А. Н. БОРОДИН

ЭЛЕМЕНТАРНЫЙ КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Издание восьмое,
стереотипное

Рекомендовано Министерством общего и профессионального образования Российской Федерации в качестве учебного пособия для студентов высших учебных заведений, обучающихся по нематематическим специальностям

ЛАНЬ®
САНКТ-ПЕТЕРБУРГ • МОСКВА • КРАСНОДАР
2011

ББК 22.17
Б 83

Бородин А. Н.

Б 83 Элементарный курс теории вероятностей и математической статистики: Учебное пособие. 8-е изд., стер. — СПб.: Издательство «Лань», 2011. — 256 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-0442-1

Учебное пособие содержит систематическое изложение основных разделов элементарного курса теории вероятностей и математической статистики. К традиционным разделам добавлен один новый — «Процедура рекуррентного оценивания», ввиду особой важности этой процедуры для приложений. Теоретический материал сопровождается большим количеством примеров и задач из разных областей знаний.

ББК 22.17

Рецензенты:
доктор физ.-матем. наук *А. Ю. ЗАЙЦЕВ*,
академик *И. А. ИБРАГИМОВ*,
профессор *Я. Ю. НИКИТИН*

Оригинал-макет подготовлен
автором в пакете $\TeX$

Обложка
С. ШАПИРО, А. ЛАПШИН

ОГЛАВЛЕНИЕ

Часть 1
ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Часть 2
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Предисловие

Изложение теории вероятностей и математической статистики сразу в строго формализованном виде на основе теории меры не способствует выработке у читателя правильных интуитивных понятий, связанных с этими дисциплинами. Существует своего рода вероятностное мышление, которое формируется поэтапно, по мере углубления в эти своеобразные разделы математики. Появление данного учебника продиктовано желанием создать краткий курс теории вероятностей и математической статистики, в котором материал был бы изложен на довольно строгом уровне и в то же время был бы понятен широкому кругу читателей. Это предъявляет жесткие требования к отбору материала. В учебник включены основополагающие результаты теории вероятностей и математической статистики, для строгого доказательства которых не требуется слишком громоздкий аналитический аппарат. Изложение сопровождается большим количеством примеров и задач, которые снабжены решениями.

В конце каждого параграфа приводится набор задач для самостоятельного решения. Задачи подобраны так, чтобы проиллюстрировать применение значительной части материала, изложенного в параграфе. В конце учебника даны ответы, а для наиболее интересных задач дано частичное или полное решение.

Учебник предназначен в основном для студентов нематематических специальностей и тех, кто решил приобрести начальные знания по теории вероятностей и математической статистике. Он также может быть полезным преподавателям элементарного курса теории вероятностей и математической статистики для систематизации материала.

Третье издание в части теоретического материала не сильно отличается от первого. Увеличено количество задач для самостоятельного решения. Устранены замеченные опечатки, в значительной степени благодаря моим коллегам, за что я им признателен. Я особенно благодарен А.Ю.Зайцеву, М.А.Козловой и В.Н.Судакову за очень полезные замечания, способствовавшие улучшению учебника.

