
У Ч Е Б Н И К И
ВЫСШЕЙ ШКОЛЫ ЭКОНОМИКИ

ВШЭ
HSE

У Ч Е Б Н И К И
ВЫСШЕЙ ШКОЛЫ ЭКОНОМИКИ

ВШЭ
HSE

А.С. Шведов

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА**

Промежуточный уровень



Издательский дом
Высшей школы экономики

Москва, 2016

УДК 519.2
ББК 22.17я7
Ш34

Рецензенты:

член-корреспондент РАН, доктор физико-математических наук, профессор
департамента прикладной экономики факультета экономических наук НИУ ВШЭ
И.Г. Поспелов;

доктор технических наук, профессор кафедры бизнес-аналитики факультета
бизнеса и менеджмента НИУ ВШЭ *А.П. Кирсанов*

Шведов, А. С. Теория вероятностей и математическая статистика:
Ш34 промежуточный уровень [Текст] : учеб. пособие / А. С. Шведов ; Нац.
исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей
школы экономики, 2016. — (Учебники Высшей школы экономики). —
280 с. — 600 экз. — ISBN 978-5-7598-1301-9 (в пер.).

Это пособие промежуточного уровня является продолжением учебного пособия А.С. Шведова «Теория вероятностей и математическая статистика» вводного уровня (М.: Изд. дом ВШЭ, 1995; 2005). Вместе эти две книги представляют собой расширенный курс лекций, в разное время читавшийся автором в бакалавриате факультетов экономики, бизнес-информатики и компьютерных наук НИУ ВШЭ, и являются базовым учебным пособием при преподавании дисциплины «Теория вероятностей и математическая статистика» на втором и третьем курсах факультета бизнес-информатики и на втором курсе факультета компьютерных наук. Выбор тем в настоящей книге традиционен для курсов теории вероятностей и математической статистики. Это совместные распределения вероятностей и моменты случайных величин, проверка гипотез, оценки параметров распределений, доверительные интервалы. Рассматриваются также исследование выборками, цепи Маркова, теория статистических решений. Включено значительное количество примеров прикладного характера, из которых многие доведены до окончательных числовых ответов. Эти примеры предназначены не только для иллюстрации теории, они приближены к тем задачам, с которыми выпускники смогут встретиться в будущей работе. В этом отношении книга похожа на учебники по статистике, используемые в западных университетах.

Для студентов и аспирантов, специализирующихся по экономике, бизнес-информатике, компьютерным наукам.

УДК 519.2
ББК 22.17я7

ISBN 978-5-7598-1301-9

© Шведов А.С., 2016
© Оформление. Издательский дом
Высшей школы экономики, 2016

ОГЛАВЛЕНИЕ

Предисловие	7
Глава 1. Совместные распределения вероятностей	11
1.1. Совместные функции распределения и маргинальные функции распределения	11
1.2. Совместные дискретные распределения вероятностей	22
1.3. Совместные непрерывные распределения вероятностей	30
1.4. Ожидания и распределения функций случайных величин	43
1.5. Условные ожидания	55
1.6. Двумерное нормальное распределение	60
Глава 2. Метод моментов и метод максимального правдоподобия ...	69
2.1. Метод моментов	71
2.2. Метод максимального правдоподобия	74
2.3. Информация Фишера	86
2.4. Достаточные статистики	98
Глава 3. Доверительные интервалы и проверка гипотез	112
3.1. Доверительные интервалы для ожидания и для дисперсии	113
3.2. Доверительные интервалы для разности ожиданий и для отношения дисперсий	121
3.3. Доверительные интервалы для доли и для разности долей	130
3.4. Доверительные множества	134
3.5. Связь проверки гипотез с построением доверительных интервалов	139
3.6. Проверка гипотез и доверительные интервалы для корреляции ..	147
3.7. Эмпирические функции распределения и статистические критерии Колмогорова–Смирнова	154
3.8. Байесовские доверительные интервалы	158
Глава 4. Датчики случайных чисел	163
4.1. Датчики равномерно распределенных случайных чисел	164
4.2. Случайные числа, соответствующие различным функциям распределения	171

Глава 5. Исследование выборками	179
5.1. Алгоритмы простого случайного выбора	180
5.2. Выборочное среднее и выборочная дисперсия при простой случайной выборке.	185
5.3. Стратифицированная случайная выборка.	190
Глава 6. Цепи Маркова.	196
6.1. Основные определения	196
6.2. Примеры и свойства	200
Глава 7. Проверка гипотез и оценка параметров как задачи теории статистических решений	212
7.1. Функции потерь и функции риска	212
7.2. Допустимые и минимаксные решающие правила.	217
7.3. Априорные и апостериорные распределения.	220
7.4. Байесовские риски и байесовские решающие правила.	224
7.5. Допустимые, минимаксные и байесовские решающие правила при двух возможных состояниях среды	227
7.6. Рандомизированные решения	232
7.7. Неинформативные априорные распределения.	234
7.8. Примеры статистических оценок, являющихся байесовскими решающими правилами	235
7.9. Задачи с конечным числом возможных решений и проверка гипотез	241
7.10. Определение оптимального размера выборки до начала наблюдений	244
7.11. Определение оптимального размера выборки в процессе наблюдений	248
Приложения. Некоторые одномерные распределения вероятностей	255
П.1. Отрицательное биномиальное распределение	255
П.2. Гипергеометрическое распределение	260
П.3. Гамма-распределение	264
П.4. Бета-распределение	268
Библиографическая справка	271
Список литературы	275
Предметный указатель	277

ПРЕДИСЛОВИЕ

Данная книга является продолжением книги автора «Теория вероятностей и математическая статистика», первое издание которой вышло в 1995 г., а второе — в 2005 г. Но именно продолжением, независимой книгой, а не второй частью, несмотря на то что при ссылках мы будем называть предыдущую книгу [ТВиМС-1]. Первое издание настоящей книги вышло в 2007 г.

Книга [ТВиМС-1] носит вводный характер и предназначена для студентов, которые не выбрали математику своей основной специальностью. Для автора было важным, чтобы студенты, прочитав книгу, испытали чувство победы, ощутили, если этого еще не произошло раньше, вкус к математике. А для этого книга должна быть написана не только понятно и ясно, но и быть достаточно короткой.

В то же время соответствие существующим международным требованиям к курсу «Statistics» при подготовке бакалавров по специальности «Economics», место курса «Теория вероятностей и математическая статистика» в учебных планах Национального исследовательского университета «Высшая школа экономики», требовали включения в данный курс еще ряда разделов. Выходом из такой ситуации стало написание еще одной книги.

Здесь необходимо пояснить взаимоотношение нашего курса «Теория вероятностей и математическая статистика» и традиционного курса «Statistics», занимающего одно из главных мест при подготовке бакалавров по специальности «Economics» в развитых странах со сложившейся рыночной системой. Центральные темы предмета «Statistics» — это проверка гипотез, оценка параметров, построение доверительных интервалов. Многочисленные западные учебники по курсу «Statistics» представляют перечисленные темы и включают в разных соотношениях следующие три блока вопросов: вероятность, математическая статистика, применение методов математической статистики. Третий блок вопросов обычно представлен в учебниках по «Statistics» очень большим числом тщательно сделанных миниатюр, частью в виде примеров, частью в виде задач, иллюстрирующих применение методов математической статистики в хозяйственной деятельности, менеджменте, социологии, финансах. В своей будущей работе, встречаясь с аналогичными задачами, специалист должен узнавать знакомые с университетских времен примеры и понимать, как и зачем в этих случаях применяются методы математической статистики. А вот вопросы, которые могут быть рассмотрены в курсах «Macroeconomics», в курсы «Statistics», как правило, не включаются. В определенной степени этот третий блок вопросов представлен и в данной, и в предыдущей книгах автора.

Негативную роль для российского экономического образования играет то, что и «Statistics» (статистическая наука), и «statistical data» (статистические данные) переводятся на русский язык одним и тем же словом «статистика». Хотя в повседневной жизни слово «statistics» и может использоваться для обозначения статистических данных, в университетах «Statistics» — это именно статистическая наука, теория статистического вывода.

Позиция автора заключается в том, что Россия пока не стала признанным законодателем мод при подготовке студентов по специальности «Economics», и к попыткам изменить существующую западную педагогическую традицию при подготовке студентов по этой специальности следует относиться очень осторожно.

Также негативную роль для российского экономического образования играет и то, что и «Economics» (экономическая наука), и «economy» (хозяйственная деятельность) опять же переводятся на русский язык одним и тем же словом «экономика». Это служит источником многих недоразумений. В странах со сложившимся рыночным хозяйством при подготовке студентов по специальности «Economics» не делается попыток включить в учебные планы все, что может оказаться полезным специалисту по хозяйственной деятельности, химию, историю, право, культурологию. Аналогично тому, как при подготовке студентов-физиков их учат именно физике, несмотря на то что в будущей работе им могут быть полезными знания и по праву, и по менеджменту, и многое другое. Но если включить в программу обучения все это, нельзя будет подготовить полноценных специалистов по физике.

Такое некороткое обсуждение взаимоотношения курса «Теория вероятностей и математическая статистика» и традиционного курса «Statistics» необходимо для того, чтобы объяснить сделанный в настоящей книге отбор материала и подход к изложению.

Хотя для обучения по данной книге студент должен быть значительно лучше подготовленным в математике, чем для обучения по книге [ТВиМС-1], это все же еще не тот читатель, которому можно давать жесткий, профессиональный математический текст. Книга написана, насколько это возможно для выбранного материала, в щадящем стиле. Выкладки проводятся достаточно подробно. Все ссылки на другие книги (в том числе и с указанием того, где можно найти доказательства, не приведенные в тексте) вынесены в библиографическую справку в конце книги. Нумеруются лишь немногие формулы. Многочисленные примеры занимают едва ли не столько же места, сколько основной текст. Но при этом, скажем, двойные индексы, от которых избавлен читатель [ТВиМС-1], здесь используются совершенно свободно. За очень небольшим исключением в данной книге не повторяется материал, содержащийся в [ТВиМС-1].

Позиция автора состоит в том, что недостаточная подготовка по теории функций нескольких действительных переменных или нетвердое знание того, что такое счетное множество, не должны превращать для студента теорию вероятностей и математическую статистику в закрытую территорию. (Пусть читатель, являющийся профессиональным математиком, подумает, может ли он сам понять формулировку теоремы, если он не уверен полностью в значении хотя бы одного термина.) В книге [ТВиМС-1] результаты, относящиеся к теории функций нескольких действительных переменных, не используются, понятие счетного множества используется очень осторожно. В настоящей книге таких ограничений нет. В главе 1 рассматриваются совместные распределения случайных величин, и здесь результаты теории функций нескольких действительных переменных становятся необходимыми. Отдельно изучаются совместное дискретное распределение вероятностей и совместное непрерывное распределение вероятностей. Условное ожидание рассматривается не только как число, но и как случайная величина.

