УЧЕБНИКИ
ВЫСШЕЙ ШКОЛЫ ЭКОНОМИКИ

ВШЭ
HSE

А.С.Шведов

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
и математическая статистика

Второе издание,
переработанное и дополненное

Издательский дом ГУ ВШЭ

Москва 2005

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Теория вероятностей относится к тем разделам математики, которые находят самое широкое применение на практике. Одна из областей, где используется теория вероятности, — это статистические исследования.

Значительная часть информации поступает к нам в виде чисел. Собрать данные и представить их в форме таблиц, графиков, диаграмм или компьютерных файлов очень важно. Рисунки и отдельные наиболее значимые числа дают ответы на большинство вопросов. Но многое из того, что содержится в числовой информации, остается скрытым и может быть выявлено только при помощи методов математической статистики.

Для чего могут использоваться статистические данные? Как извлечь из них нужную нам информацию?

Статистические данные могут быть использованы для описания и анализа экономической системы, построения прогнозов и выработки решений. Еще одно направление — это оценка возможностей выбора (оценка опционов) на финансовых рынках.

Говоря о построении прогнозов, мы, в первую очередь, имеем в виду количественные, а не качественные прогнозы. Качественные прогнозы — это ответы на вопросы, произойдет либо нет то или другое событие, или какой из возможных вариантов осуществится. Количественные прогнозы — это предсказания каких-то числовых характеристик для определенных будущих моментов или отрезков времени.

Анализ является неотъемлемой составной частью прогноза и в то же время представляет самостоятельный интерес. В этой книге речь, в основном, пойдет именно о традиционных методах статистического анализа. Методы использования статистической информации для выработки решений очень разнообразны. Однако наиболее интересные и важные подходы по своей сложности выходят за пределы этой книги и потому в ней не отражены. Некоторые примеры использования статистических данных для выработки решений приведены в параграфе 3.4.

Многие идеи высшей математики применяются при работе с числовой экономической информацией. Они превратились в инструменты, с которыми полезно уметь обращаться любому экономисту. Никто не спорит с тем, что в жизни эти инструменты нужны биржевому аналитику больше, чем директору магазина, специалисту по опционам и фьючерсам — больше, чем налоговому инспектору. Однако практически всем экономистам приходится постоянно иметь дело с числовой информацией, и понимание того, что может и чего не может дать математика при работе со статистическими данными, безусловно, полезно. Широкое распространение компьютеров обеспечило возможность применения программ, в которых реализованы методы математической статистики. И конечно, только опыт работы может научить, какой из методов и когда следует использовать.

Применение инструментов математической статистики подразумевает, что в распоряжении мастера есть достаточно большой числовой материал, т.е. набор данных, выражающих одну или несколько характеристик экономической системы для одного или нескольких моментов или отрезков времени. Это десятки, сотни, а иногда и значительно большие массивы чисел.

Инструменты математической статистики достаточно сильно отличаются друг от друга и по назначению, и по принципам действия. Цель данной книги состоит в том, чтобы рассказать, на чем основаны эти инструменты, и привести некоторые примеры их использования.

По преданию, один король попросил своего придворного мудреца обучить его математике.

— Я чувствую, что знание математики может принести мне пользу, — сказал король. — Но только у меня слишком мало свободного времени, и мне хотелось бы узнать все, что нужно, как-нибудь побыстрее.

— Ваше величество, — ответил мудрец, — к сожалению, в математике не существует королевских путей.

Фраза о том, что в математике не существует королевских путей, стала крылатой. Математику нельзя выучить ни за два часа, ни за двадцать часов.

Студенты-экономисты иногда испытывают такое же желание, как тот король:

— Нас окружает масса статистической информации, и мы чувствуем, что математика может помочь нам как-то разобраться во всех этих числах и научиться работать с ними. Пожалуйста, изложите только самую суть, не надо лишних деталей.

Но именно разбираясь в деталях, человек приобретает математическую культуру, и именно эта культура, а не хранящийся в памяти набор формулировок, должна стать основным результатом обучения. Без нее использование математических инструментов будет формальным и малоэффективным.

Большинство математических методов, используемых при работе с числовой экономической информацией, относится к математической теории, называемой теорией вероятностей. Эта теория возникла в XVII в. в работах французских математиков Б. Паскаля и П. Ферма. Многие важные результаты были получены в XVIII и XIX вв. В число основателей теории вероятностей входят Я. Бернулли, П. Лаплас, К. Гаусс, П.Л. Чебышев. Но только в XX в. теория вероятностей была превращена в безупречно строгую математическую теорию советским математиком А.Н. Колмогоровым, который применил идеи французского математика А. Лебега, относящиеся к теории интегрирования, для обоснования теории вероятностей. После появления работ А.Н. Колмогорова стало ясно, что использовавшиеся ранее в теории вероятностей на полуинтуитивном уровне понятия (и имеющие свои названия) являются общепринятыми в математике. Поэтому существует некоторое дублирование терминов, используемых в теории вероятностей и во всей остальной математике (об этом мы будем говорить в главе 1). В настоящее время в теории вероятностей принято использовать старые традиционные термины, удобные для приложений, но наполненные строгим математическим содержанием. К сожалению, конструкции А.Н. Колмогорова достаточно сложны. Студенты-математики изучают их обычно не раньше третьего курса. В этом учебнике, предназначенном для экономистов, они будут изложены коротко и в очень упрощенном виде.

Под математической статистикой, с точки зрения математиков, понимается применение методов теории вероятностей при работе со статистической информацией. По теории вероятностей и математической статистике есть много хороших учебников как для студентов-математиков, так и для студентов-нематематиков (см. список литературы в конце учебного пособия). Настоящее учебное пособие задумывалось как азбука. Оно должно быть написано просто и понятно, в ущерб общности и полноте, говоря условно, крупными буквами с картинками, где это только возможно. Но в то же время оно должно давать фундамент для последующей работы.

Основам теории вероятностей, которые необходимы для понимания учебного курса, посвящена глава 1. Наиболее важным здесь является понятие случайной величины, на котором строится все дальнейшее изложение. В главе 2 разбираются некоторые задачи комбинаторики, помогающие понять, откуда возникает закон нормального распределения вероятностей — краеугольный камень математической статистики. Теория вероятностей занимает относительно небольшую часть учебного пособия — основная его часть посвящена статистическим методам. Поэтому отбор материала, относящегося к теории вероятностей, был произведен очень жестко. Как и в любых других разделах теоретической математики, непосредственно выходящих на приложения, одни понятия и теоремы теории вероятностей используются в приложениях непосредственно, а другие нужны для того, чтобы доказать используемые теоремы. При отборе материала нами был сделан выбор в пользу понятий и теорем первого вида. Мы стремились к тому, чтобы, затратив минимальное время, читатель усвоил основные принципы применения теории вероятностей для работы со статистической информацией. Слабой стороной такого способа изложения является то, что многие важные теоремы приведены без доказательств. (О том, где можно найти доказательства этих теорем, сказано в библиографической справке.)

Главы 3 и 4 посвящены определению нормально распределенных случайных величин и различным приложениям нормального распределения. В дальнейшем вводятся другие законы распределения вероятностей и сразу же даются примеры их применения: χ^2-распределение (глава 5), t-распределение Стьюдента (глава 6), F-распределение Фишера (глава 7), распределение Пуассона (глава 8). Все эти распределения, за исключением распределения Пуассона, строятся на основе нормального распределения.

Главы 9, 10 и 11 посвящены другому важному разделу математической статистики — регрессии и корреляции. Данный математический аппарат необходим для исследования взаимного влияния и взаимной зависимости динамических рядов и наборов наблюдений.

Изложение ведется совсем просто в начале книги и постепенно усложняется в последующем. Начиная с параграфа 1.3 предполагается, что читатель знаком с основами математического анализа. В главе 10 используются операции над матрицами.

Изменения по сравнению с первым изданием носят, в основном, редакционный характер. Заменены некоторые примеры. Несколько расширены параграфы, посвященные условным вероятностям, проверке гипотез и доверительным интервалам.

В основу учебного пособия положены лекции, которые автор читал студентам Государственного университета — Высшей школы экономики в течение ряда лет. При этом необходимо подчеркнуть, что данная книга остается все же учебным пособием, ряд важных разделов курса в книге не представлен. Автор благодарит коллектив ГУ ВШЭ за помощь в организации курса и всех коллег, высказавших автору свои конструктивные и доброжелательные замечания. Отдельно автор выражает благодарность Э.Б. Ершову за большое число полезных замечаний, сделанных при подготовке еще первого издания книги.

1
глава

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ МНОЖЕСТВ И ТЕОРИИ ВЕРОЯТНОСТЕЙ

В этой главе вводится постоянно используемое в дальнейшем понятие случайной величины. Определяются также другие понятия: событие, вероятность, условная вероятность, ожидаемое значение (ожидание) случайной величины, дисперсия случайной величины, корреляция случайных величин.

1.1

Множества и функции

Строгость в математике является непременным условием. Все используемые понятия определяются, все утверждения, за исключением аксиом, доказываются. В каждом определении новые понятия опираются на понятия, введенные раньше. Но если так, то как же быть с самыми первыми определениями? Через что определяются вводимые в них понятия? Ответ на эти вопросы может быть только один. В математике есть неопределимые понятия. Такими понятиями являются множество и функция.

Прежде чем перейти к обсуждению того, что такое множество и что такое функция, подчеркнем, что роль этих понятий не ограничивается тем, что через них определяются все остальные понятия математики. Еще более важно то, что любая математическая теорема — это есть некоторое утверждение о множествах и (или) функциях.

Множество — это совокупность элементов, обладающих некоторым свойством.

Ясно, что это не определение, а пояснение. Слово "совокупность" ничем не лучше слова "множество". Примерами множеств могут быть:

множество людей, живущих в доме;

множество яблок, лежащих в корзине;

множество точек на плоскости.

Для двух множеств мы будем использовать специальные обозначения:

$\boldsymbol{N}$ — множество целых положительных чисел, т.е. множество, состоящее из чисел 1, 2, 3, 4, 5, 6 и т.д.;

$\boldsymbol{R}$ — множество действительных чисел, в которое входят все элементы множества N, а также отрицательные числа, дробные числа и иррациональные числа, например, такие, как $\sqrt{2}$, π и e.

Введем несколько символов, относящихся к множествам, необходимых нам для дальнейшего.

$$x \in A$$

(элемент x принадлежит множеству A).

Определение. Множество A принадлежит множеству B, если любой элемент множества A является Элементом множества B. В этом случае говорят, что A — *подмножество* B.

$$A \subset B$$

(множество A принадлежит множеству B; см. рис. 1.1).

В частности, $A \subset A$. Подмножество, состоящее из одного элемента x, обозначается $\{x\}$. Если $x \in A$, то $\{x\} \subset A$. Фигурные скобки используются для того, чтобы наглядно показать, из каких элементов состоит множество. Например,

$$\boldsymbol{N} = \{1, 2, 3, ...\}.$$

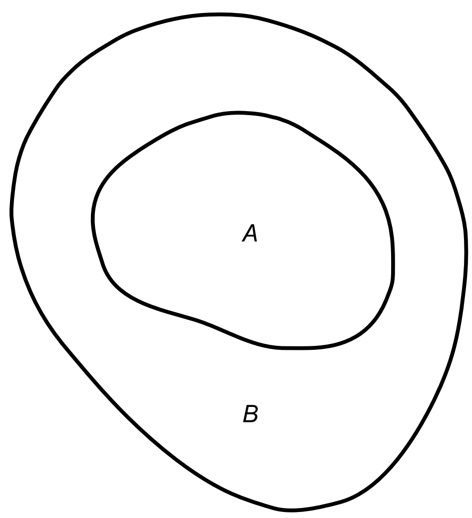


Рис. 1.1. Подмножество *A* множества *B*

Определение. Множество *C* называется *пересечением* множеств *A и B*, если множество *C* состоит из тех и только тех элементов, которые принадлежат и множеству *A,* и множеству *B*.

$$C = A \cap B$$

(C — пересечение A и B; см. рис. 1.2).

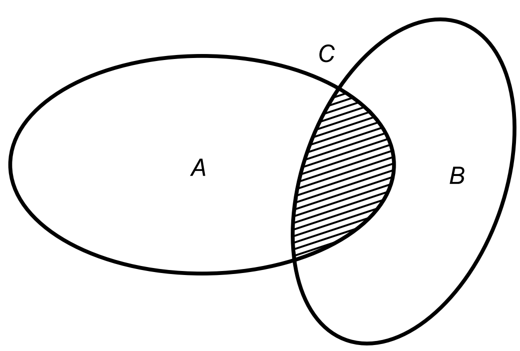


Рис. 1.2. Множество *C* — пересечение множеств *A* и *B*

Определение. Множество C называется *объединением* множеств A и B, если множество C состоит из тех и только тех элементов, которые принадлежат хотя бы одному из множеств A или B.

$$C = A \cup B$$

(C — объединение A и B; см. рис. 1.3).

