). \end{cases} \quad (11.12)$$

Подставив эти выражения вместо m и $(n-m)$ в основания степеней $m^{m+1/2}$ и $(n-m)^{n-m+1/2}$ (и сохраняя временно m в показателях степеней), получим

$$\begin{aligned}
 P_m^{(n)} &= \frac{1}{\sqrt{npq}\sqrt{2\pi}} \cdot \left(1 + \sqrt{\frac{q}{np}} Z_m\right)^{-(m+1/2)} \times \\
 &\times \left(1 - \sqrt{\frac{p}{nq}} Z_m\right)^{-(n-m+1/2)} \cdot e^{\frac{\Theta}{12n}} = \frac{1}{\sqrt{npq}\sqrt{2\pi}} \times \\
 &\times \exp \left\{ -\left(m + \frac{1}{2}\right) \cdot \ln \left(1 + \sqrt{\frac{q}{np}} Z_m\right) - \right. \\
 &\left. - \left(n - m + \frac{1}{2}\right) \ln \left(1 - \sqrt{\frac{p}{nq}} Z_m\right) + \frac{\Theta}{12n} \right\}. \quad (11.13)
 \end{aligned}$$

Так как число испытаний n , по определению, велико, то логарифмы, стоящие в показателе степени, можно приближенно заменить отрезками степенных рядов, сохранив в разложении

$$\ln(1+x) = x - \frac{1}{2} \cdot x^2 + \frac{1}{3} \cdot x^3 \dots \quad (|x| < 1)$$

только первые два члена ряда.

Выполнив указанные преобразования и учитывая (11.12), формуле (11.13) можно придать вид

$$P_m^{(n)} = \frac{1}{\sqrt{npq}\sqrt{2\pi}} \cdot \exp \left[-\frac{Z_m^2}{2} + \frac{\Theta}{12n} \right]$$

или, пренебрегая вторым слагаемым, в показателе экспоненты

$$P_m^{(n)} \approx \frac{1}{\sqrt{npq}\sqrt{2\pi}} \cdot \exp \left(-\frac{Z_m^2}{2} \right). \quad (11.14)$$

Формула (11.14), также полученная Лапласом, выражает так называемую *локальную теорему Лапласа*.

Вернемся к формуле (11.8). Подставляя в нее выражение (11.14), получаем

$$\mathbf{P}\{m_1 \leq m \leq m_2\} = \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}} \times \\ \times e^{\left(-\frac{Z_m^2}{2}\right)} \cdot e^{\left(\frac{\Theta}{12n}\right)} \cdot \frac{1}{\sqrt{npq}}. \quad (11.15)$$

Так как при увеличении m на единицу в соответствии с (11.11) переменная Z_m увеличивается на $\Delta = 1/\sqrt{npq}$, то формулу (11.15) можно рассматривать как интегральную сум-

му, написанную для интеграла от функции $\frac{1}{\sqrt{2\pi}} \cdot e^{\left(-\frac{z^2}{2}\right)}$ при ее интегрировании по переменной z в пределах от $z_1 = \frac{m_1 - np}{\sqrt{npq}}$ до $z_2 = \frac{m_2 - np}{\sqrt{npq}}$ с шагом интегрирования Δ .

Следовательно, перейдя к пределу $n \rightarrow \infty$, вместо (11.15) получим

$$\lim_{n \rightarrow \infty} \mathbf{P}\{m_1 \leq m \leq m_2\} = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{\left(-\frac{z^2}{2}\right)} dz = \\ = \frac{1}{2} \cdot \left[\Phi \left(\frac{m_2 - np}{\sqrt{npq}} \right) - \Phi \left(\frac{m_1 - np}{\sqrt{npq}} \right) \right]. \quad (11.16)$$

Это и есть теорема Лапласа, которую в отличие от равенства (11.14) называют *интегральной теоремой Лапласа* или *интегральной теоремой Муавра–Лапласа*, поскольку А. Муавр получил эту теорему для частного случая $p = 1/2$ еще в 1730 году.

Теорема Лапласа имеет значение не только в предельном случае, когда $n \rightarrow \infty$, но и для конечных n , так как при числе испытаний, равном уже нескольким десяткам, непосредствен-

ный подсчет вероятности (11.8) становится весьма утомительным из-за наличия факториалов от больших чисел, а при n порядка 1000 расчеты по этой формуле становятся практически невыполнимыми. Между тем на практике иногда приходится сталкиваться и с более длинными сериями испытаний. В таких случаях теорема Лапласа становится просто незаменимой и при этом дает высокую точность.

Формулами Лапласа в инженерных задачах пользуются и при конечном n , причем эти формулы обычно уже при $n \approx 10 - 20$ дают хорошую точность. Тем не менее представляет интерес более детально выяснить возможности уточнения этих формул и установить пределы их применимости.

Одним из возможных путей уточнения является получение асимптотического разложения для вероятности (11.15) по степеням малого параметра $1/\sqrt{n}$, что можно, например, сделать, учитывая в формуле Стирлинга (11.14) следующие слагаемые в показателе степени экспоненты [17]:

$$n! = \sqrt{2\pi n} \cdot \exp \left\{ \frac{1}{12n} - \frac{1}{360n^3} + \frac{\Theta^*}{n^5} \right\}.$$

Некоторое уточнение можно получить также, если в формуле (11.16) интегрирование производить не между m_1 и m_2 , а между $(m_1 - 1/2)$ и $(m_2 + 1/2)$, то есть пользоваться формулой

$$\mathbf{P}\{m_1 \leq m \leq m_2\} \approx \frac{1}{2} \left[\Phi \left(\frac{m_2 + 0,5 - np}{\sqrt{npq}} \right) - \Phi \left(\frac{m_1 - 0,5 - np}{\sqrt{npq}} \right) \right], \quad (11.17)$$

а вместо формулы (11.14) — формулой

$$P_m^{(n)} = \frac{1}{2} \cdot \left[\Phi \left(\frac{m + 0,5 - np}{\sqrt{npq}} \right) - \Phi \left(\frac{m - 0,5 - np}{\sqrt{npq}} \right) \right]. \quad (11.18)$$

Формулы Лапласа обычно дают приемлемую точность, если интервал $[m_1, m_2]$ изменения m не выходит за пределы интервала $[np - 3\sqrt{npq}, np + 3\sqrt{npq}]$, причем границы последнего

интервала не выходят за пределы интервала $[0, n]$, которому принадлежат значения m . Иными словами, этими формулами с уверенностью можно пользоваться, если выполняются неравенства:

$$0 \leq np - 3\sqrt{npq} \leq m \leq np + 3\sqrt{npq} \leq n, \quad (11.19)$$

которым можно придать и такой вид:

$$np \geq 9q, \quad nq \geq 9p, \quad |m - np| \geq 3\sqrt{npq}. \quad (11.20)$$

Если эти неравенства не выполняются, то применение формул Лапласа при конечном n может дать относительно большую ошибку.

В заключение данного параграфа рассмотрим несколько свойств интегральной функции Лапласа, которая используется при многих вероятностных расчетах.

Во-первых, из определения функции (11.5) следует, что функция Лапласа нечетная. Действительно, перейдя от переменной интегрирования t к переменной $t_1 = -y$, получим

$$\Phi(z) = -\Phi(-z). \quad (11.21)$$

Во-вторых, положив в формуле (11.5) $z = \infty$, обнаруживаем, что интеграл, стоящий справа, равен $\sqrt{\pi/2}$ и, следовательно,

$$\Phi(\infty) = 1. \quad (11.22)$$

Разложив под знаком интеграла экспоненту в ряд и произведя интегрирование ряда почленно, получим

$$\begin{aligned} \Phi(z) &= \frac{2}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt = \frac{2}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j}{j! 2^j} \int_0^z t^{2j} dt = \\ &= \frac{2}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j z^{2j+1}}{(2j+1) j! 2^j}. \end{aligned}$$

Таким образом, функция Лапласа представима рядом Тэйлора

$$\Phi(z) = \frac{2z}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j z^{2j}}{(2j+1) j! 2^j}. \quad (11.23)$$

Полученный ряд сходится при любом значении z , однако при $z \gg 1$ он неудобен для вычислений.

Для больших значений z может быть использовано асимптотическое разложение. Для получения этого разложения представим интеграл (11.5) в виде разности двух интегралов, положив

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-t^2/2} dt - \frac{2}{\sqrt{2\pi}} \int_z^{\infty} e^{-t^2/2} dt. \quad (11.24)$$

Первый интеграл в соответствии с (11.22) равен единице, следовательно, формуле (11.24) можно придать вид

$$1 - \Phi(z) = \frac{2}{\sqrt{2\pi}} \int_z^{\infty} e^{-t^2/2} dt. \quad (11.25)$$

Будем последовательно в (11.25) интегрировать по частям, каждый раз выбирая в качестве дифференциала следующее произведение:

$$e^{-t^2/2} t dt = -d[e^{-t^2/2}].$$

Производя интегрирование n раз и учитывая при этом, что внеинтегральные слагаемые на верхнем пределе обращаются в нуль, получим

$$1 - \Phi(z) = \frac{2}{z\sqrt{2\pi}} e^{-z^2/2} \left[1 - \frac{1}{z^2} + \frac{1 \cdot 3}{z^4} \dots + \right. \\ \left. + (-1)^{n-1} \cdot \frac{(2n-3)!!}{z^{2(n-1)}} \right] + R_n, \quad (11.26)$$

где остаток R_n представляется следующим интегралом:

$$R_n = \frac{2}{\sqrt{2\pi}} \cdot (-1)^n \cdot (2n-1)!! \int_z^{\infty} e^{-t^2/2} \cdot \frac{dt}{t^{2n}}. \quad (11.27)$$

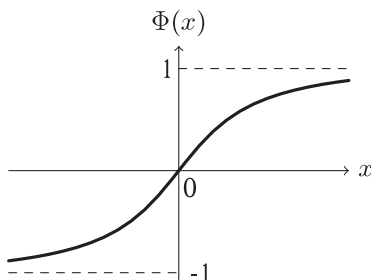


Рис. 11.1
Интеграл вероятностей $\Phi(z)$

Ряд (11.26) является асимптотическим: он расходится при любом z , однако может обеспечить достаточно высокую точность, если z велико, а сам ряд оборвать на таком числе слагаемых n , при котором члены ряда еще не начали увеличиваться вследствие факториальных множителей, стоящих в числителе.

В этом легко убедиться, оценив величину остаточного члена R_n . Имеем

$$\begin{aligned}
 |R_n| &= \frac{2}{\sqrt{2\pi}} \cdot (2n-1)!! \int_z^\infty e^{-t^2/2} \cdot \frac{dt}{t^{2n}} < \\
 &< \frac{2(2n-1)!!}{\sqrt{2\pi}} \cdot e^{-z^2/2} \int_z^\infty \frac{dt}{t^{2n}} = \\
 &= \frac{2(2n-1)!!}{(2n-1)\sqrt{2\pi}} \cdot e^{-z^2/2} \frac{1}{z^{2n-1}},
 \end{aligned}$$

откуда следует

$$|R_n| < \frac{2}{z\sqrt{2\pi}} \cdot e^{-z^2/2} \frac{(2n-3)!!}{z^{2(n-2)}}.$$

Таким образом, величина остаточного члена по модулю меньше последнего учтенного члена (11.26) и может быть весьма малой, если z велико, а число учтенных членов ряда выбрано так, чтобы последний учтенный член ряда был достаточно малым.

Ниже приведена краткая таблица значений функции $\Phi(z)$, ее график дан на рис. 11.1.

Таблица 11.1
Функция Лапласа при целочисленных
значениях аргумента

z	1	2	3	4	5
$\Phi(z)$	0,6827	0,9545	0,9973	0,999936	0,9999974

Выше речь шла об однородных испытаниях Бернулли, в каждом из которых вероятность появления события A остается неизменной. Приведенные предельные теоремы допускают обобщение и на случай неоднородных испытаний, когда вероятность появления события в k -м опыте $p_k = \mathbf{P}(A_k)$ зависит от номера испытания k . Мы ограничимся одной из таких теорем, доказанной впервые Пуассоном.

Теорема Пуассона

Если в последовательности независимых испытаний вероятность появления события A в опыте с номером k равняется p_k , то

$$\lim_{n \rightarrow \infty} \mathbf{P}\{|\bar{p} - \bar{p}_n| < \varepsilon\} = 1, \quad (11.28)$$

где $\bar{p}$ означает частоту появления события A (11.1), а $\bar{p}_n$ — среднюю вероятность появления A в первых n испытаниях.

$$\bar{p}_n = \frac{1}{n} \sum_{k=1}^n p_k. \quad (11.29)$$

Теорема Пуассона является непосредственным обобщением теоремы Бернулли и точно также как эта последняя теорема представляет элементарный частный случай теоремы Чебышева. Поэтому отдельное доказательство теоремы Пуассона здесь не приводится.

Нужно сказать, что помимо приведенных классических теорем Бернулли, Лапласа и Пуассона, основываясь на общих предельных теоремах, которые будут рассмотрены далее (им

будет посвящена отдельная глава), можно доказать и некоторые другие предельные соотношения для схем повторных испытаний. Доказательство предельных теорем представляет интересный и методологически очень важный раздел вероятностной теории. Данный параграф носит характер элементарного введения в этот круг проблем. К этим вопросам мы еще вернемся после изучения теории случайных величин, а сейчас обратимся к примерам.

Пример 11.1. Вероятность появления события A при каждом из 100 независимых испытаний равна 0,3. Используя точные и приближенные формулы, определить вероятность того, что событие A появится: а) ровно 30 раз; б) от 40 до 42 раз; в) от 21 до 45 раз.

Решение. а) Если m — число появления события A при $n = 100$ независимых испытаниях, то по точной формуле (11.8)

$$P(m = 30) = C_{100}^{30} \cdot 0,3^{30} \cdot 0,7^{70} = \frac{100!}{30! 70!} 100^{-100} \cdot 3^{30} \cdot 7^{70}.$$

Используя таблицы логарифмов для факториалов чисел, получаем

$$\begin{aligned} \lg P(m = 30) &= 157,970 - 32,424 - 100,78 - \\ &\quad - 100 + 30 \cdot 0,4771 + 70 \cdot 0,8451 = -1,062. \end{aligned}$$

Искомая вероятность $P(m = 30) = 0,0867$. В данном случае $np = 30$, $npq = 21$. Условия $\frac{|m - np|}{\sqrt{npq}} \leq 3$, $np \geq 9q$, $nq \geq 9p$ выполняются, поэтому можно использовать формулу (11.18). Имеем

$$P(m = 30) \approx \Phi\left(\frac{0,5}{\sqrt{21}}\right) = \Phi(0,109) = 0,0867.$$

В соответствии с (11.14) находим $P(m = 30) \approx \frac{1}{\sqrt{2\pi\sqrt{21}}} = 0,0870$, то есть формула (11.14) дает две верные значащие цифры, на одну меньше, чем формула (11.18).

б) По приближенной формуле (11.17) получаем

$$\begin{aligned}
 P(40 \leq m \leq 42) &= \\
 &= \frac{1}{2} \left[\Phi \left(\frac{42,5 - 30}{\sqrt{21}} \right) - \Phi \left(\frac{39,5 - 30}{\sqrt{21}} \right) \right] = \\
 &= \frac{1}{2} [\Phi(2,729) - \Phi(2,073)] = \frac{1}{2} (0,99365 - 0,96172) = \\
 &= 0,01596 \approx 0,016.
 \end{aligned}$$

По точной формуле (11.18) находим $P(40 \leq m \leq 42) = 0,017$. Отличие в последнем знаке получилось потому, что в данном случае значения m лежат вблизи одной из границ интервала (11.19).

в) Используя формулу (11.17), находим

$$\begin{aligned}
 P(21 \leq m \leq 45) &= \\
 &= \frac{1}{2} \left[\Phi \left(\frac{45,5 - 30}{\sqrt{21}} \right) - \Phi \left(\frac{20,5 - 30}{\sqrt{21}} \right) \right] = \\
 &= \frac{1}{2} [\Phi(3,382) - \Phi(2,073)] = 0,9805.
 \end{aligned}$$

По точной формуле (11.8) в этом случае получим

$$P(21 \leq m \leq 45) = \sum_{m=21}^{45} C_{100}^m \cdot 0,3^m \cdot 0,7^{100-m} = 0,9829.$$

Пример 11.2. Вероятность того, что на лотерейный билет упадет выигрыш, равна 0,2. Какое количество билетов необходимо приобрести, чтобы с вероятностью 0,90 выиграли не менее трех билетов?

Решение. Обозначим через m число выигравших билетов из общего числа билетов n . В соответствии с (11.17) имеем

$$\begin{aligned}
 P\{3 \leq m \leq n\} &= 0,90 \approx \\
 &\approx \frac{1}{2} \left[\Phi \left(\frac{n + 0,5 - n \cdot 0,2}{\sqrt{n \cdot 0,2 \cdot 0,8}} \right) - \Phi \left(\frac{2,5 - n \cdot 0,2}{\sqrt{n \cdot 0,2 \cdot 0,8}} \right) \right].
 \end{aligned}$$

Следует ожидать, что n — достаточно большое число, чтобы можно было считать

$$\Phi\left(\frac{n + 0,5 - n \cdot 0,2}{\sqrt{n \cdot 0,2 \cdot 0,8}}\right) \approx 1.$$

Тогда

$$\Phi\left(\frac{0,2n - 2,5}{0,4\sqrt{n}}\right) = 0,8.$$

Находя по таблицам функции Лапласа аргумент функции, соответствующий значению функции 0,8, получаем $\frac{0,2n - 2,5}{0,4\sqrt{n}} \approx 1,28$, то есть

$$n - 2,56\sqrt{n} - 12,5 = 0.$$

Решая квадратное уравнение, находим

$$\sqrt{n} = 1,28 + \sqrt{14,14} = 5,04,$$

то есть число билетов n должно быть близко к 26. При найденном значении n

$$\Phi\left(\frac{n + 0,5 - n \cdot 0,2}{\sqrt{n \cdot 0,2 \cdot 0,8}}\right) = \Phi(3,48) = 0,9995,$$

чем подтверждается допустимость сделанного приближения.

Пример 11.3. Проводится серия из n независимых испытаний с заранее неизвестной вероятностью p появления события A в каждом испытании. По итогам эксперимента вероятность p^* приближенно оценивается с помощью полученной частоты $\bar{p}$. Примем в качестве меры точности такой оценки вероятность

$$r = \mathbf{P}\{|\bar{p} - p| \leq \varepsilon\}.$$

Какие ограничения следует наложить на p , чтобы вероятность r была не ниже заданного значения α , близкого к единице? Построить допустимый интервал изменения p для $n = 100$ испытаниям, уровня погрешности $\varepsilon = 0,05$ и минимальной точности оценки $\alpha = 0,9545$.

Решение. Перепишем выражение r в виде

$$r = \mathbf{P}\{m_1 \leq m \leq m_2\},$$

где

$$m_1 = (p - \varepsilon)n, \quad m_2 = (p + \varepsilon)n.$$

По теореме Муавра–Лапласа с учетом нечетности функции $\Phi(z)$ получаем

$$\begin{aligned} r &= \frac{1}{2} \Phi\left(\frac{(p + \varepsilon)n - np}{\sqrt{npq}}\right) - \Phi\left(\frac{(p - \varepsilon)n - np}{\sqrt{npq}}\right) = \\ &= \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) = \Phi\left(\varepsilon\sqrt{\frac{n}{p(1-p)}}\right). \end{aligned}$$

По условию эта вероятность должна быть не меньше α , то есть

$$\Phi\left(\varepsilon\sqrt{\frac{n}{p(1-p)}}\right) \geq \alpha.$$

Обозначим через z_α корень уравнения

$$\Phi(z) = \alpha, \quad (0 \leq \alpha \leq 1).$$

Благодаря монотонности функции Лапласа (см. рис. 11.1) получаем

$$\varepsilon\sqrt{\frac{n}{p(1-p)}} \geq z_\alpha,$$

откуда искомое ограничение на p находим в виде

$$p^2 - p + \frac{\varepsilon^2 n}{z_\alpha^2} \geq 0.$$

Перепишем последнее неравенство в виде

$$\left(p - \frac{1}{2}\right)^2 + \left(\frac{\varepsilon^2 n}{z_\alpha^2} - \frac{1}{4}\right) \geq 0.$$

Тогда становится очевидным, что при соблюдении условия

$$\varepsilon \geq \frac{z_\alpha}{2\sqrt{n}}$$

вероятность p может быть произвольной, а при обратном знаке неравенства

$$\varepsilon < \frac{z_\alpha}{2\sqrt{n}}$$

получаем два допустимых интервала изменения p

$$0 \leq p \leq \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\varepsilon^2 n}{z_\alpha^2}}, \quad \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{\varepsilon^2 n}{z_\alpha^2}} \leq p \leq 1.$$

Полученный результат объясняется следующим образом. Производная вероятности r по параметру p

$$\frac{dr}{dp} = \frac{\varepsilon\sqrt{n}}{\sqrt{2\pi}} \cdot \exp\left\{-\frac{\varepsilon^2 n}{2p(1-p)}\right\} \cdot \frac{(2p-1)}{\sqrt{p^3(1-p^3)}}$$

обращается в нуль при $p = \frac{1}{2}$, причем меняет знак с минуса на плюс. Следовательно, при $p = \frac{1}{2}$ имеем минимум r , равный

$$r_{\min} = \Phi(2\varepsilon\sqrt{n}).$$

Если выполнено условие (11.31), то аргумент функции Лапласа $2\varepsilon\sqrt{n} \geq z_\alpha$, а это значит, что

$$r_{\min} \geq \alpha$$

при всех $p \in [0, 1]$. Если же выполняется неравенство (11.32), то

$$r_{\min} < \alpha,$$

и из области допустимых значений p нужно выбросить интервал

$$\left(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\varepsilon^2 n}{z_\alpha^2}}, \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{\varepsilon^2 n}{z_\alpha^2}} \right),$$

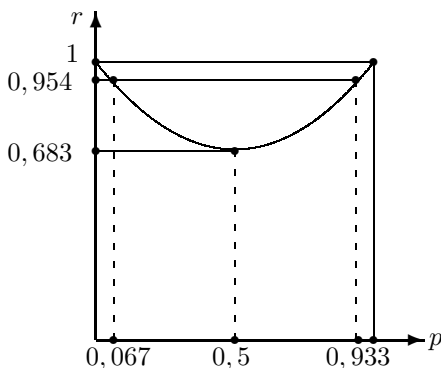


Рис. 11.2
К решению задачи 11.3

непосредственно примыкающий к точке $p = 1/2$. В указанном интервале вероятность r , характеризующая точность оценки, лежит ниже требуемого уровня α .

Для заданных в условии числовых значений параметров, согласно таблице 11.1, имеем $z_{0,9545} = 2$. При этом $\varepsilon = 0,05 < \frac{z_{0,9545}}{2\sqrt{n}} = 0,1$, так что реализуется неравенство (11.32), и имеем два допустимых интервала изменения p , а именно, $0 \leq p \leq \frac{1}{2} - \frac{\sqrt{3}}{4}$ и $\frac{1}{2} + \frac{\sqrt{3}}{4} \leq p \leq 1$ или $0 \leq p \leq 0,067$ и $0,933 \leq p \leq 1$. График зависимости r от p представлен на рис. 11.2.

Пример 11.4. Стрелок A поражает мишень с вероятностью $p_A = 0,8$, а стрелок B — с вероятностью $p_B = 0,6$. Между стрелками проводятся состязания по стрельбе, состоящие из n туров. В каждом туре выставляется новая мишень. Стрелки стреляют по ней поочередно, пока кто-либо не поразит цель. Право первого выстрела во всех нечетных турах имеет стрелок A , а в четных — стрелок B . Какова частота выигрыша каждым из стрелков отдельных туров состязания, если их общее число $n = 2m$ достаточно велико и четно?

Решение. Обозначим вероятность выигрыша одного тура соревнований спортсменом i при условии, что первый выстрел

делал спортсмен j , через $Q_i^{(j)}$, где $i, j \in \{A, B\}$. Эти вероятности, в принципе, можно было бы определить методом рекуррентных соотношений по аналогии с решением примера 10.4, однако проще сразу записать систему уравнений

$$\begin{cases} Q_A^{(A)} = 0,8 + 0,2Q_A^{(B)}, & Q_B^{(B)} = 0,6 + 0,4Q_B^{(A)}, \\ Q_A^{(A)} + Q_B^{(A)} = 1, & Q_A^{(B)} + Q_B^{(B)} = 1. \end{cases}$$

Первое уравнение означает следующее: стрелок A , начиная первым, сразу выигрывает с вероятностью $0,8$, либо промахивается с вероятностью $0,2$ и превращается в стреляющего вторым. Второе уравнение подобным же образом описывает выигрыш в данном туре второго стрелка. Два последних уравнения представляют условие нормировки для каждого из вариантов начала тура.

Решая приведенную систему, получаем

$$Q_A^{(A)} = \frac{20}{23}, \quad Q_B^{(A)} = \frac{3}{23}, \quad Q_B^{(B)} = \frac{15}{23}, \quad Q_A^{(B)} = \frac{8}{23}.$$

При чередовании права первого выстрела имеем неоднородные испытания Бернулли, в которых вероятность успеха игрока A в k -м туре стрельбы дается равенством

$$p_k = \begin{cases} \frac{20}{23}, & k = 2m + 1, \\ \frac{8}{23}, & k = 2m. \end{cases}$$

Очевидно, среднее значение p_k за $n = 2m$ туров состязания

$$\overline{p_n} = \frac{1}{2m} \sum_{k=1}^{2m} p_k = \frac{1}{2} \left(\frac{20}{23} + \frac{8}{23} \right) = \frac{14}{23}.$$

На основании теоремы Пуассона можно утверждать, что при $m \gg 1$ частота выигрышей первым игроком отдельных туров будет сходиться по вероятности к этому же значению, то есть

$$P_A \approx \frac{14}{23}, \quad P_B \approx \frac{9}{23}.$$

Таким образом, игрок A будет выигрывать примерно в полтора раза чаще, чем его соперник.

Пример 11.5 (Продолжение примера 11.4). В условиях примера 11.4 предлагается уравнивать шансы на победу обоих стрелков. Для этого стрелок A дает стрелку B фору. После каждого тура, когда A стреляет первым, право первого выстрела переходит к B сразу на следующие l туров. Допустим, что всего производится $n = m(l + 1)$ туров стрельбы. Как следует распорядиться числом l , чтобы отношение частот выигрыша обоих участников было как можно ближе к единице?

Решение. В данном случае имеем неоднородные испытания Бернулли, в которых каждое $(l + 1)$ испытание заканчивается успехом A с вероятностью $\frac{20}{23}$, а остальные испытания он выигрывает с вероятностью $\frac{8}{23}$. Поэтому

$$p_k = \begin{cases} \frac{20}{23}, & k = (l + 1)s + 1, \\ \frac{8}{23}, & (l + 1)s + 2 \leq k \leq (l + 1)(s + 1), \end{cases}$$

где $s = \overline{0, \infty}$.

Среднее значение вероятностей p_k за $n = m(l + 1)$ туров состязаний

$$\overline{p_n} = \frac{1}{m(l + 1)} \sum_{k=1}^{m(l+1)} p_k = \frac{(20 + 8l)}{23(l + 1)}.$$

Если $m \gg 1$, то тогда и $n \gg 1$, благодаря чему по теореме Пуассона заключаем, что

$$P_A \approx \frac{(20 + 8l)}{23(l + 1)}, \quad P_B = \frac{3 + 15l}{23(l + 1)}.$$

Рассмотрим отношение

$$\Delta_l = \frac{P_A}{P_B} = \frac{20 + 8l}{3 + 15l}, \quad (l = \overline{1, \infty}).$$

Таблица 11.2

Отношение частот в задаче о стрельбе

1	1	2	3	4	5	∞
Δ_l	1,556	1,091	0,917	0,825	0,769	0,533

Эта величина монотонно убывает от значения $\Delta_1 = 14/9 = 1,556$ до $\Delta_\infty = 8/15 = 0,533$. Конкретные числовые значения Δ_l приводятся в таблице 11.2.

Из приведенных числовых результатов следует, что наиболее справедливым является вариант, когда менее опытный стрелок получает право в три раза чаще стрелять первым. При этом Δ_3 отличается от единицы в меньшую сторону на 0,083. При $l = 2$ значение Δ_l будет больше единицы на 0,091. Вариант $l = 1$, разобранный в предыдущем примере, дает большее преимущество первому стрелку, все остальные значения l — второму, причем в случае $l \rightarrow \infty$ второй стрелок имеет примерно такое же преимущество над первым, как первый при $l = 1$ над вторым.

АКСИОМАТИЧЕСКОЕ ПОСТРОЕНИЕ ТЕОРИИ СЛУЧАЙНЫХ СОБЫТИЙ

§ 12. АКСИОМАТИКА КОЛМОГОРОВА. ПОНЯТИЕ ПРОСТРАНСТВА ЭЛЕМЕНТАРНЫХ СОБЫТИЙ

Введенное в § 2 классическое определение вероятности позволило установить основные правила исчисления вероятностей. Однако это определение охватывает только лишь весьма узкий класс случайных событий. К нему относятся те случайные события, для которых можно указать конечное число элементарных исходов испытаний, равноценных с точки зрения возможности их реализации на опыте. Иначе говоря, случайный эксперимент должен обладать известной симметрией.

Переход к геометрическим вероятностям (см. § 3), безусловно, несколько расширяет возможности классической трактовки случайных явлений, однако он далеко не исчерпывает все те задачи, решение которых вероятностными методами представляется возможным и к которым теория вероятностей фактически начала применяться с самого момента ее возникновения. Большинство этих приложений оказались удачными и увенчались рядом новых полезных научных результатов (вопросы страхования, анализ ошибок измерения, молекулярная физика и ряд других).

Вместе с тем были также и случаи, когда применение вероятностных методов не только не привело к положительным результатам, но даже в известной степени скомпрометировало саму теорию вероятностей в глазах части математиков и представителей прикладных наук. Назрела необходимость

в строгом построении теории, пригодной не только к рассмотрению случайных явлений, трактуемых по классическим канонам, но и гораздо более широкого класса случайных явлений, встречающихся в естествознании и технике.

Естественным способом построения такой теории является подход, обычный для математики вообще: вначале должны быть сформулированы исходные, основополагающие принципы теории (аксиомы), а затем на их базе дедуктивным путем должны быть выведены вытекающие из них остальные положения теории. При этом, как известно, при соблюдении обязательных требований полноты и непротиворечивости системы аксиом в остальном эти аксиомы могут быть выбраны неоднозначным образом и в значительной мере произвольно.

Разумеется, к построению содержательной науки может привести только та система аксиом, при формулировке которой учтены реально наблюдаемые свойства достаточно широкого класса случайных явлений. Иначе говоря, при аксиоматизации должен быть учтен опыт изучения тех случайных явлений, для которых понятие вероятности, несомненно, себя оправдало. К ним относятся случайные события, укладываемые в классическую схему случайного эксперимента, а также разнообразные массовые явления, давно и успешно изучаемые в рамках вероятностного подхода.

Важную методическую роль здесь играют предельные теоремы, простейшие варианты которых были рассмотрены в предыдущем параграфе. Если частоты появления событий в серии испытаний при определенных условиях близки к соответствующим вероятностям, значит аксиомы должны не противоречить экспериментально наблюдаемым свойствам частот.

Напомним, что первоначально теория вероятностей возникла как сугубо прикладная наука. Случайные явления рассматривались неразрывно с порождающим их физическим экспериментом. Однако в конце XIX века были установлены связи между теорией вероятностей и чистой (абстрактной) математикой. В знаменитом программном докладе, прочитанном в 1900 году в Париже Д. Гильбертом¹, проблема аксиоматиза-

¹ Гильберт Давид (1862–1943) — немецкий математик.

ции теории вероятностей была названа в числе наиболее актуальных математических проблем XX века. Так называемая «шестая проблема Гильберта» заключается в аксиоматизации тех физических дисциплин, в которых преобладающую роль играет математика. К числу последних Гильберт отнес механику и теорию вероятностей, которая тогда считалась ближе стоящей к физике, чем к математике [7].

После доклада Гильберта было предпринято несколько попыток строгого построения теории вероятностей как аксиоматизированной дисциплины. Среди них необходимо в первую очередь упомянуть Р. Мизеса и С. Н. Бернштейна. Если Мизес исходил из частотного определения вероятности, которое не получило широкого распространения, то Бернштейн в 1917 году дал первое строгое универсальное построение вероятностной теории с использованием методов теории множеств и теории меры [3].

После Бернштейна были предложены и некоторые другие системы аксиом, из которых наиболее удачной оказалась аксиоматика, разработанная в 1933 году А. Н. Колмогоровым. Она была подробно изложена в специальной брошюре [12], впервые опубликованной на немецком языке в Германии. *Аксиоматика Колмогорова* связывает теорию вероятностей с теорией множеств, алгеброй и теорией меры. Она охватывает единым подходом как случайные события, о которых шла речь в предыдущих главах, так и более сложные объекты, изучаемые теорией вероятностей, такие, как случайные величины, рассматриваемые далее в настоящем пособии, а также случайные функции, описывающие случайные явления в их динамике [22].

Существенно, что аксиоматика Колмогорова не ограничивает физическое содержание рассматриваемых случайных явлений и, что очень важно, степень сложности этих явлений. Благодаря этому она может служить строгой основой для развития самых разнообразных новых разделов вероятностной теории, постоянно возникающих под влиянием прогресса естествознания и техники. В качестве примера можно назвать теорию случайных полей, теорию случайных (нечетких) множеств, теорию случайных графов, теорию случайных

матриц и другие вновь появившиеся разделы вероятностной теории.

Прежде чем непосредственно переходить к изложению самого содержания аксиоматики, необходимо сделать одно методическое замечание о причинах возникновения случайности в физическом эксперименте. Любой эксперимент определяется условиями его постановки. Для случайных экспериментов характерно следующее. Помимо контролируемого экспериментатором комплекса заявленных условий имеются еще некоторые скрытые, умалчиваемые условия, которые не контролируются и не регистрируются, но способны в той или иной степени повлиять на исход опыта. В результате при одних и тех же контролируемых условиях исход испытания становится непредсказуемым. Если эти скрытые условия характеризуются устойчивостью и повторяемостью, то в длинной серии испытаний частоты различных исходов будут стабилизироваться, что позволяет трактовать такую задачу как вероятностную.

Например, в опыте с монетой или, скажем, с игральной костью берется одна и та же монета (игральная кость), с которой совершаются внешне очень похожие действия, однако чрезвычайно сложно обеспечить полную идентичность всех бросаний. Несколько меняется высота, с которой отпускается бросаемый предмет, придаваемый ему вращательный момент, угол наклона к поверхности стола и т. п. Все эти, казалось бы, мелкие, пренебрежимо малые отличия в начальных условиях на самом деле существенно влияют на характер движения монеты или кости и приводят к неоднозначному финальному результату, который трактуется как случайный при заданных контролируемых условиях опыта.

В аксиоматике Колмогорова вероятностная модель любого случайного эксперимента формируется из трех стандартных составных частей, которые строятся последовательно друг за другом по общим правилам, не зависящим от конкретного физического содержания решаемой задачи. Эти этапы или шаги включают:

1) формирование *пространства элементарных событий* Ω ; 2) построение *алгебры* (или *сигма-алгебры*) *случайных событий* $\mathfrak{A}$; 3) задание на $\mathfrak{A}$ *вероятностной меры* P .

Совокупность всех этих трех элементов $(\Omega, \mathfrak{A}, P)$ образует так называемое *вероятностное пространство* данной задачи.

Вначале разберем первый из элементов вероятностного пространства — *пространство элементарных событий*, которому посвятим оставшуюся часть данного параграфа.

Напомним (см. § 1), что *испытанием* или *опытом* называют воспроизведение некоторого заданного комплекса условий с целью наблюдения происходящих при этом *событий*. Под событием здесь понимается любое физическое явление или любой значимый физический факт, являющийся результатом, следствием проведенного испытания. Каждому физическому эксперименту соответствует некоторое множество связанных с ним событий. Это множество может быть конечным, бесконечным, счетным или несчетным.

Задать событие — это значит указать признаки, по которым можно однозначно определить, произошло событие или нет. Применительно к событиям возможны только два этих варианта (произошло или нет), при возникновении сомнительных ситуаций необходимо уточнить признаки события, чтобы всегда был единственный ответ. События по-разному проявляют себя в традиционном детерминированном и случайном (стохастическом) эксперименте.

Допустим вначале, что экспериментатор в состоянии контролировать все без исключения существенные факторы, влияющие на исход испытания. Тогда эксперимент превращается в детерминированный (неслучайный). В таком эксперименте появление или непоявление любого события однозначно определяется выбранными условиями эксперимента. При этом множество происходящих при испытании событий будет неизменным при любом повторении такого испытания.

Принципиально иначе ставится *случайный эксперимент*, изучаемый вероятностными методами. Здесь от экспериментатора ускользает часть скрытых, неконтролируемых условий, способных влиять на исход испытания. В результате множество происходящих в эксперименте событий будет случайным (некоторые авторы называют такие множества *нечеткими*). Если же выделить какое-то одно отдельно взятое событие, то его появление становится непредсказуемым: в некоторых

испытаниях оно появляется, а в других — нет. Именно такие события § 1 были названы случайными.

При построении абстрактной математической модели случайного эксперимента вполне естественным и логичным является начать с описания множества его возможных исходов, то есть полного комплекса всех тех событий, которые хоть когда-нибудь могли бы случиться и не противоречат условиям эксперимента. Естественно, как и во всякой теории, при формировании этого множества приходится делать некоторые упрощения, исключая из рассмотрения аномальные, экзотические исходы, свидетельствующие об ошибках в постановке опыта, действии внешних мешающих факторов и тому подобное. Например, в задаче о бросании монеты можно не учитывать, что она может встать на ребро, игральный кубик может упасть со стола, при выдаче игровых карт по ошибке можно сразу вытащить две карты и так далее. Объектом математического исследования является некий идеализированный эксперимент, в котором жестко очерчены все возможные варианты его завершения.

В аксиоматике Колмогорова множество возможных исходов опыта определенным образом структурируется. В нем выделяются так называемые *элементарные исходы* или *элементарные события*. Элементарное событие представляет собой простейший неразложимый исход испытания. Элементарные события характеризуются следующими свойствами: 1) в ходе опыта одно из них обязательно происходит; 2) появление одного из них исключает появление других; 3) элементарное событие неразложимо.

Поясним последнее требование. Условимся различать составные (или разложимые) события и элементарные (или неразложимые) события. Например, при бросании игрального кубика событие «выпало четное число очков» можно представить в виде суммы трех элементарных событий («выпало два очка», «выпало четыре очка», «выпало шесть очков»). Перечисленные три события в рамках данного эксперимента уже неразложимы, и их следует признать элементарными.

Множество всех элементарных событий называется *пространством элементарных событий*. По Колмогорову нель-

зя говорить ни о каких вероятностях вне их связи с определенным пространством элементарных событий, то есть вне связи с каким-либо (хотя бы мысленным) физическим экспериментом. Элементарные события рассматриваются как раз и навсегда заданные в решаемой задаче. Они являются первоначальными и неопределимыми понятиями всей теории. После формирования пространства элементарных событий любая физическая задача независимо от ее конкретного содержания переводится в сугубо математическую плоскость.

Для того чтобы пространство элементарных событий можно было бы использовать для математических выводов, его необходимо параметризовать. Поставим в соответствие каждому элементарному исходу испытания некоторый числовой (скалярный или векторный) параметр ω . Пространство элементарных событий условимся обозначать прописной буквой Ω .

Легко видеть, что задачи, допускающие классическое определение вероятности, без всякого труда укладываются в описанную схему (см. § 2). В этом случае в роли элементарных событий выступают несовместные и равновозможные исходы опыта E_k ($k = \overline{1, n}$). Эти элементарные события легко параметризуются, если параметр ω принять равным номеру исхода, то есть положить $\omega_i = i$, а тогда пространство элементарных событий

$$\Omega = \{i : i = \overline{1, n}\} \quad (12.1)$$

будет совпадать с отрезком натурального ряда чисел.

При геометрическом определении вероятности (см. § 3) случайный эксперимент сводится к выбору случайной точки из заданной области Ω , причем таким образом, чтобы попадание точки в равновеликие части дозволенной области было равновозможным. В данном случае элементарным событиям соответствуют различные точки $\omega \in \Omega$, а сама эта область представляет собой ничто иное, как пространство элементарных событий.

Ранее было показано, что геометрическое определение вероятности фактически сводится к классическому определению, если в последнем допустить возможность предельного

перехода при неограниченном увеличении n . Аксиоматическое определение вероятности в известном смысле обобщает геометрическое, но с рядом существенных отличий. Во-первых, пространство элементарных событий Ω может быть совершенно произвольным и не обязано допускать наглядную геометрическую интерпретацию. Во-вторых, снимаются какие-либо ограничения, касающиеся равновозможности попадания изображающей точки в различные части Ω . Случайный эксперимент не обязан обладать симметрией, а вероятности приходится строить более сложным путем, на основе теории меры, как это будет описано в дальнейшем.

Далее рассмотрим примеры, поясняющие понятие пространства элементарных событий.

Пример 12.1. Построить пространство элементарных событий при однократном подбрасывании монеты.

Решение. При математическом описании этого опыта естественным является отвлечься от таких несущественных «возможностей», как «монета встала на ребро», «монета скатилась со стола и исчезла» и тому подобные. В идеализированном эксперименте возможны лишь два исхода: O — выпал орел и P — выпала решка. Этому идеализированному эксперименту соответствует пространство элементарных событий

$$\Omega = \{O, P\}.$$

В данном случае мы отказались от нумерации исходов цифрами, как в (12.1), и сохранили для наглядности буквенные обозначения. Общее число исходов $N_\Omega = 2$.

Пример 12.2. Построить пространство элементарных событий для серии из n испытаний Бернулли, в каждом из которых подбрасывается монета.

Решение. В данном случае для обозначения элементарных событий удобно использовать векторное обозначение $\vec{\omega} = \{\omega_i\}_{i=1}^n$, где ω_i — символ, показывающий, как закончилось i -е испытание Бернулли. Каждое из этих испытаний аналогично опыту, рассмотренному в предыдущем примере, поэтому

$$\Omega = \{\vec{\omega} : \vec{\omega} = \{\omega_i\}_{i=1}^n, \omega_i \in \{O, P\}, i = \overline{1, n}\}.$$

Общее число элементарных событий, составляющих пространства Ω , дается равенством $N_\Omega = 2^n$.

Пример 12.3. Построить пространство элементарных событий при однократном бросании шестигранного игрального кубика («игральной кости»).

Решение. На гранях кубика выбиты очки от 1 до 6. Случайный эксперимент заключается в определении числа выпавших очков. Пространство элементарных событий здесь можно ввести по правилу (12.1)

$$\Omega = \{i : i = \overline{1, 6}\}.$$

Ясно, что $N_\Omega = 6$.

Пример 12.4. Проводятся сложные испытания Бернулли, в каждом из которых подбрасывается игральная кость. Требуется сформировать для серии n таких испытаний пространство Ω .

Решение. Как и в предыдущем примере нас интересует число выпавших очков. Результат эксперимента здесь определяется числом очков, полученным в каждом из опытов. Поэтому при параметризации возможных исходов следует использовать векторный параметр $\vec{\omega} = \{\omega_i\}_{i=1}^n$, где ω_i определяется результатом i -го испытания. Очевидно, в данном случае

$$\Omega = \{\vec{i} : \vec{i} = \{i_s\}_{s=1}^n, i_s = \overline{1, 6}, s = \overline{1, n}\},$$

а общее число исходов $N_\Omega = 6^n$.

Во всех приведенных примерах пространство элементарных событий оказалось конечным множеством, поскольку для всех них $N_\Omega < \infty$. Выполнение этого неравенства отнюдь не является обязательным. Во многих практически интересных вероятностных задачах приходится рассматривать случайные эксперименты, для которых $N_\Omega = \infty$. Бесконечное число возможных исходов получается, например, в задачах на геометрические вероятности, а также в задачах на повторные испытания, если число последних не ограничивается.

Пример 12.5. Проводится серия повторных испытаний, в которых монету бросают до первого появления орла. Требуется построить пространство элементарных событий.

Решение. Результатом такого эксперимента можно считать номер опыта, в котором впервые появится орел (O). Число предшествующих ему опытов i , в которых выпала решка (P), может быть произвольным целым неотрицательным числом. Принимая это число i в качестве параметра элементарного события, получим

$$\Omega = \{i : i = \overline{0, \infty}\}.$$

Если обозначать элементарные события аналогично примеру 12.2, то в данном примере i -й исход соответствует комбинации $P...PO$, где буква P перед O должна быть выписана $(i - 1)$ раз. Разумеется, здесь $N_\Omega = \infty$.

Пример 12.6 (см. пример 3.2). Вернемся к ранее рассмотренной задаче о встрече. Допустим, что каждый из встречающихся может опоздать по отношению к оговоренному моменту встречи на случайное время, которое не превосходит T . Каково в этом случае будет пространство элементарных событий?

Решение. В качестве элементарного события здесь следует принять событие, состоящее в том, что i -й участник встречи опоздал на t_i , где $0 \leq t_i \leq T$, ($i = \overline{1, 2}$). Поэтому

$$\Omega = \{(t_1, t_2) : 0 \leq t_i \leq T, i = \overline{1, 2}\},$$

то есть элементарному событию соответствует точка квадрата со стороной T . Число таких точек бесконечно, поэтому $N_\Omega = \infty$.

Пример 12.7 (см. пример 3.3). Построить пространство элементарных событий в задаче о составлении треугольника.

Решение. Напомним, что в этой задаче берутся три одинаковых отрезка прямых длиной l , от которых по случайному закону отламываются куски произвольной длины. Обозначим оставшиеся части отрезков через x, y, z . Данному случайному эксперименту можно сопоставить точку в трехмерном пространстве с координатами (x, y, z) , тогда, принимая координаты этой точки в качестве параметра элементарного события, получим

$$\Omega = \{(x, y, z) : 0 \leq x \leq l, 0 \leq y \leq l, 0 \leq z \leq l\},$$

что дает $N_\Omega = \infty$.

Несмотря на то, что во всех последних примерах $N_\Omega = \infty$, между ними есть существенное отличие. В примере 12.5 все элементы Ω можно занумеровать, то есть это множество является хотя и бесконечным, но счетным. В двух других примерах такая нумерация невозможна, а множество Ω является несчетным.

Если множество Ω является конечным или счетным (про такие множества говорят, что они не более, чем счетны), то такое пространство элементарных событий называют *дискретным*.

Во всех рассмотренных примерах пространство элементарных событий Ω является точечным множеством, то есть его можно представить как некоторую область пространства конечной размерности. Такая ситуация типична при рассмотрении случайных событий, а также случайных величин, изучаемых в следующих главах. Между тем аксиоматический подход сохраняет силу и для множеств Ω более общего вида. Подобные задачи возникают, например, в теории случайных функций [22, 23].

Рассмотрим частицу, совершающую броуновское движение [23]. Результатом опыта по наблюдению за такой частицей будет некоторая реализация ее траектории. Для такой, гораздо более сложной, чем мы сейчас рассматриваем, задачи, тем не менее, можно ввести пространство элементарных событий, которое будет включать множество всевозможных траекторий частицы.

Элементы ω этого пространства будут являться функциями времени, задающими различные возможные варианты движения частицы. Следовательно, элементы ω уже не являются точками конечномерного пространства, а принадлежат некоторому бесконечномерному пространству.

Этот пример, разумеется, выходит за рамки теории случайных событий и даже всего нашего курса, однако он иллюстрирует возможности аксиоматического подхода при рассмотрении случайных явлений.

Универсальность, общность и строгость аксиоматического подхода позволяет изучать самые разнообразные и сложные вероятностные модели.

§ 13. АЛГЕБРА И СИГМА-АЛГЕБРА СОБЫТИЙ. АКСИМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Во всякой аксиоматической теории имеются исходные, первоначальные понятия, не определяемые в рамках этой теории. В аксиоматике теории вероятностей таким понятием является элементарное событие. Строго говоря, выделение элементарных событий и их физическая трактовка лежит вне рамок теории вероятностей. Собственно вероятностный анализ начинается с того, чтобы каким-то образом связать происходящие в эксперименте сложные события с заданным в качестве исходных данных множеством элементарных исходов.

Здесь важную роль играет понятие *наблюдаемого* в данном эксперименте *события*. Как уже отмечалось, выделение пространства Ω означает некоторую идеализацию эксперимента. Выделенные элементарные события позволяют выразить далеко не все мыслимые события, а лишь некоторую их часть, называемую наблюдаемыми событиями.

Поясним это на примерах. Обратимся к примеру 12.3 (бросание игральной кости). Если интересоваться событиями, касающимися числа выпавших очков, то все такие события при задании Ω в виде, указанном в примере, наблюдаемы. Вместе с тем, событие «после падения на стол кость перевернулась k раз» уже не будет наблюдаемым, так как выделенные элементарные события не несут в себе необходимую для ответа на этот вопрос информацию.

Рассмотрим также пример 12.5 (бросание монеты до первого появления орла). Допустим, что в игре участвуют два игрока, перед каждым новым броском один из них выбирается по жребию и совершает это бросок. Рассматривается событие A — «орел выбросил первый игрок». Очевидно, это событие не является наблюдаемым при системе элементарных событий, предложенных в примере. Вместе с тем, событие A можно сделать наблюдаемым, если уточнить правила игры. Допустим, что игроки бросают монету поочередно, начиная с первого игрока. Тогда событие A эквивалентно событию «орел впервые появился при $2k + 1$ бросании ($k = \overline{0, \infty}$)». Последнее собы-

тие уже становится наблюдаемым, если задать элементарные исходы, как в примере.

Допустим, что некоторое событие A наблюдаемо в данном эксперименте. С точки зрения указанного события эксперимент может закончиться двояко: либо произошло событие A , либо событие A не произошло, то есть произошло противоположное событие $\bar{A}$. С другой стороны, эксперимент в своем идеальном, модельном виде может закончиться только появлением элементарных исходов ω , и притом всякий раз только какого-то одного из них. Следовательно, множество Ω можно разбить на два подмножества: Ω_A — множество элементарных событий благоприятствующих появлению A , и $\Omega_{\bar{A}}$ — множество ω , противоречащих A .

Ясно, что так введенные множества удовлетворяют условиям

$$\Omega_A \subset \Omega, \quad \Omega_{\bar{A}} \subset \Omega, \quad \Omega_A \cup \Omega_{\bar{A}} = \Omega, \quad \Omega_A \cap \Omega_{\bar{A}} = \emptyset, \quad (13.1)$$

где $\emptyset$ обозначает пустое множество. Про множество Ω_A говорят, что оно является *интерпретацией* события A или *отражением* события A во множестве Ω .

Вообще говоря, само событие A и его интерпретация Ω_A в пространстве Ω физически представляют совершенно разные объекты: первое — некоторое явление материального мира, а второе — область в пространстве параметров ω . Следующий шаг аксиоматизации заключается в *отождествлении* события A и множества Ω_A .

Таким образом, в дальнейшем любое событие A понимается как некоторое подмножество Ω , состоящее из всех таких точек ω , которые благоприятствуют A . Если в результате эксперимента реализовался исход $\omega \in A$, то событие A произошло, если же $\omega \notin A$ (то есть $\omega \in \bar{A}$), то с точки зрения появления события A эксперимент завершился безрезультатно (A не произошло).

Трактовка событий как некоторых множеств, являющихся подмножествами пространства Ω , позволяет перенести на множество случайных событий, относящихся к данному случайному эксперименту, алгебраические операции, принятые в теории множеств, а также ряд основополагающих понятий,

употребительных в этой теории. Разумеется, содержание этих понятий при вероятностной трактовке интерпретируется по-новому.

В теории множеств среди подмножеств данного множества Ω особо выделяют два его подмножества — само множество Ω и пустое $\emptyset$. Множество Ω при такой трактовке его как события означает событие, которое в данном эксперименте происходит всегда, то есть эквивалентно достоверному событию U . Пустым называют множество, не содержащее ни одного элемента ω , следовательно, $\emptyset$ отражает невозможное событие V , которое в рассматриваемом эксперименте не происходит никогда.

Для множеств определено отношение порядка или принадлежности. Говорят, что множество A принадлежит B и пишут $A \subset B$, если все элементы ω первого множества являются также элементами и второго. Применительно к событиям это означает, что A влечет за собой событие B , то есть при каждом наступлении A неизбежно происходит и B . Если одновременно $A \subset B$ и $B \subset A$, то такие события естественно считать эквивалентными, то есть совпадающими и не различать их, записывая при этом $A = B$.

Далее сумма событий и их произведение определяется в духе теории множеств как объединение и пересечение множеств, являющихся отражением этих событий в пространстве Ω . Именно этим объясняется возможность использования обозначения $A \cap B$ для произведения событий A и B и обозначение $A \cup B$ для их суммы. При этом полностью сохраняется математический смысл этих понятий, который вкладывается в них при элементарном определении в § 1.

Отметим некоторые свойства так определенных операций над случайными событиями. Прежде всего заметим, что для любого события A по определению выполняется двойное включение

$$\emptyset \subset A \subset \Omega. \quad (13.2)$$

По отношению к событиям $\emptyset$ и Ω результат их сложения с любым событием A и умножения на это событие таков:

$$A \cup \emptyset = A, \quad A \cap \Omega = A, \quad (13.3)$$

$$A \cap \emptyset = \emptyset, \quad A \cap \Omega = A. \quad (13.4)$$

С этой точки зрения невозможное событие напоминает нуль в алгебре чисел, а достоверное ведет себя как единица по отношению к операции умножения.

Интересно заметить, что любое событие A не изменится, если его сложить с самим собой или умножить на себя

$$A \cup A = A, \quad A \cap A = A. \quad (13.5)$$

Это принципиально отличается от свойств аналогичных операций над числами, также, как и второе равенство (13.3).

Сумма событий A и B в известном смысле слова ограничивает каждое из слагаемых, поскольку оба слагаемых оказываются включенными в сумму

$$A \subset (A \cup B), \quad B \subset (A \cup B) \quad (13.6)$$

и при этом для любого события C , в которое включены и A , и B ($A \subset C$, $B \subset C$), справедливо включение

$$(A \cup B) \subset C, \quad (13.7)$$

то есть сумма таких событий также влечет за собой наступление C .

Для произведения событий знак включения (13.6) меняется на противоположный

$$(A \cap B) \subset A, \quad (A \cap B) \subset B. \quad (13.8)$$

При этом выполнение включения

$$C \subset (A \cap B) \quad (13.9)$$

приводит к тому, что C оказывается включенным в каждый из сомножителей ($C \subset A$, $C \subset B$).

Формально алгебраически несовместные события определяются условием

$$A \cap B = \emptyset. \quad (13.10)$$

При этом аналогично § 5 можно ввести и систему несовместных событий $(A_1, \dots, A_n)$, которую составляют попарно независимые события.

Рассмотрим некоторое событие A , отражением которого в пространстве Ω является множество Ω_A . Напомним, что дополнение Ω_A до Ω согласно (13.1) обозначалось через $\Omega_{\bar{A}}$. Предположим, что $\Omega_{\bar{A}}$ само принадлежит Ω , то есть отождествляется с некоторым событием. Это событие, происходящее тогда и только тогда, когда не происходит A , называется *противоположным* A и обозначается $\bar{A}$. Ясно, что

$$A \cup \bar{A} = \Omega, \quad A \cap \bar{A} = \emptyset, \quad (13.11)$$

причем A и $\bar{A}$ образуют полную группу событий.

Часто рассматривают также разность событий. Разность $A \setminus B$ представляет собой множество, состоящее из тех элементов A , которые не входят в B . Операция вычисления обратного события (инверсия) и операция вычисления разности событий выражаются друг через друга следующим образом:

$$\bar{A} = \Omega \setminus A, \quad A \setminus B = A \cap \bar{B}. \quad (13.12)$$

Следовательно, можно определить лишь какую-то одну из этих двух операций, тогда другая выражается через первую автоматически.

Продолжим ряд формул (13.2)–(13.4), касающихся операций над событиями $\emptyset$ и Ω применительно к операциям вычитания и инвертирования. Имеем

$$\bar{\emptyset} = \Omega, \quad \bar{\Omega} = \emptyset, \quad \Omega \setminus \emptyset = \Omega, \quad \emptyset \setminus \Omega = \emptyset, \quad (13.13)$$

$$A \setminus \Omega = \emptyset, \quad \Omega \setminus A = \bar{A}, \quad A \setminus \emptyset = A, \quad \emptyset \setminus A = \emptyset. \quad (13.14)$$

Если считать операцию инвертирования определенной, то из двух операций — сложения и умножения событий, одна будет выражаться через другую. Это следует из формул де Моргана, разобранных в § 1. Например, если взять за основу операцию сложения событий, то тогда в обозначениях настоящего параграфа

$$A \cap B = \overline{\bar{A} \cup \bar{B}}, \quad (13.15)$$

то есть произведение двух событий является событием, противоположным по отношению к сумме событий, противоположных сомножителям.

Рассмотрим некоторые примеры построения случайных событий по заданному пространству элементарных событий.

Пример 13.1 (Продолжение примера 12.2). Пусть проводится серия из $n = 3$ бросаний монеты, и событие A состоит в том, что в этой серии ни разу после орла не выпала решка. Построить отражение A в пространстве Ω , введенном в примере 12.2.

Решение. В данном случае имеем $N_{\Omega} = 2^3 = 8$,

$$\Omega = \{OOO, OOP, OPO, POO, OPP, POP, PPO, PPP\}.$$

Интересующему нас событию благоприятствуют исходы ω_1 , ω_4 , ω_7 , ω_8 , следовательно,

$$A = \{OOO, POO, PPO, PPP\}.$$

Пример 13.2 (Продолжение примера 12.3). Проводится однократное бросание игральной кости. Событие A состоит в том, что выпавшее число очков является простым числом. Построить событие A .

Решение. Как было показано в примере 12.3, пространство Ω в данном случае состоит из шести младших натуральных чисел

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Из этих чисел простыми являются только числа 2, 3 и 5, поэтому

$$A = \{2, 3, 5\}.$$

Пример 13.3 (Продолжение примера 12.5). Два игрока поочередно бросают монету до тех пор, пока в первый раз не выпадет орел, причем начинает первый игрок. Пусть событие A заключается в том, что именно первый игрок выбросил орла и победил в игре. Выразить это событие через элементарные события, введенные в примере 12.5.

Решение. Напомним, что элементарные исходы нумеровались числом i неудачных бросаний до первого выигрышного. При поочередной смене бросающих первый игрок выигрывает, если i оказалось четным. Следовательно,

$$\Omega = \{i\}_{i=0}^{\infty}, \quad A = \{2i\}_{i=0}^{\infty}.$$

Пример 13.4 (Продолжение примера 12.6). Допустим, что в задаче о встрече (см. пример 3.2) каждый встречающийся ожидает другого не более, чем t_0 минут. Построить событие A , состоящее в том, что встреча все-таки состоялась.

Решение. Напомним, что пространство элементарных событий представляет здесь квадрат со стороной T (рис. 13.1), то есть

$$\Omega = \{(t_1, t_2) : 0 \leq t_i \leq T, i = \overline{1, 2}\}.$$

Интересующее нас событие A дается той частью квадрата, которая ограничена прямыми $t_2 = t_0 \pm t_1$, так что

$$A = \{(t_1, t_2) : |t_2 - t_1| \leq t_0, 0 \leq t_i \leq T, i = \overline{1, 2}\}.$$

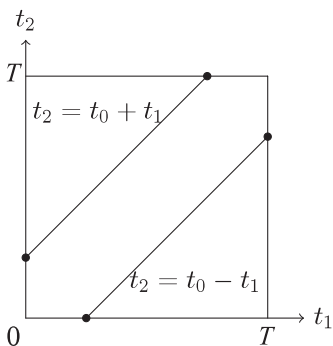


Рис. 13.1
К задаче о встрече

Рассмотрим некоторый случайный эксперимент, для которого сформировано пространство элементарных событий Ω . В соответствии с данными выше разъяснениями будем отождествлять случайные события, свойственные данному эксперименту, с подмножествами Ω . Обозначим совокупность всевозможных подмножеств Ω через $\mathfrak{A}$. Множество $\mathfrak{A}$ представляет собой полный набор всевозможных мыслимых случайных событий, которые допускает заданное пространство Ω .

Из предыдущего изложения становится очевидной возможность выполнения операций над случайными событиями. Таких операций несколько, однако все они могут быть выражены через минимальный набор из двух операций: инвертирования и суммирования. Ясно, что множество $\mathfrak{A}$ нецелесообразно брать совершенно произвольным. Удобнее всего взять его таким, чтобы оно было замкнуто относительно двух указанных операций. Это естественным образом приводит к понятиям ал-

гебры событий и сигма-алгебры событий. Остановимся на них подробнее.

Определение 13.1.

Класс $\mathfrak{A}$ подмножеств множества Ω называется алгеброй при выполнении следующих условий:

- 1) $\Omega \in \mathfrak{A}$;
- 2) для любого события $A \in \mathfrak{A}$ противоположное событие $\bar{A} \in \mathfrak{A}$;
- 3) для любой пары событий $A \in \mathfrak{A}$ и $B \in \mathfrak{A}$ их объединение $(A \cup B) \in \mathfrak{A}$.

К этому определению следует сделать ряд замечаний. Во-первых, любая алгебра $\mathfrak{A}$ будет замкнутой не только по отношению к операции суммирования двух событий, но также и любого конечного числа событий. Пусть, например, имеется система событий $\{A_i\}_{i=1}^n$, причем $A_i \in \mathfrak{A}$ при всех $i = \overline{1, n}$. Положим

$$A_* = \bigcup_{i=1}^n A_i, \quad (13.16)$$

тогда $A_* \in \mathfrak{A}$.

Во-вторых, алгебра $\mathfrak{A}$ замкнута и по отношению к операции перемножения событий. Именно из того, что $A \in \mathfrak{A}$ и $B \in \mathfrak{A}$ следует, что $(A \cap B) \in \mathfrak{A}$. Это следует из формулы де Моргана (13.15) и того факта, что операции суммирования и инвертирования не выводят за пределы алгебры.

Не выводят за пределы алгебры и другие операции, принятые в теории случайных событий, например, операция вычитания, которая согласно (13.12) сводится к инвертированию и умножению (последнее можно свести к суммированию и инвертированию).

То же самое можно сказать по отношению к операции вычисления симметрической разности

$$A \triangle B = (A \setminus B) \cup (B \setminus A). \quad (13.17)$$

Симметрическая разность двух множеств A и B включает тех их элементы, которые принадлежат только какому-то одному из них (неважно, какому). Следовательно, $A \triangle B$ — событие,

состоящее в том, что A и B не произойдут одновременно (хотя они и совместны).

Далее рассмотрим некоторые примеры алгебр событий.

Пример 13.5. Доказать, что класс подмножеств Ω простейшего вида

$$\mathfrak{A}_\Omega = \{\Omega, \emptyset\} \quad (13.18)$$

является алгеброй.

Решение. Оба указанных множества по определению являются подмножествами Ω (их иногда называют *несобственными* его подмножествами). Далее согласно (13.13) оба элемента (13.18) противоположны друг другу, так что выполняется второе требование определения. Наконец, из (13.3) следует, что сумма элементов алгебры равняется Ω , то есть принадлежит ей же. Значит, $\mathfrak{A}_\Omega$ действительно является алгеброй, она называется *тривиальной алгеброй*. Это, так сказать, самая «бедная» из всех алгебр, которые можно построить на основе данного пространства Ω .

Пример 13.6. Пусть $A \subset \Omega$ — некоторое случайное событие. Доказать, что множество

$$\mathfrak{A}_A = \{\Omega, \emptyset, A, \overline{A}\} \quad (13.19)$$

образует алгебру.

Решение. Элементы $\mathfrak{A}_A$ разбиваются на две пары взаимно противоположных событий, причем, все они принадлежат $\mathfrak{A}_A$ по определению. Суммирование элементов (13.19) по формулам (13.3) и (13.11) не выводит за пределы $\mathfrak{A}_A$. Следовательно, класс $\mathfrak{A}_A$ действительно образует алгебру, ее называют *алгеброй, порожденной событием A* . Эта алгебра является более «богатой», чем ранее рассмотренная алгебра $\mathfrak{A}_\Omega$, определяемая (13.18), поскольку

$$\mathfrak{A}_\Omega \subset \mathfrak{A}_A. \quad (13.20)$$

Пример 13.7. Рассмотрим произвольную конечную полную в Ω систему событий $\{E_k\}_{k=1}^n$. Считается, что при любом $k = \overline{1, n}$ все события системы возможны, (то есть $E_k \neq \emptyset$) и,

разумеется, $E_k \subset \Omega$. Полнота системы означает, что

$$\bigcup_{k=1}^n E_k = \Omega; \quad E_k \cap E_j = \emptyset, \quad j \neq k. \quad (13.21)$$

Иначе говоря, E_k являются единственно возможными и несовместными исходами эксперимента. Образует из множеств E_k систему множеств путем объединения конечного числа этих множеств в произвольных их комбинациях с добавлением пустого множества

$$\mathfrak{A}_E^n = \{\emptyset; \bigcup_{s=1}^j E_{i_s}, j = \overline{1, n},$$

$$i_s = \overline{1, n}, i_s \neq i_r \text{ при } s \neq r\}. \quad (13.22)$$

Показать, что так построенный класс множеств является алгеброй при любом $n \geq 2$ и любом выборе полной системы базисных событий $\{E_k\}_{k=1}^n$.

Замечание. При $n = 2$ из условия (13.21) следует, что $E_2 = \overline{E_1}$, так что, если положить $E_1 = A$, $E_2 = \overline{A}$, то рассматриваемый пример сводится к предыдущему, то есть

$$\mathfrak{A}_E^2 = \mathfrak{A}_A. \quad (13.23)$$

Решение. Ясно, что противоположное по отношению к любому E_j событие сводится к сумме остальных событий системы

$$\overline{E_j} = \bigcup_{k=1, k \neq j}^n E_k, \quad (13.24)$$

которая по округлению входит в $\mathfrak{A}_E^n$. Далее, если просуммировать два произвольных события из $\mathfrak{A}_E^n$ вида $(E_{i_1} + \dots + E_{i_s})$ и $(E_{k_1} + \dots + E_{k_r})$ при любом выборе целых чисел s, r, i_α, k_β , то мы получим некоторую сумму базисных событий $\{E_k\}_{k=1}^n$, в которой отсутствуют те события, которых нет ни в первом, ни во втором слагаемом, а повторяющиеся в обоих слагаемых E_k засчитываются в сумме только один раз. Но все подобного рода суммы по определению включены в $\mathfrak{A}_E^n$. Значит, условия

определения 13.1 выполнены, и класс (13.22) действительно есть алгебра событий.

Пример 13.8 (Продолжение примера 13.7). Рассмотрим пространство элементарных событий в виде замкнутого интервала $a \leq x \leq b$ числовой оси

$$\Omega = \{x : a \leq x \leq b\}. \quad (13.25)$$

Разобьем Ω на систему непересекающихся n подынтервалов

$$\Omega_k = \begin{cases} \{x : \alpha_{k-1} \leq x < \alpha_k\}, & k = \overline{1, n-1}, \\ \{x : \alpha_{k-1} \leq x \leq \alpha_k\}, & k = n, \end{cases} \quad (13.26)$$

где $\alpha_0 = a$, $\alpha_n = b$, а остальные α_i выбираются произвольным образом из условия $\alpha_{i-1} < \alpha_i$, ($i = \overline{1, n}$). Далее определим множество $\mathfrak{A}_n$ как совокупность всевозможных объединений подмножеств вида (13.26), дополненную пустым множеством. Обладает ли $\mathfrak{A}_n$ свойствами алгебры?

Решение. Будем трактовать случайный эксперимент, как розыгрыш случайной точки, которая попадает в интервал (13.25). Это напоминает схему введения геометрических вероятностей, однако в отличие от последней, мы отказываемся от требования равновозможности попадания в равновеликие интервалы (см. § 3). Далее определим события E_k как попадание разыгрываемой точки в интервалы Ω_k , определяемые (13.26). Как было разъяснено в примере 13.7, получаемая система событий A_n является алгеброй.

Отметим, что подобным же образом строится алгебра событий для задач на геометрические вероятности, если размерность Ω больше единицы. В этом случае, как, например, в задаче о встрече (пример 13.5), прямоугольник, задающий Ω , разбивается на более мелкие непересекающиеся прямоугольники Ω_{kj} , всевозможные конечные суммы которых дополняются пустым множеством, после чего используются результаты примера 13.7.

Понятие алгебры событий вполне достаточно при рассмотрении дискретного (то есть конечного или счетного) множества Ω . В том случае, когда множество Ω несчетно, а это

встречается, например, в задачах на геометрические вероятности, приходится использовать более тонкие модели алгебры событий, а именно, так называемые сигма-алгебры.

Определение 13.2.

Класс $\mathfrak{A}$ подмножеств множества Ω называется сигма-алгеброй при выполнении следующих условий:

- 1) $\Omega \subset \mathfrak{A}$;
- 2) для любого события $A \subset \mathfrak{A}$ противоположное ему $\overline{A} \subset \mathfrak{A}$;
- 3) для любой бесконечной системы событий $A_i \subset \mathfrak{A}$, ($i = \overline{1, \infty}$) имеем

$$\bigcup_{i=1}^{\infty} A_i \subset \mathfrak{A}. \quad (13.27)$$

Определение 13.2 отличается от определения 13.1 только тем, что требование замкнутости $\mathfrak{A}$ относительно суммы *конечного* числа слагаемых заменено на аналогичное требование, но уже для *бесконечной* суммы слагаемых. Поэтому ясно, что сигма-алгебра всегда является алгеброй, но обратное, вообще говоря, неверно.

Пример 13.9. Проверим, будет ли множество $\mathfrak{A}_{\max}$ всевозможных подмножеств множества Ω являться сигма-алгеброй.

Решение. Мы условились отождествлять событие A с множеством Ω_A тех элементарных событий $\omega \in \Omega$, которые благоприятствуют A (см. (13.1)). Элементарные исходы $\omega \in \Omega$, не благоприятствующие A , будучи элементами Ω , также образуют некоторое событие, названное нами противоположным, обозначаемое $\overline{A}$ и имеющее отражение $\Omega_{\overline{A}}$ в Ω . Поскольку в условиях примера говорится о *всевозможных* подмножествах Ω , то $\Omega_{\overline{A}}$ просто обязано войти в $\mathfrak{A}_{\max}$, значит, второе условие определения выполняется, также, как и первое, которое очевидно.

Некоторые сомнения остаются только насчет третьего условия. Введем обозначение для события

$$A_* = \bigcup_{i=1}^{\infty} A_i. \quad (13.28)$$

Будем рассуждать от противного. Допустим, что $A_* \not\subset \mathfrak{A}_{\max}$. Это значит, что в A_* имеется хотя бы один элемент, скажем, ω_* , который не входит в Ω . Но тогда среди событий A_i , образующих бесконечную сумму (13.28), должно найтись хотя бы одно событие A_{i_0} такое, что $\omega_* \in A_{i_0}$. Но все A_i являются подмножествами Ω , и такое попросту невозможно. Значит, $A_* \subset \mathfrak{A}_{\max}$, а последняя является сигма-алгеброй.

Таким образом, класс подмножеств Ω

$$\mathfrak{A}_{\max} = \{A : A \subset \Omega\} \quad (13.29)$$

всегда является сигма-алгеброй. Это наиболее «богатая» из всех сигма-алгебр. Все остальные сигма-алгебры занимают промежуточное положение между $\mathfrak{A}_{\max}$ и алгеброй (13.18) в том смысле, что для любой из них $\mathfrak{A}$ справедливо двойное включение

$$\mathfrak{A}_{\Omega} \subset \mathfrak{A} \subset \mathfrak{A}_{\max}. \quad (13.30)$$

По этой причине сигма-алгебру, составленную из всевозможных подмножеств Ω , часто называют *максимальной сигма-алгеброй*.