Как уже было сказано, центральные темы предмета «Statistics» — это проверка гипотез, оценка параметров, построение доверительных интервалов. Данные темы рассматриваются в главах 2 и 3. Мы не будем здесь перечислять все рассматриваемые в этих главах задачи, вместо этого остановимся лишь на одном вопросе, который часто задают студенты и который относится ко всем типам задач из глав 2 и 3. (Формулируя здесь этот вопрос, мы исходим из того, что читатель не впервые в жизни слышит о проверке гипотез, оценке параметров, построении доверительных интервалов.) Какой длины n должен быть набор наблюдений x_1, x_2, \dots, x_n , чтобы можно было применять ту или иную процедуру проверки гипотезы или нахождения неизвестного параметра? Отвечать приходится вопросом. А что вы имеете в виду, грубую прикидку или прецизионный расчет? Или что-то среднее? Чем выше требования к точности, тем больше должно быть n . Кроме того, при одних и тех же требованиях к точности требования к числу наблюдений n в разных задачах оказываются очень разными. Когда число наблюдений n равняется нескольким десяткам или нескольким сотням, для одних задач результаты получаются достаточно высокой точности, а для других — не очень.

Материал главы 4 «Датчики случайных чисел» необходим для понимания главы 5 «Исследование выборками».

В главе 6 рассматриваются последовательности случайных величин, называемые цепями Маркова. Такие последовательности случайных величин используются в экономической науке.

В главе 7 показана связь проверки гипотез и оценивания параметров с теорией статистических решений (этой главы не было в первом издании книги).

Автор с симпатией относится к непримиримости молодых математиков к любым нестрогостям; он и сам когда-то был таким. Но это — книга другого стиля. Она не предназначена для подготовки бакалавров по математике. Даже, можно сказать, что не годится для этого. Там именно должна воспитываться непримиримость к любым нестрогостям, разрешить использовать рассуждения на эвристическом уровне строгости означает испортить студента.

Но безупречно строгий курс по теории вероятностей — это только курс продвинутого уровня. Данную дисциплину студенты-математики изучают обычно не раньше третьего курса, и этому предшествуют два года напряженной подготовки исключительно по математике. Курс по данному предмету вводного или промежуточного уровня — это всегда компромисс между строгостью и простотой. И чересчур высоко или непродуманно поставленная планка при вероятностно-статистической подготовке бакалавров по специальности «Econometrics» приводит к провальным результатам. (Книгу [ТВиМС-1] и данную можно сравнить с двумя планками для прыжков на уроке физкультуры у школьников, поставленными на разной высоте.) Принцип, что строгость не враг простоты, а неперемное ее условие, в данном случае не срабатывает.

СОВМЕСТНЫЕ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

В этой главе рассматриваются различные функции распределения, связанные с несколькими случайными величинами: совместные, маргинальные, условные. Отдельно даются совместное дискретное распределение и совместное непрерывное распределение. Вводятся ожидания и условные ожидания функций от случайных величин. Изучается двумерное нормальное распределение.

1.1. СОВМЕСТНЫЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И МАРГИНАЛЬНЫЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

Напомним, что *случайная величина* — это функция, определенная на *пространстве элементарных событий* Ω и принимающая значения в множестве действительных чисел R . *Элементарное событие* — это элемент множества Ω , *событие* — это подмножество множества Ω . *Вероятность события* $A \subseteq \Omega$ — это число $P(A)$, удовлетворяющее условию $0 \leq P(A) \leq 1$. Подробнее об этих понятиях см. [ТВиМС-1, разделы 1.2, 1.3].

Рассматривая некоторый набор случайных величин, мы всегда будем считать, что все случайные величины X_1, X_2, \dots, X_k определены на одном и том же пространстве элементарных событий Ω , т.е.

$$X_1: \Omega \rightarrow R, \quad X_2: \Omega \rightarrow R, \quad \dots, \quad X_k: \Omega \rightarrow R.$$

Тогда набор случайных величин X_1, X_2, \dots, X_k называют также *векторнозначной случайной величиной* или *случайным вектором*.

Рассмотрим k действительных чисел x_1, x_2, \dots, x_k . Другими словами можно сказать, что рассматривается k -мерный вектор

$$(x_1, x_2, \dots, x_k) \in R^k,$$

где R^k — k -мерное евклидово пространство. Событие

$$\{\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \dots \cap \{X_k \leq x_k\}\}$$

для краткости будем обозначать

$$\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k\}.$$

Напомним, что через $\{X_j \leq x_j\}$ обозначается событие (подмножество множества Ω) вида

$$\{\omega \in \Omega: X_j(\omega) \leq x_j\}, \quad j = 1, 2, \dots, k.$$

Положим

$$F(x_1, x_2, \dots, x_k) = \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k),$$

где через $\mathbf{P}(A)$ обозначается вероятность события A (в данном случае фигурные скобки в дополнение к круглым не пишутся). Функция

$$F: R^k \rightarrow [0, 1]$$

называется *совместной функцией распределения* случайных величин X_1, X_2, \dots, X_k .

Замечание. Правильнее (что мы впоследствии и будем делать) было бы обозначать введенную функцию не $F(x_1, x_2, \dots, x_k)$, а $F_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k)$, чтобы подчеркнуть ее связь со случайными величинами X_1, X_2, \dots, X_k .

Для того чтобы сделать формулы менее громоздкими, мы ограничимся рассмотрением случая $k = 2$. Никаких трудностей принципиального характера при переходе от случая $k = 2$ к общему случаю не возникает.

Заметим, что для любых случайных величин X_1 и X_2 и при любом фиксированном $x_2 \in R$ функция

$$F_{X_1, X_2}(x_1, x_2)$$

как функция аргумента x_1 является монотонно неубывающей. Действительно, если $x'_1 < x''_1$, то для множеств

$$A' = \{\omega \in \Omega: X_1(\omega) \leq x'_1, X_2(\omega) \leq x_2\}$$

и

$$A'' = \{\omega \in \Omega: X_1(\omega) \leq x''_1, X_2(\omega) \leq x_2\}$$

имеет место соотношение $A' \subseteq A''$. Поэтому $\mathbf{P}(A') \leq \mathbf{P}(A'')$ и, значит,

$$F_{X_1, X_2}(x'_1, x_2) \leq F_{X_1, X_2}(x''_1, x_2).$$

Аналогично, при любом фиксированном $x_1 \in R$ функция

$$F_{X_1, X_2}(x_1, x_2)$$

как функция аргумента x_2 является монотонно неубывающей.

В дальнейшем нам понадобится следующая теорема, называемая *теоремой о непрерывности вероятности*.

Теорема 1.1. Для сужающейся последовательности событий

$$A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots$$

существует

$$\lim_{n \rightarrow \infty} \mathbf{P}(A_n)$$

и этот предел равен $\mathbf{P}(A)$, где событие

$$A = \bigcap_{n=1}^{\infty} A_n.$$

Для расширяющейся последовательности событий

$$B_1 \subseteq B_2 \subseteq \dots \subseteq B_n \subseteq \dots$$

существует

$$\lim_{n \rightarrow \infty} \mathbf{P}(B_n)$$

и этот предел равен $\mathbf{P}(B)$, где событие

$$B = \bigcup_{n=1}^{\infty} B_n.$$

Доказательство теоремы 1.1 мы приводить не будем.

Теорема 1.2. При любом $x_1 \in R$

$$\lim_{x_2 \rightarrow -\infty} F_{X_1, X_2}(x_1, x_2) = 0.$$

При любом $x_2 \in R$

$$\lim_{x_1 \rightarrow -\infty} F_{X_1, X_2}(x_1, x_2) = 0.$$

Доказательство. Выберем последовательность действительных чисел $x_2^{(n)}$, стремящуюся к $-\infty$ при $n \rightarrow \infty$. Положим

$$A_n = \{X_1 \leq x_1, X_2 \leq x_2^{(n)}\}.$$

Тогда

$$\bigcap_{n=1}^{\infty} A_n = \emptyset,$$

где \emptyset — пустое множество. По определению совместной функции распределения и по теореме 1.1

$$\lim_{n \rightarrow \infty} F_{X_1, X_2}(x_1, x_2^{(n)}) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\emptyset) = 0.$$

Аналогично доказывается равенство нулю и второго предела. □

Теорема 1.3.

$$\lim_{x \rightarrow \infty} F_{X_1, X_2}(x, x) = 1.$$

Доказательство. Выберем последовательность действительных чисел $x^{(n)}$, стремящуюся к ∞ при $n \rightarrow \infty$. Положим

$$A_n = \{X_1 \leq x^{(n)}, X_2 \leq x^{(n)}\}.$$

Тогда

$$\bigcup_{n=1}^{\infty} A_n = \Omega.$$

По определению совместной функции распределения и по теореме 1.1

$$\lim_{n \rightarrow \infty} F_{X_1, X_2}(x^{(n)}, x^{(n)}) = \lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\Omega) = 1. \quad \square$$

Введем обозначения для пределов на $+\infty$; при фиксированном $x_1 \in R$ положим

$$F_{X_1, X_2}(x_1, \infty) = \lim_{x_2 \rightarrow \infty} F_{X_1, X_2}(x_1, x_2)$$

и при фиксированном $x_2 \in R$ положим

$$F_{X_1, X_2}(\infty, x_2) = \lim_{x_1 \rightarrow \infty} F_{X_1, X_2}(x_1, x_2).$$

Данные пределы существуют, поскольку во всех случаях рассматриваются пределы на бесконечности для монотонно неубывающих функций, не превосходящих 1.

По теореме 1.1 при фиксированном $x_1 \in R$

$$\lim_{x_2 \rightarrow \infty} \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbf{P}(X_1 \leq x_1, X_2 < \infty) = \mathbf{P}(X_1 \leq x_1)$$

и при фиксированном $x_2 \in R$

$$\lim_{x_1 \rightarrow \infty} \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbf{P}(X_1 < \infty, X_2 \leq x_2) = \mathbf{P}(X_2 \leq x_2).$$

Вероятности в правых частях двух последних равенств — это значения функций распределения случайных величин X_1 и X_2 , рассматриваемых по отдельности:

$$F_{X_1}(x_1) = \mathbf{P}(X_1 \leq x_1), \quad F_{X_2}(x_2) = \mathbf{P}(X_2 \leq x_2).$$

Таким образом,

$$F_{X_1, X_2}(x_1, \infty) = F_{X_1}(x_1), \quad F_{X_1, X_2}(\infty, x_2) = F_{X_2}(x_2).$$

Функции $F_{X_1}(x_1)$ и $F_{X_2}(x_2)$ называются *маргинальными функциями распределения* случайных величин X_1 и X_2 .

Пример 1.1. Пусть случайная величина X_1 может принимать значения x'_1 и x''_1 , а случайная величина X_2 может принимать значения x'_2 и x''_2 , причем $x'_1 < x''_1$ и $x'_2 < x''_2$. Допустим, что вероятности событий

$$\begin{aligned} \mathbf{P}(X_1 = x'_1, X_2 = x'_2) &= 0.3, & \mathbf{P}(X_1 = x'_1, X_2 = x''_2) &= 0.4, \\ \mathbf{P}(X_1 = x''_1, X_2 = x'_2) &= 0.1, & \mathbf{P}(X_1 = x''_1, X_2 = x''_2) &= 0.2. \end{aligned}$$

Представим эти вероятности в виде следующей таблицы, где в нижней строке приведены вероятности $\mathbf{P}(X_1 = x'_1)$ и $\mathbf{P}(X_1 = x''_1)$, а в правом столбце — вероятности $\mathbf{P}(X_2 = x'_2)$ и $\mathbf{P}(X_2 = x''_2)$:

	x'_1	x''_1	
x'_2	0.3	0.1	0.4
x''_2	0.4	0.2	0.6
	0.7	0.3	

В следующей таблице приводятся значения совместной функции распределения F_{X_1, X_2} в точках (x'_1, x'_2) , (x'_1, x''_2) , (x''_1, x'_2) , (x''_1, x''_2) . Например,

$$\begin{aligned} F_{X_1, X_2}(x''_1, x'_2) &= \mathbf{P}(X_1 \leq x''_1, X_2 \leq x'_2) = \\ &= \mathbf{P}(X_1 = x'_1, X_2 = x'_2) + \mathbf{P}(X_1 = x''_1, X_2 = x'_2) = 0.3 + 0.1 = 0.4. \end{aligned}$$

В нижней строке показаны значения маргинальной функции распределения F_{X_1} в точках x'_1 и x''_1 . В правом столбце показаны значения маргинальной функции распределения F_{X_2} в точках x'_2 и x''_2 :

	x'_1	x''_1	
x'_2	0.3	0.4	0.4
x''_2	0.7	1.0	1.0
	0.7	1.0	

Из этого примера становится понятным происхождение названия «маргинальные» для функций распределения F_{X_1} и F_{X_2} . Оно происходит от английского слова «margin» — полоса. Значения этих функций распределения записываются в полосах.

