

МИНОБРНАУКИ РОССИИ

---

Санкт-Петербургский государственный электротехнический  
университет „ЛЭТИ“

---

А. И. ДАУГАВЕТ Е. В. ПОСТНИКОВ  
Н. М. ЧЕРВИНСКАЯ

## ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

Учебное пособие

Санкт-Петербург  
Издательство СПбГЭТУ „ЛЭТИ“  
2012

МИНОБРНАУКИ РОССИИ

---

Санкт-Петербургский государственный электротехнический  
университет „ЛЭТИ“

---

А. И. ДАУГАВЕТ Е. В. ПОСТНИКОВ  
Н. М. ЧЕРВИНСКАЯ

## ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

Учебное пособие

Санкт-Петербург  
Издательство СПбГЭТУ „ЛЭТИ“  
2012

УДК 519.21(075)

ББК В171

Д 21

Д21 Даугавет А. И., Постников Е. В., Червинская Н. М. Введение в теорию вероятностей: Учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2012. 80 с.

ISBN 978-5-7629-

Содержит материалы по основным разделам теории вероятностей: элементарной теории вероятностей, теории распределения случайных величин и векторов, основные предельные теоремы, входящие в программу курса дисциплины „Теория вероятностей и математическая статистика“ технических университетов. Является первой частью курса и соответствует программе дисциплины „Теория вероятностей и математическая статистика“.

Предназначено для студентов бакалавров технических факультетов, обучающихся по всем направлениям и специальностям.

Издание позволяет студентам самостоятельно и более углубленно изучать разделы курса, кратко изложенные на лекциях. Достаточное количество, подробно решенных примеров должно помочь студентам в самостоятельной работе, при выполнении текущих домашних заданий.

УДК УДК 519.21(075)

ББК В171

Рецензенты: кафедра высшей математики СПХФА;  
д-р физ.-мат. наук, проф. Я. И. Белопольская (СПбГАСУ).

Утверждено  
редакционно-издательским советом университета  
в качестве учебного пособия

ISBN 978-5-7629-

© СПбГЭТУ „ЛЭТИ“, 2012

## Введение

Данное издание предназначено для студентов второго курса всех направлений и специальностей факультетов электротехники и автоматизации, электроники, экономики и менеджмента и открытого факультета СПбГЭТУ. Пособие написано на основе курса, читаемого авторами, и является исправленным и дополненным изданием учебного пособия [1] тех же авторов. Основным дополнением является включение нового раздела „Элементы комбинаторики“. В пособии излагается теория вероятностей, соответствующая первой части новой рабочей программы для бакалавров по дисциплине „Теория вероятностей и математическая статистика“. Вторая часть данной дисциплины изложена в выходящем в этом году пособии [2].

Для более глубокого изучения некоторых разделов теории вероятностей можно рекомендовать учебное пособие для вузов [3].

Для самостоятельной работы студентов авторы рекомендуют книгу [4], в которой разбирается большое количество примеров и каждый параграф предваряется необходимыми теоретическими сведениями.

Отметим некоторые стандартные обозначения, принятые в пособии. Множество натуральных чисел обозначается символом  $\mathbb{N}$ ; множество целых чисел —  $\mathbb{Z}$ ; множество вещественных чисел —  $\mathbb{R}$ ; множество комплексных чисел —  $\mathbb{C}$ . Знак  $\blacksquare$  отмечает конец доказательства; а знак  $\bullet$  — конец примера.

# 1. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ВЕРОЯТНОСТЕЙ

## 1.1. Случайный эксперимент

В окружающем нас мире происходят явления, результат которых не вполне однозначно определен. Такие явления принято называть случайными, в отличие от детерминированных, результат которых всегда один при неизменных условиях. Явление, условия протекания которого можно контролировать, будем называть экспериментом. Таким образом, эксперимент — это выполнение некоторых действий при фиксированном наборе определенных условий. Результат эксперимента называется его исходом. У случайного эксперимента исходов несколько, у детерминированного — всегда один. Предполагается, что случайный эксперимент может быть многократно повторен, хотя бы мысленно.

Простейшим примером случайного эксперимента является подбрасывание монеты: монета может упасть либо гербом вверх, либо цифрой. Таким образом, у этого случайного эксперимента 2 исхода. Другим случайнym экспериментом является измерение некоторого постоянного физиче-

ского параметра, например сопротивления резистора, одним и тем же прибором.

Результаты нескольких повторных измерений в общем случае будут несколько отличаться друг от друга. Эти различия обусловлены влиянием многих второстепенных факторов, таких, как переходное сопротивление контактов, случайные вибрации аппаратуры, погрешности отсчета показаний и др. Совместное влияние всех второстепенных факторов, которые невозможно контролировать, приводит к различным результатам измерений, т. е. к множеству исходов эксперимента.

Приведенные примеры показывают, что несмотря на то, что основные условия эксперимента остаются неизменными, его исходы могут быть различными. Эта неоднозначность результата всегда связана с наличием некоторых второстепенных факторов, влияющих на исход эксперимента, но не заданных в числе основных условий. Совершенно очевидно, что в природе нет ни одного физического явления, в котором не присутствовал бы в той или иной мере элемент случайности. Никогда нельзя в точности повторить абсолютно все условия опыта.

Вернемся к эксперименту с подбрасыванием монеты. Замечено, что если монета „правильная“ (не имеет изъянов), то при многократных подбрасываниях стороны выпадают примерно одинаковое число раз. Допустим, что проведена серия из  $n$  подбрасываний, причем герб выпал  $k$  раз. Назовем частотой выпадения герба число  $k/n$ . Тогда оказывается, что при большом числе подбрасываний  $n$  частота  $k/n$  близка к 0.5. При этом по мере удлинения серии подбрасываний колебания значения частоты относительно 0.5 становятся меньше.

Свойство группировки частоты около некоторого числа при многократном повторении случайного эксперимента носит название статистической устойчивости частот.

Все мыслимые эксперименты можно разделить на 3 группы. К первой группе относятся детерминированные эксперименты, обладающие полной устойчивостью исходов. Ко второй группе относятся случайные эксперименты, в которых наблюдается статистическая устойчивость частот. В третью группу входят случайные эксперименты, у которых статистической устойчивости нет. Сферой применения теории вероятностей являются случайные эксперименты второй группы. Таким образом, теория вероятностей как математическая дисциплина занимается изучением математических моделей случайных экспериментов, обладающих свойством статистической устойчивости частот.

Теория вероятностей зародилась в средние века, когда делались попытки разработать выигрышные стратегии для азартных игр. Ее развитие шло по пути накопления фактов и математических результатов, и лишь в

XX в. теория вероятностей превратилась в самостоятельный раздел математики. Следует отметить, что методами теории вероятностей нельзя предсказать исход очередного случайного эксперимента, однако можно выявить основные закономерности, присущие данному эксперименту.

## 1.2. Математическая модель случайного эксперимента

Случайный эксперимент обычно описывается на языке высказываний. Основным же объектом исследований в теории вероятностей является математическая модель случайного эксперимента, которая описывается на языке теории множеств. Построение математической модели случайного эксперимента является неформализуемой процедурой, выходящей за рамки теории вероятностей и целиком относящейся к компетенции исследователя. Рассмотрим основные компоненты математической модели эксперимента в рамках элементарной теории вероятностей.

**Определение 1.1.** Все взаимоисключающие исходы эксперимента  $\omega_2$  называются элементарными событиями.

Будем считать, что множество элементарных событий  $\Omega$  конечно или счетно, т. е.

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

или

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}.$$

**Определение 1.2.** Конечное или счетное множество элементарных событий называется дискретным пространством элементарных событий.

Построение математической модели эксперимента обычно начинают с построения дискретного пространства элементарных событий  $\Omega$ .

**Пример.** Эксперимент состоит в подбрасывании игральной кости и регистрации числа очков, выпавших на верхней грани. Естественно считать элементарным событием (исходом) выпавшее число от 1 до 6. Таким образом, элементарные события – это:  $\omega_1 = 1, \omega_2 = 2, \omega_3 = 3, \omega_4 = 4, \omega_5 = 5, \omega_6 = 6$ . •

**Определение 1.3.** Любое подмножество дискретного пространства элементарных событий называется событием.

Множество всех событий на  $\Omega$  – это множество всех его подмножеств. В частности, событиями являются следующие множества: пустое множество  $\emptyset$ , которое называют невозможным событием, так как оно не имеет

благоприятных исходов; множество всех исходов  $\Omega$  – его называют достоверным событием; множество, состоящее из одного элементарного события  $\{\omega_i\}$ .

**Пример.** Эксперимент состоит в подбрасывании игральной кости. Рассмотрим событие  $A$ : „выпало число не меньше, чем 5“. Это множество содержит 2 элементарных события:  $A = \{5, 6\}$ . Исходы  $\omega_5 = 5$  и  $\omega_6 = 6$  являются благоприятными для  $A$ . Невозможное событие – „выпало число 8“. Достоверное событие – „выпало число от 1 до 6“.

Событие  $A$  наступило, если наступило любое из элементарных событий, составляющих событие  $A$ , или, другими словами, благоприятных событию  $A$ .

Итак, дискретное пространство элементарных событий  $\Omega$  есть первый компонент математической модели случайного эксперимента. •

Перейдем теперь к определению другого компонента модели – вероятностного распределения.

**Определение 1.4.** Пусть каждому элементарному событию  $\omega_i \in \Omega$  поставлено в соответствие число  $P(\omega_i)$ , причем

$$0 \leq P(\omega_i) \leq 1, \quad (1.1)$$

$$\sum_i P(\omega_i) = 1, \quad (1.2)$$

тогда число  $P(\omega_i)$  называется вероятностью элементарного события  $\omega_i$ .

Условие (1.1) выражает тот факт, что вероятность любого элементарного события есть число между нулем и единицей, а условие (1.2) требует, чтобы сумма вероятностей всех элементарных событий равнялась единице. При этом, если пространство элементарных событий  $\Omega$  конечно и количество элементарных событий равно  $n$ , то это обычная сумма  $n$  слагаемых. Если же  $\Omega$  счетно, то в условии (1.2) рассматривается сумма числового положительного ряда.

**Определение 1.5.** Вероятностью  $P(A)$  любого события  $A$  на  $\Omega$  называется сумма вероятностей всех элементарных событий, благоприятных  $A$ :

$$P(A) = \sum_{\omega_i \in A} P(\omega_i). \quad (1.3)$$

Из определения 1.5, в частности, следует, что  $P(\emptyset) = 0$ ,  $P(\Omega) = 1$ .

**Определение 1.6.** Вероятностным распределением на  $\Omega$  называется функция  $P$ , которая каждому событию  $A \subset \Omega$  ставит в соответствие его вероятность  $P(A)$ .

Каждое вероятностное распределение на  $\Omega$  удовлетворяет условиям (1.1)–(1.3). Их обычно называют аксиомами вероятностного распределения. С другой стороны, любая функция, заданная на множестве всех подмножеств  $\Omega$ , удовлетворяющая аксиомам (1.1)–(1.3), определяет на  $\Omega$  вероятностное распределение.

Возникает вопрос: откуда брать вероятностное распределение? Решение этой задачи целиком лежит на совести экспериментатора, который строит математическую модель случайного эксперимента. Вспомним, что случайные эксперименты, модели которых изучают в теории вероятностей, обязаны обладать свойством статистической устойчивости частот. Это означает, что частота элементарного события (исхода)  $\omega_i$  по мере увеличения количества проводимых экспериментов должна стабилизироваться около некоторого неизвестного числового значения. В качестве абстрактного аналога этого числового значения в математической модели как раз и вводится вероятность элементарного события  $P(\omega_i)$ . В конкретной серии экспериментов используют частоту  $\nu = k/n$ , в математической модели эксперимента вместо частоты  $\nu$ , которая может изменяться от серии к серии, введена постоянная вероятность  $P(\omega_i)$ , отражающая предположение о неизменности условий проведения эксперимента. В свою очередь, вероятность исхода  $P(\omega_i)$  тогда должна интерпретироваться как ожидаемая частота  $\nu$  наступления элементарного события  $\omega_i$  в достаточно длинной серии экспериментов. Таким образом, так как модель должна быть адекватна породившему ее эксперименту, вероятность каждого исхода следует назначать так, чтобы она соответствовала частоте этого исхода.

**Пример.** Эксперимент состоит в подбрасывании монеты. Тогда  $\Omega = \{\Gamma, \Ц\} = \{\omega_1, \omega_2\}$ . Зададим на  $\Omega$  вероятностное распределение. Ясно, что это можно сделать многими способами, например:

- 1)  $P(\omega_1) = P(\omega_2) = 1/2$ ;
- 2)  $P(\omega_1) = 1/3, P(\omega_2) = 2/3$ ;
- 3)  $P(\omega_1) = p, P(\omega_2) = 1 - p, 0 \leq p \leq 1$ .

Во всех трех случаях аксиомы вероятностного распределения выполнены. Следовательно, эти 3 варианта задания вероятностного распределения на  $\Omega$  можно расценить как 3 различные модели одного и того же эксперимента. Какую из этих моделей считать адекватной эксперименту?

Если монета „правильная“, то, как подсказывает жизненный опыт, обе ее стороны выпадают одинаково часто. Следовательно, частота выпадения герба или цифры будет группироваться около числа  $1/2$ . Английский статистик К. Пирсон, например, подбрасывал „правильную“ монету 24 000 раз, и герб выпал 12 012 раз. Логические умозаключения позволяют сделать вывод о равновозможности выпадения как герба, так и цифры. Значит, естественно считать вероятность выпадения герба „правильной“ моне-

ты равной  $1/2$ . Поэтому первая модель – это модель для „правильной“ монеты, и она будет более точно, чем другие модели, описывать эксперимент с „правильной“ монетой. Тогда модель № 2 будет описывать эксперимент с „неправильной“ монетой, у которой цифра выпадает в 2 раза чаще, чем герб.

С помощью модели № 3 можно описать эксперимент с любой монетой. Для этого достаточно в качестве числа  $p$  выбрать частоту выпадения герба, проведя серию подбрасываний выбранной монеты. Если условия эксперимента таковы, что все исходы можно считать равновозможными, то всем элементарным событиям приписываются одинаковую вероятность. Такое вероятностное распределение называют классическим, а соответствующую модель – схемой равновозможных исходов.

Задание вероятностного распределения завершает построение математической модели случайного эксперимента. Таким образом, модель включает в себя 2 элемента: дискретное пространство элементарных событий – множество  $\Omega$  и вероятностное распределение  $P$ , приписывающее каждому элементарному событию  $\omega_i \in \Omega$  его вероятность  $P(\omega_i)$ . Теперь вероятность любого события  $A \subset \Omega$  может быть вычислена в соответствии с аксиомой (1.3) суммированием вероятностей всех элементарных событий, благоприятных  $A$ .

В случае классического распределения вероятностей такой подсчет осуществляется особенно просто. Пусть  $\Omega = \{\omega_1, \dots, \omega_n\}$ , где  $n = |\Omega|$  – мощность (количество элементов)  $\Omega$ . В этом случае полагают

$$P(\omega_i) = \frac{1}{|\Omega|}, \quad i = 1, 2, \dots, n.$$

Тогда вероятность события  $A$ , согласно (1.3), вычисляется как

$$P(A) = \sum_{\omega_i \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|},$$

где  $|A|$  – число элементов множества  $A$ , т. е. число благоприятных для  $A$  исходов.

**Пример.** Куб с окрашенными гранями распилен на 27 кубиков одинакового размера (рис. 1.1). Полученные кубики тщательно перемешаны. Наудачу извлекается один кубик. Построить модель эксперимента и указать вероятность того, что извлеченный кубик имеет ровно две окрашенные грани.

Как отмечалось, модель случайного эксперимента состоит из дискретного пространства элементарных событий  $\Omega$  и заданного на  $\Omega$  вероятностного распределения  $P$ , приписывающего каждому элементарному событию

$\omega_i$  его вероятность  $P(\omega_i)$ . Эксперимент заключается в вытаскивании кубика и определении числа окрашенных граней. Как определить  $\Omega$  и что взять в качестве элементарного события?

Можно, к примеру, выбрать в качестве исхода эксперимента число окрашенных граней у выбранного кубика. Тогда

$$\Omega_1 = \{0_{\text{о. г.}}, 1_{\text{о. г.}}, 2_{\text{о. г.}}, 3_{\text{о. г.}}\}.$$

Теперь нужно задать вероятности исходов. Если действовать по схеме классической вероятности, то

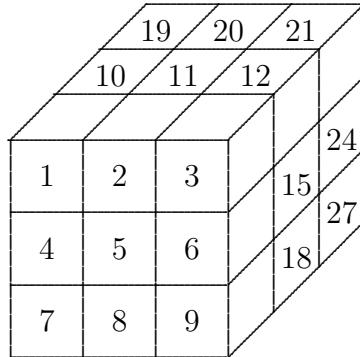


рис. 1.1

$$P(0_{\text{о. г.}}) = P(1_{\text{о. г.}}) = P(2_{\text{о. г.}}) = P(3_{\text{о. г.}}) = \frac{1}{4}.$$

Однако разумна ли гипотеза о равновозможности исходов? Очевидно, что нет, так как неокрашенный кубик один, а остальных больше, значит, частота вытаскивания неокрашенного кубика будет меньше частоты других исходов. Поэтому построенная модель плохо описывает эксперимент.

Построим другое пространство,  $\Omega_2 = \{1, 2, 3, \dots, 27\}$ , где исход – номер вытащенного кубика, который он получил при распиливании. Отметим, что исходы пространства  $\Omega_1$  являются событиями пространства  $\Omega_2$ :

$$\begin{aligned} 0_{\text{о. г.}} &= \{14\}; \\ 1_{\text{о. г.}} &= \{5, 11, 13, 15, 17, 23\}; \\ 2_{\text{о. г.}} &= \{2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 26\}; \\ 3_{\text{о. г.}} &= \{1, 3, 7, 9, 19, 21, 25, 27\}. \end{aligned}$$

Слова „тщательно перемешиваются“ и „наудачу“ в описании эксперимента позволяют сделать предположение о равновозможности исходов в пространстве  $\Omega_2$ , так как все 27 кубиков ничем друг от друга не отличаются. Значит, вероятности всех исходов можно принять одинаковыми и равными  $1/27$  в силу их равновозможности. Тогда  $P(0_{\text{о. г.}}) = 1/27$ ,  $P(1_{\text{о. г.}}) = 6/27$ ,  $P(2_{\text{о. г.}}) = 12/27$ ,  $P(3_{\text{о. г.}}) = 8/27$ .

Итак, вероятность события „ $2_{\text{о. г.}}$ “ равна  $P(2_{\text{о. г.}}) = 4/9$ . •

### 1.3. Операции над событиями

Пусть  $\Omega$  – дискретное пространство элементарных событий,  $A, B, C$  – события. Дадим вероятностную интерпретацию некоторых фактов теории множеств:

- $\omega \in A$  (элемент  $\omega$  принадлежит множеству  $A$ ) – исход  $\omega$  благоприятен для события  $A$ ;
- $A \subset B$  (множество  $A$  содержится во множестве  $B$ ) – событие  $A$  влечет за собой событие  $B$ ;
- $A = B$  (множества  $A$  и  $B$  равны) – события  $A$  и  $B$  состоят из одних и тех же элементарных событий.

**Определение 1.7.** Суммой событий  $A$  и  $B$  называется событие  $C$ , содержащее все исходы  $\omega \in \Omega$ , благоприятные либо событию  $A$ , либо событию  $B$ , обозначается  $C = A \cup B$ .

**Определение 1.8.** Произведением событий  $A$  и  $B$  называется событие  $C$ , состоящее из исходов, благоприятных как для события  $A$ , так и для события  $B$ , обозначается  $C = A \cap B$ .

**Определение 1.9.** Разностью событий  $A$  и  $B$  называется событие  $C$ , состоящее из исходов, благоприятных для события  $A$ , но неблагоприятных для события  $B$ , обозначается  $C = A \setminus B$ .

**Определение 1.10.** События  $A$  и  $B$  называются несовместными, если они не имеют общих исходов, т. е. если  $A \cap B = \emptyset$ .

**Определение 1.11.** Событие  $C$  называется противоположным событию  $A$ , если  $C$  состоит из всех исходов, неблагоприятных для  $A$  (дополняет  $A$  до  $\Omega$ ), обозначается  $C = \Omega \setminus A = \bar{A}$ .

**Пример.** Рассмотрим эксперимент – двукратное подбрасывание монеты. Тогда  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} = \{\text{ГГ}, \text{ГЦ}, \text{ЦГ}, \text{ЦЦ}\}$ . Пусть событие  $A$  – „герб выпал хотя бы один раз“, событие  $B$  – „монеты упали разными сторонами“. Значит,

$$A = \{\omega_1, \omega_2, \omega_3\}, \quad B = \{\omega_2, \omega_3\}.$$

Тогда

$$A \cup B = \{\omega_1, \omega_2, \omega_3\} = A, \quad A \cap B = \{\omega_2, \omega_3\} = B, \quad \bar{A} = \{\omega_4\}, \\ \bar{B} = \{\omega_1, \omega_4\}, \quad \bar{A} \cap B = \emptyset, \quad A \cup \bar{A} = \Omega.$$

Введенные операции над событиями позволяют конструировать новые события пространства элементарных событий. Далее покажем как вычислять вероятности этих новых событий.

#### 1.4. Теорема сложения. Вероятность противоположного события

Пусть  $\Omega$  – пространство элементарных событий,  $P$  – вероятностное распределение на  $\Omega$ . Рассмотрим как связаны вероятности  $P(A)$ ,  $P(B)$  событий  $A$  и  $B$  и вероятность  $P(A \cup B)$  их суммы  $A \cup B$ .