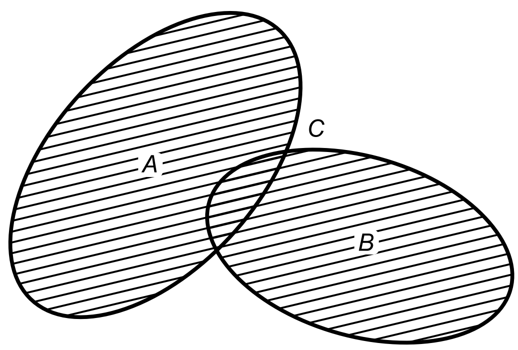


Рис. 1.3. Множество C — объединение множеств A и B

Множество называется *конечным,* если оно содержит конечное число элементов, и *бесконечным в* противоположном случае. Например, множество целых чисел, больших 0, но меньших 10, конечно (оно содержит 9 элементов), а множества $\boldsymbol{N}$ и $\boldsymbol{R}$ бесконечны.

Пример 1.1. Обозначим через [0,1] множество действительных чисел $\boldsymbol{x} \in \boldsymbol{R}$**,** удовлетворяющих условию

$$0 \le x \le 1.$$

Множество [0,1] бесконечно (хотя и ограничено).

При $a \le b$ множество точек $x \in \boldsymbol{R}$ таких, что

$$a \le x \le b,$$

называется *отрезком.* При $a < b$ множество точек $x \in \boldsymbol{R}$ таких, что

$$a < x < b,$$

называется *интервалом*.

До сих пор мы считали, что множество — это совокупность некоторых элементов. Теперь понятие множества нужно расширить. В корзине с яблоками может лежать одно яблоко, два, три и т.д. Но может не лежать ни одного. Так же можно рассмотреть совокупность, в которой нет элементов. Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом

$$\varnothing.$$

Пустое множество считается подмножеством (но не элементом!) любого множества.

Поясним на примере, почему понятие пустого множества является полезным и даже необходимым. Пусть множество A состоит из трех элементов

$$a_1,\ a_2,\ a_3.$$

Тогда у множества A есть восемь подмножеств:

$$\varnothing,\ \{a_1\},\ \{a_2\},\ \{a_3\},\ \{a_1,\ a_2\},\ \{a_1,\ a_3\},\ \{a_2,\ a_3\}\{a_1,\ a_2,\ a_3\}.$$

(В число подмножеств включаются пустое множество и само множество A.) Рассмотрим пересечение подмножеств $\{a_1, a_2\}$ и $\{a_1, a_3\}$:

$$\{a_1,\ a_2\} \cap \{a_1,\ a_3\} = \{a_1\}.$$

А что такое пересечение, например, $\{a_1, a_2\}$ и $\{a_3\}$? Если бы мы не ввели понятие пустого множества, то должны были бы признать, что в данном случае операция пересечения не определена, что очень неудобно. Но теперь ответ на поставленный вопрос не вызывает затруднений:

$$\{a_1,\ a_2\} \cap \{a_3\} = \varnothing.$$

Теперь можно перейти к обсуждению того, что такое функция.

Функция — это отображение, которое каждому элементу одного множества ставит в соответствие некоторый элемент другого множества.

Ясно, что это опять не определение, а пояснение. Слово “отображение” ничем не лучше слова “функция”. Запись

$$f : A \to B$$

означает, что функция f ставит в соответствие каждому элементу множества A некоторый элемент множества B, и читается так: "функция f определена на множестве A и принимает значения в множестве B" или так: "f действует из A в B". На рис. 1.4 и 1.5 показаны две функции; при этом каждое из множеств A и B состоит из трех элементов.

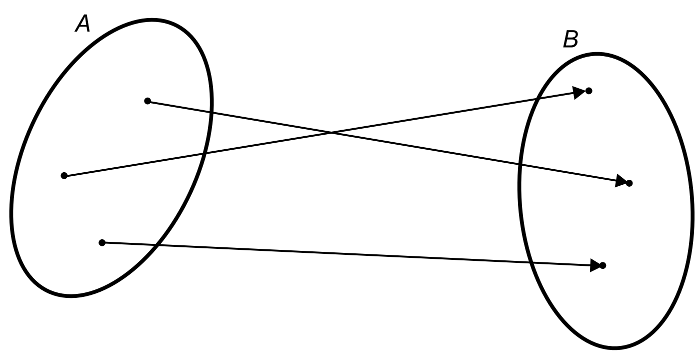


Рис. 1.4. Пример взаимно-однозначной функции

Функция, изображенная на рис. 1.4, называется *взаимно-однозначной:* каждому элементу множества A поставлен в соответствие свой элемент множества B, и каждый элемент множества B соответствует некоторому элементу множества A. Функция, изображенная на рис. 1.5, взаимно-однозначной не является.

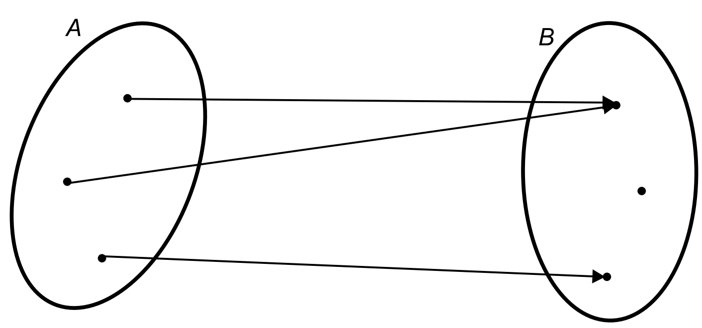


Рис. 1.5. Пример не взаимно-однозначной функции

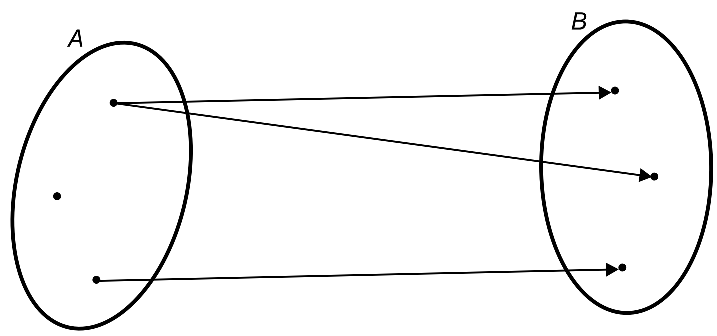


Рис. 1.6. Пример отображения, не являющегося функцией

Отображение, показанное на рис. 1.6, не является; функцией сразу по двум причинам. Во-первых, одному из элементов множества A поставлено в соответствие два элемента множества B, и, во-вторых, другому элементу множества A вообще ничего не поставлено в соответствие.

Если функция f действует из A в B, $f: A \to B$, и элемент x принадлежит множеству A, $x \in A$, то поставленный в соответствие x элемент множества B обозначается $f(x)$. Таким образом, $f(x) \in B$.

Определение. Множество пар $(x, f(x))$, где $x \in A$, называется *графиком* функции f.

З а м е ч а н и е. Иногда через $f(x)$ обозначается сама функция f, а не элемент множества значений. Как правило, это не приводит к недоразумениям, но, встретив такое выражение, необходимо понять, что имеется в виду.

Пример 1.2. Рассмотрим функцию

$$f: R \to [-1, 1],$$

действующую по правилу $f(x) = \sin x$. График этой функции (точнее, часть графика) изображен на рис. 1.7.

Еще раз подчеркнем, что график функции — это множество, а не функция. Не следует путать функцию с ее графиком.

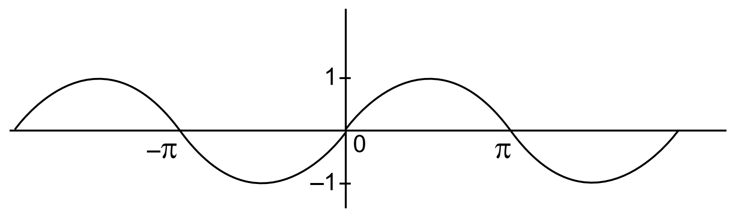


Рис. 1.7. График функции $f(x) = \sin x$

Если функция $f : A \to \boldsymbol{R}$ и функция $g : A \to \boldsymbol{R}$, то можно определить функцию, являющуюся суммой этих двух функций и обозначаемую $(f + g)$

$$(f + g) : A \to \boldsymbol{R},$$

которая каждому $x \in A$ ставит в соответствие число $f(x) + g(x)$. Аналогично можно определить и произведение функций. Если $c \in \boldsymbol{R}$, то могут быть определены функции cf и $(f + c)$. Функции, принимающие значения в $\boldsymbol{R}$, мы будем называть *числовыми* функциями.

1.2

Пространства элементарных событий и случайные величины. Независимость событий и случайных величин

В предисловии мы говорили, что только в XX в. теория вероятностей была превращена в безупречно строгую математическую теорию, и оказалось, что все используемые в теории вероятностей понятия аналогичны понятиям из других разделов математики. К этому времени в теории вероятностей сложилась своя терминология, которая осталась общепринятой, хотя можно было бы заменить существующие термины их общематематическими аналогами. Но в этом случае язык теории вероятностей стал бы менее понятен в приложениях. Соответствие между основными терминами приведено в следующей таблице.

Общематематический термин	*Термин теории вероятностей*
Множество событий	Пространство элементарных
Элемент	Элементарное событие
Подмножество	Событие
Длина, площадь, объем подмножества	Вероятность события
Числовая функция	Случайная величина

Приведенное соответствие нужно понимать следующим образом. Термином “пространство элементарных событий” мы будем обозначать некоторое множество, конечное или бесконечное. Элементы и подмножества этого множества будем называть соответственно элементарными событиями и событиями.

Если подмножества плоскости или пространства являются линиями, фигурами или телами, то для них определены длины, площади или объемы. Аналогично каждому событию (т.е. подмножеству пространства элементарных событий) ставится в соответствие число, называемое его вероятностью. Это число так же, как длина, площадь или объем является неотрицательным, и если разделить некоторое событие на два непересекающихся события, то вероятность исходного события будет равна сумме вероятностей полученных частей.

Случайная величина — это функция, определенная на пространстве элементарных событий, и принимающая значения в множестве $\boldsymbol{R}$.

З а м е ч а н и е. То, что название столь важного объекта состоит из двух слов “случайная величина”, а не из одного слова, очень нехорошо. Но на сегодняшний день это название является общепринятым.

Пространство элементарных событий (множество) принято обозначать греческой буквой Ω (омега). Сами элементарные события (элементы множества) обозначаются тоже буквами “омега”, но не прописными, а строчными:

$$\omega_1,\ \omega_2,\ \dots\ .$$

События обозначаются буквами A, B, ... То есть $\omega \in \Omega$, $A \subset \Omega$. Случайные величины чаще всего обозначаются буквами X, Y, ... Поскольку случайная величина — это числовая функция, определенная на множестве Ω, верна запись

$$X : \Omega \to \boldsymbol{R}.$$

Множество Ω может быть как конечным (т.е. содержать конечное число элементов), так и бесконечным. Для случая, когда множество Ω содержит конечное число элементов, мы построим теорию совершенно строго; для случая бесконечного множества Ω необходимо использовать значительно более сложные математические конструкции, поэтому мы ограничимся некоторыми пояснениями и не будем приводить все необходимые определения.

Итак, пусть множество Ω конечно и содержит N элементов:

$$\Omega = \{\omega_1, \omega_2, ..., \omega_N\}.$$

Предположим, что задано N положительных действительных чисел

$$p_1, p_2, \dots, p_N,$$

каждое из которых не больше 1. Пусть, кроме того,

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

(В дальнейшем мы часто будем использовать символ Σ, который обозначает сумму, в данном случае, N чисел.) Из перечисленных условий следует, что если $N > 1$, то

$$0 < p_i < 1,\ i = 1, ..., N.$$

Число p_i назовем *вероятностью элементарного события* ω_i. Если событие $A \subset \Omega$, то *вероятность события* A определим, как

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

где суммируются вероятности только тех элементарных событий ω_i, которые являются элементами подмножества A. Из этого определения ясно, что

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

Пустому множеству приписывается вероятность 0:

$$P(\varnothing) = 0.$$

Для любого другого события A

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

Следует признать, что в случае конечного пространства элементарных событий Ω аналогия между вероятностями событий и длинами, площадями, объемами подмножеств является достаточно далекой. Но нам необходимо будет рассматривать и бесконечные пространства элементарных событий Ω, и тогда данная аналогия будет очень полезна.