Лемма 1.1. При

$$-\infty < a_1 < b_1 < \infty, \quad -\infty < a_2 < b_2 < \infty$$

вероятность события

$$\{a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2\}$$

равна

$$F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_1, b_2) - F_{X_1, X_2}(b_1, a_2) + F_{X_1, X_2}(a_1, a_2).$$

Доказательство. Рассмотрим события

$$\begin{aligned} A &= \{X_1 \leq b_1, X_2 \leq b_2\}, & B &= \{X_1 \leq a_1, X_2 \leq b_2\}, \\ C &= \{X_1 \leq b_1, X_2 \leq a_2\}, & D &= \{X_1 \leq a_1, X_2 \leq a_2\}. \end{aligned}$$

Тогда

$$\begin{aligned} P(A) &= F_{X_1, X_2}(b_1, b_2), & P(B) &= F_{X_1, X_2}(a_1, b_2), \\ P(C) &= F_{X_1, X_2}(b_1, a_2), & P(D) &= F_{X_1, X_2}(a_1, a_2). \end{aligned}$$

Заметим, что событие $\{a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2\}$ имеет вид $A \setminus (B \cup C)$ и при этом $B \cup C \subseteq A$. С другой стороны, $B \cap C = D$. По теореме сложения вероятностей (см. [ТВиМС-1, разделы 1.2, 1.3])

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

Поэтому

$$\begin{aligned} P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) &= P(A) - P(B \cup C) = \\ &= P(A) - P(B) - P(C) + P(D). \quad \square \end{aligned}$$

Замечание. С использованием теоремы о непрерывности вероятности утверждение леммы 1.1 можно распространить и на случай

$$-\infty \leq a_1 < b_1 \leq \infty, \quad -\infty \leq a_2 < b_2 \leq \infty.$$

В терминах совместных функций распределения может быть дан критерий *независимости случайных величин*. Определение независимости произвольного числа случайных величин дано в [ТВиМС-1, раздел 1.2]. (Для двух случайных величин это определение фактически содержится внутри доказательства теоремы 1.4.)

Теорема 1.4. *Случайные величины X_1 и X_2 независимы тогда и только тогда, когда для любых $x_1 \in R$ и $x_2 \in R$ имеет место соотношение*

$$F_{X_1, X_2}(x_1, x_2) = F_{X_1}(x_1) \cdot F_{X_2}(x_2).$$

Доказательство. Пусть случайные величины X_1 и X_2 независимы. Согласно определению независимости случайных величин это означает, что при любых a_1, b_1 , удовлетворяющих условию:

$$-\infty \leq a_1 < \infty, \quad -\infty < b_1 \leq \infty, \quad a_1 \leq b_1,$$

и при любых a_2, b_2 , удовлетворяющих условию:

$$-\infty \leq a_2 < \infty, \quad -\infty < b_2 \leq \infty, \quad a_2 \leq b_2,$$

независимы события

$$\{a_1 \leq X_1 \leq b_1\} \quad \text{и} \quad \{a_2 \leq X_2 \leq b_2\}.$$

Взяв

$$a_1 = -\infty, \quad b_1 = x_1, \quad a_2 = -\infty, \quad b_2 = x_2,$$

получаем, что независимы события

$$\{X_1 \leq x_1\}, \quad \{X_2 \leq x_2\}.$$

Таким образом,

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \\ &= \mathbf{P}(X_1 \leq x_1) \cdot \mathbf{P}(X_2 \leq x_2) = F_{X_1}(x_1) \cdot F_{X_2}(x_2). \end{aligned}$$

Доказано, что из независимости случайных величин следует, что их совместная функция распределения представима в виде произведения маргинальных функций распределения.

Пусть, наоборот, при любых $x_1 \in R$ и $x_2 \in R$ совместная функция распределения случайных величин X_1 и X_2 представима в виде произведения маргинальных функций распределения. Рассмотрим сначала случай

$$-\infty \leq a_1 < b_1 \leq \infty, \quad -\infty \leq a_2 < b_2 \leq \infty$$

и покажем, что независимы события

$$\{a_1 < X_1 \leq b_1\}, \quad \{a_2 < X_2 \leq b_2\}.$$

По лемме 1.1

$$\begin{aligned} \mathbf{P}(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) &= \\ &= F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_1, b_2) - F_{X_1, X_2}(b_1, a_2) + F_{X_1, X_2}(a_1, a_2) = \\ &= F_{X_1}(b_1)F_{X_2}(b_2) - F_{X_1}(a_1)F_{X_2}(b_2) - F_{X_1}(b_1)F_{X_2}(a_2) + F_{X_1}(a_1)F_{X_2}(a_2) = \\ &= (F_{X_1}(b_1) - F_{X_1}(a_1))(F_{X_2}(b_2) - F_{X_2}(a_2)) = \\ &= \mathbf{P}(a_1 < X_1 \leq b_1) \cdot \mathbf{P}(a_2 < X_2 \leq b_2). \end{aligned}$$

Тем самым доказано, что события

$$\{a_1 < X_1 \leq b_1\}, \quad \{a_2 < X_2 \leq b_2\}$$

независимы.

Теперь рассмотрим случай

$$a_1 \leq b_1, \quad a_2 \leq b_2$$

и покажем, что независимы события

$$\{a_1 \leq X_1 \leq b_1\}, \quad \{a_2 \leq X_2 \leq b_2\}.$$

Заметим, что можно не рассматривать случаи $a_1 = -\infty$ и $a_2 = -\infty$, поскольку событие $\{-\infty \leq X_1 \leq b_1\}$ совпадает с событием $\{-\infty < X_1 \leq b_1\}$, а событие $\{-\infty \leq X_2 \leq b_2\}$ совпадает с событием $\{-\infty < X_2 \leq b_2\}$ (функ-

ции X_1 и X_2 принимают лишь действительные значения и не могут принимать значение $-\infty$).

Для того чтобы показать независимость указанных событий, выберем монотонно стремящуюся к нулю последовательность положительных чисел ε_n . Согласно доказанному выше, при любом n

$$\begin{aligned} \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1, a_2 - \varepsilon_n < X_2 \leq b_2) &= \\ &= \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1) \cdot \mathbf{P}(a_2 - \varepsilon_n < X_2 \leq b_2). \end{aligned}$$

Тогда

$$\bigcap_{n=1}^{\infty} \{a_1 - \varepsilon_n < X_1 \leq b_1, a_2 - \varepsilon_n < X_2 \leq b_2\} = \{a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2\}$$

и по теореме 1.1

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1, a_2 - \varepsilon_n < X_2 \leq b_2) &= \\ &= \mathbf{P}(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2); \end{aligned}$$

также имеем

$$\bigcap_{n=1}^{\infty} \{a_1 - \varepsilon_n < X_1 \leq b_1\} = \{a_1 \leq X_1 \leq b_1\}$$

и по теореме 1.1

$$\lim_{n \rightarrow \infty} \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1) = \mathbf{P}(a_1 \leq X_1 \leq b_1);$$

и также

$$\bigcap_{n=1}^{\infty} \{a_2 - \varepsilon_n < X_2 \leq b_2\} = \{a_2 \leq X_2 \leq b_2\}$$

и по теореме 1.1

$$\lim_{n \rightarrow \infty} \mathbf{P}(a_2 - \varepsilon_n < X_2 \leq b_2) = \mathbf{P}(a_2 \leq X_2 \leq b_2).$$

Но по доказанному выше при любом n

$$\begin{aligned} \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1, a_2 - \varepsilon_n < X_2 \leq b_2) &= \\ &= \mathbf{P}(a_1 - \varepsilon_n < X_1 \leq b_1) \cdot \mathbf{P}(a_2 - \varepsilon_n < X_2 \leq b_2). \end{aligned}$$

Отсюда следует, что

$$\mathbf{P}(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2) = \mathbf{P}(a_1 \leq X_1 \leq b_1) \cdot \mathbf{P}(a_2 \leq X_2 \leq b_2). \quad \square$$

З а м е ч а н и е. Аналогичный результат верен и для произвольного числа случайных величин. Случайные величины X_1, X_2, \dots, X_k независимы тогда и только тогда, когда для любых $x_1 \in R, x_2 \in R, \dots, x_k \in R$ имеет место соотношение

$$F_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_k}(x_k).$$

Доказательство проводится по той же схеме. Иногда это условие, что совместная функция распределения представима в виде произведения маргинальных функций распределения, принимают за определение независимости случайных величин.

Пример 1.2 (ср. пример из [ТВиМС-1, раздел 1.4]). Во время 5-минутного перерыва между занятиями один из студентов, Андрей, разговаривает по телефону, а другой студент, Борис, подошел к преподавателю и попросил еще раз объяснить не понятый им материал. Случайная величина X_1 используется для моделирования продолжительности разговора Андрея по телефону, а случайная величина X_2 — для моделирования продолжительности разговора Бориса с преподавателем. Можно принять, например, следующие значения для вероятностей отдельных событий, связанных со случайными величинами X_1 и X_2 , если бы Андрей и Борис действовали, ничего не зная один про другого:

$$\begin{aligned}P(X_1 \leq 0) &= 0, & P(X_1 \leq 1) &= 0.2, & P(X_1 \leq 2) &= 0.4, \\P(X_1 \leq 3) &= 0.6, & P(X_1 \leq 4) &= 0.8, & P(X_1 \leq 5) &= 1; \\P(X_2 \leq 0) &= 0, & P(X_2 \leq 1) &= 0.5, & P(X_2 \leq 2) &= 0.8, \\P(X_2 \leq 3) &= 0.9, & P(X_2 \leq 4) &= 1, & P(X_2 \leq 5) &= 1.\end{aligned}$$

Требуется построить совместную функцию распределения случайных величин X_1 и X_2 .

Оговоримся сразу, что приведенных данных недостаточно для того, чтобы построить эту совместную функцию распределения. Нужна какая-то информация о том, как эти случайные величины (и процессы, которые они моделируют) связаны.