**Теорема 1.1.** Если события  $A$  и  $B$  несовместны, то вероятность суммы событий равна сумме их вероятностей:

$$A \cap B = \emptyset, \Rightarrow P(A \cup B) = P(A) + P(B).$$

**Доказательство.** На основании аксиомы аддитивности (1.3) можно записать, что

$$P(A \cup B) = \sum_{\omega \in (A \cup B)} P(\omega).$$

С учетом условия  $A \cap B = \emptyset$  последняя сумма разбивается на две:

$$\sum_{\omega \in (A \cup B)} = \sum_{\omega \in A} P(\omega) + \sum_{\omega \in B} P(\omega) = P(A) + P(B).$$

Теорема доказана. ■

Следующая теорема, называемая теоремой сложения, позволяет вычислить вероятность суммы событий для любых событий  $A$  и  $B$ .

**Теорема 1.2. Теорема сложения.** Для любых событий  $A$  и  $B$

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

**Доказательство.** Можно записать, что

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B). \quad (1.5)$$

Для графического представления подобных тождеств в теории множеств принято пользоваться так называемыми диаграммами Венна. При этом множество изображается плоской фигурой. Диаграмма Венна для тождества (1.5) представлена на рис. 1.2.

Итак, представим сумму событий  $A$  и  $B$  в виде суммы трех попарно несовместных событий. Согласно теореме (1.1),

$$P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B), \quad (1.6)$$

С другой стороны (рис. 1.2),

$$A = A \setminus B + A \cap B, \quad B = B \setminus A + A \cap B.$$

Значит,

$$P(A) = P(A \setminus B) + P(A \cap B), \quad (1.7)$$

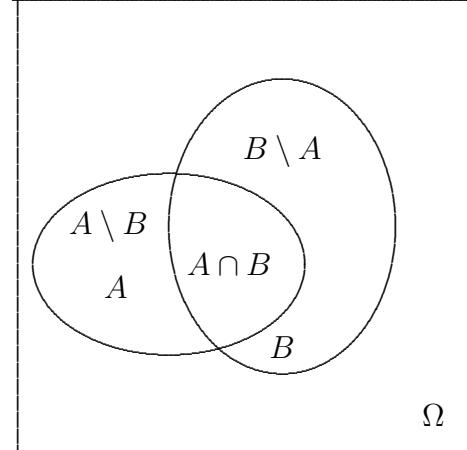


рис. 1.2

$$P(B) = P(B \setminus A) + P(A \cap B). \quad (1.8)$$

Тогда, подставляя в (1.6)  $P(A \setminus B)$  из (1.7) и  $P(B \setminus A)$  из (1.8), получим утверждение теоремы (1.4):

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

**Пример.** Из урны, содержащей 29 пронумерованных шаров, наудачу выбирается один шар. Какова вероятность того, что его номер делится на 2 или на 3?

Пусть  $\Omega = \{1, 2, \dots, 29\}$  – множество исходов, состоящее из 29 номеров шаров, которые могут быть выбраны в результате эксперимента. Естественно предположить, что все исходы равновозможны, поэтому каждому исходу приписываем вероятность  $1/29$ .

Введем события:  $A$  – „номер делится на 2“,  $B$  – „номер делится на 3“. Теперь необходимо найти вероятность события  $C = A \cup B$ . Можно непосредственно выписать все элементарные события, входящие в  $C$ . Однако это путь громоздкий. Воспользуемся теоремой сложения.

Найдем число элементарных событий, благоприятных событиям  $A$  и  $B$  и событию  $A \cap B$  – „номер делится на 6“:

$$n_A = [29/2] = 14, \quad n_B = [29/3] = 9, \quad n_{A \cap B} = [29/6] = 4$$

(здесь знак  $[ \cdot ]$  обозначает целую часть числа). Тогда  $P(A) = 14/29$ ,  $P(B) = 9/29$ ,  $P(A \cap B) = 4/29$ . Наконец,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 14/29 + 9/29 - 4/29 = 19/29. \bullet$$

Следующая теорема о вероятности противоположного события оказывается очень полезной при решении многих задач.

**Теорема 1.3.** *Сумма вероятностей противоположных событий равна 1.*

**Доказательство.** Пусть  $A$  – произвольное событие. По определению противоположного события  $A \cup \bar{A} = \Omega$  и  $A \cap \bar{A} = \emptyset$ . Тогда по теореме сложения

$$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) = 1. \blacksquare$$

**Пример.** Из колоды в 36 карт вытаскиваются 3 карты. Какова вероятность события „вытащена хотя бы одна дама или хотя бы один король“?

Пусть  $\Omega$  есть множество всех сочетаний по 3 карты из 36. Количество элементов в  $\Omega$  есть  $n = C_{36}^3$ . Так как все исходы равновозможны, приписываем им одинаковые вероятности  $P(\omega_i) = 1/n$ ,  $i = 1, 2, \dots, n$ . Введем события:  $A$  – „вытащена хотя бы одна дама“,  $B$  – „вытащена хотя бы один король“. Требуется найти вероятность  $P(A \cup B)$ .

Перейдем к противоположному для  $A \cup B$  событию  $\overline{A \cup B} = \bar{A} \cap \bar{B}$  – „не вытащены ни одна дама и ни один король“. Количество благоприятных исходов для  $\bar{A} \cap \bar{B}$  очевидно равно  $m = C_{28}^3$ . Значит,  $P(\bar{A} \cap \bar{B}) = m/n$ . Тогда, согласно теореме 1.3,

$$P(A \cup B) = 1 - P(\bar{A} \cap \bar{B}) = 1 - m/n. \bullet$$

### 1.5. Условная вероятность события.

#### Условное распределение вероятностей

Рассмотрим важное в теории вероятностей понятие условной вероятности события, позволяющее описывать целую группу случайных экспериментов со сложной структурой. Обсуждение начнем со следующей задачи.

**Пример.** Студенческая группа состоит из  $n$  студентов. При этом в ней  $n_1$  спортсменов ( $n_1 \leq n$ ),  $n_2$  отличников ( $n_2 \leq n$ ) и из спортсменов-отличников  $n_3 \leq n_1, n_3 \leq n_2$ . Наугад выбирается студент. Найти вероятности событий:  $A$  – „выбран спортсмен“;  $B$  – „выбран отличник“;  $A \cap B$  – „выбран спортсмен-отличник“.

Очевидно, все исходы, заключающиеся в выборе одного студента из  $n$  возможных, равновозможны. Значит,  $\Omega = \{\omega_1, \dots, \omega_n\}$  и  $P(\omega_i) = 1/n$ . Тогда  $P(A) = n_1/n$ ;  $P(B) = n_2/n$ ;  $P(A \cap B) = n_3/n$ .  $\bullet$

Изменим эксперимент. Пусть событие  $A$  произошло, т. е. известно, что выбран спортсмен. Какие вероятности теперь приписать исходам? Факт наступления события  $A$  будем обозначать с помощью косой черты в записи события, например,  $\{\omega_i/A\}$  – элементарное событие  $\omega_i$  при условии, что событие  $A$  наступило.

Ясно, что если событие  $A$  наступило, то любой исход в модифицированном эксперименте должен быть благоприятным для  $A$ . Поэтому исход из  $\bar{A}$  произойти не может. Следовательно,  $P(\omega_i/A) = 0$  для каждого  $\omega_i \in \bar{A}$ . Кроме того, выбор любого спортсмена равновозможен, значит, для каждого  $\omega_i \in A$   $P(\omega_i/A) = 1/n_1$ . Теперь

$$P(B/A) = \sum_{\omega_i \in A \cap B} P(\omega_i/A) = \frac{n_3}{n_1}.$$

Итак, получили новое распределение вероятности на  $\Omega$ :

$$P(\omega_i/A) = \begin{cases} 1/n_1, & \omega_i \in A, \\ 0, & \omega_i \notin A, \quad i = 1, \dots, n, \end{cases}$$

которое отличается от исходного. Такое распределение естественно назвать условным распределением, а составляющие его вероятности – условными

вероятностями. При этом условные вероятности можно выразить через безусловные:

$$\frac{n_3}{n_1} = \frac{n_3/n}{n_1/n} = \frac{P(A \cap B)}{P(A)}$$

или

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

Приведенные в примере соображения лежат в основе формального определения условной вероятности события.

**Определение 1.12.** Пусть  $\Omega$  – пространство элементарных событий с вероятностным распределением  $P$ . Пусть  $A$  и  $B$  – некоторые события на  $\Omega$ , причем  $P(A) > 0$ . Тогда число

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad (1.9)$$

называется *условной вероятностью события  $B$  при условии, что событие  $A$  наступило* (коротко говорят „*вероятностью  $B$  при условии  $A$* “).

Формула (1.9) задает правило  $P(\cdot/A) : B \rightarrow P(B/A)$ , согласно которому каждому событию  $B$  ставится в соответствие его условная вероятность. Это правило называется *условным распределением вероятностей на  $\Omega$* .

Заметим, что определение условной вероятности (1.9) является формальным и не использует какие-либо дополнительные рассуждения, но, как показывает рассмотренный пример, такое определение согласуется со здравым смыслом.

Перепишем определение (1.9) в виде

$$P(A \cap B) = P(A)P(B/A). \quad (1.10)$$

Эту формулу традиционно называют *теоремой умножения*, хотя доказывать здесь ничего не нужно, лишь бы  $P(A)$  была больше нуля. Теорема умножения позволяет записать вероятность произведения событий через произведение безусловной и условной вероятностей этих событий.

Формула (1.10) допускает простые обобщения:

$$P(A \cap B \cap C) = P(A)P(B/A)P(C/A \cap B), \quad (1.11)$$

если  $P(A \cap B) > 0$ ;

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2/A_1) \dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1}),$$

если  $P(A_1 \cap \dots \cap A_{n-1}) > 0$ .

Докажем, например, (1.11):

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

Теорема умножения играет важную роль при определении вероятностей на некоторых типах пространств элементарных событий.

**Пример.** В каждой из двух урн имеется по 2 белых и по 3 черных шара. Из первой урны наудачу достается один шар и перекладывается во вторую урну, а затем из второй урны наудачу извлекается один шар и перекладывается в первую. Найти вероятность того, что состав шаров в урнах не изменится.

В данной задаче имеем дело со сложным экспериментом, состоящим из двух частей. Пусть исходы первой части эксперимента, которые состоят в том, что из первой урны вынут либо белый, либо черный шар, составляют множество  $\Omega_1 = \{\delta_1, r_1\}$  с вероятностью распределения  $P_1(\delta_1) = \frac{2}{5}$ ,  $P_1(r_1) = \frac{3}{5}$ . Исходы второй части эксперимента обозначим как  $\Omega_2 = \{\delta_2, r_2\}$ . Для задания вероятностного распределения на  $\Omega_2$  вспомним, что оно будет определяться исходами первой части: если во вторую урну переложили белый шар из первой урны, то  $P_2(\delta_2/\delta_1) = 1/2$ ,  $P_2(r_2/\delta_1) = 1/2$ ; если же во вторую урну переложили черный шар, то  $P_2(\delta_2/r_1) = 1/3$ ,  $P_2(r_2/r_1) = 2/3$ .

Запишем теперь множество исходов  $\Omega$  полного эксперимента в виде декартова произведения множеств  $\Omega_1$  и  $\Omega_2$ :  $\Omega = \Omega_1 \Omega_2$ . При этом в  $\Omega$  будет 4 исхода:  $(\delta_1, \delta_2)$ ,  $(\delta_1, r_2)$ ,  $(r_1, \delta_2)$ ,  $(r_1, r_2)$ .

Пусть событие  $A_1$  – „из первой урны вынут белый шар“. Тогда  $A_1 = \{(\delta_1, \delta_2), (\delta_1, r_2)\}$  и  $P(A_1) = P_1(\delta_1) = \frac{2}{5}$ .

Пусть событие  $B$  – „из первой урны вынут черный шар“,  $B_1 = \{(r_1, \delta_2), (r_1, r_2)\}$  и  $P(B_1) = P_1(r_1) = 3/5$ .

Теперь для событий  $A_2$  – „из второй урны вынут белый шар“ и  $B_2$  – „из второй урны вынут черный шар“ можно записать, что  $P(A_2/A_1) = 1/2$ ;  $P(B_2/A_1) = 1/2$ ; а  $P(A_2/B_1) = 1/3$ ;  $P(B_2/B_1) = 2/3$ .

На основании теоремы умножения запишем:

$$P((\delta_1, \delta_2)) = P(A_1 \cap A_2) = P(A_1)P(A_2/A_1) = 2/5 \cdot 1/2 = 1/5;$$

$$P((\delta_1, r_2)) = P(A_1 \cap B_2) = P(A_1)P(B_2/A_1) = 2/5 \cdot 1/2 = 1/5;$$

$$P((r_1, \delta_2)) = P(B_1 \cap A_2) = P(B_1)P(A_2/B_1) = 3/5 \cdot 1/3 = 1/5;$$

$$P((r_1, r_2)) = P(B_1 \cap B_2) = P(B_1)P(B_2/B_1) = 3/5 \cdot 2/3 = 2/5.$$

Таким образом, задано вероятностное распределение на  $\Omega$  для сложного эксперимента. Теперь не составляет труда найти вероятность события  $C$  – „состав шаров в урнах не изменился“:

$$P(C) = P(\{(\delta_1, \delta_2), (r_1, r_2)\}) = 1/5 + 2/5 = 3/5. \bullet$$

## 1.6. Формула полной вероятности. Формула Байеса

Пусть  $A$  – произвольное событие, а события  $H_1, \dots, H_n$  попарно несовместны ( $H_i \cap H_j = \emptyset$  при  $i \neq j$ ,  $i, j = 1, 2, \dots, n$ ), причем  $P(H_i) > 0$ . Пусть  $A \subset H_1 \cup H_2 \cup \dots \cup H_n$ . Тогда справедлива следующая формула, которую называют формулой полной вероятности:

$$P(A) = \sum_{i=1}^n P(H_i)P(A/H_i). \quad (1.12)$$

Для доказательства (1.12) отметим, что событие  $A$  можно представить в виде суммы попарно несовместных событий:

$$A = (A \cap H_1) \cup (A \cap H_2) \cup \dots \cup (A \cap H_n).$$

Тогда, в соответствии с теоремой сложения и теоремой умножения, получим:

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

**Пример.** Три автоматические линии производят микросхемы, причем первая производит 2 % всей продукции, вторая – 30 %, третья – 50 %. Доля брака в продукции первой линии составляет 1 %, второй линии – 2 % и третьей линии – 5 %. Какова вероятность того, что случайно выбранная микросхема дефектна?

Попробуем решить задачу стандартным образом, начиная с построения пространства элементарных событий  $\Omega$ .

Естественно считать исходом выбор конкретной микросхемы из всего выпуска. Обозначим мощность  $\Omega$  через  $|\Omega|$ . Так как микросхема извлекается наудачу, то все исходы равновозможны и  $P(\omega_i) = 1/|\Omega|$ .

Введем на  $\Omega$  события  $H_i$ ,  $i = 1, 2, 3$  – „микросхема изготовлена  $i$ -й линией“. Очевидно,  $\Omega = H_1 \cup H_2 \cup H_3$ . По условию задачи

$$|H_1| = 0.2|\Omega|; \quad |H_2| = 0.3|\Omega|; \quad |H_3| = 0.5|\Omega|.$$

Пусть событие  $A$  – „микросхема дефектная“. Тогда, в соответствии со схемой решения задач на классическую вероятность,  $P(A) = |A|/|\Omega|$ . Очевидно, что событие  $A$  можно представить в виде суммы попарно несовместных событий:  $A = (A \cap H_1) \cup (A \cap H_2) \cup (A \cap H_3)$ . Так как  $|A \cap H_1| = 0.01 \cdot 0.2|\Omega|$ ;  $|A \cap H_2| = 0.02 \cdot 0.3|\Omega|$ ;  $|A \cap H_3| = 0.05 \cdot 0.5|\Omega|$ , то

$$P(A) = \frac{|A \cap H_1| + |A \cap H_2| + |A \cap H_3|}{|\Omega|} = 0.033.$$

Теперь решим задачу с использованием формулы полной вероятности. Можно записать:

$$P(A) = \sum_{i=1}^3 P(H_i)P(A/H_i).$$

Но для вероятностей  $P(H_i)$  и условных вероятностей  $P(A/H_i)$ ,  $i = 1, 2, 3$ , из условия задачи известны их значения:  $P(H_1) = 0.2$ ;  $P(H_2) = 0.3$ ;  $P(H_3) = 0.5$ ;  $P(A/H_1) = 0.01$ ;  $P(A/H_2) = 0.02$ ;  $P(A/H_3) = 0.05$ . Значит,  $P(A) = 0.2 \cdot 0.01 + 0.3 \cdot 0.02 + 0.5 \cdot 0.05 = 0.033$ . •

Результат, естественно, получился тот же. Однако, используя формулу полной вероятности, получили ответ, минуя этап построения пространства элементарных событий.

Запишем определение условной вероятности события  $H_i$  при условии, что событие  $A$  наступило:

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

Заменяя вероятность  $P(A)$  через формулу полной вероятности, получим:

$$P(H_i/A) = \frac{P(H_i)P(A/H_i)}{\sum_{k=1}^n P(H_k)P(A/H_k)}. \quad (1.13)$$

Формула (1.13) называется формулой Байеса. Ее можно интерпретировать следующим образом. Пусть перед проведением некоторого научного эксперимента имеется  $n$  гипотез  $H_1, H_2, \dots, H_n$  о природе изучаемого явления, которым соответствуют вероятности  $P(H_1), P(H_2), \dots, P(H_n)$ . Эти вероятности называются априорными.

Затем проводим научный эксперимент, в результате которого может наступить событие  $A$ . Если событие  $A$  наступило, то переоцениваем веру в гипотезы, заменяя априорные вероятности  $P(H_i)$  на апостериорные вероятности  $P(H_i/A)$ .

Для иллюстрации применения формулы Байеса рассмотрим предыдущий пример, в условиях которого решим следующую задачу: какова вероятность того, что выбранная микросхема, оказавшаяся дефектной, произведена первой, второй или третьей линией?

Другими словами, требуется найти апостериорные вероятности  $P(H_1/A)$ ,  $P(H_2/A)$ ,  $P(H_3/A)$ . В соответствии с формулой Байеса

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{0.2 \cdot 0.01}{0.033} \simeq 0.061;$$

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{0.3 \cdot 0.02}{0.033} \simeq 0.182;$$

$$P(H_3/A) = \frac{P(H_3)P(A/H_3)}{P(A)} = \frac{0.5 \cdot 0.05}{0.033} \simeq 0.757;$$

Найденные апостериорные вероятности отличаются от априорных вероятностей событий  $H_1$ ,  $H_2$  и  $H_3$ .

### 1.7. Независимые события

Естественно считать, что событие  $A$  не зависит от события  $B$ , если условная вероятность события  $A$  при условии, что событие  $B$  наступило, равна безусловной вероятности события  $A$ , т. е.  $P(A/B) = P(A)$ . Из определения условной вероятности в этом случае получаем, что  $P(A \cap B) = P(A)P(B)$ . Следовательно, если  $A$  не зависит от  $B$ , то и  $B$  не зависит от  $A$ , поскольку последнее равенство симметрично.

**Определение 1.13.** События  $A$  и  $B$  называются независимыми, если выполняется равенство

$$P(A \cup B) = P(A)P(B). \quad (1.14)$$

Смысл этого формального определения заключается в том, что если произошло одно из двух независимых событий, то это никак не влияет на вероятность другого события. Но в таком случае, если первое событие не произошло, то это также не должно влиять на вероятность второго. Этот результат сформулируем в виде следующей теоремы.

**Теорема 1.4.** Если события  $A$  и  $B$  независимы, то события  $\bar{A}$  и  $B$  также независимы.

**Доказательство.** Очевидно, что  $P(B) = P(A \cap B) + P(\bar{A} \cap B)$ . Тогда

$$\begin{aligned} P(\bar{A} \cap B) &= P(B) - P(A \cap B) = P(B) - P(A)P(B) = \\ &= (1 - P(A))P(B) = P(\bar{A})P(B). \end{aligned}$$

Теорема доказана. ■

**Следствие.** Если события  $A$  и  $B$  независимы, то события  $\bar{A}$  и  $\bar{B}$  также независимы.

Если для событий  $A$  и  $B$  не выполнено условие (1.14), то они называются зависимыми.

Определение независимости событий можно обобщить для большего, нежели 2, числа событий.

**Определение 1.14.** События  $A_1, A_2, \dots, A_n$  называются независимыми в совокупности, если для любых  $k$  из них ( $k \leq n$ ) выполняется

условие

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j}). \quad (1.15)$$

Если соотношение (1.15) выполняется только при  $k = 2$ , то события называются попарно независимыми.

Следующий пример показывает, что попарная независимость не обеспечивает независимости в совокупности.

**Пример (Бернштейна).** Три из четырех граней правильного тетраэдра окрашены в красный, синий и зеленый цвета. Четвертая грань окрашена во все 3 цвета. Эксперимент состоит в подбрасывании тетраэдра. Исход – падение тетраэдра на определенную грань. Таким образом,  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} = \{\text{к, с, з, ксз}\}$ ,  $P(\omega_i) = 1/4$ . Рассмотрим события:  $A$  – „тетраэдр упал на грань, содержащую красный цвет“,  $A = \{\omega_1, \omega_4\}$ ,  $P(A) = 1/2$ ;  $B$  – „тетраэдр упал на грань, содержащую синий цвет“,  $B = \{\omega_2, \omega_4\}$ ,  $P(B) = 1/2$ ;  $C$  – „тетраэдр упал на грань, содержащую зеленый цвет“,  $C = \{\omega_3, \omega_4\}$ ,  $P(C) = 1/2$ .