Пример 1.3. Предположим, что мы два раза бросали монету и каждый раз мог выпасть либо герб, либо решетка. Возможны следующие 4 исхода:

$$ГГ, ГР, РГ, РР.$$

Каждый из исходов назовем элементарным событием:

$$\Omega = \{ГГ, ГР, РГ, РР\}.$$

Если монета доброкачественная, то вероятности всех исходов равны:

$$p_1 = p_2 = p_3 = p_4 = \frac{1}{4}.$$

Пусть событие A заключается в том, что хотя бы один раз выпал герб, т.е.

$$A = \{ГГ, ГР, РГ\}.$$

Тогда вероятность события A

$$P(A) = \sum_{\omega_i \in A} p_i = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

Непосредственно из определения вероятности события следует, что если $A \cap B = \varnothing$, то

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

То же справедливо для любого числа событий. Если события A_1, A_2, ..., A_k попарно не пересекаются, то

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{i=1}^{k} P(A_i).$$

Данное утверждение называется *теоремой о сложении вероятностей.*

Основополагающую роль в дальнейшем будет играть следующее определение.

Определение. События A и B называются *независимыми,* если

$$P(A \cap B) = P(A)P(B).$$

Отметим, из определения независимости событий следует, что событие, являющееся пустым множеством, и любое другое событие независимы. Данное определение может быть обобщено на любое число событий.

Определение. События $A_1, A_2, ..., A_k$ называются независимыми, если

$$P\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} P(A_1).$$

Здесь мы ввели сразу два новых обозначения:

$\bigcap\limits_{i=1}^{k}$ — пересечение k множеств (т.е. множество, состоящее из тех и только тех элементов, которые принадлежат каждому из перечисленных множеств);

$\prod\limits_{i=1}^{k}$ — произведение k чисел.

Замечание 1. Используется также другое определение независимости k событий, когда требуется, чтобы при любом $m < k$ любые m событий из $A_1, A_2, ..., A_k$ были независимыми. То определение независимости k событий, которое используется нами, является более широким.

Замечание 2. Независимость событий для бесконечного множества Ω определяется так же, как и для конечного множества Ω.

Как показывает следующий пример, из попарной независимости нескольких событий не следует их независимость.

Пример 1.4. Пусть Ω состоит из 4 элементарных событий и вероятность каждого элементарного события равна $\frac{1}{4}$. Пусть

$$A_1 = \{\omega_1, \omega_2\},\ A_2 = \{\omega_1, \omega_3\},\ A_3 = \{\omega_1, \omega_4\}$$

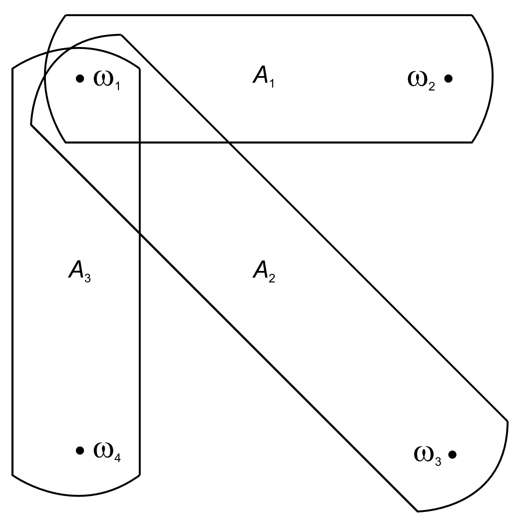


Рис. 1.8. Три события, любые два из которых независимы, а все вместе они не являются независимыми

(см. рис. 1.8). Тогда

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$$

и

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = P(\{\omega_1\}) = \frac{1}{4},$$

т.е. попарно между собой любые два из этих событий независимы. Но

$$P(A_1 \cap A_2 \cap A_3) = P(\{\omega_1\}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3).$$

То есть события A_1, A_2, A_3 не независимы.

В дальнейшем нам, в основном, будет нужно не само определение независимости событий, а построенное на нем определение независимости случайных величин. Пусть случайная величина

$$X : \Omega \to \boldsymbol{R}$$

принимает k различных значений $x_1, x_2, ..., x_k$ и случайная величина

$$Y : \Omega \to \boldsymbol{R}$$

принимает l различных значений $y_1, y_2, ..., y_l$. Поскольку множество Ω конечно и содержит всего N элементов, то

$$k \le N, \, l \le N.$$

Обозначим через A_i подмножество Ω, на котором случайная величина X равна x_i. Таким образом,

$$\Omega = \bigcup_{i=1}^{k} A_i.$$

Обозначим через B_j подмножество Ω, на котором случайная величина Y равна y_j,

$$\Omega = \bigcup_{j=1}^{l} B_j.$$

Определение. Случайные величины X и Y *независимы,* если при любых i и j ($1 \le i \le k$, $1 \le j \le l$) независимы события A_i и B_j.

Подчеркнем, что независимость событий и независимость случайных величин — это разные понятия. Использование в их названиях одного и того же слова не помогает усвоению предмета, но является общепринятым.

Поскольку о независимости событий можно говорить только в том случае, когда они являются подмножествами одного и того же пространства элементарных событий Ω, о независимости случайных величин можно говорить только тогда, когда они являются функциями, определенными на одном и том же пространстве элементарных событий Ω.

Пример 1.5. Допустим, что множество Ω состоит из kl точек, лежащих на плоскости в прямоугольнике $1 \le x \le k$, $1 \le y \le l$ и имеющих целые координаты (на рис. 1.9 изображено множество Ω при $k = 7$, $l = 4$). Будем считать, что вероятность, соответствующая каждой точке, равна $\frac{1}{kl}$. Определим случайную величину X, как абсциссу точки, а случайную величину Y — как ординату точки.

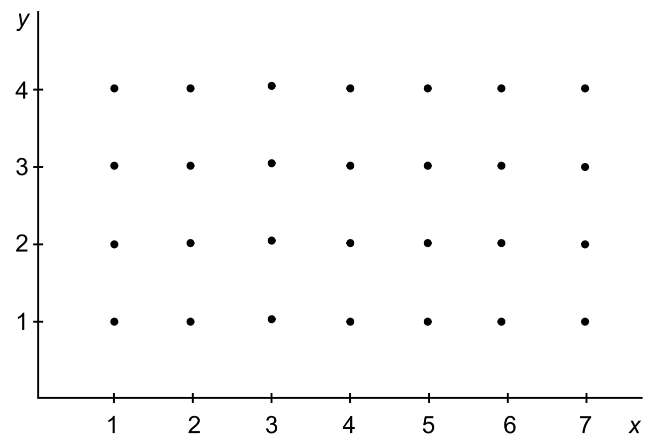


Рис. 1.9. Пространство элементарных событий Ω, состоящее из 28 точек

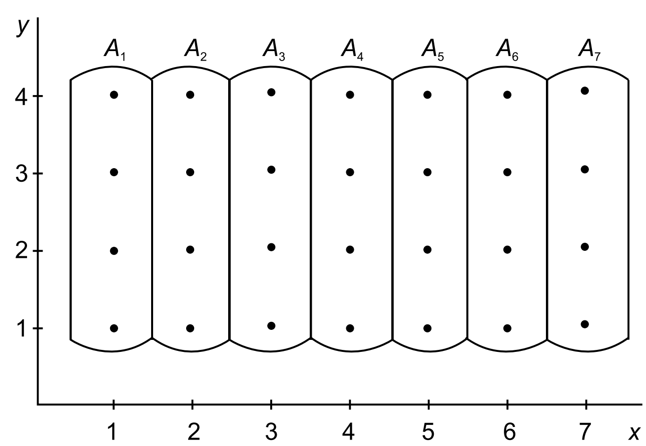


Рис. 1.10. Пространство элементарных событий Ω, разбитое на подмножества, на каждом из которых постоянна случайная величина X

На рис. 1.10 изображена разбивка множества Ω на подмножества A_i, на каждом из которых постоянна случайная величина X, а на рис. 1.11 изображена разбивка множества Ω, на подмножества B_j, на каждом из которых постоянна случайная величина Y.

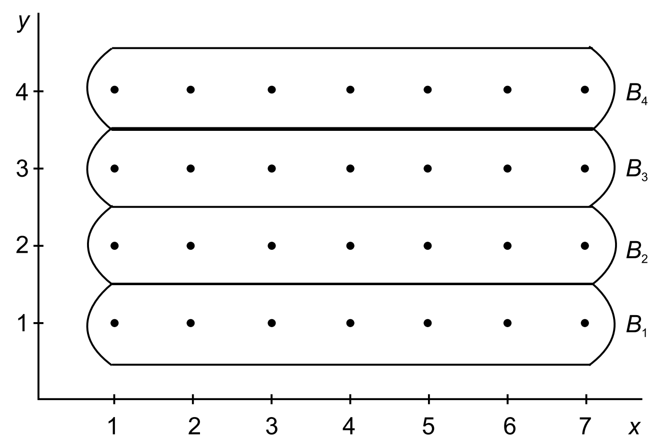


Рис. 1.11. Пространство элементарных событий Ω, разбитое на подмножества, на каждом из которых постоянна случайная величина Y

Тогда при любом i

$$P(A_i) = \frac{1}{k}$$

и при любом j

$$P(B_j) = \frac{1}{l}.$$

Но множество $A_i \cap B_j$ состоит из одной точки. Поэтому

$$P(A_i \cap B_j) = \frac{1}{kl} = \frac{1}{k} \times \frac{1}{l} = P(A_i)P(B_j).$$

То есть случайные величины X и Y независимы.

Разобранный пример можно обобщить, задав случайную величину X как некоторую функцию от абсциссы, а случайную величину

Y — как некоторую функцию от ординаты. При этом случайные величины X и Y по-прежнему будут независимыми, так как множества A_i и B_j будут строиться аналогично.

Чтобы дать пример не независимых случайных величин, допустим, что в примере 1.5 X по-прежнему является функцией только от абсциссы точки, а Y принимает во всех точках множества Ω различные значения. Тогда при любом j множество B_j состоит из одной точки. Поэтому, например, если B_1 — это множество, содержащее левую нижнюю точку на рис. 1.9, то

$$A_1 \cap B_1 = B_1$$

и

$$P(A_1 \cap B_1) > P(A_1)P(B_1).$$

Пример 1.6. Две студентки, Аня и Бетти, хотят во время 10-минутного перерыва между занятиями позвонить по телефону своим знакомым. Пусть случайная величина X используется для моделирования продолжительности разговора Ани, а случайная величина Y — для моделирования продолжительности разговора Бетти. И та, и другая случайная величина принимает целые значения от 0 до 10. Можно ли считать случайные величины X и Y независимыми?

Рассмотрим вначале случай, когда Аня и Бетти могут использовать только один телефон. Тогда случайные величины X и Y не являются независимыми. Действительно,

$$P(X = 6) > 0, \qquad P(Y = 6) > 0,$$

но

$$P((X = 6) \cap (Y = 6)) = 0.$$

Последняя вероятность равна 0, поскольку во время 10-минутного перерыва с одного и того же телефона не могут быть сделаны два звонка продолжительностью 6 минут каждый. А если нам удалось найти хотя бы одну пару чисел x и y такую, что события $X = x$ и $Y = y$ не независимы, то не независимы и случайные величины X и Y. События $X = 6$ и $Y = 6$ не независимы, поскольку

$$P((X = 6) \cap (Y = 6)) \neq P(X = 6)\, P(Y = 6).$$

Рассмотрим теперь случай, когда Аня и Бетти могут звонить с двух разных телефонов. Тогда приведенные возражения против того, чтобы считать случайные величины X и Y независимыми, отпадают.

Автору неизвестны реальные прикладные задачи, для решения которых было бы нужно моделировать продолжительность разговора студенток по телефону в перерыве между занятиями при помощи случайных величин. Данный пример, как и многие другие примеры, приведенные в этой книге, носит учебный или полуучебный характер. Но существуют важные прикладные задачи, для решения которых те или иные числовые показатели моделируются случайными величинами. Например, это задачи расчета безарбитражной цены финансового инструмента и определения стратегии хеджирования. Здесь случайными величинами моделируются цены основных активов, процентные ставки, значения обменных курсов. Другой пример — это макроэкономические задачи, связанные с изучением роста валового внутреннего продукта и выявлением влияния отдельных факторов на его изменение. Но сложность используемого для решения этих задач аппарата теории вероятностей не позволяет соединить изложение начал данного раздела математики с рассмотрением серьезных прикладных задач.

Приведем другое определение независимости случайных величин, причем не только для двух, но и для любого числа случайных величин.