Если действия Андрея и Бориса никак не связаны между собой, то можно сделать предположение, что случайные величины X_1 и X_2 независимы. Тогда совместная функция распределения этих случайных величин равна произведению маргинальных функций распределения. Например,

$$F_{X_1, X_2}(3, 1) = F_{X_1}(3) \cdot F_{X_2}(1) = 0.6 \cdot 0.5 = 0.3.$$

Но предположим теперь, что Андрею нужно срочно поговорить с Борисом, и как только Андрей видит, что Борис закончил разговор с преподавателем, он также заканчивает свой разговор по телефону. Таким образом, в задаче появляется дополнительное условие

$$X_1 \leq X_2.$$

Как построить совместную функцию распределения случайных величин X_1 и X_2 в этом случае? Из условия задачи ясно, что маргинальная функция распределения случайной величины X_2 должна совпадать с исходной функцией распределения случайной величины X_2 , поскольку действия Андрея не влияют на действия Бориса. Ясно также, что маргинальная функция

распределения случайной величины X_1 не может совпадать с той функцией распределения случайной величины X_1 , которая приведена в начале примера.

Действительно, поскольку $\mathbf{P}(X_2 \leq 4) = 1$ и $X_1 \leq X_2$, получаем

$$F_{X_1, X_2}(4, 4) = \mathbf{P}(X_1 \leq 4, X_2 \leq 4) = 1.$$

Тогда для маргинальной функции распределения случайной величины X_1 должно выполняться соотношение

$$F_{X_1}(4) = F_{X_1, X_2}(4, \infty) \geq F_{X_1, X_2}(4, 4) = 1.$$

Но для функции распределения случайной величины X_1 , приведенной в начале примера, $F_{X_1}(4) = 0.8$.

Для определения значений совместной функции распределения

$$F_{X_1, X_2}(x_1, x_2) = \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2)$$

будем исходить из равенства

$$\mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbf{P}(X_2 \leq x_2) - \mathbf{P}(X_1 > x_1, X_2 \leq x_2).$$

При $x_1 \geq x_2$ очевидно, что $\mathbf{P}(X_1 > x_1, X_2 \leq x_2) = 0$, и поэтому

$$\mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbf{P}(X_2 \leq x_2).$$

Пусть $x_1 < x_2$. Тогда

$$\begin{aligned} \mathbf{P}(X_1 > x_1, X_2 \leq x_2) &= \\ &= \mathbf{P}(X_1 > x_1, X_2 \leq x_1) + \mathbf{P}(X_1 > x_1, x_1 < X_2 \leq x_2) = \\ &= \mathbf{P}(X_1 > x_1, x_1 < X_2 \leq x_2). \end{aligned}$$

Воспользовавшись определением условной вероятности, получаем

$$\begin{aligned} \mathbf{P}(X_1 > x_1, x_1 < X_2 \leq x_2) &= \\ &= \mathbf{P}(X_1 > x_1 \mid x_1 < X_2 \leq x_2) \cdot \mathbf{P}(x_1 < X_2 \leq x_2), \end{aligned}$$

и, таким образом, окончательно,

$$\begin{aligned} \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) &= \\ &= \mathbf{P}(X_2 \leq x_2) - \mathbf{P}(X_1 > x_1 \mid x_1 < X_2 \leq x_2) \cdot \mathbf{P}(x_1 < X_2 \leq x_2). \end{aligned}$$

Поскольку распределение вероятностей случайной величины X_2 известно из условий примера, для определения вероятностей $\mathbf{P}(X_1 \leq x_1, X_2 \leq x_2)$ достаточно знать условные вероятности $\mathbf{P}(X_1 > x_1 \mid x_1 < X_2 \leq x_2)$.

Эти условные вероятности будут разными при разных вариантах поведения Андрея. При первом варианте, пока Борис не закончил разговор с преподавателем, Андрей ведет свой разговор по телефону так, как нужно ему и его собеседнику. Тогда, в соответствии с той функцией распределе-

ния случайной величины X_1 , которая приведена в начале примера, можно считать, что

$$\begin{aligned} \mathbf{P}(X_1 > 0 \mid 0 < X_2 \leq x_2) &= 1.0, & \mathbf{P}(X_1 > 1 \mid 1 < X_2 \leq x_2) &= 0.8, \\ \mathbf{P}(X_1 > 2 \mid 2 < X_2 \leq x_2) &= 0.6, & \mathbf{P}(X_1 > 3 \mid 3 < X_2 \leq x_2) &= 0.4, \\ \mathbf{P}(X_1 > 4 \mid 4 < X_2 \leq x_2) &= 0.2, & \mathbf{P}(X_1 > 5 \mid 5 < X_2 \leq x_2) &= 0.0. \end{aligned}$$

Значения совместной функции распределения $F_{X_1, X_2}(x_1, x_2)$ представлены в следующей таблице. Например,

$$\begin{aligned} \mathbf{P}(X_1 \leq 1, X_2 \leq 3) &= \\ &= \mathbf{P}(X_2 \leq 3) - \mathbf{P}(X_1 > 1 \mid 1 < X_2 \leq 3) \cdot (\mathbf{P}(X_2 \leq 3) - \mathbf{P}(X_2 \leq 1)) = \\ &= 0.9 - 0.8 \cdot (0.9 - 0.5) = 0.58. \end{aligned}$$

В верхней строке показаны значения x_1 . В нижней строке показаны значения маргинальной функции распределения $F_{X_1}(x_1)$. В левом столбце показаны значения x_2 . В правом столбце показаны значения маргинальной функции распределения $F_{X_2}(x_2)$.

$x_2 \backslash x_1$	0	1	2	3	4	5	$F_{X_2}(x_2)$
0	0.	0.	0.	0.	0.	0.	0.
1	0.	0.5	0.5	0.5	0.5	0.5	0.5
2	0.	0.56	0.8	0.8	0.8	0.8	0.8
3	0.	0.58	0.84	0.9	0.9	0.9	0.9
4	0.	0.6	0.88	0.96	1.	1.	1.
5	0.	0.6	0.88	0.96	1.	1.	1.
$F_{X_1}(x_1)$	0.	0.6	0.88	0.96	1.	1.	

При втором варианте Андрей намеренно затягивает разговор по телефону, пока Борис не закончит говорить с преподавателем. Тогда условная вероятность $\mathbf{P}(X_1 > x_1 \mid x_1 < X_2 \leq x_2) = 1$, и значения совместной и маргинальных функций распределения представлены в следующей таблице. Например:

$$\begin{aligned} \mathbf{P}(X_1 \leq 1, X_2 \leq 3) &= \\ &= \mathbf{P}(X_2 \leq 3) - \mathbf{P}(X_1 > 1 \mid 1 < X_2 \leq 3) \cdot (\mathbf{P}(X_2 \leq 3) - \mathbf{P}(X_2 \leq 1)) = \\ &= 0.9 - 1 \cdot (0.9 - 0.5) = 0.5. \end{aligned}$$

В верхней строке показаны значения x_1 . В нижней строке показаны значения маргинальной функции распределения $F_{X_1}(x_1)$. В левом столбце показаны значения x_2 . В правом столбце показаны значения маргинальной функции распределения $F_{X_2}(x_2)$.

$x_2 \backslash x_1$	0	1	2	3	4	5	$F_{X_2}(x_2)$
0	0.	0.	0.	0.	0.	0.	0.
1	0.	0.5	0.5	0.5	0.5	0.5	0.5
2	0.	0.5	0.8	0.8	0.8	0.8	0.8
3	0.	0.5	0.8	0.9	0.9	0.9	0.9
4	0.	0.5	0.8	0.9	1.	1.	1.
5	0.	0.5	0.8	0.9	1.	1.	1.
$F_{X_1}(x_1)$	0.	0.5	0.8	0.9	1.	1.	

При втором варианте время разговора Андрея по телефону совпадает со временем разговора Бориса с преподавателем, поэтому функции распределения случайных величин X_1 и X_2 оказываются одинаковыми.

Используя найденные совместные функции распределения и маргинальные функции распределения и применив теорему 1.4, нетрудно увидеть, что и при первом, и при втором варианте случайные величины X_1 и X_2 не являются независимыми.

1.2. СОВМЕСТНЫЕ ДИСКРЕТНЫЕ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

Как и в предыдущем разделе, ограничимся рассмотрением совместного распределения двух случайных величин X_1 и X_2 . Но это делается исключительно ради краткости и наглядности. По точно той же схеме можно рассматривать совместное распределение случайных величин X_1, X_2, \dots, X_k .

Пусть при любых $x_1 \in R, x_2 \in R$ совместная функция распределения случайных величин X_1 и X_2 имеет вид

$$F_{X_1, X_2}(x_1, x_2) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f(q_1, q_2).$$

Здесь Q_1 и Q_2 — конечные или счетные множества, $Q_1 \subseteq R, Q_2 \subseteq R$.
Функция

$$f: R^2 \rightarrow R$$

принимает лишь неотрицательные значения и обладает тем свойством, что

$$f(x_1, x_2) = 0, \text{ если } x_1 \notin Q_1 \text{ или } x_2 \notin Q_2$$

(т.е. положительные значения функция f может принимать лишь в точках (x_1, x_2) , таких что $x_1 \in Q_1, x_2 \in Q_2$).

Тогда говорят, что случайные величины X_1 и X_2 имеют *совместное дискретное распределение вероятностей*. Функция f называется в этом

случае *совместной дискретной функции плотности*. Согласно теореме 1.3 должно выполняться соотношение

$$\sum_{q_1 \in Q_1} \sum_{q_2 \in Q_2} f(q_1, q_2) = 1.$$

Заметим, что если пространство элементарных событий Ω , на котором определены функции X_1 и X_2 , является конечным множеством (этот случай подробно рассматривается в [ТВиМС-1]), то совместное распределение случайных величин X_1 и X_2 обязано быть совместным дискретным распределением. В этом случае множества Q_1 и Q_2 — конечные множества.

Напомним, что множество M называется *счетным*, если можно установить *взаимно-однозначное соответствие* между элементами множества M и натуральными числами. Иначе говоря, существует функция $g: M \rightarrow N$, где N — множество натуральных чисел, такая что, во-первых, различным элементам m_1 и m_2 множества M соответствуют различные натуральные числа $n_1 = g(m_1)$ и $n_2 = g(m_2)$ и, во-вторых, для любого натурального числа n найдется элемент m множества M такой, что $g(m) = n$. Например, множество натуральных чисел N счетно.

Счетные множества Q_1 и Q_2 могут иметь достаточно сложную структуру, например, содержать сходящиеся последовательности. Но во многих прикладных задачах, если рассматриваются счетные множества Q_1 и Q_2 , приходится сталкиваться с ситуацией, когда как множество Q_1 , так и множество Q_2 состоят из точек, которые достаточно далеко отстоят друг от друга. Математически это условие можно выразить, например, так. Действительную прямую R можно разбить на полуинтервалы одинаковой длины так, что каждый из этих полуинтервалов содержит не больше одной точки из множества Q_1 и не больше одной точки из множества Q_2 . Чтобы упростить изложение, будем считать данное условие выполненным. Отметим однако, что явно это условие используется только при доказательстве теоремы 1.5.