Число элементов максимальной сигма-алгебры значительно превосходит число элементов самого пространства Ω , по которому она построена. Пусть, например, N_{Ω} конечно и равняется n . Ясно, что из n элементов Ω можно составить C_n^k подмножеств по k элементов. Значит, общее число подмножеств $A \subset \Omega$, попавших в $\mathfrak{A}_{\max}$, будет таким

$$N_{\mathfrak{A}_{\max}} = \sum_{k=0}^n C_n^k = 2^n. \quad (13.31)$$

Если пространство Ω бесконечно, но счетно, то максимальная сигма-алгебра $\mathfrak{A}_{\max}$ имеет хотя и достаточно сложный, но все-таки обозримый вид. Ситуация сильно осложняется в случае несчетного множества Ω , как, например, в примерах 13.4 или 13.8. И дело здесь не просто в большом числе элементов $\mathfrak{A}_{\max}$, сама структура этого множества оказывается достаточно сложной, и на нем трудно ввести согласованную с ним меру, как того требует аксиоматика Колмогорова (см. далее).

Вообще, вопрос о сравнении численности бесконечных множеств с чисто теоретической точки зрения является достаточно сложным. В теории множеств он решается с помощью введения понятия мощности или кардинального числа множества [2]. Говорят, что два множества имеют одинаковую мощность, если между их элементами можно установить взаимно однозначное соответствие. В этом смысле конечные множества равномощны, если они имеют равное число элементов, а более мощным является то, где их больше.

В случае бесконечных множеств A и B второе считается более мощным, чем первое, если все множество A эквивалентно некоторой части B , тогда как B не эквивалентно никакой части A . В теории множеств одной из центральных является теорема Кантора², согласно которой мощность множества всех подмножеств любого непустого множества A всегда больше мощности самого множества A [2].

В связи с отмеченными принципиальными трудностями в качестве математической модели для множества случайных событий для случайного эксперимента на практике часто используют не максимальную сигма-алгебру $\mathfrak{A}_{\max}$, а какую-либо более «бедную» сигма-алгебру $\mathfrak{A} \subset \mathfrak{A}_{\max}$. Это можно сделать несколькими способами.

Наиболее простой и очевидный состоит в «дискретизации» множества Ω . Допустим, что в Ω можно построить *бесконечную* полную систему событий $\{E_k\}_{k=1}^{\infty}$. Тогда аналогично примеру 13.7 алгебра, составленная из всевозможных объединений событий E_k с добавлением пустого множества

$$\mathfrak{A}_E^{\infty} = \{\emptyset; \bigcup_{s=1}^j E_{i_s}, j = \overline{1, \infty}, \\ i_s = \overline{1, \infty}, i_s \neq i_r \text{ при } s \neq r\}, \quad (13.32)$$

будет являться сигма-алгеброй. При этом приходится довольствоваться менее детальным описанием происходящих событий. Например, все события A , включенные в E_k ($A \subset E_k$), приближенно заменяются на E_k .

² Кантор Георг (1845–1918) — немецкий математик, творец теории множеств.

Другой способ состоит в том, что с самого начала задают интересующий нас класс случайных событий $K \subset \mathfrak{A}_{\max}$. Если этот класс K сам обладает свойствами сигма-алгебры, то, естественно, полагают $\mathfrak{A} = K$. Чаще всего K нужными свойствами не обладает. Тогда строят *минимальную сигма-алгебру, содержащую класс K* , обозначаемую через $\mathfrak{A}_{\min}^K$.

Минимальной сигма-алгеброй, содержащей класс K подмножеств пространства Ω , называется такая сигма-алгебра $\mathfrak{A}_{\min}^K$, которая, во-первых, содержит в себе класс K , то есть $K \subset \mathfrak{A}_{\min}^K$ и, во-вторых, для любой другой сигма-алгебры $\mathfrak{A}$, содержащей K , оказывается, что минимальная сигма-алгебра сама содержится в ней, то есть $\mathfrak{A}_{\min}^K \subset \mathfrak{A}$. В более подробных аксиоматических курсах теории вероятностей (см., например, [8]) доказывается, что для любого класса подмножеств K из Ω минимальная сигма-алгебра $\mathfrak{A}_{\min}^K$ существует.

Нужно сказать, что в прикладных задачах обычно не возникает сложностей ни с формированием пространства элементарных событий, ни с построением нужной сигма-алгебры событий. Строгое теоретико-множественное определение этих понятий необходимо для приведения вероятностных выводов и рассуждений в соответствие с требованиями современной математики, оперирующей абстрактными и логически безупречными понятиями и конструкциями. Материал настоящей главы можно рассматривать как первоначальное, и притом достаточно упрощенное введение в этот круг проблем, рассчитанное на читателя, не являющегося профессиональным математиком и заинтересованного прежде всего в решении практических задач.

Итак, предположим, что пространство элементарных событий Ω рассматриваемого случайного эксперимента задано и на нем построена удовлетворяющая нас по своим свойствам сигма-алгебра $\mathfrak{A}$. Последний, третий шаг в построении вероятностной модели эксперимента заключается в задании на $\mathfrak{A}$ вероятностной меры.

Мерой множества называется числовая функция этого множества являющаяся естественным обобщением таких понятий, как длина отрезка, площадь плоской фигуры, объем тела. Если обратиться к физическим аналогиям, то мера

аналогична массе некоторого тела при определенном законе распределения этой массы внутри тела.

В теории меры множество, которое требуется измерить, должно быть определенным образом структурировано. Требования к этому множеству, принятые в общей теории меры [29], сводятся к тому, чтобы оно являлось некоторой алгеброй или сигма-алгеброй. При этом различают конечно-аддитивные и счетно-аддитивные меры. Дадим соответствующие определения.

Определение 13.3. Пусть $\mathfrak{A}$ является некоторой алгеброй подмножеств из Ω . Числовая функция $\mathbf{P}(A)$, определенная для $A \subset \mathfrak{A}$, называется конечно-аддитивной вероятностной мерой на $\mathfrak{A}$ при соблюдении условий:

- 1) $\mathbf{P}(A) \geq 0$ для всех $A \subset \mathfrak{A}$;
- 2) $\mathbf{P}(\Omega) = 1$;
- 3) для любых непересекающихся подмножеств $A \subset \mathfrak{A}$ и $B \subset \mathfrak{A}$ имеем

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B). \quad (13.33)$$

Отметим, что при соблюдении второго условия мера называется *нормированной*. Третье условие с помощью математической индукции легко может быть обобщено на произвольное конечное число непересекающихся подмножеств $A_1, \dots, A_n$ пространства Ω

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbf{P}(A_i). \quad (13.34)$$

Если в последнем равенстве заменить конечное n на $n = \infty$, сохранив два остальных условия, то приходим к определению *счетно-аддитивной* вероятностной меры.

Определение 13.4. Рассмотрим сигма-алгебру $\mathfrak{A}$ подмножеств пространства Ω . Числовая функция $\mathbf{P}(A)$, определенная на $\mathfrak{A}$, называется счетно-аддитивной вероятностной мерой на $\mathfrak{A}$, если выполнены условия:

- 1) $\mathbf{P}(A) \geq 0$ для любого подмножества $A \subset \mathfrak{A}$;
- 2) $\mathbf{P}(\Omega) = 1$;

3) для любой системы непересекающихся множеств $\{A_i\}_{i=1}^{\infty}$, таких, что $A_i \subset \mathfrak{A}$ при всех $i = \overline{1, \infty}$

$$\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i). \quad (13.35)$$

Число $\mathbf{P}(A)$, приписываемое каждому подмножеству A пространства Ω , то есть в нашей трактовке — случайному событию A , называется *вероятностью* этого события.

Отметим, что в случае *дискретного* пространства элементарных событий (когда множество Ω конечно или счетно) достаточно пользоваться определением 13.3. Если же Ω *несчетно*, то приходится применять более общее определение 13.4. Ясно, что из счетной аддитивности вероятностной меры следует ее конечная аддитивность, но обратное, вообще говоря, неверно.

В заключение данного параграфа заметим, что тройка $(\Omega, \mathfrak{A}, P)$, состоящая из некоторого пространства элементарных событий Ω , сигма-алгебры (или алгебры) событий $\mathfrak{A}$, трактуемых, как подмножества Ω , а также счетно-аддитивной (или конечно-аддитивной) меры P , заданной на элементах $\mathfrak{A}$, называется *вероятностным пространством* и дает полную математическую модель рассматриваемого случайного эксперимента. Построить вероятностное пространство данного эксперимента — это значит сформировать все три перечисленных элемента вероятностной модели.

§ 14. КЛАССИЧЕСКИЕ МОДЕЛИ ВЕРОЯТНОСТНОГО ПРОСТРАНСТВА

Убедимся, что введенное определение вероятности не противоречит ранее введенным ее определениям: классическому (§ 2) и геометрическому (§ 3). Рассмотрим вначале случай конечного пространства элементарных событий Ω . Допустим, что оно состоит из n элементов ω_i , ($i = \overline{1, n}$). Как отмечалось в § 12 без ущерба для общности это пространство можно параметризовать таким образом, чтобы принять $\omega_i = i$, то есть $\Omega = \{i\}_{i=1}^n$ представляет собой набор n младших натуральных

чисел. Обозначим элементарное событие, состоящее в реализации на опыте исхода $\omega_i = i$ через E_i . Ясно, что система событий $\{E_i\}_{i=1}^n$ является полной. Всевозможные случайные события A будут отражаться в множестве Ω как некоторые совокупности целых чисел $\{i_1, \dots, i_s\}$, где $0 \leq s \leq n$, причем при $s = 0$ по определению имеем пустое множество, а при $s = n$ получаем полный набор натуральных чисел от 1 до n . При этом целые числа i_k не превосходят n и не совпадают при попарно различных k . Множество всех таких наборов чисел $\{i_k\}_{k=1}^s$ при всевозможных допустимых s и указанным образом выбранных i_k образует алгебру событий $\mathfrak{A}$.

Введем на этой алгебре вероятностную меру следующим образом. Припишем каждому элементарному исходу i некоторый вес p_i таким образом, чтобы все эти веса, во-первых, были неотрицательны

$$0 \leq p_i \leq 1, \quad (i = \overline{1, n}) \quad (14.1)$$

и, во-вторых, удовлетворяли *условию нормировки*

$$\sum_{i=1}^n p_i = 1. \quad (14.2)$$

Далее для каждого события $A \subset \mathfrak{A}$ определим вероятностную меру $\mathbf{P}(A)$ как сумму весов элементарных событий E_i , благоприятствующих этому событию

$$\mathbf{P}(A) = \sum_{i \in A} p_i. \quad (14.3)$$

Так введенная вероятность, разумеется, удовлетворяет первому условию определения 13.3. Кроме того при $A = \Omega$ отражением этого события является все множество $\{i\}_{i=1}^n$, так что сумма (14.3) переходит в (14.2), и выполняется условие 2 определения. Наконец, несовместность событий A и B означает, что отражающие их наборы $\{i_k\}$ не пересекаются, а тогда выполняется третье условие определения, записываемое в виде (13.33).

Таким образом, все условия определения 13.3 выполнены и так построенная мера (14.3) является аддитивной нормированной вероятностной мерой. С математической точки зрения мера (14.3) при различных способах задания чисел p_i позволяет охватить все мыслимые конечные вероятностные пространства. Нам хорошо известен один пример таких задач, решаемый с помощью классического определения вероятности § 2.

Допустим, что все элементарные исходы E_i в силу симметрии эксперимента равновозможны. Тогда

$$p_i = \frac{1}{n}, \quad (i = \overline{1, n}). \quad (14.4)$$

Далее обозначим число элементарных исходов, входящих в множество A , то есть благоприятствующих соответствующему событию, через m , тогда по формуле (14.3) находим

$$\mathbf{P}(A) = \frac{m}{n}, \quad (14.5)$$

что полностью совпадает с классическим определением вероятности.

Вопрос о «правильности» назначаемых в том или ином эксперименте вероятностей p_i , входящих в (14.3), мы оставляем в стороне. Этот вопрос вообще не является вопросом *математическим* и должен проверяться на практике в ходе соответствующего *физического* эксперимента. Такой эксперимент может быть основан, например, на *предельных теоремах* (типа закона больших чисел), согласно которым в длинных сериях испытаний частоты должны быть близки к соответствующим вероятностям.

Теперь рассмотрим случайный эксперимент, пространство элементарных событий которого представляет собой некоторую область в многомерном евклидовом пространстве соответствующей размерности. К этому типу относятся все задачи на геометрические вероятности, которые сводятся к случайному эксперименту, заключающемуся в выборе наудачу точки из заданной области. Примем эту область за Ω , а всевозможные ее подобласти A будем трактовать как случайные события.

Допустим, что сама область Ω имеет меру Лебега $m(\Omega)$ и будем рассматривать только такие подмножества $A \subset \Omega$, которые также имеют меру Лебега $m(A)$. Напомним, что мера Лебега в случае двумерного евклидова пространства сводится к площади плоской фигуры, а в случае трехмерных тел — к их объему.

Как известно из теории меры [29], класс подмножеств, имеющих меру Лебега и принадлежащих измеримому по Лебегу множеству, образует сигма-алгебру. Будем считать, что в силу симметрии эксперимента, вероятность любого события $A \subset \mathfrak{A}$ пропорциональна лебеговой мере соответствующего множества и не зависит от формы A , то есть примем

$$\mathbf{P}(A) = C \cdot m(A), \quad (14.6)$$

где C — некоторая универсальная постоянная, неизменная для всех событий в данном эксперименте.

Из условия нормировки, записывая (14.6) для всего пространства Ω , находим

$$\mathbf{P}(\Omega) = C \cdot m(\Omega), \quad (14.7)$$

что дает

$$C = \frac{1}{m(\Omega)}, \quad (14.8)$$

а тогда для любого события A получаем

$$\mathbf{P}(A) = \frac{m(A)}{m(\Omega)}, \quad (14.9)$$

что соответствует обычному геометрическому определению вероятностей.

Первое и второе свойства вероятности $\mathbf{P}(A)$ из определения (13.4) вытекают из самого построения меры (14.9), а последнее вытекает из свойств меры Лебега, как счетно-аддитивной функции множества.

Каких-либо отдельных примеров к материалу данного параграфа не требуется, так как весь он представляет разбор двух примеров приложения общей теории, изложенной в § 12 и § 13.

§ 15. АКСИОМАТИЧЕСКИЙ ВЫВОД ОСНОВНЫХ ПОЛОЖЕНИЙ ТЕОРИИ СЛУЧАЙНЫХ СОБЫТИЙ

В предыдущих параграфах настоящей главы были достаточно подробно разобраны все элементы вероятностного пространства $(\Omega, \mathfrak{A}, P)$, представляющего общую модель случайного эксперимента. В самом начале курса, в первой главе, излагались классические его модели, которые, как оказалось, являются частным случаем, и притом одним из простейших частных случаев, общей универсальной модели. Пришло время сформулировать сами аксиомы и формально-логически вывести из них основные правила исчисления вероятностей, установленные первоначально при классической трактовке случайного эксперимента.

Напомним, что пространство элементарных событий Ω представляет произвольное множество элементов, множество событий, происходящих в эксперименте, моделируется сигма-алгеброй $\mathfrak{A}$ (не обязательно максимальной), а P является вероятностной, то есть нормированной и счетно-аддитивной мерой, заданной на элементах сигма-алгебры $\mathfrak{A}$. Приводимые ниже аксиомы являются необходимым следствием ранее изложенной элементарной теории случайных событий, а их достаточность для воссоздания содержания этой теории средствами аксиоматики показывается в настоящем параграфе.

Аксиома I (аксиома неотрицательности)

Всякому случайному событию $A \in \mathfrak{A}$ ставится в соответствие неотрицательное число

$$P(A) \geq 0, \quad (15.1)$$

называемое его вероятностью.

Аксиома II (аксиома нормировки)

Событие $A = \Omega \in \mathfrak{A}$ называется достоверным. Вероятность достоверного события

$$P(\Omega) = 1. \quad (15.2)$$

Аксиома III (аксиома сложения)

Для всякой несовместной системы событий $\{A_i\}_{i=1}^n \subset \mathfrak{A}$

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbf{P}(A_i). \quad (15.3)$$

Перечисленная система аксиом действительна для случая дискретного пространства Ω . Если же оно не является счетным, то последняя из аксиом усиливается дополнительным условием, чтобы (15.3) выполнялось также и при $n \rightarrow \infty$.

Аксиома IV (расширенная аксиома сложения)

Для всякой бесконечной несовместной системы событий $\{A_i\}_{i=1}^{\infty} \subset \mathfrak{A}$

$$\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i). \quad (15.4)$$

Используя приведенный набор аксиом нетрудно чисто аналитически с помощью стандартных алгебраических вычислений вывести уже хорошо известные нам из элементарной теории свойства случайных событий.

Начнем с получения выражения для вероятности противоположного события $\mathbf{P}(\bar{A})$. Рассмотрим систему двух событий $(A, \bar{A})$. Они несовместны ($A \cap \bar{A} = \emptyset$), поэтому обратимся к аксиоме III. По определению A и $\bar{A}$ взаимно дополняют друг друга до достоверного события ($A \cup \bar{A} = \Omega$), поэтому по формуле (15.3)

$$\mathbf{P}(A) + \mathbf{P}(\bar{A}) = \mathbf{P}(\Omega). \quad (15.5)$$

Применяя аксиому II и разрешая полученное равенство относительно $\mathbf{P}(\bar{A})$, находим

$$\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A), \quad (15.6)$$

как это и принято в элементарной теории.

Положим в равенстве (15.6) в качестве рассматриваемого события $A = \Omega$. Поскольку $\bar{\Omega} = \emptyset$, то на основании аксиомы II убеждаемся, что невозможное событие $\emptyset$ характеризуется нулевой вероятностью

$$\mathbf{P}(\emptyset) = 0. \quad (15.7)$$

Рассмотрим два случайных события A и B , из которых событие A влечет наступление события B , то есть $A \subset B$.

Ясно, что в этом случае B можно представить в виде

$$B = A \cup (B \setminus A), \quad (15.8)$$

причем событие

$$B_* = B \setminus A \quad (15.9)$$

несовместно с A

$$B_* \cap A = \emptyset \quad (15.10)$$

и, значит, к паре событий A и B_* применима аксиома III, откуда следует

$$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B \setminus A). \quad (15.11)$$

Это позволяет выразить вероятность разности вложенных друг в друга событий в виде

$$\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A), \quad (A \subset B). \quad (15.12)$$

С помощью последней формулы можно, во-первых, сравнивать числовое значение вероятности для включенных одно в другое событий

$$\mathbf{P}(A) \leq \mathbf{P}(B), \quad (A \subset B), \quad (15.13)$$

и, во-вторых, оценить максимальное значение вероятности для любого события

$$\mathbf{P}(A) \leq 1, \quad (15.14)$$

если положить в (15.13) $B = \Omega$ и учесть аксиому II.

Перейдем к выводу теоремы сложения вероятностей. Будем исходить из следующих двух тождеств

$$A \cup B = A + (B \setminus A), \quad B = (B \setminus A) + A \cap B, \quad (15.15)$$

в каждом из которых оба слагаемых в правой части несовместны. Используя аксиому сложения, находим

$$\mathbf{P}(A \cup B) = \mathbf{P}(B) + \mathbf{P}(B \setminus A), \quad (15.16)$$

$$\mathbf{P}(B) = \mathbf{P}(B \setminus A) + \mathbf{P}(A \cap B). \quad (15.17)$$

Исключая из этих равенств вероятность разности B и A , окончательно находим

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B). \quad (15.18)$$

Это утверждение представляет уже хорошо известную нам *теорему сложения вероятностей* для двух событий.

Из полученной формулы автоматически следует теорема сложения вероятностей для произвольного числа событий

$$\begin{aligned} \mathbf{P}\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n \mathbf{P}(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} \mathbf{P}(A_{i_1} \cap A_{i_2}) + \\ &+ \sum_{1 \leq i_1 < i_2 < i_3 \leq n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots \\ &\dots + (-1)^n \mathbf{P}(A_1 \cap \dots \cap A_n). \end{aligned} \quad (15.19)$$

Дело в том, что формула (15.19) выводится из (15.18) чисто алгебраически с помощью метода математической индукции. Способ, которым была получена исходная формула (15.18) при этом никакой роли не играет (см. § 5).

Теперь обратимся к *теореме умножения вероятностей*. Здесь приходится давать особое определение для условной вероятности. Рассмотрим систему двух событий (A, B) , заданных на одном и том же вероятностном пространстве $(\Omega, \mathfrak{A}, P)$, то есть относящихся к одному и тому же случайному эксперименту. Допустим, что событие A возможно, так что $\mathbf{P}(A) \neq 0$. Тогда условной вероятностью события B при условии, что A произошло, называют величину

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}. \quad (15.20)$$

При таком определении теорема умножения вероятностей уже не требует доказательства и автоматически следует из (15.20). Необходимость определения (15.20) возникает потому, что при аксиоматическом построении теории мы уже не можем апеллировать к опыту, и, поскольку понятие условной вероятности никак не следует из предыдущих аксиом, его приходится определять особо.

Здесь, впрочем, возникает одна терминологическая тонкость: говоря об отношении (15.20) как о «вероятности» (пусть даже и условной), необходимо убедиться, что так введенная новая величина $\mathbf{P}(B|A)$ действительно обладает свойствами вероятности. Речь идет о свойстве

$$0 \leq \mathbf{P}(B|A) \leq 1, \quad (15.21)$$

аналогичном (15.14) свойстве нормировки

$$\mathbf{P}(\Omega|A) = 1, \quad (15.22)$$

аналогичном (15.2), а также выполнении для любых двух событий A и B , несовместных в данном эксперименте, равенства

$$\mathbf{P}((A \cup B)|C) = \mathbf{P}(A|C) + \mathbf{P}(B|C), \quad (15.23)$$

где, естественно, считается, что $\mathbf{P}(C) \neq 0$, то есть $C \neq \emptyset$. Конечно, (15.23) вполне аналогично (15.3).

Равенство (15.22) является тривиальным следствием определения (15.20), поскольку в нашей алгебре событий

$$\mathbf{P}(A \cap \Omega) = \mathbf{P}(A). \quad (15.24)$$

Для получения (15.23) вначале записываем (15.3) для произвольной системы несовместных событий (A, B)

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B), \quad (15.25)$$

после чего делим обе части этого равенства на $\mathbf{P}(C) \neq 0$ и применяем к каждому слагаемому определение (15.20).

Несколько сложнее доказывается неравенство (15.21), точнее говоря, его правая часть (левая, очевидно, выполняется, так как $\mathbf{P}(B|A)$ выражается в виде отношения неотрицательных чисел).

Из определения понятия произведения событий как пересечения множеств следует включение

$$(A \cap B) \subset A, \quad (15.26)$$

но тогда по свойству (15.13), заменяя в нем A на $A \cap B$, а B на A получаем

$$\mathbf{P}(A \cap B) \leq \mathbf{P}(A), \quad (15.27)$$

откуда следует неравенство

$$\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \leq 1, \quad (15.28)$$

которое с учетом определения условной вероятности (15.20) эквивалентно правому неравенству (15.21).

Таким образом, введенная чисто формально с помощью определения (15.20) величина $\mathbf{P}(B|A)$ обладает всеми свойствами вероятности и ее можно трактовать как вероятность некоторого события. Далее можно ввести понятие *независимости событий* совершенно аналогично тому, как это было сделано в § 4.

Напомним, что в элементарной теории случайных событий важную роль играли также формула полной вероятности (см. § 6) и формула Байеса (см. § 7). Вывод этих формул был чисто алгебраическим и основывался на применении теорем сложения вероятностей (15.18) и умножения вероятностей

$$\mathbf{P}(A \cap B) = \mathbf{P}(B|A) \cdot \mathbf{P}(A), \quad (15.29)$$

которая в аксиоматической теории является простым следствием определения (15.20). Конечно, все формулы § 6 и § 7 в полном объеме сохраняют силу и при теоретико-множественном определении вероятности.

Итак, приведенные выше аксиомы оказываются непротиворечивыми и при этом обладают достаточной полнотой, чтобы получить на их основе все найденные ранее формулы исчисления вероятностей случайных событий. Если при этом трактовать вероятность события как некоторую нормированную счетно-аддитивную функцию, заданную на сигма-алгебре случайных событий, трактуемых как подмножества соответствующего пространства элементарных событий, то такая аксиоматика позволяет охватить с единой точки зрения широкий класс случайных экспериментов. Он включает не только

случайные события, но также случайные величины, случайные функции, а также ряд других новых объектов изучения теории вероятностей, постоянно появляющихся в прикладных научных и инженерных задачах.

Отметим, что стандартные задачи теории случайных событий, которые ранее решались с использованием наглядных представлений, можно рассматривать и решать чисто алгебраически, основываясь на аксиоматических свойствах вероятности.

Пример 15.1. Доказать, что для любого возможного события A ($P(A) \neq 0$), рассматриваемого в качестве условия для произвольного события B , выполняется неравенство

$$P(B|A) \geq 1 - \frac{P(\bar{B})}{P(A)}.$$

Решение. Рассмотрим сумму двух указанных событий $A \cup B$. Используя теорему сложения (15.18) и определение условной вероятности (15.20), переписанное в форме теоремы умножения (15.29), на основании свойства (15.14) получаем

$$P(A) + P(B) - P(B|A)P(A) \leq 1.$$

Далее учитывая свойство (15.6), записываем

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

что дает

$$P(A) - P(\bar{B}) \leq P(B|A)P(A)$$

и окончательно после деления обеих частей неравенства на $P(A)$ получаем требуемое.

Пример 15.2. Имеется два независимых события A и B . Рассмотрим симметрическую разность этих событий

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

Выразить вероятность симметрической разности через вероятности вычитаемых событий.

Решение. Из определения понятия разности событий как дополнения множеств следует

$$A \setminus B = A \cap \overline{B}, \quad B \setminus A = B \cap \overline{A}.$$

Тогда, учитывая несовместность $A \setminus B$ и $B \setminus A$ по аксиоме сложения, имеем

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

По теореме умножения в форме (15.29)

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

В силу независимости A от B (а значит, и от $\overline{B}$)

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

По свойству (15.6)

$$\begin{aligned} \mathbf{P}(A \triangle B) &= \mathbf{P}(A)(1 - \mathbf{P}(B)) + \mathbf{P}(B)(1 - \mathbf{P}(A)) = \\ &= \mathbf{P}(A) + \mathbf{P}(B) - 2\mathbf{P}(A)\mathbf{P}(B). \end{aligned}$$

Интересно заметить, что, если $\mathbf{P}(A) = \mathbf{P}(B) = p$, то тогда

$$\mathbf{P}(A \triangle B) = 2p - 2p^2 = 2p(1 - p).$$

Поскольку $0 \leq p \leq 1$, так, что

$$p(1 - p) \leq \frac{1}{4},$$

то для равновероятных независимых событий

$$\mathbf{P}(A \triangle B) \leq \frac{1}{2},$$

причем знак равенства достигается только в том случае, когда

$$\mathbf{P}(A) = \mathbf{P}(B) = \frac{1}{2}.$$

Пример 15.3. Доказать неравенство

$$\mathbf{P}(A \cap B) \geq 1 - \mathbf{P}(\bar{A}) - \mathbf{P}(\bar{B}),$$

где A и B произвольные события.

Решение. По формуле де Моргана

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

поэтому

$$\mathbf{P}(A \cap B) = 1 - \mathbf{P}(\bar{A} + \bar{B}) = 1 - \mathbf{P}(\bar{A}) - \mathbf{P}(\bar{B}) + \mathbf{P}(\bar{A} \cap \bar{B}).$$

Поскольку

$$\mathbf{P}(\bar{A} \cap \bar{B}) \geq 0,$$

отсюда получаем требуемое.

Пример 15.4. Пусть события A и B независимы. Доказать, что тогда A и $\bar{B}$ также независимы.

Решение. Рассмотрим пересечение сравниваемых событий $\mathbf{P}(A \cap \bar{B})$. Выразим его через разность событий A и $A \cap B$ так:

$$\mathbf{P}(A \cap \bar{B}) = \mathbf{P}(A \setminus (A \cap B)).$$

Поскольку $A \cap B$ включено в A , то по свойству (15.13) имеем

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

что дает в силу независимости A и B

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

а это уже означает, что A и $\bar{B}$ независимы.

Пример 15.5. Имеются два независимых события A и B , причем оба этих события не являются невозможными ($\mathbf{P}(A) \neq 0$, $\mathbf{P}(B) \neq 0$). Могут ли эти события быть несовместными?

Решение. Независимость A и B означает, что

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B),$$

то есть условные вероятности появления событий совпадают с безусловными

$$\mathbf{P}(A|B) = \mathbf{P}(A) \neq 0, \quad \mathbf{P}(B|A) = \mathbf{P}(B) \neq 0.$$

Вместе с тем несовместность означает, что

$$\mathbf{P}(A \cap B) = 0,$$

откуда следует, что

$$\mathbf{P}(A|B) = \mathbf{P}(B|A) = 0.$$

Полученное противоречие доказывает, что несовместные, но возможные события обязательно являются зависимыми, а независимые события всегда совместны.

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

§ 16. РЯД РАСПРЕДЕЛЕНИЯ И ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ДИСКРЕТНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Случайной величиной называется такая величина, которая в результате испытания может принимать различные значения в зависимости от случайного исхода испытания. В теории вероятностей рассматриваются только такие случайные величины, которые обладают *вероятностной мерой*, то есть такие, для которых можно указать вероятность попадания значения случайной величины в любой заданный интервал.

Условимся обозначать случайные величины большими буквами из конца латинского алфавита: $X, Y, Z, \dots$, снабжая их по мере необходимости индексами. Малые буквы того же наименования будем использовать для обозначения реализаций случайных величин, то есть их значений, получаемых в результате опыта ($x, y, z, \dots$). Из этого правила мы будем делать исключения только в тех случаях, когда будут употребляться буквы греческого алфавита — в этом случае и случайные величины и их реализации будут обозначаться одной буквой.

Если случайная величина X может принимать (с отличной от нуля вероятностью) только заданные дискретные значения $x_1, x_2, x_3, \dots$, то такая величина называется *случайной величиной дискретного типа* или, короче, *дискретной случайной величиной*.

Если же случайная величина может принимать любые значения в некотором интервале (или в нескольких таких интервалах), причем вероятность попадания значения случайной

величины в любой элементарный интервал $(x, x + dx)$ пропорциональна его длине dx , то такая величина называется *случайной величиной непрерывного типа* или *непрерывной случайной величиной*.

Наконец, случайные величины, которые с отличными от нуля вероятностями могут принимать заданные дискретные значения и, одновременно с этим, с вероятностями, пропорциональными dx , могут принимать также значения в любом элементарном интервале длиной dx , лежащем в заданном промежутке на числовой оси, называются *случайными величинами смешанного типа*.

Рассмотрим вначале более подробно *дискретные случайные величины*. Для определения дискретной случайной величины достаточно ввести полную группу случайных событий $A_1, A_2, \dots, A_n$ и сопоставить каждому из этих событий A_j некоторое число x_j . Введем для этой цели *индикатор случайного события A* или *характеристическую случайную величину $I(A)$* , определяемую равенством

$$I(A) = \begin{cases} 1, & \text{если при испытании имело место} \\ & \text{событие } A, \\ 0, & \text{если при испытании имело место} \\ & \text{противоположное событие } \bar{A}. \end{cases} \quad (16.1)$$

Тогда всякую дискретную случайную величину X можно будет определить равенством

$$X = \sum_{j=1}^n x_j I(A_j). \quad (16.2)$$

Возможными значениями случайной величины X при этом будут числа $x_1, x_2, \dots, x_n$, которые она принимает с вероятностями

$$p_j = \mathbf{P}(A_j), \quad (j = \overline{1, n}). \quad (16.3)$$

Число возможных значений n случайной величины X может быть и бесконечным, однако при любом значении n , как конечном, так и бесконечном, вследствие того, что случайные

события A_j составляют полную группу несовместных событий, должно выполняться условие нормировки

$$\sum_{j=1}^n p_j = 1. \quad (16.4)$$

Итак, исчерпывающей характеристикой дискретной случайной величины является ее *ряд распределения*, то есть совокупность всех возможных ее значений x_j и соответствующих им вероятностей p_j . Однако более удобно характеризовать дискретную случайную величину не рядом распределения, а *функцией распределения*¹, определяемой равенством

$$F(x) = \mathbf{P}\{X < x\}, \quad (16.5)$$

где x — любое вещественное число. Это определение не предполагает дискретность возможных значений случайной величины и, следовательно, может быть применено для случайных величин любого типа. Как мы увидим дальше, его используют для произвольных случайных величин.

Покажем, что для дискретной случайной величины ряд распределения позволяет найти функцию распределения, а зная функцию распределения, можно однозначно восстановить ряд распределения, то есть покажем, что оба этих способа задания дискретной случайной величины взаимно эквивалентны. Предположим для простоты, что возможные значения случайной величины x_j пронумерованы в порядке их возрастания, так что $x_j < x_{j+1}$ для всех $j = \overline{1, n-1}$. Выберем аргумент функции распределения x в интервале $[x_j, x_{j+1})$. В этом случае неравенство $X < x$ может быть выполнено только при осуществлении одного из j несовместных событий $X = x_k$, ($k = \overline{1, j}$). Следовательно, применяя формулу сложения вероятностей для несовместных событий, получим

$$F(x) = \sum_{k=1}^j p_k, \quad (x_j < x \leq x_{j+1}). \quad (16.6)$$

¹ Иногда функцию распределения называют *интегральным законом распределения*.

Функция распределения $F(x)$ для дискретной случайной величины является разрывной ступенчатой функцией, скачки которой имеют место в точках x_k , ($k = \overline{1, n}$), соответствующих возможным значениям случайной величины X . Величина скачка в точке $x = x_k$ равна вероятности p_k ($k = \overline{1, n}$). Сумма всех скачков равна единице (рис. 16.1). Таким образом, если задан ряд распределения, то есть известны возможные значения x_k ($k = \overline{1, n}$) случайной величины X и вероятности p_k этих значений, то можно вычислить значение функции распределения $F(x)$ при любом x . Зная функцию распределения $F(x)$, по абсциссам точек разрывов и по величинам скачков можно восстановить и сам ряд распределения. Поэтому свойства дискретной случайной величины одинаково полно определяются как рядом распределения, так и функцией распределения.

В том случае, когда приходится иметь дело с несколькими случайными величинами, условно принимается, что вид функции распределения определяется наименованием ее аргумента. Следовательно, например, $F(x)$ и $F(y)$ обозначают не одну и ту же функциональную зависимость, в которой произведена замена аргумента, а функции распределения, соответственно,

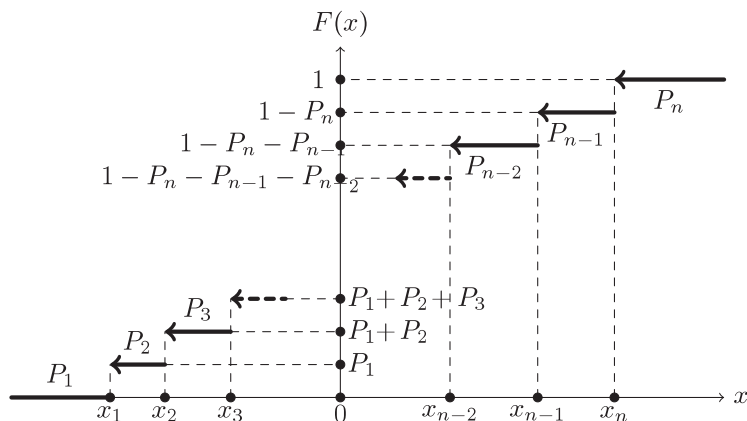


Рис. 16.1

К определению функции распределения

для случайных величин X и Y , могущие иметь совершенно различные математические выражения. Только в тех случаях, когда аргумент у функции распределения (по тем или иным причинам) может быть обозначен буквой, отличной от наименования рассматриваемой случайной величины, это наименование будем явно указывать в индексе. Так, например, $F_x(z)$ будет обозначать функцию распределения случайной величины X , взятую при аргументе z .

В приложениях часто встречаются дискретные случайные величины, способные принимать только целые неотрицательные значения. Эти случайные величины называются *целочисленными*. Значения, принимаемые такой случайной величиной, можно всегда считать равными соответственно $0, 1, \dots, n$, где n — заданное конечное или бесконечное число. В том случае, когда целочисленная случайная величина не может принимать некоторые значения k , для них следует считать $p_k = 0$. От любой дискретной случайной величины можно перейти к целочисленной случайной величине, если, например, в качестве ее возможных значений условно рассматривать номера возможных значений исходной случайной величины. При этом, например, всегда будет $p_0 = 0$, если номера возможных значений x_k начинаются с первого.

Примером целочисленной случайной величины является число m появлений события A в серии n независимых испытаний (схема Бернулли). Вообще, путем введения индикаторов вида (16.1) все задачи, решаемые в схеме случайных событий, могут быть переведены на язык дискретных случайных величин. Следовательно, понятие случайной величины является более общим, чем понятие случайного события, включающим понятие случайного события как частный случай. При исследовании целочисленных случайных величин весьма эффективным является метод производящих функций, рассмотренный в § 10 применительно к случайным событиям.

Будем называть производящей функцией $G(u)$ целочисленной случайной величины X , заданной рядом распределения

$$p_m = \mathbf{P}(X = m), \quad (m = \overline{1, n}),$$

многочлен от u , коэффициентами которого являются вероятности p_m , то есть положим, как и раньше,

$$G(u) = \sum_{m=0}^n p_m u^m. \quad (16.7)$$

В случае $n = \infty$ имеем вместо конечной суммы бесконечный ряд, который гарантированно сходится по крайней мере при $|u| \leq 1$.

Производящая функция всегда удовлетворяет условию нормировки

$$G(1) = 1, \quad (16.8)$$

являющемуся следствием равенства (16.4), причем как при конечном n , так и при $n = \infty$. Коэффициенты разложения (16.7) можно определить по формулам (см. § 10)

$$p_k = \frac{1}{k!} \cdot \left. \frac{d^k G(u)}{du^k} \right|_{u=0} = \frac{1}{2\pi i} \int_{\Gamma} \frac{G(z)}{z^{k+1}} dz, \quad (k = \overline{0, n}), \quad (16.9)$$

где интегрирование по комплексной переменной z в последнем интеграле производится по любому гладкому замкнутому контуру Γ , охватывающему начало координат.

Обозначим через Q_k вероятность того, что целочисленная случайная величина X не превосходит заданного числа k , то есть примем

$$Q_k = \mathbf{P}(X \leq k) = \sum_{j=0}^k p_j = F(k+1), \quad (k = \overline{0, n}). \quad (16.10)$$

Для вероятностей $Q_0, Q_1, \dots, Q_n$ также можно ввести свою производящую функцию, приняв

$$Q(u) = \sum_{k=0}^n Q_k u^k. \quad (16.11)$$

В отличие от производящей функции $G(u)$ для исходных вероятностей p_k , эта функция уже не удовлетворяет обычному

условию (16.8), то есть $Q(1) \neq 1$. Поэтому данную функцию будем называть *обобщенной производящей функцией*. Установим связь обобщенной производящей функции $Q(u)$ с исходной производящей функцией $G(u)$. Из (16.10) следует, что

$$Q_k = Q_{k-1} + p_k, \quad (k = \overline{0, n}),$$

где для единообразия принято, что $Q_{-1} \equiv 0$. Подставляя это выражение в (16.11), получим

$$Q(u) = \sum_{k=1}^n Q_{k-1} u^k + \sum_{k=0}^n p_k u^k. \quad (16.12)$$

Легко видеть, что

$$Q_n = \sum_{j=0}^n p_j = 1.$$

Поэтому равенство (16.12) можно переписать в виде

$$Q(u) = uQ(u) + G(u) - u^{n+1}. \quad (16.13)$$

Разрешая последнее равенство относительно $Q(u)$, имеем

$$Q(u) = \frac{G(u) - u^{n+1}}{1 - u}. \quad (16.14)$$

Обобщенную производящую функцию можно ввести и для вероятностей того, что случайная величина X превосходит заданное значение k , то есть для вероятностей

$$R_k = \mathbf{P}(X > k) = \sum_{j=k+1}^n p_j = 1 - F(k+1), \quad (k = \overline{0, n-1}). \quad (16.15)$$

Положим

$$R(u) = \sum_{k=0}^{n-1} R_k u^k. \quad (16.16)$$

С помощью (16.11) и (16.16) получаем

$$Q(u) + R(u) = \sum_{k=0}^n u^k = \frac{1 - u^{n+1}}{1 - u}. \quad (16.17)$$

Поэтому вновь введенная обобщенная производящая функция $R(u)$ и исходная производящая функция $G(u)$ самой случайной величины X будут связаны равенством

$$R(u) = \frac{1 - G(u)}{1 - u}. \quad (16.18)$$

Метод производящих функций является эффективным средством анализа распределения целочисленных случайных величин.

Пример 16.1. В партии из десяти изделий три бракованных. Наудачу берется пять изделий. Составить ряд распределения для числа X взятых бракованных изделий, построить многоугольник распределения и функцию распределения. Составить производящую функцию.

Решение. Вначале посчитаем вероятности p_k непосредственно. Возможные значения случайного числа X взятых бракованных изделий $x_k = 0, 1, 2, 3$, причем

$$p_k = \mathbf{P}(X = k) = \frac{C_3^k C_7^{5-k}}{C_{10}^5}, \quad (k = \overline{0, 3}).$$

Рассчитав по этой формуле вероятности p_k , получим ряд распределения, который представим в виде таблицы 16.1. Многоугольник распределения и график функции распределения, построенные по этим данным, приведены на рис. 16.2.

Таблица 16.1

Ряд распределения числа взятых бракованных изделий

k	0	1	2	3
x_k	0	1	2	3
p_k	$\frac{1}{12}$	$\frac{5}{12}$	$\frac{5}{12}$	$\frac{1}{12}$

Теперь воспользуемся *методом производящих функций*. Производящая функция $G(u)$ в соответствии с (16.7) будет иметь вид

$$G(u) = \frac{1}{12}(1 + 5u + 5u^2 + u^3).$$

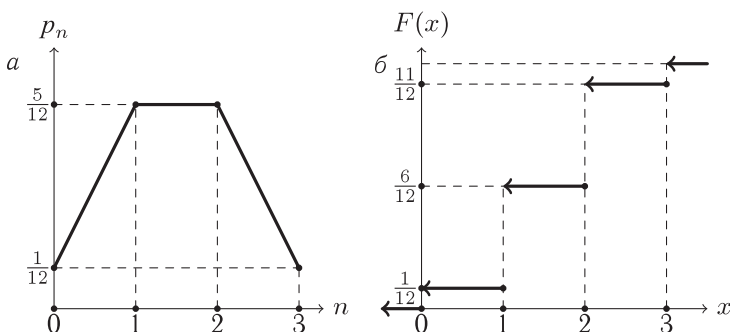


Рис. 16.2

К примеру 16.1: а — многоугольник распределения,
б — функция распределения

Пример можно решить и иначе: вначале найти производящую функцию, а затем определить вероятности p_k . Для этого обозначим через $p_k^{(n)}$ вероятность получения k бракованных изделий из числа выбранных n изделий, а через $G^{(n)}(u)$ соответствующую им производящую функцию. Очевидно, для $p_k^{(n)}$ выполняются рекуррентные соотношения

$$p_k^{(n)} = p_{k-1}^{(n-1)} \cdot \frac{4-k}{11-n} + p_k^{(n-1)} \cdot \frac{8-n+k}{11-n}.$$

Умножая последнее равенство на u^k и суммируя по k от 0 до 3, получим рекуррентное соотношение уже для производящих функций

$$G^{(n)}(u) = \frac{8-n+3u}{11-n} G^{(n-1)}(u) + \frac{u-u^2}{11-n} \cdot \frac{dG^{(n-1)}(u)}{du}.$$

Так как $G^{(1)}(u) = \frac{7}{10} + \frac{3}{10}u$, то, применяя рекуррентное соотношение для $G^{(n)}$ четыре раза, получим

$$G^{(5)}(u) = \frac{1}{12} + \frac{5}{12}u + \frac{5}{12}u^2 + \frac{1}{12}u^3,$$

откуда, конечно, следуют прежние выражения для искомых вероятностей

$$p_0 = \frac{1}{12}, \quad p_1 = \frac{5}{12}, \quad p_2 = \frac{5}{12}, \quad p_3 = \frac{1}{12}.$$

Пример 16.2 (Продолжение задачи о разорении § 10).

В качестве примера применения обобщенных производящих функций вновь рассмотрим задачу о разорении игрока, в которой предполагается, что игроки A и B играют серию партий, ставкой в которых является один рубль, причем каждую партию игрок A выигрывает с вероятностью p , проигрывает с вероятностью q и делает ничью с вероятностью r . Требуется определить вероятность разорения игрока A , если его начальный капитал равнялся a рублей, а начальный капитал игрока B составлял b рублей.

Решение. Введем в рассмотрение случайную величину X — капитал игрока A в ходе игры. Очевидно, X может принимать только целые значения m от 0 до $(a+b)$.

Обозначим $P(x)$ вероятность разорения игрока A (выигрыш игры игроком B) при начальном капитале A , равном x . Совокупность значений x и соответствующих значений вероятностей $P(x)$ не образуют ряд распределения, так как $\sum_{x=0}^{a+b} P(x) \neq 1$. Однако метод производящих функций применим и в этом случае. Введем обобщенную производящую функцию $Q(u)$, положив

$$Q(u) = \sum_{m=0}^{a+b} P(m)u^m,$$

причем

$$P(0) = 1, \quad P(a+b) = 0.$$

Так как игрок A может разориться при капитале x , или выиграв очередную партию, или проиграв ее, или сделав ничью, то справедливо рекуррентное соотношение

$$P(m) = pP(m+1) + qP(m-1) + rP(m), \quad (0 < m < a+b).$$

Умножая обе части этого равенства на u^m и суммируя по m от 1 до $(a+b-1)$, получим

$$Q(u) - 1 = \frac{p}{u} [Q(u-1 - P(1))] + qu[Q(u) - 1] + qu - quP(a+b-1)u^{a+b-1} + r[Q(u) - 1].$$

Разрешая это равенство относительно $Q(u)$, находим

$$Q(u) = \frac{qP(a+b-1)u^{a+b+1} + [pP(1) - p - q]u + p}{(u-1)(qu-p)}.$$

Так как по построению $Q(u)$ является полиномом степени $(a+b-1)$, то числитель должен делиться на знаменатель нацело, то есть корни знаменателя $u = 1$ и $u = \frac{p}{q}$ должны являться также и корнями числителя. Следовательно, вероятности $P(1)$ и $P(a+b-1)$ удовлетворяют системе уравнений

$$qP(a+b-1) + pP(1) = q,$$

$$q \left(\frac{p}{q} \right)^{a+b+1} P(a+b-1) + p \cdot \frac{p}{q} P(1) = \frac{p^2}{q},$$

решая которую получим

$$P(1) = \frac{1 - \left(\frac{p}{q} \right)^{a+b-1}}{1 - \left(\frac{p}{q} \right)^{a+b}};$$

$$P(a+b-1) = \frac{1 - \left(\frac{p}{q}\right)}{1 - \left(\frac{p}{q}\right)^{a+b}}.$$

Подставляя эти значения в формулу для $Q(u)$ и производя необходимые сокращения, будем иметь

$$Q(u) = \sum_{m=0}^{a+b} \frac{1 - \left(\frac{p}{q}\right)^{a+b-m}}{1 - \left(\frac{p}{q}\right)^{a+b}} u^m,$$

откуда следует

$$P(m) = \frac{1 - \left(\frac{p}{q}\right)^{a+b-1}}{1 - \left(\frac{p}{q}\right)^{a+b}}.$$

Полагая здесь $m = a$, получим искомую вероятность разорения первого игрока, совпадающую с (10.45)

$$P(A) = \frac{1 - \left(\frac{p}{q}\right)^b}{1 - \left(\frac{p}{q}\right)^{a+b}}.$$

Пример 16.3. Проводится серия из n испытаний Бернулли, в каждом из которых событие A появляется с вероятностью p . Построить ряд распределения для случайной величины X , дающей число появлений события A в данной серии испытаний.

Решение. Возможные значения случайной величины $x_j = j$, ($j = \overline{0, n}$) образуют начальный отрезок натурального ряда

чисел. При этом в терминах случайной величины X введенные в § 8 случайные события B_j , состоящие в том, что событие A появилось ровно j раз, могут быть выражены так:

$$B_j = \{X = j\}, \quad (j = \overline{0, n}).$$

Поэтому случайную величину X можно выразить через характеристические случайные величины (16.1) событий B_j следующим образом:

$$X = \sum_{j=1}^n jI(B_j).$$

Что касается вероятностей

$$p_j = P\{X = j\}, \quad (j = \overline{0, n}),$$

то нетрудно понять, что

$$p_j = P_j^{(n)},$$

где $P_j^{(n)}$ было получено в § 8 и соответствует биномиальному распределению

$$p_j = C_n^j p^j q^{n-j}, \quad (q = 1 - p; j = \overline{0, n}).$$

Следует сказать, что сам термин «биномиальное распределение» в первую очередь относится к случайной величине X , рассмотренной в данном примере, а уже потом — к набору вероятностей $P_j^{(n)}$, полученных в § 8.

Пример 16.4. Серия испытаний Бернулли, в каждом из которых событие A появляется с вероятностью p , проводится до первого появления события A . Требуется построить ряд распределения для случайной величины Y , равной числу неудачных опытов до первого удачного.

Решение. В данном случайном эксперименте, как было показано в § 12, мы имеем бесконечное пространство элементарных событий, которым соответствует бесконечный набор возможных значений случайной величины Y , задаваемый равенством $y_j = j$, где $j = \overline{0, \infty}$.

Обозначим, как и раньше, через A_j событие, эквивалентное появлению A в j -м опыте, а через B_j — событие, состоящее в появлении A в j -м опыте впервые в серии. Ясно, что

$$B_j = A_j \prod_{i=1}^{j-1} \overline{A_i}, \quad (j = \overline{1, \infty}),$$

где при $j = 1$ «пустое» произведение событий считается равным достоверному событию. Поскольку

$$p_j = P\{Y = j\} = \mathbf{P}(B_{j+1}), \quad (j = \overline{0, \infty}),$$

то благодаря независимости A_i друг от друга находим

$$p_j = (1 - p)^j p, \quad (j = \overline{0, \infty}).$$

Легко видеть, что условие нормировки вероятностей p_j выполняется при любом $p \in (0, 1)$, в чем легко убедиться, воспользовавшись формулой для бесконечно убывающей геометрической прогрессии. Поскольку вероятности p_j образуют геометрическую прогрессию, то такой закон распределения Y называется *геометрическим*. Сама же величина представляется в виде бесконечного ряда

$$Y = \sum_{j=0}^{\infty} j I(B_{j+1}).$$

Нулевой член в этой сумме можно опустить, заменив нижний предел суммирования на $j = 1$.

Пример 16.5. Имеется двузначное десятичное число, обе цифры которого с равной вероятностью могут принимать произвольное целое значение от 0 до 9. Построить ряд распределения для суммы цифр построенного числа.

Решение. Введем случайные события: A_i — в первом разряде случайного числа стоит i , ($i = \overline{0, 9}$), а B_j — во втором разряде числа стоит j , ($j = \overline{0, 9}$). Тогда само число X может быть представлено в виде

$$X = \sum_{i=0}^9 \sum_{j=0}^9 (i + 10j) I(A_i B_j),$$

а сумма его цифр Y представляется так:

$$Y = \sum_{i=0}^9 \sum_{j=0}^9 (i+j) I(A_i B_j).$$

Очевидно, Y является дискретной случайной величиной, возможными значениями которой y_j являются целые числа от 0 до 18. Представим Y в виде

$$Y = \sum_{k=0}^{18} k I(C_k).$$

Здесь C_k обозначают случайные события, происходящие в том случае, когда Y принимает значение k . Нетрудно понять, что событие C_k можно легко выразить с помощью операций суммирования и перемножения событий A_i и B_j :

$$C_k = \begin{cases} \sum_{i=0}^k A_i B_{k-i}, & 0 \leq k \leq 9, \\ \sum_{i=k-9}^{18} A_i B_{k-i}, & 10 \leq k \leq 18. \end{cases}$$

Действительно, если $0 \leq k \leq 9$, то нужно взять произведения $A_i B_j$ при значениях индексов i и j , в сумме дающих k , при этом каждый из индексов должен пробегать значения от 0 до k . Если же $k > 9$, то по-прежнему нужно суммировать произведения вида $A_i B_{k-i}$, но следует брать только такие комбинации индексов, для которых одновременно $0 \leq i \leq 9$ и $0 \leq k-i \leq 9$, то есть суммировать нужно по значениям $k-9 \leq i \leq 9$.

Итоговое выражение можно записать в виде одной суммы с пределами, зависящими от k ,

$$C_k = \sum_{i=\max(0, k-9)}^{\min(k, 9)} A_i B^{k-i}, \quad (0 \leq k \leq 18).$$

Очевидно, искомые вероятности

$$p_k = P(Y = k) = \mathbf{P}(C_k), \quad (0 \leq k \leq 18).$$

При вычислении вероятностей C_k следует учесть, что так как цифры числа независимы, то система событий $\{A_i\}_{i=0}^9$ не зависит от системы $\{B_j\}_{j=0}^9$. При этом внутри каждой системы все события несовместны. Согласно примеру 15.5 несовместные события всегда зависимы, поэтому нельзя утверждать, что вся система событий $\{A_i, B_j\}_{i,j=0}^9$ независима. Она всего лишь распадается на две независимые подсистемы из несовместных, но зависимых событий.

Система $\{A_i, B_j\}_{i,j=0}^9$ симметрична в том смысле, что

$$\mathbf{P}(A_i B_j) = \left(\frac{1}{10}\right)^2 = 0,01 = \text{const}(i, j).$$

Поэтому, вычисляя с помощью теорем умножения и сложения вероятностей $\mathbf{P}(C_k)$, получим

$$p_k = 0,01 \begin{cases} k+1, & 0 \leq k \leq 9, \\ 19-k, & 10 \leq k \leq 18. \end{cases}$$

Графически ряд распределения $(k, p_k; k = \overline{0, 18})$ представлен на рис. 16.3.

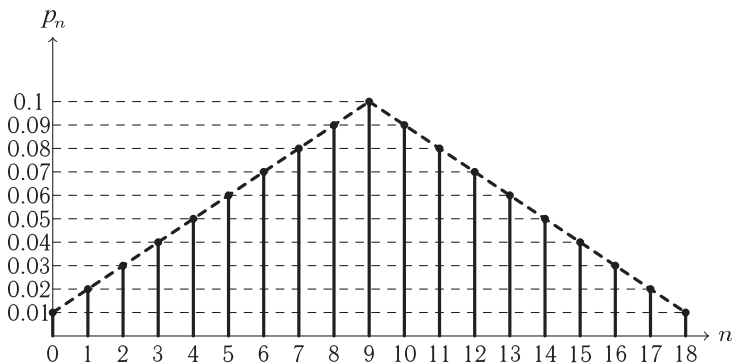


Рис. 16.3

Ряд распределения суммы цифр случайного числа

§ 17. АКСИОМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН

В предыдущем параграфе мы рассмотрели простейший тип случайных величин — величины дискретного типа. Как было показано, такие величины можно выразить через специально построенную систему событий $\{A_i\}_{i=1}^n$ с помощью введения так называемых характеристических случайных величин события, определяемых равенством (16.1). Построенная таким образом теория основывается на представлении вида (16.2) и, в общем, не выходит за рамки рассмотрения систем случайных событий. Как мы видели, такой подход позволяет решать и прикладные задачи.

Между тем, помимо дискретных существуют и более сложные по своим свойствам случайные величины, способные принимать произвольные числовые значения из некоторого интервала. Мы их называли непрерывными случайными величинами, или величинами непрерывного типа. Для них подход, изложенный в § 16, уже не срабатывает. Конечно, можно было бы пытаться представить непрерывную величину X , принимающую значения в некотором интервале $[a, b]$, в виде интеграла, обобщающего сумму (16.2)

$$X = \int_a^b f(x)I(A_x)dx, \quad (17.1)$$

где $f(x)dx$ дает элементарную вероятность попадания X в промежуток $(x, x + dx)$, A_x — случайное событие, означающее, что $X = x$, а $I(A_x)$ — характеристическая случайная величина этого события. К сожалению, такое определение является неудовлетворительным и в теории вероятностей не используется по целому ряду причин.

Во-первых, представление вида (17.1) чрезвычайно сложно обосновать математически. В предыдущей главе мы отмечали, что рассмотрение несчетного множества исходов испытания усложняет задачу и требует применения сложного математического аппарата. Во-вторых, с чисто прикладной точки зрения представления вида (17.1) являются неудобными и

не позволяют получить интересующие нас вероятностные характеристики непрерывной случайной величины X . Наконец, в-третьих, подобного рода определения не допускают обобщения на случайные эксперименты более сложного вида, например, те, в которых с помощью случайных функций описываются динамические случайные явления.

С учетом всего сказанного приходится отказаться от, казалось бы, заманчивой и логичной перспективы элементарно свести случайные величины к рассмотрению случайных событий. Требуется самостоятельная теория, причем ею обычно пользуются как для непрерывных, так и для дискретных случайных величин, хотя последние, как мы видели, можно определять и через случайные события и изучать, в общем, элементарными средствами.

Общая теория случайных величин основывается на аксиоматике А. Н. Колмогорова, изложенной в предыдущей главе. Напомним, что математическая модель любого случайного эксперимента состоит в построении вероятностного пространства $(\Omega, \mathfrak{A}, P)$, где Ω обозначает пространство элементарных событий, $\mathfrak{A}$ — некоторую сигма-алгебру событий, заданную на Ω и включающую различные подмножества Ω , трактуемые как случайные события, а P есть нормированная счетно-аддитивная вероятностная мера, заданная на элементах $\mathfrak{A}$.

Рассмотрим числовую величину X , являющуюся функцией элементарного события

$$X = \varphi(\omega), \quad (\omega \in \Omega) \quad (17.2)$$

и определенную для всех $\omega \in \Omega$. Если функция φ измерима относительно меры P , то такую числовую функцию $X = \varphi(\omega)$ называют *случайной величиной*.

Понятие измеримой функции детально обсуждается в курсах функционального анализа [13]. Напомним, что мера $\mathbf{P}(A)$ является числовой функцией множества A и служит обобщением таких понятий, как длина отрезка, площадь плоской фигуры, объем тела. Измеримость функции $\varphi(\omega)$ означает, что измеримыми должны быть все множества элементов ω , для которых выполняется условие $\varphi(\omega) < x$ при произвольном $-\infty < x < \infty$.

Поясним этот вопрос подробнее. Рассмотрим множество A_x тех элементарных событий $\omega \in \Omega$, для которых $X(\omega)$ оказывается меньшим, чем заданное число x

$$A_x = \{\omega : X(\omega) < x\}. \quad (17.3)$$

Так вот, для измеримой функции φ множество A_x обязано быть P -измеримым, то есть должно принадлежать сигмалгебре $\mathfrak{A}$. Иначе говоря, для P -измеримой функции φ все множества вида (17.3) при любом $x \in (-\infty, \infty)$ должны представлять собой некоторые случайные события, которые, в принципе, могут наблюдаться в данном случайном эксперименте.

При таком определении случайной величины с ней всегда можно связать вероятностную меру, поскольку этой мерой при любом x обладает множество A_x , определяемое (17.3), так как $A_x \subset \mathfrak{A}$. Обычно вводят в рассмотрение функцию распределения X , определяя ее равенством

$$F(x) = \mathbf{P}(A_x), \quad (-\infty < x < \infty). \quad (17.4)$$

Нужно сказать, что при элементарном изложении теории вероятностей не прибегают к описанным теоретико-множественным построениям. Случайной называют такую величину, которая в результате опыта может иметь различные значения, причем заранее не известно, какие именно. Затем говорят, что исчерпывающей характеристикой случайной величины X является ее функция распределения

$$F(x) = \mathbf{P}(X < x). \quad (17.5)$$

Случайная величина считается заданной, если задана функция F .

Элементарное определение иногда оказывается достаточным в некоторых прикладных задачах, где интересуются только законом распределения X , то есть видом ее функции распределения $F(x)$. Например, если рассматривать дискретные случайные величины, описанные в § 16, и удовольствоваться только знанием вероятностей p_k , то в принципе, можно ограничиться и элементарным определением, выразив случайную

величину X через случайные события, как это описано выше, и трактуя последние в элементарном смысле в духе главы 1. Вместе с тем, такое определение часто оказывается неудовлетворительным по целому ряду причин.

Во-первых, между законом распределения случайной величины и самой случайной величиной нет взаимной однозначной связи. Каждой случайной величине соответствует только один определенный закон распределения, но одному и тому же закону $F(x)$ могут соответствовать несколько различных случайных величин, не совпадающих друг с другом.

Рассмотрим простейший пример. Имеется дискретная случайная величина X с возможными значениями $\{x_k\}_{k=-n}^n$, симметричными относительно начала координат, так что $x_{-k} = -x_k$, ($k = \overline{1, n}$), причем $x_0 = 0$. Допустим, что вероятности p_k также симметричны относительно изменения знака индекса k , то есть $p_{-k} = p_k$, ($k = \overline{1, n}$). Такой закон распределения тоже называют симметричным. Нетрудно понять, что при этих условиях случайная величина $Y = -X$ и исходная случайная величина X имеют один и тот же закон распределения $F(x)$. Иначе говоря, две совершенно разные случайные величины, у которых в результате эксперимента всегда получаются значения, не совпадающие даже по своему знаку, оказываются неразличимыми в рамках элементарного определения случайной величины.

Во-вторых, существуют задачи, в которых недостаточно знать только закон распределения случайной величины или какие-либо его параметры. Часто требуется производить те или иные действия или преобразования с самой случайной величиной. Если сводить, как в элементарном определении, задание случайной величины к заданию ее закона распределения, то ставить и решать такие задачи непосредственно на основе элементарного определения затруднительно.

В-третьих, элементарное определение не обладает достаточной гибкостью, чтобы быть обобщенным на случайные явления более сложной природы. Между тем, аксиоматическое определение допускает самые широкие обобщения.

Рассмотрим, например, определение случайной функции [22]. Для введения понятия случайной функции $X(t)$

неслучайного аргумента t достаточно предположить, что функция φ в определении (17.2) зависит от t , как от параметра, то есть положить

$$X(t) = \varphi(\omega, t) \quad (17.6)$$

и, по-прежнему, считать, что эта функция измерима относительно $\omega \in \Omega$ при любом t . При соблюдении этого условия множество вида

$$A_{x,t} = \{\omega : X(t) < x\}, \quad (17.7)$$

которое строится по аналогии с (17.3), будет принадлежать сигма-алгебре $\mathfrak{A}$. Отличие от случайной величины будет только в том, что сами элементы сигма-алгебры будут зависеть еще и от t , как от параметра. Разумеется, сама теория оказывается гораздо более сложной, однако, общая идея определения, как видим, сохраняется.

Вернемся к рассмотрению случайных величин и попытаемся дать на аксиоматическом языке определение дискретной случайной величины X , характеризуемой рядом распределения $(x_j, p_j; j = \overline{1, n})$. Для этого рассмотрим некоторую полную систему событий $\{A_k\}_{k=1}^n \subset \Omega$, характеризующую следующими свойствами:

$$A_i \cap A_j = \emptyset, \quad (i = \overline{1, n}; j = \overline{1, n}; i \neq j), \quad (17.8)$$

$$\bigcup_{i=1}^n A_i = \Omega, \quad (17.9)$$

$$\mathbf{P}(A_i) = p_i, \quad (i = \overline{1, n}). \quad (17.10)$$

Ясно, что такая система событий может быть выбрана неоднозначным образом. Например, если пространство Ω представляет собой некоторую область на плоскости, а мера P определяется как мера Лебега, то есть площадь соответствующей плоской фигуры, то речь идет о разбиении области Ω на n подобластей, имеющих заданные площади. Такое разбиение можно произвести бесконечно большим числом способов.

Далее введем функцию, заданную при $\omega \in \Omega$ и определяемую кусочно-постоянным выражением

$$\varphi(\omega) = x_j; \omega \in A_j \quad (j = \overline{1, n}), \quad (17.11)$$

и определим случайную величину X по формуле (17.2). Очевидно, эта случайная величина будет обладать рядом распределения (x_j, p_j) , функцией распределения (16.6) и будет допускать представление (16.2). Из построения видно, что такая случайная величина с данным законом распределения будет неединственной. Если взять любые другие базовые случайные события A_i , удовлетворяющие условию (17.10), образующие полную группу событий, то ряд распределения X не изменится, хотя функция φ , сохраняя кусочно-постоянное выражение, изменит свой вид, поскольку изменятся области ее кусочного постоянства.

Далее рассмотрим некоторые примеры аксиоматического определения случайных величин, ограничившись пока рассмотрением дискретных случайных величин.

Пример 17.1. Один раз бросается монета. Пространство элементарных событий строится так, как это описано в примере 12.1.

$$\Omega = \{O, P\}.$$

В качестве сигма-алгебры $\mathfrak{A}$ принята максимальная сигма-алгебра $\mathfrak{A}_{\max}$ (совокупность всех подмножеств Ω). Вероятностная мера задается условием

$$\mathbf{P}(\{O\}) = \mathbf{P}(\{P\}) = \frac{1}{2}.$$

Функция φ принимается в виде

$$\varphi(\omega) = \begin{cases} 1, & \omega = O, \\ 0, & \omega = P. \end{cases}$$

Необходимо построить функцию распределения $X = \varphi(\omega)$, а также множество A_x , определяемое (17.3).

Решение. Вначале построим множество A_x . При $x \leq 0$ это множество пусто, при $0 < x \leq 1$ оно совпадает с элементом

$\omega = P$, а при $x > 1$ в A_x входит все пространство Ω , иначе говоря

$$A_x = \{\omega : X < x\} = \begin{cases} \emptyset, & x \leq 0, \\ \{P\}, & 0 < x \leq 1, \\ \Omega, & x > 1. \end{cases}$$

Функция распределения такой случайной величины дается равенством

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

График этой функции распределения представлен на рис. 17.1.

Пример 17.2. Построить функцию распределения для числа выигравших очков при однократном бросании игральной кости, используя пространство элементарных событий, введенное в примере 12.3 и максимальную сигма-алгебру событий.

Решение. Напомним, что пространство Ω в данном случае представляет отрезок из шести младших чисел натурального ряда

$$\Omega = \{i : i = \overline{1, 6}\}.$$

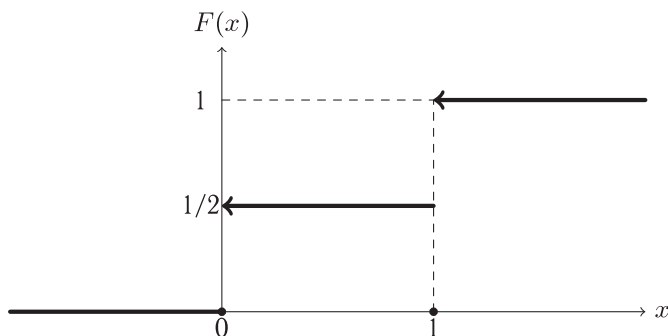


Рис. 17.1

Функция распределения в примере 17.1

Обозначим случайное число выпавших очков через X . Тогда как функция ω эта случайная величина выразится в виде

$$X = \varphi(\omega) = \omega, \quad (\omega = \overline{1, 6}).$$

Множество A_x представляется выражением

$$A_x = \{\omega : X < x\} = \begin{cases} \emptyset, & x \leq 1, \\ \{1\}, & 1 < x \leq 2, \\ \{1, 2\}, & 2 < x \leq 3, \\ \{1, 2, 3\}, & 3 < x \leq 4, \\ \{1, 2, 3, 4\}, & 4 < x \leq 5, \\ \{1, 2, 3, 4, 5\}, & 5 < x \leq 6, \\ \Omega, & x > 6. \end{cases}$$

Функция распределения, с учетом того, что

$$\mathbf{P}(X = i) = \frac{1}{6}, \quad (i = \overline{1, 6}),$$

выражается в следующем виде:

$$F(x) = \begin{cases} 0, & x \leq 1, \\ \frac{1}{6}, & 1 < x \leq 2, \\ \frac{2}{6}, & 2 < x \leq 3, \\ \frac{3}{6}, & 3 < x \leq 4, \\ \frac{4}{6}, & 4 < x \leq 5, \\ \frac{5}{6}, & 5 < x \leq 6, \\ 1, & x > 6. \end{cases}$$

Ее график можно найти на рис. 17.2.

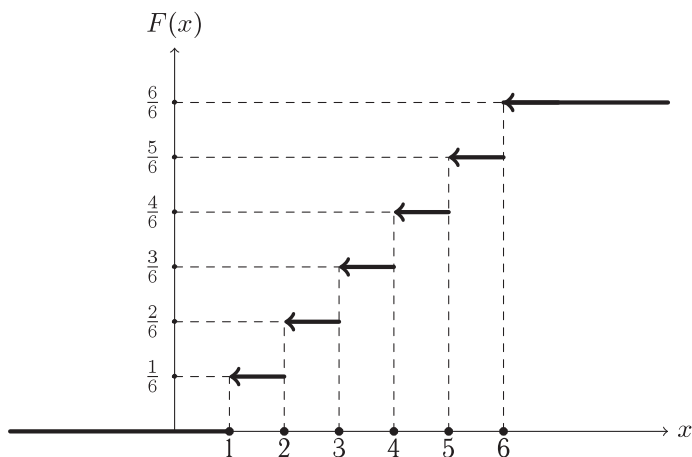


Рис. 17.2

Функция распределения для числа очков
при однократном бросании игральной кости

Пример 17.3. Получить функцию распределения для характеристической случайной величины события A , определяемой (16.1)

$$X = I(A) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Решение. Легко видеть, что в данном случае

$$A_x = \{\omega : X < x\} = \begin{cases} \emptyset, & x \leq 0, \\ \bar{A}, & 0 < x \leq 1, \\ \Omega, & x > 1. \end{cases}$$

Поэтому независимо от конкретного задания вероятностной меры P имеем

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \mathbf{P}(\bar{A}), & 0 < x \leq 1, \\ 1, & x > 1. \end{cases}$$

Пример 17.4. Допустим, что в задаче на геометрические вероятности (см. § 3) область Ω поделена на k равновеликих частей (с точки зрения их меры Лебега). Вероятностная мера вводится с помощью меры Лебега (см. § 15). Случайная величина X представляет собой номер той области A_j , в которую в результате опыта попала случайная точка. Построить функцию распределения так определенной случайной величины.

Решение. Обозначим меру Лебега произвольной области $A \subset \Omega$, измеримой по Лебегу, через $m(A)$. В качестве сигма-алгебры $\mathfrak{A}$ используем сигма-алгебру всех измеримых по Лебегу подмножеств из Ω (см. § 15). Тогда, как мы помним, для любого события A (условимся, как и раньше отождествлять случайные события с их отражением на Ω) имеем

$$\mathbf{P}(A) = \frac{m(A)}{m(\Omega)}.$$

Поскольку по условию все области A_j равновелики, то

$$m(A_j) = \text{const}(j), \quad (j = \overline{1, k}).$$

С другой стороны, учитывая, что

$$\sum_{j=1}^k m(A_j) = m(\Omega),$$

заключаем

$$p_j = \mathbf{P}(A_j) = \frac{1}{k} = \text{const}(j), \quad (j = \overline{1, k}).$$

Значит, случайная величина X характеризуется рядом распределения $(x_j = j, p_j = \frac{1}{k}, j = \overline{1, k})$. Дискретное распределение, в котором все p_j одинаковы, называют *равномерным распределением*.

Далее имеем

$$X = \varphi(\omega) = i, \quad \omega \in A_i, \quad (i = \overline{1, k}).$$

Легко видеть, что

$$A_x = \begin{cases} \emptyset, & x \leq 1, \\ \bigcup_{i=1}^j A_i, & j < x \leq j+1, (j = \overline{1, k-1}), \\ \Omega, & x > k. \end{cases}$$

Поэтому окончательно имеем

$$F(x) = \begin{cases} 0, & x \leq 1, \\ \frac{j}{k}, & j < x \leq j+1, (j = \overline{1, k-1}), \\ 1, & x > k. \end{cases}$$

Пример 17.5. Целочисленная случайная величина имеет функцию распределения

$$F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - \frac{1}{n}, & n-1 < x \leq n, (n = \overline{1, \infty}). \end{cases}$$

Построить ряд распределения этой случайной величины.