Список обозначений

P_n – число перестановок

A_n^k – число размещений

C_n^k – число сочетаний, биномиальные коэффициенты

Ω – пространство элементарных событий ω

$\overline{A}$ – событие, дополнительное к A

$\mathbf{P}(\cdot)$ – вероятность события

$\mathbf{E}(\cdot)$ – математическое ожидание случайной величины

$\mathbf{D}(\cdot)$ – дисперсия случайной величины

$\mathrm{cov}(\cdot,\cdot)$ – ковариация случайных величин

$\Phi(x)$ – функция Лапласа

$\phi(x)$ – функция Гаусса

$\mathcal{N}(x)$ – стандартное нормальное распределение

$\mathcal{K}(x)$ – функция распределения Колмогорова

$\mathcal{P}_k(x)$ – функция распределения Пирсона с k степенями свободы

$\mathcal{S}_k(x)$ – функция распределения Стьюдента с k степенями свободы

$\mathbb{I}_A(x)$ – индикатор множества A

$\psi(z)$ – производящая функция

$\varphi(z)$ – характеристическая функция

$f * g$ – свертка функций f и g

$\overline{X}_n$ – выборочное среднее

S_n^2 – выборочная дисперсия

z_p – квантиль порядка p

$\hat{z}_p$ – выборочная квантиль порядка p

$X_{(k)}$ – k-ая порядковая статистика

$I_n(\theta)$ – информационное количество Фишера

$\sim, \approx, o(1)$ – символы, определенные на стр. 65

Часть 1

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Теория вероятностей – это математическая наука, изучающая закономерности случайных явлений.

§ 1. Элементы комбинаторики

Комбинаторикой называется область математики, в которой изучаются вопросы о том, сколько различных комбинаций, подчиненных тем или иным условиям, можно составить из заданных объектов.

Определение 1. Множество (совокупность элементов) называется **занумерованным**, если каждому элементу этого множества сопоставлено свое натуральное число (номер) от 1 до n.

В случае, когда n является конечным, говорят, что множество состоит из n элементов. Если $n = \infty$, то множество называется **счетным**.

Для краткости занумерованные множества также будут называться **наборами**.

Конечное или счетное множество называется **упорядоченным**, если каждому элементу сопоставлен его порядковый номер, т.е. номер места, на котором он расположен.

В случае, когда занумерованное множество упорядочено, следует различать идентификационный номер элемента и его порядковый номер.

Определение 2. Отличающиеся друг от друга порядком наборы, составленные из всех элементов данного конечного множества, называются **перестановками** этого множества.

Пример 1. Множество, состоящее из трех элементов $\{1, 2, 3\}$, имеет следующие перестановки: (1,2,3), (1,3,2), (2,3,1), (2,1,3), (3,2,1), (3,1,2).

Число всех перестановок множества из n элементов обозначается P_n.

Теорема 1 (о числе перестановок). *Число перестановок множества из n элементов определяется по формуле $P_n = n!$, где $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$.*

Доказательство. Отведем для размещения элементов данного набора n занумерованных ячеек (мест). Первое место может занимать элемент с любым из номеров $1, 2, \ldots, n$. Пусть, например, это элемент с номером 1. На остальных $n - 1$ местах могут стоять элементы с различными наборами номеров из чисел $2, 3, \ldots, n$, отличающиеся друг от друга лишь порядком. Число возможных таких наборов равно числу перестановок P_{n-1}. Таким образом, число различных наборов, состоящих из n элементов, у которых на первом месте стоит элемент с номером 1, равно P_{n-1}. Случай, когда на первом месте стоит элемент с номером 1, ничем не отличается от случая, когда на первом месте стоит элемент с любым другим номером. Следовательно, число различных наборов с фиксированным первым элементом равно P_{n-1}. Поскольку на первое место можно поместить n различных элементов с номерами от 1 до n, то число всех отличающихся друг от друга наборов множества из n элементов будет nP_{n-1}, т. е. $P_n = nP_{n-1}$. Это равенство справедливо для любого n, поэтому можно написать цепочку равенств

$$P_n = nP_{n-1} = n(n-1)P_{n-2} = \\ = n(n-1)(n-2)P_{n-3} = \cdots = n(n-1)(n-2)\ldots 2P_1.$$

Остается лишь заметить, что $P_1 = 1$.

Задача 1. Сколькими различными маршрутами можно разнести корреспонденцию в 5 адресов?

Решение. Занумеруем адреса цифрами от 1 до 5. Каждому маршруту можно сопоставить один из наборов, состоящих из этих пяти цифр, например, (2,5,3,4,1). Такой набор означает, что сначала выбирается второй адрес, затем пятый, третий, четвертый и первый. Всего различных маршрутов, т. е. отличающихся порядком наборов пяти цифр будет $5! = 120$.