Теорема 1.5. *При любых $x_1 \in R$ и $x_2 \in R$*

$$\mathbf{P}(X_1 = x_1, X_2 = x_2) = f(x_1, x_2).$$

Доказательство. Возьмем монотонно стремящуюся к нулю последовательность положительных чисел ε_n и рассмотрим события

$$A_n = \{x_1 - \varepsilon_n < X_1 \leq x_1, x_2 - \varepsilon_n < X_2 \leq x_2\}.$$

Тогда

$$\bigcap_{n=1}^{\infty} A_n = \{X_1 = x_1, X_2 = x_2\}$$

и по теореме о непрерывности вероятности

$$\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(X_1 = x_1, X_2 = x_2).$$

С другой стороны, по лемме 1.1 с учетом вида совместной функции распределения случайных величин X_1 и X_2

$$P(A_n) = \sum_{\substack{q_1 \in Q_1 \\ x_1 - \varepsilon_n < q_1 \leq x_1}} \sum_{\substack{q_2 \in Q_2 \\ x_2 - \varepsilon_n < q_2 \leq x_2}} f(q_1, q_2).$$

С учетом условий, которым удовлетворяет функция f , а также условия, наложенного на множества Q_1 и Q_2 , при достаточно больших n

$$P(A_n) = f(x_1, x_2).$$

□

Функции

$$f_{X_1}(x_1) = \sum_{q_2 \in Q_2} f(x_1, q_2), \quad f_{X_2}(x_2) = \sum_{q_1 \in Q_1} f(q_1, x_2)$$

называются *маргинальными дискретными функциями плотности*.

В случае совместного дискретного распределения вероятностей *маргинальные функции распределения*, определенные в разделе 1.1, естественным образом выражаются через маргинальные дискретные функции плотности:

$$F_{X_1}(x_1) = F_{X_1, X_2}(x_1, \infty) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} \sum_{q_2 \in Q_2} f(q_1, q_2) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} f_{X_1}(q_1),$$

$$F_{X_2}(x_2) = F_{X_1, X_2}(\infty, x_2) = \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} \sum_{q_1 \in Q_1} f(q_1, q_2) = \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f_{X_2}(q_2).$$

Теорема 1.6. *Случайные величины X_1 и X_2 независимы тогда и только тогда, когда для любых $x_1 \in R$ и $x_2 \in R$ имеет место соотношение*

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2),$$

т.е. совместная дискретная функция плотности равна произведению маргинальных дискретных функций плотности.

Доказательство. Пусть совместная дискретная функция плотности равна произведению маргинальных дискретных функций плотности. Тогда

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f(q_1, q_2) = \\ &= \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} f_{X_1}(q_1) \cdot \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f_{X_2}(q_2) = F_{X_1}(x_1) \cdot F_{X_2}(x_2). \end{aligned}$$

Таким образом, совместная функция распределения равна произведению маргинальных функций распределения. По теореме 1.4 отсюда следует, что случайные величины X_1 и X_2 независимы.

Пусть случайные величины X_1 и X_2 независимы. Тогда по теореме 1.4 совместная функция распределения этих случайных величин равна произведению их маргинальных функций распределения, т.е.

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= F_{X_1}(x_1) \cdot F_{X_2}(x_2) = \\ &= \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} f_{X_1}(q_1) \cdot \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f_{X_2}(q_2) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f_{X_1}(q_1) \cdot f_{X_2}(q_2). \end{aligned}$$

Но по теореме 1.5 совместная дискретная функция плотности определяется вероятностями событий $\{X_1 = q_1, X_2 = q_2\}$. То есть две разные функции быть совместными дискретными функциями плотности для одной и той же пары случайных величин не могут. Поэтому

$$f(q_1, q_2) = f_{X_1}(q_1) \cdot f_{X_2}(q_2). \quad \square$$

З а м е ч а н и е. Аналогичный результат верен и для произвольного числа случайных величин. Случайные величины X_1, X_2, \dots, X_k независимы тогда и только тогда, когда для любых $x_1 \in R, x_2 \in R, \dots, x_k \in R$ имеет место соотношение

$$f(x_1, x_2, \dots, x_k) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \dots \cdot f_{X_k}(x_k).$$

Доказательство проводится по той же схеме.

Пример 1.3. В игре, связанной с бросанием кости, участвуют три человека. Первый участник получает два рубля, если выпало 5 или 6, и платит один рубль, если выпало 1, 2, 3 или 4. Второй участник получает три рубля, если выпало 1, 2 или 3, и платит три рубля, если выпало 4, 5 или 6. Третий участник расплачивается с первым и со вторым после того, как результат становится известным. При этом первый и второй участники могут не знать об условиях, на которых играет другой из них.

Пусть случайная величина X_1 моделирует выигрыш первого участника, случайная величина X_2 моделирует выигрыш второго участника.

Требуется построить совместную дискретную функцию плотности этих случайных величин, маргинальную дискретную функцию плотности каждой из случайных величин X_1 и X_2 и определить, независимы ли эти случайные величины.

В этом примере множество Q_1 состоит из двух точек: $Q_1 = \{-1, 2\}$, и множество Q_2 также состоит из двух точек: $Q_2 = \{-3, 3\}$.

Введем еще одну случайную величину. Пусть U — количество выпавших очков. Тогда

$$\mathbf{P}(X_1 = -1, X_2 = -3) = \mathbf{P}(U = 4) = \frac{1}{6},$$

$$\mathbf{P}(X_1 = -1, X_2 = 3) = \mathbf{P}(U = 1 \text{ или } U = 2 \text{ или } U = 3) = \frac{1}{2},$$

$$\mathbf{P}(X_1 = 2, X_2 = -3) = \mathbf{P}(U = 5 \text{ или } U = 6) = \frac{1}{3},$$

$$\mathbf{P}(X_1 = 2, X_2 = 3) = 0.$$

Тогда (см. теорему 1.5) значения совместной дискретной функции плотности $f(x_1, x_2)$ представляются в виде следующей таблицы. В нижней строке показаны значения маргинальной дискретной функции плотности $f_{X_1}(x_1)$. В правом столбце показаны значения маргинальной дискретной функции плотности $f_{X_2}(x_2)$.

$x_2 \backslash x_1$	-1	2	$f_{X_2}(x_2)$
-3	1/6	1/3	1/2
3	1/2	0	1/2
$f_{X_1}(x_1)$	2/3	1/3	

Поскольку, например,

$$\begin{aligned} f_{X_1}(2) \cdot f_{X_2}(3) &= \mathbf{P}(X_1 = 2) \cdot \mathbf{P}(X_2 = 3) = \frac{1}{3} \cdot \frac{1}{2} \neq \\ &\neq 0 = \mathbf{P}(X_1 = 2, X_2 = 3) = f(2, 3), \end{aligned}$$

то согласно теореме 1.6 случайные величины X_1 и X_2 не независимы.

Пример 1.4. Примером совместного дискретного распределения k случайных величин является *мультиномиальное распределение*. Частный случай мультиномиального распределения — это *биномиальное распределение*, рассмотренное в книге [ТВиМС-1, раздел 2.3]. Опишем мультиномиальное распределение в той же последовательности, в какой в книге [ТВиМС-1] описано биномиальное распределение.

Пусть проводится n независимых испытаний, результатом каждого испытания может быть один из k исходов, где k — натуральное число, $k \geq 2$. (Биномиальному распределению отвечает $k = 2$.) Пусть p_j , $j = 1, 2, \dots, k$, — это вероятность того, что результатом испытания стал исход j . Считается, что данная вероятность не меняется от испытания к испытанию. Очевидно, что

$$p_1 + p_2 + \dots + p_k = 1.$$

Пусть значение случайной величины X_j — это число исходов j при n испытаниях. Требуется найти совместное распределение случайных величин X_1, X_2, \dots, X_k , т.е. все возможные вероятности

$$\mathbf{P}(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k).$$

Очевидно, что эти вероятности ненулевые лишь при целых $x_j \geq 0$, где $j = 1, 2, \dots, k$, и таких, что

$$x_1 + x_2 + \dots + x_k = n.$$

Рассмотрим пространство элементарных событий Ω , где каждое элементарное событие — это строка $e_1 e_2 \dots e_n$, а каждое из e_i может быть любым числом от 1 до k . Например, при $k = 3, n = 2$

$$\Omega = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}.$$

Множество Ω конечно и содержит k^n элементов. Чтобы определить, какую вероятность следует приписать элементарному событию $e_1 e_2 \dots e_n$, рассмотрим события A_{ij} ; событие A_{ij} заключается в том, что при i -м испытании произошел исход j . Очевидно, что при любом i

$$\mathbf{P}(A_{ij}) = p_j, \quad j = 1, 2, \dots, k.$$

Условие независимости испытаний, данное в условиях примера на повседневном языке, на математический язык переведем так, что любые события

$$A_{1e_1}, A_{2e_2}, \dots, A_{ne_n}$$

независимы. (Определение независимости произвольного числа событий дается в [ТВиМС-1, раздел 1.2].) Тогда вероятность элементарного события $e_1 e_2 \dots e_n$, у которого x_1 символов e_i равны 1, x_2 символов e_i равны 2, ..., x_k символов e_i равны k , рассчитывается по формуле

$$\mathbf{P}(\{e_1 e_2 \dots e_n\}) = \mathbf{P}\left(\bigcap_{i=1}^n A_{ie_i}\right) = \prod_{i=1}^n \mathbf{P}(A_{ie_i}) = p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \cdot p_k^{x_k}.$$

Теперь, для того чтобы ответить на вопрос, какова вероятность события

$$\{X_1 = x_1, X_2 = x_2, \dots, X_k = x_k\},$$

следует определить, сколько существует элементарных событий $e_1 e_2 \dots e_n$, у которых x_1 символов e_i равны 1, x_2 символов e_i равны 2, ..., x_k символов e_i равны k .

Существует $C_n^{x_1}$ способов выбрать x_1 испытаний из n , исходом которых является 1 (см. [ТВиМС-1, раздел 2.1]). Напомним, что при целом неотрицательном m и целом неотрицательном $r, 0 \leq r \leq m$,

$$C_m^r = \frac{m!}{(m-r)! \cdot r!}.$$

Числа C_m^r называются *биномиальными коэффициентами*. Существует $C_{n-x_1}^{x_2}$ способов выбрать x_2 испытаний из $n - x_1$, исходом которых является 2. Наконец, существует $C_{n-x_1-\dots-x_{k-2}}^{x_{k-1}}$ способов выбрать x_{k-1} испытаний из $n - x_1 - \dots - x_{k-2}$, исходом которых является $k - 1$. Осталось заметить, что

$$C_n^{x_1} \cdot C_{n-x_1}^{x_2} \cdot \dots \cdot C_{n-x_1-\dots-x_{k-2}}^{x_{k-1}} = \frac{n!}{x_1! \cdot x_2! \cdot \dots \cdot x_k!}.$$

Таким образом,

$$\begin{aligned} \mathbf{P}(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) &= \\ &= \frac{n!}{x_1! \cdot x_2! \cdot \dots \cdot x_k!} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \cdot p_k^{x_k}. \end{aligned}$$

Предположим, что $f_{X_1}(x_1) > 0$ при любом $x_1 \in Q_1$ и $f_{X_2}(x_2) > 0$ при любом $x_2 \in Q_2$.

При $x_1 \in R$, $x_2 \in Q_2$ положим

$$f(x_1 | X_2 = x_2) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)};$$

при $x_2 \in R$, $x_1 \in Q_1$ положим

$$f(x_2 | X_1 = x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}.$$

Функции $f(x_1 | X_2 = x_2)$ и $f(x_2 | X_1 = x_1)$ называются *условными дискретными функциями плотности*.