Решение. Представленная функция является кусочно-постоянной и претерпевает скачки в точках x_j , ($j = \overline{2, \infty}$). Величина скачка

$$\begin{aligned} p_j &= F(j+0) - F(j-0) = \left(1 - \frac{1}{j}\right) - \left(1 - \frac{1}{j-1}\right) = \\ &= \frac{1}{j-1} - \frac{1}{j} = \frac{1}{(j-1)j}, \quad (j = \overline{2, \infty}) \end{aligned}$$

дает вероятность того, что $X = j$. Следовательно, ряд распределения данной случайной величины представляется следующим набором пар чисел $(j, \frac{1}{(j-1)j})$, ($j = \overline{2, \infty}$).

**§ 18. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ НЕПРЕРЫВНОГО
И СМЕШАННОГО ТИПА.
ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ
И ПЛОТНОСТЬ ВЕРОЯТНОСТИ**

Случайные величины непрерывного типа, как уже говорилось выше, не могут быть получены путем сопоставления их возможных значений случайным событиям, как это имеет место для дискретных случайных величин. Поэтому при построении теории непрерывных случайных величин от такого подхода приходится отказаться и использовать общие универсальные теоретико-множественные построения.

При теоретико-множественном построении теории вероятностей, как было отмечено в § 17, введение понятия случайной величины принципиально не отличается от введения понятия случайного события. Рассматривается вероятностное пространство $(\Omega, \mathfrak{A}, P)$, где Ω — пространство элементарных событий ω , $\mathfrak{A}$ — сигма-алгебра подмножеств этого пространства, $P = \mathbf{P}(A)$ — нормированная счетно-аддитивная мера множества A . Вводится числовая функция $X(\omega)$ элементарного события ω , обладающая тем свойством, что множество ω , для которых выполнено, что $X(\omega) < x$, $(-\infty < x < \infty)$, принадлежит сигма-алгебре $\mathfrak{A}$. В этом случае величину $X(\omega)$ называют *случайной величиной*. В соответствии с этим определением, для каждого x существует вероятность

$$\mathbf{P}\{X(\omega) < x\} = F(x). \quad (18.1)$$

В зависимости от вида функции $X(\omega)$ случайная величина будет принадлежать к различным *типам*: если числовые значения $X(\omega)$ образуют счетное множество $\{x_i\}_{i=1}^{\infty}$, то случайная величина является *дискретной*; если $X(\omega)$ может принимать любые значения в некотором интервале, причем бесконечно малая вероятность попадания $X(\omega)$ в интервал $(x, x + dx)$ имеет первый порядок относительно dx , то такая величина называется *непрерывной*; наконец, если наряду с дискретными значениями $x_1, x_2, \dots$, которые случайная величина может принимать с отличными от нуля вероятностями, эта величина $X(\omega)$ также может принимать и любые

вещественные значения в некотором интервале, причем вероятность попадания ее значения в интервал $(x, x + dx)$ пропорциональна dx , то такая величина называется случайной величиной *смешанного типа*.

Во всех перечисленных случаях полной характеристикой случайной величины является *функция ее распределения* $F(x)$. Явный вид функциональной зависимости $X(\omega)$ имеет значение только в том отношении, что он определяет тип случайной величины и ее свойства. Поэтому в обозначении случайной величины аргумент ω в дальнейшем мы указывать не будем, обозначая случайную величину просто X (или другой большой буквой латинского алфавита), указывая явно ее тип и непосредственно определяя ее вероятностные свойства, например, указывая ее функцию распределения.

Рассмотрим общие свойства функции распределения $F(x)$, которые справедливы для случайных величин произвольного типа из перечисленных выше. Независимо от типа случайной величины, ее функция распределения имеет следующие свойства:

$$0 \leq F(x) \leq 1, \quad (18.2)$$

$$F(-\infty) = 0, \quad (18.3)$$

$$F(+\infty) = 1, \quad (18.4)$$

$$\mathbf{P}\{a \leq X \leq b\} = F(b+0) - F(a-0), \quad (18.5)$$

$$F(x + \Delta) \geq F(x), \quad \text{если } \Delta > 0. \quad (18.6)$$

Первые три свойства непосредственно следуют из определения (18.1), свойство (18.6) является следствием свойства (18.5), так как

$$F(x + \Delta) - F(x) = \mathbf{P}\{x \leq X < x + \Delta\} \geq 0,$$

поскольку вероятность не может быть отрицательной.

Для доказательства свойства (18.5) рассмотрим три события: событие A — выполняется неравенство $X < a$, событие B — выполняется неравенство $a \leq X < b$ и событие C — выполняется неравенство $X < b$. Очевидно, что $C = A + B$, причем события A и B несовместны.

Следовательно,

$$\mathbf{P}(C) = \mathbf{P}(A) + \mathbf{P}(B),$$

откуда получаем

$$\mathbf{P}(B) = \mathbf{P}(C) - \mathbf{P}(A)$$

или, переписывая в явном виде,

$$\mathbf{P}\{a \leq X < b\} = \mathbf{P}\{X < b\} - \mathbf{P}\{X < a\},$$

что эквивалентно (18.5). Таким образом, функция $F(x)$ является неубывающей функцией своего аргумента.

Как было показано в § 16, справедливо следующее утверждение. Если X — дискретная случайная величина, то $F(x)$ — ступенчатая функция, непрерывная слева и имеющая скачки в точках $x = x_j$, где x_j ($j = \overline{1, n}$) — возможные значения случайной величины. Величины скачков равны вероятностям этих значений $p_j = \mathbf{P}(X = x_j)$.

Предположим теперь, что $F(x)$ — непрерывная случайная величина. Выберем некоторый интервал $[x, x + dx)$. В соответствии со свойствами (18.5) и (18.6) и определением непрерывной случайной величины, имеем

$$\mathbf{P}\{x \leq X < x + dx\} = F(x + dx) - F(x) = O(dx) \geq 0. \quad (18.7)$$

Отсюда следует, что функция распределения непрерывной случайной величины является непрерывно дифференцируемой функцией.

Обозначим производную от $F(x)$ через $f(x)$, то есть положим

$$f(x) = \frac{dF(x)}{dx}. \quad (18.8)$$

Будем называть $f(x)$ *плотностью распределения*².

Устремляя в равенстве (18.7) dx к нулю, получим

$$\mathbf{P}\{X = x\} = 0, \quad (18.9)$$

² Иногда используют термин «дифференциальный закон распределения».

то есть вероятность того, что непрерывная случайная величина примет какое-то определенное значение, равна нулю. Этот результат не противоречит практике, так как практический интерес всегда представляет не вероятность получения непрерывной случайной величины некоторого фиксированного значения, а вероятность того, что случайная величина попадет в некоторый интервал dx . Будем называть эту вероятность элементом вероятности и обозначать dP . В соответствии с (18.7) и (18.8) элемент вероятности, соответствующий интервалу $(x, x + dx)$, определяется формулой

$$dP = f(x)dx. \quad (18.10)$$

Если случайная величина X смешанного типа, то ее функция распределения будет иметь скачки первого рода в точках, соответствующих дискретным значениям x_j , ($j = \overline{1, n}$), принимаемым случайной величиной с отличными от нуля вероятностями p_j , являясь непрерывной функцией в интервалах между этими значениями.

Таким образом, можно сформулировать следующее свойство функции распределения для величин смешанного типа. Функция распределения $F(x)$ случайной величины смешанного типа, принимающей дискретные значения x_j с вероятностями $p_j > 0$, ($j = \overline{1, n}$), может быть представлена в виде

$$F(x) = (1 - a)F_1(x) + aF_2(x), \quad (18.11)$$

где $F_1(x)$ — функция распределения непрерывной случайной величины, $F_2(x)$ — функция распределения дискретной случайной величины, принимающей значения x_j с вероятностями

$$p'_j = \frac{1}{a} p_j \quad (j = \overline{1, n}); \quad a = \sum_{j=1}^n p_j < 1.$$

Например, пусть случайная величина X может принимать два значения $x_1 = -1$ и $x_2 = 1$ с вероятностями $p_1 = \frac{1}{8}$ и $p_2 = \frac{3}{8}$, а кроме того, попадание ее значения равновозможно

в интервал $[0, 2]$. В этом случае $a = \frac{1}{2}$, $p'_1 = \frac{1}{4}$, $p'_2 = \frac{3}{4}$, а функции распределения $F_1(x)$, $F_2(x)$ и $F(x)$ будут иметь вид, представленный на рис 18.1.

Рассмотрим основные свойства плотности вероятности непрерывной случайной величины. Из определения плотности вероятности (18.8) и формулы (18.6) следует, что

$$f(x) \geq 0. \quad (18.12)$$

Интегрируя (18.8) по x от $-\infty$ до x , с учетом (18.3) получим

$$F(x) = \int_{-\infty}^{\infty} f(x_1) dx. \quad (18.13)$$

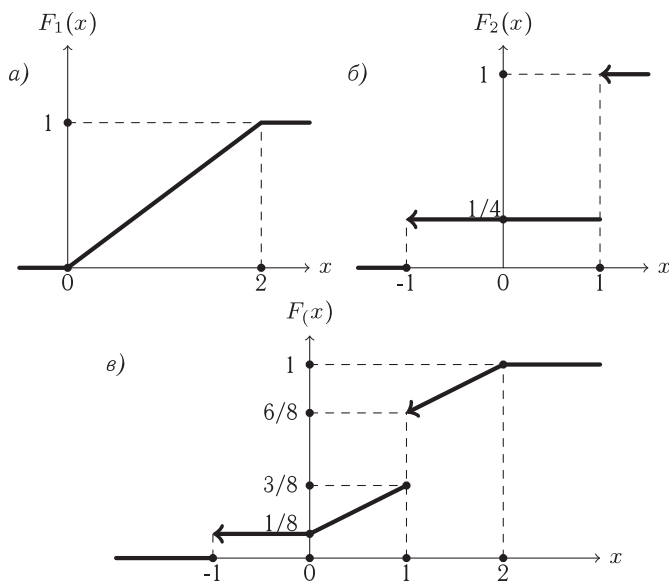


Рис. 18.1

Пример закона распределения смешанного типа:

а — равномерный закон в промежутке $[0, 2]$;

б — дискретный закон ($x_2 = -x_1 = 1$, $p_2 = 3/4$, $p_1 = 3/4$);

в — $F(x) = 1/2 F_1(x) + 1/2 F_2(x)$.

Подставляя последнее равенство в (18.5), имеем

$$\mathbf{P}\{a \leq X < b\} = \int_a^b f(x)dx. \quad (18.14)$$

Далее, положив $a = -\infty$, $b = \infty$, получим отсюда

$$\int_{-\infty}^{\infty} f(x)dx = 1. \quad (18.15)$$

Наконец, из (18.3) и (18.4) следует

$$f(\pm\infty) = 0. \quad (18.16)$$

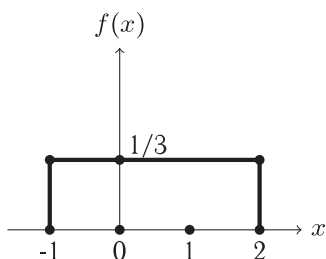


Рис. 18.2
Плотность вероятности
равномерного закона
распределения
в промежутке $[-1, 2]$

Плотность вероятности более удобна для характеристики случайной величины, чем функция распределения, так как она более чувствительна к изменению свойств случайной величины.

На рисунках 18.2, 18.3 и 18.4 в качестве примера приведено несколько типичных плотностей вероятности непрерывной случайной величины. Свойства случайных величин, имеющих такие

плотности вероятности, будут более подробно рассмотрены далее. Здесь только отметим, что общие свойства плотности вероятности допускают обращение ее ординат в бесконечность, однако при этом функция $f(x)$ должна быть интегрируема и площадь, ограниченная кривой $y = f(x)$ и осью абсцисс, в соответствии с (18.15) должна быть равной единице.

Понятие плотности вероятности, введенное для непрерывной случайной величины, можно обобщить также на дискретные и на смешанные случайные величины, если воспользоваться дельта-функцией Дирака³ $\delta(x)$. Дельта-функция, как

³ Дирак Поль Адриен Морис (1902–1984) — английский физик.

известно [6], принадлежит к классу обобщенных функций и определяется формулами

$$\delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \neq 0, \end{cases} \quad (18.17)$$

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

Определим для дискретной случайной величины, имеющей ряд распределения (x_j, p_j) , $(j = \overline{1, n})$, плотность вероятности формулой

$$f(x) = \sum_{j=1}^n p_j \delta(x - x_j). \quad (18.18)$$

Проверим, что функция распределения, найденная по этой плотности вероятности, будет иметь правильные значения.

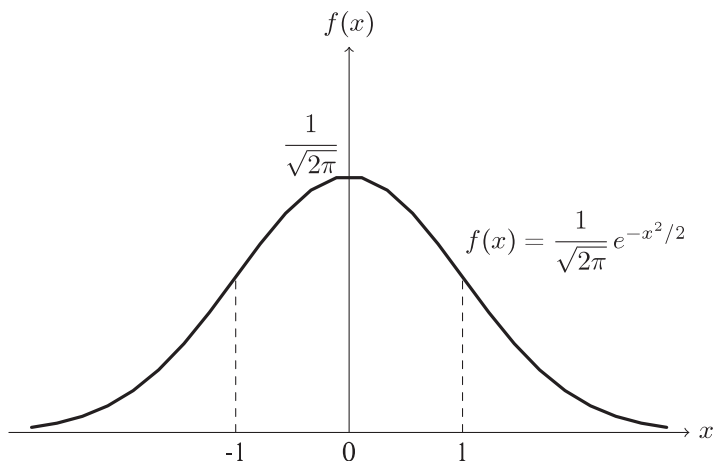


Рис. 18.3

Плотность вероятности стандартного нормального закона распределения

Проинтегрировав для этого (18.18) от $-\infty$ до x , с учетом (8.13), получим

$$F(x) = \sum_{j=1}^n p_j \int_{-\infty}^x \delta(x' - x_j) dx'. \quad (18.19)$$

При интегрировании дельта-функций те интегралы, для которых значения аргументов $x' - x_j$ на верхнем пределе будут положительными, обратятся в единицы. Остальные интегралы, в соответствии с (18.17), обратятся в нуль. Следовательно, вместо (18.19) можно записать

$$F(x) = \sum_{x_j < x} p_j, \quad (18.20)$$

где суммирование ведется только по тем значениям, для которых $x_j < x$. Формула (18.20) дает правильное значение функции распределения (см. § 16) и, следовательно, выражение (18.18) можно рассматривать, как плотность вероятности дискретной случайной величины.

Таким образом, учитывая (18.11), получим еще одно свойство плотности вероятности. Плотность вероятности непрерывной, дискретной и смешанной случайной величины может быть представлена в виде

$$f(x) = (1 - a)f_1(x) + \sum_{j=1}^n p_j \delta(x - x_j), \quad (18.21)$$

где $f_1(x)$ — плотность вероятности непрерывной случайной величины, $p_j = \mathbf{P}\{X = x_j\}$, а $a = \sum_{j=1}^n p_j \leq 1$. При $a = 1$ случайная величина X — дискретная, при $a = 0$ — непрерывная, а промежуточные значения $0 < a < 1$ отвечают смешанному типу случайной величины.

Таким образом, для величин любого типа их свойства одинаково полно характеризуются как функцией распределения, так и плотностью вероятности, между которыми имеется взаимно однозначное соответствие. Если свойства случайной величины заданы, мы будем говорить, что задан ее закон распределения. Так, плотности вероятности, изображенные на

рис. 18.2, 18.3 и 18.4 относятся к законам распределения, соответственно: равномерному, нормальному и закону арксинуса.

Теперь рассмотрим несколько примеров.

Пример 18.1. Определить значения постоянных a и b , при которых функция

$$\varphi(x) = a + b \cdot \arctg \frac{x}{a}, \quad (-\infty < x < \infty)$$

может быть функцией распределения.

Решение. Если $\varphi(x)$ является функцией распределения, то в соответствии с (18.3) и (18.4) должны выполняться равенства

$$\varphi(-\infty) = 0, \quad \varphi(\infty) = 1,$$

откуда получаем систему уравнений

$$a - \frac{\pi}{2} \cdot b, \quad a + \frac{\pi}{2} \cdot b = 1,$$

решая ее, находим

$$a = \frac{1}{2}, \quad b = \frac{1}{\pi},$$

таким образом, окончательно,

$$F(x) = \frac{1}{2} \left(1 + \frac{2}{\pi} \cdot \arctg(2x) \right),$$

а соответствующая плотность вероятности выражается в виде

$$f(x) = \frac{1}{\pi(x^2 + 0,25)}.$$

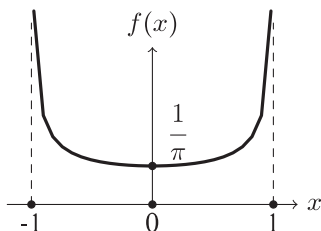


Рис. 18.4
Плотность вероятности
закона распределения
арксинуса

Пример 18.2. Сигнал Y , поступающий на выход нелинейного звена автоматического регулирования, связан с входным сигналом X соотношениями

$$Y = \begin{cases} -\alpha, & X < -\alpha, \\ X, & -\alpha \leq X \leq \alpha, \\ \alpha, & X > \alpha, \end{cases}$$

где $-\beta \leq X \leq \beta$, причем случайная величина X имеет постоянную плотность вероятности в интервале $(-\beta, \beta)$, ($\beta > \alpha$). Определить тип случайной величины Y , найти ее функцию распределения и плотность вероятности.

Решение. Так как X принимает значения только в интервале $(-\beta, \beta)$, а плотность вероятности $f(x)$ в данном интервале имеет постоянное значение, то

$$f(x) = \begin{cases} 0, & x < -\beta, \\ \frac{1}{2\beta}, & -\beta \leq x \leq \beta, \\ 0, & x > \beta. \end{cases}$$

Здесь постоянное значение плотности $f(x)$ в интервале $(-\beta, \beta)$ должно быть выбрано так, чтобы выполнялось условие нормировки (18.15).

Случайная величина Y принимает значения $-\alpha$ и $+\alpha$ в тех случаях, когда случайная величина X выйдет за пределы интервала $(-\alpha, \alpha)$ в меньшую и, соответственно, в большую сторону. Для вероятностей p_1 и p_2 этих событий имеем следующие выражения:

$$p_1 = \int_{-\beta}^{-\alpha} f(x) dx = \frac{\beta - \alpha}{2\beta}; \quad p_2 = \int_{\alpha}^{\beta} f(x) dx = \frac{\beta - \alpha}{2\beta}.$$

В интервале значений Y , лежащих в промежутке $(-\alpha, \alpha)$, для элемента вероятности получаем

$$dP = \mathbf{P}\{y \leq Y \leq y + dy\} = \int_y^{y+dy} f(x) dx = \frac{1}{2\beta} dy.$$

Следовательно, случайная величина Y является величиной смешанного типа, а ее функция распределения $F(y)$ определяется равенством

$$F(y) = \begin{cases} 0, & y < -\alpha, \\ \frac{x + \beta}{2\beta}, & -\alpha \leq y < \alpha, \\ 1, & y \geq \alpha. \end{cases}$$

График функции $F(y)$ представлен на рисунке 18.5. Плотность распределения вероятности $f(y)$ в соответствии с (18.11) и (18.18) будет иметь вид

$$f(y) = \frac{\beta + \alpha}{2\beta} f_1(y) + \frac{\beta - \alpha}{2\beta} \delta(x + \alpha) + \frac{\beta - \alpha}{2\beta} \delta(x - \alpha),$$

где $f_1(y)$ — плотность вероятности непрерывной случайной величины

$$f_1(y) = \begin{cases} \frac{1}{2\alpha}, & |x| \leq \alpha, \\ 0, & |x| > \alpha. \end{cases}$$

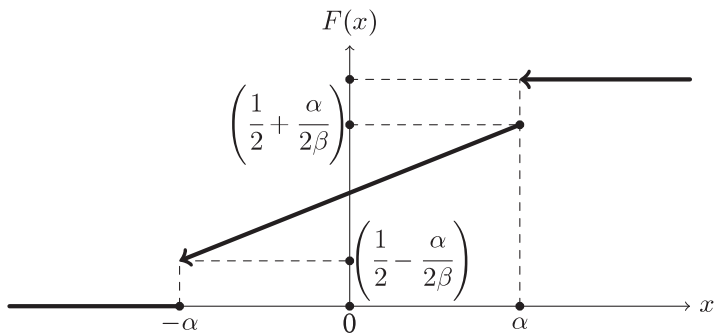


Рис. 18.5

К примеру 18.2. График функции распределения

§ 19. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ СЛУЧАЙНОЙ ВЕЛИЧИНЫ. МОМЕНТЫ И ДРУГИЕ ПАРАМЕТРЫ ЗАКОНА РАСПРЕДЕЛЕНИЯ

Пусть дискретная случайная величина X задана ее рядом распределения (x_j, p_j) , $(j = \overline{1, n})$. Введем в рассмотрение неслучайную величину $\mathbf{M}[X]$, называемую *математическим ожиданием* случайной величины X и определяемую как сумма произведений возможных значений этой величины на соответствующие им вероятности, то есть формулой

$$\mathbf{M}[X] = \sum_{j=1}^n x_j p_j. \quad (19.1)$$

Как будет показано далее при рассмотрении предельных теорем теории вероятностей, если провести большое число опытов, в каждом из которых реализуется значение случайной величины X , то среднее арифметическое полученных значений мало уклоняется от математического ожидания. Однако между математическим ожиданием и средним арифметическим имеется принципиальное отличие, заключающееся в том, что математическое ожидание является величиной неслучайной, в то время как среднее арифметическое — величина случайная. Иногда математическое ожидание называют средним значением случайной величины. Так как этот термин по звучанию близок к термину «среднее арифметическое», во избежание ошибок, в данном пособии термин «среднее» будет применяться только к экспериментальным средним.

Определение математического ожидания, данное выше для дискретной величины, может быть легко обобщено на случайные величины любого типа, если воспользоваться понятием элемента вероятности $dP = f(x)dx$ и от суммирования по всем возможным значениям в формуле (19.1) перейти к интегрированию. В результате получим общую формулу, справедливую для случайной величины любого типа

$$\mathbf{M}[X] = \int_{-\infty}^{\infty} x f(x) dx. \quad (19.2)$$

В том случае, когда возможные значения случайной величины ограничены, интеграл (19.2) имеет конечное значение. Однако, если случайная величина может принимать и бесконечные значения, интеграл (19.2) может обратиться в бесконечность. В этом случае говорят, что математическое ожидание случайной величины *не существует*.

При доказательстве ряда теорем о математическом ожидании существенно, чтобы интеграл (19.2) (или сумма (19.1)) не только сходилась, но и сходилась абсолютно. Поэтому обычно принимается, что случайная величина X имеет математическое ожидание только в том случае, если выполняется условие

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty. \quad (19.3)$$

Например, если плотность вероятности задается равенством

$$f(x) = \frac{1}{\pi(x^2 + 1)}, \quad (19.4)$$

то интеграл (19.2), рассматриваемый в смысле главного значения, обращается в нуль, однако, интеграл (19.3) расходится и, следовательно, случайная величина, имеющая плотность вероятности (19.4), не имеет математического ожидания.

Применение понятия математического ожидания в том случае, когда интеграл (19.3) расходится, недопустимо и может привести к неверным результатам. В этом отношении является показательным так называемый парадокс «петербургской

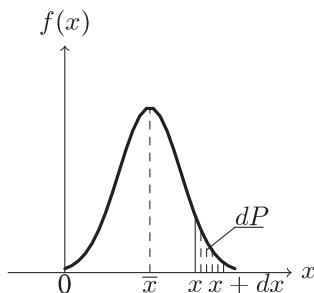


Рис. 19.1
К выражению
математического ожидания
через плотность
вероятности

игры»⁴, заключающийся в следующем. Павел делает ставку в один рубль, после чего бросается монета. Если выпадает решка, игра на этом останавливается и Петр получает ставку, если же выпадает орел — Павел удваивает ставку и снова бросается монета на тех же условиях. Требуется определить, какой предварительный взнос N должен уплатить Петр Павлу для того, чтобы игра была бы «безобидной», то есть чтобы математическое ожидание выигрыша каждого из игроков равнялось нулю? Очевидно, что N равно математическому ожиданию выигрыша Петра.

Петр может выиграть 2^{j-1} рублей с вероятностью 2^{-j} , ($j = \overline{1, \infty}$). Следовательно, математическое ожидание выигрыша Петра

$$N = \sum_{j=1}^{\infty} 2^{j-1} 2^{-j} = \infty.$$

Таким образом, если пользоваться в этом случае понятием «безобидной» игры, введенном в предположении, что соответствующие математические ожидания существуют, мы получим явно абсурдный вывод, что за право играть в данную сомнительную игру Петру следует уплатить бесконечно большую сумму денег.

Введенное выше понятие математического ожидания позволяет вычислить математическое ожидание не только самой случайной величины, но и любой (однозначной) функции этой величины. Рассмотрим некоторую функцию $\varphi(X)$ случайной величины X . По определению математического ожидания, для нахождения $\mathbf{M}[\varphi(X)]$ нужно взять сумму произведений значений этой функции $\varphi(x)$ на соответствующие вероятности, то есть на вероятности $f(x)dx$ попадания случайной величины X в интервал $[x, x + dx)$. Таким образом, получаем выражение

$$\mathbf{M}[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \quad (19.5)$$

⁴ Эта задача, предложенная Николаем Бернулли (брат Якова Бернулли), была опубликована впервые в Записках Петербургской академии.

где интеграл предполагается сходящимся абсолютно. В случае, когда величина X дискретная, интеграл (19.5) превращается в сумму

$$\mathbf{M}[\varphi(X)] = \sum_{j=1}^n \varphi(x_j) p_j,$$

которая также должна сходиться абсолютно.

Математическое ожидание случайной величины имеет простую механическую аналогию, которая часто позволяет находить его значение без специальных вычислений. Будем рассматривать плотность вероятности $f(x)$ как плотность некоторой физической величины, например, веса стержня, ориентированного вдоль оси x . В этом случае элемент вероятности $dP = f(x)dx$ будет давать вес элемента стержня dx , удаленного от начала координат на расстояние x (или, что эквивалентно, элемент площади под кривой $y = f(x)$, заштрихованный на рис. 19.1).

Определим координату $\bar{x}$ центра тяжести такого стержня. Согласно известной из механики формуле для положения центра тяжести [18]

$$\bar{x} = \frac{\int_{-\infty}^{\infty} x f(x) dx}{\int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^{\infty} x f(x) dx, \quad (19.6)$$

поскольку интеграл, стоящий в знаменателе, равен единице по свойству нормировки плотности вероятности. Сравнивая (19.6) с (19.2), замечаем, что

$$\mathbf{M}[X] = \bar{x}. \quad (19.7)$$

Поэтому для обозначения математического ожидания случайной величины в дальнейшем мы будем использовать также малую букву того же наименования с чертой сверху⁵. (Например, $\bar{x}$, $\bar{y}$, $\bar{z}$ будут обозначать математические ожидания случайной величины X , Y , Z , соответственно).

⁵ Для обозначения математического ожидания иногда, особенно в переводной литературе, используется также символ $\mathbf{E}[X]$.

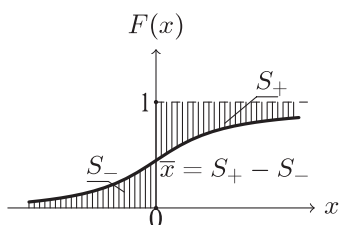


Рис. 19.2
К выражению
математического ожидания
через функцию распределения

Полученная механическая аналогия показывает, что в том случае, когда кривая $y = f(x)$ симметрична относительно некоторой оси, абсцисса оси симметрии как раз и дает математическое ожидание $\bar{x}$.

Формуле (19.2) можно придать другой вид, выразив математическое ожидание не через плотность вероятности, а через функцию распределения $F(x)$.

Представив для этого интеграл (19.5) в эквивалентной форме

$$\bar{x} = \int_{-\infty}^0 x d[F(x)] + \int_0^{\infty} x d[F(x) - 1] \quad (19.8)$$

и интегрируя каждое слагаемое в отдельности по частям, замечаем, что внеинтегральные члены обращаются в нуль и, следовательно,

$$\mathbf{M}[X] = \bar{x} = \int_0^{\infty} [1 - F(x)] dx - \int_{-\infty}^0 F(x) dx. \quad (19.9)$$

Таким образом, математическому ожиданию $\bar{x}$ можно дать и следующую геометрическую интерпретацию: математическое ожидание $\bar{x}$ равно разности площадей области, расположенной в правой полуплоскости между кривой $y = F(x)$ и прямой $y = 1$, и области, расположенной в левой полуплоскости между кривой $y = F(x)$ и осью абсцисс (рис. 19.2). Очевидно, что математическое ожидание случайной величины не дает никакой информации о форме кривой $f(x)$, а только показывает положение этой кривой относительно начала координат. Такие параметры закона распределения случайной величины иногда называют *параметрами положения*.

Для характеристики формы $f(x)$ необходимы другие параметры (*параметры формы*), которые могут быть выбраны различным образом. Наиболее удобными параметрами такого типа являются начальные и центральные моменты закона распределения различных порядков, определяемые как математические ожидания соответствующих степеней случайной величины.

Начальным моментом m_k произвольного порядка k , ($k = \overline{1, \infty}$) называется математическое ожидание k -й степени случайной величины; центральным моментом μ_k порядка k , ($k = \overline{2, \infty}$) называется математическое ожидание k -й степени разности случайной величины и ее математического ожидания. Таким образом, учитывая (19.5), для случайной величины X имеем следующие начальные и центральные моменты:

$$m_k = \mathbf{M}(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx, \quad (k = \overline{1, \infty}), \quad (19.10)$$

$$\mu_k = \mathbf{M}[(X - \bar{x})^k] = \int_{-\infty}^{\infty} (x - \bar{x})^k f(x) dx, \quad (k = \overline{2, \infty}). \quad (19.11)$$

Начальный момент нулевого порядка тождественно равен единице, а при $k = 1$ начальный момент совпадает с математическим ожиданием. Следовательно, для обозначения математического ожидания случайной величины X мы имеем три эквивалентных обозначения, которые будут по мере надобности использоваться далее

$$\mathbf{M}(X) = \bar{x} = m_1. \quad (19.12)$$

Центральный момент нулевого порядка тождественно равняется единице, а центральный момент первого порядка тождественно равняется нулю, так как положив в (19.11) $k = 1$, получим $\mu_1 = \bar{x} - \bar{x} = 0$.

В том случае, когда необходимо явно отметить наименование случайной величины, для которой вычисляются моменты,

в обозначении моментов m_k и μ_k вводится верхний индекс. Так, например, $m_k^{(x)}, m_k^{(y)}, \mu_k^{(x)}, \mu_k^{(y)}$ обозначают начальные и центральные моменты случайных величин X и Y соответственно. Иногда также пишут $m_k(X), \mu_k(X)$, если, например, момент вычисляется для функции от случайной величины.

Поскольку моменты определяются как математические ожидания степеней случайной величины, для существования момента необходима абсолютная сходимость соответствующих интегралов (19.10) и (19.11). Для моментов четного порядка это требование совпадает с требованием сходимости соответствующих интегралов, для моментов нечетного порядка эти интегралы должны сходиться абсолютно.

Между центральными и начальными моментами существует ряд алгебраических соотношений, позволяющих выражать один вид моментов через другой. Для вывода этих соотношений введем новое обозначение. Будем называть разность $(X - \bar{x})$ случайной величины и ее математического ожидания *центрированной случайной величиной* и помечать вверху кружком $\overset{\circ}{X}$, то есть положим

$$\overset{\circ}{X} = X - \bar{x}. \quad (19.13)$$

Очевидно, что для центрированной случайной величины ее начальные и центральные моменты совпадают и равны центральному моменту исходной случайной величины соответствующего порядка.

Представив в (19.10) X^k в виде $[(X - \bar{x}) + \bar{x}]^k$, раскрывая квадратные скобки и разбивая интеграл от суммы на сумму интегралов, получим

$$m_k = \sum_{j=0}^k C_k^j \int_{-\infty}^{\infty} (x - \bar{x})^{k-j} \cdot \bar{x}^j f(x) dx = \sum_{j=0}^k C_k^j \cdot \mu_{k-j} \cdot \bar{x}^j,$$

или учитывая, что $\mu_0 = 1, \mu_1 = 0$, будем иметь

$$m_k = \bar{x}^k + \sum_{j=0}^{k-2} C_k^j \cdot \bar{x}^j \cdot \mu_{k-j}. \quad (19.14)$$

Аналогичным образом, раскрывая скобки в интеграле (19.11) и интегрируя почленно полученную сумму, находим

$$\mu_k = \sum_{j=0}^k (-1)^j \cdot C_k^j \cdot \bar{x}^j \cdot m_{k-j}. \quad (19.15)$$

Формула (19.14) позволяет выразить начальные моменты любого порядка через центральные моменты и математическое ожидание $\bar{x}$, а формула (19.15) — выразить центральные моменты через начальные.

Из моментов более высокого порядка, чем первый, наибольшее значение имеет второй центральный момент, определяемый в соответствии с (19.10) формулой

$$\mu_2 = \mathbf{M}(\overset{\circ}{X}^2) = \int_{-\infty}^{\infty} (x - \bar{x})^2 \cdot f(x) dx. \quad (19.16)$$

Если продолжить механическую аналогию, использованную выше при рассмотрении математического ожидания $\bar{x}$, то можно констатировать, что второй центральный момент аналогичен моменту инерции стержня, масса которого распределена с плотностью $f(x)$, вычисленному относительно центра тяжести стержня [18]. Следовательно, при увеличении μ_2 «масса» стержня имеет тенденцию распределяться дальше от его центра тяжести — ординаты плотности $f(x)$ в области значений $x \approx \bar{x}$ будут уменьшаться за счет увеличения ординат на больших удалениях от $\bar{x}$. С вероятностной точки зрения это означает, что при росте μ_2 рассеивание значений случайной величины около ее математического ожидания будет увеличиваться. Поэтому второй центральный момент называют также дисперсией и обозначают $\mathbf{D}(X)$. Очевидно, что дисперсия имеет размерность, квадратичную относительно размерности X . Это не всегда удобно при расчетах. Поэтому вместо дисперсии чаще рассматривают положительный квадратный корень из дисперсии $\sigma = \sqrt{\mathbf{D}(X)}$, называемый средним квадратическим отклонением или стандартом.

Таким образом, для второго центрального момента также, как и для математического ожидания, используется несколько

обозначений

$$\mu_2 = \mathbf{M}(\overset{\circ}{X}^2) = \mathbf{D}(X) = \sigma^2. \quad (19.17)$$

Положив в формуле (19.14) $k = 2$, получим связь между дисперсией и вторым начальным моментом

$$m_2 = \mathbf{D}(X) + \bar{x}^2, \quad D(X) = m_2 - \bar{x}^2. \quad (19.18)$$

Из моментов более высокого порядка наиболее часто используются моменты третьего и четвертого порядков. Для этих моментов формулы (19.14) и (19.15) дают следующие соотношения между центральными и начальными моментами:

$$\mu_3 = m_3 - 3 \cdot \bar{x} \cdot m_2 + 2 \cdot \bar{x}^3, \quad (19.19)$$

$$m_3 = \mu_3 - 3 \cdot \bar{x} \cdot \mu_2 + \bar{x}^3, \quad (19.20)$$

$$\mu_4 = m_4 - 4 \cdot \bar{x} \cdot m_3 + 6 \cdot \bar{x}^2 \cdot m_2 - 3 \cdot \bar{x}^4, \quad (19.21)$$

$$m_4 = \mu_4 - 4 \cdot \bar{x} \cdot \mu_3 + 6 \cdot \bar{x}^2 \cdot \mu_2 + \bar{x}^4. \quad (19.22)$$

Центральные моменты третьего и четвертого порядков являются более тонкими характеристиками формы кривой $y = f(x)$, чем дисперсия. Так как эти моменты являются величинами размерными, то вместо них удобно перейти к безразмерным величинам Sk и Ex , называемым коэффициентом скошенности (асимметрии) и эксцессом, соответственно, и определяемым равенствами

$$Sk = \frac{\mu_3}{\sigma^3}, \quad (19.23)$$

$$Ex = \frac{\mu_4}{\sigma^4} - 3. \quad (19.24)$$

Отличие коэффициента Sk от нуля указывает на то, что кривая $f(x)$ не обладает симметрией (обратное утверждение не

верно — асимметричная кривая может иметь нулевой коэффициент скошенности). Эксцесс обращается в нуль для одного из наиболее употребительных законов — нормального закона распределения (или закона Гаусса — см. далее гл. 5). Отличие Ex от нуля является признаком, что данный закон распределения не является нормальным (конечно, Ex может равняться нулю и для некоторых законов, отличных от нормального).

Для дискретной целочисленной случайной величины X моменты любого порядка могут быть выражены через производящую функцию. Действительно, сравнивая формулу для производящей функции (16.7)

$$G(u) = \sum_{j=0}^n p_j \cdot u^j$$

с формулой для начального момента k -го порядка целочисленной случайной величины, которая в соответствии с (19.10) имеет вид

$$m_k = \sum_{j=0}^n j^k \cdot p_j, \quad (19.25)$$

замечаем, что

$$\begin{aligned} m_1 &= \left. \frac{dG(u)}{du} \right|_{u=1}, \\ m_2 &= \left. \frac{d}{du} \left[u \frac{dG(u)}{du} \right] \right|_{u=1}, \\ m_3 &= \left. \frac{d}{du} \left\{ u \frac{d}{du} \left[u \frac{dG(u)}{du} \right] \right\} \right|_{u=1}, \\ &\dots\dots\dots \end{aligned} \quad (19.26)$$

Таким образом для получения начального момента k -го порядка к производящей функции нужно последовательно k раз применить оператор дифференцирования по аргументу u и умножения на u и положить в окончательном результате $u = 1$.

Кроме моментов, в качестве параметров закона распределения могут быть выбраны и некоторые другие величины. Например, в качестве параметров положения может быть использована или *медиана* m , то есть абсцисса ординаты, делящей площадь под кривой $y = f(x)$ на две равные части, или же *мода* x_0 — абсцисса максимума кривой $y = f(x)$. Медиана определяется уравнением

$$\mathbf{P}\{X < m\} = \mathbf{P}\{X > m\}. \quad (19.27)$$

Для непрерывной случайной величины функция распределения $F(x)$ непрерывна, следовательно, уравнение (19.27) в этом случае всегда имеет решение. Однако это решение может быть неоднозначно, если при $F(x) = 1/2$ функция распределения имеет горизонтальную площадку, соответствующую невозможным значениям случайной величины.

Законы распределения, имеющие единственную моду, называют *унимодальными*. Мода x_0 существует не для любого закона распределения, так как плотность вероятности может иметь или несколько максимумов (*полимодальные* законы распределения), или не иметь вообще максимумов (*антимодальные* законы распределения). Например, закон с двумя модами называют *бимодальным*.

Для симметричного закона распределения непрерывной случайной величины в качестве меры рассеивания иногда используется половина ширины интервала, симметрично расположенного относительно математического ожидания, вероятность попадания в который равняется $1/2$. Эта величина E , называемая *срединным отклонением*⁶, определяется формулой

$$F(\bar{x} + E) - F(\bar{x} - E) = \frac{1}{2} \quad (19.28)$$

или эквивалентной ей (для симметричного закона распределения) формулой

$$F(\bar{x} + E) = 0,75. \quad (19.29)$$

⁶ Иногда употребляют также термин *вероятное отклонение*.

Наконец, в качестве параметров закона распределения можно использовать *квантили* ξ_p , определяемые уравнениями

$$F(\xi_p) = p, \quad (0 < p < 1). \quad (19.30)$$

Очевидно, что медиана есть квантиль порядка $1/2$, а срединное отклонение представляет собой разность квантилей порядка $3/4$ и $1/4$, поделенную на два

$$E = \frac{1}{2} \cdot (\xi_{3/4} - \xi_{1/4}). \quad (19.31)$$

Определенная последней формулой величина может быть использована и для несимметричных законов распределения.

В заключение заметим, что в качестве безразмерной меры рассеивания случайной величины также используют *коэффициент вариации*

$$\gamma = \frac{\sigma}{|\bar{x}|}. \quad (19.32)$$

Далее рассмотрим несколько примеров вычисления моментов случайных величин.

Пример 19.1. Производится серия n независимых испытаний, в каждом из которых событие A может появиться с вероятностью p (схема Бернулли). Определить математическое ожидание и дисперсию числа X появления события A в серии.

Решение. В соответствии с формулой (16.7) для целочисленной случайной величины X производящая функция будет иметь вид

$$G(u) = (pu + q)^n, \quad q = 1 - p.$$

Применяя формулы (19.24) и (19.16), находим

$$\bar{x} = m_1 = \left. \frac{dG(u)}{du} \right|_{u=1} = np(pu + q)^{n-1} \Big|_{u=1} = np,$$

$$m_2 = \left. \frac{d}{du} \left(u \frac{dG}{du} \right) \right|_{u=1} = [np(pu + q)^{n-1} + n(n-1)p^2(pu + q)^{n-1}] \Big|_{u=1} = npq + n^2 p^2,$$

$$\mathbf{D}(X) = \sigma^2 = m_2 - \bar{x}^2 = npq.$$

Учитывая формулы (19.16) и (19.17), имеем

$$\sigma^2 = \sum_{j=1}^n p_j q_j; \quad \mu_3 = \sum_{j=1}^n (q_j - 3p_j).$$

Следовательно, асимметрия

$$S_k = \frac{\sum_{j=1}^n (1 - 2p_j)p_j \cdot q_j}{\left(\sum_{j=1}^n p_j \cdot q_j\right)^{3/2}}.$$

При $p_j = p$, $q_j = 1 - p = q$

$$S_k = \frac{1 - 2p}{\sqrt{npq}},$$

то есть при увеличении числа испытаний асимметрия уменьшается.

Пример 19.2. В серии n независимых испытаний вероятность появления события A в испытании номер j равна p_j . Определить первые три начальных момента числа появления события A в этой серии.

Решение. Производящая функция в данном случае такова:

$$G(u) = \prod_{j=1}^n (p_j u + q_j).$$

Применение формулы (19.24) упрощается, если перейти к логарифмам. Логарифмируя, получим

$$\begin{aligned} \ln G(u) &= \sum_{j=1}^n \ln(p_j u + q_j), \\ \frac{dG(u)}{du} &= G(u) \cdot \sum_{j=1}^n \frac{p_j}{p_j u + q_j}, \\ \frac{d}{du} \left(u \frac{dG}{du} \right) &= \frac{dG(u)}{du} + u \cdot \frac{d^2 G(u)}{du^2}, \end{aligned}$$

$$\begin{aligned} \frac{d}{du} \left\{ u \frac{d}{du} \left[u \frac{dG(u)}{du} \right] \right\} &= \frac{dG(u)}{du} - u \frac{d^2 G(u)}{du^2} - u^2 \frac{d^3 G(u)}{du^3}, \\ \frac{d^2 G(u)}{du^2} &= \frac{dG(u)}{du} \cdot \sum_{j=1}^n \frac{p_j}{p_j u + q_j} - G(u) \sum_{j=1}^n \frac{p_j^2}{(p_j u + q_j)^2}, \\ \frac{d^3 G(u)}{du^3} &= \frac{d^2 G(u)}{du^2} \cdot \sum_{j=1}^n \frac{p_j}{p_j u + q_j} - \\ &- 2 \cdot \frac{dG(u)}{du} \cdot \sum_{j=1}^n \frac{p_j^2}{(p_j u + q_j)^2} + 2 \cdot G(u) \cdot \sum_{j=1}^n \frac{p_j^2}{(p_j u + q_j)^2}. \end{aligned}$$

Подставляя в последние равенства последовательно $u = 1$ и применяя формулу (19.26), получим

$$\begin{aligned} m_1 &= \left. \frac{dG(u)}{du} \right|_{u=1} = \sum_{j=1}^n p_j, \\ m_2 &= \left. \frac{d}{du} \left[u \frac{dG}{du} \right] \right|_{u=1} = \left(\sum_{j=1}^n p_j \right)^2 + \sum_{j=1}^n p_j q_j = m_1^2 + \sum_{j=1}^n p_j q_j, \\ m_3 &= \left. \frac{d}{du} \left\{ u \frac{d}{du} \left[u \frac{dG(u)}{du} \right] \right\} \right|_{u=1} = \\ &= \left[\left(\sum_{j=1}^n p_j \right)^2 + \sum_{j=1}^n p_j q_j \right] \sum_{l=1}^n p_l + \\ &+ 2 \cdot \sum_{j=1}^n p_j \cdot \sum_{l=1}^n p_l q_l + \sum_{j=1}^n (1 - 4p_j). \end{aligned}$$

Пример 19.3. Получить математическое ожидание, дисперсию и коэффициент вариации для случайной величины X , распределенной по равномерному закону, характеризуемому

$$x_k = k, \quad p_k = \frac{1}{n} \quad (k = \overline{1, n}).$$

Решение. Вначале составим производящую функцию

$$G(u) = \frac{1}{n} \cdot \sum_{j=1}^n u^j = \frac{1}{n} \cdot \frac{u(1-u^n)}{(1-u)}.$$

Применяя формулы (19.24) и (19.16), будем иметь

$$\bar{x} = m_1 = \left. \frac{dG(u)}{du} \right|_{u=1} = \frac{1}{n} \cdot \left. \sum_{j=1}^n j u^{j-1} \right|_{u=1} = \frac{1}{n} \cdot \sum_{j=1}^n j.$$

Воспользовавшись методом математической индукции, нетрудно показать, что

$$\sum_{j=1}^n j = \frac{n(n+1)}{2},$$

поэтому

$$\bar{x} = \frac{(n+1)}{2}.$$

Аналогично вычисляется второй начальный момент

$$\begin{aligned} m_2 &= \left. \frac{d}{du} \left(u \frac{dG}{du} \right) \right|_{u=1} = \left. \frac{d}{du} \left(\frac{1}{n} \cdot \sum_{j=1}^n j u^j \right) \right|_{u=1} = \\ &= \frac{1}{n} \cdot \left. \sum_{j=1}^n j^2 u^{j-1} \right|_{u=1} = \frac{1}{n} \cdot \sum_{j=1}^n j^2. \end{aligned}$$

По индукции получаем, что

$$\sum_{j=1}^n j^2 = \frac{n(n+1)(2n+1)}{6},$$

откуда следует, что

$$m_2 = \frac{n(n+1)(2n+1)}{6},$$

а тогда по формуле (19.18) находим

$$\mathbf{D}(X) = \frac{n(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n^2-1)}{12}.$$

Для коэффициента вариации γ отсюда находим

$$\sigma = \sqrt{\mathbf{D}(X)} = \frac{\sqrt{n^2-1}}{2\sqrt{3}},$$

$$\gamma = \frac{\sigma}{\bar{x}} = \sqrt{\frac{(n-1)}{3(n+1)}}.$$

Приведенные примеры показывают, что непосредственное применение формул, выражающих начальные и центральные моменты через производящую функцию целочисленной случайной величины, может быть сопряжено с утомительными и громоздкими вычислениями. Их можно несколько сократить, если воспользоваться так называемыми *факториальными моментами*.

$$\nu_k = \mathbf{M}(X(X-1)\dots(X-k+1)), \quad (k = \overline{1, \infty}). \quad (19.33)$$

Такое наименование связано с тем, что произведение, стоящее под знаком математического ожидания (19.33), напоминает определение факториала.

Факториальные моменты целочисленной случайной величины элементарно выражаются через производные производящей функции. Действительно,

$$\nu_k = \sum_{j=0}^n j(j-1)\dots(j-k+1)P_j = \left. \frac{d^k G}{du^k} \right|_{u=1}. \quad (19.34)$$

При этом факториальные моменты чрезвычайно просто связаны с начальными. Согласно (19.33)

$$\begin{aligned} \nu_1 &= m_1, \\ \nu_2 &= m_2 - m_1, \\ \nu_3 &= m_3 - 3m_2 + 2m_1, \\ &\dots\dots\dots \end{aligned} \quad (19.35)$$

Разрешая бесконечную систему уравнений (19.35) относительно начальных моментов, находим

$$\begin{aligned} m_1 &= \nu_1, \\ m_2 &= \nu_2 + \nu_1, \\ m_3 &= \nu_3 + 3\nu_2 + \nu_1, \\ &\dots\dots\dots \end{aligned} \quad (19.36)$$

Если производные (19.34) легко вычисляются, то этот метод позволяет существенно упростить вычисления.

Пример 19.4. Получить дисперсию и математическое ожидание закона Пуассона

$$\mathbf{P}(X = k) = P_k = e^{-a} \cdot \frac{a^k}{k!}, \quad (k = \overline{0, \infty}).$$

Решение. Производящая функция закона Пуассона

$$G(u) = \sum_{k=0}^{\infty} e^{-a} \cdot \frac{a^k}{k!} \cdot u^k = e^{-a} \cdot \sum_{k=0}^{\infty} \frac{(au)^k}{k!} = e^{-a} \cdot e^{au} = e^{a(u+1)}$$

очень подходит для применения описанного только что метода, поскольку

$$\left. \frac{d^k G}{du^k} \right|_{u=1} = a^k \cdot e^{a(u-1)} \Big|_{u=1} = a^k.$$

Следовательно, в данном случае

$$\nu_k = a^k,$$

а тогда по формуле (19.36) сразу получим

$$\begin{aligned} m_1 &= a, \\ m_2 &= a^2 + a, \end{aligned}$$

откуда следует, что

$$\mathbf{D}(X) = m_2 - m_1^2 = a = m_1.$$

Таким образом, оба вычисляемых момента для данного закона совпадают.

Пример 19.5. Рассмотрим геометрический закон распределения

$$\mathbf{P}(X = k) = (1 - q)q^k, \quad (k = \overline{0, \infty}),$$

в котором для удобства от ранее использованного параметра p мы перешли к параметру $q = 1 - p$, ($0 < q < 1$). Подсчитаем коэффициент вариации этого закона.

Решение. Имеем для производящей функции

$$\sigma(u) = \sum_{k=0}^{\infty} (1 - q)q^k u^k = \sum_{k=0}^{\infty} (1 - q)(qu)^k = \frac{1 - q}{1 - qu}.$$

Дифференцируя это выражение, находим

$$\nu_k = \left. \frac{(1 - q)q^k k!}{(1 - qu)^{k+1}} \right|_{u=1} = \frac{q^k \cdot k!}{(1 - q)^k},$$

откуда

$$m_1 = \frac{q}{1 - q},$$

$$m_2 = \nu_2 + \nu_1 = \frac{2q^2}{(1 - q)^2} + \frac{q}{(1 - q)},$$

$$D(X) = m_2 - m_1^2 = \frac{q^2}{(1 - q)^2} + \frac{q}{(1 - q)} = \frac{q}{(1 - q)^2} = \sigma^2.$$

Следовательно,

$$\gamma = \frac{\sigma}{\bar{x}} = \frac{\sqrt{q}(1 - q)}{(1 - q)q} = \frac{1}{\sqrt{q}},$$

то есть коэффициент вариации такого закона всегда больше единицы, так как по определению параметр $q < 1$.

§ 20. ХАРАКТЕРИСТИЧЕСКАЯ ФУНКЦИЯ. СЕМИИНВАРИАНТЫ

Характеристической функцией $E(u)$ случайной величины X называется математическое ожидание e^{iuX} , то есть

$$E(u) = M(e^{iuX}), \quad (20.1)$$

где u — вещественный аргумент, а i — мнимая единица.

Характеристическая функция, наряду с плотностью вероятности и функцией распределения, так же, как и они, полностью определяет свойства случайной величины и является ее универсальной характеристикой в том смысле, что она применима для случайных величин любого типа.

В соответствии с правилом вычисления математического ожидания функции случайной величины (19.5), вместо (20.1), можно записать

$$E(u) = \int_{-\infty}^{\infty} e^{iux} f(x) dx, \quad (20.2)$$

где $f(x)$ — плотность вероятности случайной величины X . Если X — величина дискретная, заданная рядом распределения (x_j, p_j) , $(j = \overline{1, n})$, то интеграл (20.2) превращается в сумму и формула для характеристической функции примет вид

$$E(u) = \sum_{j=1}^n p_j e^{iux_j}. \quad (20.3)$$

Наконец, если X — целочисленная случайная величина, то характеристическая функция может быть выражена через производящую функцию $G(u)$. Действительно, сравнивая в этом случае выражение для характеристической функции

$$E(u) = \sum_{j=0}^n p_j e^{iuj}, \quad (20.4)$$

с формулой для производящей функции

$$G(u) = \sum_{j=0}^n p_j u^j, \quad (20.5)$$

замечаем, что

$$E(u) = G(e^{iu}). \quad (20.6)$$

Вернемся к определению характеристической функции в общем случае. Формула (20.2) показывает, что характеристическая функция является интегральным преобразованием Фурье плотности вероятности $f(x)$. Следовательно, функция $E(u)$ обладает рядом общих свойств, присущих интегральному преобразованию Фурье [28], а также некоторыми дополнительными специфическими свойствами, связанными со свойствами плотности вероятности $f(x)$.

Из теории интегралов Фурье известно, что интеграл (20.2) существует для любого вещественного u и является непрерывной функцией u , если функция $f(x)$ абсолютно интегрируема. Плотность вероятности $f(x) \geq 0$, причем

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (20.7)$$

Следовательно, характеристическая функция существует для случайной величины любого типа, причем имеет место не только прямое, но и обратное преобразование Фурье, то есть из формулы (20.2) следует формула, позволяющая, зная E , восстановить по ней f в виде

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} E(u) du. \quad (20.8)$$

Отметим некоторые общие свойства характеристической функции. Подставив в интеграл (20.2) $u = 0$ и учитывая (20.7), получим, что всегда

$$E(0) = 1. \quad (20.9)$$

Далее из очевидного неравенства

$$\left| \int_{-\infty}^{\infty} e^{iux} f(x) dx \right| \leq \int_{-\infty}^{\infty} |e^{iux}| f(x) dx = 1$$

следует, что

$$|E(u)| \leq 1, \quad (20.10)$$

то есть характеристическая функция ограничена.

Характеристическая функция линейной функции от X может быть выражена через характеристическую функцию X . Действительно, положив

$$Y = aX + b \quad (20.11)$$

и отмечая характеристические функции X и Y соответствующими индексами, применяя формулу (20.2), имеем

$$\begin{aligned} E_y(u) &= \int_{-\infty}^{\infty} e^{iu(ax+b)} f(x) dx = \\ &= e^{iub} \cdot \int_{-\infty}^{\infty} e^{i(au)x} f(x) dx = e^{iub} E_x(au), \end{aligned}$$

то есть

$$E_y(u) = e^{iub} E_x(au). \quad (20.12)$$

Таким образом, прибавление к случайной величине неслучайного слагаемого b приводит к появлению у характеристической функции экспоненциального множителя e^{iub} , а умножение случайной величины на постоянный коэффициент a приводит к появлению этого коэффициента в качестве множителя у аргумента характеристической функции исходной случайной величины.

Если плотность вероятности $f(x)$ — четная функция своего аргумента, то характеристическая функция является вещественной⁷, так как произведя замену переменной интегрирования x на $x_1 = -x$, получим

$$E^*(u) = \int_{-\infty}^{\infty} e^{-iux} f(x) dx = \int_{-\infty}^{\infty} e^{iux_1} f(x_1) dx_1 = E(u). \quad (20.13)$$

⁷ Комплексно-сопряженные величины мы условимся отмечать звездочкой, оставляя черту сверху для обозначения математического ожидания.

Очевидно, что справедливо и обратное утверждение: если $E(u)$ вещественная, то плотность вероятности $f(x)$ — четная функция.

Кроме перечисленных выше, характеристическая функция удовлетворяет еще целому ряду важных свойств, наличие которых ограничивает класс функций, которые могут рассматриваться в качестве характеристических функций случайной величины. Например, характеристическая функция является положительно определенной в том смысле, что для любой функции $\eta(u)$ и любых пределов интегрирования (a, b) , для которых имеет смысл интеграл в формуле (20.14), справедливо неравенство

$$\int_a^b \int_a^b \eta^*(u_1) \eta(u_2) E(u_2 - u_1) du_1 du_2 \geq 0. \quad (20.14)$$

Действительно, рассмотрим заведомо неотрицательное выражение

$$\left| \int_a^b e^{iuX} \eta(u) du \right|^2 = \int_a^b \int_a^b \eta^*(u_1) \eta(u_2) e^{i(u_2 - u_1)X} du_1 du_2. \quad (20.15)$$

Умножая двойной интеграл в (20.15) при $X = x$ на $f(x)$ и интегрируя по x в бесконечных пределах, получим интеграл (20.14), который не может быть отрицательным, поскольку математическое ожидание неотрицательной величины всегда само неотрицательно.

Воспользуемся методом характеристических функций для получения ряда вспомогательных формул, относящихся к дельта-функции, которая дальше будет широко использоваться.

Рассмотрим центрированную случайную величину $\overset{\circ}{X}$, плотность вероятности которой отлична от нуля только на интервале $[-a, a]$. При $a \rightarrow 0$ отрезок оси абсцисс, на котором плотность $f(x) > 0$ стягивается в точку, а так как при этом условие нормировки плотности вероятности должно

сохраняться, то $f(0) \rightarrow \infty$. Таким образом, при $a \rightarrow 0$ плотность вероятности $f(x)$ превращается в дельта-функцию и мы получим

$$\lim_{a \rightarrow 0} f(x) = \delta(x),$$

$$\delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \neq 0, \end{cases} \quad (20.16)$$

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

Предельное значение характеристической функции $E(u)$ при $a \rightarrow 0$ согласно (20.2) при этом будет таким:

$$\lim_{a \rightarrow 0} E(u) = \int_{-\infty}^{\infty} e^{iux} [\lim_{a \rightarrow 0} f(x)] dx = \int_{-\infty}^{\infty} e^{iux} \delta(x) dx = 1. \quad (20.17)$$

Воспользовавшись равенством (20.8), получим

$$\lim_{a \rightarrow 0} f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} [\lim_{a \rightarrow 0} E(u)] du,$$

что с учетом (20.16) и (20.17) дает

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} du.$$

Так как $f(x)$ — вещественная функция, то полученное интегральное представление для дельта-функции можно переписать в виде

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iux} du. \quad (20.18)$$

Из этого равенства следует, что $\delta(x)$ — четная функция.

Производную s -го порядка от дельта-функции $\delta(x)$ можно понимать как предел производной s -го порядка от $f(x)$ при $a \rightarrow 0$. Воспользовавшись равенством (20.8), находим

$$\frac{d^s f(x)}{dx^s} = \frac{(-1)^s}{2\pi} \int_{-\infty}^{\infty} u^s e^{-iux} E(u) du. \quad (20.19)$$

Так как дельта-функция и ее производные вещественны, то для производной s -го порядка от дельта-функции имеет место следующее интегральное представление:

$$\delta^{(s)}(x) = \frac{i^s}{2\pi} \int_{-\infty}^{\infty} u^s e^{iux} du. \quad (20.20)$$

Более строгое доказательство формулы (20.20) можно найти в специальной литературе по обобщенным функциям [6].

Воспользуемся полученными формулами для дельта-функции для доказательства еще одного свойства характеристической функции, справедливого только для непрерывных случайных величин. Докажем, что для непрерывной случайной величины характеристическая функция обращается в нуль при $|u| \rightarrow \infty$, то есть

$$\lim_{|u| \rightarrow \infty} E(u) = 0. \quad (20.21)$$

Для этого рассмотрим $\int_{-\infty}^{\infty} |E(u)|^2 du$. Подставляя вместо $|E(u)|$ интеграл (20.2) и рассматривая произведение интегралов как двойной интеграл, получим

$$\int_{-\infty}^{\infty} |E(u)|^2 du = \iint_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iu(x_1+x_2)} f(x_1) f(x_2) dx_1 dx_2. \quad (20.22)$$

Выполнив вначале интегрирование по u и, учитывая формулу (20.18), имеем

$$\int_{-\infty}^{\infty} |E(u)|^2 2\pi \cdot \iint_{-\infty}^{\infty} f(x_1) f(x_2) \delta(x_2 - x_1) dx_1 dx_2. \quad (20.23)$$

Так как для любой непрерывной в точке $(x_2 - x_1)$ функции $\varphi(x_2)$ справедливо равенство

$$\int_{-\infty}^{\infty} \varphi(x_2) \delta(x_2 - x_1) dx_2 = \varphi(x_1), \quad (20.24)$$

то одно интегрирование в (20.23) можно выполнить, после чего получим

$$\int_{-\infty}^{\infty} |E(u)|^2 du = 2\pi \cdot \int_{-\infty}^{\infty} f(x)^2 dx. \quad (20.25)$$

Предположим для простоты, что плотность вероятности $f(x)$ ограничена, то есть

$$f(x) \leq M_0 < \infty. \quad (20.26)$$

В этом случае

$$\int_{-\infty}^{\infty} f^2(x) dx \leq M_0 \int_{-\infty}^{\infty} f(x) dx = M_0 < \infty, \quad (20.27)$$

или иначе говоря

$$\int_{-\infty}^{\infty} |E(u)|^2 du < \infty. \quad (20.28)$$

Необходимым условием сходимости интеграла (20.28) является обращение в нуль подынтегрального выражения на бесконечности, чем и доказывается правильность формулы (20.21).

Если $f(x)$ неограничена, но случайная величина X по-прежнему непрерывна, то формула (20.21) остается в силе, хотя ее доказательство несколько усложняется [28]. Для дискретной случайной величины свойство (20.21) не имеет места, так как в этом случае плотность вероятности $f(x)$ содержит дельта-функции и интеграл (20.22) расходится.

Весьма важным свойством характеристической функции является ее связь с моментами случайной величины. Покажем, что если начальные моменты m_j существуют до момента порядка k , то

$$m_j = \frac{1}{ij} \cdot \left. \frac{d^j E(u)}{du^j} \right|_{u=0}, \quad (j = \overline{1, k}), \quad (20.29)$$

то есть характеристическая функция является производящей функцией для моментов распределения и асимптотически при малых $|u|$ может быть представлена в этом случае в виде отрезка ряда Тэйлора

$$E(u) = \sum_{j=0}^k \frac{m_j}{j!} (iu)^j + O(u^k). \quad (20.30)$$

Для вывода формулы (20.29) достаточно заметить, что поскольку для существования момента m_j необходимо, чтобы интеграл

$$m_j = \int_{-\infty}^{\infty} x^j f(x) dx \quad (20.31)$$

сходился абсолютно, то интеграл (20.2), определяющий характеристическую функцию

$$E(u) = \int_{-\infty}^{\infty} e^{iux} f(x) dx, \quad (20.32)$$

можно дифференцировать по u как по параметру j раз. Выполнив это дифференцирование и положив затем $u = 0$, убеждаемся в справедливости формулы (20.29).

Если обозначить $\overset{\circ}{E}(u)$ характеристическую функцию центрированной случайной величины $\overset{\circ}{X} = X - \bar{x}$, то для центрального момента μ_j случайной величины X получим

$$\mu_j = \frac{1}{ij} \cdot \left. \frac{d^j \overset{\circ}{E}(u)}{du^j} \right|_{u=0}, \quad (20.33)$$

или, учитывая, что согласно (20.12)

$$\overset{\circ}{E}(u) = e^{-i\bar{x}u} \cdot E(u), \quad (20.34)$$

для центрального момента получим формулу

$$\mu_j = \frac{1}{i^j} \cdot \frac{d^j}{du^j} [e^{-i\bar{x}u} \cdot E(u)] \Big|_{u=0}. \quad (20.35)$$

В некоторых задачах удобней использовать не разложение характеристической функции в ряд по степеням ее аргумента, а разложение в этот ряд ее логарифма, то есть разложение

$$\ln E(u) = \frac{\varkappa_\nu}{\nu!} (iu)^\nu + O(u^k), \quad (20.36)$$

где вещественные коэффициенты $\varkappa_\nu$ носят название *семиинвариантов* или *кумулянтов* распределения.

Между семиинвариантами $\varkappa_\nu$ и моментами случайной величины можно установить взаимно однозначное соответствие, если на основании (20.6) записать характеристическую функцию в виде

$$E(u) = \exp \left\{ \sum_{\nu=1}^k \frac{\varkappa_\nu}{\nu!} (iu)^\nu + O(u^k) \right\}, \quad (20.37)$$

разложить правую часть последнего равенства по степеням u и сравнить коэффициенты полученного ряда с коэффициентами ряда (20.30). В результате получим формулы, выражающие начальные моменты через семиинварианты. Если сравнить коэффициенты разложения $\ln E(u)$, получаемые из формул (20.37) и (20.30), то мы получим возможность выразить семиинварианты через моменты. Опуская промежуточные выкладки, приведем окончательные формулы для первых шести семиинвариантов:

$$\begin{aligned} \varkappa_1 &= \bar{x}; \quad \varkappa_2 = \sigma^2; \quad \varkappa_3 = \mu_3; \quad \varkappa_4 = \mu_4 - 3\sigma^4; \\ \varkappa_5 &= \mu_5 - 10\sigma^2\mu_3; \quad \varkappa_6 = \mu_6 - 15\sigma^2\mu_4 - 10\mu_3^2 + 30\sigma^6. \end{aligned} \quad (20.38)$$

Формулы, выражающие первые шесть начальных моментов через семиинварианты, имеют вид

$$\begin{aligned}
 m_1 &= \kappa_1, \quad m_2 = \kappa_2 + \kappa_1^2, \quad m_3 = \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3, \\
 m_4 &= \kappa_4 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + 6\kappa_1^2\kappa_2 + \kappa_1^4, \\
 m_5 &= \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + \\
 &\quad + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5, \\
 m_6 &= \kappa_6 + 5\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + \\
 &\quad + 15\kappa_4\kappa_1^2 + 15\kappa_4\kappa_1^2 + 10\kappa_3^2 + 60\kappa_3\kappa_2\kappa_1 + \\
 &\quad + 20\kappa_3\kappa_1^3 + 15\kappa_2^3 + 45\kappa_2^2\kappa_1^2 + 15\kappa_2\kappa_1^4 + \kappa_1^6.
 \end{aligned} \tag{20.39}$$

Формулы (20.38) показывают, что определенные равенствами (19.23) и (19.24) асимметрия и эксцесс выражаются через семиинварианты следующим образом:

$$S_k = \frac{\kappa_3}{\kappa_2^{3/2}}, \quad E_x = \frac{\kappa_4}{\kappa_2^2}. \tag{20.40}$$

Рассмотрим несколько примеров.

Пример 20.1. Целочисленная случайная величина X имеет ряд распределения $p_m = \mathbf{P}(X = m) = \frac{1}{(m+1)(m+2)}$, $(m = \overline{0, \infty})$. Найти характеристическую функцию.

Решение. Так как X — дискретная случайная величина, то в соответствии с (20.3) имеем

$$E(u) = \sum_{m=0}^{\infty} e^{imu} \cdot \frac{1}{(m+1)(m+2)}.$$

Для суммирования бесконечной суммы рассмотрим вначале ряд

$$\psi(y) = \sum_{m=0}^{\infty} \frac{y^m}{(m+1)(m+2)}.$$

При $|y| < 1$ имеем

$$\begin{aligned}
 \psi(y) &= \sum_{m=0}^{\infty} \frac{y^m}{m+1} - \sum_{m=0}^{\infty} \frac{y^m}{m+2} = \\
 &= \frac{1}{y} \cdot \left(y + \frac{1}{2}y^2 + \frac{1}{3}y^3 + \dots \right) - \\
 &\quad - \frac{1}{y^2} \cdot \left(y + \frac{1}{2}y^2 + \frac{1}{3}y^3 + \dots - y \right) = \\
 &= -\frac{1}{y} \ln(1-y) - \frac{1}{y^2} [-\ln(1-y) - y] = \\
 &= \frac{1}{y} \left[1 + \left(\frac{1}{y} - 1 \right) \ln(1-y) \right].
 \end{aligned}$$

Исходный ряд сходится и при $|y| = 1$, следовательно, положив в окончательной формуле $y = e^{iu}$, получим

$$E(u) = e^{-iu} [1 + (e^{-iu} - 1) \ln(1 - e^{iu})].$$

Случайная величина X не имеет моментов начиная с первого. Поэтому формула (20.29) в данном случае неприменима, что сразу обнаруживается при попытке ее использования, так как в данном случае

$$\left. \frac{1}{i} \frac{dE(u)}{du} \right|_{u=0} = \infty.$$

Найденная функция $E(u)$ удовлетворяет общим свойствам характеристических функций, а именно, она непрерывна, удовлетворяет условию нормировки $E(0) = 1$. Поскольку величина X дискретная, то, естественно, $\lim_{|z| \rightarrow \infty} E(z) \neq 0$.

Пример 20.2. Случайная величина X , возможные значения которой ограничены интервалом $(0, 1)$, имеет плотность вероятности $f(x) = 2x$. Найти характеристическую функцию, первые четыре момента, асимметрию и эксцесс.

Решение. Для характеристической функции имеем

$$E(u) = 2 \cdot \int_0^1 e^{iux} x dx = \frac{2}{u^2} [(1 - iu) \cdot e^{iu} - 1].$$

В качестве проверки можно в найденное выражение подставить $u = 0$. Раскрывая неопределенность, получим $E(0) = 1$. Подставляя $u = \pm\infty$, находим $E(\pm\infty) = 0$.

Так как значения случайной величины X ограничены ($0 \leq X \leq 1$), то все моменты существуют и применима формула (20.29). Однако вычислений, связанных с дифференцированием $E(u)$, в данном случае можно избежать, разложив $E(u)$ в ряд по степеням u . Применив для этой цели разложение

$e^{iu} = \sum_{j=0}^{\infty} \frac{1}{j!} \cdot (iu)^j$, после очевидных преобразований, находим

$$E(u) = 2 \cdot \sum_{j=2}^{\infty} \frac{(j+1)}{(j+2)!} \cdot (-1)^j \cdot (iu)^j.$$

Сравнивая коэффициенты полученного ряда с коэффициентами ряда (20.30) (или применяя формулу (20.29)), приходим к общему выражению для начального момента j -го порядка

$$m_k = \frac{2}{k+2}.$$

Следовательно, начальные моменты даются равенством

$$m_1 = \frac{2}{3}, \quad m_2 = \frac{1}{2}, \quad m_3 = \frac{2}{5}, \quad m_4 = \frac{1}{3}.$$

Далее, применяя (20.13), находим

$$\begin{aligned} \mu_2 &= \frac{1}{2} - \frac{4}{9} = \frac{1}{18}, \quad \mu_3 = \frac{2}{5} - 3 \cdot \frac{2}{3} \cdot \frac{1}{2} + 2 \cdot \frac{8}{27} = -\frac{1}{135}, \\ \mu_4 &= \frac{1}{3} - 4 \cdot \frac{2}{3} \cdot \frac{2}{5} + 6 \cdot \frac{41}{92} - 3 \cdot \frac{16}{81} = \frac{1}{135}. \end{aligned}$$

Следовательно,

$$S_k = \frac{1 \cdot 18\sqrt{18}}{135} = -0,564; \quad E_x = \frac{18^2}{135} - 3 = -\frac{3}{5} = -0,6.$$

Пример 20.3. Доказать справедливость следующего рекуррентного соотношения между центральными моментами и семиинвариантами

$$\mu_k = \sum_{\nu=1}^k C_{k-1}^{\nu-1} \cdot \mu_{k-\nu} \kappa_{\nu},$$

где все моменты и семиинварианты предполагаются существующими.