Определение. Рассмотрим случайные величины $X_1, X_2, ..., X_k$,

$$X_i : \Omega \to \boldsymbol{R},\ i = 1, 2, ..., k.$$

Рассмотрим набор чисел a_i, $i = 1, 2, ..., k$ (допускается, чтобы a_i при некоторых i принимало значение $-\infty$) и набор чисел b_i, $i = 1, 2, ..., k$ (допускается, чтобы b_i при некоторых i принимало значение ∞) такие, что при любом i, $1 \le i \le k$,

$$-\infty \le a_i \le b_i \le \infty.$$

Рассмотрим события $A_i \subset \Omega$, $i = 1, 2, ..., k$, где A_i состоит из тех и только тех $\omega \in \Omega$, для которых

$$a_i \le X_i(\omega) \le b_i.$$

Случайные величины $X_1, X_2, ..., X_k$ называются *независимыми*, если для любых наборов $a_1, a_2, ..., a_k$ и $b_1, b_2, ..., b_k$, удовлетворяющих приведенным выше условиям, события $A_1, A_2, ..., A_k$ независимы.

З а м е ч а н и е. Первое определение независимости, данное для двух случайных величин X и Y, можно обобщить для любого числа случайных величин X_1, X_2, ..., X_k. Тогда можно показать, что для конечных пространств элементарных событий Ω первое и второе определение независимости случайных величин эквивалентны. Но (это будет видно из дальнейшего) для бесконечных пространств элементарных событий Ω допустимо использовать только второе определение независимости случайных величин.

Приведенное определение независимости случайных величин X_1, X_2, ..., X_k является достаточно сложным. Чтобы лучше понять его, можно порекомендовать читателю обобщить пример 1.6 для трех студенток и рассмотреть различные варианты, при которых случайные величины X_1, X_2, X_3 могут быть независимыми, и при которых они не могут быть независимыми.

1.3

Ожидаемое значение и дисперсия случайной величины. Ковариация и корреляция случайных величин

Начнем с рассмотрения случая, когда множество Ω конечно и содержит N элементов:

$$\Omega = \{\omega_1, \omega_2, ..., \omega_N\}.$$

Напомним, что вероятность элементарного события ω_i мы обозначаем p_i, случайная величина X — это функция

$$X : \Omega \to \boldsymbol{R}.$$

Определение. *Ожидаемым значением* случайной величины X называется следующее число

$$E(X) = \sum_{i=1}^{N} X(\omega_i) p_i.$$

Пример 1.7. В разные дни студент затрачивает разное время на дорогу до университета. Пусть Ω — множество, состоящее из N дней, каждому элементарному событию, т.е. каждому дню припишем вероятность $p_i = \frac{1}{N}$, $i = 1, 2, ..., N$.

Пусть X — это случайная величина, которая означает время, затраченное студентом на дорогу до университета в каждый из дней. Тогда среднее время, затрачиваемое студентом на дорогу до университета,

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} X(\omega_i).$$

З а м е ч а н и е. Иногда ожидаемое значение случайной величины называют ее *средним значением*, или *ожиданием*, или *математическим ожиданием*.

Определение. *Дисперсией* случайной величины X называется ожидаемое значение случайной величины

$$\left(X - E(X)\right)^2,$$

которое обозначается $D(X)$. Другими словами,

$$D(X) = \sum_{i=1}^{N}\left(X(\omega_i) - E(X)\right)^2 p_i.$$

Дисперсия характеризует разброс случайной величины, т.е. насколько сильно в различных точках множества Ω случайная величина отличается от среднего значения. Только в том случае, когда случайная величина постоянна, ее дисперсия равна нулю. Во всех других случаях (напомним, что речь идет о конечном множестве Ω) дисперсия случайной величины положительна. Часто вместо дисперсии удобно рассматривать другую меру разброса случайной величины.

Определение. *Стандартным отклонением* случайной величины называют квадратный корень из ее дисперсии:

$$\sigma_X = \sqrt{D(X)}$$

(σ — греческая буква сигма).

В рассмотренном примере 1.7 стандартное отклонение измеряется в минутах, как и среднее значение, а дисперсия — в минутах в квадрате, что менее удобно.

Поскольку случайная величина — это числовая функция, то, как говорилось в параграфе 1.1, можно рассматривать суммы и произведения случайных величин, умножать их на константы и прибавлять константы (константу также можно считать случайной величиной). Например, значение случайной величины $(X + Y)$ в точке ω_i определяется как

$$X(\omega_i) + Y(\omega_i).$$

При этом, конечно, подразумевается, что случайные величины X и Y определены на одном и том же пространстве элементарных событий Ω.

Непосредственно из определений видно, что для любой случайной величины X и для любого числа $c \in \boldsymbol{R}$ справедливы следующие равенства:

$$E(cX) = cE(X), \qquad E(X + c) = E(X) + c,$$
$$D(cX) = c^2 D(X), \qquad D(X + c) = D(X),$$
$$\sigma_{cX} = |c|\sigma_X, \qquad \sigma_{X+c} = \sigma_X.$$

Определение. *Ковариацией* случайных величин X и Y называется следующее число

$$\operatorname{Cov}(X,\ Y) = E\big((X - E(X))(Y - E(Y))\big).$$

В частности, $\operatorname{Cov}(X, X) = D(X)$.

То, что в формуле для $\operatorname{Cov}(X, Y)$ трижды используется символ математического ожидания E, не должно вызывать затруднений.

$$(X - E(X))(Y - E(Y))$$

есть случайная величина, определенная на том же самом пространстве элементарных событий Ω, что и случайные величины X и Y.

Определение. Если стандартные отклонения случайных величин X и Y положительны, то *корреляцией* X и Y называется следующее число:

$$\mathrm{Cor}(X,\ Y)=\frac{\mathrm{Cov}(X,\ Y)}{\sigma_X\sigma_Y}.$$

З а м е ч а н и е. Иногда величину Cor(X, Y) называют *коэффициентом корреляции* случайных величин X и Y.

Т е о р е м а 1.1. *Ожидаемое значение суммы, случайных величин равно сумме ожидаемых значений*

$$E(X+Y)=E(X)+E(Y).$$

Д о к а з а т е л ь с т в о.

$$E(X+Y)=\sum_{i=1}^{N}\left(X(\omega_i)+Y(\omega_i)\right)p_i=$$

$$=\sum_{i=1}^{N}X(\omega_i)p_i+\sum_{i=1}^{N}Y(\omega_i)p_i=E(X)+E(Y).$$

□

Здесь и далее символом □ мы будем обозначать конец доказательства.

Т е о р е м а 1.2. *Ожидаемое значение произведения независимых случайных величин равно произведению ожидаемых значений*

$$E(XY)=E(X)E(Y).$$

Д о к а з а т е л ь с т в о. Пусть как и при определении двух независимых случайных величин, случайная величина X принимает значение x_1 на множестве A_1, значение x_2 — на множестве A_2, ..., значение x_k — на множестве A_k; случайная величина Y принимает значение y_1 на множестве B_1, значение y_2 — на множестве B_2, ..., значение y_l — на множестве B_l. Тогда нетрудно увидеть, что

$$E(X)=\sum_{i=1}^{k}x_iP(A_i),\qquad E(Y)=\sum_{j=1}^{l}y_jP(B_j),$$

и

$$E(XY)=\sum_{i=1}^{k}\sum_{j=1}^{l}x_iy_jP(A_i\cap B_j)=\sum_{i=1}^{k}\sum_{j=1}^{l}x_iy_jP(A_i)P(B_j)=$$

$$= \sum_{i=1}^{k} x_i P(A_i) \times \sum_{j=1}^{l} y_j P(B_j).$$

Мы воспользовались независимостью случайных величин X и Y, из которой следует, что

$$P(A_i \cap B_j) = P(A_i)P(B_j).$$

□

Т е о р е м а 1.3. *Дисперсия суммы независимых случайных величин равна сумме дисперсий*

$$D(X + Y) = D(X) + D(Y).$$

Д о к а з а т е л ь с т в о. Обозначим

$$\mu_X = E(X), \quad \mu_Y = E(Y).$$

(μ — греческая буква “мю”). Пользуясь теми же обозначениями, что и при доказательстве теоремы 1.2, имеем

$$D(X+Y) = \sum_{i=1}^{k}\sum_{j=1}^{l}\left(x_i + y_j - \mu_X - \mu_Y\right)^2 P\left(A_i \cap B_j\right) =$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{l}(x_i - \mu_X)^2 P\left(A_i \cap B_j\right) + \sum_{i=1}^{k}\sum_{j=1}^{l}\left(y_j - \mu_Y\right)^2 P\left(A_i \cap B_j\right) +$$

$$+\ 2\sum_{i=1}^{k}\sum_{j=1}^{l}(x_i - \mu_X)\left(y_j - \mu_Y\right)P\left(A_i \cap B_j\right).$$

Поскольку первая сумма в последнем выражении равна

$$\sum_{i=1}^{k}(x_i - \mu_X)^2 P(A_i) = D(X)$$

и вторая сумма равна

$$\sum_{j=1}^{l}\left(y_j - \mu_Y\right)^2 P\left(B_j\right) = D(Y),$$

полученный результат можно записать так:

$$D(X + Y) = D(X) + D(Y) + 2\mathrm{Cov}(X, Y).$$

Заметим, что до сих пор мы не пользовались независимостью случайных величин X и Y. С учетом независимости случайных величин X и Y

$$\operatorname{Cov}(X,Y)=\sum_{i=1}^{k}(x_i-\mu_X)P(A_i)\times\sum_{j=1}^{l}(y_j-\mu_Y)P(B_j)=$$

$$=\left[\sum_{i=1}^{k}x_iP(A_i)-\mu_X\right]\left[\sum_{j=1}^{l}y_jP(B_j)-\mu_Y\right]=0.$$

□

Теорема 1.4. *Ковариация двух независимых случайных величин равна* 0.

Доказательство этой теоремы содержится в доказательстве теоремы 1.3.

Из теоремы 1.4 следует, что если случайные величины X и Y независимы и $\sigma_X > 0$, $\sigma_Y > 0$, то

$$\operatorname{Cor}(X, Y) = 0.$$

Теорема 1.5. *Для любых случайных величин X и Y таких, что* $\sigma_X > 0$, $\sigma_Y > 0$

$$-1 \le \operatorname{Cor}(X, Y) \le 1.$$

Доказательство. При доказательстве теоремы 1.3 было установлено, что для любых случайных величин X^* и Y^*

$$D(X^* + Y^*) = D(X^*) + D(Y^*) + 2\operatorname{Cov}(X^*, Y^*).$$

Положим

$$X^*=\frac{X-E(X)}{\sigma_X},\quad Y^*=\frac{Y-E(Y)}{\sigma_Y}.$$

Тогда с учетом теорем 1.1 и 1.3 $E(X^*) = 0$, $E(Y^*) = 0$, $D(X^*) = 1$, $D(Y^*) = 1$ и

$$\operatorname{Cov}(X^*, Y^*) = \operatorname{Cor}(X, Y).$$

Из условия (поскольку дисперсия любой случайной величины неотрицательна)

$$D(X^* + Y^*) \ge 0$$

получаем

$$2 + 2\mathrm{Cor}(X,\ Y) \geq 0,$$

или

$$\mathrm{Cor}(X,\ Y) \geq -1.$$

Аналогично из условия

$$D(X^* - Y^*) \geq 0$$

получаем

$$\mathrm{Cor}(X,\ Y) \geq 1.$$

□

Т е о р е м а 1.6. Для *любых случайных величин X и Y таких, что* $\sigma_X > 0$, $\sigma_Y > 0$, *следующие утверждения равносильны:*

1) *для некоторых чисел a и b, a* $\neq 0$

$$X = aY + b;$$

2) Cor $(X,\ Y)$ *равна* 1 *или* –1.