События  $A, B, C$  являются попарно независимыми:

$$A \cap B = A \cap C = B \cap C = \{\omega_4\}, \quad P(A \cap B) = P(A \cap C) = P(B \cap C) = 1/4,$$

$$P(A)P(B) = P(A)P(C) = P(B)P(C) = 1/4.$$

Однако события  $A, B, C$  не являются независимыми в совокупности:

$$P(A \cap B \cap C) = 1/4; \quad P(A)P(B)P(C) = 1/8.$$

Условие (1.15) не выполнено. •

## 1.8. Независимые испытания

С помощью определения независимых событий можно строить модели сложных экспериментов, состоящих из нескольких независимых частей. Рассмотрим, для простоты, эксперимент, состоящий из двух частей. Первая часть описывается пространством элементарных событий  $\Omega_1$  с вероятностным распределением  $P_1$ . Вторая часть описывается пространством  $\Omega_2$  с распределением  $P_2$ . Пусть событие  $A$  связано с первой частью эксперимента,  $A \subset \Omega_1$ , а  $B$  связано со второй частью, т. е.  $B \subset \Omega_2$ . Требуется построить модель всего эксперимента так, чтобы события  $\tilde{A}$  – „в первой части произошло событие  $A$ “ и  $\tilde{B}$  – „во второй части произошло событие  $B$ “ были независимыми.

Если модель сложного эксперимента (т. е.  $\Omega$  и  $P$ ) построена так, что для любых событий  $A \subset \Omega_1, B \subset \Omega_2$  события  $\tilde{A}$  и  $\tilde{B}$  независимы, то говорят, что эксперимент состоит из двух независимых испытаний.

**Теорема 1.5.** Пусть  $\Omega = \Omega_1 \times \Omega_2 = \{(\omega_i, \kappa_j) \mid \omega_i \in \Omega_1, \kappa_j \in \Omega_2\}$ .  
Определим  $P((\omega_i, \kappa_j)) = P_1(\omega_i)P_2(\kappa_j)$ . Тогда:

- 1)  $P(\cdot)$  есть распределение вероятностей на  $\Omega$ ;
- 2) события  $\tilde{A} = A \times \Omega_2 = \{(\omega_i, \kappa_j) \mid \omega_i \in A, \kappa_j \in \Omega_2\}$  и  $\tilde{B} = \Omega_1 \times B = \{(\omega_i, \kappa_j) \mid \omega_i \in \Omega_1, \kappa_j \in B\}$  имеют вероятности  $P(\tilde{A}) = P_1(A)$  и  $P(\tilde{B}) = P_2(B)$ ;
- 3) события  $\tilde{A}$  и  $\tilde{B}$  независимы.

**Доказательство.**

1) Согласно определению вероятности события  $0 \leq P_1(\omega_i)P_2(\kappa_j) \leq 1$ , значит  $0 \leq P((\omega_i, \kappa_j)) \leq 1$ . Тогда

$$\sum_{(\omega_i, \kappa_j) \in \Omega} P((\omega_i, \kappa_j)) = \sum_{\omega_i \in \Omega_1} \sum_{\kappa_j \in \Omega_2} P_1(\omega_i)P_2(\kappa_j) = \sum_{\omega_i \in \Omega_1} P_1(\omega_i) \sum_{\kappa_j \in \Omega_2} P_2(\kappa_j) = 1,$$

т. е. действительно  $P(\cdot)$  задает на  $\Omega$  распределение вероятностей.

2) Для вероятности события  $\tilde{A}$  можно записать:

$$P(\tilde{A}) = \sum_{(\omega_i, \kappa_j) \in \tilde{A}} P((\omega_i, \kappa_j)) = \sum_{\omega_i \in A} P_1(\omega_i) \sum_{\kappa_j \in \Omega_2} P_2(\kappa_j) = P_1(A) \cdot 1 = P_1(A).$$

Аналогично доказывается и для вероятности  $P(\tilde{B})$  события  $\tilde{B}$ .

3) Запишем вероятность произведения событий  $\tilde{A}$  и  $\tilde{B}$ :

$$\begin{aligned} P(\tilde{A} \cap \tilde{B}) &= \sum_{(\omega_i, \kappa_j) \in \tilde{A} \cap \tilde{B}} P((\omega_i, \kappa_j)) = \sum_{\substack{\omega_i \in A \\ \kappa_j \in B}} \sum_{\kappa_j \in B} P_1(\omega_i)P_2(\kappa_j) = \\ &= P_1(A)P_2(B) = P(\tilde{A})P(\tilde{B}). \end{aligned}$$

Теорема доказана. ■

Таким образом, эксперимент, состоящий из двух независимых испытаний, описывается моделью  $(\Omega, P)$ , где  $\Omega = \Omega_1 \times \Omega_2$ ,  $P(\cdot) = P_1(\cdot)P_2(\cdot)$ .

Доказанная теорема допускает обобщения. Для описания эксперимента, состоящего из  $n$  независимых испытаний  $(\Omega_i, P_i)$ , нужно взять  $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$  и  $P((\omega_i, \kappa_j, \dots, \delta_k)) = P_1(\omega_i)P_2(\kappa_j) \dots P_n(\delta_k)$ . Тогда для любых событий  $A_1 \subset \Omega_1$ ,  $A_2 \subset \Omega_2$ ,  $\dots$ ,  $A_n \subset \Omega_n$  будет справедливо  $P(\tilde{A}_1) = P_1(A_1)$ ,  $P(\tilde{A}_2) = P_2(A_2)$ ,  $\dots$ ,  $P(\tilde{A}_n) = P_n(A_n)$ , и группа событий  $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_n$  будет независима в совокупности.

**Пример.** Стрелок попадает в мишень с вероятностью  $p$ . Он стреляет до первого попадания. Описать эксперимент и найти вероятность события  $A$  – „стрелок произвел не менее двух выстрелов“.

Рассмотрим множество  $\Omega_1 = \{1, 0\}$ . Это множество включает исходы части эксперимента, состоящей из одного выстрела. „1“ означает попадание

с вероятностью  $p$  и „0“ – промах с вероятностью  $q = 1 - p$ . Множество исходов  $\Omega$  оказывается счетным:

$\Omega$					
$\omega_1$	1				
$\omega_2$	0	1			
$\omega_3$	0	0	1		
$\dots$	0	0	...	0	1
$\omega_n$	0	0	...	...	0
$\omega_\infty$	0	0	...	...	0

При этом  $\omega_2 \in \Omega_1^2 = \Omega_1 \times \Omega_1$  с вероятностным распределением  $P_2(\cdot) = P_1(\cdot)P_1(\cdot)$ , т. е.  $P_2(\omega_2) = qp$ . Следовательно, полагаем  $P(\omega_2) = qp$ . Продолжая рассуждения далее, получим, что  $\omega_n \in \Omega^n = \Omega_1 \times \Omega_1 \times \dots \times \Omega_1$  с распределением  $P_n(\cdot) = P_1(\cdot) \dots P_1(\cdot)$ . При этом  $P_n(\omega_n) = q^{n-1}p$ . Значит,  $P(\omega_n) = q^{n-1}p$ .

Итак, каждому исходу сопоставлена его вероятность  $P(\omega_n)$ , пользуясь тем, что  $\omega_n \in \Omega^n$  и имеет вероятность  $P_n(\omega_n)$ . Осталось найти вероятность исхода  $\omega_\infty$ . Запишем, что  $P(\omega_\infty) + \sum_{n=1}^{+\infty} (P(\omega_n)) = 1$ . Но

$$\sum_{n=1}^{+\infty} P(\omega_n) = \sum_{n=1}^{\infty} q^{n-1}p = \frac{p}{1-q} = \frac{p}{p} = 1.$$

Значит,  $P(\omega_\infty) = 0$ . Вероятностное распределение построено.

Найдем теперь вероятность события  $A$ . Проще найти вероятность противоположного события  $\bar{A}$  – „стрелок произвел менее двух выстрелов“. Это значит, что  $\bar{A} = \{\omega_1\}$ . Тогда  $P(\bar{A}) = p$  и  $P(A) = 1 - p = q$ . •

## 1.9. Геометрические вероятности

Рассмотрим следующую задачу: на телефонной линии длиной 10 км произошел обрыв, равновозможный в любой точке. Определить вероятность того, что обрыв произошел между пятым и шестым километрами.

Попытаемся решить эту задачу стандартным способом. В качестве исхода естественно принять координату точки обрыва. Множество исходов при этом получается несчетным, что не укладывается в модель с дискретным пространством элементарных событий. Не обращая на это внимания, зададим вероятностное распределение. Любая положительная вероятность, приписанная элементарному событию, приводит к тому, что в сумме все вероятности несчетного количества элементарных событий равны бесконечности. Единственный вариант – задать вероятность каждого исхода равной

нулю. Решить задачу, строя элементарную модель случайного эксперимента, не удалось. Это означает, что модель случайного эксперимента, состоящая из дискретного пространства элементарных событий и заданного на нем вероятностного распределения, требует дальнейшего расширения, позволяющего учесть несчетные множества исходов.

Вернемся к задаче и попытаемся решить ее без использования элементарной модели, основываясь на очевидных фактах. Обратим внимание на то, что в силу равновозможности исходов вероятность обрыва на любом промежутке  $[a, b]$  пропорциональна длине этого промежутка. Тогда естественно принять вероятность события  $A$  – „обрыв произошел на промежутке  $[a, b]$ “ – равной  $P(A) = \frac{b-a}{10}$  при любых  $a$  и  $b$ , таких, что  $0 \leq a \leq b \leq 10$ .

Значит, вероятность обрыва между пятым и шестым километрами линии равна 0.1. Задача решена.

Назначенные вероятности носят название „геометрические вероятности“. Вообще, геометрическими вероятностями называют вероятности событий, пропорциональные некоторой мере этих событий. В качестве такой меры могут выступать длина, площадь, объем и другие физические величины.

## 2. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

При нахождении вероятностей случайных событий возникают задачи, связанные с выбором, расположением и подсчетом количеств элементов некоторых множеств. Подобные задачи и методы их решения рассматриваются и изучаются в разделе математики, который называется комбинаторика или комбинаторный анализ. Рассмотрим основные понятия комбинаторики.

### 2.1. Основная формула комбинаторики

Пусть имеется 2 конечных множества  $A = \{a_1, \dots, a_m\}$  и  $B = \{b_1, \dots, b_n\}$ . Из элементов этих множеств будем составлять упорядоченные пары вида  $(a_i, b_j)$  ( $i = 1, \dots, m$ ,  $j = 1, \dots, n$ ). Первым элементом пары является элемент первого множества, вторым – элемент второго множества. Подсчитаем общее количество таких пар. Для этого будем последовательно выбирать элементы из первого множества и для каждого отобранного элемента – элементы из второго множества. Выберем сначала первый элемент  $a_1$  из множества  $A$ , а затем – элементы  $b_1, \dots, b_n$  из множества  $B$  и составим пары  $(a_1, b_1), \dots, (a_1, b_n)$ . Общее число таких пар равно  $n$ . Вместо

элемента  $a_1$  можно взять любой другой элемент множества  $A$ . Поскольку множество  $A$  содержит  $m$  элементов, то число всех упорядоченных пар  $(a_i, b_1) \dots (a_i, b_n)$  ( $i = 1, \dots, m$ ) будет равно  $mn$ .

Рассмотрим теперь  $k$  конечных множеств. Первое множество  $A_1 = \{a_1^1, \dots, a_{n_1}^1\}$  содержит  $n_1$  элементов, второе  $A_2 = \{a_1^2, \dots, a_{n_2}^2\}$  –  $n_2$  элементов, а последнее  $A_k = \{a_1^k, \dots, a_{n_k}^k\}$  –  $n_k$  элементов. Используя элементы этих множеств, будем составлять упорядоченные наборы длиной  $k$  вида  $(a_i^1, a_j^2, \dots, a_l^k)$  ( $i = 1, \dots, n_1$ ,  $j = 1, \dots, n_2$ ,  $l = 1, \dots, n_k$ ). Первым элементом набора является элемент первого множества, вторым – элемент второго, последним – элемент  $k$ -го множества. Для подсчета общего количества таких упорядоченных наборов будем выбирать из каждого множества по одному элементу и последовательно составлять наборы. Сначала рассмотрим наборы из одного элемента. Они могут составляться только из элементов множества  $A_1$ . Их количество равно  $n_1$  – совпадает с числом элементов этого множества. Наборы из двух элементов  $(a_i^1, a_j^2)$  ( $i = 1, \dots, n_1$ ,  $j = 1, \dots, n_2$ ) составляются из элементов множеств  $A_1$  и  $A_2$ . Их общее число равно  $n_1 n_2$ . Если к каждой паре элементов из элементов множеств  $A_1$  и  $A_2$  последовательно добавлять по одному элементу из множества  $A_3$ , то получатся упорядоченные тройки элементов. Поскольку множество  $A_3$  содержит  $n_3$  элементов, то количество упорядоченных троек будет равно  $n_1 n_2 n_3$ . Продолжая эти рассуждения составим упорядоченные наборы из элементов  $k$  множеств. Общее число наборов равно произведению количеств элементов этих множеств:

$$n_1 n_2 \dots n_k.$$

Полученное выражение называется основной формулой комбинаторики или правилом умножения. Эту формулу используют для вывода многих формул комбинаторики. При этом в некоторых случаях множества  $A_2, A_3, \dots$  с самого начала не задаются. Множество  $A_2$ , содержащее  $n_2$  элементов, становится известным только после того как выбирается элемент множества  $A_1$  и определяется этим элементом. Множество  $A_3$ , содержащее  $n_3$  элементов, задается после выбора элементов из множеств  $A_1$  и  $A_2$  по одному и определяется выбранными элементами и т. д.

**Пример.** Меню содержит 5 различных салатов, 3 супа, 4 вторых и 3 напитка. Сколько различных обедов из четырех блюд можно заказать?

Общее число различных обедов:  $5 \cdot 3 \cdot 4 \cdot 3 = 180$ . •

**Пример.** Сколько существует различных семизначных телефонных номеров, если при составлении номера телефона могут использоваться любые цифры и номер не может начинаться с нуля?

Цифры – это целые числа от нуля до девяти. Общее количество цифр равно десяти. Первая цифра номера телефона может быть любой, кроме

нуля, поэтому она выбирается из девяти цифр. Все остальные цифры номера телефона выбираются из десяти цифр, поэтому число семизначных номеров телефонов будет равно:  $9 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 9 \cdot 10^6$ . •

**Пример.** Сколько существует трехзначных нечетных чисел с различными цифрами?

Поскольку число нечетное, то оно оканчивается одной из пяти цифр: 1, 3, 5, 7, 9. Все цифры числа различны и число не может начинаться с нуля, поэтому первая цифра числа выбирается из множества, содержащего 8 цифр (из всего множества цифр исключают последнюю цифру трехзначного числа и цифру 0). Вторая цифра не может совпадать с первой и третьей, т. е. выбирается из множества, содержащего 8 цифр. Тогда общее число трехзначных нечетных чисел с различными цифрами равно:  $8 \cdot 8 \cdot 5 = 320$ .

•

## 2.2. Перестановки

Пусть имеется некоторое конечное множество  $A = \{a_1, \dots, a_n\}$ . Используя элементы этого множества будем составлять наборы, содержащие  $n$  элементов.

**Определение 2.1.** Любой упорядоченный набор из  $n$  элементов данного множества называется *перестановкой*.

Общее количество перестановок обозначается  $P_n$ . Подсчитаем, сколько существует различных перестановок. Будем последовательно составлять всевозможные перестановки, выбирая элементы из множества  $A$ . Первый элемент перестановки выбирается из всего множества  $A$ , содержащего  $n$  элементов. Этот элемент фиксируется, отбирается и не возвращается в множество  $A$ . Второй элемент перестановки будет выбираться из множества, содержащего  $(n - 1)$  элемент. Второй элемент также фиксируется, отбирается и в множество не возвращается, поэтому третий элемент перестановки будет выбираться из множества, содержащего  $(n - 2)$  элемента. Каждый следующий элемент перестановки будет выбираться из множества, объем которого на один элемент меньше предыдущего. Последний элемент будет выбираться из множества, содержащего один элемент. Тогда, согласно основной формуле комбинаторики, общее число перестановок элементов множества  $A$  находится по формуле

$$P_n = n(n - 1)(n - 2) \dots 1 = n!.$$

**Пример.** Сколько различных трехцветных флагов одинакового размера с продольными полосами можно сделать, используя ленты белого, синего и красного цветов?

Число различных флагов равно числу различных перестановок трех лент:  $P_3 = 3! = 6$ . •

**Пример.** Сколько четырехзначных чисел с различными цифрами можно составить, используя цифры 0,2,4,6?

Число различных перестановок четырех заданных чисел равно  $4!$ , но четырехзначные числа не могут начинаться с нуля, поэтому следует исключить наборы цифр, которые начинаются с нуля. Искомое число четырехзначных чисел равно  $P_4 - P_3 = 4! - 3! = 18$ . •

Рассмотрим теперь некоторый набор

$$B = (\underbrace{a_1, \dots, a_1}_{n_1}, \underbrace{a_2, \dots, a_2}_{n_2}, \dots, \underbrace{a_k, \dots, a_k}_{n_k}),$$

который содержит  $n_1$  элементов  $a_1$ ,  $n_2$  элементов  $a_2$ ,  $n_k$  элементов  $a_k$ , причем  $a_1, a_2, \dots, a_k$  различны. В этом наборе всего  $n = n_1 + n_2 + \dots + n_k$  элементов. Перестановки элементов такого набора называются перестановками с повторениями. Число различных перестановок с повторениями обозначается  $P_n(n_1, n_2, \dots, n_k)$ . Найдем это число. Общее число перестановок всех элементов  $B$  равно  $n!$ . Поскольку среди элементов  $B$  есть одинаковые, то любая фиксированная перестановка не изменится, если менять местами одинаковые элементы. Согласно основной формуле комбинаторики число всех перестановок одинаковых элементов между собой будет равно  $n_1!n_2! \dots n_k!$ , поэтому выполняется равенство

$$n! = P_n(n_1, n_2, \dots, n_k)n_1!n_2! \dots n_k!.$$

Значит, общее число различных перестановок элементов  $B$  находится по правилу

$$P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1!n_2! \dots n_k!}.$$

**Пример.** Сколько различных десятизначных чисел можно составить, используя пять единиц, три двойки и две семерки?

Общее число используемых цифр равно десяти, поэтому искомое количество чисел находится по правилу

$$P_{10}(5, 3, 2) = \frac{10!}{5!3!2!} = 2520. •$$

### 2.3. Размещения

Рассмотрим, как и ранее, множество  $A = \{a_1, \dots, a_n\}$ . Пусть из  $n$  элементов этого множества составляются разнообразные наборы, содержащие  $k$  элементов этого множества ( $k \leq n$ ).

**Определение 2.2.** Любой упорядоченный набор из  $k$  элементов данного множества, содержащего  $n$  элементов, называется размещением из  $n$  по  $k$ .

Общее число размещений из  $n$  по  $k$  обозначается  $A_n^k$ .

Если  $k = n$ , то число размещений  $n$  по  $n$  совпадает с числом перестановок:  $A_n^k = P_n = n!$ .

Для подсчета числа размещений из  $n$  по  $k$  так же, как и ранее, будем составлять различные упорядоченные наборы из  $k$  элементов множества  $A$ , последовательно выбирая элементы и не возвращая их. Первый элемент размещения выбирается из  $n$  элементов всего множества  $A$ , второй – из множества, содержащего  $(n - 1)$  элемент, третий – из множества, содержащего  $(n - 2)$  элемента, а значит, последний  $k$ -й элемент выбирается из множества, содержащего  $(n - k + 1)$  элемент. Тогда согласно основной формуле комбинаторики число размещений  $n$  по  $k$  определяется по правилу

$$A_n^k = n(n - 1)(n - 2) \dots (n - k + 1) = \frac{n!}{(n - k)!}.$$

**Пример.** Сколько различных трехцветных флагов одинакового размера с продольными полосами можно сделать, используя ленты семи разных цветов?

Число различных флагов равно числу различных размещений семи лент по три:  $A_7^3 = 7 \cdot 6 \cdot 5 = 210$ . •

**Пример.** Сколько способами можно распределить 4 билета на футбол и 2 билета в театр между десятью студентами?

Искомое число способов совпадает с числом упорядоченных пар  $(a_i, b_j)$ , где первый элемент пары – студенты, которым достались билеты на футбол, второй элемент пары – студенты, получившие билеты в театр. Количество элементов множества, из которого выбирается первый элемент пары, –  $A_{10}^4$ , тогда число элементов множества, из которого выбирается второй элемент пары, –  $A_6^2$ . При подсчете количеств элементов множеств, из которых выбираются элементы пары, следует учитывать, какому из выбранных студентов достался конкретный билет. Согласно основной формуле комбинаторики искомое число способов:  $A_{10}^4 \cdot A_6^2 = 151200$ . •

## 2.4. Сочетания

**Определение 2.3.** Любая неупорядоченная совокупность из  $k$  элементов множества  $A = \{a_1, \dots, a_n\}$ , содержащего  $n$  различных элементов называется сочетанием из  $n$  элементов по  $k$ .

Общее число сочетаний из  $n$  по  $k$  обозначается  $C_n^k$ .

Сочетание из  $n$  элементов по  $k$  получается, если из множества  $A$  одновременно выбираются  $k$  элементов. Для нахождения  $C_n^k$  рассмотрим сначала произвольное размещение из  $n$  по  $k$ . Оно содержит  $k$  элементов множества  $A$ . При подсчете  $A_n^k$  – общего числа размещений из  $n$  по  $k$  – учитывается порядок расположения выбранных  $k$  элементов. Общее число перестановок отобранных  $k$  элементов множества  $A$  равно  $k!$ . Значит,  $A_n^k = C_n^k k!$ , поэтому

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

Отметим некоторые свойства  $C_n^k$ .

**Теорема 2.1.** Для числа сочетаний из  $n$  по  $k$  справедливы равенства:

1.  $C_n^0 = 1$ .
2.  $C_n^k = C_n^{n-k}$ .
3.  $C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$ .