Задача 2. Цифры 0,1,2,3 написаны на четырех карточках. Сколько различных четырехзначных чисел можно составить из этих карточек?

Решение. Число различных комбинаций из четырех цифр равно 4!. Не все эти комбинации являются четырехзначными числами, т. к. есть комбинации, начинающиеся с нуля. Таких комбинаций будет 3!, и их нужно исключить В результате число различных четырехзначных чисел равно $4! - 3! = 18$.

Определение 3. Упорядоченные наборы, состоящие из k различных элементов, выбранных из данных n элементов, называются **размещениями** из n элементов по k.

Размещения могут отличаться друг от друга как элементами, так и порядком.

Пример 2. Различными размещениями множества из трех элементов $\{1,2,3\}$ по два будут наборы (1,2), (2,1), (1,3), (3,1), (2,3), (3,2).

Число всех размещений из n элементов по k обозначается A_n^k. При $k = n$ число размещений совпадает с числом перестановок.

Теорема 2 (о числе размещений). *Число размещений из n элементов по k определяется по формуле*

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

Доказательство. Отведем для элементов размещения k занумерованных мест. Поскольку размещения упорядочены, то у каждого элемента есть номер места, на котором он расположен: 1-й, 2-й, ... или k-й. На первом месте в размещении может стоять элемент с любым из номеров $1, 2, \dots, n$. Поставим на первое место, например, элемент с номером 1. Для того, чтобы получить различные размещения из n элементов по k, у которых единица стоит на первом месте, нужно из оставшихся $n-1$ элементов составить всевозможные размещения по $k-1$ элементу и поставить их на свободные $k-1$ мест. Следовательно, число различных размещений с единицей на первом месте будет равно A_{n-1}^{k-1}. На первое место можно поставить также любой из элементов 2,3,. . или n, и каждый раз число различных размещений с фиксированным первым элементом будет равно A_{n-1}^{k-1}. Так как на первое место можно поместить n различных элементов, то число всех размещений из n элементов по k будет nA_{n-1}^{k-1}, т. е.

$A_n^k = nA_{n-1}^{k-1}$. Это равенство верно для любых k, n таких, что $k \leqslant n$. Используя этот факт и очевидное соотношение $A_{n-k+1}^1 = n - k + 1$, получим

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

Задача 3. Студентам надо сдать 4 экзамена за 8 дней. Сколькими способами можно составить расписание сдачи экзаменов?

Решение. Занумеруем дни сдачи экзаменов цифрами 1,2,...,8. Составлять различные расписания можно следующим образом. Сначала выберем дни для сдачи экзаменов, например, (2,4,5,7), а затем порядок сдачи экзаменов. Таким образом, нужно составить различные наборы четырех чисел из восьми, которые отличаются друг от друга не только элементами, но и порядком. Таких наборов $A_8^4 = 8 \cdot 7 \cdot 6 \cdot 5 = 1680$.

Определение 4. Неупорядоченные наборы, состоящие из k элементов, взятых из данных n элементов, называются **сочетаниями** из n элементов по k.

Сочетания отличаются друг от друга только элементами.

Пример 3. Для множества $\{1, 2, 3\}$ сочетаниями по 2 элемента являются $\{1, 2\}, \{1, 3\}, \{2, 3\}$.

Число всех сочетаний из n элементов по k обозначается C_n^k.

Теорема 3 (о числе сочетаний). *Число сочетаний из n элементов по k определяется по формуле*

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

Доказательство. Формулу для числа сочетаний проще всего вывести, основываясь на формулах для числа размещений и перестановок. Для получения этой формулы зададимся вопросом о том, как можно образовать различные размещения из n элементов по k. Можно составить различные сочетания из n элементов по k, а потом в каждом из сочетаний различными способами поменять порядок. Таким образом, чтобы получить всевозможные размещения, нужно для каждого из C_n^k сочетаний осуществить $k!$ перестановок. Следовательно, $A_n^k = C_n^k \cdot k!$. Отсюда получаем $C_n^k = A_n^k/k! = \frac{n!}{k!(n-k)!}$.