Д о к а з а т е л ь с т в о. Пусть X^* и Y^* определяются так же, как и при доказательстве теоремы 1.5. Тогда, если выполнено условие (1),

$$X^* = \frac{X - E(X)}{\sigma_X} = \frac{aY + b - aE(Y) - b}{|a|\sigma_Y} = \frac{a}{|a|}\frac{Y - E(Y)}{\sigma_Y} = \frac{a}{|a|}Y^*,$$

т.е.

$$X^* = Y^* \text{ при } a > 0$$

и

$$X^* = -Y^* \text{ при } a < 0.$$

Поэтому при $a > 0$

$$0 = D(X^* - Y^*) = D(X^*) + D(Y^*) - 2\mathrm{Cov}(X^*,\ Y^*) = 2(1 - \mathrm{Cor}(X,\ Y)),$$

откуда

$$\mathrm{Cor}(X,\ Y) = 1.$$

Аналогично при $a < 0$

$$\mathrm{Cor}(X,\ Y) = -1.$$

Таким образом, доказано, что из условия (1) следует условие (2). Докажем, что из условия (2) следует условие (1). Из приведенных выкладок следует, что условие

$$\mathrm{Cor}(X,\ Y) = 1$$

влечет условие

$$D(X^* - Y^*) = 0,$$

а условие

$$\mathrm{Cor}(X,\ Y) = -1$$

влечет условие

$$D(X^* + Y^*) = 0.$$

Из равенства нулю дисперсии следует, что случайная величина — постоянная функция (напомним, что мы рассматриваем случай конечного множества Ω). Поэтому в первом случае

$$X^* - Y^* = c,$$

а во втором случае

$$X^* + Y^* = c,$$

где c — некоторое число. Отсюда получаем, что для первого случая

$$X = \frac{\sigma_X}{\sigma_Y} Y + \left(-\frac{\sigma_X}{\sigma_Y} E(Y) + \sigma_X c + E(X) \right)$$

и для второго случая

$$X = -\frac{\sigma_X}{\sigma_Y} Y + \left(\frac{\sigma_X}{\sigma_Y} E(Y) + \sigma_X c + E(X) \right).$$

□

Существует случай бесконечного множества Ω, который может быть рассмотрен так же, как и случай конечного множества Ω. Это случай $\Omega = \boldsymbol{N}$. (Напомним, что через $\boldsymbol{N}$ мы обозначаем множество целых положительных чисел, $\boldsymbol{N} = \{1, 2, 3...\}$.) Вместо сумм $\sum\limits_{i=1}^{N}$ надо рассматривать суммы $\sum\limits_{i=1}^{\infty}$ и накладывать условия, обеспечивающие сходимость рядов, что вызывает трудности лишь технического характера. Мы не будем рассматривать этот случай вплоть до главы 8, где он окажется необходим для изучения одного важного класса случайных величин.

Более интересным с точки зрения статистических приложений (хотя и недостаточным для математического моделирования реальных ситуаций) является случай, когда $\Omega = [0, 1]$. Отрезок $[0, 1]$ — это тоже бесконечное множество. Но здесь оказывается, что нельзя в качестве случайных величин брать любые функции

$$X : [0, 1] \to \boldsymbol{R},$$

а в качестве событий — любые подмножества $A \subset [0, 1]$.

Например, существует функция (ее конструкция достаточно сложна)

$$X : [0, 1] \to \boldsymbol{R},$$

обладающая следующим свойством. Для любого отрезка $[a, b] \subset [0, 1]$, где $a < b$, и для любого числа $c \in \boldsymbol{R}$ существует точка $z \in [a, b]$ такая, что

$$X(z) = c.$$

То есть эта функция на любом сколь угодно малом отрезке принимает любые значения. Конечно, эта функция не является непрерывной. Использование подобных функций в качестве случайных величин привело бы к непреодолимым трудностям, например оказалось бы, что для событий

$$\{\omega \in [0, 1] : X(\omega) \leq c\},$$

где c — некоторое действительное число, невозможно определить вероятность. Но в то же время чрезмерное сужение класса рассматриваемых функций приводит к невозможности построить содержательную теорию.

Еще одна неприятность заключается в следующем. Если событие является отрезком, принадлежащим $[0, 1]$ (а такие события рассматривать можно), то в качестве вероятности события наиболее естественно взять длину отрезка. Тогда вероятность отдельной точки равна нулю. Но в таком случае, каким должен быть аналог теоремы о сложении вероятностей, ведь объединение всех точек совпадает с $[0, 1]$ и имеет вероятность 1?

В случае бесконечного множества Ω в качестве событий и случайных величин приходится рассматривать только подмножества и числовые функции, удовлетворяющие определенным требованиям.

Формулировки соответствующих условий достаточно сложны, и мы приводить их не будем. Отметим только, что при $\Omega = [0, 1]$

1) в число событий входят любые отрезки, принадлежащие [0, 1], и объединения конечного числа таких отрезков;

2) в число случайных величин входят все непрерывные функции

$$X : [0, 1] \to \boldsymbol{R}$$

и все кусочно-постоянные функции.

Сначала определим ожидаемое значение случайной величины, которая является кусочно-постоянной функцией. Пусть отрезок [0, 1] разбит на N отрезков длиной $p_1, p_2, \ldots, p_N$, не имеющих попарно общих точек, кроме концов. Таким образом,

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

Рассмотрим случайную величину

$$X : \Omega \to \boldsymbol{R},$$

которая постоянна внутри каждого из этих отрезков и непрерывна справа (см. рис. 1.12).

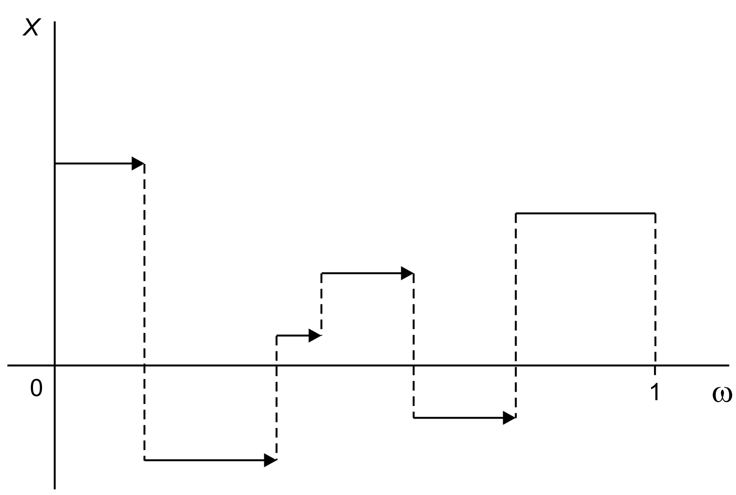


Рис. 1.12. График случайной величины, являющейся кусочно-постоянной функцией

Значения, которые принимает эта случайная величина, обозначим $x_1, x_2, \ldots, x_N$. Ожидаемое значение данной случайной величины определим по формуле

$$E(X) = \sum_{i=1}^{N} x_i p_i.$$

Заметим, что это определение очень похоже на определение ожидаемого значения случайной величины, для конечного множества Ω. Если обозначить через ω_i середину i-го отрезка, а вероятностью элементарного события считать длину данного отрезка, то определение будет совпадать в точности.

Определение ожидаемого значения случайной величины, являющейся кусочно-постоянной функцией на [0, 1], можно записать и в другом виде:

$$E(X) = \int_0^1 X(\omega) d\omega.$$

Для кусочно-постоянной функции значения интеграла и суммы совпадают. Но при помощи интеграла ожидаемое значение можно определить не только для кусочно-постоянной функции $X(\omega)$, но и для функций более общего вида, например, кусочно-непрерывных (см. рис. 1.13). Поскольку длина отрезка интегрирования равна 1, то так определенное ожидаемое значение случайной величины X — это в некотором смысле среднее значение функции X на отрезке [0, 1].

Рис. 1.13. График случайной величины, являющейся кусочно-непрерывной функцией

Если ожидаемое значение случайной величины определено, то дисперсия, ковариация и корреляция определяются по тем же формулам, что и в случае конечного множества Ω.

В качестве другого примера бесконечного пространства элементарных событий Ω возьмем единичный квадрат (см. рис. 1.14).

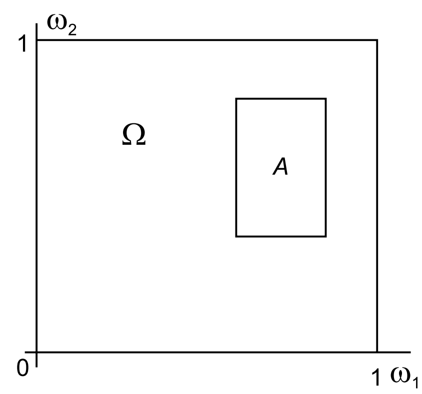


Рис. 1.14. Пространство элементарных событий Ω, являющееся единичным квадратом, и событие $A \subset \Omega$

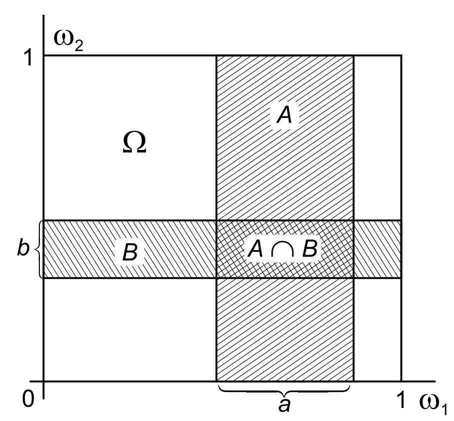


Рис. 1.15. Независимые события A и B в случае, когда пространство элементарных событий Ω является единичным квадратом

Рассмотрим события, являющиеся прямоугольниками со сторонами, параллельными осям координат. Пусть вероятность события A — это площадь прямоугольника.

На рис. 1.15 изображены независимые события A и B. Действительно, площадь множества A равна $a \times 1 = a$; площадь множества B равна $1 \times b = b$; площадь множества $A \cap B$ равна $a \times b$, т.е. условие независимости событий

$$P(A \cap B) = P(A)\, P(B)$$

выполнено.

Рассматривая единичный квадрат $\Omega = [0, 1]^2$, где

$$[0, 1]^2 = \{(x_1, x_2) : 0 \le x_1 \le 1, 0 \le x_2 \le 1\},$$

в качестве событий мы до сих пор брали только прямоугольники со сторонами, параллельными осям координат. Но можно в качестве событий рассматривать любые подмножества единичного квадрата, для которых определена площадь, и принимать эту площадь за вероятность события. Тогда событиями являются, например, круги, треугольники, параллелограммы. Более того, к событиям возможно отнести и некоторые множества, для которых площадь (в обычном смысле) не определена. Примером такого множества может быть

$$C = \bigcup_{m=1}^{\infty} K_m,$$

где K_m — круг с центром в точке $\left(\frac{1}{2m}, \frac{1}{2m}\right)$ радиуса $\frac{1}{3m(m+1)}$ (радиусы выбраны так, чтобы круги не пересекались). Вероятностью события C является сумма ряда, m-й член которого — это площадь круга K_m.

Если множества A и B являются событиями, то событием является и объединение этих множеств, причем для вероятности объединения имеет место выражение

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

Пусть теперь пространство элементарных событий Ω — это единичный куб

$$[0,1]^3 = \{(x_1, x_2, x_3) : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ 0 \le x_3 \le 1\}.$$

Если в качестве события рассмотреть подмножество множества Ω, являющееся прямоугольным параллелепипедом со сторонами, параллельными осям координат,

$$A = \{(x_1, x_2, x_3) : a_1 \le x_1 \le b_1,\ a_2 \le x_2 \le b_2,\ a_3 \le x_3 \le b_3\},$$

то за вероятность события A можно принять объем этого параллелепипеда

$$P(A) = \prod_{i=1}^{3} (b_i - a_i).$$

В качестве событий можно рассматривать и другие подмножества единичного куба, для которых определен объем, например шары, тетраэдры, непрямоугольные параллелепипеды, и принимать этот объем за вероятность соответствующего события. К событиям возможно отнести и некоторые множества, для которых объем (в обычном смысле) не определен, например множества того же вида, что и рассмотренное выше множество C, только множества K_m теперь не круги, а шары.

Аналогично при любом $n \in N$ в качестве пространства элементарных событий можно рассмотреть $\Omega = [0, 1]^n$, где

$$[0,1]^n = \{(x_1, x_2, ..., x_n) : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1,\ ...,\ 0 \le x_n \le 1\}.$$

Такое множество называется n-мерным единичным кубом. Данное пространство элементарных событий Ω достаточно для тех статистических приложений, которые рассматриваются в этой книге. В качестве событий могут быть взяты, в частности, n-мерные параллелепипеды со сторонами параллельными осям координат,

$$A = \{(x_1, x_2, ..., x_n) : a_1 \le x_1 \le b_1,\ a_2 \le x_2 \le b_2,\ a_n \le x_n \le b_n\}.$$

За вероятность $P(A)$ такого события принимается n-мерный объем параллелепипеда

$$P(A) = \prod_{i=1}^{n} (b_i - a_i).$$

Но такими параллелепипедами, как это видно из двумерного и трехмерного случаев, не исчерпывается набор всех подмножеств множества $\Omega = [0, 1]^n$, которые возможно отнести к событиям.

Числовая функция

$$X : [0, 1]^n \to \boldsymbol{R}$$

называется случайной величиной, если для любого $x \in \boldsymbol{R}$ множество $A \subset [0, 1]^n$, состоящее из тех и только тех $\omega \in [0, 1]^n$, для которых

$$X(\omega) \le x,$$

является событием. Можно показать, что любая непрерывная на $[0, 1]^n$ функция X является случайной величиной. (Здесь необходимо, чтобы читатель вернулся к случаю $n = 1$, нарисовал график некоторой функции X и посмотрел, какими при различных x являются события

$$A = \{\omega \in [0, 1] : X(\omega) \le x\}.$$

В дальнейшем такие события часто будут обозначаться сокращенно $X \le x$).