Решение. Из определения семиинвариантов (20.37) имеем

$$\ln E(u) = \sum_{\nu=1}^{\infty} \frac{\kappa_{\nu}}{\nu!} (iu)^{\nu}.$$

Дифференцируя по u , получим

$$\frac{dE(u)}{du} = E(u) \cdot \sum_{\nu=0}^{\infty} i \cdot \frac{\kappa_{\nu+1}}{\nu!} \cdot (iu)^{\nu}.$$

Считая, что характеристическая функция $E(u)$ соответствует центрированной случайной величине, подставим ряд (20.30), заменив в нем m_j на μ_j , что дает

$$\sum_{j=0}^{\infty} \frac{1}{j!} \cdot \mu_{j+1} (iu)^j = \sum_{l=0}^{\infty} \sum_{\nu=0}^{\infty} \frac{1}{l! \nu!} \cdot \mu_l \kappa_{\nu+1} (iu)^{\nu+l}.$$

Вводя новый параметр суммирования $j = \nu + l$ и производя суммирование вначале по l , а затем по j , получим

$$\sum_{j=1}^{\infty} \frac{1}{j!} \cdot \mu_{j+1} (iu)^{j-1} = \sum_{j=0}^{\infty} \left[\sum_{l=0}^j \frac{1}{l!} \cdot \mu_l \kappa_{j-l} \cdot \frac{1}{(j-l)!} \right] (iu)^j.$$

Приравнявая коэффициенты у одинаковых степеней (iu) , имеем

$$\mu_j = \sum_{l=0}^{j-1} j-1 \frac{j!}{l!(j-l)!} \cdot \mu_l \cdot \kappa_{j-l} = \sum_{\nu=1}^j C_j^{\nu} \cdot \mu_{j-\nu} \cdot \kappa_{\nu},$$

где в последней сумме суммирование производится от $\nu = 1$, так как для центрированной величины $\varkappa_1 = 0$. Полученное выражение и является искомой формулой, рассматривая которую, как систему линейных уравнений относительно неизвестных значений семиинвариантов $\varkappa_\nu$ от $\nu = 2$ до $\nu = k$, можно получить явное выражение для $\varkappa_\nu$:

$$\varkappa_\nu = (-1)^{\nu-1} \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ \mu_2 & 0 & 1 & \dots & 0 \\ \mu_3 & C_2^0 \mu_2 & 0 & \dots & 0 \\ \mu_4 & C_3^0 \mu_3 & C_3^1 \mu_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \mu_\nu & C_{\nu-1}^0 \mu_{\nu-1} & C_{\nu-1}^1 \mu_{\nu-2} & \dots & 0 \end{vmatrix},$$

$$(\nu = \overline{2, k}).$$

Придавая ν значения от 2 до 6, получаем ранее приведенные выражения (20.38).

НЕКОТОРЫЕ НАИБОЛЕЕ УПОТРЕБИТЕЛЬНЫЕ ТИПОВЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ

§ 21. ПРИМЕРЫ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Рассмотрим несколько законов распределения дискретных случайных величин, наиболее часто встречающихся в приложениях.

21.1. Равномерное распределение

Простейшим распределением дискретной случайной величины является *равномерное распределение*, при котором любое из n возможных значений x_j ($j = \overline{1, n}$) случайной величины X имеет одну и ту же вероятность p . Так как $\sum_{j=1}^n p_j = 1$,

то это означает, что $p = \frac{1}{n}$. Следовательно, для равномерно распределенной случайной величины ряд распределения имеет вид

$$p_j = \mathbf{P}(X = x_j) = \frac{1}{n}, \quad (j = \overline{1, n}). \quad (21.1)$$

Мы уже встречались с этим распределением в примере 19.3, сейчас изучим его подробнее. Функция распределения равномерного закона

$$F(x) = \begin{cases} 0, & x \leq x_1, \\ \frac{j}{n}, & x_j < x \leq x_{j+1}, \quad (j = \overline{1, n-1}), \\ 1, & x > x_n. \end{cases} \quad (21.2)$$

Начальный и центральный моменты k -го порядка для X рассчитываются по формулам:

$$m_k = \frac{1}{n} \sum_{j=1}^n x_j^k,$$

$$\mu_k = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^k.$$

Математическое ожидание $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ в данном случае совпадает со средним арифметическим возможных значений случайной величины X . Характеристическая функция случайной величины X в соответствии с (20.3) будет иметь вид

$$E(u) = \frac{1}{n} \sum_{j=1}^n e^{ix_j u}. \quad (21.3)$$

В частном случае, когда

$$x_k = k \quad (k = \overline{1, n}), \quad (21.4)$$

то есть дискретная случайная величина X является к тому же еще и целочисленной, формулы для моментов и для характеристической функции упрощаются.

$$m_k = \frac{1}{n} \sum_{j=1}^n j^k = \frac{n^k}{k+1} + \frac{n^{k-1}}{2} + \sum_{l=1}^{[k/2]} C_k^{2l-1} B_{2l} n^{k-2l} \cdot \frac{1}{2l}, \quad (21.5)$$

где B_{2l} обозначает числа Бернулли, наименьшие из которых имеют следующие значения:

$$B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30}.$$

Общее выражение для чисел Бернулли имеет вид [1]

$$B_{2j} = \frac{(-1)^{j-1} (2j)!}{2^{2j-1} \pi^{2j}} \cdot \sum_{s=1}^{\infty} \frac{1}{s^{2j}}, \quad (j \geq 1).$$

Применяя формулу (21.5), находим

$$m_1 = \frac{n+1}{2}, \quad m_2 = \frac{(n+1)(2n+1)}{6}, \quad \sigma^2 = \frac{n^2-1}{12}.$$

Характеристическая функция вычисляется с помощью формулы для суммы геометрической прогрессии и имеет вид

$$E(u) = \frac{1}{n} \sum_{m=1}^n e^{imu} = \frac{(e^{inu} - 1)e^{iu}}{n(e^{iu} - 1)}. \quad (21.6)$$

21.2. Биномиальный закон распределения

Биномиальным законом распределения называется закон распределения целочисленной случайной величины X , заданной рядом распределения

$$\mathbf{P}(X = m) = C_n^m p^m (1-p)^{n-m}, \quad (m = \overline{0, n}). \quad (21.7)$$

Такая случайная величина возникает естественным образом при рассмотрении числа появлений события A в серии n независимых испытаний при условии, что в каждом испытании $\mathbf{P}(A) = p$ (схема Бернулли). Эта схема была подробно рассмотрена во второй главе при изучении случайных событий. Величины, подчиняющиеся биномиальному закону распределения (21.7), могут рассматриваться и независимо от приводящей к их возникновению схемы Бернулли.

Найдем характеристическую функцию случайной величины, подчиняющейся биномиальному закону распределения. В соответствии с формулой (20.3) имеем

$$\begin{aligned} E(u) &= \sum_{m=0}^n e^{imu} C_n^m p^m (1-p)^{n-m} = \\ &= (pe^{iu} + q)^n, \quad q = 1-p. \end{aligned} \quad (21.8)$$

Этот же результат можно получить, если в производящую функцию $G(u) = (pu + q)^n$ в соответствии с (20.1), вместо аргумента u подставить e^{iu} .

Воспользуемся характеристической функцией (21.8) для нахождения первых моментов биномиального закона распределения. В соответствии с (20.29) имеем

$$\begin{aligned}\bar{x} = m_1 &= \frac{1}{i} \frac{d}{du} \cdot [(pe^{iu} + q)^n] \Big|_{u=0} = \\ &= npe^{iu}(pe^{iu} + q)^{n-1} \Big|_{u=0} = np, \quad (21.9)\end{aligned}$$

$$\begin{aligned}m_2 &= -\frac{d^2}{du^2} \cdot [(pe^{iu} + q)^n] \Big|_{u=0} = np[e^{iu}(pe^{iu} + q)^{n-1} + \\ &+ (n-1)pe^{2iu}(pe^{iu} + q)^{n-2}] \Big|_{u=0} = \\ &= npq + n^2p^2, \quad (21.10)\end{aligned}$$

$$\begin{aligned}m_3 &= \frac{1}{i^3} \frac{d^3 E(u)}{du^3} \Big|_{u=0} = \\ &= np[1 + 3(n-1)p + (n-1)(n-2)p^2], \quad (21.11)\end{aligned}$$

$$\begin{aligned}m_4 &= np[1 + 7(n-1)p + 6(n-1)(n-2)p^2 + \\ &+ (n-1)(n-2)(n-3)p^3]. \quad (21.12)\end{aligned}$$

Следовательно, используя соотношения (19.15) между центральными и начальными моментами имеем

$$\begin{aligned}\sigma^2 &= \mu_2 = m_2 - m_1^2 = npq, \\ \mu_3 &= npq(q-p), \\ \mu_4 &= npq(1-6pq) + 3n^2p^2q^2.\end{aligned} \quad (21.13)$$

21.3. Закон Пуассона

Целочисленная случайная величина имеет закон распределения Пуассона, если ряд распределения этой величины задан формулой

$$p_m = \mathbf{P}(X = m) = \frac{a^m}{m!} e^{-a}, \quad (m = \overline{0, \infty}), \quad (21.14)$$

где a — некоторая положительная постоянная, называемая параметром закона Пуассона. Напомним, что закон Пуассона также рассматривался в схеме случайных событий и был получен в результате предельного перехода в схеме Бернулли к бесконечному числу испытаний, когда $n \rightarrow \infty$ при сохранении постоянства $np = a$.

Характеристическая функция для закона Пуассона в соответствии с (20.3) будет такой:

$$E(u) = \sum_{m=0}^{\infty} e^{ium} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{(ae^{iu})^m}{m!},$$

что после суммирования ряда дает

$$E(u) = e^{-a} \cdot e^{ae^{iu}}. \quad (21.15)$$

Эту же формулу можно получить из выражения (21.8) для характеристической функции биномиального закона распределения, перейдя к пределу при $n \rightarrow \infty$, когда $np = a$.

Воспользуемся формулой (21.15) для нахождения моментов распределения Пуассона. Применяя формулу (20.29), моменты можно найти путем дифференцирования характеристической функции. Однако в данном случае проще вначале найти семиинварианты, а затем уже по ним определить и сами моменты. Действительно, разлагая в формуле (21.15) e^{iu} по степеням (iu) , получим

$$E(u) = \exp \left\{ \sum_{j=1}^{\infty} \frac{a}{j!} (iu)^j \right\}, \quad (21.16)$$

следовательно, все семиинварианты закона распределения Пуассона равны параметру этого закона a :

$$x_j = a, \quad (j = \overline{1, \infty}). \quad (21.17)$$

Подставляя эти значения в формулы (20.38), получим

$$\bar{x} = a; \quad \sigma^2 = a; \quad \mu_3 = a; \quad \mu_4 = a + 3a^2. \quad (21.18)$$

Таким образом, кроме постоянства семиинвариантов, закон Пуассона обладает и еще одной характерной особенностью, заключающейся в том, что все первые три его момента m_1, μ_2, μ_3 равны между собой. Воспользовавшись (21.18) для нахождения асимметрии и эксцесса, получим, что для закона Пуассона и асимметрия, и эксцесс положительны и задаются очень простыми выражениями:

$$S_k = \frac{\mu_3}{\sigma^3} = \frac{1}{\sqrt{a}}, \quad (21.19)$$

$$E_x = \frac{\mu_4}{\sigma^4} - 3 = \frac{1}{a}. \quad (21.20)$$

Распределение Пуассона является предельным не только для числа появления события A в схеме Бернулли, но и для обобщенного биномиального распределения, соответствующего числу появлений события в серии независимых испытаний, когда вероятность появления события в k -м испытании p_k не остается постоянной, если при числе испытаний $n \rightarrow \infty$ наибольшая из вероятностей $p_{\max} = \max\{p_n\}$ стремится к нулю, но при этом $\sum_{k=1}^n p_k = a$ все-таки остается постоянной.

Действительно, согласно результату, полученному во второй главе для производящей функции $G^{(n)}(u)$, в этом случае имеет место формула

$$G^{(n)}(u) = \prod_{k=1}^n (q_k + p_k u) = \prod_{k=1}^n [1 + p_k(u - 1)]. \quad (21.21)$$

Логарифмируя обе части последнего равенства, получим

$$\ln G^{(n)}(u) = \sum_{k=1}^n \ln[1 + p_k(u - 1)]. \quad (21.22)$$

Считая $|p_k(u - 1)| < 1$, можно воспользоваться разложением логарифма в ряд Тэйлора вблизи точки $u = 1$:

$$\begin{aligned} \ln[1 + p_k(u - 1)] &= p_k(u - 1) - \\ &- \frac{1}{2} p_k^2(u - 1)^2 + \frac{1}{3} p_k^3(u - 1)^3 - \dots, \end{aligned} \quad (21.23)$$

что позволяет представить (21.22) в виде

$$\ln G^{(n)}(u) = (u-1) \sum_{k=1}^n p_k - \frac{1}{2} \cdot (u-1)^2 \sum_{k=1}^n p_k^2 + \frac{1}{3} (u-1)^3 \sum_{k=1}^n p_k^3 - \dots \quad (21.24)$$

Так как здесь $\sum_{k=1}^n p_k = a$, а при $m \geq 2$ справедлива оценка

$$\sum_{k=1}^n p_k^m \leq p_{\max}^{m-1} \sum_{k=1}^n p_k = a p_{\max}^{m-1}, \quad (21.25)$$

то из (21.25) следует, что

$$\begin{aligned} |\ln G^{(n)}(u) - a(u-1)| &\leq a p_{\max} |u-1| + \\ &+ \frac{1}{2} a^2 p_{\max}^2 |u-1|^2 + \frac{1}{3} a^3 p_{\max}^3 |u-1|^3 + \dots = \\ &= -\ln[1 - a p_{\max} |u-1|] \xrightarrow{n \rightarrow \infty} 0, \end{aligned} \quad (21.26)$$

что эквивалентно

$$G(u) = \lim_{n \rightarrow \infty} G^{(n)}(u) = e^{-a(u-1)}. \quad (21.27)$$

Для того чтобы от производящей функции целочисленной случайной величины перейти к характеристической функции, согласно (20.6), нужно ее аргумент u заменить на e^{iu} . Выполнив эту замену в (21.27), для характеристической функции $E(u)$, получим

$$E(u) = e^{-a(e^{iu}-1)}, \quad (21.28)$$

то есть характеристическую функцию закона Пуассона с параметром

$$a = \sum_{k=1}^{\infty} p_k.$$

21.4. Геометрическое распределение

Геометрическим называется распределение целочисленной случайной величины X , имеющей ряд распределения

$$p_k = \mathbf{P}(X = k) = q^k p, \quad q = 1 - p, \quad (k = \overline{0, \infty}), \quad (21.29)$$

представляющий собой бесконечно убывающую геометрическую прогрессию. Этому распределению, например, подчиняется число испытаний, предшествующих появлению события A в серии независимых испытаний, если вероятность появления события A в одном испытании равна p .

Так как

$$\mathbf{P}\{X < k\} = \sum_{j=0}^{k-1} p_j = p \sum_{j=0}^{k-1} q^j = p \frac{q^k - 1}{q - 1} = 1 - q^k,$$

то функция распределения будет иметь вид

$$F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - q^k, & k - 1 < x \leq k. \end{cases} \quad (21.30)$$

Для характеристической функции получим

$$E(u) = \sum_{k=0}^{\infty} e^{iuk} q^k p = \frac{p}{1 - qe^{iu}}. \quad (21.31)$$

Дифференцируя (21.31), находим, опуская несложные промежуточные выкладки,

$$\bar{x} = m_1 = \frac{q}{p}; \quad m_2 = \frac{q}{p} + \frac{q^2}{p^2}; \quad \sigma^2 = m_2 - \bar{x}^2 = \frac{q}{p^2}. \quad (21.32)$$

21.5. Отрицательное биномиальное распределение

Предположим, что производятся независимые испытания, при каждом из которых событие A может произойти с вероятностью p . Обозначим через X число испытаний, которые нужно произвести сверх заданного их числа m для того, чтобы событие A произошло m раз. Случайная величина $X + m$

примет значение n ($n \geq m$), если событие A произойдет $m - 1$ раз при первых $n - 1$ испытаниях и произойдет также при n -м испытании, поэтому

$$\mathbf{P}\{X + m = n\} = (C_{n-1}^{m-1} p^{m-1} q^{n-m})p.$$

Положим $n - m = k$, тогда

$$\mathbf{P}\{X = k\} = C_{m+k-1}^{m-1} p^m q^k \quad (k = \overline{0, \infty}). \quad (21.33)$$

Ряд распределения (21.33) носит название отрицательного биномиального распределения. Это название связано с тем, что выражение $C_{m+k-1}^{m-1} q^k$ является общим членом разложения бинома $(1 - q)^{-m}$ в ряд по степеням q . Действительно, имеет место разложение (при $|q| < 1$)

$$(1 - q)^{-m} = \sum_{k=0}^{\infty} C_{m+k-1}^{m-1} q^k, \quad (21.34)$$

поэтому

$$p^m \sum_{k=0}^{\infty} C_{m+k-1}^{m-1} q^k = p^m (1 - q)^{-m} = 1.$$

Распределение (21.33) встречается во многих приложениях. Как частный случай при $m = 1$ из него получается ранее рассмотренное геометрическое распределение (21.30).

Определение характеристической функции по формуле (20.3) с учетом (21.34) дает

$$E(u) = \frac{p^m}{(1 - qe^{iu})^m}. \quad (21.35)$$

Дифференцируя (21.35), находим по обычным правилам

$$\bar{x} = \frac{mq}{p}; \quad m_2 = \frac{mq}{p^2} + \frac{m^2 q^2}{p^2}; \quad \sigma^2 = \frac{mq}{p^2}. \quad (21.36)$$

21.6. Гипергеометрическое распределение

Гипергеометрическим называется распределение дискретной случайной величины X , определяемое рядом распределения

$$\mathbf{P}\{X = k\} = \frac{C_m^k C_{n-m}^{s-k}}{C_n^s}, \quad (k = \overline{0, s}), \quad (21.37)$$

где n , m и s , ($m < n$, $s < n$) — заданные целые числа.

Это распределение возникает, например, при рассмотрении следующей задачи, связанной с контролем качества продукции: из n изделий некоторым признаком A (например, это может быть бракованное изделие) обладают m изделий. Обозначим через X случайную величину, означающую число изделий, обладающих признаком A , среди взятых наудачу s изделий, тогда

$$\mathbf{P}\{X = k\} = \frac{C_m^k C_{n-m}^{s-k}}{C_n^s}, \quad (k = \overline{0, s}). \quad (21.38)$$

Для вычисления моментов гипергеометрического распределения удобнее воспользоваться производящей функцией, а не характеристической функцией. Согласно определению производящей функции имеем

$$G(u) = \frac{1}{C_n^s} \cdot \sum_{k=0}^s C_m^k C_{n-m}^{s-k} u^k. \quad (21.39)$$

Далее, для вычисления моментов случайной величины X с помощью производящей функции, удобно представить выражение (21.39) в виде производной от полинома. Имеем

$$\begin{aligned} (1 + uv)^m (1 + v)^{n-m} &= \\ &= \sum_{k=0}^m (uv)^k C_m^k \sum_{j=0}^{n-m} C_{n-m}^j v^j = \\ &= \sum_{s=0}^n v^s \left[\sum_{k=0}^s C_m^k C_{n-m}^{s-k} u^k \right]. \end{aligned}$$

Коэффициент при v^s в правой части этого равенства равен $C_n^s G(u)$, поэтому

$$G(u) = \frac{1}{s! C_n^s} \left\{ \frac{\partial^s}{\partial v^s} [(1+uv)^m (1+v)^{n-m}] \right\} \Big|_{v=0}. \quad (21.40)$$

С помощью полученного выражения находим

$$G'(u) = \frac{m}{s! C_n^s} \left\{ \frac{\partial^s}{\partial v^s} [v(1+uv)^{m-1} (1+v)^{n-m}] \right\} \Big|_{v=0}.$$

Тогда

$$G'(1) = \frac{m}{s! C_n^s} \left\{ \frac{\partial^s}{\partial v^s} [v(1+v)^{n-1}] \right\} \Big|_{v=0} = \frac{m}{C_n^s} C_{n-1}^{s-1} = \frac{ms}{n};$$

$$G'(1) = \frac{m(m-1)}{C_n^s} C_{n-2}^{s-2} = \frac{m(m-1)}{n(n-1)} s(s-1).$$

Основные характеристики рассмотренных целочисленных дискретных распределений представлены в таблице 21.1, где дается ряд распределения, производящая и характеристическая функции, математическое ожидание, дисперсия и вероятнейшее значение. Отметим, что через $[x]$ с помощью квадратных скобок обозначается целая часть числа x . Далее перейдем к примерам на дискретные распределения.

Пример 21.1. Для биномиального закона распределения (21.7) определить первые четыре семиинварианта.

Решение. Для определения искоемых семиинвариантов можно воспользоваться значениями первых четырех моментов, определяемых формулой (21.7), однако проще непосредственно исходить из формулы для характеристической функции (21.2). Представив характеристическую функцию в виде

$$E(u) = e^{n \ln[1-p(1-e^{iu})]},$$

замечаем, что семиинвариант κ_j является коэффициентом в разложении функции $n \ln[1-p(1-e^x)]$ в ряд Маклорена.

Таблица 21.1
Основные характеристики дискретных
законов распределения

Название закона распределения	Ряд распределения	Производящая функция $G(u)$
Равномерное	$\mathbf{P}\{X = k\} = \frac{1}{n},$ $(k = \overline{1, n})$	$\frac{u(1 - u^n)}{n(1 - u)}$
Биномиальное	$\mathbf{P}\{X = k\} =$ $= C_n^k p^k q^{n-k},$ $(k = \overline{0, n})$	$(q + pu)^n$
Обобщенное биномиальное	$\mathbf{P}\{X = k\} =$ $= \frac{1}{k!} \left. \frac{d^k G(u)}{du^k} \right _{u=0}$	$\prod_{j=1}^n (q_j + p_j u)$
Пуассона	$\mathbf{P}\{X = k\} = \frac{a^k}{k!} e^{-a},$ $(k = \overline{0, n})$	$e^{a(u-1)}$
Геометрическое при конечном числе значений X	$\mathbf{P}\{X = k\} = q^k p,$ $(0 \leq k < n)$ $\mathbf{P}\{X = n\} = q^n$	$\frac{p[1 - (qu)^n]}{1 - qu} + (qu)^n$
Геометрическое при бесконечном числе значений X	$\mathbf{P}\{X = k\} = q^k p,$ $(0 \leq k < \infty)$	$\frac{p}{1 - qu}$
Отрицательное биномиальное (Паскаля)	$\mathbf{P}\{X = k\} =$ $= C_{k-1}^{m-1} p^m q^{k-m},$ $(k = \overline{0, \infty})$	$\frac{(pu)^n}{(1 - qu)^m}$
Гипергеометрическое	$\mathbf{P}\{X = k\} =$ $= \frac{C_m^k C_{n-m}^{s-k}}{C_n^s},$ $(k = \overline{0, s})$	$\frac{1}{s! C_n^s} \left\{ \frac{\partial^s}{\partial v^s} \times \right.$ $\left. \times [(1+uv)^m (1+v)^{n-m}] \right\}$

Продолжение табл. 21.1

Название закона распределения	Характеристическая функция	Математическое ожидание
Равномерное	$\frac{e^{iu}(1 - e^{inu})}{n(1 - e^{iu})}$	$\frac{1}{2}(n + 1)$
Биномиальное	$(q + pe^{iu})^n$	np
Обобщенное биномиальное	$\prod_{j=1}^n (q_j + p_j e^{iu})$	$\sum_{j=1}^n p_j$
Пуассона	$\exp(ae^{iu} - 1)$	a
Геометрическое при конечном числе значений X	$\frac{p[1 - q^n e^{inu}]}{1 - qe^{iu}} + q^n e^{inu}$	$\frac{q}{p}(1 - q^n)$
Геометрическое при бесконечном числе значений X	$\frac{p}{1 - qe^{iu}}$	$\frac{q}{p}$
Отрицательное биномиальное (Паскаля)	$\frac{p^m e^{imu}}{(1 - qe^{iu})^m}$	$\frac{m}{p}$
Гипергеометрическое	$G(e^{iu})$	$\frac{ms}{n}$

Продолжение табл. 21.1

Название закона распределения	Дисперсия	Вероятнейшее значение
Равномерное	$\frac{n^2 - 1}{12}$	—
Биномиальное	npq	$[(n+1)p]$ или $[(n+1)p - 1]$ и $[(n+1)p]$
Обобщенное биномиальное	$\sum_{j=1}^n p_j q_j$	
Пуассона	a	a или $a - 1$
Геометрическое при конечном числе значений X	$\frac{q}{p^2} - \frac{q^{n+1}}{p} (2n + 1 + \frac{q^{n+1}}{p})$	0 при $p > q^n$, n при $p < q^n$
Геометрическое при бесконечном числе значений X	$\frac{q}{p^2}$	0
Отрицательное биномиальное (Паскаля)	$\frac{mq}{p^2}$	$\left[\frac{m-1}{p} \right]$ или $\left[\frac{m-1}{p} - 1 \right]$ и $\left[\frac{m-1}{p} \right]$
Гипергеометрическое	$\frac{ms(n-m)(n-s)}{n^2(n-1)}$	$\left[\frac{(m+1)(s+1)}{n+2} \right]$ или $\left[\frac{(m+1)(s+1)}{n+2} \right] + 1$ и $\left[\frac{(m+1)(s+1)}{n+2} \right]$

Таким образом, имеем

$$\begin{aligned}\kappa_j &= n \frac{d^j}{dx^j} \left\{ \ln[1 - p(1 - e^x)] \right\} \Big|_{x=0}, \\ \kappa_1 &= \frac{npe^x}{1 - p(1 - e^x)} \Big|_{x=0} = np, \\ \kappa_2 &= \left[\frac{npe^x}{1 - p(1 - e^x)} - \frac{np^2 e^{2x}}{[1 - p(1 - e^x)]^2} \right] \Big|_{x=0} = npq, \\ \kappa_3 &= np(q - p), \\ \kappa_4 &= npq(1 - 6pq).\end{aligned}$$

Пример 21.2. Определить вероятность того, что в экран площадью $S = 0,3 \text{ мм}^2$, поставленный на расстоянии $r = 1 \text{ м}$ перпендикулярно потоку альфа-частиц от радиоактивного вещества, за время $\tau = 2 \text{ секунды}$ попадет не менее трех частиц, если масса вещества $M = 0,1 \text{ мг}$, атомный вес $A = 226$, время полураспада $T_{1/2} = 1622 \text{ года}$.

Решение. Число ядер радиоактивного вещества $n = N_0 \times \frac{M}{A}$, где $N_0 = 6,02 \cdot 10^{23}$ — число Авогадро, весьма велико. Следовательно, число частиц, попадающих в экран, подчиняется закону Пуассона, а вероятность попадания в него m частиц $P_m = \frac{a^m}{m!} e^{-a}$, где a — математическое ожидание числа частиц, попадающих в экран за указанное время. Если λ — вероятность распада одного атома в течение одной секунды, то математическое ожидание общего числа распадов равно $n\lambda\tau$, следовательно,

$$a = n\lambda\tau \frac{S}{4\pi r^2}.$$

Для определения λ обозначим среднее число атомов, не распавшихся к моменту t , через $n(t)$. Математическое ожидание числа распавшихся атомов за время dt равно $n(t)\lambda dt$, то

есть

$$\begin{aligned} dn(t) &= -n(t)\lambda dt; & \frac{dn(t)}{dt} &= -\lambda; \\ n(t) &= n(0)e^{-\lambda t}. \end{aligned}$$

За время полураспада $T_{1/2}$ распадается половина начального числа атомов, значит

$$\frac{1}{2}n(0) = n(0)e^{-\lambda T_{1/2}}, \quad \lambda = \frac{\lg 2}{T_{1/2} \lg e}.$$

Таким образом, находим

$$\begin{aligned} a &= \frac{\tau N_0 M \lg 2 \cdot S}{AT_{1/2} \lg e \cdot 4\pi r^2} = \\ &= \frac{2 \cdot 6,02 \cdot 10^{23} \cdot 10^{-3} \lg 2 \cdot 0,03}{226 \cdot 1622 \cdot 365 \cdot 24 \cdot 60 \cdot 60 \lg e \cdot 4\pi \cdot 10^4} \approx 1. \end{aligned}$$

Искомая вероятность

$$P = 1 - P_1 - P_2 - P_3 = 1 - \left(1 + \frac{1}{2} + \frac{1}{6}\right)e^{-1} \approx 0,39.$$

Пример 21.3. Вероятность p того, что прибор при его включении даст отказ является постоянной для каждого включения. Определить математическое ожидание $\bar{x}$ и дисперсию $D(X)$ числа включений до первого отказа.

Решение. Случайная величина X — число включений до получения первого отказа подчиняется геометрическому распределению (21.29)

$$\mathbf{P}\{X = m\} = q^m p \quad (m = \overline{0, \infty}).$$

Следовательно, математическое ожидание и дисперсия определяются формулами (21.32):

$$\bar{x} = \frac{1-p}{p}, \quad D(X) = \frac{1-p}{p^2}.$$

§ 22. НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

Из многочисленных законов распределения непрерывных случайных величин, находящих свое применение в различных приложениях, наибольшее значение имеет нормальный закон распределения или закон Гаусса. Широкое распространение этого закона связано с двумя причинами. Во-первых, как это будет показано далее при рассмотрении предельных теорем теории вероятностей, при весьма широких допущениях нормальный закон распределения является предельным законом распределения для суммы независимых случайных величин при стремлении числа слагаемых к бесконечности. Во-вторых, этот закон распределения обладает устойчивостью в том смысле, что сумма нормально распределенных случайных величин также подчиняется нормальному закону распределения. Поэтому, хотя на практике никогда не встречаются случайные величины, строго следующие нормальному закону распределения, поскольку условия предельной теоремы требуют обращения в бесконечность числа слагаемых, — этот закон во многих случаях весьма хорошо аппроксимирует реальные законы распределения.

Нормальный закон распределения можно определить по-разному. Наиболее естественным представляется следующее определение: случайная величина X , заданная в интервале $(-\infty, +\infty)$, имеет нормальный закон распределения, если все ее семиинварианты, начиная с третьего, обращаются в нуль.

Так как первый семиинвариант есть математическое ожидание $\bar{x}$, а второй — дисперсия, то, применяя общую формулу (20.37), для характеристической функции нормального закона распределения, получим

$$E(u) = e^{i\bar{x}u - \frac{1}{2}\sigma^2 u^2}. \quad (22.1)$$

Таким образом, нормальный закон распределения определяется двумя параметрами: $\bar{x}$ и σ^2 . Поэтому для указания на то, что случайная величина X подчиняется нормальному закону распределения (или, сокращенно — X является нор-

мальной случайной величиной), применяется обозначение $X \in N(\bar{x}, \sigma^2)$.

Так как при линейном преобразовании случайной величины

$$Y = aX + b$$

характеристическая функция преобразуется по формуле (20.12)

$$E_y(u) = e^{ibu} E_x(au),$$

то в том случае, когда величина X является нормальной

$$E_y(u) = \exp \left\{ -\frac{a^2 \sigma_x^2}{2} u^2 + i(\bar{x} + b)u \right\}. \quad (22.2)$$

Значит, Y есть нормальная величина с математическим ожиданием $\bar{y} = \bar{x} + b$ и средним квадратическим отклонением $\sigma_y = |a|\sigma_x$. Следовательно, линейное преобразование нормальной случайной величины не меняет тип закона распределения.

Если нормальная случайная величина центрирована, то в формуле (22.1) нужно положить $\bar{x} = 0$, и для характеристической функции центрированной нормальной величины получим

$$\dot{E}(u) = e^{-\frac{1}{2}\sigma^2 u^2}. \quad (22.3)$$

Разложив экспоненту, стоящую справа, по степеням u^2 , получим

$$\dot{E}(u) = \sum_{k=0}^{\infty} (-1)^k \frac{\sigma^{2k}}{2^k k!} u^{2k}. \quad (22.4)$$

Так как центральный момент μ_k связан с характеристической функцией центрированной случайной величины формулой (20.33), то, дифференцируя ряд (22.4) последовательно $2k$ раз, полагая $u = 0$ и умножая на $i^{2k} = (-1)^k$, получим общую формулу для центрального момента четкого порядка нормальной случайной величины

$$\mu_{2k} = \frac{(2k)!}{2^k k!} \sigma^{2k} = (2k-1)!! \sigma^{2k}, \quad (k = \overline{1, \infty}). \quad (22.5)$$

Ряд (22.4) содержит только четные степени u , поэтому применение формулы (20.33) для момента нечетного порядка дает

$$\mu_{2k+1} = 0, \quad (k = \overline{1, \infty}). \quad (22.6)$$

Положив в формуле (22.5) $k = 2$, получим

$$\mu_4 = 3\sigma^4. \quad (22.7)$$

Найдем плотность вероятности нормального закона распределения. Для центрированной нормальной величины, применяя к (22.3) обратное преобразование Фурье, в соответствии с формулой (20.8), находим

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux - \frac{1}{2} \sigma^2 u^2} du = \\ &= \frac{1}{2\pi} e^{-x^2/2} \int_{-\infty}^{\infty} e^{-\frac{1}{2} (\sigma u - \frac{ix}{\sigma})^2} du = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2}, \end{aligned} \quad (22.8)$$

где при вычислении последнего интеграла принято во внимание, что

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

Этот интеграл известен, как интеграл Пуассона.

Таким образом, плотность вероятности центрированной нормальной величины имеет вид

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2}. \quad (22.9)$$

Если же случайная величина не центрирована, то при вычислении ее плотности вероятности, в соответствии с (22.1), получим

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iu(x-\bar{x}) - \frac{1}{2} \sigma^2 u^2} du. \quad (22.10)$$

Формула (22.10) отличается только тем, что в ней x заменен на $(x - \bar{x})$. Поэтому, минуя промежуточные выкладки и делая эту замену в окончательной формуле (22.9), получим окончательно для общего случая нормальной случайной величины

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}. \quad (22.11)$$

Формулу (22.11) часто называют канонической формой записи нормального закона распределения. Эта формула позволяет дать следующее определение нормального закона распределения: случайная величина X , заданная в интервале $(-\infty, \infty)$, подчиняется нормальному закону распределения, если ее плотность вероятности $f(x)$ может быть выражена в виде экспоненты, содержащей в показателе степени полином второй степени от x . Иначе говоря, X — нормальная величина, если

$$f(x) = e^{-Q_2(x)}, \quad (22.12)$$

где $Q_2(x)$ — полином второй степени¹. Очевидно, что определение нормального закона распределения (22.12) эквивалентно определению (1).

Нормальный закон распределения впервые был введен в теорию вероятностей при доказательстве теоремы Муавра–Лапласа (см. главу 2), согласно которой число m появлений события A в схеме Бернулли, при большом числе опытов n , приближенно может рассматриваться как непрерывная величина, подчиняющаяся нормальному закону распределения.

Примерно одновременно с Лапласом нормальным законом распределения пользовался Гаусс для характеристики ошибок измерения. Поэтому нормальный закон распределения часто называют законом Гаусса, а величины, подчиняющиеся этому закону — гауссовскими случайными величинами².

¹ Очевидно, что для того, чтобы $f(\pm\infty) = 0$, коэффициент старшего члена полинома $Q_2(x)$ должен быть положителен, а остальные коэффициенты должны быть такими, чтобы выполнялось условие нормировки f .

² Термин «нормальный закон» принадлежит А. Пуанкаре.

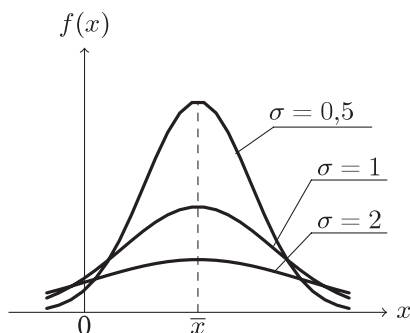


Рис. 22.1

Плотность вероятности нормального закона для $\bar{x} = 1$ при различных значениях σ : $\sigma = 0,5$, $\sigma = 1$ и $\sigma = 2$

И Лаплас, и Гаусс представляли плотность вероятности нормального закона распределения в несколько ином виде

$$f(x) = \frac{h}{\sqrt{\pi}} e^{-h^2(x-\bar{x})^2}, \quad (22.13)$$

где параметр h носит название «меры точности», поскольку при увеличении h уменьшаются вероятности отклонений случайной величины от ее математического ожидания, что соответствует повышению точности, если, например, под X понимать значение величины $\bar{x}$, искаженное ошибками измерения.

Сравнение (22.13) с (22.11) показывает, что между мерой точности h и средним квадратическим отклонением σ существует следующая связь:

$$h = \frac{1}{\sigma\sqrt{2}}, \quad \sigma = \frac{1}{h\sqrt{2}}. \quad (22.14)$$

График функции $y = f(x)$ для нормального закона распределения симметричен относительно вертикальной оси, проходящей через точку с абсциссой $x = \bar{x}$. При изменении математического ожидания $\bar{x}$ кривая $y = f(x)$, не изменяя формы, смещается вдоль оси абсцисс. Наибольшее значение плотности вероятности $f(x)$ принимает как раз при $x = \bar{x}$, то есть

нормальный закон распределения одномодальный, причем мода совпадает с математическим ожиданием и при этом

$$\max_{-\infty < x < \infty} \{f(x)\} = f(\bar{x}) = \frac{h}{\sqrt{\pi}} = \frac{1}{\sigma\sqrt{2\pi}}. \quad (22.15)$$

При изменении среднего квадратического отклонения σ сохраняется положение кривой $y = f(x)$, но изменяется ее форма (рис. 22.1). При увеличении σ максимальная ордината кривой $y = f(x)$ уменьшается. Так как площадь под этой кривой равна единице, то вследствие увеличения σ кривая становится более полой. Это означает, что будут увеличиваться вероятности больших отклонений случайной величины X от ее математического ожидания $\bar{x}$. Наоборот, с уменьшением среднего квадратического отклонения σ кривая $y = f(x)$ становится более вытянутой вверх. При $\sigma \rightarrow 0$ плотность вероятности $f(x)$ стремится к дельта-функции Дирака.

Найдем функцию распределения $F(x)$ нормальной случайной величины X . Имеем

$$F(x) = \int_{-\infty}^x f(x_1) dx_1 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x_1 - \bar{x})^2}{2\sigma^2}} dx_1.$$

Полагая $\frac{x - \bar{x}}{\sigma} = t$, получаем

$$\begin{aligned} F(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x - \bar{x}}{\sigma}} e^{-t^2/2} dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-t^2/2} dt + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x - \bar{x}}{\sigma}} e^{-t^2/2} dt. \end{aligned}$$

Так как

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-t^2/2} dt = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = \frac{1}{2},$$

то

$$F(x) = \frac{1}{2} \left(1 + \frac{2}{\sqrt{2\pi}} \int_0^{\frac{x-\bar{x}}{\sigma}} e^{-t^2/2} dt \right). \quad (22.16)$$

Интеграл $\int_0^z e^{-t^2/2} dt$ не выражается через элементарные функции, но может быть выражен через *функцию Лапласа* $\Phi(z)$, определяемую равенством

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt, \quad (22.17)$$

которая уже была введена ранее в главе 2 при рассмотрении простейших предельных теорем.

Для функции $\Phi(z)$, называемой также *интегралом вероятности*, имеются подробные таблицы [19]. Используя обозначение (22.17), функцию распределения (22.16) можно представить в виде

$$F(x) = \frac{1}{2} \left[1 + \Phi \left(\frac{x - \bar{x}}{\sigma} \right) \right]. \quad (22.18)$$

График этой функции представлен на рис. 22.2.

Зная функцию распределения $F(x)$, можно определить вероятность попадания нормальной случайной величины X в любой интервал (x_1, x_2) . Учитывая (22.18), для этой вероятности получим

$$\mathbf{P}\{x_1 \leq X \leq x_2\} = \frac{1}{2} \left[\Phi \left(\frac{x_2 - \bar{x}}{\sigma} \right) - \Phi \left(\frac{x_1 - \bar{x}}{\sigma} \right) \right]. \quad (22.19)$$

Если интервал (x_1, x_2) симметричен относительно точки $x = \bar{x}$, то можно принять $x_2 - \bar{x} = \bar{x} - x_1 = \varepsilon$, а тогда

$$\mathbf{P}\{x_1 \leq X \leq x_2\} = \mathbf{P}\{|X - \bar{x}| \leq \varepsilon\}.$$

Из (22.19) следует, что в этом простейшем случае

$$\mathbf{P}\{|X - \bar{x}| \leq \varepsilon\} = \Phi \left(\frac{\varepsilon}{\sigma} \right). \quad (22.20)$$

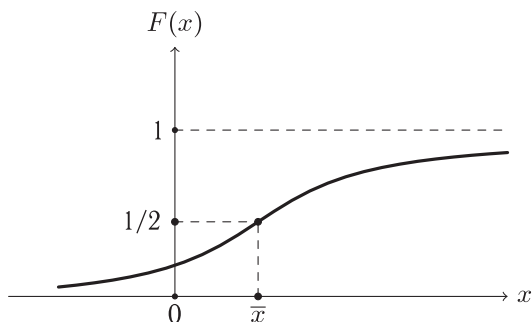


Рис. 22.2
Функция распределения нормального закона
при $\bar{x} = \sigma = 1$

Пусть, например, ε выбрано так, что $\varepsilon = k\sigma$, тогда

$$\mathbf{P}\{|X - \bar{x}| \leq k\sigma\} = \Phi(k). \quad (22.21)$$

Из таблицы [19] находим: $\Phi(1) = 0,683$; $\Phi(2) = 0,954$; $\Phi(3) = 0,997$. Поэтому с вероятностью 0,683 нормально распределенная случайная величина X не выйдет из интервала $(\bar{x} - \sigma, \bar{x} + \sigma)$ и с вероятностью 0,954 — из интервала $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$. С вероятностью 0,997 значение случайной величины принадлежит интервалу $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$ и с вероятностью 0,0027 — интервалу $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$.

Вероятность отклонения нормальной случайной величины X от ее математического ожидания $\bar{x}$ в любую сторону больше чем на три средних квадратических отклонения равна 0,0027. Так как эта вероятность мала, то часто приближенно можно считать, что интервалом возможных значений нормально распределенной случайной величины X является не вся числовая ось $(-\infty, \infty)$, а лишь конечный промежуток $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$, то есть, что случайная величина X может отклониться от ее среднего значения $\bar{x}$ не больше чем на три средних квадратических отклонения. Это приближенное правило иногда называют *правилом трех сигм*.

Так как плотность вероятности $f(x)$ нормальной случайной величины X симметрична относительно прямой $x = \bar{x}$, то в качестве характеристики рассеивания вместо среднего квадратического отклонения σ можно использовать так называемое срединное³ отклонение E , или половину длины интервала, симметричного относительно математического ожидания $\bar{x}$, вероятность попадания случайной величины X в который равна половине.

Согласно определению срединного отклонения и (22.20)

$$\mathbf{P}\{|X - \bar{x}| \leq E\} = \Phi\left(\frac{E}{\sigma}\right) = 0,5. \quad (22.22)$$

Корень уравнения $\Phi(z) = 0,5$ обозначается через $\rho\sqrt{2}$, где $\rho = 0,476936\dots$ ⁴, а множитель $\sqrt{2}$ вводится для того, чтобы при замене σ срединным отклонением плотность нормального закона распределения имела бы более простой вид. Так как $\frac{E}{\sigma}$ обозначено $\rho\sqrt{2}$, то для нормального закона имеем

$$E = \rho\sqrt{2}\sigma, \quad (22.23)$$

где $\rho\sqrt{2} = 0,6744\dots \approx \frac{2}{3}$.

Если в (22.11) заменить σ на $\frac{E}{\rho\sqrt{2}}$, то плотность вероятности для нормальной случайной величины X примет вид

$$f(x) = \frac{\rho}{E\sqrt{\pi}} e^{-\frac{\rho^2(x-\bar{x})^2}{E^2}}. \quad (22.24)$$

Именно в таком виде плотность вероятности используется в некоторых инженерных приложениях.

³ Иногда вместо термина «срединное отклонение» употребляют термин «вероятное отклонение».

⁴ Для запоминания числа ρ полезно заметить, что с точностью до трех десятичных знаков эта постоянная совпадает с $\lg 3$, то есть $\rho \approx \lg 3$.

Если в формуле (22.16) для функции распределения $F(x)$ заменить σ на $\frac{E}{\rho\sqrt{2}}$, то получим

$$F(x) = \frac{1}{2} \left[1 + \Phi \left(\rho\sqrt{2} \cdot \frac{x - \bar{x}}{E} \right) \right]. \quad (22.25)$$

Данное выражение можно упростить, если ввести новую функцию $\hat{\Phi}(z)$, которая связана с функцией Лапласа $\Phi(z)$ равенством

$$\hat{\Phi}(z) = \Phi(\rho\sqrt{2}z). \quad (22.26)$$

Функция $\hat{\Phi}(z)$, называемая *приведенной функцией Лапласа* (иногда *функцией Крампа*⁵), определяется формулой

$$\hat{\Phi}(z) = \frac{2\rho}{\sqrt{\pi}} \int_0^z e^{-\rho^2 t^2} dt. \quad (22.27)$$

Для $\hat{\Phi}(z)$, также, как и для $\Phi(z)$, составлены подробные таблицы. Производя замену (22.26) в (22.25), для функции распределения получим

$$F(x) = \frac{1}{2} \left[1 + \hat{\Phi} \left(\frac{x - \bar{x}}{E} \right) \right]. \quad (22.28)$$

Формулы (22.19) и (22.20) для вероятностей попадания нормальной случайной величины X в интервалы (x_1, x_2) и $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ с помощью приведенной функции Лапласа могут быть записаны, соответственно, в виде

$$\mathbf{P}\{x_1 \leq X \leq x_2\} = \frac{1}{2} \left[\hat{\Phi} \left(\frac{x_2 - \bar{x}}{E} \right) - \hat{\Phi} \left(\frac{x_1 - \bar{x}}{E} \right) \right], \quad (22.29)$$

$$\mathbf{P}\{|X - \bar{x}| \leq \varepsilon\} = \hat{\Phi} \left(\frac{\varepsilon}{E} \right). \quad (22.30)$$

⁵ Крамп Христиан (1760–1826) — немецкий математик.

С помощью (22.30) получаем

$$\mathbf{P}\{|X - \bar{x}| \leq kE\} = \Phi(k). \quad (22.31)$$

В приложениях часто встречаются случайные величины, законы распределения которых близки к нормальному. В этих случаях бывает целесообразно плотность вероятности представить в виде ряда, главный член которого соответствует нормальному закону распределения. Такой ряд можно получить следующим образом.

Рассмотрим случайную величину X , семиинварианты которой равны $\varkappa_j$ ($j = \overline{1, \infty}$). Характеристическая функция такой величины в соответствии с (20.37) может быть представлена в виде

$$E_x(u) = \exp \left\{ \sum_{\nu=1}^{\infty} \frac{\varkappa_{\nu}}{\nu!} (iu)^{\nu} \right\}. \quad (22.32)$$

Далее перейдем от величины X к центрированной и нормированной величине

$$Y = \frac{1}{\sigma_x} (X - \bar{x}). \quad (22.33)$$

Обозначим семиинварианты этой величины (22.23) через $\varkappa'_j$ и будем пока считать, что все семиинварианты существуют. Тогда в соответствии с (20.12)

$$\varkappa'_1 = \nu; \quad \varkappa'_2 = 1; \quad \varkappa'_j = \frac{1}{\sigma_x^j} \varkappa_j, \quad (j = \overline{3, \infty}), \quad (22.34)$$

и формула (22.32) примет вид

$$E_y(u) = e^{-\frac{u^2}{2}} \cdot \exp \left\{ \sum_{\nu=3}^{\infty} \frac{\varkappa'_{\nu}}{\nu!} (iu)^{\nu} \right\}. \quad (22.35)$$

Второй множитель в (22.35) можно формально рассматривать как характеристическую функцию случайной величины, первые два семиинварианта которой равны нулю, а остальные равны $\varkappa'_{\nu}$, то есть такой фиктивной случайной величины, у которой первые два момента равны нулю, а остальные моменты

m'_j совпадают с моментами случайной величины Y . Поэтому имеет место разложение (20.30)

$$\exp \left\{ \sum_{\nu=3}^{\infty} \frac{\kappa'_\nu}{\nu!} (iu)^\nu \right\} = 1 + \sum_{j=3}^{\infty} \frac{m'_j}{j!} (iu)^j, \quad (22.36)$$

где моменты m'_j могут быть выражены через семиинварианты κ'_j по формулам (20.39). Подставляя (22.36) в (22.35), находим

$$E_y(u) = e^{-\frac{u^2}{2}} + e^{-\frac{u^2}{2}} \cdot \sum_{j=3}^{\infty} \frac{m'_j}{j!} (iu)^j. \quad (22.37)$$

Применяя к обеим частям последнего равенства обратное преобразование Фурье, получим для плотности вероятности случайной величины Y выражение

$$\begin{aligned} f(y) = & \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iuy - \frac{u^2}{2}} du + \\ & + \sum_{j=3}^{\infty} \frac{m'_j (-1)^j}{j!} \cdot \frac{d^j}{dy^j} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iuy - \frac{u^2}{2}} du \right]. \end{aligned} \quad (22.38)$$

Первое слагаемое в (22.38) есть плотность вероятности $\varphi(y)$ нормированной центрированной случайной величины, определяемой так:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iuy - \frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} \equiv \varphi(y). \quad (22.39)$$

Следовательно, формуле (22.38) можно придать вид

$$f(y) = \varphi(y) + \sum_{j=3}^{\infty} \frac{(-1)^j}{j!} m'_j \varphi^{(j)}(y), \quad (22.40)$$

где производные $\varphi^{(j)}(y)$ связаны с полиномами Чебышева–Эрмита⁶ [17]

$$H_j(x) = (-1)^j e^{x^2} \frac{d^j}{dx^j} (e^{-x^2}) \quad (22.41)$$

⁶ Эрмит Шарль (1822–1901) — французский математик.

соотношением

$$\varphi^{(j)}(y) = \frac{1}{2^j \sqrt{2\pi}} e^{-y^2/2} H_j \left(\frac{y}{\sqrt{2}} \right). \quad (22.42)$$

Обращаясь к формулам (20.38), с учетом (22.34), получим

$$\begin{aligned} \kappa'_3 &= Sk; & \kappa'_4 &= Ex; \\ \kappa'_5 &= \frac{\mu_5}{\sigma_x^5} - 10Sk; & \kappa'_6 &= \frac{\mu_6}{\sigma_x^6} - 15Ex - 10Sk^2 - 15. \end{aligned} \quad (22.43)$$

Подставляя эти выражения в (22.40), получим искомый ряд для плотности вероятности нормированной и центрированной случайной величины $Y = \frac{X - \bar{x}}{\sigma_x}$:

$$\begin{aligned} f(y) &= \varphi(y) + \frac{1}{6} Sk \varphi^{(3)}(y) + \\ &+ \frac{1}{24} Ex \varphi^{(4)}(y) - \frac{1}{120} \left(\frac{\mu_5}{\sigma_x^5} - 10Sk \right) \varphi^{(5)}(y) + \\ &+ \frac{1}{720} \left(\frac{\mu_6}{\sigma_x^6} - 15Ex - 10Sk^2 - 15 \right) \varphi^{(6)}(y), \end{aligned} \quad (22.44)$$

где асимметрия Sk , эксцесс Ex и все центральные моменты μ_j относятся к исходной случайной величине X .

Ряд (22.44) называется рядом Грама⁷–Шарлье⁸ типа А. Учитывая (22.42), этот ряд является разложением функции $f(y)$ по ортогональным функциям Эрмита [17] $e^{-y^2/2} H_j(y)$ и, следовательно, сходимость этого ряда может быть доказана с помощью признаков, существующих для разложений подобного типа. Обычно необходимости в подобном доказательстве не возникает, так как этот ряд используется в качестве приближенного представления плотности вероятности $f(y)$ и приме-

⁷ Грам Йорген (1850–1916) — датский математик.

⁸ Шарлье Карл Вильгельм Людвиг (1862–1934) — шведский астроном и математик.

няется только для тех распределений, у которых первые три-четыре члена ряда дают достаточную точность. Использование ряда в этом случае упрощается тем, что функции $\varphi^{(j)}(y)$ для первых значений j индекса табулированы и приводятся в таблицах для вероятностных расчетов [19].

Рассмотрим примеры.

Пример 22.1. Дана плотность вероятности случайной величины X

$$f(x) = k \cdot 2^{-0,1x^2+2x}, \text{ где } k — \text{некоторая постоянная.}$$

Определить $\bar{x} = M(X)$, $\sigma^2 = D(X)$ и представить плотность вероятности в канонической форме.

Решение. Случайная величина X — нормальная, так как $f(x)$ можно представить в виде экспоненты, имеющей полином второй степени в показателе. Воспользовавшись канонической формой плотности вероятности нормального закона распределения (22.11), имеем

$$k \cdot 2^{-0,1x^2+2x} = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\bar{x})^2/2\sigma^2}.$$

Логарифмируя, получим

$$-\frac{(x-\bar{x})^2}{2\sigma^2} \cdot \lg e + \lg \left(\frac{1}{\sigma\sqrt{2\pi}} \right) = (-0,1x^2 + 2x) \lg 2 + \lg k.$$

Приравнявая коэффициенты в обеих частях равенства при одинаковых степенях x , находим

$$\begin{aligned} -\frac{1}{2\sigma^2} \cdot 0,4343 &= -0,1 \cdot 0,3010; \\ \frac{\bar{x}}{\sigma^2} \cdot 0,4343 &= 2 \cdot 0,3010, \end{aligned}$$

откуда следует:

$$\begin{aligned} \sigma^2 &= \frac{0,4343}{2 \cdot 0,1 \cdot 0,3010} \approx 7,239, \quad \sigma \approx 2,62; \\ \bar{x} &= 2 \cdot \frac{0,3010}{0,4343} \cdot \sigma^2 = 10. \end{aligned}$$

Таким образом, окончательно

$$f(x) = \frac{1}{2,62\sqrt{2\pi}} \exp \left\{ -\frac{(x-10)^2}{2 \cdot 2,62^2} \right\}.$$

Пример 22.2.

Измерение расстояния между двумя точками на местности сопровождается систематической ошибкой $\bar{x} = -20$ см и случайной ошибкой $\overset{\circ}{X}$, которую можно считать центрированной нормальной величиной, имеющей дисперсию $\sigma_x^2 = 25$ см². Определить вероятность того, что суммарная ошибка по модулю будет меньше $\varepsilon = 25$ см.

Решение. Так как прибавление к нормальной случайной величине постоянной не меняет тип закона распределения, то суммарная ошибка $X = \overset{\circ}{X} + \bar{x}$ — нормальная величина с математическим ожиданием $\bar{x}$ и дисперсией σ_x^2 . Требуется определить $\mathbf{P}\{|X| \leq \varepsilon\}$. Используя формулу (22.19), имеем

$$\begin{aligned} \mathbf{P}\{|X| \leq \varepsilon\} &= \mathbf{P}\{-\varepsilon \leq X \leq \varepsilon\} = \\ &= \frac{1}{2} \cdot \left[\Phi\left(\frac{\varepsilon - \bar{x}}{\sigma_x}\right) - \Phi\left(\frac{\varepsilon - \bar{x}}{\sigma_x}\right) \right] = \\ &= \frac{1}{2} \cdot \left[\Phi\left(\frac{25}{5}\right) + \Phi\left(\frac{5}{5}\right) \right] = \frac{1}{2} \cdot [1 + \Phi(1)] = 0,842, \end{aligned}$$

где учтено, что $\Phi(5) \approx 1$, $\Phi(1) = 0,683$.

Пример 22.3. Вывести ряд для обратной функции Лапласа, то есть выразить зависимость x от y , если $y = \Phi(x)$. Пользуясь полученным рядом, выразить срединное отклонение нормального закона распределения E через среднее квадратическое отклонение σ .

Решение. Так как функция $\Phi(x)$ — нечетная, то искомым ряд можно представить в виде

$$x = y(a_0 + a_1 y^2 + a_2 y^4 + \dots).$$

Подставляя этот ряд в ряд для функции Лапласа

$$y = \frac{2}{\sqrt{2\pi}} \cdot \sum_{j=0}^{\infty} (-1)^j \frac{1}{(2j+1)2^j j!} x^{2j+1}$$

и приравнявая коэффициенты при одинаковых степенях y , для первых трех коэффициентов находим

$$a_0 = \sqrt{\frac{\pi}{2}}, \quad a_1 = \frac{1}{6} a_0^3 = \frac{\pi}{12} \cdot \sqrt{\frac{\pi}{2}},$$

$$a_2 = \frac{7}{120} \cdot \frac{\pi^2}{4} \sqrt{\frac{\pi}{2}}, \quad a_3 = \left(\frac{\pi}{2}\right)^3 \cdot \sqrt{\frac{\pi}{2}} \cdot \frac{127}{24 \cdot 210}.$$

Если $x = E$, то $y = \Phi\left(\frac{E}{\sigma}\right) = \frac{1}{2}$. Следовательно, положив в полученном ряде $x = \frac{E}{\sigma} = \rho\sqrt{2}$, а $y = \frac{1}{2}$, получаем младшие члены ряда для определения ρ :

$$\rho = \frac{\sqrt{\pi}}{4} \left(1 + \frac{\pi}{48} + \frac{7\pi^2}{120 \cdot 4 \cdot 16} + \frac{127 \cdot \pi^3}{8 \cdot 24 \cdot 210 \cdot 64} + \dots \right).$$

Подставляя заданные числа, находим

$$\rho \cong 0,4768.$$

Точное значение $\rho = 0,4769$.

§ 23. ДРУГИЕ ПРИМЕРЫ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Рассмотрим еще несколько законов распределения непрерывных случайных величин, которые наряду с нормальным законом часто встречаются в приложениях.

23.1. Равномерное распределение

Равномерным распределением называется такое распределение непрерывной случайной величины X , при котором плотность вероятности $f(x)$ имеет постоянное значение в заданном конечном интервале (a, b) и равна нулю вне этого интервала.

В силу выполнения условия нормировки для f имеем

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x < a, \quad x > b. \end{cases} \quad (23.1)$$

Равномерному закону распределения подчиняются непрерывная случайная величина, вероятность попадания значения которой в любые равновеликие отрезки интервала (a, b) одинаковы. Примером таких случайных величин являются ошибки округления, значение координаты точки, выбранной наудачу на заданном отрезке прямой и т. п.

Вероятность попадания равномерно распределенной случайной величины X в интервал (x_1, x_2) , где $a \leq x_1 \leq x_2 \leq b$, определяется равенством

$$\mathbf{P}\{x_1 < X < x_2\} = F(x_2) - F(x_1) = \frac{x_2 - x_1}{b - a}. \quad (23.2)$$

Подставляя выражение (23.1) для $f(x)$ в общую формулу для начального момента k -го порядка m_k , получим

$$m_k = \int_a^b x^k f(x) dx = \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)}, \quad (k = \overline{1, \infty}). \quad (23.3)$$

При $k = 1$ из (23.3) получаем математическое ожидание случайной величины X

$$\bar{x} = \frac{1}{2}(a+b). \quad (23.4)$$

Для центрального момента k -го порядка μ_k аналогичным образом находим

$$\begin{aligned} \mu_k &= \int_a^b (x - \bar{x})^k f(x) dx = \\ &= \frac{1}{(k+1)(b-a)} \cdot \left[\left(\frac{b-a}{2} \right)^{k+1} - \left(\frac{a-b}{2} \right)^{k+1} \right]. \end{aligned}$$

Из этого равенства следует, что

$$\mu_{2k+1} = 0, \quad \mu_{2k} = \frac{(b-a)^{2k}}{2^{2k}(2k+1)} \quad (k = \overline{0, \infty}). \quad (23.5)$$

Полагая в последнем равенстве $k = 1$, получим выражение для дисперсии $D(X) = \mu_2$ и среднего квадратичного отклонения равномерно распределенной случайной величины

$$D(X) = \frac{(b-a)^2}{12}, \quad \sigma = \frac{b-a}{2\sqrt{3}}. \quad (23.6)$$

Интегрируя плотность вероятности (23.1) от $-\infty$ до x , получим функцию распределения $F(x)$ случайной величины, равномерно распределенной в интервале (a, b) в форме

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases} \quad (23.7)$$

Графики функций $f(x)$ и $F(x)$ приведены на рис. 23.1.

Для характеристической функции $E(u)$ с помощью (23.1) получим

$$E(u) = \frac{1}{b-a} \int_a^b e^{iux} dx = \frac{1}{iu(b-a)} (e^{ibu} - e^{iau}).$$

Вместо параметров a и b , полностью определяющих равномерный закон распределения, часто удобно использовать параметры $\bar{x} = 1/2 (a + b)$ и $l = 1/2 (b - a)$, из которых второй обычно называют просто параметром закона равномерного распределения, а первый — параметром сдвига. В этом случае плотность вероятности (23.1) может быть записана в виде

$$f(x) = \begin{cases} \frac{1}{2l} & \text{при } |x - \bar{x}| \leq l, \\ 0 & \text{при } |x - \bar{x}| > l. \end{cases} \quad (23.8)$$

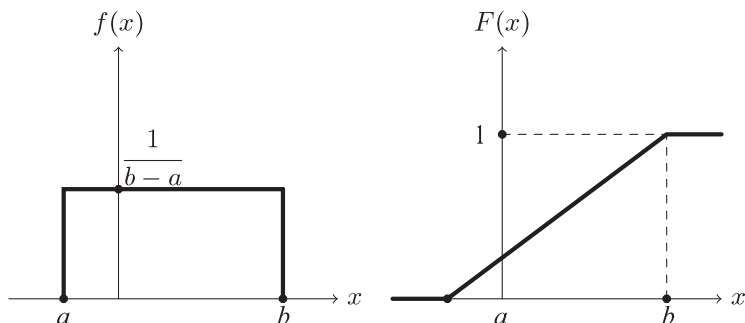


Рис. 23.1

*Плотность вероятности и функция распределения
равномерного закона*

Для характеристической функции $E(u)$ при этом получим более компактное выражение:

$$E(u) = e^{i\bar{x}u} \frac{\sin lu}{lu}. \quad (23.9)$$

Формула (23.6) для дисперсии и среднего квадратического отклонения здесь принимает вид

$$D(X) = \frac{l^2}{3}; \quad \sigma = \frac{l}{\sqrt{3}}. \quad (23.10)$$

Нетрудно видеть, что срединное отклонение $E = \frac{l}{2}$. Следовательно, для равномерно распределенной случайной величины X срединное отклонение E и среднее квадратическое отклонение σ связаны равенством

$$E = \frac{\sqrt{3}}{2} \sigma. \quad (23.11)$$

Равномерное распределение не имеет моды, а его медиана совпадает с математическим ожиданием $\bar{x}$.

Для равномерного закона распределения можно вывести общую формулу для семиинвариантов, если воспользоваться

определением чисел Бернулли, как коэффициентов в разложении функции $\frac{t}{e^t - 1}$ по степеням t [1]:

$$\frac{t}{e^t - 1} = \sum_{j=0}^{\infty} \frac{B_j}{j!} t^j, \quad (23.12)$$

из которой путем разложения левой части можно получить, что

$$\begin{aligned} B_0 &= 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \\ B_4 &= -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_{2j+1} = 0, \quad (j \geq 1). \end{aligned} \quad (23.13)$$

Если первые два члена суммы в (23.12) перенести в левую часть, то после деления обеих частей равенства на t получим

$$\frac{1}{e^t - 1} - \frac{1}{t} + \frac{1}{2} = \sum_{j=2}^{\infty} \frac{B_j}{j!} t^{j-1},$$

что эквивалентно (23.14)

$$\frac{e^{t/2} + e^{-t/2}}{2(e^{t/2} - e^{-t/2})} - \frac{1}{t} = \sum_{j=2}^{\infty} \frac{B_j}{j!} t^{j-1}.$$

Интегрирование равенства (23.14) по t от 0 до $2ilu$ дает

$$\ln \left(\frac{\sin lu}{lu} \right) = \sum_{j=2}^{\infty} \frac{B_j}{j \cdot j!} (2l)^j (iu)^j. \quad (23.15)$$

Так как в соответствии с (23.9) $\frac{\sin lu}{lu} = \overset{\circ}{E}(u)$ — характеристической функции центрированной равномерно распределенной случайной величины с параметром l , то, учитывая (20.37) и (23.13), будем иметь

$$\sum_{j=2}^{\infty} \frac{\kappa_j}{j!} (iu)^j = \sum_{j=2}^{\infty} \frac{B_j}{2j \cdot (2j)!} (2l)^{2j} (iu)^{2j}. \quad (23.16)$$

Таким образом, поскольку все нечетные числа Бернулли с индексами большими единицы, обращаются в нуль, все нечетные семиинварианты равномерного закона распределения равны нулю, а для четных семиинвариантов получено выражение

$$\kappa_{2j} = \frac{(2l)^{2j}}{2j} B_{2j}, \quad (j = \overline{1, \infty}). \quad (23.17)$$

В частности, отсюда находим

$$\kappa_2 = \frac{l^2}{3}, \quad \kappa_4 = -\frac{2l^4}{15}, \quad (23.18)$$

что дает для эксцесса следующее значение:

$$Ex = \frac{\kappa_4}{\kappa_2^2} = -1,2.$$

23.2. Распределение Коши

Примером непрерывной случайной величины, для которой не существуют моменты, уже начиная с первого, является случайная величина X , распределенная по закону Коши⁹ (распределение арктангенса). В этом случае плотность вероятности имеет вид

$$f(x) = \frac{a}{\pi(x^2 + a^2)}, \quad (-\infty < x < \infty), \quad (23.19)$$

где a — заданная положительная постоянная. Функция распределения в данном случае определяется формулой

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctg \frac{x}{a}. \quad (23.20)$$

Графики функций $f(x)$ и $F(x)$ приведены на рис. 23.2, из которых видно, что качественно эти кривые имеют такой же вид, что и графики аналогичных функций для нормального закона распределения (см. рис. 23.1).

⁹ Коши Огюст Луи (1789–1857) — французский математик.

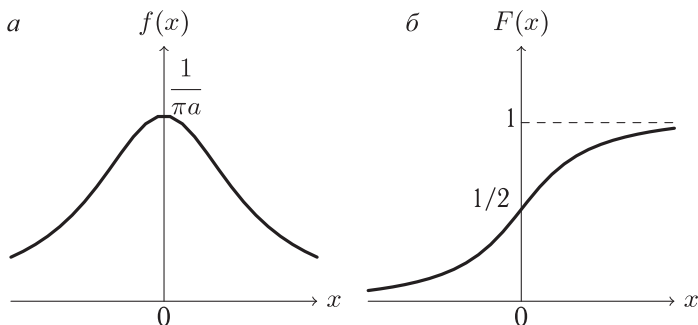


Рис. 23.2

Плотность вероятности и функция распределения
закона Коши (арктангенса)

Характеристическая функция распределения Коши будет такой:

$$E(u) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{iux}}{x^2 + a^2} dx. \quad (23.21)$$

Подынтегральное выражение в (23.21) имеет два простых полюса в точках $x = \pm ia$. Дополняя вещественную ось полуокружностью при $u > 0$ в верхней полуплоскости, а при $u < 0$ — в нижней полуплоскости, в соответствии с общей теорией вычетов [7] получим при $u \geq 0$

$$\int_{-\infty}^{\infty} \frac{e^{iux}}{x^2 + a^2} dx = 2\pi i \left. \frac{e^{iux}}{x + ia} \right|_{x=ia} = \frac{\pi}{a} e^{-au}.$$

При $u \leq 0$ тем же способом находим

$$\int_{-\infty}^{\infty} \frac{e^{iux}}{x^2 + a^2} dx = -2\pi i \left. \frac{e^{iux}}{x - ia} \right|_{x=-ia} = \frac{\pi}{a} e^{au}.$$

Значит, при любом u характеристическая функция распределения Коши определяется формулой

$$E(u) = e^{-a|u|}. \quad (23.22)$$

Распределение Коши имеет, например, случайная величина $X = a \operatorname{tg} \Phi$, где угол Φ является случайной величиной, равномерно распределенной в интервале $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Такое же распределение имеет отношение $X = \frac{Y}{Z}$, в котором Y и Z — независимые нормальные центрированные случайные величины (в этом случае параметр $a = \frac{\sigma_y}{\sigma_z}$).

Используя равенство (23.20), для вероятности попадания случайной величины X в интервал (x_1, x_2) , получим

$$\mathbf{P}\{x_1 < X < x_2\} = \frac{1}{\pi} \left(\operatorname{arctg} \frac{x_2}{a} - \operatorname{arctg} \frac{x_1}{a} \right)$$

или, применяя известные формулы тригонометрии,

$$\mathbf{P}\{x_1 < X < x_2\} = \frac{1}{\pi} \operatorname{arctg} \frac{a(x_2 - x_1)}{a^2 + x_1 x_2}. \quad (23.23)$$

Из рис. 23.2а видно, что мода и медиана распределения Коши равны нулю. Несмотря на симметричный вид кривой $f(x)$, закон распределения Коши не имеет математического ожидания, так как соответствующий интеграл расходится:

$$\begin{aligned} \int_{-\infty}^{\infty} |x| f(x) dx &= \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{x^2 + a^2} dx = \\ &= \frac{2a}{\pi} \int_0^{\infty} \frac{x}{x^2 + a^2} dx = \frac{2a}{\pi} \ln(x^2 + a^2) \Big|_0^{\infty} = \infty. \end{aligned}$$

Очевидно, что для рассматриваемой случайной величины тем более не существуют никакие моменты более высокого порядка. Таким образом, несмотря на кажущееся сходство кривых $f(x)$ и $F(x)$ для закона распределения Коши с аналогичными кривыми для нормального закона распределения, свойства этих законов резко отличаются друг от друга.

23.3. Распределение арксинуса

Примером *антимодального* закона распределения является распределение арксинуса. В этом случае плотность вероятности и функция распределения определяются формулами

$$f(x) = \begin{cases} \frac{1}{\pi\sqrt{a^2 - x^2}}, & |x| < a, \\ 0, & |x| \geq a, \end{cases} \quad (23.24)$$

$$F(x) = \begin{cases} 0, & x \leq -a, \\ \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin \frac{x}{a}, & |x| < a, \\ 1, & x > a. \end{cases} \quad (23.25)$$

Графики этих функций приведены на рис. 23.3.

Характеристическая функция случайной величины X будет иметь вид

$$E(u) = \frac{1}{\pi} \int_0^{\pi} e^{iua \cos \varphi} d\varphi = J_0(au),$$

где $J_0(z)$ обозначает функцию Бесселя первого рода нулевого порядка [17]. В результате для характеристической функции

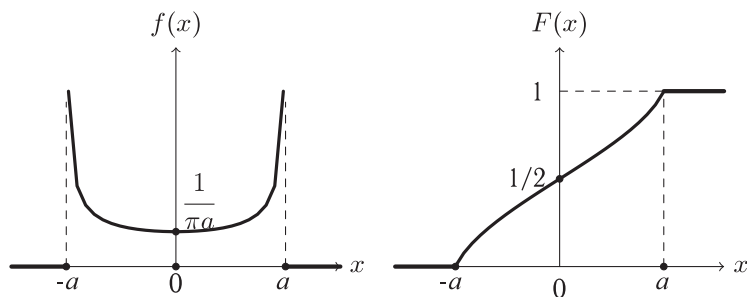


Рис. 23.3

Плотность вероятности и функция распределения
закона арксинуса

получим

$$E(u) = J_0(au). \quad (23.26)$$

Воспользовавшись рядом для функции Бесселя

$$J_0(z) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{z}{2}\right)^{2k}}{(k!)^2},$$

характеристическую функцию (23.26) можно представить в виде

$$E(u) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k!)} \cdot \frac{(2k)!}{2^{2k} \cdot (k!)^2} \cdot u^{2k}. \quad (23.27)$$

В соответствии с (20.30) характеристическая функция является производящей функцией для начальных моментов (считаем все моменты существующими)

$$E(u) = \sum_{k=0}^{\infty} \frac{(iu)^k}{k!} m_k. \quad (23.28)$$

Сравнивая (23.28) с (23.27), заключаем, что все нечетные моменты распределения арксинуса равны нулю. Следовательно, начальные моменты совпадают с центральными, для которых справедлива формула

$$\mu_{2k} = m_{2k} = \frac{(2k)!}{2^{2k} (k!)^2} a^{2k}; \quad \mu_{2k+1} = m_{2k+1} = 0. \quad (23.29)$$

В частности, дисперсия $\sigma^2 = \mu_2 = \frac{a^2}{2}$, а четвертый момент $\mu_4 = \frac{3}{8} a^4$ и, следовательно, эксцесс, также, как для равномерного закона, будет отрицательным

$$E_x = \frac{\mu_4}{\mu_2^2} - 3 = -1,5.$$

Для закона арксинуса срединное отклонение E равно среднему квадратическому σ . Действительно, учитывая определение

срединного отклонения (19.24) и выражение для функции распределения (23.25), получим

$$\frac{2}{\pi} \arcsin \frac{E}{a} = \frac{1}{2}, \quad \text{откуда} \quad \arcsin \frac{E}{a} = \frac{\pi}{4},$$

$$\frac{E}{a} = \frac{\sqrt{2}}{2}, \quad E = \frac{\sqrt{2}}{2} a = \sigma.$$

23.4. Распределение Рэлея

Распределением Рэлея¹⁰ называется закон распределения случайной величины X , для которой плотность вероятности определяется формулой

$$f(x) = \begin{cases} \frac{x}{a^2} e^{-x^2/2a^2}, & x \geq 0; \\ 0, & x < 0. \end{cases} \quad (23.30)$$

Интегрируя данное выражение от $-\infty$ до x , получим функцию распределения

$$F(x) = \begin{cases} 0, & x < 0; \\ 1 - e^{-x^2/2a^2}, & x \geq 0. \end{cases} \quad (23.31)$$

Распределению Рэлея подчиняется, например, расстояние случайной точки на плоскости от начала координат, если прямоугольные координаты этой точки являются независимыми центрированными нормальными случайными величинами с одинаковыми дисперсиями.