Задача 4. В хоккейном турнире участвуют 6 команд. Каждая команда должна сыграть с каждой одну игру. Сколько игр сыграно в турнире?

Решение. Различные пары команд образуют сочетания из 6 по 2, поскольку порядок среди двух команд, играющих в одной игре, нам безразличен. Следовательно, число игр будет равно $C_6^2 = \frac{6!}{2!4!} = 15$.

Задача 5. Имеется прямоугольник, разбитый на клетки. При этом вдоль одной стороны n клеток, а вдоль другой – m.

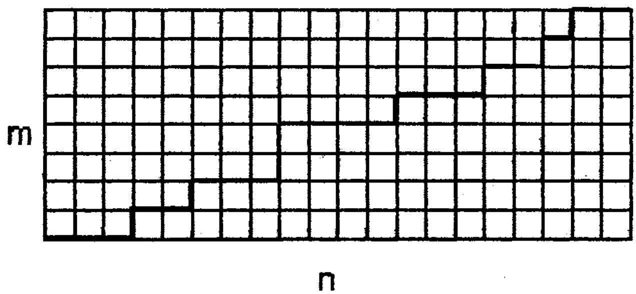

Рис. 1

Пусть можно двигаться только по сторонам клеток либо вправо, либо вверх. Сколько существует различных путей из левого нижнего угла в правый верхний?

Решение. Сопоставим ходам вдоль клеток цифры 0 и 1. При этом 0 означает, что мы идем вправо, а 1 – вверх. Тогда каждому пути соответствует набор из $m + n$ цифр, например, 00110...10, в котором будет ровно n нулей и m единиц. Сколько таких различных наборов цифр? Всего в таком наборе имеется $m+n$ позиций, и надо среди них разместить m единиц, а остальные места оставить для нулей. Выбрать m позиций среди $m + n$ можно C_{m+n}^m способами. Столько существует различных путей.

Коэффициенты C_n^k называются **биномиальными коэффициентами**, так как они входят в формулу **бинома Ньютона**

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

Свойства коэффициентов C_n^k:

$$1)\ \sum_{k=0}^{n} C_n^k = 2^n; \qquad 2)\ C_n^k = C_n^{n-k}; \qquad 3)\ C_n^k = C_{n-1}^{k-1} + C_{n-1}^k,$$

где полагаем $0! = 1$ и $C_n^0 = C_n^n = 1$.

Первое свойство следует из формулы бинома Ньютона при $a = 1, b = 1$. Второе свойство непосредственно вытекает из формулы для числа сочетаний C_n^k. Третье свойство можно проверить следующим образом:

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

Третье свойство позволяет последовательно вычислять биномиальные коэффициенты C_n^k с помощью так называемого **треугольника Паскаля**:

$$\begin{array}{ccccccccccc}
&&&&& C_0^0 &&&&& \\
&&&& C_1^0 && C_1^1 &&&& \\
&&& C_2^0 && C_2^1 && C_2^2 &&& \\
&& C_3^0 && C_3^1 && C_3^2 && C_3^3 && \\
& C_4^0 && C_4^1 && C_4^2 && C_4^3 && C_4^4 & \\
C_5^0 && C_5^1 && C_5^2 && C_5^3 && C_5^4 && C_5^5 \\
\multicolumn{11}{c}{\dots\dots\dots\dots\dots\dots\dots\dots}
\end{array}
\qquad
\begin{array}{ccccccccccc}
&&&&& 1 &&&&& \\
&&&& 1 && 1 &&&& \\
&&& 1 && 2 && 1 &&& \\
&& 1 && 3 && 3 && 1 && \\
& 1 && 4 && 6 && 4 && 1 & \\
1 && 5 && 10 && 10 && 5 && 1 \\
\multicolumn{11}{c}{\dots\dots\dots\dots\dots\dots\dots\dots}
\end{array}$$

Здесь каждое число, кроме крайних единиц, является суммой двух вышерасположенных.