Напомним, что при рассмотрении случайных величин

$$X : [0, 1] \to \boldsymbol{R}$$

ожидаемое значение случайной величины X определялось, как интеграл от этой функции по отрезку $[0, 1]$. Аналогично при $n = 2$ ожидаемым значением случайной величины X называется двойной интеграл от функции $X(\omega)$ по единичному квадрату $[0, 1]^2$, при $n = 3$ — тройной интеграл по единичному кубу $[0, 1]^3$. При произвольном n ожидаемым значением случайной величины X называется n-кратный интеграл от функции $X(\omega)$ по n-мерному единичному кубу $[0, 1]^n$.

Но если для конечного пространства элементарных событий Ω любая случайная величина

$$X : \Omega \to \boldsymbol{R}$$

имеет ожидаемое значение, то для бесконечных пространств элементарных событий Ω существуют случайные величины, у которых нет ожидаемого значения.

Если все ожидаемые значения существуют, то теоремы 1.1—1.6 остаются справедливыми и в случае, когда пространство элементарных событий Ω бесконечно. Остаются справедливыми и формулы

$$E(cX) = cE(X),\ E(X + c) = E(X) + c,$$

а также соответствующие формулы для дисперсий и стандартных отклонений, приведенные в начале параграфа.

1.4

Функции распределения и функции плотности случайных величин

Пусть Ω — пространство элементарных событий, конечное или бесконечное, случайная величина

$$X : \Omega \to \boldsymbol{R}$$

и $x \in \boldsymbol{R}$. Через A обозначим множество элементов $\omega \in \Omega$ таких, что

$$X(\omega) \le x.$$

Положим

$$F(x) = P(A).$$

Функция

$$F : \boldsymbol{R} \to [0, 1]$$

называется *функцией распределения* случайной величины X.

З а м е ч а н и е. Правильнее (но более громоздко) было бы обозначить эту функцию не $F(x)$, а $F_X(x)$, чтобы подчеркнуть ее связь со случайной величиной X.

Ясно, что F — неубывающая функция. Действительно, если $x_1 < x_2$, то для множеств

$$A_1 = \{\omega \in \Omega : X(\omega) \le x_1\}$$

и

$$A_2 = \{\omega \in \Omega : X(\omega) \le x_2\}$$

имеет место соотношение $A_1 \subset A_2$. Отсюда следует, что $P(A_1) \le P(A_2)$ и, значит,

$$F(x_1) \le F(x_2).$$

Посмотрим, как устроена функция распределения в том случае, когда случайная величина определена на конечном множестве Ω,

$$\Omega = \{\omega_1, \omega_2, ..., \omega_N\},$$

и вероятности элементарных событий равны

$$p_1, p_2, ..., p_N.$$

Пусть $X(\omega_i) = x_i$, $i = 1, 2, ..., N$, причем

$$x_1 < x_2 < \cdots < x_N.$$

Тогда функция распределения случайной величины X имеет вид, показанный на рис. 1.16; значение функции F при переходе через точку x_i, возрастает на p_i.

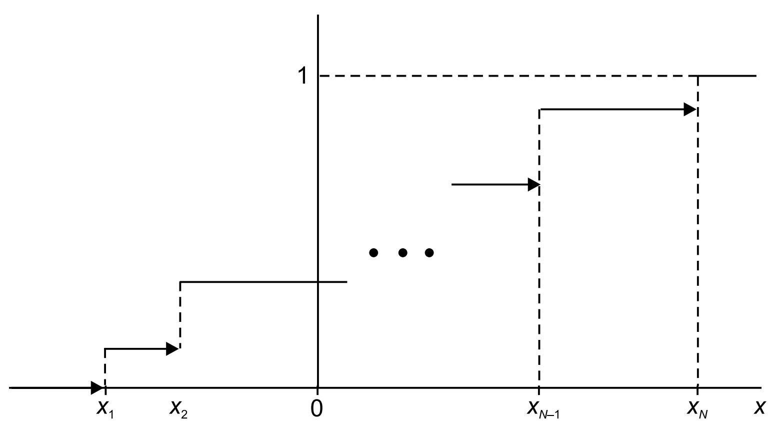


Рис. 1.16. Функция распределения случайной величины, определенной на конечном пространстве элементарных событий Ω

Для случая, когда Ω — конечное множество, F является неубывающей ступенчатой функцией. Полезно сравнить рис. 1.16 с рис. 1.12. На рис. 1.12 ось x расположена вертикально, а на рис. 1.16 — горизонтально. По вертикальной оси на рис. 1.16 откладывается сумма длин отрезков, изображенных на рис. 1.12, на которых $X(\omega) \le x$.

Нетрудно понять, что если две случайные величины, определенные на конечных множествах Ω, имеют одинаковые функции распределения, то математические ожидания и дисперсии этих случайных величин совпадают. Действительно, функция распределения содержит всю информацию о значениях, принимаемых случайной величиной, и о вероятностях соответствующих событий.

Если множество Ω бесконечно, то функция $F(x)$ может быть как непрерывной (и даже дифференцируемой), так и разрывной. В любом случае она остается неубывающей функцией (см. рис.1.17).

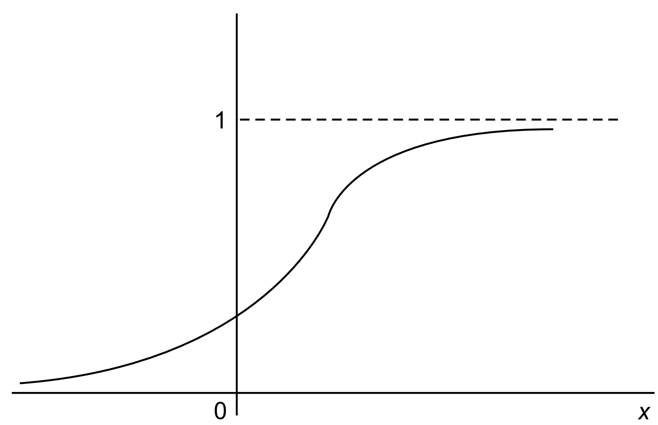


Рис. 1.17. Дифференцируемая функция распределения случайной величины (пространство элементарных событий бесконечно)

Пример 1.8. Три студента, Антон, Борис и Валерий, по-разному использовали 5-минутный перерыв между занятиями. Антон позвонил по телефону. Борис подошел к преподавателю и попросил еще раз объяснить непонятный им материал. Валерий решил попить чаю в буфете. Пусть случайная величина X используется для моделирования продолжительности разговора Антона по телефону, случайная величина Y — для моделирования продолжительности разговора Бориса с преподавателем, случайная величина Z — для моделирования времени, которое Валерий провел в буфете. Будем считать, что случайные величины X, Y и Z принимают любые действительные значения от 0 минут до 5 минут.

Можно принять, например, следующие значения для вероятностей отдельных событий, связанных со случайными величинами X, Y и Z. Для случайной величины X

$$P(X \le 0) = 0;\ P(X \le 1) = 0{,}2;\ \ P(X \le 2) = 0{,}4;$$
$$P(X \le 3) = 0{,}6;\ P(X \le 4) = 0{,}8;\ \ P(X \le 5) = 1.$$

Для случайной величины Y

$$P(Y \le 0) = 0;\ P(Y \le 1) = 0{,}5;\ \ P(Y \le 2) = 0{,}8;$$
$$P(Y \le 3) = 0{,}9;\ P(Y \le 4) = 0{,}95;\ \ P(Y \le 5) = 1.$$

Для случайной величины Z

$$P(Z \le 0) = 0;\ P(Z \le 1) = 0;\ \ P(Z \le 2) = 0{,}1;$$
$$P(Z \le 3) = 0{,}2;\ P(Z \le 4) = 0{,}5;\ \ P(Z \le 5) = 1.$$

Графики функций распределения случайных величин X, Y и Z тогда могут иметь вид, приведенный на рис.1.18.

Функция распределения содержит в себе значительную часть информации о поведении случайной величины. Но, зная отдельно функции распределения случайных величин X и Y, нельзя ответить, например, на вопрос, будут ли эти случайные величины независимыми.

Если функция F в каждой точке $x \in \boldsymbol{R}$ имеет производную

$$F'(x) = f(x),$$

где f — непрерывная функция, то функция f называется *функцией плотности* случайной величины X. При любом $x \in \boldsymbol{R}$

$$f(x) \geq 0,$$

поскольку $F(x)$ является неубывающей функцией. По формуле Ньютона — Лейбница получаем

$$F(b) - F(a) = \int_a^b f(x)\mathrm{d}x$$

— вероятность события A, состоящего из тех элементов $\omega \in \Omega$, для которых

$$a \leq X(\omega) \leq b.$$

Из соотношений

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to \infty} F(x) = 1$$

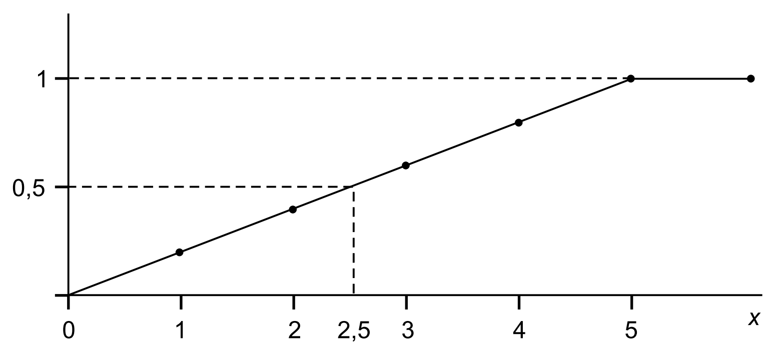
следует, что для любой функции плотности $f(x)$

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

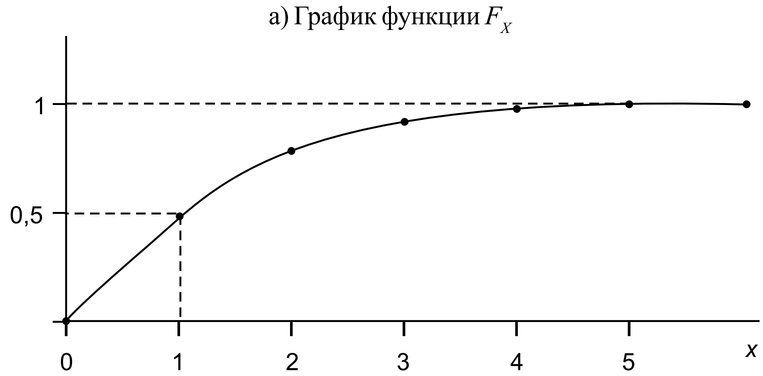
и

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

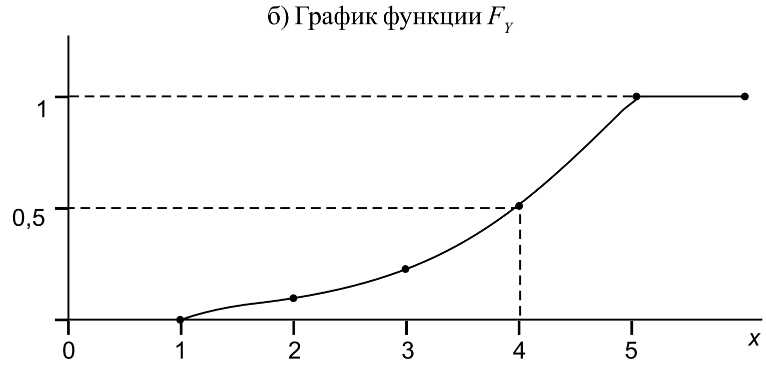
Если случайная величина имеет функцию плотности, то, конечно, эта случайная величина определена на бесконечном множестве Ω, поскольку функция распределения случайной величины, определенной на конечном множестве Ω, как мы установили выше, является кусочно-постоянной и, следовательно, не является дифференцируемой на всей прямой.



а) График функции F_X



б) График функции F_Y



в) График функции F_Z

Рис. 1.18. Графики функций распределения случайных величин X, Y и Z

Т е о р е м а 1.7. *Если случайная величина X имеет функцию плотности f, то ожидаемое значение случайной величины X выражается следующей формулой*

$$E(X) = \int_{-\infty}^{\infty} u\, f(u) du.$$

В конце параграфа 1.3 мы определили ожидаемое значение для кусочно-непрерывных случайных величин

$$X : [0, 1] \to \boldsymbol{R},$$

как

$$E(X) = \int_{0}^{1} X(\omega) d\omega.$$

Как уже говорилось в параграфе 1.3, определение ожидаемого значения для случайных величин общего вида является обобщением данного определения.