Из теоремы 1.5 и определения маргинальных дискретных функций плотности f_{X_1} и f_{X_2} следует, что

$$f(x_1 | X_2 = x_2) = \frac{\mathbf{P}(X_1 = x_1, X_2 = x_2)}{\mathbf{P}(X_2 = x_2)} = \mathbf{P}(X_1 = x_1 | X_2 = x_2)$$

— условная вероятность события $X_1 = x_1$ при условии события $X_2 = x_2$;

$$f(x_2 | X_1 = x_1) = \frac{\mathbf{P}(X_1 = x_1, X_2 = x_2)}{\mathbf{P}(X_1 = x_1)} = \mathbf{P}(X_2 = x_2 | X_1 = x_1)$$

— условная вероятность события $X_2 = x_2$ при условии события $X_1 = x_1$.

Из теоремы 1.6 следует, что если случайные величины X_1 и X_2 независимы, то при любом $x_2 \in Q_2$

$$f(x_1 | X_2 = x_2) = f_{X_1}(x_1)$$

и при любом $x_1 \in Q_1$

$$f(x_2 | X_1 = x_1) = f_{X_2}(x_2).$$

Функции

$$F(x_1 | X_2 = x_2) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} f(q_1 | X_2 = x_2)$$

и

$$F(x_2 | X_1 = x_1) = \sum_{\substack{q_2 \in Q_2 \\ q_2 \leq x_2}} f(q_2 | X_1 = x_1)$$

называются *условными функциями распределения*.

Пример 1.5. Допустим, что для большого числа корпоративных облигаций, имеющих кредитные рейтинги A , B и C , собраны данные по их доходности к погашению. Случайная величина X_1 показывает доходность к погашению облигации и может принимать значения 5, 6, 7, 8%. Случайная величина X_2 показывает кредитный рейтинг облигации и равна 0 для облигаций с кредитным рейтингом A , равна 1 для облигаций с кредитным рейтингом B , равна 2 для облигаций с кредитным рейтингом C . В таблице для каждой пары возможных значений случайных величин X_1 и X_2 показана доля облигаций, имеющих указанную доходность к погашению и указанный кредитный рейтинг. На основании закона больших чисел будем считать соответствующую долю вероятностью для облигации попасть в данную группу по доходности к погашению и в данную группу по кредитному рейтингу.

$x_2 \backslash x_1$	5%	6%	7%	8%
0	0.08	0.10	0.06	0.00
1	0.11	0.17	0.16	0.05
2	0.02	0.09	0.10	0.06

Требуется найти маргинальную дискретную функцию плотности $f_{X_1}(x_1)$, маргинальную дискретную функцию плотности $f_{X_2}(x_2)$, условную дискретную функцию плотности $f(x_1 | X_2 = x_2)$, условную дискретную функцию плотности $f(x_2 | X_1 = x_1)$.

Ответ получается непосредственно из определений маргинальных дискретных функций плотности и условных дискретных функций плотности и приведен в следующих таблицах.

	$x_1 = 5\%$	$x_1 = 6\%$	$x_1 = 7\%$	$x_1 = 8\%$
$f_{X_1}(x_1)$	0.21	0.36	0.32	0.11
$f(x_1 X_2 = 0)$	8/24	10/24	6/24	0
$f(x_1 X_2 = 1)$	11/49	17/49	16/49	5/49
$f(x_1 X_2 = 2)$	2/27	9/27	10/27	6/27

	$x_2 = 0$	$x_2 = 1$	$x_2 = 2$
$f_{X_2}(x_2)$	0.24	0.49	0.27
$f(x_2 X_1 = 5)$	8/21	11/21	2/21
$f(x_2 X_1 = 6)$	10/36	17/36	9/36
$f(x_2 X_1 = 7)$	6/32	16/32	10/32
$f(x_2 X_1 = 8)$	0	5/11	6/11

Замечание. Для полноты изложения упомянем про случай $k = 1$. Если функция распределения случайной величины X_1 имеет вид

$$F_{X_1}(x_1) = \sum_{\substack{q_1 \in Q_1 \\ q_1 \leq x_1}} f(q_1),$$

то распределение вероятностей случайной величины X_1 называют *дискретным распределением*, а функцию f — *дискретной функцией плотности*. Говорить о маргинальных функциях распределения или об условных функциях распределения при $k = 1$ не имеет смысла.

1.3. СОВМЕСТНЫЕ НЕПРЕРЫВНЫЕ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

Вновь мы ограничимся рассмотрением совместного распределения двух случайных величин X_1 и X_2 . Но аналогичные результаты имеют место и для произвольного числа случайных величин X_1, X_2, \dots, X_k .

Если при любых $x_1 \in R$, $x_2 \in R$ совместная функция распределения случайных величин X_1 и X_2 имеет вид

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(q_1, q_2) dq_1 dq_2,$$

где интегрируемая функция

$$f: R^2 \rightarrow R$$

принимает лишь неотрицательные значения, то говорят, что случайные величины X_1 и X_2 имеют *совместное непрерывное распределение вероятностей*. Функция f называется в этом случае *совместной функцией плотности*. Согласно теореме 1.3 должно выполняться соотношение

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(q_1, q_2) dq_1 dq_2 = 1.$$

Из определения совместной функции распределения следует, что в тех точках, где совместная функция распределения дифференцируема, имеет место равенство

$$\frac{\partial^2 F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2} = f(x_1, x_2).$$

Теорема 1.7. При любых $-\infty \leq a_1 < b_1 \leq \infty$, $-\infty \leq a_2 < b_2 \leq \infty$

$$\mathbf{P}(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(q_1, q_2) dq_1 dq_2.$$

Доказательство. Пусть сначала

$$-\infty < a_1 < b_1 < \infty, \quad -\infty < a_2 < b_2 < \infty.$$

Рассмотрим следующие подмножества множества R^2 :

$$\begin{aligned} S_1 &= \{x_1 \leq a_1, x_2 \leq a_2\}, & S_2 &= \{x_1 \leq a_1, a_2 < x_2 \leq b_2\}, \\ S_3 &= \{a_1 < x_1 \leq b_1, x_2 \leq a_2\}, & S_4 &= \{a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2\} \end{aligned}$$

и соответствующие интегралы

$$I_j = \iint_{S_j} f(q_1, q_2) dq_1 dq_2, \quad j = 1, 2, 3, 4.$$

По определению совместной функции распределения

$$\begin{aligned} F_{X_1, X_2}(a_1, a_2) &= I_1, & F_{X_1, X_2}(a_1, b_2) &= I_1 + I_2, \\ F_{X_1, X_2}(a_2, b_1) &= I_1 + I_3, & F_{X_1, X_2}(b_1, b_2) &= I_1 + I_2 + I_3 + I_4. \end{aligned}$$

По лемме 1.1

$$\begin{aligned} \mathbf{P}(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) &= \\ &= F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_2, b_1) - F_{X_1, X_2}(a_1, b_2) + F_{X_1, X_2}(a_1, a_2) = \\ &= (I_1 + I_2 + I_3 + I_4) - (I_1 + I_3) - (I_1 + I_2) + I_1 = I_4 = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(q_1, q_2) dq_1 dq_2. \end{aligned}$$

Для случая

$$-\infty < a_1 < b_1 < \infty, \quad -\infty < a_2 < b_2 < \infty$$

теорема доказана.

Доказательство для случаев $a_1 = -\infty$, $a_2 = \infty$, $b_1 = -\infty$, $b_2 = \infty$ проводится с использованием теоремы 1.1. \square

Пусть S — некоторое множество, принадлежащее R^2 , например, многоугольник или круг. Через $\{(X_1, X_2) \in S\}$ обозначим событие, состоящее из тех и только тех $\omega \in \Omega$, для которых

$$(X_1(\omega), X_2(\omega)) \in S.$$

Тогда

$$\mathbf{P}((X_1, X_2) \in S) = \iint_S f(q_1, q_2) dq_1 dq_2. \quad (1.1)$$

Этот результат совпадает с утверждением теоремы 1.7, когда S является полуоткрытым прямоугольником вида

$$\{(x_1, x_2): a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2\}.$$

Вообще же данный результат представляет собой серьезное обобщение теоремы 1.7.

Мы не будем давать доказательство формулы (1.1). Скажем только, что оно основано на возможности приближения множества S множествами, состоящими из непересекающихся прямоугольников, с последующим использованием теорем 1.7 и 1.1.

В частности, как это следует из формулы (1.1), если S — множество, имеющее нулевую площадь, то вероятность события $\{(X_1, X_2) \in S\}$ равна нулю. Это принципиальное отличие совместного непрерывного распределения от совместного дискретного распределения.

Известно, что для интегрируемой функции f двойной интеграл выражается через интегралы по отдельным переменным:

$$\int_{a_2}^{b_2} \int_{a_1}^{b_1} f(q_1, q_2) dq_1 dq_2 = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(q_1, q_2) dq_1 \right) dq_2 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(q_1, q_2) dq_2 \right) dq_1.$$

Функции

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, q_2) dq_2, \quad f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(q_1, x_2) dq_1$$

называются *маргинальными функциями плотности*.

В случае совместного непрерывного распределения вероятностей *маргинальные функции распределения*, определенные в разделе 1.1, естественным образом выражаются через маргинальные функции плотности:

$$F_{X_1}(x_1) = F_{X_1, X_2}(x_1, \infty) = \int_{-\infty}^{x_1} \left(\int_{-\infty}^{\infty} f(q_1, q_2) dq_2 \right) dq_1 = \int_{-\infty}^{x_1} f_{X_1}(q_1) dq_1$$

и

$$F_{X_2}(x_2) = F_{X_1, X_2}(\infty, x_2) = \int_{-\infty}^{x_2} \left(\int_{-\infty}^{\infty} f(q_1, q_2) dq_1 \right) dq_2 = \int_{-\infty}^{x_2} f_{X_2}(q_2) dq_2.$$

Это означает, что маргинальные функции плотности f_{X_1} и f_{X_2} являются функциями плотности случайных величин X_1 и X_2 соответственно.

Пример 1.6. Рассмотрим функцию

$$f: R^2 \rightarrow R$$

следующего вида:

$$f(x_1, x_2) = \begin{cases} 6(1 - x_1 - x_2) & \text{при } 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_1 + x_2 \leq 1; \\ 0 & \text{при остальных } x_1, x_2. \end{cases}$$

Требуется проверить, является ли функция f совместной функцией плотности. В случае положительного ответа найти маргинальные функции плотности, отвечающие данной совместной функции плотности. Для случайных величин X_1 и X_2 , совместной функцией плотности которых является функция f , найти вероятность события $\{X_1 + X_2 \leq 0.8\}$.

Нетрудно убедиться, что функция f принимает неотрицательные значения при любых $x_1 \in R, x_2 \in R$ и является интегрируемой. Поэтому для того чтобы функция f была совместной функцией плотности, достаточно, чтобы выполнялось условие

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1.$$

Поскольку функция f равна нулю всюду вне единичного квадрата $\{0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$, а внутри единичного квадрата равна нулю при $x_1 > 1 - x_2$, то

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 &= \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 = \\ &= \int_0^1 \int_0^{1-x_2} f(x_1, x_2) dx_1 dx_2 = 6 \cdot \int_0^1 \int_0^{1-x_2} (1 - x_1 - x_2) dx_1 dx_2. \end{aligned}$$

Вычислим внутренний интеграл:

$$\begin{aligned} \int_0^{1-x_2} (1 - x_1 - x_2) dx_1 &= (1 - x_2) \int_0^{1-x_2} dx_1 - \int_0^{1-x_2} x_1 dx_1 = \\ &= (1 - x_2)^2 - \frac{x_1^2}{2} \Big|_0^{1-x_2} = (1 - x_2)^2 - \frac{(1 - x_2)^2}{2} = \frac{(1 - x_2)^2}{2}. \end{aligned}$$

Поэтому

$$6 \int_0^1 \int_0^{1-x_2} (1 - x_1 - x_2) dx_1 dx_2 = 6 \int_0^1 \frac{(1 - x_2)^2}{2} dx_2 = 3 \int_0^1 u^2 du = 3 \frac{u^3}{3} \Big|_0^1 = 1.$$

Тем самым установлено, что функция f является совместной функцией плотности.