Характеристическая функция распределения Рэлея не сводится к элементарным функциям, но может быть выражена через хорошо изученную и табулированную функцию, называемую интегралом вероятности от комплексного аргумента и определяемую равенством [17]

$$W(z) = e^{-z^2} \left(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2} dt \right). \quad (23.32)$$

Графики функций $f(x)$ и $F(x)$ приведены на рис. 23.4.

¹⁰ Рэлей (Стретт) Джон Уильям (1842–1919) — английский физик.

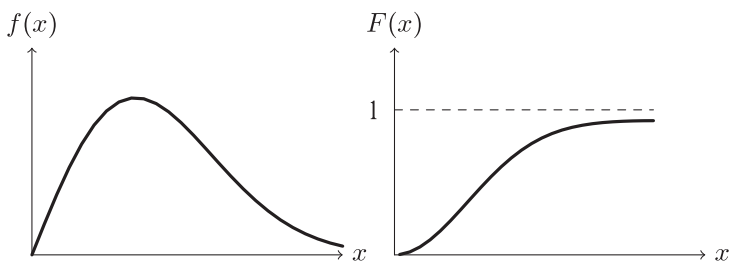


Рис. 23.4

Плотность вероятности и функция распределения
закона Рэля

Действительно, для характеристической функции распределения Рэля имеем

$$E(u) = \frac{1}{a^2} \int_0^{\infty} \exp \left\{ iux - \frac{x^2}{2a^2} \right\} x dx. \quad (23.33)$$

Положив $x = -ia\sqrt{2}t + ia^2u$, после простых преобразований получим

$$E(u) = 1 + iau\sqrt{\frac{\pi}{2}} e^{-a^2u^2/2} \left[1 + \frac{2i}{\sqrt{\pi}} \int_0^{au/\sqrt{2}} e^{v^2} dv \right] \quad (23.34)$$

или, учитывая (23.32),

$$E(u) = 1 + iau\sqrt{\frac{\pi}{2}} W \left(\frac{au}{\sqrt{2}} \right). \quad (23.35)$$

Вероятность попадания случайной величины X в интервал (x_1, x_2) , где $x_1 \geq 0$, для распределения Рэля определяется простой формулой:

$$\mathbf{P}\{x_1 < X < x_2\} = e^{-x_1^2/2a^2} - e^{-x_2^2/2a^2}. \quad (23.36)$$

Начальный момент k -го порядка будет таким:

$$m_k = \frac{1}{a^2} \int_0^{\infty} x^{k+1} e^{-x^2/2a^2} dx.$$

Для вычисления последнего интеграла примем $\frac{x^2}{2a^2} = y$, тогда

$$m_k = (a\sqrt{2})^k \int_0^{\infty} y^{k/2} e^{-y} dy = (a\sqrt{2})^k \Gamma\left(\frac{k}{2} + 1\right).$$

Воспользовавшись свойствами гамма-функции [17] и, учитывая, что $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, окончательно получим

$$m_{2k} = 2^k k! a^{2k}; \quad m_{2k+1} = (2k+1)!! \sqrt{\frac{\pi}{2}} \cdot a^{2k+1}. \quad (23.37)$$

Для математического ожидания $\bar{x} = m_1$ и дисперсии $D(X) = m_2 - \bar{x}^2$ из (23.37) получим следующее:

$$\bar{x} = a \sqrt{\frac{\pi}{2}}; \quad D(X) = 2a^2 \left(1 - \frac{\pi}{4}\right). \quad (23.38)$$

Третий начальный момент $m_3 = 3 \sqrt{\frac{\pi}{2}} a^3$, поэтому третий центральный момент будет выражен в виде

$$\mu_3 = m_3 - 3\bar{x}m_2 + 2\bar{x}^3 = a^3(\pi - 3) \frac{\pi}{2},$$

а коэффициент асимметрии не зависит от a и дается равенством

$$Sk = \frac{\mu_3}{\sigma^3} = \frac{(\pi - 3)\sqrt{\pi}}{4(1 - \pi/4)^{3/2}} \approx 0,631. \quad (23.39)$$

Четвертый центральный момент оказывается равным

$$\mu_4 = m_4 - 4\bar{x}m_3 + 6\bar{x}^2m_2 - 3\bar{x}^4 = a^4 \left(8 - \frac{3}{4} \pi^2 \right),$$

а потому эксцесс также не зависит от a

$$E_x = \frac{\mu_4}{\sigma^4} - 3 = \frac{3/2\pi \left(1 - \frac{\pi}{4} \right) - 1}{\left(1 - \frac{\pi}{4} \right)^2} \approx 0,245. \quad (23.40)$$

Плотность вероятности $f(x)$ принимает максимальное значение при $x = a$, поэтому мода распределения Рэля всегда равна a . Условие $\mathbf{P}\{X \leq x\} = 0,5$ в данном случае имеет вид $1 - e^{-x^2/(2a^2)} = 0,5$ или $x^2/(2a^2) = \ln 2$. Из последнего выражения следует, что медиана распределения Рэля $M_2 = a\sqrt{2 \ln 2}$.

23.5. Распределение Максвелла

Случайная величина X подчиняется закону Максвелла¹¹, если ее плотность вероятности определяется формулой

$$f(x) = \begin{cases} \frac{2x^2}{a^3\sqrt{2\pi}} e^{-x^2/2a^2}, & x \geq 0; \\ 0, & x < 0. \end{cases} \quad (23.41)$$

Такое распределение имеет, например, модуль трехмерного вектора, если все три его составляющие являются независимыми нормальными случайными величинами с одинаковыми дисперсиями и нулевыми математическими ожиданиями.

При $x < 0$ функция распределения $F(x)$ равна нулю, а при $x \geq 0$ она дается интегралом

$$F(x) = \frac{2}{a^3\sqrt{2\pi}} \int_0^x x_1^2 e^{-x_1^2/2a^2} dx_1.$$

¹¹ Максвелл Джеймс Клерк (1831–1879) — шотландский физик.

Интегрируя здесь по частям, получим

$$\begin{aligned} \int_0^x x_1^2 e^{-x_1^2/2a^2} dx_1 &= \\ &= xa^2 e^{-x^2/2a^2} + a^3 \int_0^{x/a} e^{-x_1^2/2a^2} dx_1 = \\ &= -xa^2 e^{-x^2/2a^2} + \frac{a^3 \sqrt{2\pi}}{2} \Phi\left(\frac{x}{a}\right), \end{aligned}$$

где $\Phi(x)$ обозначает интеграл вероятностей.

Следовательно, окончательно

$$F(x) = \begin{cases} 0, & x < 0; \\ \Phi\left(\frac{x}{a}\right) - \frac{2x}{a\sqrt{2\pi}} e^{-x^2/2a^2}, & x \geq 0. \end{cases} \quad (23.42)$$

Графики функций $f(x)$ и $F(x)$ приведены на рис. 23.5.

Характеристическая функция распределения Максвелла дается интегралом

$$E(u) = \frac{2}{a^3 \sqrt{2\pi}} \int_0^\infty x^2 \exp\left\{iux - \frac{x^2}{2a^2}\right\} dx.$$

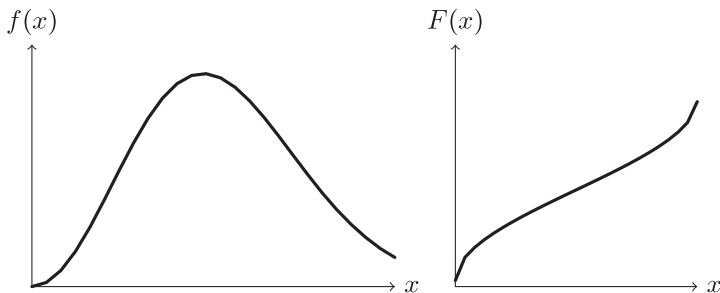


Рис. 23.5

Плотность вероятности и функция распределения
закона Максвелла

Сделав замену переменных $x = -ia\sqrt{2}t + ia^2u$, после простых преобразований, получим

$$E(u) = (1 - a^2u^2)W\left(\frac{au}{\sqrt{2}}\right) + iau\sqrt{\frac{2}{\pi}}, \quad (23.43)$$

где $W(z)$ по-прежнему обозначает интеграл вероятности от комплексного аргумента, определяемый формулой (23.32).

Вероятность попадания случайной величины X , подчиняющейся закону распределения Максвелла, в интервал (x_1, x_2) , где $x_1 \geq 0$ определяется формулой

$$\begin{aligned} \mathbf{P}\{x_1 < X < x_2\} &= \Phi\left(\frac{x_2}{a}\right) - \Phi\left(\frac{x_1}{a}\right) - \\ &- \frac{a}{a\sqrt{2\pi}}\left(x_2 e^{-x_2^2/2a^2} - x_1 e^{-x_1^2/2a^2}\right). \end{aligned} \quad (23.44)$$

Начальный момент k -го порядка m_k будет таким:

$$\begin{aligned} m_k &= \frac{2}{a^3\sqrt{2\pi}} \int_0^\infty x^{k+2} \exp\left\{-\frac{x^2}{2a^2}\right\} dx = \\ &= \frac{a^k 2^{k/2+1}}{\sqrt{\pi}} \int_0^\infty e^{-t} t^{(k+1)/2} dt = \\ &= \frac{a^k 2^{k/2+1}}{\sqrt{\pi}} \Gamma\left(\frac{k+3}{2}\right), \quad (k = \overline{1, \infty}). \end{aligned}$$

Учитывая свойства гамма-функции, получим

$$\begin{aligned} m_{2k} &= a^{2k}(2k+1)!!, \\ m_{2k+1} &= \frac{a^{2k+1}}{\sqrt{\pi}} 2^{k+3/2}(k+1)!, \quad (k = \overline{1, \infty}). \end{aligned} \quad (23.45)$$

Из (23.45) определим $m_1 = \frac{2\sqrt{2}}{\sqrt{\pi}} a$, $m_2 = 3a^2$, откуда легко находим

$$\bar{x} = \frac{2\sqrt{2}}{\sqrt{\pi}} a; \quad D(X) = \frac{3\pi - 8}{\pi} a^2. \quad (23.46)$$

Третий начальный момент $m_3 = \frac{8\sqrt{2}}{\sqrt{\pi}} a^3$, а третий центральный момент

$$\mu_3 = m_3 - 3\bar{x}m_2 + 2\bar{x}^3 = \frac{2\sqrt{2}(16 - 5\pi)}{\pi\sqrt{\pi}} a^3.$$

Коэффициент асимметрии не зависит от параметра a

$$Sk = \frac{\mu_3}{\sigma^3} = \frac{2\sqrt{2}(16 - 5\pi)}{(3\pi - 8)^{3/2}} \approx 0,486,$$

и несколько меньше, чем в законе Рэлея. Четвертый начальный момент $m_4 = 15a^4$, а четвертый центральный момент

$$\mu_4 = (15\pi^2 + 16\pi - 192) \cdot \frac{a^4}{\pi^2}.$$

Поэтому эксцесс

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{160\pi - 12\pi^2 - 384}{(3\pi - 8)^2} \approx 0,108,$$

что также меньше, чем в случае закона Рэлея.

Максимального значения функция $f(x)$ достигает при $x = a\sqrt{2}$, поэтому мода распределения Максвелла равна $a\sqrt{2}$. Условие $\mathbf{P}\{X \leq x\} = 0,5$ в данном случае имеет вид

$$\Phi\left(\frac{x}{a}\right) - \frac{2x}{a\sqrt{2\pi}} e^{-x^2/2a^2} = 0,5.$$

Приближенное значение решения этого уравнения (медиана) таково:

$$M_e \approx 1,54a.$$

23.6. Гамма-распределение

В теории надежности и теории массового обслуживания широко используется так называемое гамма-распределение, плотность вероятности которого определяется формулой

$$f(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, & x \geq 0; \\ 0, & x < 0, \end{cases} \quad (23.47)$$

где α и λ — заданные положительные постоянные, а $\Gamma(\alpha)$ — обозначение гамма-функции от аргумента α [17].

Интегрируя плотность вероятности (23.47) по x от $-\infty$ до x , получим функцию распределения

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{\gamma(\alpha, \lambda x)}{\Gamma(\alpha)}, & x \geq 0, \end{cases} \quad (23.48)$$

где

$$\gamma(\alpha, x) = \int_0^x e^{-t} t^{\alpha-1} dt \quad (23.49)$$

обозначает неполную гамма-функцию, для которой имеются таблицы с входами x и α [19].

Функцию распределения можно также выразить через другую табличную функцию

$$\gamma_1(\beta, x) = \frac{1}{\Gamma(\beta+1)} \int_0^{x\sqrt{1+\beta}} e^{-t} t^\beta dt,$$

которая также иногда называется неполной гамма-функцией. При этом оказывается

$$F(x) = \begin{cases} 0, & x < 0, \\ \gamma_1\left(\alpha-1, \frac{\lambda x}{\sqrt{\alpha}}\right), & x \geq 0; \end{cases} \quad (23.50)$$

Имеем при $\alpha \geq 1$

$$\gamma(\alpha, x) = -x^{\alpha-1}e^{-x} + (\alpha-1) \int_0^x e^{-t} t^{\alpha-2} dt.$$

Поэтому при целом α

$$\gamma(\alpha, x) = -(\alpha-1)! e^{-x} \left[\frac{x^{\alpha-1}}{(\alpha-1)!} + \frac{x^{\alpha-2}}{(\alpha-2)!} + \dots + \frac{x}{1!} + 1 - e^x \right].$$

Следовательно, функция распределения $F(x)$ при целом α определяется выражением

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x} \sum_{k=0}^{\alpha-1} \frac{(\lambda x)^k}{k!}, & x \geq 0. \end{cases} \quad (23.51)$$

Графики функций $f(x)$ и $F(x)$ при $\alpha = 2$ приведены на рис. 23.6.

Если $\alpha < 1$, то плотность вероятности асимптотически приближается к оси ординат. При $\alpha > 1$ эта функция имеет экстремальное значение в точке $x = \frac{\alpha-1}{\lambda}$. Поэтому мода

гамма-распределения существует при $\alpha > 1$ и равна $\frac{\alpha-1}{\lambda}$.

Характеристическая функция гамма-распределения будет такой:

$$\begin{aligned} E(u) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-x(\lambda-iu)} x^{\alpha-1} dx = \\ &= \frac{1}{\Gamma(\alpha)} \left(\frac{\lambda}{\lambda-iu} \right)^\alpha \int_0^\infty e^{-t} t^{\alpha-1} dt. \end{aligned}$$

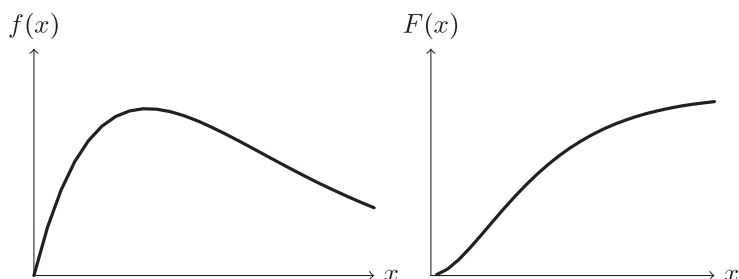


Рис. 23.6

*Плотность вероятности и функция распределения
гамма-распределения*

Так как фигурирующий здесь интеграл легко выражается через гамма-функцию, получим

$$E(u) = \left(1 - \frac{iu}{\lambda}\right)^{-\alpha}. \quad (23.52)$$

Начальный момент k -го порядка m_k для гамма-распределения определяется равенством

$$\begin{aligned} m_k &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda x} x^{k+\alpha-1} dx = \\ &= \frac{1}{\lambda^k \Gamma(\alpha)} \int_0^\infty e^{-t} t^{k+\alpha-1} dt, \end{aligned}$$

что дает после выражения последнего интеграла через гамма-функцию

$$m_k = \frac{\Gamma(k + \alpha)}{\lambda^k \Gamma(\alpha)}, \quad (k = \overline{1, \infty}). \quad (23.53)$$

С помощью (23.53) находим математическое ожидание и дисперсию случайной величины X :

$$\bar{x} = \frac{\alpha}{\lambda}; \quad D(X) = \frac{\alpha}{\lambda^2}. \quad (23.54)$$

Третий начальный момент $m_3 = \frac{\alpha(\alpha+1)(\alpha+2)}{\lambda^3}$. Поэтому третий центральный момент $\mu_3 = 2 \frac{\alpha}{\lambda^3}$, а коэффициент асимметрии $Sk = \frac{2}{\sqrt{\alpha}} > 0$. Далее, четвертый начальный момент $m_4 = \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{\lambda^4}$. Поэтому четвертый центральный момент $\mu_4 = \frac{\alpha(\alpha+2)}{\lambda^4}$, откуда находим эксцесс $Ex = \frac{\sigma}{\alpha} > 0$.

23.7. Экспоненциальное распределение

Частным видом гамма-распределения, когда $\alpha = 1$, является экспоненциальное (показательное) распределение. В этом случае плотность вероятности

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases} \quad (23.55)$$

а функция распределения

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases} \quad (23.56)$$

Графики этих функций распределения приведены на рис. 23.7.

Характеристическая функция экспоненциального распределения в соответствии с (23.52) имеет вид

$$E(u) = \frac{\lambda}{\lambda - iu}. \quad (23.57)$$

Используя равенство (23.53), получаем, что в данном случае начальный момент k -го порядка m_k оказывается равным

$$m_k = \frac{k!}{\lambda^k}, \quad (k = \overline{1, \infty}). \quad (23.58)$$

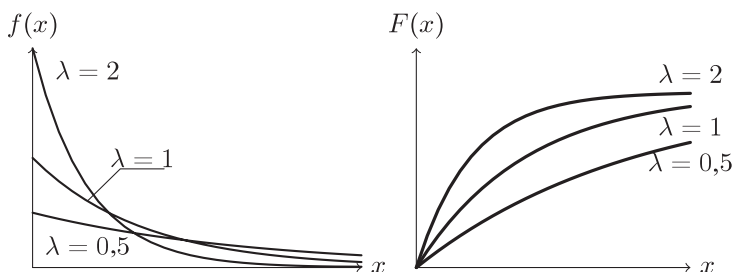


Рис. 23.7

Плотность вероятности и функция распределения показательного закона при различных значениях параметра $\lambda = 0,5, 1, 2$

Математическое ожидание и дисперсия случайной величины следующие:

$$\bar{x} = \frac{1}{\lambda}; \quad D(X) = \frac{1}{\lambda^2}. \quad (23.59)$$

Интересно заметить, что коэффициент вариации $\gamma = \frac{\sigma}{\bar{x}} \equiv 1 = \text{const}(\lambda)$.

Третий центральный момент $\mu_3 = \frac{2}{\lambda^3}$, четвертый центральный момент $\mu_4 = \frac{9}{\lambda^4}$. Коэффициент асимметрии $Sk = 2$, а эксцесс $Ex = 6$.

23.8. Распределение Вейбулла

При аппроксимации законов распределения, получаемых в результате обработки экспериментальных данных, иногда используется закон распределения Вейбулла¹², плотность вероятности которого имеет вид

$$f(x) = \begin{cases} \lambda \alpha x^{\alpha-1} e^{-\lambda x^\alpha}, & x \geq 0; \\ 0, & x < 0, \end{cases} \quad (23.60)$$

¹² Вейбулл Эрнст Хьяльмар Вэлоди (1887–1979) — шведский инженер и математик.

где λ и α — заданные положительные постоянные. Если $\alpha = 1$, то данное распределение совпадает с экспоненциальным распределением, а при $\alpha = 2$ — с распределением Рэлея.

Функция распределения в данном случае определяется формулой

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x^\alpha}, & x \geq 0. \end{cases} \quad (23.61)$$

Если $\alpha < 1$, то график функции $f(x)$ с уменьшением x монотонно приближается к оси ординат. При $\alpha > 1$ эта кривая имеет максимум в точке $x = \left(\frac{\alpha - 1}{\lambda \alpha}\right)^{1/\alpha}$. Указанное значение

аргумента x является модой распределения Вейбулла. Графики функций $f(x)$ и $F(x)$ при $\lambda = 1$ и $\alpha = 0,5, 0,75, 1, 1,5$ и 2 приведены на рис. 23.8.

Характеристическая функция распределения Вейбулла имеет вид

$$E(u) = \lambda \alpha \int_0^{\infty} e^{iux - \lambda x^\alpha} x^{\alpha-1} dx,$$

или

$$E(u) = 1 + iu \int_0^{\infty} e^{iux - \lambda x^\alpha} dx. \quad (23.62)$$

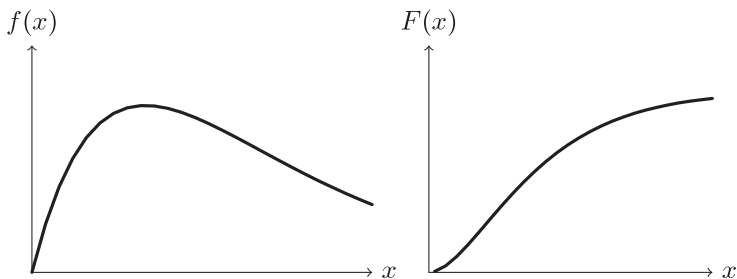


Рис. 23.8

Плотность вероятности и функция распределения закона Вейбулла при различных значениях параметра $\alpha = 0,5, 0,75, 1,5$ и 2

Последний интеграл может быть выражен через табличные функции только при некоторых частных значениях α . Вероятность попадания случайной величины α в интервал (x_1, x_2) , где $x_1 \geq 0$, в данном случае определяется формулой

$$\mathbf{P}\{x_1 \leq X \leq x_2\} = e^{-\lambda x_1^\alpha} - e^{-\lambda x_2^\alpha}. \quad (23.63)$$

Начальный момент k -го порядка m_k будет определяться равенством

$$m_k = \lambda \alpha \int_0^\infty e^{-\lambda x^\alpha} x^{k+\alpha-1} dx.$$

Положим $\lambda x^\alpha = t$, $x = \left(\frac{t}{\lambda}\right)^{1/\alpha}$, тогда $\lambda \alpha x^{\alpha-1} dx = dt$, поэтому

$$m_k = \lambda^{-k/\alpha} \int_0^\infty e^{-t} t^{k/\alpha} dt.$$

Воспользовавшись выражением для гамма-функции, получаем

$$m_k = \lambda^{-k/\alpha} \Gamma\left(\frac{k}{\alpha} + 1\right), \quad (k = \overline{1, \infty}). \quad (23.64)$$

В частности,

$$\begin{aligned} \bar{x} &= \lambda^{-1/\alpha} \Gamma\left(1 + \frac{1}{\alpha}\right); \\ D(X) &= \lambda^{-2/\alpha} \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right) \right]. \end{aligned} \quad (23.65)$$

23.9. Усеченное нормальное распределение

В некоторых случаях принимают во внимание только такие значения нормальной случайной величины, которые не выходят за пределы заданного интервала (a, b) . Получающаяся таким образом случайная величина подчиняется *усеченному нормальному закону распределения*, плотность вероятности для которого определяется формулой

$$f(x) = \begin{cases} \frac{1}{\gamma\alpha\sqrt{2\pi}} \cdot \exp\left\{-\frac{(x-\beta)^2}{2\alpha^2}\right\}, & a \leq x \leq b, \\ 0, & x < a, \quad x > b, \end{cases} \quad (23.66)$$

где α и β — заданные постоянные, причем $\alpha > 0$. Постоянная γ выражается через известные параметры a , b , α и β , так как интеграл от f по промежутку $[a, b]$ должен равняться единице. Из этого условия следует, что

$$\begin{aligned} \gamma &= \frac{1}{\alpha\sqrt{2\pi}} \int_a^b \exp\left\{-\frac{(x-\beta)^2}{2\alpha^2}\right\} dx = \\ &= \frac{1}{2} \cdot \left[\Phi\left(\frac{b-\beta}{\alpha}\right) - \Phi\left(\frac{a-\beta}{\alpha}\right) \right]. \end{aligned} \quad (23.67)$$

Функция распределения $F(x)$ теперь имеет вид

$$f(x) = \begin{cases} 0, & x < a; \\ \frac{1}{2\gamma} \cdot \left[\Phi\left(\frac{x-\beta}{\alpha}\right) - \Phi\left(\frac{a-\beta}{\alpha}\right) \right], & a \leq x \leq b, \\ 1, & x > b. \end{cases} \quad (23.68)$$

Если, например, принять $a = \beta$, а $b = \infty$, то $\gamma = \frac{1}{2}$, при этом

$F(x) = \Phi\left(\frac{x-\beta}{\alpha}\right)$, когда $x \geq a$. Очевидно, график $f(x)$ будет иметь такой же вид, как правая половина соответствующей

кривой для нормального закона распределения, но с удвоенной ординатой.

Характеристическая функция усеченного нормального распределения будет записываться в интегральной форме так:

$$E(u) = \frac{1}{\gamma\alpha\sqrt{2\pi}} \int_a^b e^{iux - (x-\beta)^2/2\alpha^2} dx. \quad (23.69)$$

Для вычисления интеграла сделаем подстановку

$$v = \frac{\alpha^2 u + i(x - \beta)}{\alpha\sqrt{2}},$$

с помощью которой интеграл (23.69) преобразуется к виду

$$E(u) = -\frac{i}{\gamma\alpha\sqrt{\pi}} e^{iu\beta - (\alpha \cdot u)^2/2\alpha^2} \int_{z_1}^{z_2} e^{v^2} dv,$$

где введены обозначения

$$z_1 = \frac{\alpha^2 u + i(a - \beta)}{\alpha\sqrt{2}}, \quad z_2 = \frac{\alpha^2 u + i(b - \beta)}{\alpha\sqrt{2}}.$$

Если теперь воспользоваться интегралом вероятности от комплексного аргумента (23.32), то характеристическую функцию $E(u)$ при конечных a и b можно выразить через данный интеграл следующим образом:

$$E(u) = \frac{1}{2\gamma} e^{i\beta u - (\alpha u)^2/2} \left[e^{z_1^2} W(z_1) - e^{z_2^2} W(z_2) \right]. \quad (23.70)$$

Ниже приводится таблица, содержащая основные характеристики некоторых наиболее часто встречающихся законов распределения, разобранных в настоящей главе.

Примечание:

1) $W(z) = e^{1z^2} \cdot \left(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2} dt \right)$ — интеграл вероятности от комплексного аргумента.

2) $\gamma(\alpha, x) = \int_0^x e^{-t} t^{\alpha-1} dt$ — неполная гамма-функция.

Таблица 23.1
Основные характеристики непрерывных законов
распределения

Характеристика закона	Нормальный закон	Равномерный закон
Область значений	$(-\infty; \infty)$	$(\bar{x} - l; \bar{x} + l)$
Плотность вероятности	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma^2}\right\}$	$\frac{1}{2l}$
Функция распределения	$\frac{1}{2} \left[1 + \Phi\left(\frac{x - \bar{x}}{\sigma}\right) \right]$	$\frac{x - \bar{x} + l}{2l}$
Дисперсия	σ^2	$1/3 l^2$
Характеристическая функция	$\exp\left\{-\frac{\sigma^2 u^2}{2} + i\bar{x}u\right\}$	$e^{i\bar{x}u} \frac{\sin(lu)}{(lu)}$
Математическое ожидание	$\bar{x}$	$\bar{x}$
Моменты m_k ; μ_k	$\mu_{2j+1} = 0,$ $\mu_{2j} = \sigma^{2j} (2j - 1)!!$	$\mu_k = 1/(k + 1) l^k$
Асимметрия	0	0
Экссесс	0	-1,2
Семиинварианты	$\varkappa_1 = \bar{x}; \varkappa_2 = \sigma^2;$ $\varkappa_j = 0, j > 2$	$\varkappa_1 = \bar{x}; \varkappa_{2j+1} = 0$ $\varkappa_{2j} = \frac{(2l)^{2j}}{2j} B_{2j}; j \geq 1$
Срединное отклонение	$\rho\sqrt{2}\sigma \approx 0,674\sigma$	0,5
Мода	$\bar{x}$	—
Медиана	$\bar{x}$	$\bar{x}$

Продолжение табл. 23.1

Характеристика закона	Закон арксинуса	Закон Рэлея
Область значений	$(-a; a)$	$(0, \infty)$
Плотность вероятности	$\frac{1}{\pi\sqrt{a^2 - x^2}}$	$\frac{x}{a^2} e^{-x^2/2a^2}$
Функция распределения	$\frac{1}{2} + \frac{1}{\pi} \arcsin\left(\frac{x}{a}\right)$	$1 - \exp\left(-\frac{x^2}{2a^2}\right)$
Дисперсия	$a^2/2$	$2a^2 (1 - \pi/4)$
Характеристическая функция	$J_0(au)$	$1 + iau \sqrt{\frac{\pi}{2}} \cdot W\left(\frac{au}{\sqrt{2}}\right)$
Математическое ожидание	0	$a\sqrt{\frac{\pi}{2}}$
Моменты m_k ; μ_k	$\mu_{2j+1} = 0,$ $\mu_{2j} = \frac{(2j)!}{2^{2j}(j!)^2} a^{2j}$	$m_k = (a\sqrt{2})^k \Gamma\left(\frac{k+2}{2}\right)$
Асимметрия	0	$\frac{(\pi-3)\sqrt{\pi}}{4(1-\pi/4)^{3/2}} \approx 0,631$
Эксцесс	-1,5	$\frac{1,5\pi(1-0,25\pi)-1}{(1-0,25\pi)^2} \approx$ $\approx 0,245$
Семиинварианты	$\kappa_{2j+1} = 0$	
Срединное отклонение	$\frac{\sqrt{2}}{2}a$	$(\sqrt{\ln 2} - \sqrt{\ln(2/\sqrt{3})})a \approx$ $\approx 0,453a$
Мода	0	a
Медиана	0	$\sqrt{2 \ln 2} a$

Продолжение табл. 23.1

Характеристика закона	Закон Максвелла	Закон Коши
Область значений	$(0; \infty)$	$(-\infty; \infty)$
Плотность вероятности	$\frac{2x^2}{a^3\sqrt{2\pi}} e^{-\frac{x^2}{2a^2}}$	$\frac{a}{\pi(a^2 + x^2)}$
Функция распределения	$\Phi\left(\frac{x}{a}\right) - \frac{2x}{a\sqrt{2\pi}} e^{-\frac{x^2}{2a^2}}$	$\frac{1}{2} + \frac{1}{\pi} \cdot \operatorname{arctg} \frac{x}{a}$
Характеристическая функция	$(1 - a^2 u^2)W(au/\sqrt{2}) +$ $+ iau\sqrt{2/\pi}$	$e^{-a u }$
Математическое ожидание	$2\sqrt{2/\pi}a$	—
Дисперсия	$\frac{3\pi - 8}{\pi} \cdot a^2$	—
Моменты m_k ; μ_k	$m_k = \frac{2^{k/2+1}}{\sqrt{\pi}} \times$ $\times \Gamma\left(\frac{k+3}{2}\right) \cdot a^k$	—
Асимметрия	$\frac{2\sqrt{2}(16 - 5\pi)}{(3\pi - 8)^{3/2}} \approx 0,486$	—
Эксцесс	$\frac{160\pi - 12\pi^2 - 384}{(3\pi - 8)^2} \approx$ $\approx 0,108$	—
Семиинварианты		—
Срединное отклонение		$\frac{\pi}{4}$
Мода	$a\sqrt{2}$	0
Медиана	$\approx 1,538a$	0

Продолжение табл. 23.1

Характеристика закона	Показательный закон (Лапласа)	Гамма-закон распределения
Область значений	$(-\infty; \infty)$	$(0, \infty)$
Плотность вероятности	$\frac{1}{2a} e^{-\frac{ x }{a}}$	$\frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}$ $\alpha > 0, \lambda > 0$
Функция распределения	$\frac{1}{2} \cdot e^{\frac{x}{a}}, x \leq 0;$ $1 - \frac{1}{2} \cdot e^{-\frac{x}{a}}, x \geq 0$	$\frac{\gamma(\alpha, \lambda x)}{\Gamma(\alpha)}$
Характеристическая функция	$\frac{1}{a^2 u^2 + 1}$	$\left(1 - \frac{i u}{\lambda}\right)^{-\alpha}$
Математическое ожидание	0	α/λ
Дисперсия	$2a^2$	$\frac{\alpha}{\lambda^2}$
Моменты m_k ; μ_k	$\mu_{2j+1} = 0,$ $\mu_{2j} = (2j)! a^{2j}$	$m_k = \frac{\Gamma(\alpha + k)}{\lambda^k \Gamma(\alpha)}$
Асимметрия	0	$\frac{2}{\sqrt{\alpha}}$
Экссесс	3	$\frac{6}{\alpha}$
Семиинварианты	$\varkappa_{2j+1} = 0,$ $\varkappa_{2j} = \frac{(2j)!}{j} a^{2j}$	$\varkappa_k = \frac{\alpha(k-1)!}{\lambda^k}$
Срединное отклонение	$a \ln 2$	
Мода	—	$\frac{\alpha - 1}{\lambda}, \alpha > 1$
Медиана	0	

Продолжение табл. 23.1

Характеристика закона	Закон Вейбулла
Область значений	$(0, \infty)$
Плотность вероятности	$\lambda \alpha x^{\alpha-1} e^{-\lambda x^\alpha} \quad \alpha > 0, \lambda > 0$
Функция распределения	$1 - e^{-\lambda x^\alpha}$
Характеристическая функция	$1 + iu \int_0^\infty e^{iux - \lambda x^\alpha} dx$
Математическое ожидание	$\lambda^{-1/\alpha} \Gamma(1 + 1/\alpha)$
Дисперсия	$\lambda^{-2/\alpha} \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right) \right]$
Моменты m_k ; μ_k	$m_k = \lambda^{-k/\alpha} \Gamma\left(\frac{k\alpha}{\alpha}\right)$
Асимметрия	$\frac{\mu_3}{\mu_2^{3/2}}$
Экссесс	$\frac{\mu_4}{\mu_2^2} - 3$
Семиинварианты	
Срединное отклонение	$\frac{2^{1/\alpha-1}}{\lambda^{1/\alpha}} \left((\ln 2)^{1/\alpha} - (\ln(2/\sqrt{3}))^{1/\alpha} \right)$
Мода	$\left(\frac{\alpha-1}{\lambda\alpha} \right)^{1/\alpha}, \quad \alpha > 1$
Медиана	$\left(\frac{1}{\lambda} \cdot \ln 2 \right)^{1/\alpha}$

§ 24. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ И ФОРМУЛА БАЙЕСА В СХЕМЕ СЛУЧАЙНЫХ ВЕЛИЧИН

Полученная в схеме случайных событий формула полной вероятности (см. § 16)

$$\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(H_k) \mathbf{P}(A|H_k) \quad (24.1)$$

может быть распространена и на схему случайных величин, если под гипотезой H_k понимать получение дискретной случайной величиной X возможного значения x_k , ($k = \overline{1, n}$). В этом случае в (24.1) нужно подставить

$$\mathbf{P}(H_k) = \mathbf{P}\{X = x_k\} = p_k; \quad \mathbf{P}(A|H_k) = \mathbf{P}(A|x_k),$$

поэтому указанную формулу можно переписать в виде

$$\mathbf{P}(A) = \sum_{k=1}^n p_k \mathbf{P}(A|x_k), \quad (24.2)$$

где для краткости введено обозначение

$$\mathbf{P}(A|x_k) = \mathbf{P}(A|X = x_k).$$

Аналогичное преобразование можно произвести и с формулой Байеса, полученной в схеме случайных событий на основе формулы полной вероятности

$$\mathbf{P}(H_k|A) = \frac{\mathbf{P}(H_k) \mathbf{P}(A|H_k)}{\sum_{j=1}^n \mathbf{P}(H_j) \mathbf{P}(A|H_j)}. \quad (24.3)$$

В схеме дискретных случайных величин

$$\mathbf{P}(H_k|A) = \mathbf{P}(X = x_k|A) = \mathbf{P}(x_k|A),$$

а поэтому формула (24.3) принимает вид

$$\mathbf{P}(x_k|A) = \frac{p_k \mathbf{P}(A|x_k)}{\sum_{j=1}^n p_j \mathbf{P}(A|x_j)}, \quad (k = \overline{1, n}). \quad (24.4)$$

Таким образом, переход в формуле полной вероятности и в формуле Байеса от схемы случайных событий к схеме *дискретных* случайных величин не встречает никаких осложнений.

Произведем обобщение формул (24.1) и (24.3) на тот случай, когда вероятность появления события A определяется полученным значением x непрерывной случайной величины X . Разобьем всю область возможных значений случайной величины X на n малых интервалов длиной Δx . Под гипотезой H_k будем понимать попадание случайной точки в k -й интервал, то есть выполнение условия $x_k \leq X < x_k + \Delta x$ ($k = \overline{1, n}$), где $F(x)$ — функция распределения случайной величины X . При достаточно малом Δx данную вероятность выразим через плотность вероятности $f(x)$, приняв $\mathbf{P}(H_k) \approx f(x_k^*) \Delta x$, где x_k^* — некоторое промежуточное значение из интервала $(x_k; x_k + \Delta x)$.

Обозначим через $\mathbf{P}(A|x)$ вероятность появления события A при условии, что случайная величина X приняла в опыте значение x , и будем приближенно считать, что для любого значения x , попадающего в k -й интервал, благодаря малости последнего условия вероятность $\mathbf{P}(A|x)$ события A имеет одинаковое значение, равное $\mathbf{P}(A|x_k^*)$. Тогда формулу (24.1) можно записать в виде

$$\mathbf{P}(A) \approx \sum_{k=1}^n \mathbf{P}(A|x_k^*) f(x_k^*) \Delta x.$$

Переходя в последнем выражении к пределу при $\Delta x \rightarrow 0$, получим следующее обобщение формулы полной вероятности:

$$\mathbf{P}(A) = \int_{-\infty}^{\infty} \mathbf{P}(A|x) f(x) dx. \quad (24.5)$$

Аналогичные рассуждения применимы и для обобщения формулы Байеса (24.3). Заменяя в ней $\mathbf{P}(H_k|A)$, $\mathbf{P}(H_k)$ и $\mathbf{P}(A|H_k)$, соответственно, на $f(x_k^*|A)\Delta x$, $f(x_k^*)\Delta x$ и $\mathbf{P}(A|x_k^*)$,

с учетом (24.5) получим

$$f(x_k^*|A) \triangle x = \frac{\mathbf{P}(A|x_k^*)f(x_k^*) \triangle x}{\int_{-\infty}^{\infty} \mathbf{P}(A|x)f(x)dx}, \quad (k = \overline{1, n}).$$

После сокращения на $\triangle x$ и предельного перехода при $\triangle x \rightarrow 0$ последнее выражение принимает вид

$$f(x|A) = \frac{f(x)\mathbf{P}(A|x)}{\int_{-\infty}^{\infty} \mathbf{P}(A|x)f(x)dx}. \quad (24.6)$$

Формула (24.6) дает выражение для плотности вероятности случайной величины X после наблюдения события A через плотность вероятности $f(x)$ этой случайной величины и вероятность $\mathbf{P}(A|x)$ появления события A при заданном значении x случайной величины X , имевшей место до испытания.

Рассмотрим несколько примеров на применение полученных формул.

Пример 24.1. Вероятность попадания в цель при каждом из независимых выстрелов равна p . В ходе стрельбы стрелок корректирует прицеливание, поэтому вероятность поражения цели при k попаданиях в нее определяется формулой

$$\mathbf{P}(k) = 1 - \left(1 - \frac{1}{\omega}\right)^k, \text{ где } \omega - \text{ заданная постоянная } (\omega > 1).$$

В результате стрельбы цель поражена. Определить вероятность того, что было k попаданий ($k = \overline{0, n}$).

Решение. Случайное событие A — поражение цели, случайная величина X — число попаданий. Тогда

$$\mathbf{P}(A|k) = 1 - \left(1 - \frac{1}{\omega}\right)^k,$$

а X распределена по биномиальному закону

$$p_k = \mathbf{P}\{X = k\} = C_n^k p^k (1-p)^{n-k}.$$

Используя формулу полной вероятности, находим

$$\mathbf{P}(A) = \sum_{k=0}^n p_k \mathbf{P}(A|k) = 1 - \left(1 - \frac{p}{\omega}\right)^n.$$

Искомая вероятность

$$\mathbf{P}(k|A) = \frac{p_k \mathbf{P}(A|k)}{\mathbf{P}(A)} = C_n^k p^k (1-p)^{n-k} \cdot \frac{1 - \left(1 - \frac{1}{\omega}\right)^k}{1 - \left(1 - \frac{p}{\omega}\right)^n}.$$

Пример 24.2. Событие A произошло m раз при n независимых испытаниях, причем вероятность того, что оно произойдет при каждом отдельно взятом испытании одинакова и равна случайной величине P , равномерно распределенной в интервале $(0, 1)$. Определить вероятность того, что событие A произойдет l раз при следующих s испытаниях.

Решение. По условию случайная величина P характеризуется плотностью вероятности

$$f(p) = \begin{cases} 1, & 0 \leq p \leq 1, \\ 0, & p < 0, \quad p > 1. \end{cases}$$

При заданной вероятности $P = p$ вероятность того, что событие A произойдет m раз при n испытаниях, определяется биномиальным законом

$$\mathbf{P}_m^{(n)}(A|p) = C_n^m p^m (1-p)^{n-m}.$$

Вероятность того, что событие A произойдет m раз, выразится в виде

$$\mathbf{P}_m^{(n)}(A) = \int_0^1 f(p) \mathbf{P}_m^{(n)}(A|p) dp = C_n^m B(m+1, n-m+1),$$

где $B(\alpha, \beta)$ обозначает бета-функцию. Хорошо известно, что бета-функция выражается через гамма-функцию

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

а гамма-функция при целом значении аргумента сводится к факториалу $\Gamma(m) = (m - 1)!$, поэтому

$$\mathbf{P}_m^{(n)}(A) = \frac{n! m! (n - m)!}{m! (n - m)! (n + 1)!} = \frac{1}{n + 1}.$$

Согласно обобщенной формуле Байеса, плотность вероятности случайной величины P после появления m раз события A при n испытаниях будет такой:

$$\begin{aligned} f(p|A) &= \frac{f(p) \cdot \mathbf{P}_m^{(n)}(A|p)}{\mathbf{P}_n^{(m)}(A)} = \\ &= (n + 1) \cdot C_n^m \cdot p^m \cdot (1 - p)^{n-m}, \quad (0 \leq p \leq 1). \end{aligned}$$

При фиксированном p вероятность появления события A ровно l раз при s независимых испытаниях определяется формулой биномиального распределения, поэтому

$$\mathbf{P}_l^{(s)}(A|p) = C_s^l \cdot p^l \cdot (1 - p)^{s-l},$$

тогда искомая вероятность

$$\begin{aligned} \mathbf{P}_s^{(l)} &= \int_0^1 f(p|A) \cdot \mathbf{P}_s^{(l)}(A|p) dp = \\ &= (n + 1) C_n^m \cdot C_s^l \int_0^1 p^{m+l} (1 - p)^{n-m+s-l} dp = \\ &= \frac{(n + 1) \cdot n! \cdot s! \cdot (m + l)! \cdot (n - m + s - l)!}{m! \cdot (n - m)! \cdot (s - l)! \cdot (n + s + 1)!} = \\ &= \frac{C_{m+l}^l \cdot C_{n+s-m-l}^{s-l}}{C_{n+s+1}^s}. \end{aligned}$$

Пример 24.3. Вероятность появления события A зависит от значения случайной величины X и выражается в виде

$$\mathbf{P}(A|x) = \begin{cases} 1 - e^{-\alpha x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Найти полную вероятность события A , считая, что случайная величина X распределена по показательному закону

$$f(x) = \lambda e^{-\lambda x}.$$

Решение. Полная вероятность события A дается формулой (24.5), которая после подстановки данных из условия примет вид

$$\mathbf{P}(A) = \int_0^{\infty} \lambda e^{-\lambda x} (1 - e^{-\alpha x}) dx.$$

Разбивая этот интеграл на два интеграла, получим

$$\mathbf{P}(A) = 1 - \frac{\lambda}{\lambda + \alpha} = \frac{\alpha}{\lambda + \alpha}.$$

Перейдем в этом выражении от параметра λ к математическому ожиданию случайной величины X

$$\bar{x} = \int_0^{\infty} x \cdot f(x) dx = \frac{1}{\lambda}.$$

Тогда получим

$$\mathbf{P}(A) = \frac{\alpha \bar{x}}{1 + \alpha \bar{x}}.$$

Легко видеть, что свойства события A кардинально меняются при увеличении $\bar{x}$ от нуля до бесконечности. При $\bar{x} \rightarrow 0$ событие A превращается в невозможное, а при неограниченном увеличении $\bar{x}$ оно становится достоверным.

Пример 24.4. Отклонение размера детали от середины поля допуска, имеющего полуширину h , может возникнуть при

выполнении любой из двух технологических операций по обработке детали. Ошибки X и Y изготовления детали на каждой из этих операций можно считать заданными нормальными случайными величинами, причем систематическая ошибка для каждой из них отсутствует, так что

$$f_x(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma_x} \cdot e^{-x^2/(2\sigma_x^2)}, \quad f_y(y) = \frac{1}{\sqrt{2\pi} \cdot \sigma_y} \cdot e^{-y^2/(2\sigma_y^2)},$$

где σ_x и σ_y характеризуют точность обработки на соответствующей операции. Определить плотность вероятности X для годных деталей, для которых полная ошибка $X + Y$ удовлетворяет условию $|X + Y| \leq h$, считая, что распределение Y не зависит от того, какое значение приняла случайная величина X .

Решение. Пусть событие A означает, что изготовленная деталь оказалась годной

$$A = \{|X + Y| \leq h\}.$$

Условная вероятность $\mathbf{P}(A|x)$ получения годной детали при условии, что $X = x$ равняется вероятности события $|Y + x| \leq h$, то есть события $-x - h \leq Y \leq -x + h$, значит,

$$\begin{aligned} \mathbf{P}(A|x) &= \int_{-x-h}^{-x+h} \frac{1}{\sqrt{2\pi}\sigma_y} \cdot e^{-y^2/(2\sigma_y^2)} \cdot dy = \\ &= \frac{1}{2} \cdot \left[\Phi\left(\frac{x+h}{\sigma_y}\right) - \Phi\left(\frac{x-h}{\sigma_y}\right) \right], \end{aligned}$$

где $\Phi(x)$ обозначает интеграл вероятностей

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

По формуле (24.6) получаем

$$\begin{aligned} f(x|A) &= \frac{f(x) \cdot \mathbf{P}(A|x)}{\int_{-\infty}^{\infty} f(x) \cdot \mathbf{P}(A|x) dx} = \\ &= \frac{\frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \left[\Phi\left(\frac{x+h}{\sigma_y}\right) - \Phi\left(\frac{x-h}{\sigma_y}\right) \right]}{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \left[\Phi\left(\frac{x+h}{\sigma_y}\right) - \Phi\left(\frac{x-h}{\sigma_y}\right) \right] dx}. \end{aligned}$$

Поскольку функция $\Phi(x)$ является нечетной

$$\Phi(-x) = -\Phi(x),$$

то нетрудно показать, что

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \Phi\left(\frac{x-h}{\sigma_y}\right) dx &= \\ &= - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \Phi\left(\frac{x+h}{\sigma_y}\right) dx. \end{aligned}$$

Рассмотрим интеграл

$$I(h) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \Phi\left(\frac{x+h}{\sigma_y}\right) dx.$$

Ясно, что поскольку $\Phi(x)$ нечетна, то $I(0) = 0$. Вычислим производную

$$I'(h) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} \cdot \frac{2}{\sqrt{2\pi}\sigma_y} \cdot e^{-(x+h)^2/(2\sigma_y^2)} dx.$$

Если выделить в показателе экспоненты

$$e^{-x^2/(2\sigma_x^2) - (x+h)^2/(2\sigma_y^2)}$$

полный квадрат относительно x , а затем проинтегрировать полученный интеграл, используя табличный интеграл Пуассона

$$\int_{-\infty}^{\infty} e^{-x^2/(2a^2)} dx = \sqrt{2\pi}a,$$

то получим

$$I'(h) = \frac{2}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \cdot e^{-h^2/2(\sigma_x^2 + \sigma_y^2)}.$$

Используя начальное условие $I(0) = 0$, находим

$$I(h) = \int_0^h \frac{2}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \cdot e^{-t^2/2(\sigma_x^2 + \sigma_y^2)} dt.$$

Прделав в этом интеграле замену переменной интегрирования

$$s = \frac{t}{\sqrt{(\sigma_x^2 + \sigma_y^2)}},$$

находим

$$I(h) = \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\frac{h}{\sqrt{\sigma_x^2 + \sigma_y^2}}} e^{-s^2/2} ds = \Phi\left(\frac{h}{\sqrt{\sigma_x^2 + \sigma_y^2}}\right).$$

Таким образом, окончательно

$$f(x|A) = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}\sigma_x} \cdot e^{-x^2/(2\sigma_x^2)} \cdot \frac{\left[\Phi\left(\frac{x+h}{\sigma_y}\right) - \Phi\left(\frac{x-h}{\sigma_y}\right) \right]}{\Phi\left(\frac{h}{\sqrt{\sigma_x^2 + \sigma_y^2}}\right)},$$

то есть условный закон уже не будет нормальным.

СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 25. ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТЬ ВЕРОЯТНОСТИ СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

Совокупность нескольких случайных величин $X_1, X_2, \dots, X_n$ называется *системой случайных величин* и обозначается $(X_1, X_2, \dots, X_n)$.

Определение системы случайных величин можно получить из основных аксиом теории вероятностей, как и понятие одной случайной величины (см. гл. 3). Для этого достаточно рассмотреть вероятностное пространство $(\Omega, \mathfrak{A}, P)$ и сопоставить каждому из элементов множества $\mathfrak{A}$ вектор $\vec{X}$ с компонентами $X_1, X_2, \dots, X_n$, значения которых определяют положение случайной точки M в n -мерном евклидовом пространстве. Поскольку множество $\mathfrak{A}$ имеет вероятностную меру P , множество случайных точек, характеризуемых значениями вектора $\vec{X}$, также имеет вероятностную меру, а совокупность компонент этого вектора можно рассматривать как систему случайных величин. Каждая из этих величин X_j в отдельности полностью характеризуется своей функцией распределения $F(x_j)$ или своей плотностью вероятности $f(x_j)$, как это было подробно рассмотрено в главе 4, однако при изучении системы случайных величин в целом этих характеристик уже недостаточно, так как они не содержат никакой информации о взаимной зависимости случайных величин, образующих рассматриваемую систему.

Исчерпывающей характеристикой системы случайных величин $X_1, X_2, \dots, X_n$ является функция распределения системы $F(x_1, x_2, \dots, x_n)$, определяемая аналогично функции распределения одной отдельно взятой случайной величины следующим образом:

$$F(x_1, x_2, \dots, x_n) = \mathbf{P}\{X_1 < x_1, X_2 < x_2, \dots, X_n < x_n\}. \quad (25.1)$$

Функция распределения $F(x_1, x_2, \dots, x_n)$ позволяет определить вероятность попадания случайной точки с координатами $(X_1, X_2, \dots, X_n)$ в любую область n -мерного пространства. Действительно, рассмотрим для наглядности случай $n = 2$. Тогда значения случайных величин X_1, X_2 можно рассматривать как декартовы координаты случайной точки M на плоскости, а функция распределения $F(x_1, x_2)$ будет давать вероятность попадания точки M внутрь прямого угла, заштрихованного на рис. 25.1. Рассмотрим элементарный прямоугольник dS , координаты вершин которого имеют значения: (x_1, x_2) , $(x_1 + dx_1, x_2)$, $(x_1 + dx_1, x_2 + dx_2)$, $(x_1, x_2 + dx_2)$. Обозначим вероятность попадания случайной точки внутрь этого прямоугольника через $dP(x_1, x_2)$, то есть положим

$$dP(x_1, x_2) = \mathbf{P}\{x_1 \leq X_1 < x_1 + dx_1; \\ x_2 \leq X_2 < x_2 + dx_2\}. \quad (25.2)$$

Используя определение функции распределения и применяя формулу сложения вероятностей, будем иметь

$$dP(x_1, x_2) = [F(x_1 + dx_1, x_2 + dx_2) - \\ - F(x_1, x_2 + dx_2)] - [F(x_1 + dx_1, x_2) - F(x_1, x_2)], \quad (25.3)$$

где первая квадратная скобка дает сумму вероятности попадания случайной точки в рассматриваемый элемент плоскости и вероятности попадания в бесконечную вертикальную полосу, отмеченную на рис. 25.1 горизонтальной штриховкой, а вторая квадратная скобка — вероятность попадания в указанную полосу. Так как произвольную область S на плоскости можно с любой точностью аппроксимировать элементарными прямоугольниками, то вероятность попадания случайной

точки в данную область можно определить, просуммировав элементарные вероятности $dP(x_1, x_2)$ по всем элементам площади dS .

Таким образом, функция распределения $F(x_1, x_2)$ позволяет определить вероятность попадания случайной точки $M(X_1, X_2)$ в любую область, а значит полностью определяет свойства системы случайных величин (X_1, X_2) . Так как рассуждения, приведенные выше для двумерного случая, обобщаются очевидным образом на пространство любого числа измерений, то отсюда можно заключить, что функция распределения действительно является исчерпывающей характеристикой любой системы случайных величин. Далее установим основные свойства этой функции.

Из определения (25.1) следует, что подобно случаю $n = 1$ справедливо двойное неравенство

$$0 \leq F(x_1, x_2, \dots, x_n) \leq 1. \quad (25.4)$$

Если хотя бы один аргумент функции распределения стремится к $-\infty$, то хотя бы одно неравенство, стоящее в (25.1) под знаком вероятности, становится невозможным событием и, следовательно, функция распределения обращается в нуль. Таким образом,

$$F(x_1, x_2, \dots, x_n)|_{x_j = -\infty} = 0, \quad (j = \overline{1, n}). \quad (25.5)$$

Если r любых аргументов функции распределения ($r < n$) обращаются в $+\infty$, то r неравенств в (25.1) соответствуют достоверным событиям и функция распределения системы случайных величин превращается в функцию распределения оставшейся подсистемы величин. Например, если выбраны

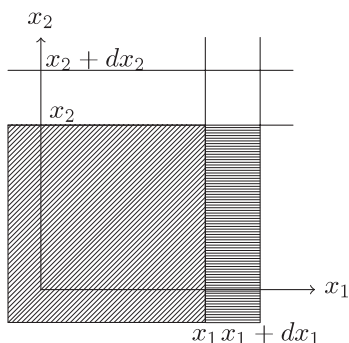


Рис. 25.1

К определению функции распределения системы

первые r аргументов, то

$$F(\infty, \infty, \dots, \infty, x_{r+1}, \dots, x_n) = F(x_{r+1}, x_{r+2}, \dots, x_n), \quad (25.6)$$

где справа стоит функция распределения системы $(n-r)$ случайных величин $(X_{r+1}, X_{r+2}, \dots, X_n)$. Условие (25.6) часто называют *условием согласования Колмогорова*. Оно показывает, что закон распределения подсистемы случайных величин однозначно выражается через закон распределения всей системы. Поэтому часто говорят об *иерархии законов распределения*, в которой законы младших порядков (меньшие n) выражаются через законы старших порядков (большие n).

Из определения функции распределения (25.1) следует, что при любом $\Delta > 0$

$$\begin{aligned} F(x_1, x_2, \dots, x_j + \Delta, \dots, x_n) - F(x_1, \dots, x_n) = \\ = \mathbf{P}\{X_1 < x_1, \dots, X_{j-1} < x_{j-1}, X_j < x_j + \Delta, \dots, X_n < x_n\}. \end{aligned}$$

Так как вероятность, стоящая в правой части равенства, не может быть отрицательной, то отсюда следует, что функция распределения системы — неубывающая функция по каждому из своих аргументов:

$$F(x_1, x_2, \dots, x_j + \Delta, \dots, x_n) \geq F(x_1, x_2, \dots, x_n), \quad \Delta > 0. \quad (25.7)$$

Все перечисленные выше свойства функции распределения имеют место для случайных величин любого типа. Для системы *непрерывных* случайных величин к свойствам (25.4)–(25.7) нужно добавить требование дифференцируемости функции $F(x_1, x_2, \dots, x_n)$ относительно каждого из ее аргументов. В этом случае элементарная вероятность dP попадания случайной точки $M(X_1, X_2, \dots, X_n)$ в элемент $dV = dx_1 dx_2 \dots dx_n$ n -мерного объема («элемент вероятности») может быть представлена в виде произведения многомерной *плотности вероятности* $f(x_1, x_2, \dots, x_n)$ на элемент объема, так, что

$$dP = f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n. \quad (25.8)$$

Плотность вероятности $f(x_1, x_2, \dots, x_n)$ связана с функцией распределения $F(x_1, x_2, \dots, x_n)$ соотношением

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}, \quad (25.9)$$

которое для системы двух случайных величин $n = 2$ сразу следует из формулы (25.3), если рассматривать dx_1 и dx_2 как бесконечно малые величины и учесть дифференцируемость функции F по всем ее аргументам (при $n > 2$ формула (25.9) может быть получена путем аналогичных рассуждений).

Если среди случайных величин, входящих в систему, имеются величины дискретного или смешанного типа, то функция распределения F уже не будет дифференцируема по всем своим аргументам. Однако понятие плотности вероятности по-прежнему можно сохранить, если ввести в рассмотрение не только обычные функции анализа, но и добавить к ним обобщенные функции — в этом случае при выполнении дифференцирования в формуле (25.9) появятся дельта-функции.

Рассмотрим в качестве примера систему двух независимых случайных величин (X, Y) , из которых первая X — непрерывная, а вторая Y — дискретная случайная величина. Предположим, что случайная величина X имеет плотность вероятности $f(x)$, а дискретная компонента характеризуется рядом распределения

$$\mathbf{P}\{Y = y_j\} = p_i, \quad (i = \overline{1, m}), \quad (25.10)$$

где, согласно сделанному предположению, y_i и p_i будем считать независимыми от значения x случайной величины X . В этом случае

$$F(x, y) = \int_{-\infty}^x f_x(x_1) dx_1 \cdot \sum_{(y_i < y)} p_i, \quad (25.11)$$

$$f(x, y) = f_x(x) \cdot \sum_{i=1}^m p_i \delta(y - y_i). \quad (25.12)$$

Аналогичные формулы имеют место и при большем числе измерений, и при наличии зависимости возможных значений

дискретных случайных величин и соответствующих им вероятностей от значений непрерывных величин. Поэтому аналогично тому, как это было ранее принято для одной случайной величины (см. главу 4), будем пользоваться плотностью вероятности для характеристики системы случайных величин любого типа, учитывая при этом, что если в системе имеются дискретные компоненты, то плотность вероятности будет содержать дельта-функции и, следовательно, будет обладать рядом специфических свойств, отличающих ее от плотности вероятности системы непрерывных случайных величин. Далее установим свойства плотности вероятности системы случайных величин, справедливые для величин любого типа.

Интегрируя равенство (25.9) последовательно по каждой из переменных $x_1, x_2, \dots, x_n$ от $-\infty$ до произвольных значений и учитывая при этом свойство (25.5.) функции распределения, получаем

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} f(x_1^*, \dots, x_n^*) dx_1^* \dots dx_n^*, \quad (25.13)$$

где переменные интегрирования отмечены звездочкой, чтобы отличить их от верхних пределов интегрирования.

Полагая в (25.13) $x_j = \infty$ для одного или нескольких значений индекса и учитывая свойство (25.6) функции распределения, получаем выражение функции распределения подсистемы через функцию распределения системы. Например, положив $x_1 = x_2 = \dots = x_r = \infty$, имеем

$$\begin{aligned} F(x_{r+1}, x_{r+2}, \dots, x_n) &= \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_{r+1}} f(x_1^*, \dots, x_n^*) dx_{r+1}^* \dots dx_n^* dx_1^* \dots dx_r^*. \end{aligned} \quad (25.14)$$

Дифференцируя последнее равенство по переменным $x_{r+1}, x_{r+2}, \dots, x_n$ и учитывая определение плотности вероятности

(25.9), получим

$$\begin{aligned} f(x_{r+1}, x_{r+2}, \dots, x_n) &= \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1, \dots, dx_r. \end{aligned} \quad (25.15)$$

Таким образом, можем сформулировать следующее правило: плотность вероятности подсистемы случайных величин равняется интегралу от плотности вероятности системы, взятому в бесконечных пределах по переменным, соответствующим случайным величинам системы, не вошедшим в выбранную подсистему.

Из свойства (25.7) функции распределения системы следует, что плотность вероятности $f(x_1, \dots, x_n)$ — неотрицательная функция, то есть при любых значениях аргументов x_j , ($j = \overline{1, n}$) имеем

$$f(x_1, \dots, x_n) \geq 0. \quad (25.16)$$

Наконец, положив все аргументы функции распределения в формуле (25.13), равными $+\infty$, и учитывая, что в этом случае, на основании ее определения, функция распределения должна обращаться в единицу, получим

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1, \dots, dx_n = 1. \quad (25.17)$$

Иначе говоря, плотность вероятности системы случайных величин нормирована к единице во всей области изменения этих величин.

Рассмотрим теперь плотность вероятности подсистемы случайных величин $X_1, X_2, \dots, X_r$ при условии, что остальные величины системы $X_{r+1}, X_{r+2}, \dots, X_n$ имеют заданные значения $x_{r+1}, x_{r+2}, \dots, x_n$. Такая плотность вероятности носит название *условной* и обозначается $f(x_1, x_2, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_n)$.

Для определения этой плотности вероятности рассмотрим три события: событие A — выполняются неравенства $x_j \leq X_j < x_j + dx_j$, ($j = \overline{1, r}$), событие B — выполняются

неравенства $x_l \leq X_l < x_l + dx_l$, ($l = \overline{(r+1), n}$) и событие C — выполняются неравенства $x_i \leq X_i < x_i + dx_i$, ($i = \overline{1, n}$). Очевидно, что C является произведением событий A и B :

$$C = A \cap B. \quad (25.18)$$

Следовательно, применяя формулу умножения вероятностей, имеем

$$\mathbf{P}(C) = \mathbf{P}(A|B) \cdot \mathbf{P}(B). \quad (25.19)$$

С другой стороны, из определения плотности вероятности системы, условной вероятности подсистемы и ее безусловной плотности, имеем

$$\begin{cases} \mathbf{P}(C) = f(x_1, \dots, x_n) dx_1 \dots dx_n, \\ \mathbf{P}(A|B) = f(x_1, \dots, x_r | x_{r+1}, \dots, x_n) dx_1 \dots dx_r, \\ \mathbf{P}(B) = f(x_{r+1}, x_{r+2}, \dots, x_n) dx_{r+1}, dx_{r+2} \dots dx_n. \end{cases} \quad (25.20)$$

Подставляя (25.20) в (25.19) и сокращая на дифференциалы dx_j , ($j = \overline{1, n}$), получаем

$$f(x_1, \dots, x_n) = f(x_1, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_n) \times \\ \times f(x_{r+1}, x_{r+2}, \dots, x_n), \quad (25.21)$$

то есть условная плотность вероятности определяется равенством

$$f(x_1, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_n) = \frac{f(x_1, \dots, x_n)}{f(x_{r+1}, x_{r+2}, \dots, x_n)}, \quad (25.22)$$

где плотность вероятности, стоящая в знаменателе, определяется формулой (25.15). Разумеется, здесь считается, что знаменатель выражения (25.22) не обращается в нуль.

Если условная плотность вероятности

$$f(x_1, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_n)$$

не зависит от значений $x_{r+1}, x_{r+2}, \dots, x_n$, так что

$$f(x_1, \dots, x_r | x_{r+1}, x_{r+2}, \dots, x_n) = f(x_1, \dots, x_r), \quad (25.23)$$

то говорят, что система случайных величин $(X_1, \dots, X_r)$ *не зависит* от системы $(X_{r+1}, X_{r+2}, \dots, X_n)$. В этом случае в соответствии с формулой (25.21)

$$f(x_1, \dots, x_n) = f(x_1, \dots, x_r) f(x_{r+1}, x_{r+2}, \dots, x_n), \quad (25.24)$$

то есть плотность вероятности системы случайных величин $(X_1, \dots, X_n)$ распадается на произведение плотностей вероятности ее подсистем $(X_1, \dots, X_r)$ и $(X_{r+1}, X_{r+2}, \dots, X_n)$.

Легко видеть, что, если система $(X_1, X_2, \dots, X_r)$ не зависит от системы $(X_{r+1}, X_{r+2}, \dots, X_n)$, то и вторая система не зависит от первой, так что две системы случайных величин могут быть или зависимыми, или взаимно независимыми. Действительно, применив формулу (25.22) для определения условной плотности вероятности второй подсистемы при условии задания величин, входящих в первую подсистему, получим

$$f(x_{r+1}, x_{r+2}, \dots, x_n | x_1, \dots, x_r) = \frac{f(x_1, \dots, x_n)}{f(x_1, x_2, \dots, x_r)}. \quad (25.25)$$

Если предположить, что первая система случайных величин не зависит от второй, то выполняется равенство (25.23), а в числитель (25.25) можно подставить произведение (25.24), что после сокращения на $f(x_1, x_2, \dots, x_r)$ дает

$$f(x_{r+1}, x_{r+2}, \dots, x_n | x_1, \dots, x_r) = f(x_{r+1}, x_{r+2}, \dots, x_n), \quad (25.26)$$

откуда следует, что из независимости первой системы от второй следует и независимость второй системы от первой.

Если все величины системы $X_1, \dots, X_n$ взаимно независимы, то плотность вероятности системы распадается на произведение плотностей вероятностей всех величин, входящих в систему, и имеет место равенство

$$f(x_1, \dots, x_n) = \prod_{j=1}^n f(x_j). \quad (25.27)$$

Справедливо и обратное утверждение: если выполняется равенство (25.27), то случайные величины X_j , ($j = \overline{1, n}$) взаимно независимы. В этом легко убедиться, подсчитав условную

плотность вероятности по формуле (25.22). Подставив (25.27) в (25.13) и представив n -кратный интеграл в виде произведения n однократных интегралов, получим

$$F(x_1, \dots, x_n) = \prod_{j=1}^n F(x_j), \quad (25.28)$$

откуда вытекает, что функция распределения системы взаимно независимых случайных величин равняется произведению функций распределения величин, входящих в систему.

Рассмотрим несколько типичных примеров на системы случайных величин.

Пример 25.1. Дана функция распределения $F(x, y)$ системы двух случайных величин (X, Y) . Определить вероятность P попадания случайной точки $M(X, Y)$ в область, заштрихованную на рис. 25.2.

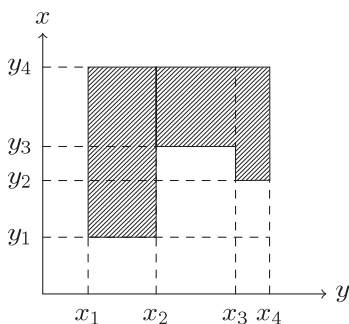


Рис. 25.2
К задаче 25.1

Решение. Разбивая область на прямоугольники, указанные на рисунке, находим

$$\begin{aligned} P = & F(x_4, y_4) - f(x_4, y_3) - F(x_1, y_4) + F(x_1, y_3) + \\ & + F(x_4, y_3) - F(x_4, y_2) - F(x_3, y_3) + F(x_3, y_2) + \\ & + F(x_2, y_3) - F(x_2, y_4) - F(x_1, y_3) + F(x_1, y_1). \end{aligned}$$

После приведения подобных членов, получим

$$P = F(x_4, y_4) + F(x_3, y_2) + \\ + F(x_2, y_3) + F(x_1, y_1) - F(x_4, y_2) - F(x_3, y_3) - \\ - F(x_2, y_3) - F(x_2, y_1) - F(x_1, y_1).$$

Пример 25.2. Система случайных величин (X, Y) имеет плотность вероятности

$$f(x, y) = \begin{cases} ce^{-a^2x^2-b^2y^2} & \text{при } a^2x^2 + b^2y^2 \leq R^2, \\ 0 & \text{при } a^2x^2 + b^2y^2 > R^2, \end{cases}$$

где a, b, R — известные постоянные, а c — нормировочная постоянная. Определить вероятность для $0 \leq r \leq R$:

$$P(r) = \mathbf{P}\{a^2X^2 + b^2Y^2 \leq r^2\}.$$

Решение. Интегрируя элемент вероятности $dP(x, y) = f(x, y)dx dy$ по указанной области и, переходя к полярным координатам, получаем

$$P(r) = c \cdot \iint_{a^2x^2+b^2y^2 \leq r^2} e^{-a^2x^2-b^2y^2} dx dy = \\ = \frac{c}{ab} \cdot \iint_{x^2+y^2 \leq r^2} e^{-x_1^2-y_1^2} dx_1 dy_1 = \\ = \frac{c}{ab} \cdot \int_0^{2\pi} \int_0^r e^{-\rho^2} \rho d\rho d\varphi = \frac{\pi c}{ab} \cdot (1 - e^{-r^2}).$$

При $r = R$ вследствие нормировки плотности вероятности $\mathbf{P}(R) = 1$. Следовательно,

$$c = \frac{ab}{\pi(1 - e^{-R^2})},$$

а тогда

$$\mathbf{P}(r) = \frac{1 - e^{-r^2}}{1 - e^{-R^2}}.$$

Пример 25.3. Система непрерывных случайных величин $(X_1, X_2, \dots, X_n)$ равномерно распределена в n -мерном параллелепипеде $a_j \leq x_j \leq b_j$, $(j = \overline{1, n})$. Найти плотность вероятности $f(x_1, \dots, x_n)$ и функцию распределения $F(x_1, \dots, x_n)$.

Решение. Случайные величины X_j взаимно независимые, причем все они, каждая в отдельности, распределены по равномерному закону. Следовательно,

$$f(x_1, \dots, x_n) = \prod_{j=1}^n f(x_j), \quad F(x_1, \dots, x_n) = \prod_{j=1}^n F(x_j),$$

$$f(x_j) = \begin{cases} \frac{1}{b_j - a_j}, & a_j \leq x_j \leq b_j, \\ 0, & x_j < a_j, x_j > b_j, \end{cases}$$

$$F(x_j) = \begin{cases} 0, & x_j < a_j, \\ \frac{x_j - a_j}{b_j - a_j}, & a_j \leq x_j \leq b_j, \\ 1, & x_j > b_j. \end{cases}$$

Пример 25.4. Координаты (X, Y) случайной точки на плоскости равномерно распределены внутри круга радиуса a , центр которого совпадает с началом координат. Определить $f(x, y)$, $f(x)$, $f(y)$, $f(x|y)$, $f(y|x)$. Являются ли координаты точки независимыми случайными величинами?