Теорему 1.7 мы примем без доказательства: это один из трудных результатов теории вероятностей. В дальнейшем, оперируя ожидаемыми значениями случайных величин, мы будем пользоваться либо выражением для ожидаемого значения из теоремы 1.7 (если случайная величина имеет функцию плотности), либо непосредственно определением ожидаемого значения (если случайная величина определена на конечном множестве Ω). Других случайных величин мы рассматривать не будем (за исключением главы 8, где рассматривается случай, когда $\Omega = \boldsymbol{N}$, но этот случай, как отмечалось выше, аналогичен случаю конечных множеств Ω).

Т е о р е м а 1.8. *Если случайная величина X имеет функцию плотности f, то дисперсия случайной величины X выражается следующей формулой*:

$$D(X) = \int_{-\infty}^{\infty} \left(u - E(X)\right)^2 f(u) du.$$

Теорема 1.8 (которую мы также не будем доказывать) и теорема 1.7 являются частными случаями одной теоремы об ожидаемом значении функции от случайной величины (напомним, что $D(X) = E((X - E(X))^2)$). Мы ограничились изложением этих двух частных случаев, поскольку общий результат в дальнейшем в этой книге не понадобится.

З а м е ч а н и е. Чтобы быть строгими, в формулировках теорем 1.7 и 1.8 следовало потребовать сходимости соответствующих интегралов. Если интеграл, участвующий в определении ожидаемого значения или дисперсии расходится, то несмотря на наличие функции плотности, конечного ожидаемого значения или конечной дисперсии у случайной величины нет. Мы не стали упоминать об этом требовании раньше, чтобы не утяжелять формулировки теорем. В дальнейшем, если не оговорено противное, интегралы будем считать сходящимися.

Подчеркнем, что из сходимости интеграла

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

вообще говоря, не следует сходимость интеграла

$$\int_{-\infty}^{\infty} x\, f(x)dx.$$

Пусть, например,

$$f(x) = \begin{cases} 0 & \text{при } x < 1 \\ \dfrac{1}{x^2} & \text{при } x \geq 1. \end{cases}$$

Тогда

$$\int_{-\infty}^{\infty} f(x)dx = \int_{1}^{\infty} \frac{1}{x^2} dx = 1,$$

но

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

Рассмотренный пример функции плотности важен для нас еще в одном отношении. Функция

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

не дифференцируема в точке $x = 1$. Это означает, что формально мы не имели даже права называть f функцией плотности. Но это возражение легко обходится, если в определении функции плотности потребовать, чтобы f была не непрерывна, а кусочно-непрерывна, а равенство

$$F'(x) = f(x)$$

было справедливо для тех точек $x \in \boldsymbol{R}$, в которых функция f непрерывна.

При таком обобщении понятия функции плотности сохраняются все ее свойства (разумеется, за исключением непрерывности) и все приведенные выше соотношения. В случае необходимости мы будем использовать это более общее определение функции плотности без специальных оговорок.

З а м е ч а н и е. На интуитивном уровне понять теорему 1.7 достаточно легко. Выберем числа a_0 и a_k так, что

$$\int_{-\infty}^{\infty} x\, f(x)dx \approx \int_{a_0}^{a_k} x\, f(x)dx.$$

В данном случае символ “≈” не имеет строгого математического смысла и означает приближенное равенство. Разобьем отрезок $[a_0, a_k]$ на k равных частей точками $a_1, ..., a_{k-1}$;

$$a_0 < a_1 < a_2 < ... < a_{k-1} < a_k.$$

Через A_i обозначим событие, состоящее из тех $\omega \in \Omega$, для которых

$$a_{i-1} \le X(\omega) < a_i;$$

$i = 1, 2, ..., k$. Тогда

$$P(A_i) = \int_{a_{i-1}}^{a_i} f(x)dx.$$

Положим

$$x_i = \frac{a_{i-1} + a_i}{2}.$$

Тогда

$$\int_{a_0}^{a_k} x\, f(x)dx = \sum_{i=1}^{k} \int_{a_{i-1}}^{a_i} x\, f(x)dx \approx \sum_{i=1}^{k} x_i \int_{a_{i-1}}^{a_i} f(x)dx = \sum_{i=1}^{k} x_i P(A_i),$$

что аналогично выражению для ожидаемого значения случайной величины X, определенной на конечном пространстве элементарных событий Ω.

Из теорем 1.7 и 1.8 следует, что если две случайные величины имеют одинаковые дифференцируемые функции распределения, то они имеют одинаковые математические ожидания и одинаковые дисперсии. Выше мы установили такой же результат для кусочно-постоянных функций распределения. Этот результат остается справедливым для любых функций распределения. Обратное утверждение неверно: случайные величины могут иметь одинаковые ожидаемые значения и одинаковые дисперсии, но совершенно разные функции распределения.

Иногда вместо слов "случайная величина имеет заданную функцию распределения" мы будем говорить, что случайная величина имеет заданное *распределение* или заданное распределение вероятностей.

Пусть случайная величина X имеет функцию плотности $f(x)$.

Определение. Пусть $0 < p < 1$. Точка x_p называется *квантилью* порядка p случайной величины X, если

$$\int_{-\infty}^{x_p} f(y)dy = p.$$

Понятие квантили случайной величины будет широко использоваться в дальнейшем при проверке различных гипотез. Квантиль порядка 0,5 называется *медианой* случайной величины X.

Определение. Точка x_0 называется *модой* случайной величины X, если в точке x_0 функция $f(x)$ имеет локальный максимум.

На рис. 1.19 показаны функция плотности случайной величины, квантиль порядка p и мода. Здесь случайная величина имеет одну моду; в принципе мод может быть и больше. Однако чаще используются случайные величины с одной модой.

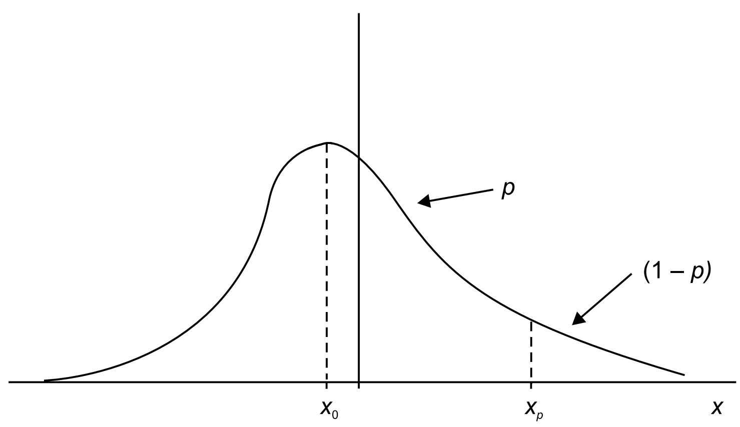


Рис. 1.19. Функция плотности случайной величины:
x_0 — мода случайной величины;
x_p — квантиль порядка p;
p — площадь фигуры, лежащей слева от прямой $x = x_p$;
$(1 - p)$ — площадь фигуры, лежащей справа от прямой $x = x_p$

1.5

Условные вероятности. Формула полной вероятности. Формула Байеса

Пусть события A и H принадлежат пространству элементарных событий Ω, причем $P(H) > 0$.

Определение. Число

$$P(A|H) = \frac{P(A \cap H)}{P(H)}$$

называется *условной вероятностью* события A при условии H (при гипотезе H).

Проиллюстрируем это определение для трех частных случаев.

1) Если $A \cap H = \varnothing$ (рис. 1.20), то

$$P(A \mid H) = 0.$$

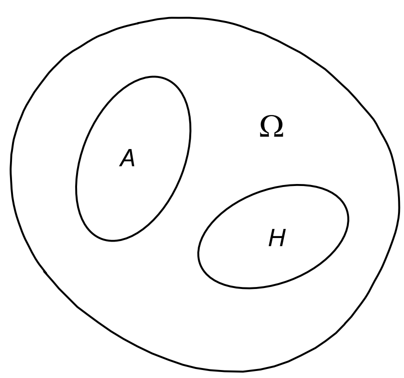


Рис. 1.20. Несовместимые события A и H

2) Если $H \subset A$ (см. рис. 1.21), то

$$A \cap H = H$$

и

$$P(A \mid H) = 1.$$

Случаи (1) и (2) являются крайними. В случае (1), если произошло событие H, то событие A произойти не может никак, и вероятность A при условии H равна 0. В случае (2), если произошло событие H, то заведомо произошло и A, поэтому вероятность A при условии H равна 1. Вообще же

$$0 \le P(A \mid H) \le 1,$$

поскольку $A \cap H \subset H$. Употребляя словосочетание “произошло событие”, мы говорим на повседневном языке, помогающем понять, откуда взялось определение условной вероятности, а не на языке строгой математической теории.

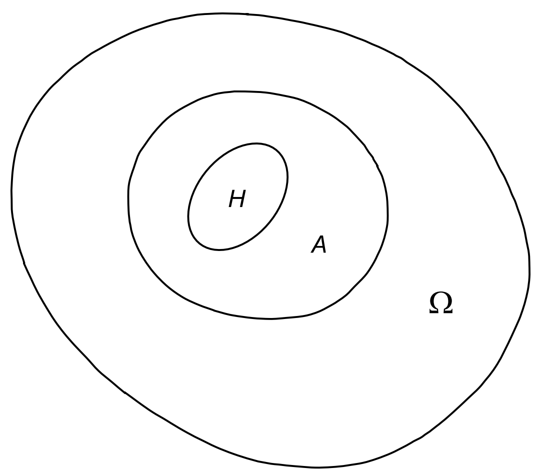


Рис. 1.21. Событие A является следствием события H

3) Если события A и H независимы, то

$$P(A \cap H) = P(A) \times P(H)$$

и

$$P(A \mid H) = P(A).$$

Пример 1.9. Одинаковы ли ответы на следующие вопросы.

1. Какова вероятность, что товар будет продан в течение дня?

2. Какова вероятность, что товар будет продан в течение дня, при условии, что он продается по цене на 20% ниже средней?

По-видимому, во втором случае вероятность выше. В данном случае событие A состоит в том, что товар будет продан в течение дня, а событие H состоит в том, что товар продается по цене на 20% ниже средней (см. рис. 1.22).

В ситуации, изображенной на рис. 1.22, событие $A \cap H$ (товар продается по цене на 20% ниже средней и будет продан в течение дня) почти что совпадает с событием H, поэтому вероятность

$$P(A \mid H) = \frac{P(A \cap H)}{P(H)}$$

близка к 1. В то же время вероятность $P(A)$ значительно меньше 1.

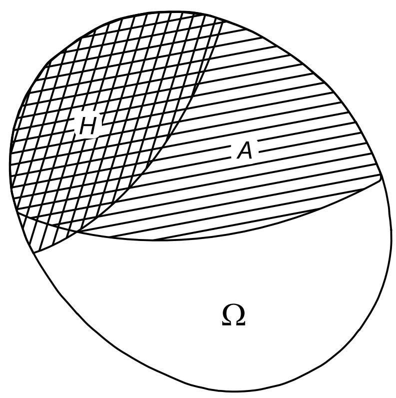


Рис. 1.22. События A и H, обладающие тем свойством, что вероятность наступления события A при условии H близка к 1

Пусть события H_1, H_2, ..., H_k попарно не пересекаются и

$$\bigcup_{i=1}^{k} H_i = \Omega$$

(см. рис. 1.23, где $k = 4$). Для любого события $A \subset \Omega$ справедлива формула

$$A = \bigcup_{i=1}^{k} (A \cap H_i).$$

Поскольку события $A \cap H_i$ попарно не пересекаются, по теореме о сложении вероятностей

$$P(A) = \sum_{i=1}^{k} P(A \cap H_i).$$

Воспользовавшись тем, что

$$P(A \cap H_i) = P(A \mid H_i)\, P(H_i),$$

получаем

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

Данная формула называется *формулой полной вероятности.*

Эта формула оказывается очень полезной в ряде случаев.

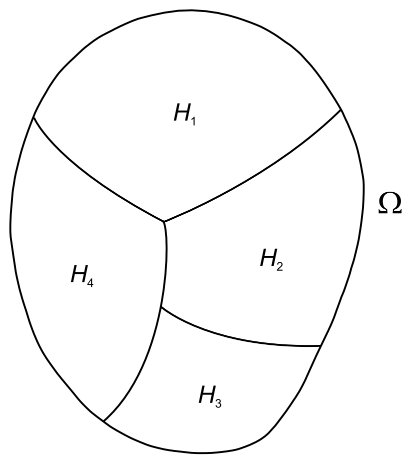


Рис. 1.23. Разбивка пространства элементарных событий Ω на попарно несовместимые события H_1, H_2, H_3, H_4

Пример 1.10. При бросании монеты с вероятностью p выпадает герб и с вероятностью $q = 1 - p$ выпадает решетка. Игрок, имеющий m руб., при каждом броске либо выигрывает 1 руб., если выпадает герб, либо проигрывает 1 руб., если выпадает решетка. Цель игрока состоит в том, чтобы увеличить свой капитал до M руб. Игра прекращается либо когда капитал игрока составит M руб. (игрок выиграл), либо когда капитал игрока составит 0 руб. (игрок проиграл). Какова вероятность, что игрок проиграет?