Здесь необходимо обратить внимание на одну тонкость. В основном тексте изложение начинается с того, что рассматриваются случайные величины

$$X_1: \Omega \rightarrow R, \quad X_2: \Omega \rightarrow R,$$

и при условии, что их совместная функция распределения выражается в виде двойного интеграла от некоторой функции f , эта функция f называется совместной функцией плотности случайных величин X_1 и X_2 .

В примере же изложение начинается с некоторой функции f , и до сих пор единственное, что установлено, это что функция f обладает теми же свойствами, какими обладают совместные функции плотности случайных величин. Следует ли отсюда, что существуют случайные величины X_1 и X_2 такие, что их совместная функция плотности совпадает с выбранной функцией f ? Ответ на этот вопрос утвердительный. Оказывается, что отталкиваясь от функции f , можно построить пространство элементарных событий Ω , построить числовые функции

$$X_1: \Omega \rightarrow R, \quad X_2: \Omega \rightarrow R,$$

определить вероятности для событий $A \subseteq \Omega$ и сделать все это таким способом, что данная функция f будет совместной функцией плотности случайных величин X_1 и X_2 . Описание этой конструкции нами не приводится. Отметим только, что аналогичный вопрос возник в книге [ТВиМС-1, раздел 3.1] после определения функции стандартного нормального распределения, а существуют ли случайные величины с такой функцией распределения? Утвердительный ответ в [ТВиМС-1] дается рисунком 3.3.

При $0 \leq x_1 \leq 1$ маргинальная функция плотности

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f(x_1, q_2) dq_2 = \int_0^{1-x_1} f(x_1, q_2) dq_2 = \\ &= 6 \int_0^{1-x_1} (1 - x_1 - q_2) dq_2 = 3(1 - x_1)^2. \end{aligned}$$

Очевидно, что $f_{X_1}(x_1) = 0$ при $x_1 < 0$ и при $x_1 > 1$, поскольку в этих случаях подынтегральная функция равна нулю.

При $0 \leq x_2 \leq 1$ маргинальная функция плотности

$$\begin{aligned} f_{X_2}(x_2) &= \int_{-\infty}^{\infty} f(q_1, x_2) dq_1 = \int_0^{1-x_2} f(q_1, x_2) dq_1 = \\ &= 6 \int_0^{1-x_2} (1 - q_1 - x_2) dq_1 = 3(1 - x_2)^2. \end{aligned}$$

$f_{X_2}(x_2) = 0$ при $x_2 < 0$ и при $x_2 > 1$, поскольку в этих случаях подынтегральная функция равна нулю.

Вероятность события $\{X_1 + X_2 \leq 0.8\}$ совпадает с вероятностью события $\{X_1 \geq 0, X_2 \geq 0, X_1 + X_2 \leq 0.8\}$, поскольку вероятности событий $\{X_1 < 0\}$ и $\{X_2 < 0\}$ равны нулю. Иначе событие $\{X_1 \geq 0, X_2 \geq 0, X_1 + X_2 \leq 0.8\}$ можно записать, как

$$\{(X_1, X_2) \in S\},$$

где треугольник $S \subseteq R^2$ имеет вид

$$\{x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 0.8\},$$

Согласно формуле (1.1) вероятность события $\{(X_1, X_2) \in S\}$ равна

$$\begin{aligned} \iint_S f(q_1, q_2) dq_1 dq_2 &= \int_0^{0.8} \int_0^{0.8-q_2} f(q_1, q_2) dq_1 dq_2 = \\ &= 6 \int_0^{0.8} \int_0^{0.8-q_2} (1 - q_1 - q_2) dq_1 dq_2 = 0.896. \end{aligned}$$

Вычисление последнего двойного интеграла совершенно аналогично проведенному выше вычислению двойного интеграла

$$\int_0^1 \int_0^{1-x_2} (1 - x_1 - x_2) dx_1 dx_2.$$

Теорема 1.8. Пусть совместная функция распределения случайных величин X_1 и X_2 дифференцируема во всех точках множества R^2 . Тогда случайные величины X_1 и X_2 независимы в том и только том случае, когда для любых $x_1 \in R$ и $x_2 \in R$ имеет место соотношение

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2),$$

т.е. совместная функция плотности равна произведению маргинальных функций плотности.

Доказательство. Пусть совместная функция плотности равна произведению маргинальных функций плотности. Тогда

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(q_1, q_2) dq_1 dq_2 = \\ &= \int_{-\infty}^{x_1} f_{X_1}(q_1) dq_1 \cdot \int_{-\infty}^{x_2} f_{X_2}(q_2) dq_2 = F_{X_1}(x_1) \cdot F_{X_2}(x_2). \end{aligned}$$

Таким образом, совместная функция распределения равна произведению маргинальных функций распределения. По теореме 1.4 отсюда следует, что случайные величины X_1 и X_2 независимы. Отметим, что при доказательстве теоремы в эту сторону условие дифференцируемости совместной функции распределения не используется. Пусть случайные величины X_1 и X_2 независимы. Тогда по теореме 1.4 совместная функция распределения этих случайных величин равна произведению их маргинальных функций распределения, т.е. для любых $x_1 \in R$ и $x_2 \in R$

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= F_{X_1}(x_1) \cdot F_{X_2}(x_2) = \\ &= \int_{-\infty}^{x_1} f_{X_1}(q_1) dq_1 \cdot \int_{-\infty}^{x_2} f_{X_2}(q_2) dq_2 = \\ &= \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{X_1}(q_1) \cdot f_{X_2}(q_2) dq_1 dq_2. \end{aligned}$$

Это означает, что

$$\frac{\partial^2 F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2} = f_{X_1}(x_1) \cdot f_{X_2}(x_2).$$

С другой стороны,

$$\frac{\partial^2 F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2} = f(x_1, x_2),$$

где f — совместная функция плотности случайных величин X_1 и X_2 . Поэтому

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2). \quad \square$$

Замечание 1. Аналогичный результат верен не только для двух, но и для произвольного числа случайных величин X_1, X_2, \dots, X_k . Доказательство проводится по той же схеме.

Замечание 2. Условие теоремы 1.8, что совместная функция распределения случайных величин X_1 и X_2 должна быть дифференцируемой во всех точках множества R^2 , можно несколько ослабить, допустив, что

существуют некоторые подмножества множества R^2 нулевой площади, где эта совместная функция распределения не дифференцируема. Однако мы не будем давать точных формулировок.

Пример 1.7. Служба охраны окружающей среды проводит замеры загрязнения воздуха тремя цехами завода. Пусть X_1 — вес примесей в единице объема воздуха, выбрасываемого первым цехом, X_2 — вес примесей в единице объема воздуха, выбрасываемого вторым цехом, X_3 — вес примесей в единице объема воздуха, выбрасываемого третьим цехом. Служба охраны окружающей среды установила, что совместная функция плотности случайных величин X_1, X_2, X_3 имеет вид

$$f(x_1, x_2, x_3) = \begin{cases} \frac{3}{2} & \text{при } (x_1, x_2, x_3) \in S, \\ 0 & \text{при } (x_1, x_2, x_3) \notin S, \end{cases}$$

где $S \subseteq R^3$ — призма с треугольным основанием

$$S = \left\{ (x_1, x_2, x_3) : x_1 \geq 0, x_2 \geq 0, 2x_1 + x_2 \leq 2, \frac{1}{3} \leq x_3 \leq 1 \right\}.$$

Требуется найти вероятность события

$$\{X_1 \leq 0.5, X_2 \leq 0.5, X_3 \leq 0.5\}.$$

Найти маргинальные функции плотности. Определить, являются ли случайные величины X_1, X_2, X_3 независимыми.

Вероятность события $\{X_1 \leq 0.5, X_2 \leq 0.5, X_3 \leq 0.5\}$ обозначим через p . Тогда

$$p = \mathbf{P}((X_1, X_2, X_3) \in T),$$

где $T \subseteq R^3$ куб

$$T = \{(x_1, x_2, x_3) : 0 \leq x_1 \leq 0.5, 0 \leq x_2 \leq 0.5, 0 \leq x_3 \leq 0.5\}.$$

Поэтому в соответствии с формулой (1.1)

$$p = \int \int \int_T f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \frac{3}{2} \int \int \int_{S \cap T} dx_1 dx_2 dx_3.$$

Нетрудно увидеть, что $S \cap T$ — это прямоугольный параллелепипед

$$S \cap T = \{(x_1, x_2, x_3) : 0 \leq x_1 \leq 0.5, 0 \leq x_2 \leq 0.5, \frac{1}{3} \leq x_3 \leq 0.5\}.$$

Тогда

$$p = \frac{3}{2} \int_0^{1/2} dx_1 \int_0^{1/2} dx_2 \int_{1/3}^{1/2} dx_3 = \frac{3}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{16}.$$

Маргинальные функции плотности имеют вид

$$f_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{\infty} f(x_1, x_2, q_3) dq_3 = \frac{3}{2} \int_{1/3}^1 dq_3 = 1$$

при $x_1 \geq 0$, $x_2 \geq 0$, $2x_1 + x_2 \leq 2$;

$$f_{X_1, X_3}(x_1, x_3) = \int_{-\infty}^{\infty} f(x_1, q_2, x_3) dq_2 = \frac{3}{2} \int_0^{2-2x_1} dq_2 = 3(1-x_1)$$

при $0 \leq x_1 \leq 1$, $\frac{1}{3} \leq x_3 \leq 1$;

$$f_{X_2, X_3}(x_2, x_3) = \int_{-\infty}^{\infty} f(q_1, x_2, x_3) dq_1 = \frac{3}{2} \int_0^{1-0.5x_2} dq_1 = \frac{3}{2} \left(1 - \frac{1}{2}x_2\right)$$

при $0 \leq x_2 \leq 2$, $\frac{1}{3} \leq x_3 \leq 1$;

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, q_2, q_3) dq_2 dq_3 = \int_{1/3}^1 3(1-x_1) dq_3 = 2(1-x_1)$$

при $0 \leq x_1 \leq 1$;

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(q_1, x_2, q_3) dq_1 dq_3 = \int_{1/3}^1 \frac{3}{2} \left(1 - \frac{1}{2}x_2\right) dq_3 = 1 - \frac{1}{2}x_2$$

при $0 \leq x_2 \leq 2$;

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(q_1, q_2, x_3) dq_1 dq_2 = \int_0^2 \frac{3}{2} \left(1 - \frac{1}{2}q_2\right) dq_2 = \frac{3}{2}$$

при $\frac{1}{3} \leq x_3 \leq 1$. При других значениях аргументов, кроме приведенных выше, значения всех маргинальных функций плотности равны нулю.