Решение. По условию

$$f(x, y) = \begin{cases} c, & x^2 + y^2 \leq a^2, \\ 0, & x^2 + y^2 > a^2, \end{cases}$$

где постоянная c определена из условия нормировки (25.17), которое в данном случае записывается для круга $x^2 + y^2 \leq a^2$. Применяя это условие, имеем

$$c = \frac{1}{\pi a^2}.$$

Далее, используя формулу (25.15), находим

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy = \frac{1}{\pi a} \cdot \int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} dy = \frac{2\sqrt{a^2-x^2}}{\pi a^2}$$

и, совершенно аналогично

$$f(y) = \frac{2\sqrt{a^2-y^2}}{\pi a^2}.$$

Применяя формулу (25.22), будем иметь

$$f(x|y) = \frac{f(x, y)}{f(y)} \cdot \begin{cases} \frac{1}{2\sqrt{a^2-y^2}}, & x^2 \leq a^2 - y^2, \\ 0, & x^2 > a^2 - y^2. \end{cases}$$

Тем же способом легко получим

$$f(y|x) = \begin{cases} \frac{1}{2\sqrt{a^2-x^2}}, & y^2 \leq a^2 - x^2, \\ 0, & y^2 > a^2 - x^2. \end{cases}$$

Поскольку условные законы отличаются от безусловных, то случайные величины X и Y зависят друг от друга.

Пример 25.5. Дана система случайных величин (X, Y) . Определить плотность вероятности случайной величины X после опыта, в результате которого установлено, что $a \leq Y < b$.

Решение. Обозначим событие, наблюдавшееся на опыте, через A . По условию задачи надо определить $f(x|A)$. Назовем попадание случайной величины X в интервал $(x, x + dx)$ событием B , а совместное осуществление событий A и B — событием C . Имеем

$$\mathbf{P}(C) = dx \int_a^b f(x, y) dy,$$

$$\mathbf{P}(A) = \int_a^b f(y)dy = \int_a^b \int_{-\infty}^{\infty} f(x, y)dx dy,$$

$$\mathbf{P}(B) = f(x)dx = dx \int_{-\infty}^{\infty} f(x, y)dy.$$

Подставляя эти выражения в формулу умножения вероятностей $\mathbf{P}(C) = \mathbf{P}(B|A)\mathbf{P}(A)$, после сокращения на dx и очевидных преобразований, получаем

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

§ 26. ХАРАКТЕРИСТИЧЕСКАЯ ФУНКЦИЯ И МОМЕНТЫ СИСТЕМЫ СЛУЧАЙНЫХ ВЕЛИЧИН

Определение математического ожидания случайной величины как суммы произведений возможных значений случайной величины на соответствующие им вероятности (см. главу 4) остается в силе и при рассмотрении величины, являющейся функцией нескольких случайных величин, образующих систему.

Пусть, например, случайная величина Z связана со случайными величинами $(X_1, \dots, X_n)$ функциональной зависимостью:

$$Z = \varphi(X_1, \dots, X_n). \quad (26.1)$$

В этом случае для математического ожидания Z имеем

$$\begin{aligned} \mathbf{M}(Z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi(x_1, \dots, x_n) dP(x_1, \dots, x_n) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n, \end{aligned} \quad (26.2)$$

где $f(x_1, \dots, x_n)$ — плотность вероятности системы $(X_1, \dots, X_n)$.

Положим здесь

$$\varphi(x_1, \dots, x_n) = e^{i \sum_{j=1}^n u_j X_j},$$

где $u_1, u_2, \dots, u_n$ — произвольные вещественные числа. В этом случае $\mathbf{M}(Z)$ будет функцией n переменных $u_1, u_2, \dots, u_n$. Обозначим эту функцию $E(u_1, u_2, \dots, u_n)$ и будем называть *характеристической функцией системы случайных величин* $(X_1, \dots, X_n)$, то есть положим

$$\begin{aligned} E(u_1, \dots, u_n) &= \mathbf{M}[e^{i \sum_{j=1}^n u_j X_j}] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{j=1}^n u_j x_j} f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \vec{u}^T \vec{x}} f(\vec{x}) d\mathbf{x}, \quad (26.3) \end{aligned}$$

где под $\vec{u}$ и $\vec{x}$ понимаются вектора с компонентами $(u_1, \dots, u_n)$ и $(x_1, \dots, x_n)$, соответственно, под $d\mathbf{x}$ — сокращенное обозначение произведения дифференциалов $dx_1 dx_2 \dots dx_n$, а T обозначает символ транспонирования. Математическое ожидание (26.3) существует для любой системы случайных величин, так как является n -кратным интегральным преобразованием Фурье неотрицательной функции $f(x_1, \dots, x_n)$, интеграл от которой по всему n -мерному пространству ограничен вследствие выполнения условия нормировки плотности вероятностей (25.17):

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n = 1, \quad (26.4)$$

которое в новых обозначениях переписывается в виде

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(\vec{x}) d\mathbf{x} = 1.$$

Таким образом, характеристическая функция существует для любой системы случайных величин.

Как известно из теории интегралов Фурье [28], в случае существования прямого n -кратного преобразования Фурье (2.3) существует и обратное преобразование, определяемое формулой

$$f(\vec{x}) = \frac{1}{(2\pi)^n} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\vec{u}^T \vec{x}} E(\vec{u}) d\mathbf{u}, \quad (26.5)$$

где по-прежнему для краткости обозначено $d\mathbf{u} = \prod_{j=1}^n du_j$.

Таким образом, зная характеристическую функцию системы, можно вначале вычислить плотность вероятности, а затем по формуле (25.13) и функцию распределения системы.

Итак, наряду с плотностью вероятности и функцией распределения, характеристическая функция является исчерпывающей характеристикой системы случайных величин: любую систему случайных величин можно характеризовать или функцией распределения, или плотностью вероятности, или характеристической функцией. Выбор того или иного способа описания системы случайных величин зависит от существа решаемой задачи. Во многих случаях применение характеристических функций упрощает вычисления. Далее установим несколько общих свойств характеристической функции.

Подобно характеристической функции одномерного закона (см. § 20) многомерная характеристическая функция так же будет ограниченной

$$|E(\vec{u})| = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\vec{u}^T \vec{x}} f(\vec{x}) d\mathbf{x} \right| \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\vec{x}) d\mathbf{x} = 1,$$

то есть

$$|E(\vec{u})| \leq 1. \quad (26.6)$$

Если один из аргументов характеристической функции $E(\vec{u})$ системы n величин (например, u_n) положить равным

нулю, то из определения (26.3) получим (переобозначая временно $E(\vec{u})$ через $E_n(u_1, \dots, u_n)$):

$$\begin{aligned} E_n(u_1, \dots, u_{n-1}, 0) &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{j=1}^{n-1} u_j x_j} \cdot \left[\int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_n \right] dx_1 \dots dx_{n-1} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{j=1}^{n-1} u_j x_j} \cdot f(x_1, \dots, x_{n-1}) dx_1 \dots dx_{n-1} = \\ &= E_{n-1}(u_1, \dots, u_{n-1}). \end{aligned} \quad (26.7)$$

Следовательно, обращение в нуль одного (или нескольких) аргумента характеристической функции дает характеристическую функцию подсистемы случайных величин, соответствующей отличным от нуля аргументам исходной характеристической функции. Если же положить равными нулю **все** аргументы характеристической функции, то получим условие

$$E(\vec{0}) = 1, \quad (26.8)$$

являющееся прямым следствием условия нормировки (26.4).

Для независимых случайных величин характеристическая функция системы $(X_1, \dots, X_n)$ равняется произведению характеристических функций $E_j(u_j)$ отдельных случайных величин, образующих эту систему,

$$E(\vec{u}) = \prod_{j=1}^n E_j(u_j).$$

Это следует из того, что для независимых случайных величин плотность вероятности системы распадается на произведение плотностей вероятности каждой из случайных величин, входящих в систему, и, следовательно, из определения характеристической функции (26.3) имеем

$$\begin{aligned}
 E(\vec{u}) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{j=1}^n u_j x_j} \prod_{j=1}^n [f(x_j) dx_j] = \\
 &= \prod_{j=1}^n \left[\int_{-\infty}^{\infty} e^{i u_j x_j} f(x_j) dx_j \right] = \prod_{j=1}^n E_j(u_j), \quad (26.9)
 \end{aligned}$$

то есть выписанное выше равенство.

Будем рассматривать систему случайных величин $(X_1, \dots, X_n)$ как вектор $\vec{X}$. Если от случайного вектора $\vec{X}$ перейти к случайному вектору $\vec{Y}$, соответствующему системе случайных величин $(Y_1, \dots, Y_n)$, получаемых путем линейного преобразования

$$\vec{Y} = \mathbf{A}\vec{X} + \vec{b}, \quad (26.10)$$

где $\mathbf{A} = \|a_{jl}\|_{j,l=1}^n$, $\vec{b} = \{b_i\}_{i=1}^n$ — неслучайные матрица и вектор, то характеристическая функция $E_y(\vec{u})$ новой системы выражается через характеристическую функцию исходной системы $E_x(\vec{u})$ по формуле

$$E_y(\vec{u}) = e^{i\vec{u}^T \vec{b}} E_x(\mathbf{A}^T \vec{u}), \quad (26.11)$$

которая непосредственно следует из определения характеристической функции (26.3).

Если линейное преобразование имеет более простой вид, а именно

$$Y_j = \alpha_j X_j + \beta_j, \quad (j = \overline{1, n}), \quad (26.12)$$

то

$$E_y(\vec{u}) = e^{i\vec{\beta}^T \vec{u}} E_x(\vec{v}), \quad (26.13)$$

где $\vec{\beta} = \{\beta_i\}_{i=1}^n$, а $\vec{v} = \{v_i\}_{i=1}^n$ имеет компоненты $v_i = \alpha_i u_i$. В частном случае, если $\alpha_i = 1$, $\beta_j = -\bar{x}_j$, то $Y_j = X_j - \bar{x}_j$, то есть равны центрированным значениям случайных величин $\overset{\circ}{X}_j$. Обозначая в этом случае $E_y(\vec{u})$ через $\overset{\circ}{E}_x(\vec{u})$, получаем

$$\overset{\circ}{E}_x(\vec{u}) = e^{-i \sum_{j=1}^n \bar{x}_j u_j} E_x(\vec{u}). \quad (26.14)$$

Можно доказать, что для системы *непрерывных* случайных величин характеристическая функция системы обращается в нуль при стремлении к бесконечности модуля хотя бы одного из ее аргументов, то есть справедливо равенство [28]

$$E(\vec{u})|_{|u_j|=\infty} = 0, \quad (j = \overline{1, n}). \quad (26.15)$$

Для дискретных случайных величин это свойство не имеет места, так как характеристическая функция в этом случае выражается через сумму конечного числа экспонент с чисто мнимым показателем, модуль которых не может тождественно обратиться в нуль при стремлении аргумента к бесконечности. Действительно, предположим для примера, что имеется система двух дискретных случайных величин (X, Y) , которые могут принимать значения (x_j, y_l) с вероятностями p_{jl} , $(j = \overline{1, n}; l = \overline{1, m})$. Вычисляя в этом случае характеристическую функцию $E(u_1, u_2)$ системы по формуле (26.3), получим

$$E(u_1, u_2) = \sum_{j=1}^n \sum_{l=1}^m e^{i(u_1 x_j + u_2 y_l)} p_{jl}, \quad (26.16)$$

и это выражение вовсе не обращается в нуль при значениях $|u_1|$ и $|u_2|$, стремящихся к бесконечности, а колеблется около нуля по гармоническому закону.

Характеристические функции особенно удобны при вычислении моментов системы случайных величин, определяемых, как обычно, через математические ожидания произведения целых степеней случайных величин, входящих в систему. Для системы случайных величин, точно также как для одной величины, рассматривают начальные и центральные моменты.

Начальным моментом k -го порядка $m_{i_1, i_2, \dots, i_n}$ системы случайных величин $(X_1, \dots, X_n)$ называют математическое ожидание произведения $X_1^{i_1}, \dots, X_n^{i_n}$, то есть

$$m_{i_1, i_2, \dots, i_n} = \mathbf{M} \left(\prod_{j=1}^n X_j^{i_j} \right) = \\ = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{j=1}^n x_j^{i_j} f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n, \quad (26.17)$$

где $\sum_{j=1}^n i_j = k$, $i_j \geq 0$, $(j = \overline{1, n}; k = \overline{1, \infty})$.

Центральным моментом k -го порядка $\mu_{i_1, i_2, \dots, i_n}$ называют математическое ожидание произведения $\overset{\circ}{X}_1^{i_1}, \dots, \overset{\circ}{X}_n^{i_n}$ следующего вида:

$$\mu_{i_1, i_2, \dots, i_n} = \mathbf{M} \left(\prod_{j=1}^n \overset{\circ}{X}_j^{i_j} \right) = \\ = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{j=1}^n (x_j - \bar{x}_j)^{i_j} f(\vec{x}) d\mathbf{x}, \quad (26.18)$$

где $\sum_{j=1}^n i_j = k$, $i_j \geq 0$, $\overset{\circ}{X}_j = X_j - \bar{x}_j$, $(j = \overline{1, n}; k = \overline{2, \infty})$.

Для существования моментов заданного порядка соответствующие интегралы в правых частях равенств (26.17) и (26.18) должны сходиться абсолютно. Рассмотрим подробнее младшие моменты первого и второго порядка.

Для начальных моментов первого порядка $k = 1$, и, следовательно, все индексы момента $m_{i_1, i_2, \dots, i_n}$ должны равняться нулю, исключая какой-то один индекс, равный единице. Следовательно, как это ясно из определения (26.17), система n случайных величин $(X_1, \dots, X_n)$ имеет n моментов первого порядка (у каждого из которых отличный от нуля индекс стоит на j -м месте), совпадающих с математическими ожиданиями этих величин

$$m_{\underbrace{0 \dots 0}_{j-1} \underbrace{1 0 \dots 0}_{n-j}} = \bar{x}_j. \quad (26.19)$$

Моменты второго порядка ($k = 2$) могут быть двух типов: моменты, у которых отличным от нуля является один индекс,

равный двум, и моменты, имеющие два отличных от нуля индекса, каждый из которых равняется единице. Моменты первого типа, как это непосредственно следует из определений (26.17) и (26.18), совпадают со вторыми моментами случайных величин, входящих в систему

$$\begin{aligned} m_{\underbrace{0\dots 0}_{j-1}\underbrace{20\dots 0}_{n-j}} &= m_2^{(x_j)}, \\ \mu_{\underbrace{0\dots 0}_{j-1}\underbrace{20\dots 0}_{n-j}} &= \mu_2^{(x_j)} = \mathbf{D}(X_j), \quad (j = \overline{1, n}). \end{aligned} \quad (26.20)$$

Следовательно, вторые моменты этого типа не содержат никакой информации относительно связи между величинами, входящими в систему.

Центральные моменты второго типа называются *корреляционными моментами* или *ковариациями* между двумя величинами. Для обозначения корреляционного момента используется символ k_{jl} или $\text{cov}(X_j, X_l)$

$$\begin{aligned} k_{jl} &= \mathbf{M}(\overset{\circ}{X}_j \overset{\circ}{X}_l) = \text{cov}(X_j, X_l) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_j - \bar{x}_j)(x_l - \bar{x}_l) f(x_j, x_l) dx_j dx_l, \end{aligned} \quad (26.21)$$

где последний двойной интеграл может быть получен из n -кратного интеграла (26.17) путем интегрирования по остальным $n - 2$ переменным с учетом (25.15). В дальнейшем для обозначения корреляционных моментов будет использоваться только первое обозначение. Совокупность корреляционных моментов k_{jl} , соответствующих рассматриваемой системе случайных величин, называется *корреляционной матрицей* этих величин

$$\mathbf{K} = \|k_{ij}\|_{i,j=1}^n.$$

Установим ряд основных свойств корреляционных моментов. Во-первых, из определения (26.21) следует, что диагональные моменты совпадают с дисперсией компонент

$$k_{jj} = \mathbf{D}(X_j) = \sigma_{x_j}^2. \quad (26.22)$$

Во-вторых, из того же определения следует, что корреляционная матрица $\mathbf{K}$ симметрична

$$k_{jl} = k_{lj}, \quad (j = \overline{1, n}; j = \overline{1, n}), \quad (26.23)$$

то есть $\mathbf{K} = \mathbf{K}^T$.

Для независимых случайных величин X_j, X_l корреляционный момент равен нулю¹, так как в этом случае ($j \neq l$)

$$f(x_j, x_l) = f(x_j)f(x_l),$$

$$\begin{aligned} k_{jl} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_j - \overline{x_j})(x_l - \overline{x_l}) f(x_j) f(x_l) dx_j dx_l = \\ &= \left[\int_{-\infty}^{\infty} (x_j - \overline{x_j}) f(x_j) dx_j \right] \times \\ &\quad \times \left[\int_{-\infty}^{\infty} (x_l - \overline{x_l}) f(x_l) dx_l \right] = 0. \end{aligned} \quad (26.24)$$

Применяя к интегралу (26.21) неравенство Коши–Буняковского

$$\left[\int_{\Omega} g(\vec{x}) q(\vec{x}) d\mathbf{x} \right]^2 \leq \int_{\Omega} g^2(\vec{x}) d\mathbf{x} \int_{\Omega} q^2(\vec{x}) d\mathbf{x}, \quad (26.25)$$

где Ω — произвольная область n -мерного евклидова пространства, и полагая при этом $g = (x_j - \overline{x_j}) \cdot \sqrt{f(x_j, x_l)}$,

¹ Обратное утверждение не имеет места, т. е. из равенства нулю k_{jl} еще нельзя заключить, что величины X_j и X_l независимы.

$q = (x_l - \overline{x_l}) \cdot \sqrt{f(x_j, x_l)}$, получаем

$$k_{jl}^2 \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_j - \overline{x_j})^2 f(x_j, x_l) dx_j dx_l \times \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_l - \overline{x_l})^2 f(x_j, x_l) dx_j dx_l = \sigma_{x_j}^2 \sigma_{x_l}^2,$$

откуда следует

$$|k_{jl}| \leq \sigma_{x_j} \sigma_{x_l}. \quad (26.26)$$

Отношение корреляционного момента k_{jl} к произведению $\sigma_{x_j} \sigma_{x_l}$ называют *коэффициентом корреляции* между случайными величинами X_j и X_l . Обозначим коэффициент корреляции через r_{jl} , определив его формулой

$$r_{jl} = \frac{k_{jl}}{\sigma_{x_j} \sigma_{x_l}}. \quad (26.27)$$

В соответствии с (26.26)

$$-1 \leq r_{jl} \leq 1. \quad (26.28)$$

Своего наибольшего по модулю значения коэффициент корреляции r_{jl} достигает в том случае, когда случайные величины X_j и X_l связаны линейной зависимостью. Действительно, пусть

$$X_j = aX_l + b. \quad (26.29)$$

Тогда в соответствии с формулами (26.12) и (26.21) получим

$$k_{jl} = a\sigma_l^2, \quad \sigma_{x_j} = a^2\sigma_{x_l}^2, \quad (26.30)$$

а по формуле (26.27) оказывается, что

$$r_{jl} = \frac{a\sigma_l^2}{|a|\sigma_l^2} = \text{sign } a. \quad (26.31)$$

Следовательно, корреляционный момент равен единице, если $a > 0$ и минус единице, если $a < 0$. Обращение в нуль

коэффициента корреляции свидетельствует об отсутствии линейной зависимости между рассматриваемыми величинами.

Центральные моменты системы случайных величин могут быть выражены через начальные моменты, а начальные моменты могут быть выражены через центральные моменты и математические ожидания случайных величин, входящих в систему. Конкретный вид этих зависимостей аналогичен соответствующим формулам главы 4.

Например, обозначив начальный момент второго порядка через m_{jl} , имеем

$$\begin{aligned} m_{jl} &= \mathbf{M}(X_j X_l) = \mathbf{M}(\overset{\circ}{X}_j + \bar{x}_j)(\overset{\circ}{X}_l + \bar{x}_l) = \\ &= \mathbf{M}(\overset{\circ}{X}_j \overset{\circ}{X}_l + \bar{x}_j \overset{\circ}{X}_l + \bar{x}_l \overset{\circ}{X}_j + \bar{x}_j \bar{x}_l) = k_{jl} + \bar{x}_j \bar{x}_l. \end{aligned} \quad (26.32)$$

Таким образом,

$$m_{jl} = k_{jl} + \bar{x}_j \bar{x}_l, \quad k_{jl} = m_{jl} - \bar{x}_j \bar{x}_l. \quad (26.33)$$

Аналогично могут быть получены подобные (26.33) соотношения и между моментами более высокого порядка.

Характеристическая функция системы случайных величин является производящей функцией для моментов системы. Действительно, дифференцируя интеграл (26.3) k раз по всей совокупности аргументов, причем i_j раз частным образом по u_j ($j = \overline{1, n}$) и положив после дифференцирования все переменные u_j равными нулю, получаем

$$m_{i_1, i_2, \dots, i_n} = \frac{1}{i^k} \frac{\partial^k}{\partial u_1^{i_1} \dots \partial u_n^{i_n}} E(\vec{u})|_{\vec{u}=\vec{0}}, \quad (26.34)$$

где $k = \overline{1, \infty}$, а $\vec{0} = \{0, \dots, 0\}$ обозначает нулевой вектор.

Так как для центрированных случайных величин начальные моменты совпадают с центральными, подставляя в последнюю формулу вместо характеристической функции системы $E(\vec{u})$ характеристическую функцию центрированного вектора $\vec{X}$, обозначаемую $\overset{\circ}{E}(\vec{u})$ и определяемую формулой

(26.14), для центральных моментов системы случайных величин получаем выражение

$$\begin{aligned}\mu_{i_1, i_2, \dots, i_n} &= \frac{1}{i^k} \frac{\partial^k}{\partial u_1^{i_1+1} \dots \partial u_n^{i_n}} \overset{\circ}{E}(\vec{u})|_{\vec{u}=\vec{0}} = \\ &= \frac{1}{i^k} \frac{\partial^k}{\partial u_1^{i_1+1} \dots \partial u_n^{i_n}} [e^{-i\vec{u}^T \vec{m}} E(\vec{u})]|_{\vec{u}=\vec{0}}, \quad (26.35)\end{aligned}$$

где $\vec{m} = \{\bar{x}_1, \dots, \bar{x}_n\}$ — вектор, составленный из математических ожиданий компонент.

При выводе формул (26.34) и (26.35) было использовано правило Лейбница² дифференцирования под знаком интеграла по параметру. Как известно, это правило справедливо, если рассматриваемый интеграл сходится равномерно. Применительно к вычислению моментов это означает, что формулы (26.34) и (26.35) справедливы только в том случае, если соответствующий момент существует.

При вычислении математических ожиданий функций случайных величин, образующих систему, некоторые из величин системы можно считать фиксированными. Математическое ожидание, вычисляемое при этих условиях, называется *условным*, а фиксированные величины записываются справа от вертикальной черты под знаком математического ожидания. Если фиксированные величины заданы, то эти величины будем обозначать малыми буквами, если же они заменяются на случайную величину, — то, как обычно, большими буквами. Так, например, $\mathbf{M}[\varphi(X_1, \dots, X_n)|x_1]$ обозначает математическое ожидание функции $\varphi(X_1, \dots, X_n)$, найденное при условии, что случайная величина X_1 фиксирована и равна x_1 , а $\mathbf{M}[\varphi(X_1, \dots, X_n)|X_1]$ обозначает математическое ожидание той же функции, но вычисленное при фиксированном значении X_1 , которое считается случайным. Условное математическое ожидание первого типа является постоянной величиной, а условное математическое ожидание второго типа, естественно, будет величиной случайной.

² Лейбниц Готфрид Вильгельм (1646–1716) — немецкий математик, физик и философ.

Рассмотрим несколько соотношений, которым удовлетворяют моменты случайных величин первых двух порядков. Эти соотношения часто называют *теоремами о математическом ожидании и дисперсии*.

Теоремами о математическом ожидании в нашем курсе будем называть следующие свойства этой операции:

$$\begin{cases} \mathbf{M}(a) = a, \\ \mathbf{M}(aX) = a\mathbf{M}(X), \\ \mathbf{M}(X + Y) = \mathbf{M}(X) + \mathbf{M}(Y), \\ \mathbf{M}(X) = \mathbf{M}_y[\mathbf{M}_x(X|Y)], \\ \mathbf{M}(XY) = \mathbf{M}(X)\mathbf{M}(Y) + k_{xy}, \end{cases} \quad (26.36)$$

где a — неслучайная величина.

Первое и *второе* равенства непосредственно следуют из формулы, определяющей математическое ожидание. Для доказательства *третьей* теоремы (третьего равенства) имеем

$$\begin{aligned} \mathbf{M}(X + Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy = \\ &= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x, y) dx \right] dy = \\ &= \int_{-\infty}^{\infty} x f(x) dx + \int_{-\infty}^{\infty} y f(y) dy = \mathbf{M}(X) + \mathbf{M}(Y). \end{aligned} \quad (26.37)$$

Как ясно из доказательства, третья теорема сохраняет силу и в том случае, когда величины X и Y зависимы. Эта теорема легко обобщается на сумму любого числа слагаемых, поскольку сумму произвольного числа слагаемых всегда можно вначале привести к сумме двух слагаемых, из которых второе включает в себя остальные члены суммы, а затем применить доказанную теорему в комбинации с методом математической

индукции. Таким образом, для любой системы независимо от числа и входящих в нее величин имеем

$$\mathbf{M} \left(\sum_{j=1}^n X_j \right) = \sum_{j=1}^n \mathbf{M}(X_j). \quad (26.38)$$

Наконец, для доказательства *четвертой* теоремы достаточно представить на основании (25.21) плотность вероятности в виде произведения $f(x|y)f(y)$, после чего выполнить следующие преобразования интегралов:

$$\begin{aligned} \mathbf{M}(X) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [xf(x|y)dx]f(y) dy = \\ &= \int_{-\infty}^{\infty} \mathbf{M}_x(X|y)f(y) dy = \mathbf{M}_y[\mathbf{M}_x(X|Y)], \end{aligned} \quad (26.39)$$

где $\mathbf{M}_x$ обозначают математическое ожидание по распределению X , а $\mathbf{M}_y$ — ту же операцию по распределению Y .

Из характера доказательства теорем о математическом ожидании ясно, что эти теоремы остаются в силе и в том случае, когда случайные величины X и Y зависят от некоторой случайной величины Z , или от нескольких таких величин, а речь идет о вычислении условных математических ожиданий при фиксированном Z (или подобных им величин).

Наконец, для доказательства последней, *пятой* теоремы выводим

$$\begin{aligned} \mathbf{M}(XY) &= \mathbf{M}[(\overset{\circ}{X} + \bar{x})(\overset{\circ}{Y} + \bar{y})] = \\ &= \mathbf{M}[\overset{\circ}{X}\overset{\circ}{Y} + \bar{x}\overset{\circ}{Y} + \bar{y}\overset{\circ}{X} + \bar{x}\bar{y}] = k_{xy} + \bar{x}\bar{y}. \end{aligned}$$

Если X и Y независимые, то $k_{xy} = 0$ и тогда

$$\mathbf{M}(XY) = \mathbf{M}(X)\mathbf{M}(Y).$$

Аналогичные теоремы имеют место для дисперсии

$$\begin{cases} \mathbf{D}(a) = 0, \\ \mathbf{D}(aX) = a^2 \mathbf{D}(X), \\ \mathbf{D}(X + Y) = \mathbf{D}(X) + \mathbf{D}(Y) + 2k_{xy}, \\ \mathbf{D}(X) = \mathbf{M}_y[\mathbf{D}_x(X|Y)] + \mathbf{D}_y[\mathbf{M}_x(X|Y)], \end{cases} \quad (26.40)$$

причем, если X и Y независимые, то выполняется еще одно равенство

$$\mathbf{D}(XY) = \mathbf{D}(X)\mathbf{D}(Y) + \bar{x}^2 \mathbf{D}(Y) + \bar{y}^2 \mathbf{D}(X).$$

Здесь a — по-прежнему неслучайная величина, k_{xy} — корреляционный момент между случайными величинами X и Y , а $\mathbf{D}_x(X|Y)$ — обозначение условной дисперсии X , найденной при фиксированном Y .

Первые две теоремы непосредственно следуют из определения дисперсии, а для доказательства *третьей* теоремы в соответствии с определением дисперсии и третьей теоремой о математическом ожидании имеем

$$\begin{aligned} \mathbf{D}(X + Y) &= \mathbf{M}\{([X + Y] - \mathbf{M}(X + Y))^2\} = \\ &= \mathbf{M}\{(\overset{\circ}{X} + \overset{\circ}{Y})^2\} = \mathbf{M}[\overset{\circ}{X}^2 + \overset{\circ}{Y}^2 + 2\overset{\circ}{X}\overset{\circ}{Y}] = \\ &= \mathbf{D}(X) + \mathbf{D}(Y) + 2k_{xy}, \end{aligned} \quad (26.41)$$

что и требовалось доказать.

Если случайные величины X и Y независимы, то $k_{xy} = 0$ и последняя формула упрощается

$$\mathbf{D}(X + Y) = \mathbf{D}(X) + \mathbf{D}(Y). \quad (26.42)$$

Следовательно, дисперсия суммы двух независимых случайных величин равняется сумме дисперсий слагаемых. Эта формула непосредственно обобщается на сумму любого числа независимых случайных величин. При любом числе независимых случайных величин $(X_1, \dots, X_n)$ имеем

$$\mathbf{D}\left(\sum_{j=1}^n X_j\right) = \sum_{j=1}^n \mathbf{D}(X_j). \quad (26.43)$$

Отметим, что формулы (26.42) и (26.43) остаются в силе и для зависимых величин, если все корреляционные моменты k_{jl} при $j \neq l$ обращаются в нуль³. В общем случае формула (26.43) усложняется и приобретает вид

$$\mathbf{D} \left(\sum_{j=1}^n X_j \right) = \sum_{j=1}^n \sum_{i=1}^n k_{ij} = \sum_{j=1}^n \mathbf{D}(X_j) + 2 \sum_{j=2}^n \sum_{l=1}^{j-1} k_{jl}. \quad (26.44)$$

Для доказательства *четвертой* теоремы на основании четвертой теоремы о математическом ожидании имеем

$$\mathbf{D}(X) = \mathbf{M}(\overset{\circ}{X})^2 = \mathbf{M}_y[\mathbf{M}_x(\overset{\circ}{X}^2 | Y)]. \quad (26.45)$$

Представив центрированную случайную величину $\overset{\circ}{X}$ в виде суммы $[X - \mathbf{M}_x(X|Y)] + [\mathbf{M}_x(X|Y) - \bar{x}]$, последней формуле можно придать вид

$$\begin{aligned} \mathbf{D}(X) = \mathbf{M}_y(\mathbf{M}_x\{[X - \mathbf{M}_x(X|Y)]^2 + [\mathbf{M}_x(X|Y) - \bar{x}]^2 + \\ + 2[\mathbf{M}_x(X|Y)][\mathbf{M}_x(X|Y) - \bar{x}]Y\}). \end{aligned} \quad (26.46)$$

Применяя при вычислении условного математического ожидания третью теорему о математическом ожидании и замечая при этом, что $\mathbf{M}_x(X|Y)$ при вычислении условного математического ожидания следует рассматривать как постоянную (неслучайную) величину и, следовательно,

$$\begin{aligned} \mathbf{M}_x\{[X - \mathbf{M}_x(X|Y)][\mathbf{M}_x(X|Y) - \bar{x}]Y\} = \\ = [\mathbf{M}_x(X|Y) - \bar{x}]\mathbf{M}_x\{[X - \mathbf{M}_x(X|Y)]Y\} = \\ = [\mathbf{M}_x(X|Y) - \bar{x}][\mathbf{M}_x(X|Y) - \mathbf{M}_x(X|Y)] = 0, \end{aligned}$$

получаем

$$\mathbf{D}(X) = \mathbf{M}_y(\mathbf{M}_x\{[X - \mathbf{M}_x(X|Y)]^2|Y\}) + \mathbf{D}_y[\mathbf{M}_x(X|Y)],$$

то есть несколько иначе переписанную четвертую формулу (26.40).

³ Такие величины называются *некоррелированными*.

Наконец, последняя, *пятая* теорема следует из теорем о математическом ожидании, если учесть, что в данном случае $\mathbf{M}(YX) = \bar{x} \bar{y}$ и, следовательно,

$$\begin{aligned} \mathbf{D}(XY) &= \mathbf{M}[(XY - \bar{x} \bar{y})^2] = \\ &= \mathbf{M}[\overset{\circ}{X} \overset{\circ}{Y} + 2\overset{\circ}{X}\overset{\circ}{Y}(\bar{x}\overset{\circ}{Y} + \bar{y}\overset{\circ}{X}) + \bar{x}^2\overset{\circ}{Y}^2 + 2\bar{x}\bar{y}\overset{\circ}{X}\overset{\circ}{Y} + \bar{y}^2\overset{\circ}{X}^2] = \\ &= \mathbf{D}(X)\mathbf{D}(Y) + \bar{x}^2\mathbf{D}(Y) + \bar{y}^2\mathbf{D}(X). \end{aligned}$$

Коэффициент корреляции недостаточно чувствителен к наличию или отсутствию функциональной зависимости между случайными величинами, если эта зависимость *нелинейна*. Так, при нелинейной зависимости коэффициент корреляции k_{jl} может обращаться в нуль, даже если такая зависимость налицо (например, в случае центральной симметрии плотности вероятности $f(x_j, x_l)$ и четной функциональной зависимости X_j от X_l). Поэтому для числовой характеристики степени зависимости между двумя случайными величинами применяют и другие, более тонкие методы. Одним из наиболее употребительных здесь является показатель связи, называемый *корреляционным отношением*. Рассмотрим его подробнее.

Рассмотрим систему (X, Y) . Назовем корреляционным отношением $\eta_{x|y}$ величину, определяемую формулой

$$\eta_{x|y} = \frac{1}{\sigma_x} \sqrt{\mathbf{D}_y[\mathbf{M}_x(X|Y)]}. \quad (26.47)$$

Заменяя в (26.47) дисперсию условного математического ожидания $\mathbf{D}_y[\mathbf{M}_x(X|Y)]$ в соответствии с четвертой теоремой о дисперсии (26.40) разностью $\{\sigma_x^2 - \mathbf{M}_y[\mathbf{D}_x(X|Y)]\}$, получим для корреляционного отношения другое выражение:

$$\eta_{x|y} = \sqrt{1 - \frac{1}{\sigma_x} \mathbf{M}_y[\mathbf{D}_x(X|Y)]}. \quad (26.48)$$

Из формул (26.47) и (26.48) следует, что в общем случае корреляционное отношение не симметрично относительно своих индексов, и для выполнения равенства $\eta_{x|y} = \eta_{y|x}$

плотность вероятности $f(x, y)$ должна обладать особыми свойствами.

При наличии однозначной функциональной зависимости X от Y , имеем, очевидно, $\mathbf{D}(X|Y) = 0$ так как при заданном Y величина X перестает быть случайной. Следовательно, $\eta_{x|y} = 1$, если X является однозначной функцией Y . Наоборот, если $\eta_{x|y} = 1$, в соответствии с (26.48) $\mathbf{M}[\mathbf{D}(X|Y)] = 0$, а поскольку дисперсия не может быть отрицательной, то отсюда следует, что $\mathbf{D}(X|Y) = 0$, то есть величины X и Y связаны однозначной зависимостью.

Очевидно, что для независимых случайных величин $\eta_{x|y} = 0$. Можно доказать, что для некоррелированных случайных величин $\eta_{x|y} = 0$. Значит, обращение в нуль корреляционного отношения гарантирует некоррелированность величин X и Y , которые тем не менее могут быть и зависимыми. Следовательно, этот параметр не решает полностью проблему выявления зависимостей, однако он гораздо более чувствителен, чем обычно применяемый коэффициент корреляции.

Рассмотрим несколько примеров, поясняющих формулы данного параграфа.

Пример 26.1. В условиях примера 25.4 (равномерное распределение декартовых координат X, Y точки в круге радиуса a) определить моменты первого и второго порядка системы (X, Y) , найти характеристическую функцию системы и вычислить корреляционное отношение $\eta_{x|y}$.

Решение. Так как плотность вероятности

$$f(x, y) = \frac{1}{\pi a^2} \quad (x^2 + y^2 \leq a^2)$$

обладает центральной симметрией, то первые начальные моменты равны нулю $\bar{x} = \bar{y} = 0$, а следовательно, центральные моменты совпадают с начальными. При вычислении вторых моментов, переходя к полярной системе координат, имеем

$$\sigma_x^2 = \frac{1}{\pi a^2} \cdot \int_0^{2\pi} \int_0^a \rho^2 \cos^2 \varphi \rho d\rho d\varphi = \frac{a^2}{4},$$

$$\sigma_y^2 = \frac{1}{\pi a^2} \cdot \int_0^{2\pi} \int_0^a \rho^2 \sin^2 \varphi \rho d\rho d\varphi = \frac{a^2}{4},$$

$$k_{xy} = \frac{1}{\pi a^2} \cdot \int_0^{2\pi} \int_0^a \rho^2 \cos \varphi \sin \varphi \rho d\rho d\varphi = 0.$$

Таким образом, рассматриваемые величины, хотя и являются зависимыми (см. пример 25.4), имеют нулевой корреляционный момент. Применяя общую формулу (26.3), для характеристической функции рассматриваемой системы имеем

$$E(u_1, u_2) = \frac{1}{\pi a^2} \cdot \iint_{x^2+y^2 \leq a^2} e^{i(u_1 x + u_2 y)} dx dy.$$

Переходя и в этом случае к полярным координатам и обозначая для краткости $u = \sqrt{u_1^2 + u_2^2}$, получаем

$$E(u_1, u_2) = \frac{1}{\pi a^2} \cdot \int_0^a \int_0^{2\pi} e^{iu\rho \cos(\varphi-\gamma)} d\varphi d\rho, \quad \gamma = \arctg \frac{u_2}{u_1}.$$

Так как по свойствам бесселевых функций [17]

$$\int_0^{2\pi} e^{iz \cos(\varphi-\gamma)} d\varphi = \int_0^{2\pi} e^{iz \cos(\Psi)} d\Psi = J_0(z),$$

где $J_0(z)$ — функция Бесселя первого ряда и, кроме того

$$\int J_0(z) z dz = z J_1(z),$$

то для искомой характеристической функции находим

$$E(u_1, u_2) = \frac{2}{au} \cdot J_1(au) = \frac{2}{a\sqrt{u_1^2 + u_2^2}} J_1(a\sqrt{u_1^2 + u_2^2}).$$

Условная плотность вероятности $f(x|y) = \frac{1}{\sqrt{a^2 - y^2}}$ (см. пример 25.4), то есть условный закон распределения случайной

величины X при заданном $Y = y$ есть закон равномерного распределения с параметром $l_y = a^2 - y^2$. Условная дисперсия этого закона согласно результатам главы 5 равняется

$\mathbf{D}(X|Y) = \frac{1}{3} l_y^2 = \frac{1}{3} (a^2 - Y^2)$, а математическое ожидание

условной дисперсии $\mathbf{M}[\mathbf{D}(X|Y)] = \frac{1}{3} \left(a^2 - \frac{a^2}{4} \right) = \frac{a^2}{4}$. Под-

ставляя это значение в (26.48), получаем $\eta_{x|y} = 0$, то есть корреляционное отношение в данной задаче обращается в нуль, хотя величины X и Y зависимые.

Пример 26.2. Определить математическое ожидание и дисперсию числа появления события A в серии n независимых испытаний, если вероятность появления события A в j -м испытании равна p_j .

Решение. Введем систему независимых случайных величин (индикаторов), X_j , равных единице, если в j -м опыте произошло событие A и нулю в противном случае. Тогда

$$X = \sum_{j=1}^n X_j.$$

Применяя теоремы о математическом ожидании и дисперсии, находим

$$\mathbf{M}(X) = \sum_{j=1}^n \mathbf{M}(X_j), \quad \mathbf{D}(X) = \sum_{j=1}^n \mathbf{D}(X_j).$$

Так как величина X_j может принимать только два значения: единица с вероятностью p_j и нуль с вероятностью $q_j = 1 - p_j$, то

$$\mathbf{M}(X_j) = p_j, \quad \mathbf{D}(X_j) = \mathbf{M}(X_j^2) - p_j^2 = p_j - p_j^2 = p_j q_j$$

и окончательно получим

$$\mathbf{M}(X) = \sum_{j=1}^n p_j, \quad \mathbf{D}(X) = \sum_{j=1}^n p_j q_j.$$

В частном случае, когда $p_j = p$, ($j = \overline{1, n}$) (обычная схема Бернулли), имеем хорошо известный нам результат

$$\mathbf{M}(X) = np, \quad \mathbf{D}(X) = np(1 - p).$$

Пример 26.3. Число заявок N , поступающих в систему массового обслуживания в течение времени t , подчиняется закону Пуассона с параметром $a = \Delta t$. Определить математическое ожидание и дисперсию числа поступающих заявок, если параметр Δ (интенсивность потока заявок) — случайная величина, плотность вероятности которой $f(\lambda) = \frac{\lambda}{b} e^{-\frac{\lambda^2}{2b^2}}$, $\lambda \geq 0$ (закон Рэлея), где b — известная постоянная.

Решение. Применяя теоремы о математическом ожидании, имеем

$$\mathbf{M}(N) = \mathbf{M}[\mathbf{M}(N|\Delta)] = \mathbf{M}(\Delta t) = t \int_0^{\infty} \frac{\lambda^2}{b} \cdot e^{-\frac{\lambda^2}{2b^2}} d\lambda = \sqrt{\frac{\pi}{2}} bt.$$

Теоремы о дисперсии дают

$$\begin{aligned} \mathbf{D}(N) &= \mathbf{M}[\mathbf{D}(N|\Delta)] + \mathbf{D}[\mathbf{M}(N|\Delta)] = \mathbf{M}(\Delta t) + \mathbf{D}(\Delta t) = \\ &= \sqrt{\frac{\pi}{b}} bt + t^2 \left[\int_0^{\infty} \frac{\lambda^3}{b} e^{-\frac{\lambda^2}{2b^2}} d\lambda - \left(\sqrt{\frac{\pi}{2}} b \right)^2 \right] = \\ &= \sqrt{\frac{\pi}{2}} bt + 2b^2 t^2 \left(1 - \frac{\pi}{4} \right). \end{aligned}$$

Пример 26.4. Показать, что корреляционная матрица $\mathbf{K} = \|k_{jl}^{(x)}\|_{i,j=1}^n$ является положительно определенной, то есть для любых вещественных чисел a_j

$$\sum_{j=1}^n \sum_{l=1}^n a_j a_l k_{jl}^{(x)} \geq 0,$$

причем знак равенства достигается только при $a_j \equiv 0$, ($j = \overline{1, n}$). Здесь предполагается, что хотя бы для одного значения j имеем $\sigma_j^2 > 0$.

Решение. Обозначим $Y_j = a_j X_j$ и положим $Y = \sum_{j=1}^n a_j Y_j$.

В соответствии с (26.44) получаем

$$\mathbf{D}(Y) = \sum_{j=1}^n \sum_{l=1}^n k_{jl}^{(y)},$$

где $k_{jl}^{(y)} = \mathbf{M}(\overset{\circ}{Y}_j \overset{\circ}{Y}_l) = a_j a_l k_{jl}^{(x)}$. Следовательно, квадратичная форма (26.49) совпадает с дисперсией введенной величины Y

$$\mathbf{D}(Y) = \sum_{j=1}^n \sum_{l=1}^n a_j a_l k_{jl}^{(x)}.$$

Поскольку $\mathbf{D}(Y) > 0$, когда среди a_j не все $a_j = 0$, то заключаем, что матрица $\mathbf{K}$ положительно определена.

§ 27. МНОГОМЕРНЫЙ НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

Наиболее важным примером закона распределения системы произвольного числа n непрерывных случайных величин $X_1, \dots, X_n$ является нормальный закон распределения n -го порядка, или, сокращенно, многомерный нормальный закон распределения.

Будем называть многомерным нормальным законом распределения закон, характеристическая функция которого $E(\vec{u})$ может быть представлена в виде экспоненты, содержащей в показателе степени полином второй степени относительно аргументов характеристической функции $u_1, \dots, u_n$

$$E(\vec{u}) = \exp \left[\sum_{j=1}^n \alpha_j u_j + \sum_{j=1}^n \sum_{l=1}^n \beta_{jl} u_j u_l \right], \quad (27.1)$$

где без нарушения общности можно считать $\beta_{jl} = \beta_{lj}$, а свободный член в показателе степени отсутствует вследствие нормировки (26.8), согласно которой $E(\vec{0}) = 1$.

Выразим постоянные α_j и β_{jl} через моменты рассматриваемой системы случайных величин. Используя формулу (26.34),

устанавливающую связь между начальными моментами и характеристической функцией системы, для математических ожиданий $\overline{x_j}$ случайных величин X_j получим

$$\overline{x_j} = \left(\frac{1}{i} \alpha_j + 2 \sum_{l=1}^n \beta_{jl} u_l E(\vec{u}) \right) \Big|_{\vec{u}=\vec{0}} = \frac{1}{i} \alpha_j. \quad (27.2)$$

Следовательно, характеристическая функция многомерного нормального закона распределения может быть представлена в виде

$$E(\vec{u}) = e^{i \sum_{j=1}^n x_j u_j + \sum_{j=1}^n \sum_{l=1}^n \beta_{jl} u_j u_l}. \quad (27.3)$$

Воспользовавшись формулой (26.14) для характеристической функции $\overset{\circ}{E}(\vec{u})$ центрированной системы нормальных величин $(\overset{\circ}{X}_1, \dots, \overset{\circ}{X}_n)$, получим

$$\overset{\circ}{E}(\vec{u}) = e^{i \sum_{j=1}^n \sum_{l=1}^n \beta_{jl} u_j u_l}. \quad (27.4)$$

Дифференцируя $E(\vec{u})$ дважды по u_j и u_l , в соответствии с формулой (26.35) для корреляционного момента k_{jl} , находим

$$\begin{aligned} k_{jl} &= - \frac{\partial^2 \overset{\circ}{E}(\vec{u})}{\partial u_j \partial u_l} \Big|_{\vec{u}=\vec{0}} = \\ &= - \frac{\partial^2}{\partial u_j \partial u_l} \left\{ \sum_{r=0}^{\infty} \frac{1}{r!} \left[\sum_{j,l=0}^n \beta_{jl} u_j u_l \right]^r \right\} \Big|_{\vec{u}=\vec{0}} = -2\beta_{jl}, \end{aligned} \quad (27.5)$$

где экспонента (27.4) представлена ее разложением в степенной ряд. Таким образом, $\beta_{jl} = -\frac{1}{2} k_{jl}$ и формулу (27.1) для характеристической функции системы нормальных случайных величин окончательно можно представить в виде

$$E(\vec{u}) = e^{i \sum_{j=1}^n \overline{x_j} u_j - 1/2 \sum_{j=1}^n \sum_{l=1}^n k_{jl} u_j u_l}. \quad (27.6)$$

Очевидно, что введенное определение многомерного нормального закона распределения не противоречит введенному ранее определению нормального закона распределения одной случайной величины, так как, положив в (27.6) все аргументы $u_j = 0$, кроме, скажем, аргумента u_r , в соответствии с (26.7) получим для характеристической функции $E_r(u_r)$ случайной величины X_r

$$E_r(u_r) = e^{i\bar{x}_r u_r - 1/2\sigma_r^2 u_r^2}, \quad (27.7)$$

что совпадает с определением характеристической функции нормального закона распределения отдельно взятой случайной величины, данным в главе 4.

Если все корреляционные моменты k_{jl} при $j \neq l$ обращаются в нуль, то формула (27.6) принимает вид

$$E(\vec{u}) = \prod_{j=1}^n e^{i\bar{x}_j u_j - 1/2\sigma_j^2 u_j^2},$$

и в этом случае характеристическая функция распадается на произведение характеристических функций случайных величин X_j , входящих в систему. Следовательно, если нормальные величины *некоррелированы*, то они и *независимы*.

Покажем, что от произвольной системы нормальных случайных величин путем линейного преобразования можно перейти к системе независимых нормальных величин, то есть покажем, что всегда можно определить матрицу преобразования $\mathbf{C} = \|c_{jl}\|_{j,l=1}^n$ так, чтобы случайные величины Y_j , определяемые формулами

$$Y_j = \sum_{l=1}^n c_{jl} x_l, \quad (j = \overline{1, n}), \quad (27.8)$$

или в векторной форме

$$\vec{Y} = \mathbf{C}\vec{X},$$

были бы взаимно независимыми и нормальными.

Перепишав показатель степени (27.6) в матричной форме и снабжая характеристическую функцию системы $(X_1, \dots, X_n)$

индексом x , имеем

$$E_x(\vec{u}) = e^{i\vec{u}^T \vec{m}_x - 1/2 \vec{u}^T \vec{K}_x \vec{u}}, \quad (27.9)$$

где $K_x = \|k_{jl}^{(x)}\|_{j,l=1}^n$ — корреляционная матрица системы $X_1, \dots, X_n$; $\vec{u} = \{u_i\}_{i=1}^n$ — векторный аргумент характеристической функции, а $\vec{m}_x$ — вектор, составленный из математических ожиданий $\bar{x}_j$: $\vec{m}_x = \{\bar{x}_j\}_{j=1}^n$.

В соответствии с формулой (26.11) характеристическая функция $E_y(\vec{u})$ системы $(Y_1, \dots, Y_n)$ равняется характеристической функции $E_x(\vec{u})$, в которой вектор $\vec{u}$ заменяется вектором $\vec{v}$, связанным с $\vec{u}$ линейным соотношением

$$\vec{v} = \mathbf{C}^T \vec{u}, \quad (27.10)$$

$$E_y(\vec{u}) = e^{i\vec{u}^T \mathbf{C} \vec{m}_x - 1/2 \vec{u}^T \mathbf{C} \mathbf{K}_x \mathbf{C}^T \vec{u}}. \quad (27.11)$$

Показатель степени в (27.11) — полином второй степени относительно аргументов характеристической функции $(u_1, \dots, u_n)$. Следовательно, система величин $(Y_1, \dots, Y_n)$ также нормальна. Запишем для этой системы характеристическую функцию в каноническом виде, согласно которому

$$E_y(\vec{u}) = e^{i\vec{u}^T \vec{m}_y - 1/2 \vec{u}^T \mathbf{K}_y \vec{u}}, \quad (27.12)$$

где вектор $\vec{m}_y$ — вектор математических ожиданий, а K_y — корреляционная матрица системы $(Y_1, \dots, Y_n)$. Сравнивая (27.12) с (27.11), получаем

$$\vec{m}_y = \mathbf{C} \vec{m}_x, \quad (27.13)$$

$$\mathbf{K}_y = \mathbf{C} \mathbf{K}_x \mathbf{C}^T. \quad (27.14)$$

По условию случайные величины Y_j ($j = \overline{1, n}$) независимы. Следовательно, корреляционная матрица $\mathbf{K}_y$ — диагональная и ее элементы могут быть представлены в виде

$$k_{jl}^{(y)} = \sigma_j^2 \delta_{jl}, \quad (j, l = \overline{1, n}), \quad (27.15)$$

где $\sigma_j^2 = \mathbf{D}(Y_j)$, а δ_{jl} обозначает дельта-символ Кронекера⁴.

⁴ Кронекер Леопольд (1823–1891) — немецкий математик.

Таким образом, задача свелась к приведению матрицы $\mathbf{K}_x$ к диагональной форме. Поскольку матрица $\mathbf{K}_x$ является положительно-определенной (см. формулу (2.49)), эта задача имеет однозначное решение [10]. При этом дисперсии $\sigma_j^2 = \lambda_j$, где λ_j — корни характеристического уравнения матрицы $\mathbf{K}_x$

$$\det(\mathbf{K}_x - \lambda \mathbf{I}) = 0, \quad (27.16)$$

где $\mathbf{I}$ — единичная матрица.

Пусть $\mathbf{C}$ — вещественная матрица, которую можно рассматривать как матрицу, соответствующую вращению декартовой системы координат в n -мерном пространстве. Матрица $\mathbf{C}$ обладает свойствами:

$$\mathbf{C}^T = \mathbf{C}^{-1}, \quad \det(\mathbf{C}) = 1.$$

Такие матрицы называют *ортогональными*. Для нахождения элементов этой матрицы, в соответствии с (27.14) и (27.15), нужно решить систему линейных алгебраических уравнений

$$\mathbf{C}\mathbf{K}_x = \mathbf{K}_y\mathbf{C}, \quad (27.17)$$

которая всегда имеет решение, поскольку матрица $\mathbf{K}_x$ — положительно-определенная.

Возвращаясь к формуле (27.12), замечаем, что с учетом (27.15) характеристическую функцию $E_y(\vec{u})$ можно записать в виде

$$E_y(\vec{u}) = e^{i \sum_{j=1}^n \bar{y}_j u_j - 1/2 \sum_{j=1}^n \sigma_j^2 u_j^2} = \prod_{j=1}^n e^{(i \bar{y}_j u_j - 1/2 \sigma_j^2 u_j^2)}, \quad (27.18)$$

как произведение характеристических функций одномерных нормальных законов распределения. Следовательно, система случайных величин $(Y_1, \dots, Y_n)$ является системой независимых нормальных величин, что и требовалось доказать.

Плотность вероятности этой системы является произведением плотностей распределения нормальных величин Y_j и может быть представлена в виде

$$\begin{aligned}
 f(\vec{y}) &= \prod_{j=1}^n \frac{1}{\sigma_j \sqrt{2\pi}} \cdot e^{-\frac{(y_j - \bar{y}_j)^2}{2\sigma_j^2}} = \\
 &= \frac{1}{(2\pi)^2 \sqrt{\det(\mathbf{K}_y)}} e^{-\frac{1}{2}(\vec{y} - \vec{m}_y)^T \mathbf{K}_y^{-1} (\vec{y} - \vec{m}_y)}, \quad (27.19)
 \end{aligned}$$

где учтено, что на основании (27.16)

$$\prod_{j=1}^n \sigma_j = \prod_{j=1}^n \sqrt{\lambda_j} = \sqrt{\det(\mathbf{K}_y)}, \quad \mathbf{K}_y^{-1} = \left\| \frac{1}{\sigma_j^2} \delta_{jl} \right\|_{j,l=1}^n. \quad (27.20)$$

Элементарная вероятность, соответствующая элементу объема $\mathbf{dy}$, равна элементарной вероятности, соответствующей элементу объема $\mathbf{dx}$ (вследствие однозначности преобразования и выполнения условия $\det(\mathbf{C}) = 1$), и, следовательно, должно соблюдаться равенство

$$f_x(\vec{x}) \mathbf{dx} = f_y(\vec{y}) \mathbf{dy} = \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} f_y(\vec{y}) \mathbf{dy}, \quad (27.21)$$

где координаты $\vec{x}$ и $\vec{y}$ связаны соотношением (27.8). Поскольку при преобразовании декартовой системы координат якобиан преобразования будет равен единице, то из (27.21) следует

$$f_x(\vec{x}) = f_y(\vec{y}), \quad (27.22)$$

где вектор $\vec{y}$ в правой части предполагается выраженным через компоненты вектора $\vec{x}$ в соответствии с (27.8).

Таким образом, для получения плотности вероятности исходной системы случайных величин $(X_1, \dots, X_n)$ достаточно в формуле (27.19) перейти от переменных y_j к переменным x_j , ($j = \overline{1, n}$). Так как на основании (27.8) и (27.13)

$$(\vec{y} - \vec{m}_y) = \mathbf{C}(\vec{x} - \vec{m}_x),$$

то, учитывая (27.14) квадратичную форму, стоящую в показателе степени в (27.19), можно представить в виде

$$\begin{aligned}
 (\vec{y} - \vec{m}_y)^T \mathbf{K}_y^{-1} (\vec{y} - \vec{m}_y) &= \\
 &= (\vec{x} - \vec{m}_x)^T \mathbf{C}^{-1} \mathbf{K}_y^{-1} \mathbf{C} (\vec{x} - \vec{m}_x) = \\
 &= (\vec{x} - \vec{m}_x)^T \mathbf{K}_x^{-1} (\vec{x} - \vec{m}_x), \quad (27.23)
 \end{aligned}$$

где учтено, что $\mathbf{C}^T = \mathbf{C}^{-1}$. Подставляя (27.23) в (27.19) и учитывая (27.22), получаем

$$f_x(\vec{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\mathbf{K}_x)}} \cdot e^{-1/2(\vec{x} - \vec{m}_x)^T \mathbf{K}_x^{-1} (\vec{x} - \vec{m}_x)} \quad (27.24)$$

или, учитывая явное выражение элементов $k_{jl}^{(-1)}$ матрицы $\mathbf{K}_x^{-1}$ через элементы корреляционной матрицы k_{jl} :

$$k_{jl}^{(-1)} = \frac{A_{jl}}{\Delta}, \quad (27.25)$$

где $\Delta = \det(\mathbf{K}_x)$, а A_{jl} — алгебраическое дополнение элемента k_{jl} определителя Δ , получим окончательно

$$f_x(\vec{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\Delta}} \cdot e^{-\frac{1}{2\Delta} \left(\sum_{j=1}^n \sum_{l=1}^n A_{jl} (x_j - \bar{x}_j) \cdot (x_l - \bar{x}_l) \right)}. \quad (27.26)$$

Формула (27.26), как и (27.24), показывает, что плотность вероятности n -мерного нормального закона распределения может быть представлена в виде экспоненты, содержащей в показателе полином второй степени от своих аргументов. Это свойство может быть взято как определение многомерного нормального закона распределения вместо определения (27.1), однако в этом случае для получения явного выражения плотности вероятности (27.26) требуются более сложные рассуждения.

Показателю степени в (27.26) можно придать другой вид, если ввести в рассмотрение определитель $(n+1)$ -го порядка $\Psi(\vec{x})$, положив

$$\Psi(\vec{x}) = \begin{vmatrix} k_{11} & k_{12} & \dots & k_{1n} & (x_1 - \bar{x}_1) \\ k_{21} & k_{22} & \dots & k_{2n} & (x_2 - \bar{x}_2) \\ \dots & \dots & \dots & \dots & \dots \\ k_{n1} & k_{n2} & \dots & k_{nn} & (x_n - \bar{x}_n) \\ (x_1 - \bar{x}_1) & (x_2 - \bar{x}_2) & \dots & (x_n - \bar{x}_n) & 0 \end{vmatrix}. \quad (27.27)$$

Разложив определитель $\Psi(\vec{x})$ по элементам последнего столбца, а затем каждое из получающихся при этом алгебраических дополнений — по элементам последней строки, получим, что

$$\Psi(\vec{x}) = - \sum_{j=1}^n \sum_{l=1}^n A_{jl} (x_j - \bar{x}_j) (x_l - \bar{x}_l), \quad (27.28)$$

где знак минус появляется потому, что при первом разложении вычисляется определитель $(n+1)$ -го порядка, а при втором разложении — определитель n -го порядка, следовательно, четность мест, на которых стоят элементы $(x_j - \bar{x}_j)$ с одинаковым значением индекса j , будет обратная. Учитывая (27.28) вместо (27.26), плотность n -мерного нормального закона распределения можно представить в виде

$$f(\vec{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\Delta}} \cdot e^{\frac{1}{2\Delta} \cdot \Psi(\vec{x})}. \quad (27.29)$$

В некоторых случаях удобно, чтобы в выражение плотности многомерного нормального закона входили только безразмерные величины. Этого можно добиться, перейдя от случайных величин X_j к нормированным и центрированным случайным величинам

$$\xi_j = \frac{X_j - \bar{x}_j}{\sigma_j}$$

и рассматривая вместо матрицы **K** корреляционных моментов матрицу **R**, составленную из безразмерных коэффициентов корреляции

$$r_{jl} = \frac{k_{jl}}{\sigma_j \sigma_l}. \quad (27.30)$$

Вынося при вычислении определителя Δ из каждого j -го столбца и из каждой j -й строки матрицы **K** член σ_j , ($j = \overline{1, n}$), получаем

$$\Delta = \det(\mathbf{K}) = \prod_{j=1}^n \sigma_j^2 \det(\mathbf{R}). \quad (27.31)$$

Далее, подставляя (27.31) в (27.27), деля строки и столбцы определителя $\Psi(\vec{x})$ на σ_j , приравнивая элементы вероятности $dP(\vec{\xi}) = dP(\vec{x})$ и учитывая при этом, что $d\mathbf{x} = \prod_{j=1}^n \sigma_j d\xi$, получаем

$$f(\vec{\xi}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\mathbf{R})}} \cdot e^{\frac{\psi(\vec{\xi})}{2 \det(\mathbf{R})}}, \quad (27.32)$$

где ψ представляет собой промасштабированный определитель Ψ

$$\psi(\vec{\xi}) = \begin{vmatrix} r_{11} & r_{12} & \dots & r_{1n} & \xi_1 \\ r_{21} & r_{22} & \dots & r_{2n} & \xi_2 \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nn} & \xi_n \\ \xi_1 & \xi_2 & \dots & \xi_n & 0 \end{vmatrix}. \quad (27.33)$$

Выше уже было отмечено, что закон распределения любой подсистемы нормальных случайных величин нормальный. Покажем, что и условные законы распределения подсистемы нормальных величин, вычисленные при фиксированных значениях остальных величин, также являются нормальными.

Рассмотрим систему нормальных случайных величин $(X_1, \dots, X_n)$. Предположим, что $(n - \nu)$ последних величин этой системы фиксированы. Тогда плотность вероятности системы $f(\vec{x})$ может быть выражена в виде произведения условной плотности подсистемы первых ν величин на плотность вероятности подсистемы фиксированных величин, то есть в виде

$$f(\vec{x}) = f(x_1, \dots, x_\nu | x_{\nu+1}, \dots, x_n) f(x_{\nu+1}, \dots, x_n). \quad (27.34)$$

Так как плотность $f(x_1, \dots, x_\nu | x_{\nu+1}, \dots, x_n)$, определяемая (27.34), может быть представлена в виде экспоненты, содержащей в показателе степени полином второй степени относительно переменных $x_1, \dots, x_\nu$, то рассматриваемая условная плотность вероятности является плотностью некоторого нормального закона распределения. Следовательно, и условная характеристическая функция $E(u_1, \dots, u_\nu | x_{\nu+1}, \dots, x_n)$ также

может быть представлена в виде

$$E(u_1, \dots, u_\nu | x_{\nu+1}, \dots, x_n) = \exp \left\{ i \sum_{r=1}^{\nu} \bar{x}_r^* u_r - \frac{1}{2} \sum_{r=1}^{\nu} \sum_{l=1}^{\nu} k_{rl}^* u_r u_l \right\}, \quad (27.35)$$

где условные математические ожидания $\bar{x}_r^*$ и условные корреляционные моменты k_{rl}^* , $(j, l = \overline{1, \nu})$ подлежат определению. Для их нахождения подставим в формулу (26.3) для характеристической функции $E(\vec{u})$ системы вместо плотности вероятности $f(\vec{x})$ ее выражение из (27.34), что дает

$$E(\vec{u}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\vec{u}^T \vec{x}} f(x_1, \dots, x_\nu | x_{\nu+1}, \dots, x_n) \times \\ \times f(x_{\nu+1}, \dots, x_n) d\mathbf{x}. \quad (27.36)$$

Выполнив интегрирование по $x_1, \dots, x_\nu$, последнему равенству можно придать вид $(n - \nu)$ -кратного интегрального преобразования Фурье:

$$E(\vec{u}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \cdot \sum_{j=\nu+1}^n x_j u_j} \times \\ \times E(u_1, \dots, u_\nu | x_{\nu+1}, \dots, x_n) \times \\ \times f(x_{\nu+1}, \dots, x_n) dx_{\nu+1} dx_{\nu+2} \dots dx_n. \quad (27.37)$$

Применяя $(n - \nu)$ раз обратное преобразование Фурье и заменяя характеристическую функцию системы n нормальных величин $E(\vec{u})$ ее выражением (27.6), получим

$$E(u_1, \dots, u_\nu | x_{\nu+1}, \dots, x_n) \cdot f(x_{\nu+1}, \dots, x_n) = \\ = \frac{1}{(2\pi)^{n-\nu}} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \cdot \sum_{\alpha=\nu+1}^n x_\alpha u_\alpha + i \vec{u}^T \vec{m}_x - 1/2 \vec{u}^T \mathbf{K}_x \vec{u}} \times$$

$$\begin{aligned}
& \times dx_{\nu+1} dx_{\nu+2} \dots dx_n = e^{i \cdot \sum_{j=1}^{\nu} \bar{x}_j u_j - 1/2 \sum_{j=1}^{\nu} \sum_{l=1}^{\nu} k_{jl} u_j u_l} \cdot \frac{1}{(2\pi)^{n-\nu}} \times \\
& \times \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -i \cdot \sum_{\alpha=\nu+1}^n x_{\alpha} u_{\alpha} + i \cdot \sum_{\alpha=\nu+1}^n (\bar{x}_{\alpha} + \sum_{r=1}^{\nu} k_{\alpha r} u_r) u_{\alpha} - \right. \\
& \quad \left. - \frac{1}{2} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n k_{\alpha\beta} u_{\beta} u_{\alpha} \right\} dx_{\nu+1} \dots dx_n. \quad (27.38)
\end{aligned}$$

Плотность вероятности $f(x_{\nu+1}, \dots, x_n)$ в соответствии с (27.26) определяется формулой

$$\begin{aligned}
f_x(\vec{x}) &= \frac{1}{(2\pi)^{(n-\nu)/2} \sqrt{\Delta_1}} \times \\
& \times e^{-\frac{1}{2\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} (x_{\alpha} - \bar{x}_{\alpha}) \cdot (x_{\beta} - \bar{x}_{\beta})}, \quad (27.39)
\end{aligned}$$

где Δ_1 — определитель корреляционной матрицы рассматриваемой подсистемы случайных величин

$$\Delta_1 = \left\| \begin{array}{cccc} k_{\nu+1, \nu+1} & k_{\nu+1, \nu+2} & \dots & k_{\nu+1, n} \\ \dots & \dots & \dots & \dots \\ k_{n, \nu+1} & k_{n, \nu+2} & \dots & k_{n, n} \end{array} \right\|, \quad (27.40)$$

а $a_{\alpha\beta}$ — алгебраическое дополнение элементов этого определителя, соответствующее элементу $k_{\alpha\beta}$, ($\alpha, \beta = \overline{\nu+1, n}$).

Последний интеграл в (27.38) отличается от обратного преобразования характеристической функции подсистемы $(X_{\nu+1}, \dots, X_n)$ только тем, что в выражении характеристической функции математические ожидания $\bar{x}_{\alpha}$ заменены на $(\bar{x}_{\alpha} + i \sum_{r=1}^{\nu} k_{\alpha r} u_r)$. Следовательно, произведя эту замену в

формуле (27.39), получим

$$\begin{aligned}
 & \frac{1}{(2\pi)^{n-\nu}} \cdot \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\left\{-i \cdot \sum_{\alpha=\nu+1}^n x_{\alpha} u_{\alpha} + \right. \\
 & \quad \left. + i \cdot \sum_{\alpha=\nu+1}^n (\bar{x}_{\alpha} + i \sum_{r=1}^{\nu} k_{\alpha r} u_r) u_{\alpha} - \right. \\
 & \quad \left. - \frac{1}{2} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n k_{\alpha\beta} u_{\beta} u_{\alpha}\right\} dx_{\nu+1} \dots dx_n = \frac{1}{(2\pi)^{(n-\nu)/2}} \times \\
 & \times \exp\left\{-\frac{1}{2\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} (x_{\alpha} - \bar{x}_{\alpha} - i \sum_{r=1}^{\nu} k_{\alpha r} u_r) \times \right. \\
 & \quad \left. \times (x_{\beta} - \bar{x}_{\beta} - i \sum_{s=1}^{\nu} k_{\beta s} u_s)\right\}. \quad (27.41)
 \end{aligned}$$

Подставляя (27.39) и (27.41) в (27.38) и определяя явно условную характеристическую функцию, после сокращения на общие множители и очевидных преобразований имеем

$$\begin{aligned}
 & E(u_1, \dots, u_{\nu} | x_{\nu+1}, \dots, x_n) \cdot f(x_{\nu+1}, \dots, x_n) = \\
 & = \exp\left\{-\frac{1}{2\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} [(x_{\alpha} - \bar{x}_{\alpha} - i \sum_{r=1}^{\nu} k_{\alpha r} u_r) \times \right. \\
 & \quad \times (x_{\beta} - \bar{x}_{\beta} - i \sum_{s=1}^{\nu} k_{\beta s} u_s) - (x_{\alpha} - \bar{x}_{\alpha})(x_{\beta} - \bar{x}_{\beta})] + \\
 & \quad \left. + i \sum_{j=1}^{\nu} \bar{x}_j u_j - \frac{1}{2} \sum_{j=1}^{\nu} \sum_{l=1}^{\nu} k_{jl} u_j u_l\right\}. \quad (27.42)
 \end{aligned}$$

Группируя в выражении, стоящем в квадратных скобках, слагаемые по степеням переменных u_{α} , получаем

$$\begin{aligned}
& E(u_1, \dots, u_\nu | x_{\nu+1}, \dots, x_n) \cdot f(x_{\nu+1}, \dots, x_n) = \\
& = \exp \left\{ i \sum_{r=1}^{\nu} \left[\bar{x}_r + \frac{1}{\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} k_{\alpha r} (x_\beta - \bar{x}_\beta) \right] u_r - \right. \\
& \left. - \frac{1}{2} \sum_{r=1}^{\nu} \sum_{s=1}^{\nu} \left[k_{rs} - \frac{1}{\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} k_{\alpha r} k_{\beta s} \right] u_r u_s \right\}. \quad (27.43)
\end{aligned}$$

Сравнивая правую часть (27.43) с (27.35), заключаем, что

$$\begin{cases} \bar{x}_r^* = \bar{x}_r + \frac{1}{\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} k_{\alpha r} (x_\beta - \bar{x}_\beta), \\ k_{rs}^* = k_{rs} - \frac{1}{\Delta_1} \sum_{\alpha=\nu+1}^n \sum_{\beta=\nu+1}^n a_{\alpha\beta} k_{\alpha r} k_{\beta s}. \end{cases} \quad (27.44)$$

Полученные формулы, в частности, показывают, что условные математические ожидания подсистемы нормальных случайных величин линейно зависят от фиксированных величин системы, а условная корреляционная матрица от этих величин вообще не зависит.

Рассмотрим подробнее случай $n = 2$, обозначив $X_1 = X$, $X_2 = Y$, $k_{12} = k_{xy} = k$, $\frac{k}{\sigma_x \sigma_y} = r$. Для корреляционной матрицы рассматриваемой системы (X, Y) получим выражение

$$\mathbf{K} = \left\| \begin{array}{cc} \sigma_x^2 & k \\ k & \sigma_y^2 \end{array} \right\|. \quad (27.45)$$

Определитель $\Delta = \det(\mathbf{K})$ и его алгебраические дополнения A_{jl} будут следующими:

$$\Delta = \sigma_x^2 \sigma_y^2 - k^2 = \sigma_x^2 \sigma_y^2 (1 - r^2),$$

$$A_{11} = \sigma_y^2, \quad A_{22} = \sigma_x^2,$$

$$A_{12} = A_{21} = -k = -\sigma_x \sigma_y r.$$

Следовательно, согласно формуле (27.26),

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \cdot \exp \left\{ -\frac{1}{2(1-r^2)} \times \right. \\ \left. \times \left[\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - 2r \frac{(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} \right] \right\}. \quad (27.46)$$

Плотность вероятности $f(x, y)$ имеет постоянное значение при постоянном значении показателя степени в (27.46), то есть на кривых вида

$$\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - 2r \frac{(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} = c^2 = \text{const.} \quad (27.47)$$

Если (x, y) рассматривать как координаты случайной точки на плоскости, то равенство (27.47) определяет некоторую кривую второго порядка.