**Доказательство.**

1. Поскольку  $0! = 1$ , то  $C_n^0 = \frac{n!}{(n-0)!0!} = 1$ .
2. Найдем  $C_n^{n-k} = \frac{n!}{(n-(n-k))!(n-k)!} = \frac{n!}{k!(n-k)!}$ . Значит выполняется равенство  $C_n^k = C_n^{n-k}$ .
3. Покажем, что выполняется цепочка равенств:

$$\begin{aligned} C_n^k + C_n^{k+1} &= \frac{n!}{(n-k)!k!} + \frac{n!}{(n-k-1)!(k+1)!} = \frac{n!(k+1) + n!(n-k)}{(n-k)!(k+1)!} = \\ &= \frac{n!(n+1)}{(n-k)!(k+1)!} = \frac{(n+1)!}{(n+1-(k+1))!(k+1)!} = C_{n+1}^{k+1}. \blacksquare \end{aligned}$$

**Пример.** Лаборант хочет поставить на две полки 4 амперметра и 6 вольтметров. Сколькими способами он может это сделать так, чтобы на каждой полке стояло равное количество приборов, и хотя бы один из них – вольтметр. Приборы на полке стоят в беспорядке.

Найдем сначала общее число способов поставить по 5 приборов на каждую полку. Это число совпадает с числом способов выбора пяти приборов из десяти:  $C_{10}^5$ . Затем найдем в скольких случаях на первой полке будут стоять только амперметры:  $C_6^5$ . В стольких же случаях вторая полка будет заполнена только амперметрами. Искомое количество способов расстановки найдем как разность:  $C_{10}^5 - 2C_6^5 = 240$ . •

**Пример.** В кошельке 5 десятирублевых и 9 двухрублевых монет (монеты считаются различными). Сколькими способами можно набрать 24 р., используя эти монеты?

Возможны только 2 варианта:  $24 = 10 \cdot 2 + 2 \cdot 2$  (выбираются две десятирублевые и две двухрублевые монеты) или  $24 = 10 + 2 \cdot 7$  (выбирается одна десятирублевая и 7 двухрублевых монет). Поскольку порядок выбора монет не важен, то искомое число способов находится по правилу:  $C_5^2 \cdot C_9^2 + C_5^1 \cdot C_9^7 = 540$ . •

## 2.5. Размещения с повторениями

Предположим, что из  $n$  различных элементов множества  $A$  последовательно выбирают  $k$  элементов. При этом каждый отобранный элемент фиксируется и возвращается в исходное множество перед выбором следующего. Каждый новый элемент выбирается из всего множества  $A$ . Подобный упорядоченный выбор  $k$  элементов из  $n$  элементов множества  $A$  называют размещениями с повторениями. При этом  $k$  может быть больше  $n$ . Общее число различных размещений с повторениями обозначается  $\bar{A}_n^k$ . Поскольку каждый раз новый элемент выбирается из всего множества  $A$ , содержащего  $n$  элементов, то согласно основной формуле комбинаторики число  $\bar{A}_n^k$  находится по правилу

$$\bar{A}_n^k = n^k.$$

**Пример.** Сколькими способами можно разложить 7 различных монет по четырем карманам?

Первая монета может попасть в любой из четырех карманов, для нее есть 4 варианта выбора. Для второй и каждой следующей монеты также будет 4 варианта выбора карманов, значит, общее количество способов разложить 7 различных монет по четырем карманам будет равно:  $4^7 = 16\,384$ . •

## 3. АКСИОМАТИКА ТЕОРИИ ВЕРОЯТНОСТЕЙ А. Н. КОЛМОГОРОВА

### 3.1. Вероятностное пространство

Как было показано в 1.9, элементарные вероятностные модели, использующие дискретное пространство элементарных событий, не могут описывать случайные эксперименты с несчетным множеством исходов. Универсальная вероятностная модель на основе теоретико-множественного аксиоматического подхода была предложена академиком А. Н. Колмогоровым в 20-х гг. XX в. и получила название „аксиоматика Колмогорова“. С ее появлением теория вероятностей превратилась в строгую математическую

дисциплину. В настоящее время аксиоматика Колмогорова общепринята в мире как фундаментальный принцип построения вероятностных моделей случайных экспериментов. Познакомимся с аксиоматикой Колмогорова подробнее.

Как и в элементарной теории вероятностей, множество всех взаимоисключающих исходов эксперимента будем называть множеством элементарных событий  $\Omega$ . Только теперь  $\Omega$  может быть конечным, счетным или несчетным.

В элементарной теории любое подмножество множества  $\Omega$  называлось событием. В аксиоматической же теории вероятностей событиями являются не все, а лишь некоторые подмножества множества  $\Omega$ , составляющие так называемую сигма-алгебру событий  $F$ .

**Определение 3.1.** Пусть  $F$  – некоторая система подмножеств множества  $\Omega$ ,  $F = \{A, B, \dots\}$ ,  $A \subset \Omega$ ,  $B \subset \Omega, \dots$  Множество  $F$  называется алгеброй, если выполнены следующие условия (аксиомы):

1.  $\Omega \in F$ .
2. Если  $A \in F$ , то  $\bar{A} \in F$ .
3. Если  $A \in F$  и  $B \in F$ , то  $A \cup B \in F$  и  $A \cap B \in F$ .

Если аксиома 3 выполняется в счетном варианте, т. е. если  $A_1, A_2, \dots \in F$ ,  $\bigcup_{i=1}^{+\infty} A_i \in F$  и  $\bigcap_{i=1}^{+\infty} A_i \in F$ , то алгебра  $F$  называется  $\sigma$ -алгеброй (сигма-алгеброй).

**Пример.** Пусть  $\Omega$  – конечное множество,  $F_0 = \{\emptyset, \Omega\}$  – самая бедная алгебра на  $\Omega$ . Для любого множества  $A \subset \Omega$  можно построить наименьшую алгебру, содержащую  $A$ :

$$F_A = \{\Omega, \emptyset, A, \bar{A}\}.$$

Вообще, для любого непустого набора подмножеств множества  $\Omega$  можно построить минимальную  $\sigma$ -алгебру, содержащую данные подмножества. •

**Пример.** Пусть  $\Omega = \mathbb{R}$ ,  $\mathfrak{M}$  – набор всех промежутков в  $\mathbb{R}$ :  $\mathfrak{M} = \{ ]a, b[; ]a, b]; [a, b[; [a, b]; -\infty \leq a \leq b \leq +\infty\}$ .

Легко видеть, что  $\mathfrak{M}$  не есть алгебра (объединение двух промежутков промежутком, в общем случае, не является). Но если дополнить  $\mathfrak{M}$  всевозможными объединениями промежутков, то получится алгебра. Наименьшая  $\sigma$ -алгебра, порожденная множеством  $\mathfrak{M}$ , называется борелевской  $\sigma$ -алгеброй на прямой. •

**Определение 3.2.** Пусть на множестве элементарных событий  $\Omega$  задана  $\sigma$ -алгебра (или алгебра  $F$  в случае конечного множества событий) его подмножеств. Элементы  $\sigma$ -алгебры (алгебры  $F$ ) называются событиями на  $\Omega$ .

Основной смысл введения  $\sigma$ -алгебры  $F$  событий на  $\Omega$  состоит в том, что теперь вероятностное распределение на  $\Omega$  задается только на событиях-элементах  $F$ .

**Определение 3.3.** *Говорят, что на алгебре событий  $F$  задано вероятностное распределение  $P$ , если каждому событию  $A \in F$  поставлено в соответствие число  $P(A)$ , называемое вероятностью события  $A$ , причем выполнены следующие аксиомы:*

4.  $0 \leq P(A) \leq 1$ .
5.  $P(\Omega) = 1$ .
6. Если  $A, B \in F$  и  $A \cap B = \emptyset$ , то  $P(A \cup B) = P(A) + P(B)$ .

Если вероятностное распределение задается на  $\sigma$ -алгебре, то аксиома 6, называемая аксиомой аддитивности, формулируется в счетном варианте (счетная аддитивность).

- 6а. Если  $A_1, A_2, \dots \in F$  и  $A_i \cap A_j = \emptyset$  при  $i \neq j$ , то

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i),$$

где справа стоит сумма сходящегося положительного числового ряда.

**Определение 3.4.** *Тройка  $(\Omega, F, P)$  обычно называется вероятностным пространством.*

Вероятностное пространство служит математической моделью случайного эксперимента.

Все основные результаты, полученные в элементарной теории, включая операции над событиями, теорему сложения, условную вероятность, теорему умножения, независимость событий, формулы полной вероятности и Байеса, без каких-либо затруднений автоматически переносятся на случай вероятностного пространства.

### 3.2. Классификация вероятностных пространств

В зависимости от содержания случайного эксперимента для его описания используют три основных типа вероятностных пространств: дискретное, непрерывное и смешанное.

*Дискретное вероятностное пространство* по сути соответствует случаю дискретного пространства элементарных событий, рассмотренному в элементарной теории. Множество элементарных событий  $\Omega = \{\omega_1, \omega_2, \dots\}$  – дискретное множество, т. е. конечное или счетное. В качестве  $\sigma$ -алгебры  $F$  всегда берется  $F = 2^\Omega$  – множество всех подмножеств  $\Omega$ . Вероятностное распределение  $P$  определяется набором вероятностей элементарных

событий  $\{P(\omega_1), P(\omega_2), \dots\}$ ,  $0 \leq P(\omega_i) \leq 1$ ,  $\sum_{i=1}^{+\infty} P(\omega_i) = 1$  и правилом вычисления вероятностей любого события

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

В *непрерывном вероятностном пространстве* в качестве множества элементарных событий обычно рассматривается  $\Omega = \mathbb{R}^n$ .  $\sigma$ -алгебра событий  $F$  есть борелевская  $\sigma$ -алгебра  $\mathcal{B}_n$ , порожденная  $n$ -мерными клетками  $\{(x_1, x_2, \dots, x_n)\}$ ,  $a_i \leq x_i \leq b_i$ ,  $-\infty \leq a_i \leq b_i \leq +\infty$ ,  $i = 1, 2, \dots, n$ . Вероятностное распределение задается на элементах  $F$ .

Среди непрерывных вероятностных пространств особый интерес представляют так называемые абсолютно непрерывные пространства. В них вероятностное распределение задается с помощью специальной функции – плотности распределения.

**Определение 3.5.** Говорят, что функция  $f(x_1, \dots, x_n)$ , отображающая  $\mathbb{R}^n$  в  $\mathbb{R}$ , является плотностью распределения вероятности в следующих случаях:

1.  $f(x_1, \dots, x_n) \geq 0 \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n$ .
2.  $f(x_1, \dots, x_n)$  интегрируема на любом множестве  $A \in F$ .
3.  $\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$ .

Тогда вероятностное распределение на  $F$  определяется так:

$$\forall A \in F, \quad P(A) = \int_A f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

**Пример.** Рассмотрим равномерное распределение на  $\Omega = [a, b]$ . В качестве  $F$  берем борелевскую  $\sigma$ -алгебру  $\mathcal{B}_1$ , порожденную отрезками прямой.

Вероятность каждого события  $A$  из  $\mathcal{B}_1$  определим по формуле

$$P(A) = \frac{\alpha(A)}{b - a},$$

где  $\alpha(A)$  – суммарная длина промежутков, составляющих множество  $A$ . Тогда

$$P(A) = \int_A \frac{1}{b - a} dx, \quad \forall A \in \mathcal{B}_1.$$

Тогда при  $x \in [a, b]$  плотность распределения вероятности  $f(x) = \frac{1}{b - a}$ .

Для приведения полученной модели к виду  $(R, \mathcal{B}_1, P)$  нужно доопределить функцию  $f(x)$  нулями вне  $[a, b]$ . Таким образом, плотность распределения

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in [a, b], \\ 0 & \text{при } x \in \mathbb{R} \setminus [a, b]. \end{cases} \bullet$$

*Смешанное вероятностное пространство* конструируется объединением дискретного и непрерывного вероятностных пространств. Ввиду некоторой искусственности подобной конструкции в дальнейшем этот тип вероятностного пространства рассматриваться не будет.

## 4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

### 4.1. Случайная величина и ее распределение

Часто встречается ситуация, когда результатом случайного эксперимента является число. Это может быть, например, число очков, выпавших на верхней грани игральной кости, или результат измерения прибором некоторого физического параметра. В таких случаях говорят, что в результате случайного эксперимента получается значение случайной величины. Естественно, разные значения случайной величины соответствуют разным исходам случайного эксперимента, каждому набору возможных значений случайной величины соответствует некоторое случайное событие вероятностного пространства, соответствующего рассматриваемому эксперименту. Итак, можно сформулировать следующее определение.

**Определение 4.1.** *Случайной величиной  $X(\omega)$  называется функция  $X : \Omega \rightarrow \mathbb{R}$ , заданная на множестве элементарных событий  $\Omega$  вероятностного пространства  $(\Omega, F, P)$ , такая, что для любого промежутка  $[a, b]$  событие  $\{\omega \mid X(\omega) \in [a, b]\}$  принадлежит  $F$ .*

В дальнейшем будем обозначать случайные величины большими латинскими буквами  $X, Y, Z$ , возможно, с индексами, а аргумент  $\omega$  будем опускать.

Если задана случайная величина  $X$ , то, согласно определению 4.1, может быть вычислена вероятность любого события вида  $\{X \in [a, b]\}$ , которую можно приписать непосредственно промежутку  $[a, b]$ , т. е. построить вероятностное пространство на  $\mathbb{R}$ , где  $\sigma$ -алгебра событий  $F$  есть  $\sigma$ -алгебра борелевских множеств, порожденная интервалами на прямой. Другими словами, каждой случайной величине соответствует вероятностное распределение на прямой. Оно называется распределением случайной величины. В любой реальной задаче интерес прежде всего представляют ве-

роятности того, что случайная величина принимает то или иное значение или попадает в некоторый промежуток, т. е. те данные, которые определяются ее распределением.

Задать распределение любой случайной величины можно при помощи функции распределения.

**Определение 4.2.** *Функцией распределения случайной величины  $X$  называется функция  $F : \mathbb{R} \rightarrow \mathbb{R}$ , которая каждому значению  $x$  ставит в соответствие число  $F(x) = P(X < x)$ .*

По заданной функции распределения легко вычислить вероятность попадания случайной величины в любой промежуток  $[a, b)$ . Действительно,

$$F(b) = P(X < b) = P(X < a) + P(a \leq X < b) = F(a) + P(a \leq X < b),$$

следовательно,

$$P\{X \in [a, b[\} = F(b) - F(a).$$

Функция распределения обладает следующими свойствами:

- 1)  $0 \leq F(x) \leq 1$ ;
- 2)  $F(x)$  – неубывающая функция;

3)  $F(x)$  непрерывна слева, т. е. в любой точке  $a$  левосторонний предел равен значению функции:

$$\lim_{x \rightarrow a-0} F(x) = F(a);$$

$$4) \lim_{x \rightarrow -\infty} F(x) = 0; \lim_{x \rightarrow +\infty} F(x) = 1.$$

Свойство 1 очевидно, так как значение  $F(x)$  представляет собой вероятность. Свойство 2 очевидным образом следует из только что полученного равенства для вероятности попадания случайной величины в промежуток. Свойство 3 следует из счетной аддитивности вероятности. Действительно, рассмотрим произвольную монотонно возрастающую последовательность  $\{x_m\}$ , сходящуюся к  $a$ . Тогда

$$\begin{aligned} F(a) - F(x_1) &= P(X \in [x_1, a]) = P\left(\bigcup_{m=1}^{\infty} \{X \in [x_m, x_{m+1}[\}\right) = \\ &= \sum_{m=1}^{\infty} P(X \in [x_m, x_{m+1}[). \end{aligned}$$

Таким образом, получили сходящийся числовой ряд. Частная сумма этого ряда сходится к его сумме, т. е.

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n P(X \in [x_m, x_{m+1}[) = F(a) - F(x_1)$$

или

$$\lim_{n \rightarrow \infty} (F(x_{n+1}) - F(x_1)) = F(a) - F(x_1),$$

откуда следует, что

$$\lim_{n \rightarrow \infty} F(x_n) = F(a)$$

для любой монотонно возрастающей сходящейся к  $a$  последовательности. Согласно одному из возможных определений левостороннего предела функции в точке, это и доказывает свойство 3.

Аналогично доказывается свойство 4. Возьмем монотонно убывающую последовательность  $\{x_m\}$ , такую, что  $\lim_{m \rightarrow +\infty} x_m = -\infty$ . Тогда

$$F(x_1) = P(X < x_1) = \sum_{m=1}^{\infty} P(X \in [x_{m+1}, x_m[).$$

Частная сумма этого ряда

$$\sum_{m=1}^n P(X \in [x_{m+1}, x_m[) = F(x_1) - F(x_{n+1}).$$

Следовательно,  $\lim_{n \rightarrow \infty} F(x_{n+1}) = 0$ , что и доказывает первое равенство свойства 4. Для доказательства второго равенства возьмем монотонно возрастающую последовательность  $\{x_m\}$ , такую, что  $\lim_{m \rightarrow +\infty} x_m = +\infty$ . Тогда

$$P(X \geq x_1) = \sum_{m=1}^{\infty} P(X \in [x_m, x_{m+1}[),$$

а частная сумма

$$\sum_{m=1}^n P(X \in [x_m, x_{m+1}[) = P(X \geq x_1) - P(X \geq x_{n+1}).$$

Следовательно,  $\lim_{n \rightarrow \infty} P(X \geq x_{n+1}) = 0$ .

Далее, согласно формуле вычисления вероятности противоположного события,

$$F(x_{n+1}) = P(X < x_{n+1}) = 1 - P(X \geq x_{n+1}),$$

откуда следует второе равенство свойства 4.

Функция распределения полностью определяет распределение случайной величины, т. е. по ней может быть вычислена вероятность попадания случайной величины в любое борелевское множество. В частности, вероятность попадания в точку  $P(X = a) = \lim_{x \rightarrow a+0} F(x) - F(a)$ .

Свойства 1–4 являются характеристическими свойствами функции распределения: любая функция, удовлетворяющая этим свойствам, может рассматриваться как функция распределения некоторой случайной величины.

## 4.2. Случайная величина с дискретным распределением

Случайная величина, множество значений которой является дискретным (конечным или счетным), называется дискретной случайной величиной. Соответствующее ей вероятностное распределение на прямой также называется дискретным. Поскольку дискретное вероятностное пространство имеет конечный или счетный набор элементарных событий, дискретная случайная величина может принимать только конечное или счетное множество значений. Пусть дискретная случайная величина ставит в соответствие каждому элементарному событию  $\omega_i$  число  $x_i$  и пусть  $P(\omega_i) = p_i$ . В терминах вероятностного распределения на прямой это означает, что случайная величина принимает только значения  $x_i$  и  $P(X = x_i) = p_i$ .

**Определение 4.3.** *Набор значений  $x_i$  вместе с вероятностями  $p_i$  называется рядом распределения дискретной случайной величины.*

Очевидно, что ряд распределения должен обладать следующими свойствами:

- 1)  $p_i \geq 0$ ;
- 2)  $\sum_i p_i = 1$ .

Свойства 1–2 являются характеристическими: всякий набор пар чисел  $(x_i, p_i)$ , где  $p_i$  удовлетворяют свойствам 1–2, может рассматриваться как ряд распределения некоторой дискретной случайной величины.

По ряду распределения дискретной случайной величины нетрудно восстановить ее функцию распределения. Пусть значения  $x_i$  расположены в порядке возрастания. Тогда  $F(x_1) = P(X < x_1) = 0$  и, следовательно,  $F(x) = 0$  при любом  $x < x_1$ . Далее, при  $x \in [x_1, x_2[$  имеем

$$F(x) = P(X = x_1) = p_1,$$

при  $x \in [x_2, x_3[$  получим

$$F(x) = P(X = x_1) + P(X = x_2) = p_1 + p_2$$

и т. д. Таким образом, функция распределения дискретной случайной величины представляет собой кусочно-постоянную функцию с разрывами в точках  $x_i$ , причем в каждой точке  $x_i$  значение функции увеличивается на  $p_i$ . В самой точке  $x_i$  значение  $F(x)$  равно значению до разрыва, что согласуется со свойством непрерывности слева функции распределения.

### 4.3. Случайная величина с абсолютно непрерывным распределением

Случайная величина  $X$  называется абсолютно непрерывной, если существует такая функция  $f(x)$ , что для любого борелевского множества  $\mathcal{A}$  на прямой  $P(x \in \mathcal{A}) = \int_{\mathcal{A}} f(x)dx$ . Согласно определению 3.5, вероятност-

ное пространство, на котором задана такая случайная величина, является абсолютно непрерывным с плотностью  $f(x)$ . Эта плотность удовлетворяет следующим свойствам:

- 1)  $f(x) \geq 0$  для любого  $x \in R^1$ ;
- 2)  $\int_{-\infty}^{+\infty} f(x)dx = 1$ .

Функция  $f(x)$  называется плотностью распределения абсолютно непрерывной случайной величины. По ней вычисляется вероятность попадания случайной величины в любой промежуток  $[a, b]$ :

$$P\{X \in [a, b]\} = \int_a^b f(x)dx.$$

По плотности может быть однозначно восстановлена функция распределения. Действительно,

$$F(x) = P(X \in (-\infty, x]) = \int_{-\infty}^x f(x)dx.$$

С другой стороны, по известной теореме об интеграле с переменным верхним пределом из последней формулы следует, что во всех точках непрерывности  $f(x)$  у функции абсолютно непрерывного распределения существует производная  $F'(x) = f(x)$ .

### 4.4. Интеграл Стильтьеса

Для того чтобы в наиболее общем виде определить числовые характеристики случайных величин, будем использовать понятие „интеграла по распределению“, некоторым образом обобщающее понятие определенного интеграла Римана.

Пусть  $F(x)$  – некоторая неубывающая кусочно-непрерывная и непрерывная слева функция, заданная на числовой прямой, и пусть  $g(x)$  – некоторая непрерывная функция, заданная на промежутке  $[a, b]$ . Разобьем про-

межуток  $[a, b]$  на  $n$  частей точками  $x_i$ :  $a = x_0 < x_1 < \dots < x_n = b$ , и пусть  $\Delta = \max_{1 \leq i \leq n} (x_i - x_{i-1})$ .

**Определение 4.4.** *Интеграл Стильтьеса*

$$\int_a^b g(x) dF(x) = \lim_{\Delta \rightarrow 0} \sum_{i=1}^{n(\Delta)} g(y_i) (F(x_i) - F(x_{i-1})),$$

где точки  $y_i \in [x_{i-1}, x_i]$ .