Для исследования независимости случайных величин X_1 , X_2 и X_3 используем теорему 1.8:

$$f_{X_1, X_2}(x_1, x_2) \neq f_{X_1}(x_1) \cdot f_{X_2}(x_2),$$

поэтому случайные величины X_1 и X_2 не независимы;

$$f_{X_1, X_3}(x_1, x_3) = 3(1-x_1) = 2(1-x_1) \cdot \frac{3}{2} = f_{X_1}(x_1) \cdot f_{X_3}(x_3),$$

поэтому случайные величины X_1 и X_3 независимы;

$$f_{X_2, X_3}(x_2, x_3) = \frac{3}{2} \left(1 - \frac{1}{2} x_2\right) = \left(1 - \frac{1}{2} x_2\right) \cdot \frac{3}{2} = f_{X_2}(x_2) \cdot f_{X_3}(x_3),$$

поэтому случайные величины X_2 и X_3 независимы. В соответствии с определением независимости произвольного числа случайных величин (см. [ТВиМС-1, раздел 1.2]), случайные величины X_1 , X_2 , X_3 не могут быть независимыми, если не независимы случайные величины X_1 и X_2 .

З а м е ч а н и е. Рассмотренный пример интересен для нас еще в одном отношении. Ни одна из маргинальных функций распределения не является дифференцируемой во всех точках множества R^2 . Так, например, маргинальная функция распределения

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{X_1, X_2}(q_1, q_2) dq_1 dq_2$$

не имеет конечной производной

$$\frac{\partial^2 F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

на границе треугольника $x_1 \geq 0$, $x_2 \geq 0$, $2x_1 + x_2 \leq 2$. Это означает, что формально мы не имели права применять теорему 1.8 для исследования независимости случайных величин. Однако, как уже было сказано в замечании 2 после теоремы 1.8, условие дифференцируемости всюду совместной функции распределения может быть ослаблено. И эта более общая теорема может быть применена в рассматриваемом примере.

Отметим также, что функцию $F_{X_1, X_2}(x_1, x_2)$, которая является совместной функцией распределения случайных величин X_1 и X_2 , мы называем маргинальной функцией распределения, если рассматриваются случайные величины X_1 , X_2 и X_3 . То же относится и к другим таким же совместным функциям распределения и совместным функциям плотности.

Вернемся к случаю $k = 2$. Обозначим через Q_1 множество тех $x_1 \in R$, для которых $f_{X_1}(x_1) > 0$, и через Q_2 — множество тех $x_2 \in R$, для которых $f_{X_2}(x_2) > 0$. При $x_1 \in R$, $x_2 \in Q_2$ положим

$$f(x_1 | X_2 = x_2) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)};$$

при $x_2 \in R$, $x_1 \in Q_1$ положим

$$f(x_2 | X_1 = x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}.$$

Функции $f(x_1 | X_2 = x_2)$ и $f(x_2 | X_1 = x_1)$ называются *условными функциями плотности*. Из определения маргинальных функций плотности следует, что

$$\int_{-\infty}^{\infty} f(x_1 | X_2 = x_2) dx_1 = 1, \quad \int_{-\infty}^{\infty} f(x_2 | X_1 = x_1) dx_2 = 1.$$

Из теоремы 1.8 следует, что если случайные величины X_1 и X_2 независимы, то при любых $x_1 \in R$, $x_2 \in Q_2$

$$f(x_1 | X_2 = x_2) = f_{X_1}(x_1)$$

и при любых $x_2 \in R$, $x_1 \in Q_1$

$$f(x_2 | X_1 = x_1) = f_{X_2}(x_2).$$

При $x_1 \in R$, $x_2 \in Q_2$ положим

$$F(x_1 | X_2 = x_2) = \int_{-\infty}^{x_1} f(q_1 | X_2 = x_2) dq_1;$$

при $x_2 \in R$, $x_1 \in Q_1$ положим

$$F(x_2 | X_1 = x_1) = \int_{-\infty}^{x_2} f(q_2 | X_1 = x_1) dq_2.$$

Функции $F(x_1 | X_2 = x_2)$ и $F(x_2 | X_1 = x_1)$ называются *условными функциями распределения*.

Определение условных функций плотности внешне совпадает с определением условных дискретных функций плотности, данным в разделе 1.2. Но в случае совместных дискретных распределений из предположения $f_{X_2}(x_2) > 0$ следует, что вероятность события $\{X_2 = x_2\}$ больше нуля. Поэтому, как показано в разделе 1.2, в случае совместных дискретных распределений $f(x_1 | X_2 = x_2)$ — это условная вероятность события $\{X_1 = x_1\}$ при условии $\{X_2 = x_2\}$.

В случае совместных непрерывных распределений при любом $x_2 \in Q_2$ вероятность события $\{X_2 = x_2\}$ равна нулю. Поэтому использовать для расчетов условных вероятностей при условии $\{X_2 = x_2\}$ обычную формулу

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

нельзя. Однако и в случае совместных непрерывных распределений к определению условной функции плотности $f(x_1 | X_2 = x_2)$ можно придти, начиная с условных вероятностей. На нестрогом уровне проведем соответствующие рассуждения.

Пусть $x_2 \in Q_2$ и $\delta > 0$ — некоторое малое положительное число. Если к условию $f_{X_2}(x_2) > 0$ добавить некоторые условия регулярности функции f_{X_2} , например, положительность в некоторой окрестности точки x_2 , то

$$\mathbf{P}\left(x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right) = \int_{x_2 - \delta/2}^{x_2 + \delta/2} f_{X_2}(q_2) dq_2 > 0.$$

Пусть $x_1 \in R$ и $\varepsilon > 0$ — некоторое малое положительное число. Тогда можно определить условную вероятность

$$\begin{aligned} \mathbf{P}\left(x_1 - \frac{\varepsilon}{2} < X_1 \leq x_1 + \frac{\varepsilon}{2} \mid x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right) &= \\ &= \frac{\mathbf{P}\left(x_1 - \frac{\varepsilon}{2} < X_1 \leq x_1 + \frac{\varepsilon}{2}, x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right)}{\mathbf{P}\left(x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right)}. \end{aligned}$$

Воспользовавшись теоремой 1.7, получаем сначала точное, а затем упрощенное приближенное выражение для числителя:

$$\begin{aligned} \mathbf{P}\left(x_1 - \frac{\varepsilon}{2} < X_1 \leq x_1 + \frac{\varepsilon}{2}, x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right) &= \\ &= \int_{x_2 - \delta/2}^{x_2 + \delta/2} \int_{x_1 - \varepsilon/2}^{x_1 + \varepsilon/2} f(q_1, q_2) dq_1 dq_2 \approx f(x_1, x_2) \cdot \delta \cdot \varepsilon. \end{aligned}$$

Аналогично для знаменателя:

$$\mathbf{P}\left(x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right) = \int_{x_2 - \delta/2}^{x_2 + \delta/2} f_{X_2}(q_2) dq_2 \approx f_{X_2}(x_2) \cdot \delta.$$

Таким образом,

$$\begin{aligned} \mathbf{P}\left(x_1 - \frac{\varepsilon}{2} < X_1 \leq x_1 + \frac{\varepsilon}{2} \mid x_2 - \frac{\delta}{2} < X_2 \leq x_2 + \frac{\delta}{2}\right) &\approx \\ &\approx \frac{f(x_1, x_2)}{f_{X_2}(x_2)} \cdot \varepsilon \approx \int_{x_1 - \varepsilon/2}^{x_1 + \varepsilon/2} \frac{f(q_1, x_2)}{f_{X_2}(x_2)} dq_1. \end{aligned}$$

Но поскольку правая часть не зависит от δ , появляется возможность определить условную вероятность:

$$\mathbf{P}\left(x_1 - \frac{\varepsilon}{2} < X_1 \leq x_1 + \frac{\varepsilon}{2} \mid X_2 = x_2\right) = \int_{x_1 - \varepsilon/2}^{x_1 + \varepsilon/2} \frac{f(q_1, x_2)}{f_{X_2}(x_2)} dq_1.$$

Отсюда соответствующая условная функция плотности:

$$f(q_1 | X_2 = x_2) = \frac{f(q_1, x_2)}{f_{X_2}(x_2)}.$$

Еще раз подчеркнем, что приведенные нестрогие рассуждения не служат никаким доказательством, а только помогают понять определение условной функции плотности.

Пример 1.7 (продолжение). Требуется найти условную функцию плотности $f(x_1 | X_2 = 1)$. Найти условную вероятность: $\mathbf{P}(X_1 > 1/4 | X_2 = 1)$.

В определении условной функции плотности $f(x_1 | X_2 = 1)$ случайная величина X_3 не участвует, поэтому роль совместной функции плотности играет маргинальная функция плотности $f_{X_1, X_2}(x_1, x_2)$. Выражения для $f_{X_1, X_2}(x_1, x_2)$ и для $f_{X_2}(x_2)$ ранее были найдены. Напомним, что

$$f_{X_1, X_2}(x_1, x_2) = 1$$

при $x_1 \geq 0, x_2 \geq 0, 2x_1 + x_2 \leq 2$;

$$f_{X_2}(x_2) = 1 - \frac{1}{2}x_2$$

при $0 \leq x_2 \leq 2$. При других значениях аргументов значения функций f_{X_1, X_2} и f_{X_2} равны нулю. Таким образом,

$$f(x_1 | X_2 = 1) = \frac{f_{X_1, X_2}(x_1, 1)}{f_{X_2}(1)} = \frac{1}{0.5} = 2$$

при $0 \leq x_1 \leq 0.5$ и

$$f(x_1 | X_2 = 1) = 0$$

при остальных x_1 .

Найдя условную функцию плотности, определяем условную вероятность:

$$\mathbf{P}\left(X_1 > \frac{1}{4} \mid X_2 = 1\right) = \int_{1/4}^{\infty} f(q_1 | X_2 = 1) dq_1 = \int_{1/4}^{1/2} 2 dq_1 = 0.5.$$

Замечание. Для полноты изложения упомянем про случай $k = 1$. Если функция распределения случайной величины X_1 имеет вид

$$F_{X_1}(x_1) = \int_{-\infty}^{x_1} f(q_1) dq_1,$$

то распределение вероятностей случайной величины X_1 называют *непрерывным распределением*, а функцию f — *функцией плотности*. Говорить о маргинальных функциях распределения или об условных функциях распределения при $k = 1$ не имеет смысла.

- Функция потерь 214
— — 0–1 218
— — при последовательном анализе 249
— правдоподобия 76
— риска 216, 249
- Центральная предельная теорема 97
Центральный момент 47
- Цепь Маркова 197
- Экспоненциальное распределение 59, 174
Элементарное событие 11
Эллипс рассеивания 66
Эмпирическая функция распределения 154, 155
Эффективная оценка 94

Учебное издание

Серия «Учебники Высшей школы экономики»

Алексей Сергеевич Шведов

**Теория вероятностей и математическая статистика:
промежуточный уровень**

Зав. редакцией *Е.А. Бережнова*

Редактор *И.Л. Легостаева*

Художественный редактор *А.М. Павлов*

Компьютерная верстка и графика: *И.Г. Андреева*

Корректор *Н.М. Дмуховская*

Подписано в печать 06.04.2016. Формат 70×100/16

Печать офсетная. Усл.-печ. л. 17,5. Уч.-изд. л. 18,0

Тираж 600 экз. Изд. № 1875

Национальный исследовательский университет «Высшая школа экономики»

101000, Москва, ул. Мясницкая, 20

Тел./факс: (499) 611-15-52

Отпечатано в ГУП ЧР «ИПК «Чувашия» Мининформполитики Чувашии

428019, г. Чебоксары, пр. И. Яковлева, 13

Тел.: (8352) 56-00-23