Формула Стирлинга (без доказательства):

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\frac{\alpha_n}{n}\right),$$

где $|\alpha_n| \leqslant 1/12$, а $e \approx 2.718$ – основание натурального логарифма.

Эта формула является полезной при больших n. Например, из нее следует, что $\ln(n!)$ с точностью до $1/(12n)$ приближается выражением

$$\frac{1}{2}\ln(2\pi) + \left(n+\frac{1}{2}\right)\ln n - n.$$

Рассмотрим вопрос о числе комбинаций элементов, составленных из элементов различных групп. Считаем, что в состав комбинации входит по одному элементу из каждой группы и порядок элементов безразличен.

Теорема 4 (о числе комбинаций). *Число различных комбинаций элементов вида $(a^1, a^2, \ldots, a^r)$, где a^l – некоторый элемент l-й группы, состоящей из n_l элементов, равно $n_1 n_2 \ldots n_r$.*

Доказательство. Докажем сначала теорему для комбинаций элементов вида (a^1, a^2), составленных из элементов двух групп, т. е. рассмотрим сначала случай $r = 2$. Составим из пар элементов (a_i^1, a_j^2), $i = 1, 2, \ldots, n_1$, $j = 1, 2, \ldots, n_2$, прямоугольную таблицу (матрицу), содержащую n_1 строк и n_2 столбцов, так чтобы пара (a_i^1, a_j^2) стояла на пересечении i-й строки и j-го столбца:

$$\begin{array}{ccccc}
(a_1^1, a_1^2) & (a_1^1, a_2^2) & (a_1^1, a_3^2) & \ldots & (a_1^1, a_{n_2}^2) \\
(a_2^1, a_1^2) & (a_2^1, a_2^2) & (a_2^1, a_3^2) & \ldots & (a_2^1, a_{n_2}^2) \\
(a_3^1, a_1^2) & (a_3^1, a_2^2) & (a_3^1, a_3^2) & \ldots & (a_3^1, a_{n_2}^2) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(a_{n_1}^1, a_1^2) & (a_{n_1}^1, a_2^2) & (a_{n_1}^1, a_3^2) & \ldots & (a_{n_1}^1, a_{n_2}^2)
\end{array}$$

Каждая из пар (a_i^1, a_j^2) встречается в этой таблице один раз, и число таких пар равно произведению $n_1 n_2$. Тем самым, утверждение теоремы справедливо при $r = 2$. Для случая произвольного числа групп элементов формула о числе комбинаций доказывается по индукции. Предположим, что она верна для $r = k$ и докажем ее для

$r = k + 1$. Первые k элементов можно рассматривать как один элемент вида $b^1 = (a^1, a^2, \ldots, a^k)$. По предположению число различных элементов этой объединенной группы равно $m = n_1 n_2 \ldots n_k$. Любой элемент $(a^1, a^2, \ldots, a^{k+1})$ из группы, состоящей из $k + 1$ элемента, представим в виде $(a^1, a^2, \ldots, a^{k+1}) = (b^1, a^{k+1})$. Используя полученную формулу для числа различных элементов, составленных из двух групп, получим, что число комбинаций элементов вида $(a^1, a^2, \ldots, a^{k+1})$ определяется равенством $N = mn_{k+1} = n_1 n_2 \ldots n_k n_{k+1}$. Тем самым, формула о числе элементов верна для $r = k+1$. Индукционный переход от $r = k$ к $r = k+1$ завершен, а значит, доказана и теорема 4.

Задача 6. Из трех классов спортивной школы нужно составить команду для соревнований, взяв по одному ученику от класса. Сколько различных команд можно составить, если в одном классе учатся 18, в другом – 20, а в третьем – 22 ученика?

Решение очевидно: $18 \cdot 20 \cdot 22 = 7920$.