Иными словами, интеграл Стильтьеса существует, если существует предел, не зависящий от выбора точек  $x_i$  и  $y_i$ .

Заметим, что при  $F(x) = x$  интеграл Стильтьеса – это обычный определенный интеграл. Кроме того, в случае непрерывно дифференцируемой  $F(x)$  интеграл Стильтьеса также сводится к обычному интегралу:

$$\int_a^b g(x) dF(x) = \int_a^b g(x) F'(x) dx.$$

Действительно, согласно теореме о среднем, на промежутках  $[x_{i-1}, x_i]$  существуют такие точки  $y_i$ , что  $F(x_i) - F(x_{i-1}) = F'(y_i)(x_i - x_{i-1})$ .

Можно выбрать эти же точки  $y_i$  в качестве аргументов  $g(x)$  в формуле определения интеграла Стильтьеса. В результате получим одно из возможных определений определенного интеграла с подынтегральной функцией  $g(x)F'(x)$ .

В частности, если на некотором интервале  $F(x)$  равна константе, то интеграл Стильтьеса от любой функции по этому интервалу равен нулю.

Если в интеграле Стильтьеса в качестве подынтегральной функции взять 1, то в правой части формулы определения будет стоять просто суммарное приращение функции  $F(x)$  на  $[a, b]$ . Таким образом,

$$\int_a^b 1 dF(x) = F(b) - F(a)$$

для любой  $F(x)$ .

Можно показать, что интеграл Стильтьеса, как и обычный интеграл, существует для любой непрерывной подынтегральной функции  $g(x)$  и обладает теми же свойствами:

1) свойством линейности: для любых чисел  $c_1, c_2$  и любых непрерывных функций  $g(x)$  и  $h(x)$  справедливо равенство

$$\int_a^b (c_1 g(x) + c_2 h(x)) dF(x) = c_1 \int_a^b g(x) dF(x) + c_2 \int_a^b h(x) dF(x);$$

2) свойством аддитивности: для любых  $a, c, b$ , таких, что  $a < c < b$ , и любой функции  $g(x)$ , непрерывной и ограниченной на  $[a, b]$  и  $[c, b]$ , справедливо равенство

$$\int_a^b g(x)dF(x) = \int_a^c g(x)dF(x) + \int_c^b g(x)dF(x).$$

Используя аддитивность, интеграл Стильтьеса можно определить для любой кусочно-непрерывной и непрерывной слева подынтегральной функции как сумму интегралов по полуоткрытым интервалам, на которых эта функция непрерывна. Наконец, при помощи предельного перехода можно определить интеграл Стильтьеса по всей числовой оси:

$$\int_{-\infty}^{+\infty} g(x)dF(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b g(x)dF(x).$$

#### 4.5. Числовые характеристики случайных величин

Ранее рассматривалось распределение случайной величины, которое может быть задано при помощи ряда распределения, функции распределения или плотности распределения. Однако в некоторых случаях интерес представляет не само распределение (которое к тому же часто не известно), а лишь некоторые числа, определенным образом его характеризующие, – так называемые числовые характеристики. Среди числовых характеристик распределения наиболее часто используются интегральные характеристики, получающиеся в общем виде как интеграл Стильтьеса по функции распределения от некоторой функции. Простейшей и наиболее часто используемой из них является математическое ожидание.

**Определение 4.5.** *Математическим ожиданием случайной величины  $X$  называется число*

$$M\{X\} = \int_{-\infty}^{\infty} x dF(x),$$

где  $F(x)$  – функция распределения случайной величины  $X$ .

В случае, когда  $X$  – абсолютно непрерывная случайная величина, существует плотность  $f(x) = F'(x)$  и математическое ожидание можно вы-

числить через плотность по формуле

$$M\{X\} = \int_{-\infty}^{\infty} xf(x)dx.$$

Если  $X$  – дискретная случайная величина, функция распределения имеет приращения только в точках  $x_i$  и эти приращения равны вероятностям  $p_i$ . Поэтому, подставив в формулу определения интеграла Стильтьеса  $g(x) = x$ , получим формулу для вычисления математического ожидания дискретной случайной величины через ряд распределения:

$$M\{X\} = \sum_i x_i p_i.$$

Сумма, стоящая в правой части последнего равенства, является суммой конечного числа слагаемых, если случайная величина может принимать только конечное число значений, или суммой ряда, если случайная величина может принимать бесконечное, но счетное число значений. Математическое ожидание существует, если соответствующий числовый ряд сходится абсолютно.

Для того чтобы понять смысл числовой характеристики  $M\{X\}$ , представим себе, что вместо вероятностного распределения задано распределение масс на прямой. Соответственно, плотность  $f(x)$  – функция зависимости физической плотности от координаты  $x$ , а  $p_i$  – точечные массы, расположенные в точках  $x_i$ . Тогда математическое ожидание представляет собой координату центра тяжести распределения масс. Т. е. п математическое ожидание – это „координата центра тяжести“ вероятностного распределения. С другой стороны, рассмотрим случай дискретной случайной величины, которая может принимать  $N$  значений с одинаковыми вероятностями. Тогда, поскольку сумма всех вероятностей равна 1, каждая из вероятностей  $p_i = 1/N$ , и математическое ожидание будет равно среднему арифметическому возможных значений случайной величины  $x_i$ . Таким образом, математическое ожидание представляет собой в некотором смысле среднее значение случайной величины.

Отметим, что из приведенных формул не следует, что математическое ожидание всегда существует: соответствующий интеграл или ряд может расходиться. Действительно, известны примеры вероятностных распределений, у которых математическое ожидание не существует.

Всякая интегральная числовая характеристика случайной величины  $X$  может быть представлена как интеграл Стильтьеса по ее функции распределения от некоторой функции  $g(x)$ . С другой стороны, эта же характеристика может рассматриваться как математическое ожидание случайной

величины  $g(X)$ , т. е. существует равенство

$$M\{g(X)\} = \int_{-\infty}^{\infty} g(x)dF(x).$$

В частности, в случае абсолютно непрерывной случайной величины с плотностью распределения  $f(x)$

$$M\{g(X)\} = \int_{-\infty}^{\infty} g(x)f(x)dx,$$

а в случае дискретной случайной величины с рядом распределения  $\{(x_i, p_i)\}$

$$M\{g(X)\} = \sum_i g(x_i)p_i.$$

Чтобы обосновать это равенство, возьмем в качестве  $g(x)$  кусочно постоянную функцию, т. е. выберем на числовой прямой точки  $x_1 < x_2 < \dots < x_{n-1}$  и примем

$$g(x) = \begin{cases} y_1, & \text{если } x < x_1, \\ y_2, & \text{если } x_1 \leq x < x_2, \\ \dots \\ y_n, & \text{если } x \geq x_{n-1}. \end{cases}$$

Тогда  $g(X)$  представляет собой дискретную случайную величину, принимающую значения  $y_1, y_2, \dots, y_n$  с вероятностями:

$$\begin{aligned} p_1 &= P(X < x_1) = F(x_1); \\ p_i &= P(x_{i-1} \leq X \leq x_i) = F(x_i) - F(x_{i-1}), \quad i = 2, \dots, n-1; \\ p_n &= P(X \geq x_n) = 1 - F(x_{n-1}). \end{aligned}$$

Согласно формуле для вычисления математического ожидания дискретной случайной величины,

$$M\{g(X)\} = y_1 F(x_1) + \sum_{i=2}^{n-1} y_i (F(x_i) - F(x_{i-1})) + y_n (1 - F(x_{n-1})).$$

С другой стороны, используя свойства линейности и аддитивности интеграла Стильтьеса, получим:

$$\int_{-\infty}^{+\infty} g(x)dF(x) = \int_{-\infty}^{x_1} y_1 dF(x) + \sum_{i=2}^{n-1} \int_{x_{i-1}}^{x_i} y_i dF(x) + \int_{x_{n-1}}^{+\infty} y_n dF(x) =$$

$$= y_1 F(x_1) + \sum_{i=2}^{n-1} y_i (F(x_i) - F(x_{i-1})) + y_n (1 - F(x_{n-1})).$$

Таким образом, для кусочно-постоянной функции требуемое равенство получено. Не будем приводить здесь полного доказательства этого равенства для произвольной  $g(x)$ , отметим лишь, что оно основывается на том факте, что любая кусочно непрерывная функция может быть сколь угодно точно приближена кусочно постоянной функцией.

Теперь нетрудно показать, что математическое ожидание обладает свойством линейности, а именно: для любых вещественных чисел  $a$  и  $b$   $M\{aX + b\} = aM\{X\} + b$ .

Действительно,

$$\begin{aligned} M\{aX + b\} &= \int_{-\infty}^{+\infty} (ax + b) dF(x) = a \int_{-\infty}^{+\infty} x dF(x) + b \int_{-\infty}^{\infty} 1 dF(x) = \\ &= aM\{X\} + b \left( \lim_{x \rightarrow +\infty} F(x) - \lim_{x \rightarrow -\infty} F(x) \right) = aM\{X\} + b. \end{aligned}$$

В качестве интегральных числовых характеристик случайной величины обычно используются начальные и центральные моменты.

**Определение 4.6.** Начальным моментом порядка  $k$  называется число

$$\alpha_k = M\{X^k\}.$$

Центральным моментом порядка  $k$  называется число

$$\mu_k = M\{(X - M\{X\})^k\}.$$

Таким образом,  $\alpha_1 = M\{X\}$ ,  $\mu_1 = 0$ .

Кроме математического ожидания наиболее часто используется центральный момент второго порядка, который называется дисперсией:

$$D\{X\} = M\{(X - M\{X\})^2\} = \int_{-\infty}^{\infty} (x - M\{X\})^2 dF(x).$$

Таким образом, в случае абсолютно непрерывной случайной величины с плотностью распределения  $f(x)$

$$D\{X\} = \int_{-\infty}^{\infty} (x - M\{X\})^2 f(x) dx,$$

а в случае дискретной случайной величины с рядом распределения  $\{(x_i, p_i)\}$

$$D\{X\} = \sum_i (x_i - M\{X\})^2 p_i.$$

Если математическое ожидание – это в некотором смысле среднее значение случайной величины, то о дисперсии можно сказать, что это – „средний квадрат отклонения от среднего“. Таким образом, дисперсия характеризует разброс значений случайной величины относительно ее математического ожидания.

Отметим, что дисперсия всегда неотрицательна, как, очевидно, неотрицательно математическое ожидание любой случайной величины, принимающей только неотрицательные значения. Кроме того, для любых вещественных чисел  $a$  и  $b$  справедливо  $D\{aX + b\} = a^2 D\{X\}$ .

Действительно,

$$\begin{aligned} D\{aX + b\} &= M\{(aX + b - M\{aX + b\})^2\} = \\ &= M\{(aX + b - aM\{X\} - b)^2\} = \\ &= M\{(a(X - M\{X\}))^2\} = a^2 M\{(X - M\{X\})^2\} = a^2 D\{X\}. \end{aligned}$$

Для вычисления дисперсии иногда удобнее применять формулу

$$D\{X\} = M\{X^2\} - (M\{X\})^2 = \alpha_2 - \alpha_1^2.$$

Справедливость этой формулы видна из следующей цепочки равенств:

$$\begin{aligned} D\{X\} &= \int_{-\infty}^{\infty} (x - M\{X\})^2 dF(x) = \int_{-\infty}^{\infty} (x^2 - 2xM\{X\} + (M\{X\})^2) dF(x) = \\ &= \int_{-\infty}^{\infty} x^2 dF(x) - 2M\{X\} \int_{-\infty}^{\infty} x dF(x) + (M\{X\})^2 \int_{-\infty}^{\infty} 1 dF(x) = \\ &= M\{X^2\} - 2M\{X\}M\{X\} + (M\{X\})^2 = M\{X^2\} - (M\{X\})^2. \end{aligned}$$

Заметим, что дисперсия имеет размерность квадрата размерности самой случайной величины, поэтому в качестве характеристики разброса значений чаще используют квадратный корень из дисперсии, который называется среднеквадратичным отклонением (СКО) случайной величины  $X$ .

Поскольку дисперсия характеризует разброс значений случайной величины относительно ее математического ожидания, естественно предположить, что вероятность больших отклонений случайной величины от ее математического ожидания оценивается через дисперсию. Для нахождения такой оценки получим сначала одно вспомогательное неравенство, касающееся математического ожидания.

**Теорема 4.1.** Пусть  $X$  – случайная величина, принимающая только неотрицательные значения. Тогда для любого  $t > 0$  справедливо неравенство

$$P(X > t) \leq \frac{1}{t} M\{X\}.$$

**Доказательство.** Имеем

$$\begin{aligned} P(X > t) &= \int_t^{+\infty} 1 dF(x) \leq \int_t^{+\infty} \frac{x}{t} dF(x) = \\ &= \frac{1}{t} \int_t^{+\infty} x dF(x) \leq \frac{1}{t} \int_0^{+\infty} x dF(x) = \frac{1}{t} M\{X\}. \end{aligned}$$

Последнее равенство справедливо в силу неотрицательности случайной величины  $X$ . ■

Докажем теперь основное неравенство, которое называется неравенством Чебышева.

**Теорема 4.2.** Пусть  $X$  – случайная величина с конечной дисперсией  $D\{X\}$ . Тогда для любого  $t > 0$  справедливо неравенство

$$P(|X - M\{X\}| \geq t) \leq \frac{D\{X\}}{t^2}.$$

**Доказательство.** Применив теорему 4.1 к неотрицательной случайной величине  $(X - M\{X\})^2$ , получим:

$$\begin{aligned} P\{|X - M\{X\}| \geq t\} &= P\{(X - M\{X\})^2 \geq t^2\} \leq \\ &\leq \frac{M\{(X - M\{X\})^2\}}{t^2} = \frac{D\{X\}}{t^2}. \end{aligned}$$

Неравенство доказано. ■

Кроме интегральных числовых характеристик случайной величины существуют и другие, не выражющиеся через интеграл Стильтьеса. Из таких характеристик чаще всего используется медиана.

**Определение 4.7.** Число  $t$  называется медианой случайной величины  $X$ , если

$$P(X \leq m) \geq \frac{1}{2}$$

и

$$P(X \geq m) \geq \frac{1}{2}.$$

Иными словами, медиана – в определенном смысле средняя точка распределения. (В случае абсолютно непрерывного распределения медиана определяется равенством  $F(m) = 0.5$ .) По сравнению с математическим ожиданием медиана обладает тем преимуществом, что она существует у любой случайной величины.

#### 4.6. Примеры распределений дискретных случайных величин

**Двуточечное распределение.** Пусть случайная величина может принимать только 2 значения 0 и 1 с вероятностями  $(1 - p)$  и  $p$  соответственно. Ее ряд распределения задается таблицей

$x_i$	0	1
$p_i$	$1 - p$	$p$

, а график функции распределения изображен на рис. 4.1.

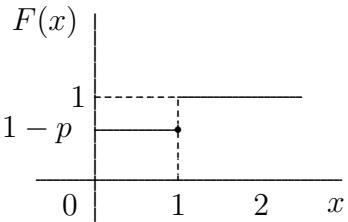


Рис. 4.1

Найдем числовые характеристики этой величины. Согласно формуле вычисления математического ожидания дискретной случайной величины,  $\alpha_1 = M\{X\} = 0(1 - p) + 1p = p$ . Начальный второй момент  $\alpha_2 = M\{X^2\} = 0^2(1 - p) + 1^2p = p$ . Дисперсия  $D\{X\} = \alpha_2 - \alpha_1^2 = p - p^2 = p(1 - p)$ .

**Биномиальное распределение.** Пусть многократно повторяется один и тот же случайный эксперимент с двумя возможными исходами  $\omega_1$  и  $\omega_2$  и пусть  $P(\omega_1) = p$ , соответственно,  $P(\omega_2) = 1 - p$ . Эксперимент повторяется  $n$  раз, причем результаты разных экспериментов не зависят друг от друга, т. е. события вида  $\{\omega_j - \text{исход } i\text{-го эксперимента}\}$  ( $i = 1, \dots, n$ ) являются независимыми в совокупности. Такая модель называется последовательностью независимых испытаний Бернулли. Построим соответствующее этой модели вероятностное пространство. В качестве одного элементарного события следует взять набор из  $n$  исходов отдельных экспериментов  $\omega = (u_1, u_2, \dots, u_n)$ , где  $u_i$  может принимать значение  $\omega_1$  или  $\omega_2$ . Согласно предположению о независимости экспериментов, каждому такому исходу следует приписать вероятность  $P(\omega) = P(u_1)P(u_2)\dots P(u_n)$ , где  $P(u_i)$  в зависимости от значения  $u_i$  может принимать значение  $p$  или  $1 - p$ .

Введем теперь случайную величину  $X$  – число наступлений исхода  $\omega_1$ . Это – дискретная случайная величина, которая может принимать значения:  $0, 1, \dots, n$ . Построим ряд распределения данной случайной величины, т. е. найдем вероятности  $p_k = P(X = k)$  ( $k = 0, 1, \dots, n$ ).

Событие  $\{X = k\}$  представляет собой объединение тех и только тех элементарных событий вероятностного пространства, в которых  $u_i$  прини-

мает значение  $\omega_1$  ровно  $k$  раз и, соответственно,  $n-k$  раз принимает значение  $\omega_2$ . Вероятность каждого из таких событий равна  $p^k(1-p)^{n-k}$ . Количество таких событий равно числу способов выбора мест для подстановки  $\omega_1$  в набор  $(u_1, u_2, \dots, u_n)$ , т. е. числу сочетаний  $C_n^k$ . Таким образом, просуммировав вероятности этих событий, получим  $P(X = k) = C_n^k p^k (1-p)^{n-k}$ . В частности,  $P(X = 0) = (1-p)^n$ ,  $P(X = n) = p^n$ .

Построенное распределение называется биномиальным, его числовые характеристики будут найдены в 5.5.

**Пример.** Симметричная монета подбрасывается 5 раз,  $X$  – число выпадений герба. Здесь  $n = 5$ ,  $p = 1 - p = 1/2$ , поэтому  $P(X = k) = C_5^k \left(\frac{1}{2}\right)^n$ , и распределенная по биномиальному закону случайная величина  $X$  имеет ряд распределения

$x_i$	0	1	2	3	4	5
$p_i$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

**Геометрическое распределение.** Пусть независимые испытания Бернулли повторяются неограниченное число раз, по крайней мере до тех пор, пока первый раз не наступит исход  $\omega_1$ . Пусть  $X$  – номер испытания, во время которого это произошло. Множеством возможных значений этой величины является множество всех натуральных чисел. Событие  $\{X = k\}$  означает, что в первых  $k-1$  испытаниях наступил исход  $\omega_2$ , а в последнем,  $k$ -м испытании –  $\omega_1$ . Поэтому, в силу независимости испытаний,  $P(X = k) = p(1-p)^{k-1}$  ( $k = 1, 2, \dots$ ).

Используя формулу суммы членов бесконечно убывающей геометрической прогрессии, нетрудно убедиться, что сумма всех этих вероятностей действительно равна 1. Построенное распределение называется геометрическим. Найдем математическое ожидание случайной величины  $X$ , которое можно интерпретировать как среднее время ожидания первого наступления исхода  $\omega_1$ . Для этого введем функцию  $h(z)$ , представляющую собой сумму членов бесконечно убывающей геометрической прогрессии со знаменателем  $z$ , т. е.

$$h(z) = \sum_{k=0}^{\infty} z^k = \frac{1}{1-z},$$

и заметим, что по теореме о почленном дифференцировании степенного ряда

$$h'(z) = \sum_{k=1}^{\infty} kz^{k-1} = \frac{1}{(1-z)^2}.$$

Используя формулу вычисления математического ожидания дискретной случайной величины и последнее равенство, находим

$$M\{X\} = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = ph'(1-p) = p \frac{1}{p^2} = \frac{1}{p}.$$

**Распределение Пуассона.** Говорят, что случайная величина  $X$  имеет распределение Пуассона с параметром  $\lambda > 0$ , если она принимает только целые неотрицательные значения, причем

$$p_k = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (k = 0, 1, 2, \dots).$$

Таким образом,  $X$  – дискретная случайная величина, имеющая бесконечно много возможных значений. Прежде всего проверим выполнение необходимых свойств указанного ряда распределения:

- 1)  $p_k > 0$  – очевидно;
- 2) согласно формуле разложения экспоненты в степенной ряд,

$$\sum_{k=0}^{\infty} p_k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Найдем теперь числовые характеристики  $X$ . Согласно формуле вычисления математического ожидания дискретной случайной величины,

$$\begin{aligned} M\{X\} &= \sum_{k=0}^{\infty} kp_k = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

Начальный второй момент

$$\begin{aligned} M\{X^2\} &= \sum_{k=0}^{\infty} k^2 p_k = e^{-\lambda} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} = \\ &= e^{-\lambda} \lambda \sum_{m=0}^{\infty} (m+1) \frac{\lambda^m}{m!} = \lambda \left( e^{-\lambda} \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} + e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \right) = \\ &= \lambda(M\{X\} + 1) = \lambda(\lambda + 1). \end{aligned}$$

Тогда дисперсия

$$D\{X\} = M\{X^2\} - (M\{X\})^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda.$$

## 4.7. Примеры распределений абсолютно непрерывных случайных величин

**Равномерное распределение.** Говорят, что случайная величина  $X$  имеет равномерное распределение на отрезке  $[a, b]$ , если ее плотность распределения постоянна на отрезке  $[a, b]$  и равна нулю вне этого отрезка. Поскольку интеграл от плотности по всей числовой оси, который в данном случае равен просто произведению постоянного значения плотности на длину отрезка, должен быть равен 1, получаем формулу для плотности:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{если } x \in [a, b], \\ 0, & \text{если } x \notin [a, b]. \end{cases}$$

График плотности равномерного распределения изображен на рис. 4.2.