Будем считать случайные величины X и Y линейно независимыми. В этом случае $r < 1$ и равенство (27.47) определяет семейство подобных эллипсов с центрами, расположенными в точке $(\bar{x}, \bar{y})$. При $c^2 = 1$ соответствующий эллипс называется единичным эллипсом (рис. 27.1). Главные полуоси этого эллипса a_1, a_2 направлены под углами α_1, α_2 к оси Ox , определяемыми равенством

$$\operatorname{tg} 2\alpha_{1,2} = \frac{2k}{\sigma_x^2 - \sigma_y^2}, \quad (27.48)$$

и имеют значения

$$a_{1,2} = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sin^2 \alpha_{1,2} + \sigma_y^2 \cos^2 \alpha_{1,2} - k \sin 2\alpha_{1,2}} = \\ = \frac{1}{2} \left[\sigma_x^2 + \sigma_y^2 \pm \sqrt{\sigma_x^2 - \sigma_y^2 + 4k^2} \right]. \quad (27.49)$$

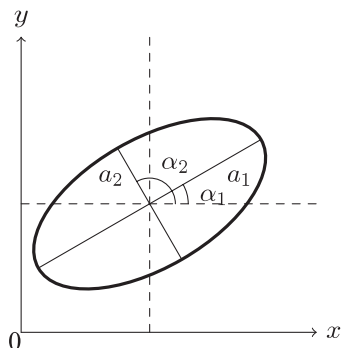


Рис. 27.1
Эллипс рассеяния

В соответствии с формулой (27.8) от системы зависимых величин (X, Y) можно перейти к системе независимых величин (X_1, Y_1) . Легко убедиться, что величины (X_1, Y_1) имеют геометрический смысл координат случайной точки $M(X, Y)$ в координатной системе Ox_1y_1 , повернутой относительно координатной системы Oxy на угол α_1 , задаваемый (27.49). Применяя формулу (27.44) для системы двух случайных величин (X, Y) , получаем, что условная плотность вероятности $f(x|y)$ является плотностью нормального закона распределения с математическим ожиданием

$$\bar{x}_y = \bar{x} + r \cdot \frac{\sigma_x}{\sigma_y} (y - \bar{y}) \quad (27.50)$$

и дисперсией

$$\sigma_{x|y}^2 = \sigma_x^2 (1 - r^2). \quad (27.51)$$

Зависимость условного математического ожидания $\bar{x}_y$ от значения y называется *линией регрессии*. Как показывает формула (27.50), в случае нормальной системы величин линия регрессии — прямая линия, причем условное математическое ожидание $\bar{x}_y$ совпадает с безусловным $\bar{x}$, если фиксированное значение другой случайной величины равно его математическому ожиданию $\bar{y}$. В отличие от условного математического

ожидания условная дисперсия $\sigma_{x|y}^2$ не зависит от значения y и всегда в $(1-r^2)$ раз меньше безусловной дисперсии. Таким образом, давая аргументу y ряд значений, получаем геометрическое место точек центров условного закона распределения на линии регрессии, причем условная плотность вероятности для любой точки будет иметь одинаковую форму, но ее максимум будет сдвинут относительно $x = \bar{x}$ на величину, зависящую от r, σ_x и σ_y .

Аналогично, применяя общую формулу (27.26), для плотности вероятности нормальной системы трех случайных величин (X, Y, Z) получаем

$$f(x, y, z) = \frac{1}{(2\pi)^{3/2} \sqrt{\Delta}} \times \\ \times \exp \left\{ -\frac{1}{2d} [(1-r_{yz})^2 \xi^2 + (1-r_{xy})^2 \zeta^2 - 2(r_{xy} - r_{xz}r_{yz}) \xi \eta - \right. \\ \left. - 2(r_{xy}r_{yz}) \xi \zeta - 2(r_{yz} - r_{xy}r_{xz}) \eta \zeta] \right\}, \quad (27.52)$$

где в данном случае определитель корреляционной матрицы

$$\begin{cases} \Delta = \sigma_x^2 \sigma_y^2 \sigma_z^2 \cdot d, d = (1 + 2r_{xy}r_{yz}r_{zx} - r_{xy}^2 - r_{yz}^2 - r_{zx}^2), \\ \xi = \frac{x - \bar{x}}{\sigma_x}, \quad \eta = \frac{y - \bar{y}}{\sigma_y}, \quad \zeta = \frac{z - \bar{z}}{\sigma_z}. \end{cases} \quad (27.53)$$

Как следует из общей формулы (27.43), двумерные условные законы распределения системы любых двух случайных величин, входящих в систему (X, Y, Z) при фиксированном значении третьей величины, так же как и закон распределения любой величины системы при фиксированных значениях двух других, являются нормальными.

Для условий плотности вероятности $f(x, y|z)$ параметры закона распределения определяются формулами:

$$\begin{cases} \bar{x}_z = \bar{x} + r_{xz} \frac{\sigma_x}{\sigma_z} (z - \bar{z}), \quad \bar{y}_z = \bar{y} + r_{zy} \frac{\sigma_y}{\sigma_z} (z - \bar{z}), \\ \sigma_{x|z} = \sigma_x \sqrt{1 - r_{xz}^2}, \quad \sigma_{y|z} = \sigma_y \sqrt{1 - r_{zy}^2}, \\ r_{xy|z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}, \end{cases} \quad (27.54)$$

применяя к которым круговую перестановку индексов x, y, z , можно получить параметры остальных двух условных законов распределения $f(y, z|x)$ и $f(z, x|y)$.

Для условной плотности вероятности $f(z|x, y)$ условное математическое ожидание и условная дисперсия определяются формулами:

$$\begin{aligned} \bar{z}_{xy} &= \bar{z}_x + r_{yz|x} \frac{\sigma_{z|x}}{\sigma_{y|x}} (y - \bar{y}_x) = \\ &= \bar{z} + \frac{\sigma_z}{(1 + r_{xy}^2)} \left[(r_{xz} - r_{xy}r_{yz}) \cdot \frac{(x - \bar{x})}{\sigma_x} + \right. \\ &\quad \left. + (r_{yz} - r_{xy}r_{xz}) \cdot \frac{(y - \bar{y})}{\sigma_y} \right], \end{aligned} \quad (27.55)$$

$$\begin{aligned} \sigma_{z|xy} &= \sigma_{z|x} \sqrt{1 - r_{yz|x}^2} = \\ &= \frac{\sigma_z}{\sqrt{1 - r_{xy}^2}} \sqrt{1 + 2r_{xy}r_{yz}r_{zx} - r_{xy}^2 - r_{yz}^2 - r_{xz}^2} = \\ &= \frac{\sigma_z \sqrt{d}}{\sqrt{1 - r_{xy}^2}}. \end{aligned} \quad (27.56)$$

Параметры условных плотностей вероятности $f(x|y, z)$ и $f(y|x, z)$ могут быть получены по формулам (27.55) и (27.56) круговой перестановкой индексов x, y, z .

Рассмотрим два примера, касающихся многомерного нормального закона распределения.

Пример 27.1. Дана нормальная система случайных величин (X, Y) , характеризуемая параметрами $\bar{x}$, $\bar{y}$, σ_x^2 , σ_y^2 , $r = k_{xy}(\sigma_x \sigma_y)^{-1}$. Определить плотность вероятности случайной величины X после испытания, показавшего, что $(a < Y < b)$ (событие A).

Решение. Используя результат решения примера 25.5 и формулу (27.46) для двумерного нормального закона распределения, получаем

$$f(x|A) = \frac{\int_a^b f(x, y) dy}{\int_a^b \int_{-\infty}^{\infty} f(x, y) dx dy} = \frac{f(x) \int_a^b f(y|x) dy}{\int_a^b f(y) dy},$$

где $f(x)$ и $f(y)$ — плотности нормальных законов распределения с параметрами $(\bar{x}, \sigma_x^2)$ и $(\bar{y}, \sigma_y^2)$, соответственно, а $f(y|x)$ — плотность нормального закона распределения с параметрами $(\bar{y}_x, \sigma_{y|x}^2)$, определяемыми формулами, аналогичными (27.50) и (27.51). Выполнив интегрирования, находим

$$f(x|A) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left[-\frac{(x - \bar{x})^2}{2\sigma_x^2} \right] \times \frac{\left[\Phi \left(\frac{b - \bar{y}_x}{\sigma_{y|x}} \right) - \Phi \left(\frac{a - \bar{y}_x}{\sigma_{y|x}} \right) \right]}{\left[\Phi \left(\frac{b - \bar{y}}{\sigma_y} \right) - \Phi \left(\frac{a - \bar{y}}{\sigma_y} \right) \right]},$$

где $\bar{y}_x = \bar{y} + r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$; $\sigma_{y|x}^2 = (1 - r^2) \sigma_y^2$.

Таким образом, условная плотность вероятности нормальной случайной величины X в данном случае, конечно, уже не является нормальной.

Пример 27.2. Декартовы координаты X, Y, Z (X_1, X_2, X_3) случайной точки в пространстве — центрированные нормальные величины, корреляционная матрица $\mathbf{K}$ которых известна. Определить направление осей декартовой системы координат

$Ox'y'z'$ ($Ox'_1x'_2x'_3$), в которой координаты случайной точки будут независимыми величинами, и найти дисперсии координат случайной точки в новой системе координат.

Решение. В соответствии с (27.16) дисперсии новых координат случайной точки равны корням уравнения

$$\det(\mathbf{K} - \lambda \mathbf{I}) = 0,$$

которое в данном случае является кубическим уравнением

$$\lambda^3 - a\lambda^2 + b\lambda - \Delta = 0,$$

где

$$\begin{aligned} \Delta &= \det(\mathbf{K}); & a &= \sigma_x^2 + \sigma_y^2 + \sigma_z^2; \\ b &= \sigma_y^2 \sigma_z^2 (1 - r_{yz}^2) + \sigma_z^2 \sigma_x^2 (1 - r_{xz}^2) + \sigma_x^2 \sigma_y^2 (1 - r_{xy}^2). \end{aligned}$$

Произведя замену

$$\lambda = \mu + \frac{1}{3} a,$$

кубическое уравнение для λ можно преобразовать к виду

$$\mu^3 + p\mu + q = 0,$$

где

$$p = b - \frac{1}{3} a^3; \quad q = -\Delta + \frac{1}{3} ab - \frac{2}{27} a^3.$$

Уравнение для μ имеет корни [16]

$$\begin{aligned} \mu_1 &= -2\kappa \cos\left(\frac{\varphi}{3}\right), & \mu_2 &= 2\kappa \cos\left(\frac{\pi - \varphi}{3}\right), \\ \mu_3 &= 2\kappa \cos\left(\frac{\pi + \varphi}{3}\right), \end{aligned}$$

где $\kappa = \sqrt{\frac{1}{3} |p| \sin q}; \quad \cos \varphi = \frac{q}{2a^3}.$

Для определения направления осей новой координатной системы⁵ необходимо решить систему уравнений (27.17), которая в данном случае дает систему трех уравнений для нахождения направляющих косинусов $\alpha_j, \beta_j, \gamma_j$ оси x'_j относительно осей x_1, x_2, x_3 , соответственно ($j = \overline{1, 3}$):

$$\begin{cases} (\sigma_x^2 - \lambda_j)\alpha_j + k_{xy}\beta_j + k_{xz}\gamma_j = 0, \\ k_{xy}\alpha_j + (\sigma_y^2 - \lambda_j)\beta_j + k_{yz}\gamma_j = 0, \\ k_{xz}\alpha_j + k_{yz}\beta_j + (\sigma_z^2 - \lambda_j)\gamma_j = 0. \end{cases}$$

Определитель полученной системы равняется нулю и, следовательно, одно из уравнений системы является следствием двух других. Дополняя систему равенством

$$\alpha^2 + \beta^2 + \gamma^2 = 1,$$

следующим из теоремы косинусов, получаем систему трех независимых уравнений, определяющих $\alpha_j, \beta_j, \gamma_j$ с точностью до знака (неопределенность знака соответствует произволу в выборе направлений осей координат).

⁵ Если от λ перейти к переменной ν , положив $\lambda = -\frac{q}{p}\nu + \frac{1}{3}$, то для определения ν получим уравнение, содержащее один числовой параметр $A = p^3/q^2$; $\nu^3 + A\nu - A = 0$. Для корней этого уравнения составлены специальные таблицы [30].

ФУНКЦИОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 28. ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

28.1. Общие положения

Рассмотрим следующую величину Y , связанную с системой случайных величин $(X_1, X_2, \dots, X_n)$ однозначной функциональной зависимостью

$$Y = \varphi(X_1, \dots, X_n) = \varphi(\vec{X}). \quad (28.1)$$

Выразим вероятностные характеристики Y через вероятностные характеристики $\vec{X}$.

Воспользовавшись определением характеристической функции как математического ожидания экспоненты e^{iuY} , имеем

$$E_y(u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{iu\varphi(\vec{x})} \cdot f(\vec{x}) d\mathbf{x}. \quad (28.2)$$

Поскольку характеристическая функция является исчерпывающей характеристикой случайной величины, формула (28.2) полностью определяет закон распределения случайной величины Y , в частности, позволяет вычислить ее плотность вероятности $f_y(y)$ и функцию распределения $F_y(y)$. Однако во многих задачах целесообразно располагать формулами, непосредственно выражающими плотность вероятности $f_y(y)$ через плотность вероятности $f_x(\vec{x})$ системы случайных компонент вектора $\vec{X}$, являющихся аргументами рассматриваемой

функциональной зависимости. Эти формулы можно получить или путем подстановки в формулу

$$f_y(y) = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} E_y(u) e^{-iuy} du \quad (28.3)$$

характеристической функции $E_y(u)$ из (28.2), или рассматривая непосредственно связь плотностей вероятности $f_y(y)$ и $f_x(\vec{x})$.

Рассмотрим вначале применение первого из упомянутых способов в общем случае функциональной зависимости. Подставляя (28.2) в (28.3) и меняя порядок интегрирования, получаем

$$f_y(y) = \int_{-\infty}^{\infty} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu(\varphi(\vec{x})-y)} du \right\} f_x(\vec{x}) d\mathbf{x}. \quad (28.4)$$

Интеграл, стоящий в фигурных скобках, равняется дельта-функции от аргумента $[\varphi(\vec{x}) - y]$ (см. главу 4):

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu(\varphi(\vec{x})-y)} du = \delta(\varphi(\vec{x}) - y), \quad (28.5)$$

следовательно, вместо (28.4), будем иметь

$$f_y(y) = \int_{-\infty}^{\infty} \delta[\varphi(\vec{x}) - y] f_x(\vec{x}) d\mathbf{x}. \quad (28.6)$$

Наличие дельта-функции под знаком интеграла позволяет выполнить одно интегрирование. Предположим, что относительно одного из своих аргументов x_j функция $\varphi(\vec{x})$ монотонна. Примем для определенности, что таким аргументом является x_1 , причем будем считать, что производная $\frac{\partial \varphi}{\partial x_1}$

существует. В этом случае вместо переменной интегрирования x_1 можно ввести новую переменную интегрирования ξ_1 , положив

$$\xi_1 = \varphi(\vec{x}) - y, \quad d\xi_1 = \frac{\partial \varphi(\vec{x})}{\partial x_1} dx_1. \quad (28.7)$$

Тогда интеграл (28.6) примет вид

$$f_y(y) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \delta(\xi_1) f_x(\vec{x}) \left| \frac{\partial \varphi(\vec{x})}{\partial x_1} \right|^{-1} d\xi_1, \dots, d\xi_n, \quad (28.8)$$

где x_1 , входящий в качестве аргумента функции $f_x(\vec{x})$ и $\frac{\partial \varphi(\vec{x})}{\partial x_1}$, предполагается выраженным через ξ_1 в соответствии с (28.7). Интегрирование по ξ_1 ведется в пределах $(-\infty, +\infty)$, так как, вследствие наличия множителя $\delta(\xi_1)$ под знаком интеграла, отличный от нуля результат дает только интегрирование по области вблизи значения $\xi_1 = 0$, а знак модуля у производной $\frac{\partial \varphi}{\partial x_1}$ отражает тот факт, что если $\frac{\partial \varphi}{\partial x_1} > 0$, то при возрастании x_1 значение ξ_1 переходит через нуль увеличиваясь, а при $\frac{\partial \varphi}{\partial x_1} < 0$ — уменьшаясь.

Воспользовавшись свойствами дельта-функции [6], интеграл по ξ_1 в (28.8) можно взять, и мы получим

$$f_y(y) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{\left| \frac{\partial \varphi}{\partial x_1} \right|} f_x(\vec{x}) dx_1 \dots dx_n, \quad (28.9)$$

где x_1 соответствует значению $\xi_1 = 0$, то есть определяется путем решения уравнения

$$y = \varphi(\vec{x}) \quad (28.10)$$

относительно x_1 .

В том случае, когда из аргументов функции $\varphi(\vec{x})$ нет ни одного, от которого функция $\varphi(\vec{x})$ зависела бы монотонно,

область изменения одного из аргументов, например, x_1 можно разбить на интервалы, в которых рассматриваемая функция монотонна; в этом случае в окончательной формуле для плотности вероятности $f_y(y)$ появится сумма интегралов типа (28.9), каждый из которых соответствует одному интервалу монотонности.

Рассмотрим применение общей формулы (28.9) к ряду частных случаев.

28.2. Функция одной случайной величины

Если

$$Y = \varphi(X), \quad (28.11)$$

где $\varphi(X)$ — монотонная функция, то формула (28.9) дает сразу

$$f_y(y) = \frac{1}{\left| \frac{d\varphi(x)}{dx} \right|} f_x(x) = \left| \frac{d\psi(y)}{dy} \right| f_x[\psi(y)], \quad (28.12)$$

где через $\psi(y)$ обозначена функция, обратная по отношению к $\varphi(x)$, так что

$$X = \Psi(Y). \quad (28.13)$$

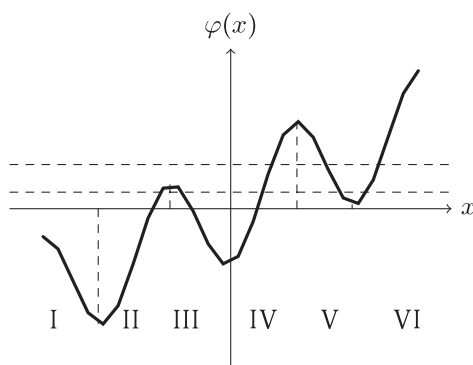


Рис. 28.1
Случай монотонной функции

Если функция $\varphi(x)$ не является монотонной, то, разбив область изменения x на интервалы, в каждом из которых функция $\varphi(x)$ монотонна (рис. 28.1), и обозначая для j -го интервала функцию, обратную $\varphi(x)$, через $\psi_j(y)$, получаем

$$f_y(y) = \sum_j \left| \frac{d\psi_j(y)}{dy} \right| f_x[\psi_j(y)], \quad (28.14)$$

где число слагаемых в сумме

равно числу интервалов монотонности, пересекаемых прямой $y = \text{const}$ (на рис. 28.1 изображен случай, когда пересекаются 4 интервала — III, IV, V, VI).

Формулы (28.12) и (28.14) могут быть получены путем простых геометрических рассуждений. Рассмотрим для примера монотонно убывающую функцию (рис. 28.2). Обратная функция $\psi(y)$ в этом случае также будет монотонно убывающей. Рассмотрим бесконечно малый интервал $(y, y + dy)$, вероятность попадания в который случайной величины Y равна $f_y(y)dy$. Однако при попадании Y в указанный интервал случайная величина X попадает в интервал $(\psi(y) + \frac{d\psi}{dy} dy, \psi(y))$ (рис. 28.2). Вероятность этого события равна $-f_x[\psi(y)] \frac{d\psi}{dy} dy$.

Таким образом, мы получили два выражения для вероятности одного и того же события. Приравнявая их друг к другу и сокращая на dy , получаем

$$f_y(y) = -\frac{d\psi}{dy} f_x[\psi(y)] = \left| \frac{d\psi}{dy} \right| f_x[\psi(y)], \quad (28.15)$$

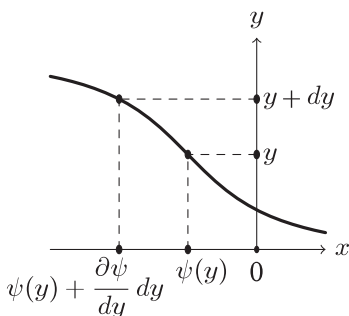


Рис. 28.2

Случай монотонной функции

что совпадает с формулой (28.12). Аналогично могут быть рассмотрены и случай монотонно возрастающей функции $\varphi(x)$, и общий случай, соответствующий формуле (28.14), представленный на рис. 28.1.

28.3. Сумма двух случайных величин

Если

$$Z = X + Y, \quad (28.16)$$

то характеристическая функция $E_z(u)$ выражается через характеристическую функцию системы (X, Y) следующим образом:

$$E_z(u) = \mathbf{M}[e^{iu_1 X + iu_2 Y}]|_{u_1=u_2=u} = E(u, u). \quad (28.17)$$

Поскольку Z является однозначной функцией обоих своих аргументов, причем в обозначениях формулы (28.1)

$$\frac{\partial \varphi}{\partial x_1} = \frac{\partial \varphi}{\partial x} = \frac{\partial \varphi}{\partial y} = 1,$$

то применение формулы (28.9) позволяет сразу выразить плотность вероятности $f_z(z)$ в двух эквивалентных формах:

$$f_z(z) = \int_{-\infty}^{\infty} f(x, z-x) dx = \int_{-\infty}^{\infty} f(z-y, y) dy. \quad (28.18)$$

Действительно, исключаемая переменная интегрирования должна быть выражена через функцию z и оставшуюся переменную из равенства (28.16). Значит, в первом случае y должно быть заменено на $(z-x)$, а во втором случае x — на $(z-y)$.

В том случае, когда случайные величины X и Y независимые, $f(x, y) = f_x(x)f_y(y)$ и формула (28.18) принимает вид

$$f_z(z) = \int_{-\infty}^{\infty} f(x)f_y(z-x) dx = \int_{-\infty}^{\infty} f(z-y)f_y(y) dy. \quad (28.19)$$

Нахождение закона распределения суммы независимых случайных величин называется *композицией* (сложением) законов распределения и условно обозначается так:

$$f_z(z) = f_x(z) * f_y(z). \quad (28.20)$$

Эта операция называется *сверткой* функций f_x и f_y .

Для суммы независимых случайных величин характеристическая функция равна произведению характеристических функций слагаемых

$$E_z(u) = E_x(u)E_y(u). \quad (28.21)$$

Действительно, по определению характеристической функции имеем

$$\begin{aligned} E_z(u) &= \mathbf{M}[e^{iuZ}] = \mathbf{M}[e^{iu(X+Y)}] = \\ &= \mathbf{M}[e^{iuX}] \mathbf{M}[e^{iuY}] = E_x(u)E_y(u), \end{aligned} \quad (28.22)$$

где переход от математического ожидания произведения экспонент к произведению математических ожиданий возможен на основании последней теоремы о математическом ожидании произведения независимых случайных величин (26.36). (Формула (28.21), естественно, может быть получена и из формулы (28.19), если к обеим ее частям применить преобразование Фурье.)

Формула (28.19) также может быть получена путем наглядных геометрических рассуждений. Будем рассматривать X и Y как координаты случайной точки на плоскости. Найдем вероятность попадания этой точки в область, определяемую неравенствами

$$z \leq Z \leq z + dz. \quad (28.23)$$

Легко видеть, что этой областью является бесконечно узкая полоса, ограниченная двумя прямыми: $x + y = z$ и $x + y = z + dz$. Следовательно, искомая вероятность может быть определена или как вероятность попадания случайной величины Z в интервал (28.23), или вероятность попадания случайной точки в бесконечно узкую полосу. В первом случае эта

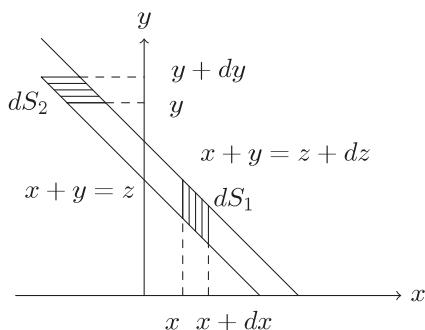


Рис. 28.3

К композиции законов распределения

вероятность может быть найдена по формуле

$$dP = f_z(z)dz, \quad (28.24)$$

во втором случае — по формуле

$$dP = dz \int_{-\infty}^{\infty} f(x, z - x)dx = dz \int_{-\infty}^{\infty} f(z - y, y)dy, \quad (28.25)$$

где первый интеграл соответствует разбиению области интегрирования на параллелограммы с вертикальными основаниями (обозначено dS_1 на рис. 28.3), во втором случае — с горизонтальными основаниями (обозначено dS_2). Приравнивая (28.24) и (28.25), получаем формулу (28.19). Особый интерес представляют случайные величины, при сложении которых вид закона распределения не меняется. Такие законы распределения называются *устойчивыми*. Из устойчивых законов распределения в приложениях наиболее часто встречается для дискретных случайных величин закон Пуассона и для непрерывных величин — нормальный закон распределения.

Действительно, если величины X и Y подчиняются закону Пуассона с параметрами a_x и соответственно a_y , то их харак-

теристические функции будут иметь вид (см. главу 4)

$$\begin{aligned} E_x(u) &= e^{ia_x[e^{iu}-1]}, \\ E_y(u) &= e^{ia_y[e^{iu}-1]}. \end{aligned}$$

Характеристическая функция суммы $Z = X + Y$ имеет вид

$$E_z(u) = E_x(u)E_y(u) = e^{ia_z[e^{iu}-1]},$$

следовательно, Z — случайная величина, подчиняющаяся закону Пуассона с параметром $a_z = a_x + a_y$.

Аналогично, если X и Y — независимые нормальные величины с математическими ожиданиями $\bar{x}$, $\bar{y}$ и дисперсиями σ_x^2 , σ_y^2 соответственно, то

$$E_z(u) = E_x(u)E_y(u) = e^{i(\bar{x}+\bar{y})u-1/2(\sigma_x^2+\sigma_y^2)u^2} = e^{iu\bar{z}-1/2\sigma_z^2u^2},$$

а это значит, что Z — нормальная случайная величина с параметрами $\bar{z} = (\bar{x} + \bar{y})$, $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$.

Сумма будет нормальной и в том случае, когда слагаемые являются зависимыми, но образуют нормальную систему величин, так как, подставляя, в соответствии с формулой (28.17), вместо аргументов u_1 и u_2 характеристической функции $E(u_1, u_2)$ системы (X, Y) один аргумент u , вновь получаем экспоненту, содержащую в показателе степени полином второй степени от u , то есть характеристическую функцию некоторого нормального закона распределения.

28.4. Частное двух случайных величин

Если рассматривается отношение

$$Z = \frac{Y}{X}, \quad (28.26)$$

то, выбирая в формуле (28.9) в качестве X_1 случайную величину X или Y , имеем соответственно

$$\frac{\partial \varphi}{\partial x_1} = \frac{\partial \varphi}{\partial x} = -\frac{y}{x^2}, \quad \frac{\partial \varphi}{\partial x_1} = \frac{\partial \varphi}{\partial y} = \frac{1}{x}, \quad (28.27)$$

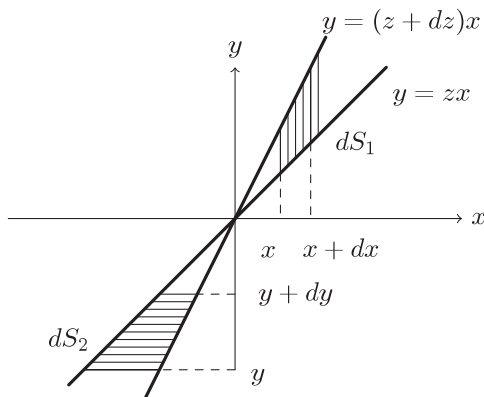


Рис. 28.4
К закону распределения частного

что после подстановки в (28.5) дает для плотности вероятности $f_z(z)$ два эквивалентных выражения:

$$\begin{aligned}
 f_z(z) &= \int_{-\infty}^{\infty} f(x, zx) |dx| dx = \\
 &= \frac{1}{z^2} \int_{-\infty}^{\infty} f\left(\frac{y}{z}, y\right) |y| dy. \quad (28.28)
 \end{aligned}$$

Рассматривая X и Y как координаты случайной точки на плоскости, формулы (28.28) можно получить и геометрически. Отличие от вывода формулы (28.19) в данном случае будет заключаться только в том, что область, определяемая неравенствами $z \leq Z \leq z + dz$, в данном случае имеет вид угла, ограниченного двумя прямыми, исходящими из начала координат с угловыми коэффициентами z и $z + dz$ (рис. 28.4), а элементы площади, вычисляемые с точностью до малых второго порядка, будут такими:

$$dS_1 = |x| dz; dx, \quad dS_2 = \frac{|y|}{z^2} dz dy.$$

28.5. Произведение двух случайных величин

Если рассмотреть произведение

$$Z = XY, \quad (28.29)$$

то и в этом случае в качестве исключаемой переменной интегрирования в формуле (28.9) можно выбрать любой из сомножителей:

$$\frac{\partial \varphi}{\partial x} = y \quad \text{для } x_1 = x, \quad \frac{\partial \varphi}{\partial y} = x \quad \text{для } x_1 = y.$$

Подставив эти значения в (28.9), получим

$$f_z(z) = \int_{-\infty}^{\infty} \left(\frac{z}{y}, y \right) |dy|^{-1} dy = \int_{-\infty}^{\infty} f \left(x_1, \frac{z}{x} \right) |x|^{-1} dx. \quad (28.30)$$

Эта формула также легко может быть получена геометрически. Представляем возможность читателю убедиться в этом самостоятельно.

28.6. Система случайных величин, функционально связанных с заданной системой величин

Рассмотрим систему случайных величин $(Y_1, \dots, Y_m)$, каждая из которых связана с заданной системой случайных величин $(X_1, \dots, X_n)$ однозначными функциональными зависимостями:

$$Y_l = \varphi_l(X_1, \dots, X_n), \quad (l = \overline{1, m}). \quad (28.31)$$

Характеристическая функция $E_y(u_1, \dots, u_n)$ системы $(Y_1, \dots, Y_m)$ определяется формулой

$$E_y(\vec{u}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{l=1}^m \varphi_l(\vec{x}) u_l} \cdot f(\vec{x}) d\mathbf{x}, \quad (28.32)$$

и, следовательно, тем самым определяется закон распределения исследуемой системы случайных величин. Однако во многих задачах удобнее иметь явное выражение для плотности

вероятности $f_y(\vec{y}) = f_y(y_1, \dots, y_m)$, а не для характеристической функции системы.

Рассмотрим вначале случай $m = n$ и будем считать, что система n равенств (28.31) однозначно определяет n обратных функциональных зависимостей:

$$X_j = \psi_j(Y_1, \dots, Y_n), \quad (j = \overline{1, n}). \quad (28.33)$$

В этом случае переход от системы случайных компонентов вектора $\vec{X}$ к системе случайных компонентов $\vec{Y}$ можно рассматривать как переход от системы старых декартовых координат $\vec{X}$ случайной точки в n -мерном евклидовом пространстве к системе новых декартовых координат $\vec{Y}$ в этом же пространстве.

Подставим в формулу, связывающую плотность вероятности $f_y(\vec{y})$ с характеристической функцией $E_y(\vec{u})$,

$$f_y(\vec{y}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} e^{-i \sum_{l=1}^n (y_l u_l)} E_y(\vec{u}) d\mathbf{u} \quad (28.34)$$

вместо $E_y(\vec{u})$ интеграл (28.32). Получим

$$f_y(\vec{y}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i \sum_{l=1}^n (\varphi_l(\vec{x}) - y_l) u_l} f_x(\vec{x}) d\mathbf{x} d\mathbf{u}. \quad (28.35)$$

Выполняя в $2n$ -кратном интеграле (28.35) вначале n интегрирований по $u_1, \dots, u_n$ и каждый раз пользуясь интегральным представлением для дельта-функции, получаем

$$f_y(\vec{y}) = \int_{-\infty}^{\infty} \prod_{l=1}^n \delta(\varphi_l(\vec{x}) - y_l) f_x(\vec{x}) d\mathbf{x}. \quad (28.36)$$

Взаимно-однозначные соотношения (28.31) и (28.32) позволяют в (28.36) перейти к переменным интегрирования ξ_l , положив

$$\xi_l = \varphi_l(\vec{x}), \quad (l = \overline{1, n}). \quad (28.37)$$

Выполнив этот переход, формулу (28.34) можно представить в виде

$$f_y(\vec{y}) = \int_{-\infty}^{\infty} (\ddot{n}) \int_{-\infty}^{\infty} \prod_{l=1}^n \delta(\xi_l - y_l) f_x[\psi(\vec{x})] |J| d\xi, \quad (28.38)$$

где

$$J = \frac{\partial(x_1, \dots, x_n)}{\partial(\xi_1, \dots, \xi_n)}$$

обозначает якобиан преобразования от переменных $\vec{x}$ к $\vec{\xi}$ и при вычислении этого якобиана в соответствии с (28.37) и (28.38) учитывается, что $x_j = \psi_j(\vec{\xi})$. Наличие под знаком интеграла произведения дельта-функций позволяет в (28.38) выполнить все n интегрирований, заменяя каждый раз после интегрирования по ξ_l в остальных множителях подынтегральной функции ξ_l на y_l . Таким образом, окончательно получим

$$f_y(\vec{y}) = f_x[\vec{\psi}(\vec{y})] \cdot \left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right|. \quad (28.39)$$

Формула (28.39) имеет простой геометрический смысл, состоящий в том, что элемент вероятности $dP(\vec{y}) = f_y(\vec{y}) d\mathbf{y}$, вычисленный в координатной системе $\vec{y}$, вследствие взаимной однозначности координат $\vec{x}$ и $\vec{y}$ равен элементу вероятности $dP(\vec{x}) = f_x(\vec{x}) d\mathbf{x}$, где значения $\vec{x}$ соответствуют выбранным значениям $\vec{y}$, а элемент n -мерного объема $d\mathbf{x}$ равен элементу объема $d\mathbf{y}$, умноженному на модуль соответствующего якобиана преобразования J .

Если величины $\vec{X}$ не могут быть выражены однозначно через величины $\vec{Y}$, то при вычислении интеграла (28.36) для перехода к новым переменным интегрирования ξ_l по формуле (28.37) необходимо область интегрирования по $\vec{x}$ разбить на подобласти, в которых имеет место взаимно однозначное соответствие между $\vec{X}$ и $\vec{Y}$, в результате вместо формулы (28.39) в правой части равенства будем иметь сумму однотипных слагаемых, каждое из которых соответствует своей подобласти и своей ветви обратной функции.

Если $m < n$, а равенства (28.31) могут быть однозначно разрешены только относительно части аргументов x_l , то в формуле (28.36) могут быть выполнены m интегрирований. Предположим, что такими аргументами являются, скажем, $x_1, \dots, z_m$, причем обратная зависимость включает оставшиеся «свободные» переменные $x_{m+1}, \dots, x_n$, как параметры

$$x_l = \psi(y_1, \dots, y_m, x_{m+1}, \dots, x_n), \quad (l = \overline{1, m}). \quad (28.40)$$

Тогда, заменяя $\{x_l\}_{l=1}^m$ переменными $\{\xi\}_l^m$ по формулам (28.37), после выполнения интегрирований получим

$$f_y(\vec{y}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_x[\vec{\psi}_*(\vec{y}, \vec{x}_*)] |J_*| d\mathbf{x}_*, \quad (28.41)$$

где компоненты, вектор-функции $\vec{\psi}_*$ даются равенством

$$\vec{\psi}_* = \begin{cases} \psi_l(y_1, \dots, y_m, x_{m+1}, \dots, x_n), & l = \overline{1, m}, \\ x_l, & l = \overline{m+1, n}, \end{cases}$$

введен вектор $\vec{x}_* = \{x_l\}_{l=m+1}^n$, $d\mathbf{x}_* = dx_{m+1} \dots dx_n$, а через J_* обозначен якобиан преобразования (28.40)

$$J_* = \left| \frac{\partial(x_1, \dots, x_m)}{\partial(y_1, \dots, y_m)} \right|.$$

Если $m > n$ и n уравнений (например, первые n уравнений) могут быть разрешены относительно x_j :

$$x_j = \psi_j(y_1, \dots, y_n), \quad (j = \overline{1, n}), \quad (28.42)$$

то, выполняя аналогичные преобразования, получаем

$$f_y(\vec{y}) = f_x(\vec{\psi}(\vec{y})) \left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right| \times \\ \times \prod_{l=n+1}^m \delta(y_l - \varphi_l), \quad (m > n), \quad (28.43)$$

где

$$\psi_j = \psi_j(y_1, \dots, y_n), \quad (j = \overline{1, n});$$

$$\varphi_l = \varphi_l(\psi_1, \dots, \psi_n), \quad (l = \overline{n+1, m}).$$

Наличие δ -функций в правой части равенства (28.43) показывает, что при $m > n$ ровно $(m - n)$ случайных величин $Y_1, \dots, Y_m$ могут быть выражены через остальные величины этой системы.

Если во всей области изменения величин $X_1, \dots, X_n$ не могут быть получены однозначные обратные зависимости типа (28.40) и (28.42), то формулы (28.41) и соответственно (28.43) усложняются — в их правых частях появляются суммы однотипных слагаемых, соответствующих областям, в которых такие однозначные зависимости имеют место.

Рассмотрим примеры вычисления законов распределения функций случайных величин.

Пример 28.1. Величина X , равномерно распределенная в интервале $[0, 1]$, связана со случайной величиной Y равенством

$$X = F(Y),$$

где $F(y)$ — функция распределения некоторой заданной случайной величины Z . Определить закон распределения Y .

Решение. Применяя формулу (28.15), замечаем, что роль функции $\psi(y)$ в данном случае играет сама функция $F(y)$, а так как $f_x(x) = 1$ ($0 \leq x \leq 1$), то

$$f_y(y) = \frac{dF(y)}{dy},$$

то есть Y имеет тот же закон распределения, что и Z .

Пример 28.2. Найти композицию n законов равномерного распределения, каждый из которых имеет нулевое математическое ожидание и параметр l .

Решение. Как показано в главе 5, характеристическая функция центрированной равномерно распределенной в интервале $(-l, l)$ величины дается равенством

$$E_0(u) = \frac{\sin(ul)}{ul}.$$

Следовательно, для характеристической функции $E(u)$ искомой величины X имеем

$$E(u) = \left(\frac{\sin ul}{ul} \right)^n = \frac{1}{2^n i^n} \sum_{j=0}^n \frac{n!}{j!(n-j)!} (-1)^j e^{iul(n-2j)}.$$

Так как плотность вероятности получается с помощью обратного преобразования Фурье

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} E(u) du,$$

то

$$\begin{aligned} \frac{d^n f(x)}{dx^n} &= \frac{(-1)^n}{(2l)^n} \sum_{j=0}^n \frac{n!}{j!(n-j)!} (-1)^j \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu[l(n-2j)-x]} du = \\ &= \frac{(-1)^n}{(2l)^n} \sum_{j=0}^n \frac{n!}{j!(n-j)!} (-1)^j \delta[l(n-2j) - x], \end{aligned}$$

где использовано обычное интегральное представление дельта-функции.

Таким образом, производная от плотности $f(x)$ порядка $n-l$ имеет разрывы первого рода в n точках: $nl, (n-2)l, \dots, -(n-2)l, -nl$, а все производные меньшего порядка и сама плотность $f(x)$ непрерывны. Следовательно,

$$f(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_0,$$

где значения a_s меняются скачками в указанных точках разрыва $f^{(n-1)}(x)$. Если $n = 2k+1$, то в интервале $-l < x < l$ функция $f(x)$ непрерывна и для коэффициентов a_s имеем

$$\begin{aligned} a_s &= \frac{1}{s!} f^{(s)}(0) = \frac{(-i)^s}{2\pi s! l^{2k+1}} \int_{-\infty}^{\infty} \frac{\sin^{2k+1}(ul)}{u^{2k-s+1}} du = \\ &= \frac{(-i)^s}{s! l^{s+1}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin^{2k+1}(y)}{y^{2k-s+1}} dy. \end{aligned}$$

При нечетном s все коэффициенты a_s обращаются в нуль. При четном s для вычисления интеграла достаточно заменить $\sin y = \frac{1}{2i}(e^{iy} - e^{-iy})$ и учесть, что

$$\frac{1}{\pi i} V \cdot p \cdot \int_{-\infty}^{\infty} \frac{e^{iay}}{y} dy = \text{sign}(a),$$

где интеграл понимается в смысле главного значения по Коши

$$V \cdot p \cdot \int_{-\infty}^{\infty} e^{iay} \frac{dy}{y^{n+1}} = \frac{(ia)^n}{(n-1)!} \text{sign}(a), \quad (a \neq 0).$$

Зная коэффициенты a_s для центрального интервала, значения a_{2k} для остальных интервалов определяем с помощью формулы (28.44), на основании которой скачки этого коэффициента на границе j -го интервала непрерывности (считая центральный интервал за нулевой) равны

$$\frac{1}{(2k)!} f^{(n)}(2jl - l) = \frac{(-1)^{k-j-1} (2k+1)}{(2l)^{2k+1} (k-j)!(K+1+j)!}.$$

Остальные коэффициенты определяются из условий непрерывности $f(x)$ и первых ее $(2k-1)$ производных. Например, если $n=3$, то $f(x) = a_2 x^2 + a_1 x + a_0$. Для центрального интервала

$$a_0^{(0)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin^3(y)}{y^3 l} dy = \frac{2}{8l}, \quad a_1^{(0)} = 0,$$

$$a_2^{(0)} = \frac{1}{2} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin^3(y)}{y l^3} dy = -\frac{2}{8l^3}.$$

Скачок значения a_2 в точке $x=l$ равен $+\frac{1}{16l^3}$, то есть сам коэффициент $a_2^{(1)} = \frac{1}{16l^3}$. Из непрерывности $f^{(1)}(x)$ при

$x = l$ следует $a_1^{(0)} + 2la_2^{(0)} = a_1^{(1)} + 2la_2^{(1)}$, откуда $a_1^{(1)} = -\frac{3}{8l^2}$. Наконец, из непрерывности $f(x)$ следует $a_0^{(0)} + la_1^{(0)} + l^2a_2^{(0)} = a_0^{(1)} + la_1^{(1)} + l^2a_2^{(1)}$, а тогда $a_2^{(1)} = \frac{1}{16l^3}$. Таким образом, окончательно, для промежутка $|x| \leq 3l$ имеем

$$f(x) = \begin{cases} \frac{1}{8l^3} (3l^2 - x^2), & |x| \leq l, \\ \frac{1}{16l^3} (9l^2 - 6lx + x^2), & l \leq |x| \leq 3l. \end{cases}$$

При $|x| > 3l$, очевидно, имеем $f(x) \equiv 0$.

Задачу можно решить и непосредственно, вычисляя плотность вероятности $f(x)$ по формуле (28.19): вначале сложить два закона равномерного распределения, потом к полученной композиции прибавить третий закон равномерного распределения и т. д. Однако при большом n такой способ вычислений менее рационален.

Пример 28.3. Найти композицию n экспоненциальных законов распределения

$$f(x) = \lambda e^{-\lambda x} \quad (x \geq 0).$$

Решение. Характеристическая функция экспоненциального закона распределения

$$E_1(u) = \int_0^{\infty} e^{iux} \lambda e^{-\lambda x} dx = \frac{1}{1 - iu/\lambda}.$$

Следовательно, характеристическая функция $E_n(u)$ искомого закона распределения

$$E_1(u) = E_1^n(u) = \left(1 - i\frac{u}{\lambda}\right)^{-n}.$$

Плотность вероятности этого закона распределения выражается формулой обратного преобразования Фурье

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iux} \left(1 - i\frac{u}{\lambda}\right)^{-n} du = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}, \quad (x \geq 0).$$

Для вычисления обратного преобразования Фурье используется теория вычетов так, как это было описано в § 23. Таким образом, композиция n экспоненциальных законов распределения дает гамма-распределение с параметрами n и λ , которое путем замены $\lambda x = \frac{1}{2} z$ может быть преобразовано в закон распределения χ^2 при $2n$ степенях свободы (см. далее (29.17)). Полученный результат, в частности, показывает, что композиция двух гамма-распределений с одинаковым параметром λ и параметрами n_1 и n_2 дает гамма-распределение с параметрами λ и $(n_1 + n_2)$, (n_1 и n_2 могут быть и дробными).

§ 29. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ, СВЯЗАННЫЕ С НОРМАЛЬНЫМ ЗАКОНОМ РАСПРЕДЕЛЕНИЯ

В качестве примеров вычисления законов распределения функций случайных величин рассмотрим несколько законов распределения, представляющих особый интерес в ряде приложений, в частности, в математической статистике.

29.1. Законы распределения χ и χ^2

Рассмотрим систему независимых, нормированных центрированных нормальных случайных величин ξ_j ($j = \overline{1, n}$). Плотность вероятности $f_{\xi}(\vec{\xi})$ этой системы определяется формулой

$$f_{\xi}(\vec{\xi}) = \frac{1}{(2\pi)^{n/2}} e^{-1/2 \sum_{j=1}^n \xi_j^2}. \quad (29.1)$$

Образует случайную величину χ , положив

$$\chi = \sqrt{\sum_{j=1}^n \xi_j^2}, \quad (29.2)$$

где число слагаемых n называют *числом степеней свободы*. В обозначениях формулы (28.1)

$$\chi = Y, \quad \xi_j = X_j, \quad (j = \overline{1, n}).$$

Обратная зависимость

$$\xi_l = \pm \sqrt{\chi^2 - \sum_{j=1}^n \xi_j^2} \quad (29.3)$$

двузначная. Поэтому применив формулу (28.4) для этого случая, получим

$$f(\chi) = 2 \int \int_{\sum_{j=2}^n \xi_j^2 \leq \chi}^{(n-1)} \frac{1}{(2\pi)^{n/2}} e^{-\chi^2} \times \\ \times \left| \frac{\partial \sqrt{\chi^2 - \sum_{j=1}^n \xi_j^2}}{\partial \chi} \right| d\xi_2 \dots d\xi_n. \quad (29.4)$$

Выписанный интеграл необходимо удвоить, так как нужно учесть неоднозначность обратной зависимости (29.3). Обеим ветвям обратной функции соответствуют равные слагаемые.

Вычисляя производную $\frac{\partial \xi_1}{\partial \chi}$, получаем в итоге

$$f(\chi) = \frac{1}{(2\pi)^{n/2}} \int \int_{\sum_{j=2}^n \xi_j^2 \leq \chi}^{(n-1)} e^{-\chi^2/2} \cdot \frac{\chi}{\sqrt{\chi^2 - \sum_{j=2}^n \xi_j^2}} d\xi_2 \dots d\xi_n. \quad (29.5)$$

Если теперь вместо переменных интегрирования ξ_j , ($j = \overline{1, n}$) ввести новые переменные η_j , положив

$$\xi_j = \chi \eta_j, \quad (j = \overline{1, n}),$$

то интегралу (29.5) можно придать вид

$$f(\chi) = e^{-\chi/2} \frac{1}{(2\pi)^{n/2}} \int_{\sum_{j=2}^n \eta_j^2 < 1}^{(n-1)} \int \frac{d\eta_2 \dots d\eta_n}{\sqrt{1 - \sum_{j=2}^n \eta_j^2}}. \quad (29.6)$$

Так как интеграл в (29.5) никак не зависит от χ , то плотность вероятности $f(\chi)$ может быть представлена в виде

$$f(\chi) = ce^{-\chi^2/2} \chi^{n-1}, \quad (29.7)$$

где постоянный коэффициент c может быть определен из условия нормировки

$$\int_0^\infty f(\chi) d\chi = 1. \quad (29.8)$$

Выполняя очевидные преобразования, основанные на преобразовании полученного интеграла к гамма-функции [17]

$$\begin{aligned} c \int_0^\infty e^{-\chi^2/2} \chi^{n-1} d\chi &= c \int_0^\infty e^{-t} 2^{n/2} \cdot \frac{1}{2} t^{n/2-1} dt = \\ &= c \cdot 2^{n/2-1} \Gamma(n/2) = 1, \end{aligned} \quad (29.9)$$

находим

$$c = \frac{2}{2^{n/2} \Gamma\left(\frac{n}{2}\right)}. \quad (29.10)$$

Далее, подставляя (29.10) в (29.7), окончательно получим¹

$$f(\chi) = \frac{c}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} e^{-\frac{\chi^2}{2}}, \quad 0 \leq \chi < \infty. \quad (29.11)$$

¹ Формула (29.11) может быть получена и геометрически, если рассматривать ξ_j как прямоугольные координаты случайной точки в n -мерном пространстве: перейдя от декартовых координат к сферическим $\chi, \theta_1, \dots, \theta_{n-1}$ и учитывая при этом, что якобиан преобразования пропорционален χ^{n-1} , получим (29.7), что после выполнения условия нормировки (29.8) дает (29.11).

При $n = 2$, согласно (29.11), имеем

$$f(\chi) = \chi \cdot e^{-\chi^2/2}, \quad (29.12)$$

то есть плотность вероятности хорошо известного нам закона распределения Релея (см. § 23).

При $n = 3$, учитывая, что $\Gamma\left(\frac{n}{2}\right) = \frac{\sqrt{3}}{2}$, получаем

$$f(\chi) = \frac{2}{\sqrt{\pi}} \chi^2 e^{-\frac{\chi^2}{2}},$$

что совпадает с плотностью вероятности закона распределения Максвелла, также разобранный в § 23.

Вернемся к общему случаю произвольного целого n . В соответствии с (29.11) для начального момента m_k порядка k имеем

$$\begin{aligned} m_k &= \frac{c}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} e^{-\frac{\chi^2}{2}} \chi^{n+k-1} d\chi = \\ &= 2^{k/2} \frac{\Gamma\left(\frac{n+k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \end{aligned} \quad (29.13)$$

так как рассматриваемый интеграл может быть получен из интеграла (29.9) простой заменой n на $(n+k)$.

Полагая вначале $k = 1$, получаем

$$m_1 = \bar{\chi} = \sqrt{2} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}. \quad (29.14)$$

Далее при $k = 2$ приходим к чрезвычайно простому выражению и, следовательно, дисперсия закона распределения χ

$$\mathbf{D}(\chi) = n - \bar{\chi}^2 = n - 2 \left[\frac{\Gamma\left(\frac{n+k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \right]^2. \quad (29.15)$$

Иногда бывает удобнее рассматривать не саму случайную величину χ , а ее квадрат, то есть величину

$$Z = \chi^2. \quad (29.16)$$

Используя формулу (28.12) и учитывая (29.11), получим для указанной величины

$$f(z) = \frac{1}{2^{n/2}\Gamma(n/2)} \cdot e^{-z/2} z^{n/2-1}, \quad (0 \leq z < \infty). \quad (29.17)$$

Закон распределения величины Z , определяемой (29.16), называется *законом распределения хи-квадрат* с n степенями свободы. Так как по определению

$$Z = \sum_{j=1}^n \xi_j^2, \quad (29.18)$$

то, воспользовавшись теоремами о математическом ожидании и дисперсии, получим

$$\bar{z} = \mathbf{M}(Z) = n, \quad \mathbf{D}(Z) = 2n, \quad (29.19)$$

поскольку для центрированных и нормированных нормальных случайных величин ξ_j имеем (см. § 22)

$$\mathbf{M}(\xi_j^2) = \mathbf{D}(\xi_j) = 1, \quad \mathbf{D}(\xi_j^2) = 2.$$

Для начального момента m_k закона распределения χ^2 , подставив в (29.13) значение $2k$ вместо k , получим

$$m_k = 2^k \frac{\Gamma\left(k + \frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = (n + 2k - 2) \dots (n + 2)n.$$

29.2. Закон распределения Стьюдента

При рассмотрении одной важной задачи математической статистики Стьюдент² ввел в рассмотрение случайную величину T , определяемую равенством

$$T = \frac{\sqrt{n}\xi_0}{\sqrt{\sum_{j=1}^n \xi_j^2}}, \quad (29.20)$$

где ξ_l ($l = \overline{0, n}$) — взаимно независимые, центрированные, нормированные нормальные случайные величины. Закон распределения T называется *законом распределения Стьюдента* с n степенями свободы.

Определим плотность вероятности $f(t)$ этого закона распределения. Обозначив

$$Y = \sqrt{n}\xi_0 \quad (29.21)$$

и учитывая (29.2), вместо (29.20) можно записать

$$T = \frac{Y}{\chi}, \quad (29.22)$$

где Y — нормальная центрированная случайная величина, дисперсия которой равна n , а плотность вероятности знаменателя, не зависящего от Y , определяется формулой (29.11).

² Стьюдент — псевдоним английского статистика Вильяма Госсета (1876–1937).

Случайная величина Y не зависит от χ , так как ξ_0 не зависит от случайных величин ξ_j , входящих в знаменатель (29.20). Следовательно, плотность вероятности системы (Y, χ) является произведением плотностей вероятности случайных величин Y и χ

$$f(y, \chi) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{y^2}{2n}} \frac{2}{2^n \Gamma(n/2)} e^{-\frac{\chi^2}{2}} \chi^{n-1}. \quad (29.23)$$

Поэтому, применяя для нахождения плотности вероятности $f(t)$ формулу (28.26), для плотности вероятности частного, получаем

$$f(t) = \frac{2}{2^n \Gamma(n/2) \sqrt{2\pi n}} \int_0^\infty e^{-\frac{1}{2}(1+\frac{t^2}{n})\chi^2} \chi^n d\chi. \quad (29.24)$$

Перейдя от переменной интегрирования χ к новой переменной интегрирования η , положив

$$\chi = \sqrt{2} \left(1 + \frac{t^2}{n}\right)^{-1/2} \eta^{1/2}, \quad (29.25)$$

последний интеграл можно выразить через гамма-функцию в виде

$$f(t) = \frac{2}{2^n \Gamma(n/2) \sqrt{2\pi n}} 2^{\frac{n+1}{2}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \times \\ \times \frac{1}{2} \cdot \int_0^\infty e^{-\eta} \eta^{\frac{n+1}{2}-1} d\eta, \quad (29.26)$$

откуда получаем

$$f(t) = \frac{2}{2^n \Gamma\left(\frac{n+1}{2}\right) \sqrt{2\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad (29.27) \\ (-\infty < t < \infty).$$

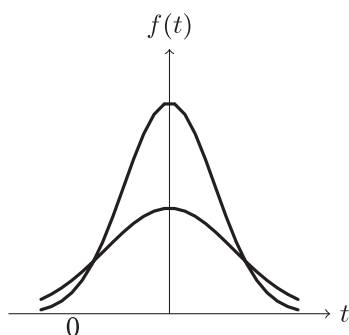


Рис. 29.1
Сравнение плотностей
вероятностей нормального
закона и закона
Стьюдента

График плотности $f(t)$ напоминает график плотности нормального закона распределения, однако между законом распределения Стьюдента и нормальным законом имеется ряд существенных отличий. Во-первых, так как при больших значениях $|t|$ плотность $f(t) \sim \frac{1}{|t|^{n+1}}$, то моменты закона распределения Стьюдента существуют только до порядка $k = n - 1$. Во-вторых, при больших $|t|$ плотность $f(t)$ стремится к нулю медленнее, чем плотность нормального за-

кона распределения. Поэтому начальная ордината плотности $f(t)$ меньше начальной ординаты плотности нормального закона распределения. Схематически отличие плотности $f(t)$ от плотности нормального закона распределения представлено на рис. 29.1, при построении которого принято, что дисперсии нормального закона распределения и закона распределения Стьюдента одинаковы.

При больших значениях n различие между $f(t)$ и плотностью нормального закона распределения исчезает, так как

$$\lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{n} \right)^{-\frac{n+1}{2}} = e^{-\frac{t^2}{2}},$$

следовательно,

$$\lim_{n \rightarrow \infty} f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}. \quad (29.28)$$

При $k < n$ все нечетные моменты m_k распределения Стьюдента обращаются в нуль вследствие четности плотности вероятности $f(t)$, а четные начальные моменты совпадают с цен-

тральными моментами. Положив $k = 2s$, имеем

$$m_{2s} = \mu_{2s} = 2 \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{\pi n}} \int_0^\infty \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} t^{2s} dt. \quad (29.29)$$

Последний интеграл может быть сведен к бета-функции $B(p, q)$, определяемой равенством (см. [17]):

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \quad (29.30)$$

Действительно, введя в (29.29) новую переменную интегрирования x по правилу

$$t = \sqrt{\frac{nx}{1-x}}, \quad dt = \frac{\sqrt{n}}{2} x^{-1/2} (1-x)^{-3/2} dx, \quad (29.31)$$

получим

$$\begin{aligned} m_{2s} = \mu_{2s} &= \frac{n^s \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{\pi}} \int_0^1 x^{s-\frac{1}{2}} (1-x)^{-\frac{n+1}{2}-\frac{3}{2}-s} dx = \\ &= \frac{n^s \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{\pi}} B\left(s + \frac{1}{2}, \frac{n}{2} - s\right) = \\ &= \frac{n^s \Gamma\left(s + \frac{1}{2}\right) \Gamma\left(\frac{n}{2} - s\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{\pi}}. \end{aligned}$$

Таким образом, окончательно

$$m_{2s} = \mu_{2s} = \frac{n^s \Gamma\left(s + \frac{1}{2}\right) \Gamma\left(\frac{n}{2} - s\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{\pi}}, \quad 2s \leq n - 1. \quad (29.32)$$

В частном случае, если $n > 2$, для дисперсии имеем

$$\mu_2 = \mathbf{D}(T) = \frac{n}{n - 2}. \quad (29.33)$$

Если $n > 4$, то для распределения Стюдента существуют моменты до четвертого порядка включительно. В этом случае коэффициент асимметрии и эксцесс выражаются в виде

$$Sk = \frac{\mu_3}{\sigma_3} = 0, \quad Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{6}{n - 4}.$$

29.3. Закон распределения Фишера

В математической статистике часто используется случайная величина

$$\Phi = \frac{m \sum_{j=1}^n \xi_j^2}{n \sum_{l=1}^m \eta_l^2}, \quad (29.34)$$

где ξ_j и η_l , ($j = \overline{1, n}$; $l = \overline{1, m}$) — независимые, центрированные, нормированные, нормальные величины. Закон распределения так построенной случайной величины Φ называется *законом распределения Фишера*³, впервые рассматривавшего такие величины.

Учитывая определение (29.2) случайной величины χ с n степенями свободы, формуле (29.34) можно придать вид

$$\Phi = \frac{m}{n} \frac{\chi_n^2}{\chi_m^2} = \frac{m}{n} \Phi_1, \quad \Phi_1 = \frac{Z_n}{Z_m}, \quad (29.35)$$

³ Фишер Рональд (1890–1968) — английский статистик.

где $Z_n = \chi_n^2$, $Z_m = \chi_m^2$ — взаимно независимые случайные величины, распределенные по закону хи-квадрат.

Используя формулу (28.26) для плотности вероятности частного случайных величин и учитывая (29.17) для плотности вероятности χ^2 распределения, получим

$$f_{\varphi_1}(\varphi_1) = \frac{\varphi_1^{\frac{n}{2}-1}}{2^{\frac{n+m}{2}} \Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{m}{2}\right)} \times \\ \times \int_0^\infty e^{-\frac{(1+\varphi_1)z_m}{2}} z_m^{\frac{n+m}{2}-1} dz_m. \quad (29.36)$$

Последний интеграл подстановкой $t = \frac{1+\varphi_1}{2} z_m$ сводится к гамма-функции $\Gamma\left(\frac{n+m}{2}\right)$. Выполнив необходимые преобразования, получим

$$f_{\varphi_1}(\varphi_1) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{m}{2}\right)} \varphi_1^{\frac{n}{2}-1} (1+\varphi_1)^{-\frac{n+m}{2}}, \quad (\varphi \geq 0). \quad (29.37)$$

Так как $\Phi = \frac{m}{n} \Phi_1$, то в соответствии с (28.12) находим

$$f(\varphi) = \frac{m}{n} f_{\varphi_1}\left(\frac{n}{m} \varphi\right),$$

или в явном виде

$$f_{\varphi}(\varphi) = \left(\frac{n}{m}\right)^{n/2} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \times \\ \times \frac{\frac{n}{2}^{-1}}{\left(1 + \frac{n}{m}\varphi\right)^{\frac{n+m}{2}}}, \quad \varphi \geq 0. \quad (29.38)$$

При $\varphi \rightarrow \infty$ плотность вероятности распределения Фишера убывает как $\varphi^{-\frac{m+2}{2}}$. Следовательно, начальные моменты этого распределения существуют только до значений $k < \frac{m}{2}$. Для таких k имеем

$$m_k = \left(\frac{n}{m}\right)^{n/2} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \times \\ \times \int_0^{\infty} \frac{\varphi^{\frac{n}{2}+k-1}}{\left(1 + \frac{n}{m}\varphi\right)^{\frac{n+m}{2}}} d\varphi. \quad (29.39)$$

Последний интеграл сводится к бета-функции, так как при

$$\varphi = \frac{m}{n} \frac{x}{1-x}, \quad (29.40)$$

он приводится к виду

$$\begin{aligned} \int_0^{\infty} \frac{\varphi^{\frac{n}{2}+k-1}}{\left(1 + \frac{n}{m}\varphi\right)^{\frac{n+m}{2}}} d\varphi &= \\ &= \left(\frac{n}{2} + k\right)^{n/2} \int_0^{\infty} x^{\frac{n}{2}+k-1} (1-x)^{\frac{m}{2}-k-1} dx = \\ &= \left(\frac{n}{2} + k\right)^{n/2} B\left(\frac{n}{2} + k, \frac{m}{2} - k\right), \end{aligned}$$

откуда

$$m_k = \left(\frac{m}{n}\right)^k \frac{\Gamma\left(\frac{n}{2} + k\right) \Gamma\left(\frac{m}{2} - k\right)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{m}{2}\right)}, k < \frac{m}{2}. \quad (29.41)$$

При $m > 4$ распределение Фишера имеет математическое ожидание и дисперсию, причем

$$\begin{aligned} \bar{\varphi} &= \frac{m}{m-2}, \\ \mathbf{D}(\Phi) &= m_2 - \bar{\varphi}^{-2} = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}. \end{aligned} \quad (29.42)$$

Рассмотрим два числовых примера.

Пример 29.1. Для случайной величины T , подчиняющейся закону распределения Стьюдента, определить вероятность

$$P = \mathbf{P}\{|T| \leq 3\sigma\}, \quad n = 5; 10; 20; 30.$$

Решение. В соответствии с формулой (29.33) дисперсия σ^2 закона распределения Стьюдента равна $\frac{n}{n-2}$. Следовательно,

ширина указанного интервала 3σ , для условий примера будет соответственно

$$3\sqrt{\frac{5}{3}} = 3,873, \quad 3\sqrt{\frac{10}{8}} = 3,354,$$

$$3\sqrt{\frac{20}{18}} = 3,162, \quad 3\sqrt{\frac{30}{28}} = 3,105.$$

Так как закон распределения Стьюдента симметричен, то, выразив искомую вероятность через функцию распределения $F(y)$, получим

$$P = 2F(3\sigma) - 1.$$

Пользуясь таблицами функции распределения Стьюдента [4] по заданным значениям n и найденным выше значениям 3σ находим искомую вероятность, значения которой приведены ниже. Как видно, при малых n вероятность P заметно отличается от вероятности попадания центрированной нормальной величины в интервал $(-3\sigma, 3\sigma)$, равной 0,9973.

Таблица 29.1

Вероятность невыхода за уровень 3σ

n	5	10	20	30
P	0,9842	0,9928	0,9951	0,9962

Пример 29.2. Для закона распределения Фишера при $n = 5$, $m = 10$ определить значения коэффициента асимметрии и эксцесса.

Решение. В соответствии с формулой (29.41) для заданных числовых значений имеем

$$m_1 = \frac{m}{m-2} = \frac{5}{4}, \quad m_2 = \frac{m^2(n+2)}{n(m-2)(m-4)} = \frac{35}{12},$$

$$m_3 = \frac{m^3(n+4)(n+2)}{n^2(m-2)(m_4)(m-6)} = \frac{105}{8},$$

$$m_4 = \frac{m^4(n+6)(n+4)(n+2)}{n^3(m-2)(m-4)(m-6)(m-8)} = \frac{1155}{8},$$

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^2 = \frac{195}{32},$$

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 = \frac{13805}{128}.$$

По формуле (29.42) имеем $\sigma^2 = \frac{65}{48}$. Следовательно, коэффициент асимметрии $Sk = 3,9$, а эксцесс $Ex = \frac{13805 \cdot 48 \cdot 48}{128 \cdot 65 \cdot 65} - 3 = 2,9$.

§ 30. МЕТОД ЛИНЕАРИЗАЦИИ ФУНКЦИЙ СЛУЧАЙНЫХ ВЕЛИЧИН

Формулы предыдущего параграфа позволяют вычислить закон распределения функции случайных величин, если закон распределения случайных аргументов известен. Однако вычисление по этим формулам часто является достаточно сложным. Поэтому во многих приложениях применяется метод линеаризации функций случайных величин, не только позволяющий упростить вычисления, но и дающий возможность приближенно выразить моменты функции случайных величин через моменты ее аргументов, независимо от вида их закона распределения.

Рассмотрим вначале функцию Y одной случайной величины X :

$$Y = \varphi(X), \quad (30.1)$$

где функцию $\varphi(x)$ в области изменения случайной величины X по-прежнему будем считать однозначной, но, кроме того, предположим, что в точке $x = \bar{x}$ функция $\varphi(x)$ имеет необходимое число производных. Тогда в достаточно малой окрестности точки $\bar{x}$ функцию $\varphi(X)$ можно разложить в ряд Тэйлора,

положив

$$Y = \varphi(\bar{x}) + \frac{d\varphi(\bar{x})}{d\bar{x}} \overset{\circ}{X} + \frac{1}{2} \cdot \frac{d^2\varphi(\bar{x})}{d\bar{x}^2} \overset{\circ}{X}^2 + \dots \\ \dots + \frac{1}{n!} \cdot \frac{d^n\varphi(\bar{x})}{d\bar{x}^n} \overset{\circ}{X}^n + R_n, \quad (30.2)$$

где $\overset{\circ}{X} = X - \bar{x}$, а остаточный член R_n вследствие случайности X будет случайной величиной.

Первые два члена разложения (30.2) дают замену функциональной зависимости $\varphi(X)$ линейной функцией, учет следующих слагаемых соответствует переходу к квадратичной, кубичной и другим более точным аппроксимациям.

Если вероятность получения значений X , для которых линейная аппроксимация дает заметные ошибки, мала, то вместо (30.2) можно приближенно написать

$$Y \approx \varphi(\bar{x}) + \frac{d\varphi(\bar{x})}{d\bar{x}} \overset{\circ}{X}. \quad (30.3)$$

Ошибка такого приближенного равенства, разумеется, требует оценки.

Предположим вначале, что такая оценка произведена и точность линейной аппроксимации (30.3) является приемлемой для рассматриваемой задачи. В этом случае на основании теорем о математическом ожидании и дисперсии легко находим

$$\bar{y} \approx \varphi(\bar{x}), \quad \sigma_y^2 \approx \left[\frac{d\varphi(\bar{x})}{d\bar{x}} \right]^2 \sigma_x^2, \quad \sigma_y \approx \left| \frac{d\varphi(\bar{x})}{d\bar{x}} \right| \sigma_x. \quad (30.4)$$

Кроме того, если, например, случайная величина X нормальна, то вследствие линейности соотношения (30.3) и случайная величина Y будет также нормальной.

Оценим допустимость метода линеаризации. Во-первых, очевидно, что этот метод недопустим, если $\frac{d\varphi(\bar{x})}{d\bar{x}}$ обращается в бесконечность или не существует, либо $\frac{d\varphi(\bar{x})}{d\bar{x}} = 0$. В первом случае разложение в ряд Тейлора в точке $x = \bar{x}$ недопустимо,

а во втором случае удержание только линейных слагаемых не учитывает случайный характер Y , поскольку первое случайное слагаемое в ряду Тэйлора будет квадратично.

Точная оценка ошибок приближенных формул (30.4), как правило, связана с громоздкими вычислениями и требует знания закона распределения $f(x)$. Однако обычно достаточно надежную оценку точности можно получить, сохраняя в ряде (30.2) следующие члены разложения и сравнивая полученный результат с первоначальным.

Так, удерживая квадратичные члены, получим

$$\left\{ \begin{array}{l} \bar{y} \approx \varphi(\bar{x}) + \frac{1}{2} \cdot \frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \sigma_x^2, \\ \sigma_y^2 \approx \left[\frac{d\varphi(\bar{x})}{d\bar{x}} \right]^2 \sigma_x^2 + \frac{d\varphi(\bar{x})}{d\bar{x}} \frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \mu_3^{(x)} + \\ \quad + \frac{1}{4} \left[\frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \right]^2 (\mu_4^{(x)} - \sigma_x^4). \end{array} \right. \quad (30.5)$$

Отсюда находим поправки $\delta_{\bar{y}}$ и $\delta(\sigma_y^2)$ к формулам (30.4), которые приближенно определяются равенствами:

$$\left\{ \begin{array}{l} \delta_{\bar{y}} \approx \frac{1}{2} \cdot \frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \sigma_x^2, \\ \delta(\sigma_y^2) \approx \frac{d\varphi(\bar{x})}{d\bar{x}} \cdot \frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \mu_3^{(x)} + \\ \quad + \frac{1}{4} \left[\frac{d^2 \varphi(\bar{x})}{d\bar{x}^2} \right]^2 \cdot (\mu_4^{(x)} - \sigma_x^4). \end{array} \right. \quad (30.6)$$

Необходимость в численной проверке ошибок, связанных с линеаризацией, в большинстве практических задач не возникает, так как убедиться в ее допустимости удастся из физических соображений. При этом иногда может оказаться, что метод линеаризации дает для $\bar{y}$ и σ_y^2 конечные и согласующиеся с практикой результаты, в то время как точных значений математического ожидания и дисперсии $\varphi(X)$ при заданном

виде плотности вероятности не существует вследствие расходимости соответствующих интегралов. Происходит это потому, что в качестве закона распределения случайной величины X в этих случаях принимается не истинный закон распределения, имеющий сложный вид или не известный точно, а его приближенное выражение, достаточно хорошо аппроксимирующее истинный закон распределения, однако оказывающееся непригодным при исследовании случайной величины Y вследствие специфического вида функции $\varphi(X)$.

Рассмотрим этот вопрос подробнее на конкретном примере. Пусть, например,

$$Y = \frac{1}{X}, \quad (30.7)$$

где под случайной величиной X будем понимать значение величины математического ожидания, полученное в результате измерения. В большинстве случаев можно считать, что величина X — нормальная с математическим ожиданием $\bar{x}$ и дисперсией σ_x^2 , характеризующей точность измерения, хотя в действительности закон распределения ошибок измерения отличается от нормального, поскольку реальные ошибки измерения всегда ограничены некоторым конечным интервалом, а нормальный закон распределения дает конечную вероятность для любого отклонения величины X от ее математического ожидания. Тем не менее допущение нормального закона распределения при исследовании величины X обычно не приводит к каким-либо ощутимым противоречиям с наблюдениями.

Положение меняется, если мы перейдем к исследованию величины Y , определяемой формулой (30.7). Если принять, что X — нормальная величина, то есть

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}},$$

то для величины Y не существует уже первый момент, так как интеграл $\int_{-\infty}^{\infty} \frac{1}{|x|} f(x) dx$ расходится. Между тем, если $\sigma_x \ll \bar{x}$, то применимость метода линеаризации не вызывает сомнения

и формулы (30.4) дают

$$\bar{y} \approx \frac{1}{\bar{x}}, \quad \sigma_y^2 = \frac{1}{\bar{x}^4} \sigma_x^2, \quad \sigma_y = \frac{1}{\bar{x}^2} \sigma_x.$$

Полученное противоречие является кажущимся и связано с тем, что допущение нормальности величины X , являющееся удобной математической абстракцией в других случаях, при вычислении моментов обратной величины $1/X$ не может быть принято.