Задачи

Задача 1.1. Имеется пять видов конвертов без марок и четыре вида марок одного достоинства. Сколькими способами можно выбрать конверт с маркой для посылки письма?

Задача 1.2. В жюри 10 человек. Трое голосуют за, семеро – против. Сколько возможно комбинаций при голосовании?

Задача 1.3. Группа студентов из 30 человек решила обменяться фотокарточками. Сколько фотокарточек понадобилось для этого?

Задача 1.4. Сколько различных вариантов хоккейной команды можно составить из 9 нападающих, 5 защитников и 3 вратарей, если в состав команды должны войти 3 нападающих, 2 защитника и 1 вратарь?

Задача 1.5. Сколько прямых линий можно провести через 8 точек, из которых ровно 3 лежат на одной прямой?

Задача 1.6. Сколько словарей нужно издать, чтобы переводить с любого из 5 языков на любой другой из этих языков.

Задача 1.7. Есть пятиразрядный цифровой замок. Кодовое устройство замка состоит из пяти вращающихся дисков, каждый из которых имеет шесть цифр от 0 до 5. Только одна комбинация из пяти цифр позволяет открыть замок. Сколько таких комбинаций?

Задача 1.8. Сколькими способами можно упорядочить множество $\{1, 2, \ldots, 2n\}$ так, чтобы каждое четное число имело четный номер?

Задача 1.9. Сколькими способами можно упорядочить множество $\{1, 2, \ldots, n\}$ так, чтобы числа 1,2,3 стояли рядом и в порядке возрастания?

Задача 1.10. Какое количество различных символов (букв, чисел и т. д.) можно передать не более чем пятью знаками кода Морзе, использующего точку $(\cdot)$ и тире $(-)$?

Задача 1.11. Автомобильные номера состоят из трех букв и четырех цифр. Найти число таких номеров, если используются 32 буквы алфавита.

Задача 1.12. Сколько шестизначных чисел, все цифры у которых различны?

Задача 1.13.(сказка) Жил-был странный правитель. Решил он своих подданных различать не по именам, а по зубам. Себе все 32 зуба оставил, как и были, белыми. Ближайшим подданным повелел один зуб на разных позициях окрасить в черный цвет, чтобы их отличать. Далее шли вассалы с двумя черными зубами на разных позициях, и так далее. В самых низших слоях были люди с одним белым зубом на разных местах, и был один только с черными. Сколько было подданных у правителя?

Задача 1.14. Сколько машинных (различающихся по написанию и не обязательно имеющих смысл) слов можно составить из букв слова КОЛОКОЛ, из букв слова ВОДОРОД?

Задача 1.15. Сколькими способами можно посадить за круглый стол 7 мужчин и 7 женщин так, чтобы никакие две женщины не сидели рядом?

Задача 1.16. Сколькими способами 9 одинаковых конфет можно разложить по пяти различным пакетам, если ни один из пакетов не должен быть пустым? Тот же вопрос, но пакеты могут быть пустыми.

§ 2. Случайные события

В основе теории вероятностей лежит понятие случайного эксперимента. Эксперимент считается **случайным**, если он может закончиться любым из совокупности известных результатов, но до осуществления эксперимента нельзя предсказать, каким именно.

Примеры случайного эксперимента: бросание монеты, бросание игральной кости, проведение лотереи, азартные игры, стрельба по цели, поступление звонков на телефонную станцию.

Различные результаты эксперимента мы будем называть **исходами**.

Определение 1. Множество всех взаимно исключающих исходов эксперимента называется **пространством элементарных событий**.

Взаимно исключающие исходы – это те, которые не могут наступить одновременно. В дальнейшем под термином "исход" подразумеваются только такие исходы.

Пространство элементарных событий мы будем обозначать буквой Ω, а его исходы – буквой ω с различными индексами и без них или другими понятными из контекста символами. Термины "элементарное событие" и "исход" будем считать синонимами.