Соответственно, функция распределения определяется по формуле

$$F(x) = \int_{-\infty}^x f(t)dt = \begin{cases} 0, & \text{если } x < a, \\ \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}, & \text{если } x \in [a, b], \\ 1, & \text{если } x > b. \end{cases} \quad (4.1)$$

График функции равномерного распределения изображен на рис. 4.3.

Ясно, что „центр тяжести“ равномерного распределения находится в середине отрезка  $[a, b]$ , т. е.  $M\{X\} = \frac{a+b}{2}$ .

В этом также нетрудно убедиться, вычислив

$$\int_{-\infty}^{\infty} xf(x)dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}.$$

Для вычисления дисперсии обозначим через  $h$  половину длины про-

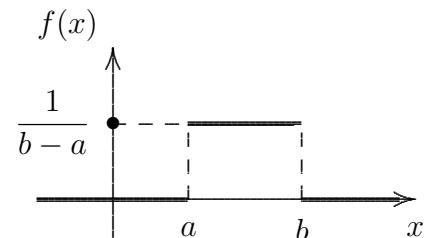


Рис. 4.2

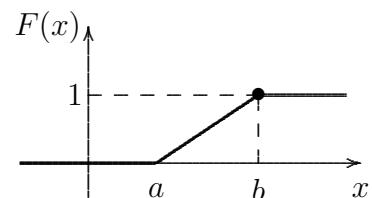


Рис. 4.3

межутка. Тогда

$$\begin{aligned}
 D\{X\} &= \int_{-\infty}^{\infty} (x - M\{X\})^2 f(x) dx = \frac{1}{b-a} \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx = \\
 &= \frac{1}{b-a} \int_{-h}^h y^2 dy = \frac{1}{b-a} \frac{2h^3}{3} = \frac{(b-a)^2}{12}.
 \end{aligned}$$

Соответственно, среднеквадратичное отклонение

$$\sqrt{D\{X\}} = \frac{b-a}{2\sqrt{3}}.$$

**Нормальное распределение.** Стандартным нормальным распределением называется абсолютно непрерывное распределение с плотностью

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

График плотности стандартного нормального распределения показан на рис. 4.4.

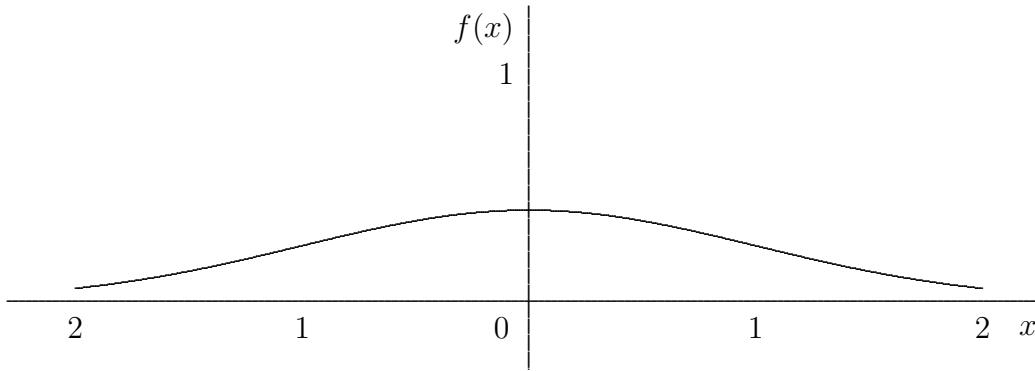


Рис. 4.4

Функция  $f(x)$ , очевидно, положительна. Требуется еще проверить, что величина

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

Вычислим квадрат этого интеграла, сведя его к двойному интегралу и использовав переход к полярным координатам:

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy =$$

$$= \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^\infty e^{-\frac{r^2}{2}} r dr = \frac{1}{2\pi} 2\pi \left( -e^{-\frac{r^2}{2}} \right) \Big|_0^{+\infty} = 1.$$

Отсюда, так как величина  $I$ , очевидно, положительна, следует, что  $I = 1$ .

Кроме того, плотность  $f(x)$  является четной функцией, достигает максимального значения при  $x = 0$ , а с удалением  $x$  от нуля убывает достаточно быстро для того, чтобы при любом натуральном  $k$  существовали интегралы

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

т. е. стандартное нормальное распределение имеет конечные начальные моменты любого порядка.

Если распределение имеет четную плотность, то оно называется симметричным. У такого распределения момент любого нечетного порядка, если он существует, равен нулю, как интеграл от нечетной функции  $x^k f(x)$  по симметричному промежутку. Таким образом, у стандартного нормального распределения математическое ожидание равно нулю. Для вычисления дисперсии воспользуемся формулой интегрирования по частям и сведем соответствующий интеграл к интегралу от плотности:

$$\begin{aligned} D\{X\} &= M\{(X - M\{X\})^2\} = M\{X^2\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \left( x \left( -e^{-\frac{x^2}{2}} \right) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left( -e^{-\frac{x^2}{2}} \right) dx \right) = 0 + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1. \end{aligned}$$

Итак, стандартное нормальное распределение имеет нулевое математическое ожидание и единичную дисперсию.

Соответствующая функция распределения не может быть выражена через элементарные функции, поэтому ее приходится записывать как интеграл от плотности:

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

Для вычисления этой функции при положительных значениях аргумента существуют специальные таблицы. Кроме того, функция  $\Phi(x)$  легко

вычисляется через специальную функцию  $\text{Erf}(x)$  – так называемую функцию ошибок:

$$\Phi(x) = \frac{1}{2} \left( 1 + \text{Erf} \left( \frac{x}{\sqrt{2}} \right) \right).$$

Для вычисления  $\Phi(x)$  при отрицательных значениях аргумента используется симметричность стандартного нормального распределения, благодаря которой  $\Phi(-x) = 1 - \Phi(x)$ .

Рассмотрим теперь случайную величину, являющуюся линейной функцией от  $X$ :  $Y = \sigma X + a$ , где  $\sigma > 0$ ,  $a$  – любое вещественное число,  $X$  имеет стандартное нормальное распределение. Согласно свойствам математического ожидания и дисперсии,

$$\begin{aligned} M\{Y\} &= \sigma M\{X\} + a = a, \\ D\{Y\} &= \sigma^2 D\{X\} = \sigma^2. \end{aligned}$$

Функция распределения этой случайной величины

$$F(x) = P(Y < x) = P(\sigma X + a < x) = P \left( X < \frac{x - a}{\sigma} \right) = \Phi \left( \frac{x - a}{\sigma} \right).$$

Соответственно, плотность распределения

$$f(x) = \left( \Phi \left( \frac{x - a}{\sigma} \right) \right)' = \frac{1}{\sigma} \Phi' \left( \frac{x - a}{\sigma} \right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Распределение с такой плотностью называется нормальным распределением с параметрами  $(a, \sigma^2)$ . Как только что было показано, параметр  $a$  является математическим ожиданием соответствующей случайной величины, а  $\sigma^2$  – ее дисперсией.

**Экспоненциальное распределение.** Экспоненциальным (или показательным) называется распределение, функция распределения которого задается формулой

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{если } x > 0, \end{cases}$$

где  $\lambda > 0$  – некоторый параметр. Эта функция непрерывна, монотонно возрастает и ее предел на бесконечности равен 1, т. е. удовлетворяет всем свойствам функции распределения. Вычисленная по этой функции распределения вероятность попадания случайной величины на отрицательную полуось равна нулю, т. е. соответствующая случайная величина может принимать только положительные значения. Кроме того, данная функция распределения дифференцируема при всех  $x$  (кроме  $x = 0$ ), следовательно, распределение является абсолютно непрерывным и его плотность

$$f(x) = F'(x) = \begin{cases} 0, & \text{если } x < 0, \\ \lambda e^{-\lambda x}, & \text{если } x > 0. \end{cases}$$

Как и в случае нормального распределения, плотность убывает достаточно быстро, для того чтобы у случайной величины существовали моменты любого порядка. Используя формулу интегрирования по частям, найдем математическое ожидание:

$$\begin{aligned} M\{X\} &= \int_{-\infty}^{\infty} xf(x)dx = \int_0^{+\infty} x\lambda e^{-\lambda x}dx = \\ &= -xe^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x}dx = \frac{e^{-\lambda x}}{-\lambda}\Big|_0^{+\infty} = \frac{1}{\lambda}. \end{aligned}$$

Начальный второй момент

$$M\{X^2\} = \int_0^{+\infty} x^2\lambda e^{-\lambda x}dx = -x^2e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} 2xe^{-\lambda x}dx = \frac{2}{\lambda}M\{X\} = \frac{2}{\lambda^2}.$$

Следовательно, дисперсия

$$D\{X\} = M\{X^2\} - (M\{X\})^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

**Распределение Коши.** Распределением Коши называется распределение, функция распределения которого задается формулой

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} x.$$

Эта функция непрерывна, монотонно возрастает на всей числовой оси от 0 до 1, т. е. удовлетворяет всем свойствам функции распределения. Кроме того, данная функция распределения дифференцируема при всех  $x$ , следовательно, распределение является абсолютно непрерывным и его плотность

$$f(x) = \frac{1}{\pi} \frac{1}{x^2 + 1}.$$

Математического ожидания у этого распределения не существует, так как первообразная соответствующей подынтегральной функции

$$xf(x) = \frac{1}{\pi} \frac{x}{1 + x^2}$$

не имеет конечного предела на бесконечности.

#### 4.8. Характеристическая функция случайной величины

Ранее было показано, что распределение случайной величины полностью характеризуется либо функцией распределения, либо плотностью, либо рядом распределения. В некоторых задачах более удобным является еще

одно представление распределения – через некоторое интегральное преобразование от него, которое называется характеристической функцией.

**Определение 4.8.** *Характеристической функцией случайной величины  $X$  называется функция  $\varphi(t) = M\{e^{itX}\}$ , где  $i$  – мнимая единица.*

В частности, в случае абсолютно непрерывной случайной величины с плотностью  $f(x)$

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

Для дискретной случайной величины с рядом распределения  $\{(x_k, p_k)\}$

$$\varphi(t) = \sum_k e^{itx_k} p_k.$$

Таким образом, характеристическая функция представляет собой комплекснозначную функцию вещественного аргумента. При помощи формулы Эйлера

$$e^{iu} = \cos u + i \sin u$$

из характеристической функции легко выделить вещественную и мнимую части. Например, в случае абсолютно непрерывной случайной величины

$$\varphi(t) = \int_{-\infty}^{\infty} \cos(tx) f(x) dx + i \int_{-\infty}^{\infty} \sin(tx) f(x) dx.$$

Поскольку синус и косинус – ограниченные функции, характеристическая функция существует у любой случайной величины (соответствующий интеграл или ряд всегда сходится). Кроме того, заметим, что в случае четной функции плотности второй интеграл в последней формуле обращается в 0, т. е. характеристическая функция симметричного распределения принимает только вещественные значения.

Подставив в формулу определения характеристической функции  $t = 0$ , получим  $\varphi(0) = M\{1\} = 1$ .

Нетрудно показать, что 1 является максимальным значением не только для самой характеристической функции, но и для ее модуля. Например, в дискретном случае

$$|\varphi(t)| = \left| \sum_k e^{itx_k} p_k \right| \leq \sum_k |e^{itx_k}| p_k = \sum_k p_k = 1.$$

Если некоторая случайная величина является линейной функцией от  $X$ , т. е.  $Y = aX + b$ , то характеристическая функция величины  $Y$

$$\tilde{\varphi}(t) = M\{e^{it(aX+b)}\} = M\{e^{itaX} e^{itb}\} = e^{itb} \varphi(at).$$

Установим связь производных характеристической функции с моментами случайной величины. Пусть у случайной величины  $X$  существует начальный момент порядка  $n$ . Тогда, рассмотрев формулы вычисления характеристической функции в дискретном и непрерывном случаях, можно убедиться в допустимости  $n$ -кратного дифференцирования под знаком математического ожидания, т. е.  $\varphi^{(n)}(t) = M\{(iX)^n e^{itX}\}$ , следовательно,  $\varphi^{(n)}(0) = i^n M\{X^n\}$ .

На самом деле по характеристической функции могут быть восстановлены не только моменты, но и все распределение в целом. Например, в случае абсолютно непрерывного распределения плотность можно восстановить по характеристической функции по формуле

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

Рассмотрим некоторые примеры характеристических функций:

1. Пусть случайная величина  $X$  имеет равномерное распределение на отрезке  $[-1, 1]$ . Тогда, в силу симметричности распределения,

$$\varphi(t) = \int_{-1}^1 \frac{1}{2} \cos(tx) dx = \left( \frac{1}{2} \frac{\sin(tx)}{t} \right) \Big|_{-1}^1 = \frac{\sin t}{t}.$$

2. Пусть случайная величина  $X$  имеет стандартное нормальное распределение. Тогда

$$\begin{aligned} \varphi(t) &= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2} - \frac{t^2}{2}} dx = \\ &= e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = e^{-\frac{t^2}{2}}. \end{aligned} \quad (4.2)$$

#### 4.9. Моделирование случайных величин

Пусть  $X$  – случайная величина с функцией распределения  $F(x)$  и пусть  $h(x)$  – некоторая строго монотонно возрастающая функция. В этом случае у функции  $h(x)$  существует обратная, и нетрудно построить функцию распределения  $G(x)$  случайной величины  $Y = h(X)$ . Действительно,

$$G(x) = P(h(X) < x) = P(X < h^{-1}(x)) = F(h^{-1}(x)).$$

Это равенство можно использовать при решении задачи моделирования случайных величин с заданной функцией распределения. Существуют

датчики случайных чисел, моделирующие случайные величины, имеющие равномерное распределение на отрезке  $[0, 1]$ . Например, в качестве такой случайной величины может использоваться дробная часть числа, получающегося в результате работы некоторого вычислительного алгоритма с постоянно меняющимися входными данными. Если случайная величина  $X$  имеет равномерное распределение на  $[0, 1]$ , то, по формуле (4.1),  $F(x) = x$ , если  $x \in [0, 1]$ .

Пусть требуется смоделировать абсолютно непрерывную случайную величину, функция распределения которой  $G(x)$  является строго возрастающей и, следовательно, имеет обратную. Примем  $h(x) = G^{-1}(x)$ . Тогда  $P(h(X) < x) = F(h^{-1}(x)) = F(G(x)) = G(x)$ , так как значение  $G(x)$  находится в интервале  $[0, 1]$ , где  $F(x) = x$ .

Таким образом, чтобы смоделировать случайную величину  $Y$  с заданной функцией распределения  $G(x)$ , нужно смоделировать равномерно распределенную на  $[0, 1]$  случайную величину  $X$  и принять  $Y = G^{-1}(X)$ .

**Пример.** Пусть требуется смоделировать случайную величину, имеющую экспоненциальное распределение, т. е.  $G(x) = 1 - e^{-\lambda x}$  при  $x \geq 0$ .

Выразив в этом равенстве  $x$  через  $G(x)$ , найдем обратную функцию:

$$G^{-1}(y) = -\frac{1}{\lambda} \ln(1 - y) = \frac{1}{\lambda} \ln \left( \frac{1}{1 - y} \right).$$

Таким образом, экспоненциально распределенная случайная величина моделируется по формуле

$$Y = \frac{1}{\lambda} \ln \left( \frac{1}{1 - X} \right),$$

где  $X$  имеет равномерное распределение на  $[0, 1]$ . •

В заключение отметим, что функцию, обратную к функции распределения, не всегда удается явно выписать, как это было сделано в данном примере. В этом случае можно искать требуемое значение  $Y$  как корень уравнения  $G(Y) = X$ .

## 5. СЛУЧАЙНЫЙ ВЕКТОР

### 5.1. Случайный вектор с дискретным распределением

Пусть имеется  $n$  случайных величин, заданных на одном вероятностном пространстве. Вектор, состоящий из этих величин, называется случайным вектором, а соответствующие случайные величины – компонентами

случайного вектора. Поскольку в  $\sigma$ -алгебру событий данного вероятностного пространства входят события, состоящие в попадании каждой случайной величины в некоторое борелевское множество на прямой, то в нее входит и пересечение таких событий. Это пересечение является событием, заключающимся в попадании случайного вектора в прямое произведение соответствующих множеств. Множество таких прямых произведений порождает  $\sigma$ -алгебру борелевских множеств в пространстве  $\mathbb{R}^n$ . Итак, если задан случайный вектор, то любому борелевскому множеству в  $\mathbb{R}^n$  можно приписать вероятность попадания в него этого случайного вектора, т. е. построить вероятностное пространство, где  $\sigma$ -алгебра событий есть  $\sigma$ -алгебра борелевских множеств в  $\mathbb{R}^n$ . Другими словами, каждому случайному вектору соответствует вероятностное распределение в  $\mathbb{R}^n$ .

Заметим, что только по распределениям отдельных компонент нельзя восстановить распределение случайного вектора, так как остаются не описанными связи между его компонентами. С другой стороны, по распределению случайного вектора можно восстановить распределение каждой его компоненты. (Событие „одна компонента попадает в заданный интервал, а остальные принимают произвольные значения“ принадлежит  $\sigma$ -алгебре событий исходного вероятностного пространства.)

Говорят, что случайный вектор имеет дискретное распределение, если множество его значений является дискретным (конечным или счетным). Очевидно, что все компоненты случайного вектора с дискретным распределением также имеют дискретное распределение. Естественно, верно и обратное: если множества всех значений компонент вектора дискретны, то и множество значений самого вектора дискретно.

Для того чтобы полностью описать дискретное распределение случайного вектора, нужно указать все значения, которые он может принимать, и вероятности, с которыми эти значения принимаются, т. е. построить многомерный аналог ряда распределения. Естественно, сумма всех заданных вероятностей должна быть равна 1. По дискретному распределению случайного вектора может быть восстановлен ряд распределения каждой его компоненты. Выпишем конкретные формулы для случая двухмерного случайного вектора  $(X_1, X_2)$ . Распределение такого вектора удобно задавать в виде таблицы:

$x_i$	$y_i$				
	$y_1$	$y_2$	$\dots$	$y_m$	$\dots$
$x_1$	$s_{11}$	$s_{12}$	$\dots$	$s_{1m}$	$\dots$
$x_2$	$s_{21}$	$s_{22}$	$\dots$	$s_{2m}$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$x_n$	$s_{n1}$	$s_{n2}$	$\dots$	$s_{nm}$	$\dots$

Каждая из строк этой таблицы соответствует одному значению первой

компоненты  $x_i$ , а каждый из столбцов – одному значению второй компоненты  $y_j$ . Соответственно, в самой таблице записываются вероятности

$$s_{ij} = P(X_1 = x_i; X_2 = y_j).$$

При суммировании столбцов этой таблицы получаются вероятности

$$p_i = P(X_1 = x_i) = \sum_j s_{ij},$$

которые вместе с значениями  $x_i$  определяют ряд распределения случайной величины  $X_1$ . При суммировании строк этой таблицы получаются вероятности

$$q_j = P(X_2 = y_j) = \sum_i s_{ij},$$

которые вместе с значениями  $y_j$  определяют ряд распределения случайной величины  $X_2$ .

**Пример.** Кубик (игральная кость) подбрасывается 2 раза. При каждом бросании фиксируется число очков, выпавших на верхней грани. По этому эксперименту естественным образом строится вероятностное пространство с 36 равновероятными элементарными исходами, каждому из которых соответствует упорядоченная пара выпавших чисел (от 1 до 6). Рассмотрим на этом вероятностном пространстве случайный вектор, состоящий из случайных величин  $X_1$  – количества выпаданий „1“ и  $X_2$  – количества выпаданий „2“. Найдем распределение этого случайного вектора. Возьмем, например, событие  $\{X_1 = 1; X_2 = 0\}$ , что означает „1 выпало ровно 1 раз, а 2 не выпало ни разу“. Это событие содержит все элементарные исходы вида  $(1, k)$  и  $(k, 1)$ , где  $k$  принимает все значения, кроме 2 и 1. Таких исходов 8, следовательно, вероятность данного события равна  $8/36$ . Аналогичным образом вычисляются вероятности для остальных возможных пар значений  $X_1$  и  $X_2$  и в результате заполняется таблица, задающая распределение случайного вектора  $(X_1, X_2)$ .

$x_i$	$y_i$		
	0	1	2
0	$\frac{16}{36}$	$\frac{8}{36}$	$\frac{1}{36}$
1	$\frac{8}{36}$	$\frac{2}{36}$	0
2	$\frac{1}{36}$	0	0

Сложив вероятности в каждом столбце этой таблицы, получим ряд распределения случайной величины  $X_1$ :

$x_i$	0	1	2
$p_i$	$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$

Соответственно, сложив вероятности в строках, получим такой же ряд распределения для  $X_2$ . •

## 5.2. Случайный вектор с абсолютно непрерывным распределением

Говорят, что случайный вектор имеет абсолютно непрерывное распределение, если существует функция  $n$  переменных  $f(x_1, \dots, x_n)$ , такая, что для любого борелевского множества  $A$  из  $\mathbb{R}^n$

$$P(\bar{X} \in A) = \int_A f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Функция  $f(x_1, \dots, x_n)$  называется плотностью распределения случайного вектора. Она обладает следующими свойствами:  $f(x_1, \dots, x_n) \geq 0$  для любого вектора  $(x_1, \dots, x_n) \in \mathbb{R}^n$ ;  $\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$ . Эти свойства легко получаются из аксиом вероятности.

Таким образом, случайному вектору с абсолютно непрерывным распределением соответствует абсолютно непрерывное вероятностное пространство.

Если случайный вектор имеет абсолютно непрерывное распределение, то все его компоненты также имеют абсолютно непрерывные распределения, причем соответствующие плотности могут быть записаны как функции одной переменной, получающиеся интегрированием по остальным переменным по пространству  $\mathbb{R}^{(n-1)}$  плотности распределения случайного вектора. (При этом номер оставшейся переменной равен номеру соответствующей компоненты случайного вектора.) Докажем этот факт для случая  $n = 2$ .