Введем следующие обозначения:

$\pi(k)$ — вероятность того, что имея k руб., игрок проиграет (греческая буква “пи”);

H_1 — событие, состоящее в том, что при первом броске выпал герб;

H_2 — событие, состоящее в том, что при первом броске выпала решетка;

A — событие, состоящее в том, что игрок проиграл.

Тогда

$$P(A \mid H_1) = \pi(m+1), \;\; P(A \mid H_2) = \pi(m-1)$$

и по формуле полной вероятности

$$\pi(m) = \pi(m+1)\,p + \pi(m-1)q.$$

Заметим, что

$$\pi(0) = 1, \; \pi(M) = 0.$$

Если монета доброкачественная, т.е. $p = q = \frac{1}{2}$, то задача допускает точное решение. В этом случае π — линейная функция от m, так как точка (m, π (m)) является серединой отрезка с концами в точках (m — 1, π (m — 1)) и (m + 1, π (m + 1)). Действительно,

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

и

$$\pi(m) = \pi(m-1)\frac{1}{2} + \pi(m+1)\frac{1}{2}.$$

Поскольку значения функции π на концах отрезка $[0, M]$ известны, находим

$$\pi(m) = 1 - \frac{m}{M}.$$

Вероятность проигрыша найдена. Естественно, что она тем больше, чем меньше отношение m к M. Как может быть использована найденная вероятность для определения наиболее выгодного рубежа M? Пусть случайная величина X означает выигрыш игрока. Тогда на подмножестве A случайная величина X принимает значение $(-m)$, а на оставшейся части пространства элементарных событий — значение $(M - m)$. Ожидаемое значение выигрыша

$$E(X) = (-m)\left(1 - \frac{m}{M}\right) + (M - m)\frac{m}{M} = 0.$$

То есть при доброкачественной монете выбор рубежа M не влияет на ожидаемое значение выигрыша.

Пусть теперь p не обязательно равно q. Тогда система уравнений

$$\begin{cases} \pi(m-1)q - \pi(m) + \pi(m+1)p = 0, \ 0 < m < M \\ \pi(0) = 1, \ \pi(M) = 0 \end{cases}$$

может быть решена с использованием компьютера.

В следующей таблице приведены полученные значения вероятностей проигрыша π (m) при $M = 10$, соответствующие начальному капиталу в m рублей ($0 \le m \le 10$) и вероятностям выпадения герба p, равным 0,6; 0,55; 0,52; 0,5; 0,48; 0,45; 0,4. Все вероятности приведены с точностью до двух десятичных знаков.

$p \backslash m$	0	1	2	3	4	5	6	7	8	9	10
0,60	1,00	0,66	0,43	0,28	0,18	0,12	0,07	0,04	0,02	0,01	0,00
0,55	1,00	0,79	0,62	0,48	0,36	0,27	0,19	0,13	0,08	0,03	0,00
0,52	1,00	0,86	0,73	0,61	0,50	0,40	0,31	0,22	0,14	0,07	0,00
0,50	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,00
0,48	1,00	0,93	0,86	0,78	0,69	0,60	0,50	0,39	0,27	0,14	0,00
0,45	1,00	0,97	0,92	0,87	0,81	0,73	0,64	0,52	0,38	0,21	0,00
0,40	1,00	0,99	0,98	0,96	0,93	0,88	0,82	0,72	0,57	0,34	0,00

Естественно, что при $p = 0{,}5$ результаты расчета совпадают с результатами, полученными из точной формулы

$$\pi(m) = 1 - \frac{m}{M}.$$

Интересно, что уменьшение вероятности выпадения герба всего с $p = 0{,}52$ до $p = 0{,}48$ увеличивает вероятность проигрыша при $m = 9$ примерно в два раза, а уменьшение вероятности выпадения герба с $p = 0{,}6$ до $p = 0{,}4$ — примерно в тридцать раз. Отметим также, что при $p = 0{,}6$ и $p = 0{,}55$ вероятность проигрыша меньше 0,5 даже при $m = 2$ и $m = 3$ соответственно.

Закончив с примером 1.10, снова вернемся к формулам

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

где события $H_1, H_2, \ldots, H_k$ попарно не пересекаются и при объединении дают все множество Ω. Заметим, что если $P(A) > 0$, то

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

Так как $P(A \cap H_i) = P(A \mid H_i)P(H_i)$, получаем

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

Выражая $P(A)$ по формуле полной вероятности, получаем

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

Данная формула называется *формулой Байеса.*

Пример 1.11. В начале года аналитик рекомендовал некоторые акции для приобретения. По результатам года определилось, что для 40% от общего числа акций доходность оказалась выше, чем средняя по рынку, а для 60% от общего числа акций доходность оказалась ниже, чем средняя по рынку. Условимся называть первые акции акциями с высокой доходностью, а вторые акции — акциями с низкой доходностью. 20% из числа акций с высокой доходностью были рекомендованы аналитиком для приобретения и 5% из числа акций с низкой доходностью были рекомендованы аналитиком для приобретения. Какова вероятность, что акция, рекомендованная аналитиком для приобретения, оказалась акцией с высокой доходностью?

Пусть

A — множество акций, рекомендованных аналитиком для приобретения;

H_1 — множество акций с высокой доходностью;

H_2 — множество акций с низкой доходностью.

Тогда

$$P(H_1) = 0{,}4 \qquad P(H_2) = 0{,}6,$$

$$P(A \mid H_1) = 0{,}2 \qquad P(A \mid H_2) = 0{,}05.$$

Поэтому

$$P(H_1|A) = \frac{P(A|H_1)P(H_1)}{P(A|H_1)P(H_1) + P(A|H_2)P(H_2)} = \frac{0{,}2 \times 0{,}4}{0{,}2 \times 0{,}4 + 0{,}5 \times 0{,}6} = \frac{8}{11}.$$

Пусть случайная величина

$$X : \Omega \to \boldsymbol{R}$$

принимает k различных значений x_1, x_2, ..., x_k. Обозначим через A_i подмножество Ω, на котором случайная величина X равна x_i;

$$\Omega = \bigcup_{i=1}^{k} A_i.$$

Из определения ожидаемого значения случайной величины нетрудно увидеть, что

$$E(X) = \sum_{i=1}^{k} x_i P(A_i).$$

Пусть $H \subset \Omega$ и $P(H) > 0$.

Определение. *Условным ожиданием* случайной величины X при заданном H называется следующее число

$$E(X|H) = \sum_{i=1}^{k} x_i P(A_i|H).$$

Пример 1.12. Доходность первого актива может быть либо –5%, либо 20%. Доходность второго актива может быть либо 0%, либо 10%. *Совместное распределение вероятностей* доходностей двух активов приведено в следующей таблице.

	–5%	20%
0%	0,1	0,3
10%	0,4	0,2

Какова ожидаемая доходность первого актива при условии, что доходность второго актива равна 0%?

Рассмотрим следующие события.

A_1 — доходность первого актива равна –5%;

A_2 — доходность первого актива равна 20%;

H — доходность второго актива равна 0%.

Тогда, поскольку $P(H) = P(A_1 \cap H) + P(A_2 \cap H) = 0,1 + 0,3 = 0,4$, имеем

$$P(A_1|H) = \frac{P(A_1 \cap H)}{P(H)} = \frac{0,1}{0,4} = 0,25,$$

$$P(A_2|H) = \frac{P(A_2 \cap H)}{P(H)} = \frac{0,3}{0,4} = 0,75$$

и

$$E(X \mid H) = -5 \times 0,25 + 20 \times 0,75 = 13,75\%.$$

В то же время

$$E(X) = -5 \times 0,5 + 20 \times 0,5 = 7,5\%.$$

1.6

Задача кавалера де Мере

Эта задача заслуживает подробного изложения по двум причинам. Во-первых, она вошла в историю математики, как одна из первых задач теории вероятностей[1]. Во-вторых, она представляет собой пример применения вероятностных методов для решения совсем неочевидной прикладной задачи, хотя и относящейся к области азартных игр.

Кавалер де Мере не был математиком. Он был заядлым игроком. Его любимая игра (назовем ее игрой № 1) состояла в следующем. Кавалер де Мере и его партнер бросают кость 4 раза. Если хотя бы один раз выпадает шестерка, выигрывает кавалер де Мере. Если нет — его партнер. Кавалер де Мере не мог объяснить в чем дело, но знал, что эта игра (хотя и не очень заметно) идет на пользу его кошельку. Однако со временем и другие игроки заметили, что эта игра идет на пользу кошельку кавалера де Мере больше, чем их собственным, поэтому находилось все меньше желающих играть с ним.

Тогда кавалер де Мере придумал игру № 2. Две кости бросаются 24 раза, и если хоть раз выпадают две шестерки — побеждает кавалер де Мере, если нет — его партнер. Увеличив число бросков в шесть раз, кавалер де Мере считал, что он просто играет в игру № 1 шесть раз подряд. Однако ситуация с кошельками изменилась на противоположную.

Расстроенный кавалер де Мере обратился к математику Паскалю за разъяснениями. Паскаль решил эту задачу и объяснил, почему вероятность выигрыша в первой игре больше 1/2, а вероятность выигрыша во второй игре меньше 1/2, обессмертив тем самым имя де Мере и сделав важный шаг в развитии теории вероятностей.

Игра № 1. Множество Ω состоит из 6^4 элементарных событий:

$$(1, 1, 1, 1), (1, 1, 1, 2), ..., (1, 1, 1, 6),$$

---

[1] См.: Никифоровский В.А, Фрейман Л.С. Рождение новой математики. М.: Наука, 1976. С. 116—117.

$$(1, 1, 2, 1), ..., (6, 6, 6, 5), (6, 6, 6, 6),$$

где для каждого элементарного события на первом месте стоит результат первого броска, на втором — второго, на третьем — третьего, на четвертом — четвертого. При этом событие A (кавалер де Мере проиграл) содержит 5^4 элементарных событий, так как неблагоприятных исходов при каждом броске может быть пять. Поэтому

$$P(A) = \frac{5^4}{6^4} = \frac{625}{1296} \approx 0,491 < \frac{1}{2}.$$

Игра № 2. Исходя из тех же соображений, что и в игре № 1, в этом случае

$$P(A) = \left(\frac{35}{36}\right)^{24}.$$

Чтобы сравнить это число с 1/2, следует перейти к логарифмам. Ответ следует из цепочки неравенств

$$\ln\left(\frac{36}{35}\right)^{24} = 24\ln\frac{36}{35} < 0,68 < \ln 2,$$

поэтому

$$\left(\frac{36}{35}\right)^{24} < 2$$

и

$$P(A) > \frac{1}{2}.$$

а именно $P(A) \approx 0,509$.

Этим и объясняется причина неудач кавалера де Мере.

Шведов, А. С.

Ш 34 Теория вероятностей и математическая статистика [Текст] : учеб. пособие для вузов / А. С. Шведов; Гос. ун-т — Высшая школа экономики. — 2-е изд., перераб. и доп. — М.: Изд. дом ГУ ВШЭ, 2005. — 254, [I, II] с. — (Учебники Высшей школы экономики). — Литерат.: с. 247—249. — Предм. указ.: с. 250—253. — 3000 экз. — ISBN 5-7598-0214-3 (в пер.).

Книга является учебным пособием по курсу "Теория вероятностей и математическая статистика". Изложение ведется в доступной форме, при этом понятия "событие", "вероятность", "случайная величина" используются в том же смысле, что и в научных трудах по математике. В пособие включено большое количество примеров, позволяющих понять принципы применения многих методов теории вероятностей и математической статистики для решения экономических и других задач.

Для студентов и аспирантов специальностей "Экономика" и "Менеджмент", других экономических специальностей, а также для интересующихся возможностями применения теории вероятностей и математической статистики в условиях рыночной экономики.

УДК 519.2(075)
ББК 22.17

Учебное издание

Шведов Алексей Сергеевич

**Теория вероятностей
и математическая статистика**

Редактор *Е.А. Рязанцева*
Художественный редактор *А.М. Павлов*
Корректор *Е.Е. Андреева*
Компьютерная верстка и графика *Н.Е. Пузанова*

ЛР № 020832 от 15 октября 1993 г. продлена до 14 октября 2003 г.
Подписано в печать 15.11.2004. Формат 60×88 $^{1}/_{16}$. Бумага офсетная.
Гарнитура Times New Roman. Печать офсетная. Тираж 3000 экз. Усл. печ. л. 15,46.
Уч.-изд. л. 7,57. Заказ № . Изд. № 254

ГУ ВШЭ. 125319, Москва, Кочновский проезд, 3
Тел.: (095) 134-16-41; 134-08-77
Факс: (095) 134-08-31