Определение 2. Произвольное подмножество пространства элементарных событий называется **событием**.

Событие может состоять из одного или нескольких элементарных событий, а также состоять из счетного или несчетного числа элементарных событий. События будут обозначаться заглавными латинскими или русскими буквами $A, B, C, \ldots$, Г,Р,Ч, $\ldots$, или записываться словами, например, {выпало четное число очков на игральной кости}. Особо следует выделить событие, не содержащее исходов. Такое событие, состоящее из пустого множества исходов, будет обозначаться символом $\varnothing$.

Говорят, что **событие A наступило**, если эксперимент заканчивается одним из исходов, входящих в событие A.

Примеры: 1) Бросание монеты. Предполагаем, что монета достаточно тонкая и при бросании не встает на ребро. Пространство элементарных событий состоит из двух исходов: Г = {выпал герб}, Р = {выпала решетка}, т. е. $\Omega = \{Г, Р\}$.

2) Бросание игральной кости, т. е. кубика, сделанного из однородного материала, грани которого занумерованы цифрами $1, 2, 3, 4, 5, 6$. Число очков, выпавшее при бросании игральной кости – цифра на верхней грани кубика. Пространство элементарных событий $\Omega = \{1, 2, 3, 4, 5, 6\}$. Событие Ч = {выпало четное число} состоит из трех исходов, т. е. Ч $= \{2, 4, 6\}$. Считаем, что Ч наступило, если выпало либо 2, либо 4, либо 6.

В зависимости от задачи в одном и том же эксперименте можно по-разному выбирать пространство элементарных событий. Так, при бросании игральной кости, если нас интересует лишь то, что выпало четное или нечетное число, можно считать $\Omega = \{Ч, Н\}$, где Ч ={четное число}, Н ={нечетное число}.

Пример неправильно выбранного пространства элементарных событий. Пусть при бросании игральной кости Ч ={четное число очков}, Т={число очков, кратное трем}. Тогда $\Omega = \{Ч, Т, 1, 5\}$ составляет все исходы эксперимента, однако исходы Ч и Т могут наступать одновременно.

Пример пространства с несчетным числом элементарных событий. Пусть есть проволока длиной, например, 1 метр. Мы растягиваем ее за концы, в результате чего происходит разрыв в какой-то точке. Множество исходов – это все точки на проволоке, которые математически можно задать отрезком $[0, 1]$, т. е. $\Omega = [0, 1]$, а каждому исходу ω соответствует координата точки разрыва.

Определение 3. **Суммой** двух событий A и B (обозначается $A + B$) называется событие, состоящее из всех исходов, входящих либо в A, либо в B.

Другими словами, под $A + B$ понимают следующее событие: произошло или событие A, или событие B, или они произошли одновременно, т. е. произошло хотя бы одно из событий A или B.

Определение 4. **Произведением** двух событий A и B (обозначается AB) называется событие, состоящее из тех исходов, которые входят как в A, так и в B.

Иными словами, AB означает событие, при котором события A и B наступают одновременно.

Определение 5. **Разностью** двух событий A и B (обозначается $A-B$) называется событие, состоящее из исходов, входящих в A, но не входящих в B.

Смысл события $A-B$ состоит в том, что событие A наступает, но при этом не наступает событие B.

Определение 6. **Симметрической разностью** двух событий A и B (обозначается $A\Delta B$) называется событие, состоящее из исходов, входящих в A *или* в B, но не входящих в A и B одновременно.

Смысл события $A\Delta B$ состоит в том, что наступает одно и только одно из событий A и B.

Определение 7. **Противоположным (дополнительным)** для события A (обозначается $\overline{A}$) называется событие, состоящее из всех исходов, которые не входят в A.

Наступление события $\overline{A}$ означает просто, что событие A не наступило.