Пусть  $f(x_1, x_2)$  – плотность распределения случайного вектора  $(X_1, X_2)$  и пусть  $[a, b]$  – произвольный интервал. Тогда

$$\begin{aligned} P(X_1 \in [a, b]) &= P((X_1, X_2) \in [a, b] \times \mathbb{R}^1) = \\ &= \int_a^b \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = \int_a^b \left( \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \right) dx_1. \end{aligned}$$

Если ввести функцию

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2, \quad (5.1)$$

то для любого интервала  $[a, b[$

$$P(X_1 \in [a, b[) = \int_a^b f_1(x_1) dx_1,$$

следовательно,  $f_1(x_1)$  является плотностью распределения случайной величины  $X_1$ .

Обратное неверно: распределение случайного вектора, состоящего из компонент, имеющих абсолютно непрерывное распределение, не только не восстанавливается по распределениям компонент, но даже может не быть абсолютно непрерывным. Примером этого может служить случайный вектор, состоящий из двух тождественно равных случайных величин, распределение которого сосредоточено на прямой  $x_1 = x_2$  и поэтому не может иметь плотности, заданной на  $\mathbb{R}^2$ .

### 5.3. Функция распределения случайного вектора

Функцией распределения случайного вектора  $(X_1, \dots, X_n)$  называется функция  $n$  переменных, определяемая формулой

$$F(x_1, \dots, x_n) = P(X_1 < x_1; X_2 < x_2; \dots; X_n < x_n).$$

При  $n = 1$  эта функция является функцией распределения случайной величины, а в общем случае обладает сходными с ней свойствами. В частности, она принимает значения только из отрезка  $[0, 1]$ , является неубывающей функцией по любой из переменных, стремится к 0, если все переменные стремятся к  $-\infty$ , и стремится к 1, если все переменные стремятся к  $+\infty$ .

По функции распределения случайного вектора легко восстанавливаются функции распределения всех его компонент. Для этого следует считать, что остальные компоненты могут принимать произвольные значения, т. е. в формуле определения функции распределения устремить все переменные, кроме той, которая соответствует выбранной компоненте, к  $+\infty$ .

По функции распределения случайного вектора вычисляются вероятности его попадания в любой „прямоугольный параллелепипед“ (прямое произведение отрезков). Посмотрим, как это делается в случае  $n = 2$ . Из рисунка видно, что

$$P((X_1, X_2) \in [a, b[ \times [c, d]) =$$

$$= P(X_1 < b; X_2 < d) - P(X_1 < a; X_2 < d) - P(X_1 < b; X_2 < c) + P(X_1 < a; X_2 < c) = \\ = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Если распределение случайного вектора является абсолютно непрерывным, то плотность распределения есть смешанная частная производная функции распределения по всем переменным:

$$f(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F(x_1, \dots, x_n).$$

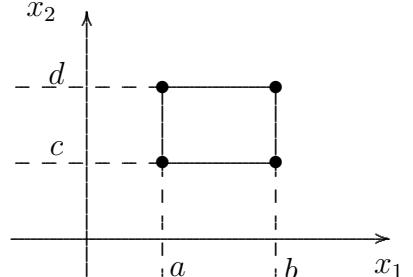


Рис. 5.1

**Пример (равномерное распределение на прямоугольнике).** Пусть плотность распределения случайного вектора  $(X_1, X_2)$  имеет вид

$$f(x_1, x_2) = \begin{cases} \frac{1}{(b-a)(d-c)}, & \text{если } (x_1, x_2) \in [a, b] \times [c, d]; \\ 0, & \text{если } (x_1, x_2) \notin [a, b] \times [c, d]. \end{cases}$$

Нетрудно проверить, что  $f(x_1, x_2)$  удовлетворяет всем свойствам плотности двухмерного случайного вектора.

Ей соответствует функция распределения

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u_1, u_2) du_1 du_2 = \\ = \begin{cases} 0, & \text{если } x_1 < a \text{ или } x_2 < c, \\ \frac{(x_1 - a)(x_2 - c)}{(b - a)(d - c)}, & \text{если } x_1 \in [a, b] \text{ и } x_2 \in [c, d], \\ \frac{x_1 - a}{b - a}, & \text{если } x_1 \in [a, b] \text{ и } x_2 > d, \\ \frac{x_2 - c}{d - c}, & \text{если } x_1 > b \text{ и } x_2 \in [c, d], \\ 1, & \text{если } x_1 > b \text{ и } x_2 > d. \end{cases}$$

Нетрудно проверить, что эта функция удовлетворяет всем свойствам функции распределения случайного вектора и

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

При  $x_2 \rightarrow +\infty$  функция  $F(x_1, x_2)$  становится функцией распределения  $X_1$ , которая является функцией равномерного распределения на  $[a, b]$ . Аналогично, случайная величина  $X_2$  имеет равномерное распределение на  $[c, d]$ .

#### 5.4. Числовые характеристики случайного вектора

Пусть известно распределение некоторого случайного вектора  $\bar{X} = (X_1, \dots, X_n)$ . Введем случайную величину  $g(\bar{X})$ , где  $g(\bar{x}) = g(x_1, \dots, x_n)$  – некоторая функция  $n$  переменных. Математическое ожидание этой величины может рассматриваться как числовая характеристика случайного вектора. Для вычисления числовых характеристик в случае дискретного или абсолютно непрерывного распределения применяются формулы, аналогичные соответствующим формулам в одномерном случае. В случае дискретного распределения

$$M\{g(X_1, \dots, X_n)\} = \sum_k g(x_1^{(k)}, \dots, x_n^{(k)})p_k,$$

где  $p_k = P(X_1 = x_1^{(k)}; \dots; X_n = x_n^{(k)})$ .

В случае абсолютно непрерывного распределения

$$M\{g(X_1, \dots, X_n)\} = \int_{\mathbb{R}^n} g(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

где  $f(x_1, \dots, x_n)$  – плотность распределения случайного вектора.

Часто рассматриваются наборы однотипных числовых характеристик. Так, математическим ожиданием случайного вектора называется вектор, составленный из математических ожиданий его компонент. (Одна компонента тоже может рассматриваться как функция вектора  $g(x_1, \dots, x_n) = x_i$ , поэтому математическое ожидание каждой компоненты случайного вектора подходит под данное определение числовой характеристики.)

Используя понятие распределения случайного вектора, нетрудно получить еще одно свойство математического ожидания.

**Теорема 5.1.** *Пусть  $X_1$  и  $X_2$  – любые случайные величины с конечными математическими ожиданиями. Тогда*

$$M\{X_1 + X_2\} = M\{X_1\} + M\{X_2\}.$$

Рассмотрим случайный вектор  $(X_1, X_2)$ . При доказательстве теоремы ограничимся случаем, когда этот вектор имеет абсолютно непрерывное распределение. Пусть  $f(x_1, x_2)$  – плотность распределения данного вектора.

Тогда

$$\begin{aligned}
M\{X_1 + X_2\} &= \int_{\mathbb{R}^2} (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 = \\
&= \int_{-\infty}^{\infty} x_1 \left( \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \right) dx_1 + \int_{-\infty}^{\infty} \left( x_2 \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 \right) dx_2 = \\
&= \int_{-\infty}^{\infty} x_1 f_1(x_1) dx_1 + \int_{-\infty}^{\infty} x_2 f_2(x_2) dx_2,
\end{aligned}$$

где, согласно формуле вычисления одномерной плотности (5.1), функции  $f_1(x_1)$  и  $f_2(x_2)$  являются плотностями распределений случайных величин  $X_1$  и  $X_2$ . Отсюда, пользуясь формулой вычисления математического ожидания абсолютно непрерывной случайной величины, получим утверждение теоремы.

**Определение 5.1.** *Ковариационным моментом компонент случайного вектора (случайных величин  $X_i$  и  $X_j$ ) называется числовая характеристика*

$$\text{cov}(X_i, X_j) = M\{(X_i - M\{X_i\})(X_j - M\{X_j\})\}.$$

*Матрица  $R$ , составленная из элементов  $r_{ij} = \text{cov}(X_i, X_j)$ , где  $X_i$  и  $X_j$  – компоненты случайного вектора, называется ковариационной матрицей случайного вектора.*

Если в определении ковариационного момента раскрыть скобки под знаком математического ожидания и воспользоваться свойствами математического ожидания, то получится другая формула для вычисления ковариационного момента:

$$\text{cov}(X_i, X_j) = M\{X_i X_j\} - M\{X_i\} M\{X_j\}. \quad (5.2)$$

В силу очевидного равенства  $r_{ij} = r_{ji}$  ковариационная матрица является симметричной. Заметим также, что при  $i = j$  ковариационный момент становится дисперсией случайной величины  $X_i$ . Таким образом, на главной диагонали ковариационной матрицы стоят дисперсии соответствующих компонент случайного вектора. Кроме того, по ковариационной матрице может быть вычислена дисперсия любой линейной комбинации компонент случайного вектора. Действительно, пусть случайная величина

$$Z = \sum_{i=1}^n a_i X_i.$$

Обозначим для краткости записи  $Y_i = X_i - M\{X_i\}$ . Тогда

$$\begin{aligned}
D\{Z\} &= M\{(Z - M\{Z\})^2\} = M\left\{\left(\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i M\{x_i\}\right)^2\right\} = \\
&= M\left\{\left(\sum_{i=1}^n a_i Y_i\right)^2\right\} = M\left\{\sum_{i=1}^n \sum_{j=1}^n a_i a_j Y_i Y_j\right\} = \\
&= \sum_{i=1}^n \sum_{j=1}^n a_i a_j M\{Y_i Y_j\} = \sum_{i=1}^n \sum_{j=1}^n r_{ij} a_i a_j.
\end{aligned}$$

Полученное равенство может быть записано также в матричной форме:

$$D\{(\bar{a}, \bar{X})\} = (R\bar{a}, \bar{a}), \quad (5.3)$$

где  $\bar{a}$  – постоянный вектор;  $(\bar{a}, \bar{X})$  – скалярное произведение вектора  $\bar{a}$  и случайного вектора  $\bar{X}$ .

Поскольку дисперсия неотрицательна, из последнего равенства следует, что  $(R\bar{a}, \bar{a}) \geq 0$  для любого вектора  $\bar{a}$ . Матрицы, удовлетворяющие этому свойству, называются неотрицательно-определенными. (Все собственные числа таких матриц неотрицательны.) Таким образом, ковариационная матрица является симметричной неотрицательно-определенной матрицей.

Ковариационные моменты характеризуют определенную связь между компонентами случайного вектора. Их размерность равна произведению размерностей соответствующих случайных величин. Безразмерной характеристикой связи двух случайных величин является коэффициент корреляции

$$\rho_{ij} = \frac{\text{cov}(X_i, X_j)}{\sqrt{D\{X_i\}D\{X_j\}}}.$$

Если случайные величины линейно связаны (одна из них является линейной функцией другой), то их коэффициент корреляции по модулю равен единице. Действительно, пусть  $X_2 = aX_1 + b$ . Тогда

$$\begin{aligned}
\text{cov}(X_1, X_2) &= M\{(aX_1 + b - aM\{X_1\} - b)(X_1 - M\{X_1\})\} = \\
&= M\{a(X_1 - M\{X_1\})^2\} = aD\{X_1\}.
\end{aligned}$$

С другой стороны,  $D\{aX_1 + b\} = a^2 D\{X_1\}$ . Таким образом,

$$\rho_{12} = \frac{aD\{X_1\}}{\sqrt{a^2 D\{X_1\} D\{X_1\}}} = \frac{a}{|a|}.$$

Итак, в случае линейно связанных величин коэффициент корреляции равен 1, если эта зависимость прямая ( $a > 0$ ), и равен  $-1$  в случае обратной

зависимости ( $a < 0$ ). На самом деле эти значения являются крайними: справедливо неравенство  $|\rho_{ij}| \leq 1$ .

Для доказательства этого неравенства рассмотрим случайный вектор, составленный из двух случайных величин  $(X_1, X_2)$ . В силу неотрицательной определенности ковариационной матрицы этого вектора ее определитель (который равен произведению собственных чисел) является неотрицательным, т. е.  $D\{X_1\}D\{X_2\} - (\text{cov}(X_1, X_2))^2 \geq 0$ .

Таким образом,  $|\text{cov}(X_1, X_2)| \leq \sqrt{D\{X_1\}D\{X_2\}}$ , откуда следует требуемое неравенство для коэффициента корреляции.

## 5.5. Независимость случайных величин

**Определение 5.2.** Случайные величины  $X_1$  и  $X_2$  называются независимыми, если независимыми являются любые события вида  $\{X_1 \in A_1\}$  и  $\{X_2 \in A_2\}$ , где  $A_1$  и  $A_2$  – произвольные борелевские множества на прямой.

Напомним, что независимость событий означает

$$P(X_1 \in A_1; X_2 \in A_2) = P(X_1 \in A_1) P(X_2 \in A_2).$$

В частности, для дискретных случайных величин

$$P(X_1 = x_i; X_2 = y_j) = P(X_1 = x_i) P(X_2 = y_j).$$

Взяв в качестве  $A_1$  и  $A_2$  интервалы  $]-\infty, x_1[$  и  $]-\infty, x_2[$ , получим, что функция распределения случайного вектора  $(X_1, X_2)$  равна произведению функций распределения входящих в него независимых случайных величин:  $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ .

Если распределения независимых случайных величин являются абсолютно непрерывными, то, проинтегрировав последнее равенство по переменным  $x_1$  и  $x_2$ , получим соответствующую связь плотности распределения случайного вектора с плотностями компонент:  $f(x_1, x_2) = f_1(x_1)f_2(x_2)$ .

Полученные формулы разложения в произведение для функций распределения или плотностей являются не только необходимыми, но и достаточными условиями независимости: при их выполнении для всех значений переменных соответствующие случайные величины являются независимыми.

Покажем теперь, что математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий. Доказательство проведем для абсолютно непрерывных случайных

величин. Итак, пусть  $X_1$  и  $X_2$  – независимые случайные величины. Тогда

$$\begin{aligned} M\{X_1X_2\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1x_2f(x_1, x_2)dx_1dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1x_2f(x_1)f(x_2)dx_1dx_2 = \\ &= \int_{-\infty}^{\infty} x_1f(x_1)dx_1 \int_{-\infty}^{\infty} x_2f(x_2)dx_2 = M\{X_1\}M\{X_2\}. \end{aligned}$$

Из этой формулы вытекает одно важное следствие: ковариационный момент независимых случайных величин равен 0. Действительно, согласно формуле (5.2), ковариационный момент

$$\begin{aligned} \text{cov}(X_1, X_2) &= M\{X_1X_2\} - M\{X_1\}M\{X_2\} = \\ &= M\{X_1\}M\{X_2\} - M\{X_1\}M\{X_2\} = 0. \end{aligned}$$

**Определение 5.3.** Случайные величины называются некоррелированными, если их ковариационный момент (и, соответственно, коэффициент корреляции) равен нулю.

Таким образом, из независимости случайных величин следует их некоррелированность. (Обратное в общем случае неверно.)

В 5.4 (теорема 5.1) было показано, что математическое ожидание суммы случайных величин равно сумме их математических ожиданий. В случае некоррелированных (в частности, независимых) случайных величин аналогичная формула справедлива и для дисперсий.

**Теорема 5.2.** Пусть  $X_1$  и  $X_2$  – некоррелированные случайные величины. Тогда

$$D\{X_1 + X_2\} = D\{X_1\} + D\{X_2\}.$$

Для доказательства этой теоремы применим к дисперсии суммы двух случайных величин формулу (5.3):

$$\begin{aligned} D\{X_1 + X_2\} &= r_{11} \cdot 1 \cdot 1 + r_{12} \cdot 1 \cdot 1 + r_{12} \cdot 1 \cdot 1 + r_{22} \cdot 1 \cdot 1 = \\ &= D\{X_1\} + 2r_{12} + D\{X_2\}. \end{aligned}$$

Отсюда вследствие некоррелированности случайных величин ( $r_{12} = 0$ ) получаем требуемую формулу.

Если известны распределения независимых случайных величин, то по ним могут быть определены не только числовые характеристики, но и распределение их суммы. Проще всего это сделать в терминах характеристических функций. Прежде всего заметим, что из определения независимости

случайных величин следует, что функции от них также являются независимыми случайными величинами. Итак, пусть  $\varphi_1(t)$  и  $\varphi_2(t)$  – характеристические функции независимых случайных величин  $X_1$  и  $X_2$ . Тогда, используя формулу математического ожидания произведения независимых случайных величин, получим характеристическую функцию их суммы:

$$\begin{aligned}\varphi(t) &= M \left\{ e^{it(X_1+X_2)} \right\} = M \left\{ e^{itX_1} e^{itX_2} \right\} = \\ &= M \left\{ e^{itX_1} \right\} M \left\{ e^{itX_2} \right\} = \varphi_1(t)\varphi_2(t),\end{aligned}$$

т. е. характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций.

Рассмотрим для примера характеристическую функцию суммы двух независимых случайных величин, имеющих стандартное нормальное распределение. Согласно формуле (4.2)

$$\varphi(t) = e^{-\frac{t^2}{2}} e^{-\frac{t^2}{2}} = e^{-t^2} = e^{-\frac{(\sqrt{2}t)^2}{2}},$$

т. е. характеристическая функция суммы двух независимых случайных величин  $X_1$  и  $X_2$ , имеющих стандартное нормальное распределение, равна характеристической функции случайной величины  $\sqrt{2}X_1$ . Этот факт нетрудно обобщить на случай произвольных нормально распределенных случайных величин и переформулировать в терминах распределений: сумма независимых нормально распределенных случайных величин имеет нормальное распределение. В частности, в случае нормального распределения для независимых одинаково распределенных случайных величин  $X_1$  и  $X_2$  существуют такие константы  $a$  и  $b$ , что  $a(X_1 + X_2) + b$  имеет то же распределение, что и  $X_1$ . Это свойство называется свойством устойчивости распределения. Уникальность нормального распределения заключается в том, что оно является единственным распределением с конечной дисперсией, обладающим свойством устойчивости.

Введенное понятие независимости двух случайных величин нетрудно обобщить для любого числа случайных величин: соответствующие события, заключающиеся в попадании каждой случайной величины в некоторое множество на прямой, должны быть независимыми в совокупности. Формулы произведения функций распределения и плотностей, а также формулы для дисперсии и характеристической функции суммы независимых случайных величин естественным образом обобщаются для любого числа независимых случайных величин.

Формулы для математического ожидания и дисперсии суммы случайных величин дают возможность простым способом определить эти числовые характеристики для случайной величины, имеющей биномиальное

распределение (см. 4.6). Определим для каждого из независимых испытаний Бернулли случайную величину:

$$Y_i = \begin{cases} 1, & \text{если в } i\text{-м испытании наступил исход } \omega_1; \\ 0 & \text{в противном случае.} \end{cases}$$

Это – двуточечные случайные величины, принимающие значение 1 с вероятностью  $p$ . Они, как было показано в 4.6, имеют  $M\{Y_i\} = p$  и  $D\{Y_i\} = p(1 - p)$ . Кроме того, эти случайные величины определяются по результатам независимых испытаний и поэтому независимы. Случайная величина  $X_n$  с биномиальным распределением равна числу наступлений события  $\omega_1$  в последовательности  $n$  испытаний Бернулли, т. е. количеству единиц в последовательности значений случайных величин  $Y_i$ . Таким образом,  $X_n = \sum_{i=1}^n Y_i$ , поэтому

$$\begin{aligned} M\{X_n\} &= M\left\{\sum_{i=1}^n Y_i\right\} = \sum_{i=1}^n M\{Y_i\} = np; \\ D\{X_n\} &= D\left\{\sum_{i=1}^n Y_i\right\} = \sum_{i=1}^n D\{Y_i\} = np(1 - p). \end{aligned}$$

Например, если  $X_n$  – число выпадений герба при 100 бросаниях симметричной монеты ( $p = 0.5$ ), то  $M\{X_n\} = 50$  и  $D\{X_n\} = 25$ .

## 5.6. Многомерное нормальное распределение

Пусть  $\bar{X}$  – случайный вектор с абсолютно непрерывным распределением в  $\mathbb{R}^n$ . Говорят, что этот вектор имеет  $n$ -мерное нормальное распределение, если он имеет математическое ожидание  $\bar{m}$ , невырожденную ковариационную матрицу  $R$  и плотность его распределения имеет вид

$$f(\bar{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det R}} e^{-\frac{1}{2}(\bar{x}-\bar{m})^T R^{-1}(\bar{x}-\bar{m})}.$$

Здесь  $R^{-1}$  – матрица, обратная к  $R$ ;  $\det R$  – определитель матрицы  $R$ . В показателе экспоненты записана квадратичная форма от компонент вектора  $\bar{x} - \bar{m}$ , построенная по симметричной матрице  $R^{-1}$ .

Прежде всего заметим, что при  $n = 1$  ковариационная матрица и ее определитель становятся дисперсией, обратная матрица – обратным числом, и получаем обычное нормальное распределение общего вида, определенное в 4.6.

Рассмотрим теперь случай, когда компоненты случайного вектора некоррелированы, т. е. ковариационная матрица – диагональная. Ее диагональные элементы, которые являются дисперсиями компонент случайного вектора, обозначим через  $\sigma_i^2$ . Обратная матрица в этом случае также будет диагональной, с числами, обратными дисперсиям, на диагонали. Таким образом, в случае некоррелированных компонент плотность нормального распределения

$$\begin{aligned} f(\bar{x}) &= \frac{1}{(2\pi)^{n/2} \sqrt{\sigma_1^2 \sigma_2^2 \dots \sigma_n^2}} e^{-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - m_i)^2}{\sigma_i^2}} = \\ &= \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(x_1 - m_1)^2}{2\sigma_1^2}} \dots \frac{1}{\sqrt{2\pi\sigma_n}} e^{-\frac{(x_n - m_n)^2}{2\sigma_n^2}}, \end{aligned} \quad (5.4)$$

т. е. плотность распадается на произведение  $n$  функций, каждая из которых зависит от своей переменной и является плотностью одномерного нормального распределения. Таким образом, проинтегрировав данную плотность распределения случайного вектора по всем переменным, кроме одной, получим плотность распределения одной компоненты случайного вектора, которая является одним сомножителем полученного произведения, т. е. плотность  $f(x)$  равна произведению плотностей компонент случайного вектора. Отсюда можно сделать вывод, что в случае некоррелированных компонент нормально распределенного случайного вектора эти компоненты являются также и независимыми и имеют нормальные распределения. Другими словами, в случае нормального распределения некоррелированность эквивалентна независимости.