Метод линеаризации допускает очевидное обобщение на случай функции нескольких переменных. Предположим, что $\vec{X} = (X_1, \dots, X_n)$ — вектор, составленный из независимых случайных величин, и рассматривается некоторая функция

$$Y = \varphi(\vec{X}), \quad (30.8)$$

причем функция $\varphi(\vec{x})$ допускает дифференцирование по каждому из своих аргументов необходимое число раз. Считая, что и в этом случае можно из области возможных значений случайных величин выделить подобласть, вероятность попадания в которую значений $\vec{X}$ близка к единице и для которой в ряде Тэйлора можно ограничиться только линейными членами, имеем

$$Y \approx \varphi(\vec{m}_x) + \sum_{j=1}^n \frac{\partial \varphi(\vec{x})}{\partial x_j} \bigg|_{\vec{x}=\vec{m}_x} \overset{\circ}{X}_j. \quad (30.9)$$

Применяя к (30.9) теоремы о математическом ожидании и дисперсии, получаем формулы, аналогичные (30.4):

$$\begin{cases} \bar{y} \approx \varphi(\vec{m}_x), & \sigma_y^2 \approx \sum_{j=1}^n \left[\frac{\partial \varphi(\vec{x})}{\partial x_j} \bigg|_{\vec{x}=\vec{m}_x} \right]^2 \sigma_{x_j}^2, \\ \sigma_y = \sqrt{\sigma_y^2}, & \vec{m}_x = \{\bar{x}_j\}_{j=1}^n. \end{cases} \quad (30.10)$$

Формулы (30.10) неприменимы, если хотя бы одна из частных производных $\frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_j}$ не существует или обращается

в бесконечность, или же, если все эти частные производные обращаются в нуль (так как в этом случае первый по малости случайный член ряда Тэйлора зависит от $\overset{\circ}{X}_j$ квадратично).

Как и в случае функции одной переменной, точность формулы (30.10) можно приближенно оценить, найдя поправки $\delta_{\bar{y}}$ и $\delta(\sigma_y^2)$ во втором приближении, то есть учитывая квадратичные члены в ряде Тэйлора, положив вместо (30.9)

$$Y \approx (\varphi(\vec{m}_x) + \sum_{j=1}^n \frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_j} \overset{\circ}{X}_j + \frac{1}{2} \sum_{j=1}^n \sum_{l=1}^n \frac{\partial^2 \varphi(\vec{m}_x)}{\partial \bar{x}_j \partial \bar{x}_l} \overset{\circ}{X}_j \overset{\circ}{X}_l). \quad (30.11)$$

Очевидные вычисления дают

$$\left\{ \begin{aligned} \delta_{\bar{y}} &\approx \frac{1}{2} \cdot \sum_{j=1}^n \frac{\partial^2 \varphi(\vec{m}_x)}{\partial \bar{x}_j^2} \sigma_{x_j}^2, \\ \delta(\sigma_y^2) &\approx \frac{1}{2} \sum_{j=1}^n \frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_j} \frac{\partial^2 \varphi(\vec{m}_x)}{\partial \bar{x}_j^2} \mu_3^{(x_j)} + \\ &+ \frac{1}{4} \cdot \left\{ \sum_{j=1}^n \left[\frac{\partial^2 \varphi(\vec{m}_x)}{\partial \bar{x}_j^2} \right] \cdot (\mu_4^{(x_j)} - \sigma_{x_j}^4) \right. \\ &\left. + 2 \sum_{j=1}^n \sum_{l=j+1}^n \left[\frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_j} \right]^2 \left[\frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_l} \right]^2 \sigma_{x_j}^2 \sigma_{x_l}^2 \right\}, \end{aligned} \right. \quad (30.12)$$

где $\mu_4^{(x_j)}$ и $\mu_3^{(x_j)}$ — третий и, соответственно, четвертый моменты случайной величины X_j .

Если компоненты вектора $\vec{X}$ не являются независимыми, то первая формула (30.10) не изменится, а вторая примет вид

$$\sigma_y^2 \approx \sum_{j=1}^n \sum_{l=1}^n \frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_j} \frac{\partial \varphi(\vec{m}_x)}{\partial \bar{x}_l} k_{jl}^{(x)}, \quad (30.13)$$

где $k_{jl}^{(x)}$ — корреляционные моменты системы $\vec{X}$.

Как и в случае функции одной случайной величины, если система компонент $\vec{X}$ нормальная, а метод линеаризации допустим, то случайная величина Y может считаться нормальной, поскольку ее приближенное значение (30.9) является линейной функцией нормальных величин.

Метод линеаризации допускает геометрическую интерпретацию, которая в ряде случаев упрощает вычисление вторых моментов функций случайных величин. Рассмотрим вначале скалярную функцию, считая компоненты случайного вектора $\vec{X}$ независимыми

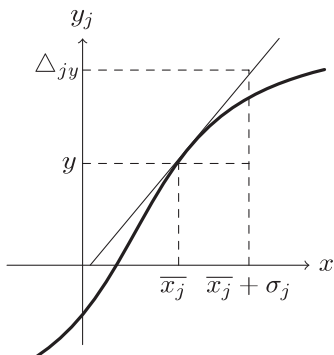


Рис. 30.1

Случай монотонной функции

$$Y = \varphi(\vec{X}). \quad (30.14)$$

Изобразим графически зависимость Y от X_j при значениях остальных случайных величин, равных их математическим ожиданиям (рис. 30.1). Указанную зависимость обозначим через $\varphi_j(x_j)$.

Приведем в точке $x_j = \bar{x}_j$ касательную к кривой $\varphi_j(x_j)$. Обозначим отрезок этой касательной до пересечения с ординатой, проведенной в точке $(\bar{x}_j + \sigma_{x_j})$, через Δ_j . Проекция этого отрезка Δ_{jy} на ось y равна $\frac{\partial \varphi(\bar{x})}{\partial \bar{x}_j} \sigma_{x_j}$. Выполнив подобное построение для каждого из аргументов x_j ($j = \overline{1, n}$), замечаем, что формулу (30.10) для дисперсии можно переписать в виде

$$\sigma_y^2 = \sum_{j=1}^n (\Delta_{jy})^2. \quad (30.15)$$

Таким образом, получаем следующее правило: для нахождения дисперсии функции нескольких случайных аргументов

необходимо для каждого случайного аргумента определить отрезок касательной Δ_j , соответствующий приращению аргумента X_j от его значения $\bar{x}_j$ до значения $\bar{x}_j + \sigma_{xj}$ (считая остальные аргументы функции равными их математическим ожиданиям), и найти сумму квадратов проекций этих отрезков для различных значений $j = \overline{1, n}$ на ось y .

Так как проекции Δ_{jy} входят в окончательную формулу во второй степени, знак отрезка Δ_j не играет роли (то есть к значению можно или прибавить или отнять σ_{xj}), то эти отрезки не являются векторами в обычном смысле слова, хотя и характеризуются направлением, совпадающим с направлением касательной к кривой $\varphi_j(x_j)$ в точке $x = \bar{x}_j$. Поэтому эти отклонения иногда называют *векториальными*.

Рассмотрим числовые примеры.

Пример 30.1. Случайная величина X распределена нормально, причем $\bar{x} = 1$, $\sigma_x = 10^{-3}$. Пользуясь методом линеаризации, найти дисперсию $Y = X^3$. Сравнить полученный результат с точным значением дисперсии.

Решение. Применяя формулу (30.4), имеем

$$\sigma_y^2 \approx \left(\frac{\bar{x}^3}{d\bar{x}} \right)^2 \sigma_x^2 = 9 \cdot 10^{-6}.$$

Точное значение σ_y^2 может быть выражено через моменты μ_j рассматриваемой случайной величины. Действительно, по определению

$$\begin{aligned} \sigma_y^2 &= \mathbf{M}\{[X^3 - \mathbf{M}(X^3)]^2\} = \\ &= \mathbf{M}\{[(\overset{\circ}{X} + \bar{x})^3 - \mathbf{M}(\overset{\circ}{X}^3 + 3\bar{x}\overset{\circ}{X}^2 + 3\bar{x}^2\overset{\circ}{X} + \\ &\quad + \bar{x}^3)]^2\} = \mu_6^{(x)} - 9\bar{x}^2\sigma_x^4 + 15\bar{x}^2\mu_4^{(x)} + 9\bar{x}^4\sigma^2, \end{aligned}$$

где учтено, что нечетные центральные моменты нормальной величины равны нулю. Так как для нормальной величины, как было показано в § 22

$$\mu_{2k} = \frac{(2k)!}{2^k k!} \sigma_x^{2k}, \quad \mu_6^{(x)} = 10\sigma_x^6, \quad \mu_4^{(x)} = 3\sigma_x^4$$

и для точного значения σ_y^2 получим

$$\begin{aligned}\sigma_y^2 &= 9\sigma_x^4 \left(1 + \frac{2}{3}\sigma_x^2 + \frac{10}{9}\sigma_x^4 \right) = \\ &= 9 \cdot 10^{-6} \left(1 + \frac{2}{3} \cdot 10^{-6} + \frac{10}{9} \cdot 10^{-12} \right),\end{aligned}$$

что отличается от приближенного значения менее, чем на 0,0001%.

Пример 30.2. Пусть X — нормальная случайная величина, характеризующаяся $\bar{x} = 1$, $\sigma_x^2 = 10^{-4}$. Требуется определить σ_y для случайной величины $Y = 10 \sin \left(\frac{\pi}{2} X \right)$.

Решение. Так как

$$\left. \frac{d \left[\sin \left(\frac{\pi}{2} x \right) \right]}{dx} \right|_{x=\bar{x}} = \frac{\pi}{2} \cos \frac{\pi}{2} = 0,$$

то формулы (30.4) не могут быть использованы и, следовательно, необходимо воспользоваться формулами (30.6), учитывающими квадратичные члены в ряде Тейлора. Имеем

$$\begin{aligned}\sigma_y^2 &\approx \left(\frac{10\pi^2}{16} \right)^2 (\mu_4^{(x)} - \sigma_x^4) = 2 \left(\frac{10\pi^2}{16} \right)^2 \cdot 2\sigma_x^4, \\ \sigma_y &\approx \frac{5\pi^2}{4} \sigma_x^2 = \frac{5\pi^2}{4} \cdot 10^{-4} \approx 0,00123.\end{aligned}$$

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

§ 31. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Предельные теоремы представляют собой самостоятельный, и притом очень важный, раздел вероятностной теории. На эту тему публикуется много научных работ, издаются специальные монографии и учебные курсы, формулировки теорем постоянно уточняются и совершенствуются с тем, чтобы доказать их справедливость при возможно более широких и общих предположениях. При появлении новых разделов теории вероятностей, таких, например, как теория случайных функций или теория случайных полей, исследователи, как правило, предпринимают серьезные усилия, чтобы доказать и соответствующие предельные теоремы, без которых никакая вероятностная теория не может считаться в полной мере законченной.

Предельные теоремы включают в себе значительную часть прикладной ценности научных результатов, полученных теорией вероятности. Важнейшее их значение состоит в том, что эти теоремы открывают возможность *экспериментальной проверки* основных положений теории вероятностей.

Физический эксперимент является универсальным средством проверки правильности любой теории. При этом в *детерминированном* эксперименте все происходящие события (в вероятностной терминологии) относятся к двум видам: либо они достоверны, либо невозможны. Поэтому достаточно провести один эксперимент (опыт, испытание) и зарегистрировать

все произошедшие события. При повторении эксперимента результат будет идентичным.

Иначе обстоит дело в *случайном* эксперименте. Здесь появление или непоявление события в отдельно взятом опыте лишь частично характеризует это событие. Поскольку события в таком эксперименте сами являются *случайными*, то для их описания необходимо задать некоторую объективную количественную меру возможности их возникновения — *вероятность*. Одного опыта оказывается недостаточно, поскольку он несет слишком мало информации (разумеется, если событие не является практически невозможным или практически достоверным). Приходится проводить целую серию, состоящую из n опытов, причем желательно брать n как можно большим. При этом нужно стремиться к тому, чтобы испытания были независимыми, тогда каждое новое испытание будет нести в себе дополнительную информацию об исследуемом явлении.

Сказанное объясняет обычную постановку задачи в предельных теоремах: берется серия из неограниченного числа, как правило, независимых испытаний. В такой постановке в § 11 рассматривались простейшие предельные теоремы. Их отличительной особенностью было то, что они относились к опытам, касающимся *случайных событий*. Сейчас мы переходим к более сложным предельным теоремам, относящимся к схеме *случайных величин*.

Выше мы отметили предельный случай, когда стохастический эксперимент превращается в детерминированный. Он соответствует тому, что наблюдаемые события практически достоверны или практически невозможны. Такой предельный случай есть и в схеме случайных величин. Если наблюдаемая случайная величина X имеет пренебрежимо малую дисперсию σ_x^2 , то достаточно всего лишь одного опыта, чтобы составить мнение о ней. Критерием здесь является малость коэффициента вариации $\gamma_x \ll 1$. Иначе говоря, если $\sigma_x \ll |\bar{x}|$, то можно ограничиться небольшим числом повторений опыта. Если σ_x соизмеримо с $\bar{x}$, то следует стремиться увеличивать n и обеспечивать независимость испытаний, делая их более информативными, точно так же, как и в схеме испытаний со случайными событиями.

Различают два вида предельных теорем. К первому типу относятся так называемые *законы больших чисел*. Они устанавливают определенные соотношения между средним арифметическим результатом наблюдений за некоторой случайной величиной и математическим ожиданием этой величины. Ко второй группе причисляют так называемые *центральные предельные теоремы*. Последние трактуют предельные свойства закона распределения сумм большого числа случайных слагаемых. Как уже отмечалось в § 11, примером теорем первого вида могут служить теоремы Бернулли и Пуассона, а второго вида — теорема Муавра–Лапласа.

В этом параграфе мы ограничимся обсуждением основных классических теорем типа закона больших чисел. При этом, как мы увидим, ранее разобранные теоремы Бернулли и Пуассона оказываются простейшим частным случаем, следствием доказываемых сейчас теорем.

Предварительно докажем одно вспомогательное неравенство, впервые установленное П. Л. Чебышевым.

Неравенство Чебышева

Пусть случайная величина X имеет математическое ожидание $\bar{x}$ и дисперсию σ_x^2 . Тогда при любом $\varepsilon > 0$

$$\mathbf{P}(|X - \bar{x}| \geq \varepsilon) \leq \frac{\sigma_x^2}{\varepsilon^2}. \quad (31.1)$$

Доказательство

Доказательство проведем для случайной величины произвольного типа, не делая оговорок, что она непрерывная, хотя будем использовать ее плотность вероятности $f(x)$. Если X — дискретная или смешанного типа, то $f(x)$ будет включать дельтаобразные слагаемые (см. главу 4).

Запишем цепочку неравенств

$$\begin{aligned} \mathbf{P}(|X - \bar{x}| \geq \varepsilon) &= \\ &= \int_{|x - \bar{x}| \geq \varepsilon} f(x) dx \leq \int_{|x - \bar{x}| \geq \varepsilon} \frac{(x - \bar{x})^2}{\varepsilon^2} f(x) dx \leq \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx = \frac{\sigma_x^2}{\varepsilon^2}. \end{aligned} \quad (31.2)$$

В первой оценке под знак интеграла добавлен множитель, бóльший единицы, а во второй оценке интервал $\bar{x} - \varepsilon \leq x \leq \bar{x} + \varepsilon$ расширен до всей числовой оси.

Теперь перейдем к формулировке и доказательству самой теоремы.

Теорема Чебышева

Если $X_1, \dots, X_n, \dots$ представляет последовательность попарно независимых случайных величин, имеющих конечные математические ожидания $\bar{x}_j$ и конечные дисперсии σ_j^2 , ограниченные одной и той же постоянной

$$\sigma_j^2 \leq \sigma_{\max}^2, \quad (j = \overline{1, \infty}), \quad (31.3)$$

то тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P}(|\tilde{x} - \bar{x}| \geq \varepsilon) = 0, \quad (31.4)$$

где

$$\tilde{x} = \frac{1}{n} \sum_{j=1}^n x_j, \quad \bar{x} = \frac{1}{n} \sum_{j=1}^n \bar{x}_j. \quad (31.5)$$

Доказательство

Положим $X = \tilde{x}$. Применяя теоремы о математическом ожидании и дисперсии, получим

$$\begin{aligned} \mathbf{M}(X) &= \mathbf{M}(\tilde{x}) = \mathbf{M}\left(\frac{1}{n} \sum_{j=1}^n X_j\right) = \\ &= \frac{1}{n} \sum_{j=1}^n \mathbf{M}(X_j) = \frac{1}{n} \sum_{j=1}^n \bar{x}_j = \bar{x}, \end{aligned} \quad (31.6)$$

$$\begin{aligned} \mathbf{D}(X) &= \mathbf{D}(\tilde{x}) = \frac{1}{n^2} \mathbf{D}\left(\sum_{j=1}^n X_j\right) = \\ &= \frac{1}{n^2} \sum_{j=1}^n \mathbf{D}(X_j) = \frac{1}{n^2} \sum_{j=1}^n \sigma_j^2 \leq \frac{\sigma_{\max}^2}{n}. \end{aligned} \quad (31.7)$$

Далее по неравенству Чебышева

$$\mathbf{P}(|X - \bar{x}| \geq \varepsilon) \leq \frac{\sigma_{\max}^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0, \quad (31.8)$$

что и требовалось доказать.

Теорема Чебышева очень часто применяется в теории ошибок измерения. Допустим, что имеется некоторая физическая величина x_0 , подвергаемая n -кратному измерению. Измерения содержат некоторую систематическую ошибку, так что в j -м измерении они в среднем группируются относительно значения $\bar{x}_j$. Измерения неравноточны, причем точность j -го измерения характеризуется дисперсией σ_j^2 . Измерения считаются независимыми. Обозначим через Δ среднее значение систематической ошибки по всем наблюдениям и положим $\bar{x} = x_0 + \Delta$. Тогда среднее арифметическое $\tilde{x}$ измеренных значений X_j будет удовлетворять неравенству (31.8), причем вероятность (31.8) будет стремиться к нулю при неограниченном увеличении числа измерений.

Легко видеть, что сформулированная в § 11 теорема Бернулли является следствием теоремы Чебышева. Рассмотрим последовательность независимых случайных событий $\{A_j\}_{j=1}^{\infty}$, каждое из которых появляется в опыте с вероятностью p . Введем характеристические случайные величины событий A_j

$$X_j = I(A_j), \quad (j = \overline{1, \infty}), \quad (31.9)$$

где $I(A)$ определяется (16.1). Все эти величины одинаково распределены, причем

$$\begin{aligned} \bar{x}_j &= \mathbf{M}(X_j) = 1 \cdot p + 0 \cdot (1 - p) = p, \\ \sigma_j^2 &= (1 - p)^2 p + (0 - p)^2 (1 - p) = (1 - p)p. \end{aligned} \quad (31.10)$$

Очевидно, что частота появления события в серии испытаний может быть представлена в виде

$$\tilde{p} = \frac{1}{n} \sum_{j=1}^n X_j. \quad (31.11)$$

Поэтому применяя к последовательности $\{X_j\}_{j=1}^{\infty}$ теорему Чебышева, получаем требуемое.

Точно также доказывается и теорема Пуассона. Разница только в том, что события A_j будут появляться в опытах с

разными вероятностями p_j , так что

$$\begin{aligned}\bar{x}_j &= 1 \cdot p_j + 0 \cdot (1 - p_j) = p_j, \\ \sigma_j^2 &= (1 - p_j)^2 \cdot p_j + (0 - p_j)^2(1 - p_j) = (1 - p_j)p_j\end{aligned}\tag{31.12}$$

оказываются зависящими от номера испытаний. Поскольку все $\bar{x}_j$ по-прежнему ограничены (например, единицей) и все σ_j^2 также ограничены (например, одной четвертью), то теорема Чебышева опять дает нужный результат.

Если предположить, что все X_j *одинаково распределены*, то требования теоремы Чебышева можно несколько ослабить, ограничившись существованием только математического ожидания.

Теорема Хинчина

Если последовательность случайных величин $\{X_i\}_{i=1}^n$ образована попарно независимыми, одинаково распределенными величинами с конечным математическим ожиданием $\bar{x}$, то

$$\lim_{n \rightarrow \infty} \mathbf{P}(|\tilde{x} - \bar{x}| \geq \varepsilon) = 0.\tag{31.13}$$

Эту теорему можно доказать непосредственным применением теоремы Чебышева, если воспользоваться так называемым методом урезания [9]. Его сущность состоит в том, что случайные величины X_j , представляются в виде суммы двух величин, первая из которых совпадает с X_j в некотором конечном интервале, а вторая является добавкой к первой, чтобы сумма совпадала с X_j . При этом уровень урезания зависит от n .

Для экономии места мы не будем приводить это доказательство, отсылая читателя к [9], однако поясним идею доказательства, основанного на методе характеристических функций (см. § 20). Из существования математического ожидания $\bar{x}$ следует, что характеристическая функция X может быть представлена в виде

$$E_x(u) = 1 + iu\bar{x}[1 + \theta(u)],\tag{31.14}$$

где $\theta(u)$ — некоторая бесконечно малая величина первого порядка малости относительно u .

Как мы помним (см. (20.1)), характеристическая функция X определяется равенством

$$E_x(u) = \mathbf{M}(e^{iuX}). \quad (31.15)$$

Поэтому согласно определению (31.5)

$$E_{\tilde{x}}(u) = \left[E_x \left(\frac{1}{n} u \right) \right]^n. \quad (31.16)$$

Подстановка разложения (31.14) в (31.16) дает

$$E_{\tilde{x}}(u) = \left\{ 1 + \frac{iu}{n} \bar{x} \left[1 + \theta \left(\frac{u}{n} \right) \right] \right\}^n. \quad (31.17)$$

Если теперь перейти в этом равенстве к пределу при $n \rightarrow \infty$ и учесть, что $\theta(u)$ стремится к нулю при $u \rightarrow 0$, находим

$$\lim_{n \rightarrow \infty} E_{\tilde{x}}(u) = e^{iu\bar{x}}, \quad (31.18)$$

а это означает, что соответствующий предельный закон распределения величины $\tilde{x}$ является дельтаобразным

$$\lim_{n \rightarrow \infty} f_{\tilde{x}}(x) = \delta(x - \bar{x}), \quad (31.19)$$

то есть в пределе среднее арифметическое таких X_j превращается в детерминированную величину $\bar{x}$.

Еще одно обобщение теоремы Чебышева связано с возможностью рассмотрения зависимых величин X_j .

Теорема Маркова

Если последовательность $\{X_i\}_{i=1}^{\infty}$ такова, что

$$\lim_{u \rightarrow \infty} \frac{1}{n^2} \mathbf{D} \left(\sum_{j=1}^n X_j \right) = 0, \quad (31.20)$$

то для любого $\varepsilon > 0$ выполняется (31.4), где $\tilde{x}$ и $\bar{x}$ задаются (31.5).

Доказательство этой теоремы является очевидным следствием неравенства Чебышева, и поэтому мы предоставляем

возможность читателю выполнить его самостоятельно. Отметим лишь, что теорема Чебышева является частным случаем теоремы Маркова. Действительно, в случае попарной независимости X_j по теоремам о дисперсии имеем

$$D(\sum_{j=1}^n X_j) = \sum_{j=1}^n \sigma_j^2, \quad (31.21)$$

после чего повторяются все рассуждения доказательства теоремы Чебышева.

Отметим еще одно интересное обобщение и уточнение теоремы Чебышева.

Теорема Бернштейна

Допустим, что все случайные величины последовательности $\{X_i\}_{i=1}^{\infty}$ попарно независимы, характеризуются одной и той же дисперсией σ^2 , а их центральные моменты удовлетворяют неравенствам

$$\mu_k \leq \frac{1}{2} k! H^{k-2} \sigma^2, \quad (k = \overline{2, \infty}), \quad (31.22)$$

где H — некоторая постоянная, не зависящая от k . Тогда при любом $\varepsilon > 0$.

$$P(|\tilde{x} - \bar{x}| \geq \varepsilon) \leq 2e^{-\frac{n\varepsilon^2}{4\sigma^2}}. \quad (31.23)$$

Смысл условия (31.22) заключается в том, что центральные моменты случайных величин должны не слишком быстро возрастать. Если это требование выполнено, то для вероятности (31.23) получается гораздо более точная оценка, чем, например, в неравенстве Чебышева (31.1), которое при $X = \tilde{x}$ дает результат, всего лишь обратно пропорциональный числу испытаний.

В заключение сделаем следующее терминологическое замечание. Если последовательность случайных величин $\{X_n\}_{n=1}^{\infty}$ по отношению к некоторой другой случайной величине X_* , заданной на том же вероятностном пространстве $(\Omega, \mathfrak{A}, P)$, что и величины X_j , обладает тем свойством, что

$$\lim_{n \rightarrow \infty} P(|X_n - X_*| \geq \varepsilon) = 0 \quad (31.24)$$

при любом $\varepsilon > 0$, то говорят, что *последовательность* $\{X_n\}_{n=1}^\infty$ *сходится к* X_* *по вероятности*.

В этой терминологии все упомянутые варианты закона больших чисел сводятся к тому, что среднее арифметическое $\tilde{x}$ наблюдавшихся значений X_j сходится по вероятности к среднему арифметическому математических ожиданий $\bar{x}$. При этом говорят о законе больших чисел в форме Чебышева, в форме Маркова и так далее.

Иногда также рассматривают и иной вид сходимости, а именно *сходимость почти наверное*. Рассмотрим последовательность случайных величин $\{X_n(\omega)\}_{n=1}^\infty$, заданных на одном и том же вероятностном пространстве $(\Omega, \mathfrak{A}, P)$, а также предельную величину $X_*(\omega)$. Обозначим через Ω_n множество тех элементарных событий ω , для которых имеем $X_n(\omega) \neq X_*(\omega)$. Тогда, если

$$P(\Omega_n) \xrightarrow{n \rightarrow \infty} 0, \quad (31.25)$$

то говорят, что X_n сходится к x_* *почти наверное*. В свою очередь, если для некоторой последовательности случайных величин оказывается, что $\tilde{x} \rightarrow \bar{x}$ почти наверное, то говорят, что данная последовательность удовлетворяет *усиленному закону больших чисел* [9].

Рассмотрим примеры применения закона больших чисел.

Пример 31.1. Производится суммирование 100 вещественных чисел, каждое из которых округляется до ближайшего целого. Ошибка округления равномерно распределена в промежутке $[-0,5, 0,5]$. Оценить вероятность того, что эта ошибка превзойдет 15 единиц, с помощью неравенства Чебышева.

Решение. По условию все складываемые величины равномерно распределены со средним $\bar{x}_j$ и параметром $l_j = 0,5$. Как показано в главе 5 для этого случая

$$\sigma_j^2 = \frac{1}{3} l_j^2 = \frac{1}{12} = \text{const}(j) = \sigma^2.$$

Рассмотрим среднее арифметическое

$$\tilde{x} = \frac{1}{100} \sum_{j=1}^{100} X_j.$$

Ясно, что $\bar{x} \equiv 0$, а величина $\varepsilon = \frac{15}{100} = 0,15$. Применяя неравенство Чебышева к $\tilde{x}$, получаем

$$\mathbf{P}(|\tilde{x} - \bar{x}| \geq \varepsilon) \leq \frac{\sigma_x^2}{\varepsilon^2 \cdot n} = \frac{1/12}{(0,15)^2 \cdot 100} = \frac{1}{27} \approx 0,0370.$$

Пример 31.2 (Продолжение примера 31.1). Решить задачу из примера 31.1 с помощью теоремы Бернштейна.

Решение. Как было показано в главе 5, для равномерного закона степенная оценка вида (31.22) выполняется. Поэтому можно воспользоваться неравенством

$$\mathbf{P}(|\tilde{x} - \bar{x}| \geq \varepsilon) = \mathbf{P}(|\tilde{x}| \geq 0,15) \leq 2e^{-\frac{n\varepsilon^2}{4\sigma_x^2}}.$$

Подставляя в это равенство

$$n = 100, \quad \varepsilon = 0,15, \quad \sigma^2 = \frac{1}{12},$$

получаем

$$\mathbf{P}(|\tilde{x}| \geq 0,15) \leq 2e^{-\frac{100 \cdot (0,15)^2}{4 \cdot 1/12}} = 2 \cdot e^{-6,75} \approx 0,00468.$$

Таким образом, по Бернштейну оценка улучшилась в десять раз.

§ 32. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Напомним вначале основное содержание доказанной в § 11 теоремы Муавра–Лапласа. Если переформулировать ее в терминах случайных величин, то указанную теорему можно выразить в виде предельного соотношения

$$\mathbf{P} \left(a \leq \frac{\sum_{j=1}^n (X_j - \bar{x}_j)}{\sqrt{\sum_{j=1}^n \sigma_j^2}} < b \right) \xrightarrow{n \rightarrow \infty} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad (32.1)$$

где X_j — независимые случайные величины, принимающие два значения — 0 и 1 с вероятностями

$$\mathbf{P}(X_j = 0) = 1 - p, \quad \mathbf{P}(X_j = 1) = p, \quad (j = \overline{1, n}), \quad (32.2)$$

$$\begin{aligned} \bar{x}_j &= \mathbf{M}(X_j) = 0 \cdot (1 - p) + 1 \cdot p, \\ \sigma_j^2 &= \mathbf{D}(X_j) = (0 - p)^2(1 - p) + (1 - p)^2p = p(1 - p). \end{aligned} \quad (32.3)$$

Иначе говоря, при таком специальном выборе случайных величин X_j оказывается, что указанным образом центрированная и нормированная сумма X_j в пределе при $n \rightarrow \infty$ асимптотически характеризуется нормальным законом распределения.

При этом возникает естественный вопрос: насколько существенным является для выполнения (32.1) именно такой конкретный выбор свойств случайных величин X_j ? Эта проблема, поставленная в свое время П. Л. Чебышевым, была впоследствии решена его учениками А. А. Марковым и А. М. Ляпуновым¹. Как оказалось, соотношения вида (32.1) имеют место лишь при самых общих ограничениях на свойства случайных величин X_j , смысл которых сводится к тому, что ни одно из слагаемых не должно превалировать в общей сумме.

Причина, по которой этот результат приобрел столь важное значение для приложений, лежит в самой сути массовых явлений, происходящих в природе и технических системах. Дело в том, что реальные физические процессы протекают обычно под влиянием большого числа независимо действующих факторов, которые обычно можно трактовать, как случайные, причем каждый из них способен лишь в небольшой степени повлиять на исход процесса в целом. Наблюдается же лишь интегральный эффект действия всех факторов в целом, и согласно центральной предельной теореме этот эффект выражается в нормальности отклика системы на весь спектр приложенных к ней случайных возмущений.

Рассмотрим n независимых случайных величин X_j , каждая из которых имеет конечное математическое ожидание $\bar{x}_j$ и конечную дисперсию σ_j^2 , причем эти величины обладают свойством, не позволяющим какой-либо одной из них преобладать

¹ Ляпунов Александр Михайлович (1857–1918) — российский математик и механик.

в сумме. Ляпунов показал, что если сформулировать новую величину

$$Z = \frac{\sum_{j=1}^n (X_j - \bar{x}_j)}{\sqrt{\sum_{j=1}^n \sigma_j^2}}, \quad (32.4)$$

то она будет при $n \rightarrow \infty$ асимптотически нормальна с нулевым математическим ожиданием и единичной дисперсией при выполнении особого условия.

Обозначим через σ знаменатель выражения (32.4)

$$\sigma = \sqrt{\sum_{j=1}^n \sigma_j^2}, \quad (32.5)$$

а через $\beta_j^{(s)}$ — следующий абсолютный центральный момент X_j

$$\beta_j^{(s)} = \mathbf{M}[|X_j - \bar{x}_j|^s]. \quad (32.6)$$

Тогда условие Ляпунова сводится к тому, что должно существовать такое $\delta > 0$, что

$$\lim_{n \rightarrow \infty} \frac{\sum_{j=1}^n \beta_j^{(\delta+2)}}{\sigma^{\delta+2}} = 0. \quad (32.7)$$

Идея доказательства Ляпунова сводится к следующему. Предположим для простоты, что все X_j одинаково распределены, и обозначим через $E_0(u)$ характеристическую функцию центрированных величин $\overset{\circ}{X}_j = \bar{X}_j - \bar{x}_j$. Тогда аналогично (31.16) можно записать для характеристической функции величины Z , определяемой (32.4)

$$E_z(u) = \left[E_0 \left(\frac{u}{\sigma} \right) \right]^n = \left[E_0 \left(\frac{u}{\sigma_0 \sqrt{n}} \right) \right]^n, \quad (32.8)$$

где квадрат величины

$$\sigma_0 = \frac{1}{\sqrt{n}} \sigma = \sqrt{\frac{1}{n} \sum_{j=1}^n \sigma_j^2} \quad (32.9)$$

обозначает среднее значение дисперсии X_j . Для одинаково распределенной величины σ_0^2 совпадает с дисперсией любого из X_j .

Представим характеристическую функцию каждого из X_j в виде отрезка степенного разложения

$$E_0(u) = 1 + \frac{1}{2} \sigma_0^2 (iu)^2 + u^2 \theta(u), \quad (32.10)$$

где $\theta(u)$ является бесконечно малой величиной относительно u при $u \rightarrow 0$. Линейный член в разложении (32.10) отсутствует, потому что характеристическая функция $E_0(u)$ соответствует центрированной величине $\tilde{X}_j = X_j - \bar{x}_j$.

Подставляя разложение (32.10) в (32.8), будем иметь

$$E_z(u) = \left\{ 1 - \frac{u^2}{\sigma_0^2 n} \left[\frac{1}{2} \sigma_0^2 - \theta \left(\frac{u}{\sigma_0 \sqrt{n}} \right) \right] \right\}^n \quad (32.11)$$

или после перехода к пределу при $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} E_z(u) = e^{-\frac{u^2}{2}}, \quad (32.12)$$

а это соответствует нормальному закону распределения с единичной дисперсией и нулевым математическим ожиданием.

Отметим, что именно для доказательства этой теоремы Ляпунов и ввел характеристические функции. Ему удалось доказать (32.12) не только для случая одинаково распределенных X_j , но также и в общем случае произвольных X_j , удовлетворяющих условию (32.7). Доказательство является достаточно громоздким и здесь не приводится.

Требование отсутствия доминирования в последовательности $\{X_n\}_{n=1}^\infty$ одной случайной величины или группы величин (как конечной, так и бесконечной по численности) является

обязательным для выполнения центральной предельной теоремы. Приведенное выше *условие Ляпунова* является лишь одним из возможных вариантов постановки такого рода ограничений. Уточнением и доработкой условий занимались впоследствии многие ученые.

Это объяснялось важностью и прикладной значимостью предельных теорем вообще. Кроме того, ставилась задача получить наиболее общие, широкие условия, менее всего ограничивающие свойства последовательности $\{X_n\}_{n=1}^{\infty}$. Удачная постановка условий могла несколько упростить доказательство теоремы, которое у Ляпунова было достаточно сложным. Было желательно также найти наиболее наглядные, физически понятные условия.

Наиболее удобным из имеющихся вариантов дополнительного ограничения в центральной предельной теореме является *условие Линдеберга*², предложенное этим финским математиком в 1922 году. Условие Линдеберга требует, чтобы при любом $\tau > 0$

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma^2} \sum_{j=1}^n \int_{|x_j - \bar{x}_j| > \tau \sigma} (x_j - \bar{x}_j)^2 f(x_j) dx_j = 0, \quad (32.13)$$

где, по-прежнему

$$\bar{x}_j = \mathbf{M}(X_j), \quad \sigma^2 = \sum_{j=1}^n \mathbf{D}(X_j). \quad (32.14)$$

Выясним смысл условия (32.13). Для этого введем случайные события A_j , состоящие в том, что $|x_j - \bar{x}_j| > \tau \sigma$. Рассуждая аналогично доказательству неравенства Чебышева, получаем

$$\begin{aligned} \mathbf{P}(A_j) &= \mathbf{P}(|X_j - \bar{x}_j| > \tau \sigma) = \int_{|x_j - \bar{x}_j| > \tau \sigma} f(x_j) dx_j \leq \\ &\leq \frac{1}{\tau^2 \sigma^2} \int_{|x_j - \bar{x}_j| > \tau \sigma} (x_j - \bar{x}_j)^2 f(x_j) dx_j. \end{aligned} \quad (32.15)$$

² Линдеберг Ярл Вольдемар (1876–1932) — финский математик.

Затем введем событие B , означающее, что наибольшая из разностей $|x_j - \bar{x}_j| > \tau\sigma$. Очевидно, что

$$\begin{aligned} \mathbf{P}(B) &= \mathbf{P}\left(\max_{1 \leq j \leq n} |X_j - \bar{x}_j| > \tau\sigma\right) = \\ &= \mathbf{P}\left(\sum_{j=1}^n A_j\right) \leq \sum_{j=1}^n \mathbf{P}(A_j), \quad (32.16) \end{aligned}$$

причем в последнем нестрогом неравенстве знак равенства достигается только тогда, когда все события A_j несовместны.

Из последнего неравенства следует, что

$$\begin{aligned} \mathbf{P}\left(\max_{1 \leq j \leq n} |X_j - \bar{x}_j| > \tau\sigma\right) &\leq \\ &\leq \frac{1}{\sigma^2 \tau^2} \sum_{j=1}^n \int_{|x_j - \bar{x}_j| > \tau\sigma} (x_j - \bar{x}_j)^2 f(x_j) dx_j. \quad (32.17) \end{aligned}$$

В соответствии с условием Линдеберга правая часть (32.16) стремится к нулю при любом фиксированном $\tau > 0$. Следовательно, условие Линдеберга представляет требование равномерной малости всех центрированных величин $\overset{\circ}{X}_j = X_j - \bar{x}_j$ относительно σ .

Прежде чем доказывать теорему, установим ряд вспомогательных неравенств:

$$\begin{aligned} |e^{i\alpha} - 1| &\leq \alpha, \\ |e^{i\alpha} - i\alpha - 1| &\leq \frac{\alpha^2}{2!}, \\ |e^{i\alpha} - \frac{1}{2}(i\alpha)^2 - i\alpha - 1| &\leq \frac{\alpha^3}{3!}. \end{aligned} \quad (32.18)$$

Первое неравенство доказывается элементарно

$$|e^{i\alpha} - 1| = \left| \int_0^\alpha (e^{ix} - 1) dx \right| \leq \int_0^\alpha |e^{ix}| dx = \int_0^\alpha 1 dx = \alpha. \quad (32.19)$$

Для доказательства второго неравенства его левая часть выражается с помощью интеграла аналогично (32.18), после чего применяется первое неравенство

$$\begin{aligned} |e^{i\alpha} - i\alpha - 1| &= \left| \frac{1}{i} \int_0^\alpha (e^{ix} - 1) dx \right| = \\ &= \left| \int_0^\alpha (e^{ix} - 1) dx \right| \leq \frac{\alpha^2}{2}. \end{aligned} \quad (32.20)$$

Точно также доказывается и третье неравенство, только подынтегральная функция здесь берется равной $e^{ix} - 1 - ix$.

Перейдем к доказательству самой теоремы. Для сокращения записей примем, что все величины X_j центрированы ($\bar{x}_j = 0$, $j = \overline{1, \infty}$). Это допущение не является принципиальным и легко может быть отброшено. При этом меняются только пределы интегрирования в интегралах типа (32.13).

Теорема Ляпунова. Если последовательность взаимно независимых центрированных случайных величин $\{X_j\}_{j=1}^\infty$ удовлетворяет условию

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma^2} \sum_{j=1}^n \int_{|x_j| > \tau \sigma} x_j^2 f(x_j) dx_j = 0 \quad (32.21)$$

при любом постоянном $\tau > 0$, где

$$\sigma = \sqrt{\sum_{j=1}^n \sigma_j^2}, \quad (32.22)$$

то тогда при $n \rightarrow \infty$ случайная величина

$$Z = \frac{1}{\sigma} \sum_{j=1}^n X_j \quad (32.23)$$

асимптотически распределена по центрированному и нормированному нормальному закону

$$\lim_{n \rightarrow \infty} f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \quad (32.24)$$

Доказательство. Закону (32.24) соответствует характеристическая функция (см. § 22)

$$\lim_{n \rightarrow \infty} E_z(u) = e^{-\frac{u^2}{2}}, \quad (32.25)$$

что эквивалентно

$$\lim_{n \rightarrow \infty} \ln E_z(u) = -\frac{u^2}{2}. \quad (32.26)$$

Благодаря независимости X_j по свойствам характеристических функций (см. § 20) можно записать

$$E_z(u) = \prod_{j=1}^n E_j\left(\frac{u}{\sigma}\right). \quad (32.27)$$

Следовательно, необходимо показать, что

$$\lim_{n \rightarrow \infty} \ln(E_z(u)) = -\frac{u^2}{2} \quad (32.28)$$

или

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n \ln \left(E_j \left(\frac{u}{\sigma} \right) \right) = -\frac{u^2}{2}. \quad (32.29)$$

Прологарифмировав равенство (32.26) и воспользовавшись рядом для логарифма, имеем

$$\begin{aligned} \ln E_z(u) &= \sum_{j=1}^n \ln \left(E_j \left(\frac{u}{\sigma} \right) \right) = \\ &= \sum_{j=1}^n \ln \left(1 + \left(E_j \left(\frac{u}{\sigma} \right) - 1 \right) \right) = \\ &= \sum_{j=1}^n \left(E_j \left(\frac{u}{\sigma} \right) - 1 \right) + R_n. \end{aligned} \quad (32.30)$$

Поскольку

$$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}, \quad |x| < 1,$$

то здесь

$$R_n = \sum_{j=1}^n \sum_{k=2}^{\infty} \frac{(-1)^{k-1}}{k} \left(E_j \left(\frac{u}{\sigma} \right) - 1 \right)^k. \quad (32.31)$$

Оценим разность $E_j \left(\frac{u}{\sigma} \right) - 1$. Вспоминая определение характеристической функции

$$E_j(u) = \int_{-\infty}^{\infty} e^{iux_j} f_j(x_j) dx_j, \quad (32.32)$$

можно записать

$$\begin{aligned} E_j \left(\frac{u}{\sigma} \right) - 1 &= \int_{-\infty}^{\infty} (e^{i \frac{u}{\sigma} x_j} - 1) f_j(x_j) dx_j = \\ &= \int_{-\infty}^{\infty} (e^{i \frac{u}{\sigma} x_j} - 1 - i \frac{u}{\sigma} x_j) f_j(x_j) dx_j. \end{aligned}$$

Третье слагаемое в скобках здесь можно добавить, так как $\bar{x}_j = 0$.

Используя второе неравенство (32.17), находим

$$\left| E_j \left(\frac{u}{\sigma} \right) - 1 \right| \leq \frac{u^2}{2\sigma^2} \int_{-\infty}^{\infty} x_j^2 f_j(x_j) dx_j, \quad (32.33)$$

и эта величина, очевидно, стремится к нулю, когда $n \rightarrow \infty$. Действительно, разбивая интеграл (32.33) на два интеграла

$$\begin{aligned} & \frac{1}{\sigma^2} \int_{-\infty}^{\infty} x_j^2 f_j(x_j) dx_j = \\ & = \frac{1}{\sigma^2} \left[\int_{|x_j| \leq \tau\sigma} x_j^2 f_j(x_j) dx_j + \int_{|x_j| > \tau\sigma} x_j^2 f_j(x_j) dx_j \right], \quad (32.34) \end{aligned}$$

обнаруживаем, что первое слагаемое не превосходит τ^2 , а второе также может быть сделано меньше τ^2 при достаточно большом n благодаря выполнению условия Линдберга (32.20). Поскольку τ — любое, сколь угодно малое, то отсюда следует, что

$$\lim_{n \rightarrow \infty} E_j \left(\frac{u}{\sigma} \right) = 1, \quad (32.35)$$

и, значит, ряд (32.30) сходится, по крайней мере, при достаточно больших значениях n .

При таких n его сумма может быть оценена по формуле для суммы бесконечно убывающей геометрической прогрессии

$$\begin{aligned} |R_n| & \leq \sum_{j=1}^n \sum_{k=2}^{\infty} \left| E_j \left(\frac{u}{\sigma} \right) - 1 \right|^k = \\ & = \frac{1}{2} \sum_{j=1}^n \frac{\left| E_j \left(\frac{u}{\sigma} \right) - 1 \right|^2}{1 - \left| E_j \left(\frac{u}{\sigma} \right) - 1 \right|}. \quad (32.36) \end{aligned}$$

Возьмем n достаточно большим, для того, чтобы выполнялось неравенство

$$\left| E_j \left(\frac{u}{\sigma} \right) - 1 \right| < \frac{1}{2}, \quad (32.37)$$

тогда

$$|R_n| \leq \sum_{j=1}^n \left| E_j \left(\frac{u}{\sigma} \right) - 1 \right|^2. \quad (32.38)$$

Продолжая оценку (32.38) с помощью оценки (32.33), в которой интеграл в правой части заменим на σ_j^2 , находим, учитывая (32.21)

$$\begin{aligned} |R_n| &\leq \max_{1 \leq j \leq n} \left| E_j \left(\frac{u}{\sigma} \right) - 1 \right| \sum_{j=1}^n \frac{u^2 \sigma_j^2}{2\sigma^2} = \\ &= \max_{1 \leq j \leq n} \left| E_j \left(\frac{u}{\sigma} \right) - 1 \right| \cdot \frac{u^2}{2}, \end{aligned} \quad (32.39)$$

причем правая часть последнего неравенства стремится к нулю согласно (32.35).

Теперь вернемся к сумме (32.29), в которой, как мы установили, $R_n \rightarrow 0$. Представим главный член этой суммы в виде

$$\sum_{j=1}^n \left(E_j \left(\frac{u}{\sigma} \right) - 1 \right) = -\frac{u^2}{2} + R_n^*, \quad (32.40)$$

где согласно (32.32)

$$R_n^* = \frac{u^2}{2} + \sum_{j=1}^n \int_{-\infty}^{\infty} \left(e^{i \frac{u}{\sigma} x_j} - 1 - i \frac{u}{\sigma} x_j \right) f_j(x_j) dx_j. \quad (32.41)$$

Вспоминая определение (32.21) для σ , это выражение можно переписать в виде

$$R_n^* = \sum_{j=1}^n \int_{-\infty}^{\infty} \left(e^{i \frac{u}{\sigma} x_j} - 1 - i \frac{u}{\sigma} x_j - \left(i \frac{u}{\sigma} x_j \right)^2 \right) f_j(x_j) dx_j, \quad (32.42)$$

а тогда с помощью третьего неравенства (32.17) получим

$$|R_n^*| \leq \sum_{j=1}^n \frac{u^3}{6\sigma^3} \int_{-\infty}^{\infty} |x_j|^3 f_j(x_j) dx_j. \quad (32.43)$$

Оценивая далее интегралы (32.43) тем же способом, как это было сделано для интегралов (32.34), убеждаемся, что все они стремятся, будучи поделенными на σ^3 , к нулю, а тогда $R_n^* \rightarrow 0$, что и завершает доказательство теоремы.

В заключение заметим, что условие Ляпунова является более жестким, чем условие Линдеберга. Если выполнено условие Ляпунова, то условие Линдеберга из него следует автоматически, но бывают ситуации, когда второе из указанных условий имеет место, а первое нарушается. Действительно, пусть при некотором $\delta > 0$ выполнено условие Ляпунова (32.7), тогда при любом $\varepsilon > 0$ имеем

$$\begin{aligned} \frac{1}{\sigma^2} \sum_{j=1}^n \int_{|x_j| \geq \varepsilon \sigma} x_j^2 f_j(x_j) dx_j &\leq \\ &\leq \frac{1}{\varepsilon^\delta \sigma^{2+\delta}} \sum_{j=1}^n \int_{|x_j| \geq \varepsilon \sigma} |x_j|^{2+\delta} f_j(x_j) dx_j \leq \\ &\leq \frac{1}{\varepsilon^\delta} \frac{\sum_{j=1}^n \beta_j^{(\delta+2)}}{\sigma^{\delta+2}} \xrightarrow{n \rightarrow \infty} 0, \end{aligned} \quad (32.44)$$

то есть выполняется и условие Линдеберга (32.21). Рассмотрим некоторые примеры применения центральной предельной теоремы.

Пример 32.1 (Продолжение примеров 31.1 и 32.2).

Складываются 100 случайных чисел, каждое из которых округляется до ближайшего целого, причем ошибки округления считаются независимыми случайными величинами, равномерно распределенными в промежутке $[-1/2, 1/2]$. Определить вероятность того, что суммарная ошибка превзойдет 15 единиц.

Решение. Поскольку все X_j одинаково распределены, то условия центральной предельной теоремы выполнены, поэтому сумма

$$X = \sum_{j=1}^{100} X_j$$

будет асимптотически нормальна ($n = 100$ можно считать достаточно большим). Для каждого числа $\bar{x}_j = 0$, а дисперсия равномерного закона с параметром $l = \frac{1}{2}$ дается равенством $\sigma_j^2 = \frac{1}{3} l^2$. Поэтому

$$\bar{x} = a, \quad \sigma_x = \sqrt{\frac{1}{3} l^2 n} = \sqrt{\frac{1}{3} \cdot \frac{1}{4} \cdot 100} = \frac{5}{\sqrt{3}}.$$

Для нормального закона

$$\begin{aligned} \mathbf{P}(|x| \geq 15) &= 1 - \mathbf{P}(|x| < 15) = 1 - \Phi\left(\frac{15\sqrt{3}}{5}\right) = \\ &= 1 - \Phi(3\sqrt{3}) \approx 1 - 0,99999 = 0,00001, \end{aligned}$$

то есть значительно точнее, чем по теореме Бернштейна, не говоря уже об оценке с помощью неравенства Чебышева. Но следует помнить, что все-таки данная оценка является приближенной, поскольку нормальность в центральной предельной теореме является только лишь асимптотической и, строго говоря, имеет место лишь в пределе при $n \rightarrow \infty$.

Пример 32.2. Сколько нужно произвести выстрелов, чтобы стрелок, который поражает мишень с вероятностью 95%, с вероятностью 0,8 промахнулся бы не менее 5 раз.

Решение. Обозначим через A_j событие, состоящее в том, что стрелок промахнется при j -м выстреле. Тогда случайные величины X_j определяются, как характеристические случайные величины событий A_j , то есть

$$X_j = I(A_j), \quad (j = \overline{1, n}),$$

причем

$$\bar{x}_j = p, \quad \sigma_j^2 = np(1-p),$$

где по условию $p = 0,05$. Общее число промахов

$$X = \sum_{j=1}^n X_j$$

характеризуется

$$\bar{x} = np, \quad \sigma_x^2 = np(1-p)$$

и по теореме является асимптотически нормальным. Искомая вероятность

$$\mathbf{P}(X \geq 5) = \frac{1}{2} \left[\Phi(\infty) - \Phi\left(\frac{5 - \bar{x}}{\sigma_x}\right) \right] = 0,8,$$

откуда находим

$$\Phi\left(\frac{5 - 0,05n}{\sqrt{0,0475n}}\right) = -0,6.$$

По таблицам функции Лапласа определяем

$$\frac{s - 0,05n}{\sqrt{0,0475n}} = -0,8416.$$

Следовательно, n определяется квадратным уравнением относительно $x = \sqrt{n}$ следующего вида:

$$0,05n - 0,183\sqrt{n} - 5 = 0,$$

решая которое, находим $\sqrt{n} \approx 12$, $n \approx 144$.

ЛИТЕРАТУРА

1. Абрамовиц М., Стиган И. Справочник по специальным функциям. М.: Наука, 1979.
2. Александров П. С. Введение в общую теорию множеств и функций. М.: Гостехиздат, 1948.
3. Бернштейн С. Н. Опыт аксиоматического обоснования теории вероятностей // Сообщения Харьковского математического общества, сер. 2, 1917, т. 15, с. 209–274.
4. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. М.: Наука, 1965.
5. Гасс С. Линейное программирование. М.: Мир, 1964.
6. Гельфанд И. М., Шилев Г. Е. Обобщенные функции и действия над ними. М.: Гостехиздат, 1959.
7. Гильберт Д. Избранные труды. М.: Факториал пресс, 1998, т. 2.
8. Гихман И. И. Скороход А. В., Ядренко М. И. Теория вероятностей и математическая статистика. Киев: Вища школа, 1979.
9. Гнеденко Б. В. Курс теории вероятностей. М.: Наука, 1969.
10. Гантмахер Ф. Р. Теория матриц. М.: Наука, 1967.
11. Кендалл М., Моран П. Геометрические вероятности. М.: Наука, 1972.
12. Колмогоров А. Н. Основные понятия теории вероятностей. М.–Л., ОНТИ, 1936.
13. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. М.: Наука, 1972.
14. Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф. Справочник по теории вероятностей и математической статистике. Киев: Наукова думка, 1978.
15. Крылов В. И. Приближенное вычисление интегралов. М.: Физматиздат, 1969.

16. Курош А. Г. Курс высшей алгебры. СПб.: Лань, 2008.
17. Лебедев Н. Н. Специальные функции и их приложения. СПб.: Лань, 2010.
18. Лойцянский Л. Г., Лурье А. И. Курс теоретической механики. М.: Гостехиздат, 1958.
19. Митропольский А. К. Интеграл вероятностей. Л.: ЛГУ, 1972.
20. Прохоров Ю. В. Розанов Ю. А. Теория вероятностей. М.: Наука, 1973.
21. Сборник задач по теории вероятностей, математической статистике и теории случайных функций. Под общей редакцией А. А. Свешникова. СПб.: Лань, 2008.
22. Свешников А. А. Прикладные методы теории случайных функций. СПб.: Лань, 2011.
23. Свешников А. А. Прикладные методы теории марковских процессов. СПб.: Лань, 2007.
24. Свешников А. Г., Тихонов А. Н. Теория функций комплексной переменной. М.: Наука, 1970.
25. Свешников А. А. Теория вероятностей и случайных функций. Случайные события. Л.: ЛПИ, 1979.
26. Свешников А. А. Теория вероятностей и случайных функций. Случайные величины. Л.: ЛПИ, 1980.
27. Свешников А. А. Теория вероятностей и случайных функций. Системы случайных величин. Л.: ЛПИ, 1982.
28. Титчмарш Э. Ч. Введение в теорию интегралов Фурье. М.: Гостехиздат, 1948.
29. Халмош П. Теория меры. М.: ИЛ, 1953.
30. Шумягский Б. М. Таблицы для решений кубического уравнения. М.: Гостехиздат, 1950.
31. Чжун К. Л. Однородные цепи Маркова. М.: Мир, 1964.

ИМЕННОЙ УКАЗАТЕЛЬ

- Байес** Т. (Bayes T.) 84
Башелье Л. (Bachelier L.) 10
Бернулли Я. (Bernoulli J.) 7, 13, 92, 104–105, 117, 119, 122, 138, 141, 168, 213, 214, 251, 291
Бертран Ж. (Bertrand J.L.F.) 38–41
Бернштейн С. Н. 12, 163, 445, 447, 459
Броун Р. (Brown R.) 10
Буняковский В. Я. 9
Бюффон Ж. (Buffon G.L.) 42, 43
- Венн** Дж. (Venn J.) 18
Вейбулл В. (Weibull E.H.W.) 324
- Гаусс** К. Ф. (Gauss C.F.) 9, 291
Гильберт Д. (Hilbert D.) 162–163
Грам Й. (Gram J.P.) 300
Гюйгенс Х. (Huygens Ch.) 6, 28
- Дирак** П. (Dirac P.A.M.) 234
- Кантор** Г. (Cantor G.) 185
Колмогоров А. Н. 10, 12, 161, 163, 167
Коши О. (Cauchy A.L.) 308
Крамп Х. (Kramp Ch.) 297
Кронекер Л. (Cronecker L.) 384
- Лаплас** П. С. (Laplace P.S.) 9, 12, 13, 139, 141–146, 460
Лейбниц Г. (Leibnitz G.W.) 367
Линдеберг Я. В. (Lindeberg J.W.) 452, 453, 458
Лоран П. А. (Laurent P.A.) 125
- Ляпунов** А. М. 9, 448, 449, 453
- Максвелл** Д. К. (Maxwell J.C.) 316
Марков А. А. 9, 444, 448
Мизес Р. (Mises R.) 14, 163
Морган де А. (de Morgan A.) 25, 176, 179
Муавр де А. (de Moivre A.) 146, 291, 440
- Паскаль** Б. (Pascal B.) 6, 28, 283
Планк М. (Planck M.) 10
Пуанкаре А. (Poincare H.J.) 41, 63
Пуассон С. (Poisson S.D.) 9, 98, 120, 256, 275–278, 283–285, 290
- Релей** Дж. (Rayleigh J.) 313
- Слуцкий** Е. Е. 11
Стирлинг Д. (Stirling J.) 103, 109, 144
- Ферма** П. (Fermat P.) 6, 28
Фишер Р. (Fisher R.A.) 424
- Хинчин** А. Я. 11, 443
- Чебышев** П. Л. 9, 440–441, 447–448, 459
- Шарлье** К. (Charlier C.L.) 300
- Эйнштейн** А. (Einstein A.) 10
Эрмит Ш. (Hermit Ch.) 299

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аксиома неотрицательности** 192
- нормировки 192
 - сложения 192, 193
 - — расширенная 193
- Аксиоматика Колмогорова** 12, 161, 163, 164, 166, 192–199
- теории вероятностей 12, 14, 161
- Алгебра множеств** 179
- событий 164, 172–188
 - — , порождённая заданным событием 180
 - — — полной системой событий 180–182
 - — — разбиением интервала 182, 183
 - — тривиальная 180
- Асимметрия (скошенность)** 248, 252, 300, 323
- Байеса формула** 84, 85, 334, 336
- Бернулли блуждания** 101–103
- испытания 92–100, 119
 - — неоднородные 96, 121, 151–152, 252, 375, 376
 - — рекуррентные 104–106
 - — сложные 108–114
 - теорема 140, 141
 - числа 307
- Бертрана парадоксы** 38–41
- Бесселя функции** 374
- Больших чисел закон** 139–141, 151–152, 438–447
- — — в форме Бернулли 140–141
 - — — — Чебышева 441
 - — — — Пуассона 151–152
 - — — усиленный 446
- Броуновское движение** 10, 171
- Бюффона задача** 42–43
- Величина случайная** 7–9, 219–220
- — гауссовская 235, 288–303
 - — дискретная 202–206, 222
 - — непрерывная 229
 - — смешанного типа 232, 236
 - — характеристическая (см. индикатор случайного события)
 - — центрированная 246, 261, 300, 384, 450
- Величины случайные зависимые** 350
- — — коррелированные 364
 - — — независимые 351
 - — — некоррелированные 371
- Вероятнейшее число появления события** 95
- Вероятностное пространство** 165, 188
- Вероятность апостериорная** 84
- априорная 84
 - безусловная 49–54
 - геометрическая 36–49
 - гипотез 74
 - полная 74–76
 - попадания в интервал 231

- , аксиоматическое определение 12, 13, 188
- , геометрическое определение 36–38, 51, 167, 190–191
- , классическое определение 7, 28–32, 51, 167, 189–190
- , частотное определение 14
- Включение событий 17
- Включения-исключения формула 61–63
- Гамма-распределение** 320–323
- Гамма-функция** 322
 - — неполная 320
- Геометрическая вероятность** 36–49
- Гипотезы** 74, 80, 84–91
- Грама-Шарлье ряд** 300
- Группа событий** полная 21, 22, 28, 74, 84, 86, 185
- Движение броуновское** 10
- Дельта-функция** 235, 262, 398–399
- Диаграмма Венна** 18, 19
- Дисперсия** 247, 370
 - , свойства 370–371
 - условная 392
- Единственно возможные исходы** 28
- Задача об игле Бюффона** 42–43
 - — очереди в билетную кассу 34–36
 - о встрече 44, 178
 - — жюри 69–74, 102, 111–113
 - — разорении 128–130, 211–213
 - — случайном блуждании частицы 101–103
 - — совпадениях 65–68
 - — составлении треугольника 44–45
 - — «счастливых» билетах 134–135
- — хорде окружности Бертрана 38–40
- Закон больших чисел** (см. больших чисел закон)
- распределения случайной величины (см. также на соответствующее распределение)
- — — дифференциальный (см. плотность вероятности)
- — — интегральный (см. функция распределения)
- — — условный (см. условная плотность вероятности)
- переместительный 21
- распределительный 21
- сочетательный 21
- Игла Бюффона** 42–43
- Игра вероятностная** 79–81
- Измеримость** 219
- Инвертирование (инверсия) события** 21–27
- Индикатор случайного события** 203, 214, 215, 218
- Интеграл вероятностей** 143, 148–151, 294, 297, 460
- Интегральная предельная теорема Лапласа** 146
- Интерпретация случайного события** 173
- Испытание** 16, 164–166
 - Бернулли (см. Бернулли испытания)
- Исход испытания** 164–165
 - — элементарный 28, 37
 - — — , благоприятствующий событию 29
- Квантиль** 251
- Композиция законов распределения** 402–404

- Корреляционная матрица 363, 376–377
Корреляционное отношение 372
Корреляционный момент 363
Коэффициент корреляции 365
Кroneкера символ 384
Кумулянты (семиинварианты) 266–267
- Лапласа функция 143, 148–151, 294, 297, 460
— теорема 141–145
Линдеберга условие 451–458
Линеаризация функции от случайной величины 429–436
Линия регрессии 391
Локальная предельная теорема 145
Ляпунова условие 449, 458
- Марковская цепь 9, 76, 84
Массовые явления 14
Математическое ожидание 240, 368
— — свойства 368–369
— — условное 391
Медиана 250
Мера вероятностная 38, 186, 187
— — конечно-аддитивная 186–187
— — нормированная 187
— — счётно-аддитивная 186–187
— — точности 292
Метод линеаризации 429–436
— производящих функций 114–127, 206–209
Мода 250
Модель испытаний, связанных в цепь 84
Момент случайной величины начальный 245, 265, 361–362
— — — центральный 245, 265–266, 270, 362
— — — факториальный 255
Мощность множества 185
- Независимость случайных событий** 49–54
— — — в совокупности 53–54, 57–59
— — — величин 351
— — — взаимная 53
— — — попарная 53, 57–59
- Операции над случайными событиями** 16–22, 174–179
Определение вероятности
— — аксиоматическое 12, 13, 188
— — геометрическое 36–38, 51, 167, 190–191
— — классическое 7, 28–32, 51, 167, 190–191
— — частотное (статистическое) 14
Опыт (см. испытание)
Оси рассеяния 390–391
Отклонение срединное 250, 296, 298, 306, 312, 313
— среднеквадратическое 247, 306, 313
Отношение корреляционное 372
Отражение события 173, 176
Отрицательное биномиальное распределение 279–280
Оценка вероятности в схеме Бернулли 154–157
- Парадокс петербургской игры** 241–242
Парадоксы Бертрана 38–41
Перестановочная (симметричная) группа событий 41
Плотность вероятности случайной величины 213
— — совместная системы случайных величин 347

- — условная 351, 393
- Полной вероятности формула 74–76
- Полная группа событий 21–22, 28, 74, 84–86, 185
- Правило «трёх сигм» 296
- Произведение случайных величин 407
- событий 17, 19
- Пространство вероятностное 165, 188
- элементарных событий 164–170
- Пуанкаре формула 61–63
- Пуассона приближение биномиального распределения 97–98
- теорема 151–152
- Равенство событий** 17
- Равновозможные исходы 28, 38, 41
- Разность событий 20
- — симметрическая 73, 198
- Распределение арксинуса 237, 311–313, 330
- биномиальное 94, 122, 251, 274–275, 282–285
- Вейбулла 324–326, 333
- гамма 320–323, 332
- геометрическое 257, 279–280, 283–285
- гипергеометрическое 281
- дискретное 202–217
- Коши 308–310, 331
- Максвелла 316–319, 331
- мультиномиальное 108–114, 123–127
- непрерывное 327
- нормальное 235, 288–303, 329
- — многомерное 377–396
- — усечённое 326–328
- отрицательное биномиальное 279–280
- полиномиальное 108–114, 123–127
- Пуассона 98, 120, 256, 275–278, 283–285, 329
- Релея 313–316, 330
- Стьюдента 420
- равномерное дискретное 253, 272–274, 283–285
- — непрерывное 234, 303–308, 329
- Фишера 426–428
- хи-квадрат 415–420
- экспоненциальное (показательное) 323–324, 332
- Регрессия 390
- Ряд распределения 202–206
- Свёртка законов распределения** (см. композиция законов распределения)
- Семиинварианты 266–267, 270
- Сигма-алгебра 183
- — максимальная 183–184
- — минимальная 186
- Событие 16
- достоверное 16, 19, 29
- наблюдаемое 172
- невозможное 16, 19, 29
- практически достоверное 29, 30
- практически невозможное 29, 30
- противоположное 21
- случайное 6, 16, 166
- элементарное 28, 166
- События зависимые 51
- независимые 51
- — в совокупности 53–54, 57–59
- — взаимно 53, 57–59
- — попарно 53, 57–59

- несовместные (непересекающиеся) 19, 20, 28, 59
- равновозможные 21, 22, 28
- совместные (пересекающиеся) 19, 59
- тождественные (равные) 17
- Сумма случайных величин 402–403
- — событий 18, 19

Теорема Бернулли 140–141

- Бернштейна 445
- Маркова 444
- Муавра–Лапласа 145–146
- Пуассона 151–152
- сложения вероятностей 59–63
- умножения вероятностей 51–54
- Хинчина 443
- центральная предельная в форме Лапласа 141–145
- — — — — Линдеберга 451–458
- — — — — Ляпунова 453
- Чебышева 441
- Типы случайных величин 229

Условная плотность вероятности 351, 385, 390

- дисперсия 391–393

Условное математическое ожидание 391–393

- Условие Линдеберга 451–458
- Ляпунова 449, 458

Формула Байеса 84–85, 334, 336

- включения-исключения 61–63
- де Моргана 25–27
- инверсии 25–27
- полной вероятности 74–76, 335
- Пуанкаре 61–63

- сложения вероятностей (см. теорема сложения вероятностей)
- Стирлинга 103, 109, 144
- умножения вероятностей (см. теорема умножения вероятностей)
- Функция Лапласа 143, 148–151, 294, 297, 460
- ошибок (см. функция Лапласа)
- производящая 115, 206–209, 249
- распределения 204, 220, 223–228
- регрессии 390
- — многомерная 343–346

Характеристическая функция 258, 356–360**Хи-квадрат распределение 415–420****Хинчина теорема 443****Центральная предельная теорема 447–460**

- — — в форме Лапласа 141–145, 447–448
- — — — — Линдеберга 451–458

- — — — — Ляпунова 453

- — — для зависимых величин 444

Центрированная случайная величина 246, 261, 300, 384, 450**Цепь Маркова 9, 84****Частное случайных величин 405–406****Частость события 14, 31****Чебышева неравенство 440–441**

- теорема 441

Число кардинальное 185

- случайное 33–34, 46–49,

82–83, 131–133, 215–217

Чистая стратегия 81

Эксперимент

детерминированный

164–165

— стохастический(случайный)

164–166

Эксцесс 248, 300

Элемент вероятности 232

Эллипс рассеяния 390–391

Явление массовое 14, 448

— случайное 448

Якобиан 409–411

ОГЛАВЛЕНИЕ

<i>Предисловие редактора</i>	3
<i>Введение</i>	6
<i>Глава 1. Элементарная теория случайных событий</i>	16
§ 1. Соотношения между случайными событиями	16
§ 2. Классическое определение вероятности	28
§ 3. Геометрические вероятности	36
§ 4. Условная вероятность. Теорема умножения вероятностей. Независимость	49
§ 5. Теорема сложения вероятностей. Несовместные события . .	59
§ 6. Формула полной вероятности	74
§ 7. Теорема гипотез (формула Байеса)	84
<i>Глава 2. Вычисление вероятности появления события при повторных испытаниях</i>	92
§ 8. Вычисление вероятностей появления события в серии независимых испытаний с двумя возможными исходами . .	92
§ 9. Вычисление вероятностей появления событий в серии испытаний с числом возможных исходов, большим двух . .	106
§ 10. Применение производящих функций и рекуррентных уравнений для вычисления вероятностей сложных событий	114
§ 11. Предельные теоремы теории вероятностей	139
<i>Глава 3. Аксиоматическое построение теории случайных событий</i>	161
§ 12. Аксиоматика Колмогорова. Понятие пространства элементарных событий	161
§ 13. Алгебра и сигма-алгебра событий. Аксиоматическое определение вероятности	172
§ 14. Классические модели вероятностного пространства	188
§ 15. Аксиоматический вывод основных положений теории случайных событий	192
<i>Глава 4. Случайные величины и их распределения вероятностей</i>	202
§ 16. Ряд распределения и функция распределения дискретной случайной величины	202
§ 17. Аксиоматическое определение случайных величин	218
§ 18. Случайные величины непрерывного и смешанного типа. Функция распределения и плотность вероятности	229
§ 19. Математическое ожидание случайной величины. Моменты и другие параметры закона распределения	240
§ 20. Характеристическая функция. Семиинварианты	258

<i>Глава 5. Некоторые наиболее употребительные типовые законы</i>	
распределения	272
§ 21. Примеры законов распределения	
дискретных случайных величин	272
21.1. Равномерное распределение	272
21.2. Биномиальный закон распределения	274
21.3. Закон Пуассона	275
21.4. Геометрическое распределение	279
21.5. Отрицательное биномиальное распределение	279
21.6. Гипергеометрическое распределение	281
§ 22. Нормальный закон распределения	288
§ 23. Другие примеры законов распределения	
непрерывных случайных величин	303
23.1. Равномерное распределение	303
23.2. Распределение Коши	308
23.3. Распределение арксинуса	311
23.4. Распределение Рэлея	313
23.5. Распределение Максвелла	316
23.6. Гамма-распределение	320
23.7. Экспоненциальное распределение	323
23.8. Распределение Вейбулла	324
23.9. Усеченное нормальное распределение	327
§ 24. Формула полной вероятности и формула Байеса	
в схеме случайных величин	334
<i>Глава 6. Системы случайных величин</i>	343
§ 25. Функция распределения и плотность вероятности	
системы случайных величин	343
§ 26. Характеристическая функция и моменты системы	
случайных величин	356
§ 27. Многомерный нормальный закон распределения	377
<i>Глава 7. Функциональные преобразования случайных величин</i>	397
§ 28. Функции случайных величин	397
28.1. Общие положения	397
28.2. Функция одной случайной величины	400
28.3. Сумма двух случайных величин	402
28.4. Частное двух случайных величин	405
28.5. Произведение двух случайных величин	407
28.6. Система случайных величин, функционально	
связанных с заданной системой величин	407
§ 29. Законы распределения, связанные с нормальным	
законом распределения	415
29.1. Законы распределения χ и χ^2	415
29.2. Закон распределения Стюдента	420
29.3. Закон распределения Фишера	424
§ 30. Метод линеаризации функций случайных величин	429
<i>Глава 8. Предельные теоремы</i>	438
§ 31. Закон больших чисел	438
§ 32. Центральная предельная теорема	447
<i>Литература</i>	461
<i>Именной указатель</i>	463
<i>Предметный указатель</i>	464

Арам Арутюнович СВЕШНИКОВ

**ПРИКЛАДНЫЕ МЕТОДЫ
ТЕОРИИ ВЕРОЯТНОСТЕЙ**
У ч е б н и к

Зав. редакцией
физико-математической литературы *О. А. Митрофанова*
Художественный редактор *С. Ю. Малахов*
Корректор *А. М. Плетьева*
Верстка *А. Г. Сандомирская*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812) 412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 20.02.12.
Бумага офсетная. Формат 84×108¹/₃₂. Печать офсетная.
Усл. п. л. 25,20. Тираж 1500 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Издательско-полиграфическое предприятие «Правда Севера».
163002, г. Архангельск, пр. Новгородский, д. 32.
Тел./факс (8182) 64-14-54; www.ippps.ru