Кроме того, можно показать, что также и в случае коррелированных компонент все компоненты случайного вектора имеют нормальное распределение. Действительно, поскольку ковариационная матрица является симметричной, существует ортогональное преобразование, приводящее ее к диагональному виду. Этому ортогональному преобразованию соответствует линейная замена переменных, а ей – линейное преобразование компонент случайного вектора. Таким образом, компоненты любого случайного вектора могут быть представлены как линейные функции от случайного вектора с некоррелированными компонентами, плотность распределения которого, в случае нормального распределения, описывается формулой (5.4). Но, как следует из рассуждений, проведенных в 5.5, любая линейная комбинация независимых нормально распределенных случайных величин также имеет нормальное распределение. Таким образом, все компоненты данного случайного вектора имеют нормальное распределение.

Рассмотрим теперь случай, когда  $n = 2$ . При некоррелированных ком-

понентах

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x_1-m_1)^2}{2\sigma_1^2} - \frac{(x_2-m_2)^2}{2\sigma_2^2}}.$$

(Как было показано ранее, к этому случаю всегда можно прийти ортогональной заменой переменных, т. е. поворотом координатных осей. Отметим только, что при таком преобразовании значения  $\sigma_1^2$  и  $\sigma_2^2$  отличаются от исходных дисперсий и представляют собой собственные числа ковариационной матрицы.)

Линии уровня этой плотности (линии, на которых плотность принимает одно и то же значение) – эллипсы с центром в точке  $(m_1, m_2)$  и полуосами, пропорциональными среднеквадратичным отклонениям  $\sigma_1$  и  $\sigma_2$ . Фигуры, ограниченные этими эллипсами, называются эллипсами рассеяния. Вычислим вероятность попадания нормально распределенного двухмерного случайного вектора в эллипс рассеяния  $B(k)$  с полуосами  $k\sigma_1$  и  $k\sigma_2$ :

$$P((X_1, X_2) \in B(k)) = \int \int_{B(k)} f(x_1, x_2) dx_1 dx_2.$$

Сделав в последнем интеграле замену переменных

$$\begin{aligned} x_1 &= m_1 + \sigma_1 \rho \cos \varphi, \\ x_2 &= m_2 + \sigma_2 \rho \sin \varphi, \end{aligned}$$

якобиан которой равен  $\sigma_1 \sigma_2 \rho$ , получим:

$$P((X_1, X_2) \in B(k)) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^k \rho e^{-\frac{\rho^2}{2}} d\rho = 1 - e^{-\frac{k^2}{2}}.$$

Так, например, вероятность попадания в эллипс рассеяния с полуосами, равными утроенным среднеквадратичным отклонениям, равна

$$1 - e^{(-9/2)} = 0.989.$$

## 6. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

### 6.1. Закон больших чисел

Рассмотрим поведение среднего арифметического независимых случайных величин. Пусть независимые случайные величины

$$X_1, X_2, \dots, X_n$$

имеют конечные дисперсии. Найдем дисперсию их среднего арифметического. Поскольку для независимых случайных величин дисперсия суммы равна сумме дисперсий, получим:

$$D \left\{ \frac{1}{n} \sum_{i=1}^n X_i \right\} = \frac{1}{n^2} \sum_{i=1}^n D\{X_i\}. \quad (6.1)$$

В частности, в случае одинаковых дисперсий  $D(X_i) = \sigma^2$  получим:

$$D \left\{ \frac{1}{n} \sum_{i=1}^n X_i \right\} = \frac{\sigma^2}{n}.$$

Таким образом, дисперсия среднего арифметического для независимых случайных величин приближается к нулю с увеличением числа слагаемых. Это означает, что с ростом числа слагаемых уменьшается разброс возможных значений среднего арифметического относительно его математического ожидания, т. е. среднее арифметическое независимых случайных величин в определенном смысле стремится к своему математическому ожиданию. Для того чтобы строго сформулировать соответствующий результат, необходимо ввести понятие сходимости для последовательности случайных величин.

**Определение 6.1.** *Последовательность случайных величин  $X_1, X_2, \dots, X_n, \dots$  сходится по вероятности к числу  $a$ , если для любого  $t \geq 0$*

$$\lim_{n \rightarrow \infty} P(|X_n - a| > t) = 0.$$

Теперь сформулируем основной результат, который называется законом больших чисел.

**Теорема 6.1.** *Пусть  $X_1, X_2, \dots, X_n, \dots$  – последовательность независимых случайных величин с одинаковыми математическими ожиданиями  $M\{X_i\} = t$  и равномерно ограниченными дисперсиями  $D[X_j] \leq C$ . Тогда последовательность средних арифметических первых  $n$  величин сходится по вероятности к математическому ожиданию  $t$ .*

**Доказательство.** Из свойств математического ожидания тривиальным образом следует, что математическое ожидание среднего арифметического равно среднему арифметическому математических ожиданий слагаемых, т. е. в данном случае числу  $t$ . Воспользовавшись неравенством Чебышева (теорема 4.1) и формулой (6.1) для дисперсии среднего арифметического, получим, что для любого  $t \geq 0$

$$P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i - t \right| > t \right) \leq \frac{D \left\{ \frac{1}{n} \sum_{i=1}^n X_i \right\}}{t^2} \leq \frac{C}{nt^2}.$$

Поскольку правая часть последнего неравенства стремится к нулю при  $n \rightarrow \infty$ , получили сходимость по вероятности последовательности средних арифметических исходных случайных величин к числу  $m$ . Теорема доказана. ■

*Замечание.* В случае одинаково распределенных случайных величин закон больших чисел остается справедливым и без предположения о конечности дисперсий, которое понадобилось для простоты доказательства. Достаточно потребовать конечности математических ожиданий.

Получим теперь одно важное следствие из закона больших чисел.

Пусть имеется последовательность независимых случайных экспериментов, в результате каждого из которых с вероятностью  $p$  может произойти событие  $A$ . Сопоставим каждому эксперименту случайную величину, принимающую значение 1, если произошло  $A$ , и равную 0 в противном случае. Все эти величины являются независимыми и имеют одно и то же двуточечное распределение с математическим ожиданием  $p$ . Сумма  $n$  таких величин представляет собой число наступлений события  $A$  при  $n$  экспериментах. Соответственно, их среднее арифметическое представляет собой частоту наступления  $A$   $v_n$ . Согласно закону больших чисел это среднее арифметическое сходится по вероятности к математическому ожиданию одной величины, т. е. для любого  $t \geq 0$

$$\lim_{n \rightarrow \infty} P(|v_n - p| > t) = 0.$$

Таким образом, закон больших чисел дает теоретическое обоснование свойства статистической устойчивости частот.

## 6.2. Центральная предельная теорема

Рассмотрим предельное поведение распределений сумм независимых одинаково распределенных случайных величин с конечными дисперсиями. Поскольку математические ожидания и дисперсии таких сумм равны суммам соответствующих характеристик, распределение самой такой суммы с ростом  $n$  будет „уходить на бесконечность“ (при ненулевом математическом ожидании) и иметь возрастающий „разброс значений“. Следовательно, сходимость распределений таких сумм возможна только при вычитании из них определенных величин и введении нормирующих множителей. Сделаем это таким образом, чтобы каждая преобразованная сумма имела стандартные числовые характеристики: нулевое математическое ожидание и единичную дисперсию. Итак, пусть  $X_1, X_2, \dots, X_n, \dots$  – последовательность независимых одинаково распределенных величин с математическими ожиданиями  $M\{X_i\} = m$  и дисперсиями  $\sigma^2$ . Построим по ней последовательность

ность нормированных сумм

$$Z_n = \frac{\sum_{i=1}^n X_i - nm}{\sigma\sqrt{n}}.$$

Используя свойства математического ожидания и дисперсии нетрудно проверить, что  $M\{Z_n\} = 0$  и  $D\{Z_n\} = 1$ .

Если случайные величины  $X_i$  имеют нормальное распределение, благодаря свойству устойчивости нормального распределения все величины  $Z_n$  будут иметь стандартное нормальное распределение. Оказывается, что в случае произвольного распределения с конечной дисперсией распределения величин  $Z_n$  с ростом  $n$  также приближаются к стандартному нормальному. Этот факт называется центральной предельной теоремой. Прежде чем точно сформулировать эту теорему, необходимо строго определить понятие сходимости распределений, или, как принято говорить, сходимости последовательности случайных величин по распределению.

**Определение 6.2.** *Последовательность случайных величин  $Y_1, Y_2, \dots, Y_n, \dots$  с функциями распределения  $F_1(x), F_2(x), \dots, F_n(x), \dots$  сходится по распределению к случайной величине  $Y_0$  с функцией распределения  $F_0(x)$ , если*

$$\lim_{n \rightarrow \infty} F_n(x) = F_0(x)$$

для всех точек  $x$ , в которых непрерывна функция  $F_0(x)$ .

Для доказательства центральной предельной теоремы понадобится следующий результат, который приведем без доказательства: сходимость по распределению случайной величины эквивалентна сходимости характеристических функций во всех точках.

Итак, центральная предельная теорема может быть сформулирована следующим образом.

**Теорема 6.2.** *Последовательность  $Z_n$  нормированных сумм независимых одинаково распределенных случайных величин с конечными дисперсиями сходится по распределению к случайной величине, имеющей стандартное нормальное распределение.*

**Доказательство.** Пусть  $\varphi(t)$  – характеристическая функция случайных величин  $Y_i = X_i - m$  и пусть  $\psi_n(t)$  – характеристические функции случайных величин  $Z_n$ . Поскольку  $X_n = \frac{\sum_{i=1}^n Y_i}{\sigma\sqrt{n}}$ , используя свойства харак-

теристических функций, полученные в 4.8 и 5.5, находим

$$\psi_n(t) = \varphi^n \left( \frac{t}{\sigma\sqrt{n}} \right).$$

Запишем для функции  $\varphi(t)$  формулу Тейлора второго порядка с центром в 0:

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2}t^2 + o(t^2).$$

Используя связь производных характеристической функции и моментов и равенства  $M\{Y_i\} = 0$  и  $M\{Y_i^2\} = \sigma^2$  получаем:

$$\varphi(t) = 1 - \frac{\sigma^2}{2}t^2 + o(t^2),$$

следовательно,

$$\psi_n(t) = \left( 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n\sigma^2}\right) \right)^n$$

и

$$\ln(\psi_n(t)) = n \ln \left( 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n\sigma^2}\right) \right).$$

Отсюда, устремляя к бесконечности число  $n$  и заменяя  $\ln(1 + x)$  на эквивалентную при  $x \rightarrow 0$  величину  $x$ , находим, что при любом  $t$

$$\lim_{n \rightarrow \infty} \ln(\psi_n(t)) = -\frac{t^2}{2},$$

откуда следует

$$\lim_{n \rightarrow \infty} \psi_n(t) = e^{-\frac{t^2}{2}}.$$

Таким образом, последовательность характеристических функций нормированных сумм  $Z_n$  сходится к функции, которая, как было показано в 4.8, является характеристической функцией случайной величины, имеющей стандартное нормальное распределение. Отсюда, благодаря эквивалентности сходимости характеристических функций и распределений, следует утверждение теоремы. ■

Итак, пусть  $F_n(x)$  – последовательность функций распределения нормированных сумм  $Z_n$ , а  $\Phi(x)$  – функция распределения стандартной нормальной случайной величины. Тогда, согласно центральной предельной теореме,

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x)$$

при любом  $x$ .

При практическом использовании центральной предельной теоремы возникает вопрос, каким должно быть число слагаемых  $n$ , чтобы распределение нормированной суммы было действительно близким к стандартному

нормальному, т. е. вопрос о скорости сходимости  $F_n(x)$  к  $\Phi(x)$ . Ответить на него помогает следующая оценка. Пусть исходные случайные величины имеют конечные трети абсолютные центральные моменты

$$\mu_3^* = M \{ |X_i - m|^3 \}.$$

Тогда при всех  $x$  и всех  $n$  справедливо неравенство

$$|F_n(x) - \Phi(x)| \leq 0.8 \frac{\mu_3^*}{\sigma^3 \sqrt{n}}.$$

### 6.3. Предельные теоремы для биномиального распределения

Пусть  $X_n$  – случайная величина, представляющая собой число наступлений события  $\omega_1$  в серии из  $n$  независимых испытаний Бернулли. Как было показано в 4.6, эта величина имеет биномиальное распределение, для нее может быть выписан ряд распределения при любых конкретных значениях  $n$  и  $p = P(\omega_1)$ . Однако при большом количестве испытаний  $n$  практическое вычисление вероятности попадания случайной величины в заданный интервал по ее ряду распределения становится очень сложным. В связи с этим были бы полезными приближенные оценки таких вероятностей.

Как было показано в 5.5, случайная величина  $X_n$  может быть представлена в виде суммы  $n$  независимых одинаково распределенных (двуточечное распределение) случайных величин с математическими ожиданиями  $p$  и дисперсиями  $p(1 - p)$ . Соответствующие нормированные суммы имеют вид

$$Z_n = \frac{X_n - np}{\sqrt{np(1 - p)}}.$$

Согласно центральной предельной теореме наблюдается сходимость последовательности распределений величин  $Z_n$  к стандартному нормальному распределению, т. е. для последовательности биномиально распределенных случайных величин  $X_n$  справедливо соотношение

$$\lim_{n \rightarrow \infty} P \left( \frac{X_n - np}{\sqrt{np(1 - p)}} < x \right) = \Phi(x)$$

при всех  $x$ , где  $\Phi(x)$  – функция распределения стандартной нормальной случайной величины. Это утверждение называется теоремой Моавра–Лапласа. (Исторически оно было получено раньше, чем центральная предельная теорема в общем случае.)

Теорема Моавра–Лапласа означает, что при больших  $n$  и при любых натуральных  $k$  и  $l$  ( $0 \leq k \leq l \leq n$ ) справедливо приближенное равенство

$$P(k \leq X_n \leq l) = P(k - 0.5 < X_n < l + 0.5) =$$

$$\begin{aligned}
&= P \left( \frac{k - 0.5 - np}{\sqrt{np(1-p)}} < Z_n < \frac{l + 0.5 - np}{\sqrt{np(1-p)}} \right) \approx \\
&\approx \Phi \left( \frac{l + 0.5 - np}{\sqrt{np(1-p)}} \right) - \Phi \left( \frac{k - 0.5 - np}{\sqrt{np(1-p)}} \right).
\end{aligned}$$

Для того чтобы установить, при каких условиях этим приближенным равенством действительно можно пользоваться, применим к данному случаю оценку скорости сходимости в центральной предельной теореме, приведенную в 6.2. Для двуточечных случайных величин  $\sigma = \sqrt{p(1-p)}$ ,

$$\begin{aligned}
\mu_3^* &= M \{ |Y_i - p|^3 \} = |0 - p|^3(1 - p) + |1 - p|^3p = \\
&= p(1 - p) (p^2 + (1 - p)^2) = p(1 - p)(1 - 2p(1 - p)).
\end{aligned}$$

Таким образом, полученным приближенным равенством можно пользоваться в случае малости величины

$$\frac{\mu_3^*}{\sigma^3 \sqrt{n}} = \frac{1 - 2p(1 - p)}{\sqrt{np(1 - p)}},$$

или, поскольку числитель последней дроби всегда принимает значения в интервале от 0.5 до 1, данное приближение применимо в случае, когда величина  $pr(1 - p)$  составляет хотя бы несколько десятков.

**Пример.** Симметричная монета подбрасывается 400 раз. Найти вероятность того, что герб выпадет не менее 180 и не более 220 раз. Здесь  $p = 0.5$ ,  $np = 200$ ,  $np(1 - p) = 100$ , следовательно,

$$P(180 \leq X_{400} \leq 220) \approx \Phi \left( \frac{20.5}{10} \right) - \Phi \left( -\frac{20.5}{10} \right) = 2\Phi(2.05) - 1. \bullet$$

Итак, в случае, когда значение вероятности  $p$  мало или близко к 1, теоремой Моавра – Лапласа нельзя пользоваться даже при достаточно больших значениях  $n$ . В этом случае, однако, можно использовать другое приближение для биномиального распределения, основанное на распределении Пуассона (см. 4.6).

**Теорема 6.3 (теорема Пуассона).** Пусть задана  $X_n$  – последовательность случайных величин с биномиальными распределениями, в которой параметр  $p$  меняется таким образом, что величина  $\lambda = np$  остается постоянной. Тогда эта последовательность сходится по распределению к случайной величине, имеющей распределение Пуассона с параметром  $\lambda$ .

Утверждение этой теоремы по существу означает, что при любом целом неотрицательном  $k$  справедливо соотношение

$$\lim_{n \rightarrow \infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

т. е. при больших значениях  $n$  в случае, когда величина произведения  $pr$  ограничивается сверху несколькими единицами, можно пользоваться соответствующим приближенным равенством, взяв в качестве параметра  $\lambda$  величину  $pr$ .

**Доказательство теоремы Пуассона.** Согласно формуле, полученной в 4.6,

$$\begin{aligned} P(X_n = k) &= \frac{n(n-1)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} = \\ &= \frac{(pn)^k}{k!} 1 \frac{n-1}{n} \dots \frac{n-k+1}{n} (1-p)^{n-k}. \end{aligned}$$

Подставив  $np = \lambda$ , получим:

$$P(X_n = k) = \frac{\lambda^k}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

Перейдя в последнем равенстве к пределу при  $n \rightarrow \infty$ , находим

$$\lim_{n \rightarrow \infty} P(X_n = k) = \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Теорема доказана. ■

В заключение заметим, что использовать теорему Пуассона можно также в случае, когда значение  $p$  близко к 1. При этом соответствующее приближенное равенство можно записывать для случайной величины  $n - X_n$ , представляющей собой число наступлений события  $\omega_2$ . Соответственно, вероятность  $p$  заменится на  $1 - p$ .

**Пример.** Пусть вероятность безошибочной передачи некоторого сообщения равна 0.99. Найти вероятность того, что при передаче 100 однотипных сообщений не менее 99 из них будут безошибочными. Здесь  $X_n$  – число безошибочных сообщений,  $p = 0.99$ ,  $n(1-p) = 1$ , следовательно, можно применить теорему Пуассона при  $\lambda = 1$  к случайной величине  $n - X_n$ . При этом получим:

$$\begin{aligned} P\{X_n \geq 99\} &= P\{n - X_n = 0\} + P\{n - X_n = 1\} \sim \\ &\sim e^{-\lambda} + \frac{\lambda}{1} e^{-\lambda} = 2e^{-1} \sim 0.74. \bullet \end{aligned}$$

## Список литературы

1. Даугавет А. И., Постников Е. В., Червинская Н. М. Теория вероятностей: Учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2001.
2. Даугавет А. И., Постников Е. В., Солынин А. А. Математическая статистика: Учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2012.
3. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики: Учеб. пособие для вузов. Спб.: Лань, 2008.
4. Кельберт М. Я., Сухов Ю. М. Вероятность и статистика в примерах и задачах. В 3 т. Т. 1: Основные понятия теории вероятностей и математической статистики. М.: Изд-во МЦНМО, 2007.

# ОГЛАВЛЕНИЕ

Введение .....	3
1. Элементарная теория вероятностей .....	3
1.1. Случайный эксперимент .....	3
1.2. Математическая модель случайного эксперимента .....	5
1.3. Операции над событиями .....	9
1.4. Теорема сложения. Вероятность противоположного события ..	10
1.5. Условная вероятность события. Условное распределение вероятностей .....	13
1.6. Формула полной вероятности. Формула Байеса .....	16
1.7. Независимые события .....	18
1.8. Независимые испытания .....	19
1.9. Геометрические вероятности .....	21
2. Элементы комбинаторики .....	22
2.1. Основная формула комбинаторики .....	22
2.2. Перестановки .....	24
2.3. Размещения .....	25
2.4. Сочетания .....	26
2.5. Размещения с повторениями .....	28
3. Аксиоматика теории вероятностей Колмогорова .....	28
3.1. Вероятностное пространство .....	28
3.2. Классификация вероятностных пространств .....	30
4. Случайные величины .....	32
4.1. Случайная величина и ее распределение .....	32
4.2. Случайная величина с дискретным распределением .....	35
4.3. Случайная величина с абсолютно непрерывным распределением .....	36
4.4. Интеграл Стильтеса .....	36
4.5. Числовые характеристики случайных величин .....	38
4.6. Примеры распределений дискретных случайных величин ..	44
4.7. Примеры распределений абсолютно непрерывных случайных величин .....	47
4.8. Характеристическая функция случайной величины .....	51
4.9. Моделирование случайных величин .....	53
5. Случайный вектор .....	54
5.1. Случайный вектор с дискретным распределением .....	54
5.2. Случайный вектор с абсолютно непрерывным распределением	57
5.3. Функция распределения случайного вектора .....	58
5.4. Числовые характеристики случайного вектора .....	60
5.5. Независимость случайных величин .....	63

5.6. Многомерное нормальное распределение .....	66
6. Предельные теоремы теории вероятностей .....	68
6.1. Закон больших чисел .....	68
6.2. Центральная предельная теорема .....	70
6.3. Предельные теоремы для биномиального распределения .....	73
Список литературы .....	76

Даугавет Александр Игоревич  
Постников Евгений Валентинович  
Червинская Нина Михайловна

ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ  
Учебное пособие

Редактор Э. К. Долгатов

---

Подписано в печать

Формат 60×84 1/16. Бумага офсетная.

Печать офсетная. Гарнитура „Times“. Печ. л. 5.

Тираж 173 экз. Заказ

---

Издательство СПбГЭТУ „ЛЭТИ“  
197376, С.-Петербург, ул. Проф. Попова, 5