Примеры. Пусть при бросании игральной кости $A = \{$выпало четное число$\}$, $B = \{$выпало число, кратное трем$\}$. Тогда $A+B = \{2,4,6\} + \{3,6\} = \{2,3,4,6\}$, $AB = \{2,4,6\}\{3,6\} = \{6\}$, $A-B = \{2,4,6\} - \{3,6\} = \{2,4\}$, $A\Delta B = \{2,3,4\}$, $\overline{B} = \{1,2,4,5\}$.

Если события изображать множествами на плоскости, то результат определенных операций над событиями выглядит следующим образом:

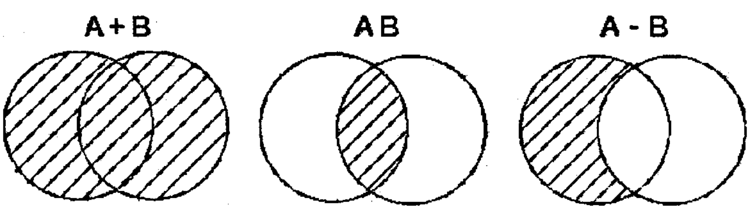

Рис. 2

Событие Ω, состоящее из всех исходов эксперимента, называется **достоверным** событием. Оно обязательно происходит, так как эксперимент всегда заканчивается каким-нибудь исходом.

Пустое множество исходов эксперимента называется **невозможным** событием и обозначается символом $\varnothing$.

Пример невозможного события. При бросании игрального кубика на ровную поверхность событие $\varnothing$ состоит в том, что кубик встал на вершину.

Определение 8. События A и B называются **несовместными**, если нет исходов, входящих как в A, так и в B, т. е. $AB = \varnothing$.

Это означает, что события не могут наступать одновременно.

Определение 9. Говорят, что событие A **содержится** в событии B (обозначается $A \subset B$), если все исходы события A входят в событие B.

Иными словами, включением $A \subset B$ обозначается такая ситуация, при которой наступление события A обязательно влечет наступление события B.

Свойства операций над событиями.

1) $A + B = B + A$; 2) $AB = BA$; 3) $A + \overline{A} = \Omega$;

4) $A\Omega = A$; 5) $AB \subset A$; 6) $A\overline{A} = \varnothing$; 7) $\overline{\overline{A}} = A$;

8) $A - B = A\overline{B}$; 9) $A\Delta B = A\overline{B} + \overline{A}B$;

10) $(A + B)\,C = AC + BC$; 11) $\overline{A + B} = \overline{A}\,\overline{B}$;

12) $\overline{AB} = \overline{A} + \overline{B}$; 13) $(A - B)\,C = AC - BC$.

Свойства 1) – 9) непосредственно следуют из определения операций над событиями.

Доказательство свойства 10). Возьмем произвольный исход ω из события $(A + B)C$ (будем обозначать это следующим образом: $\omega \in (A+B)C$). Тогда $\omega \in A+B$ и $\omega \in C$. Из включения $\omega \in A + B$ следует, что ω принадлежит хотя бы одному слагаемому. Пусть, например, $\omega \in A$. Из $\omega \in A$ и $\omega \in C$ следует, что $\omega \in AC$, и, следовательно, $\omega \in AC + BC$. Таким образом, любой исход события $(A + B)C$ является исходом события $AC + BC$, т. е. $(A + B)C \subset AC + BC$. Аналогично можно показать, что любой исход события $AC+BC$ является исходом события $(A + B)C$. Отсюда следует свойство 10), так как события его левой и правой частей состоят из одних и тех же исходов.

Доказательство свойства 11). Принадлежность исхода ω событию $\overline{A+B}$ означает, что он не принадлежит событию $A+B$, то есть не принадлежит ни A, ни B. Не принадлежит A, значит, принадлежит $\overline{A}$, не принадлежит B, значит, принадлежит $\overline{B}$, а это влечет то, что ω принадлежит $\overline{A}$ и $\overline{B}$ одновременно, т. е. $\omega \in \overline{A}\,\overline{B}$. Следовательно, выполняется включение $\overline{A+B} \subset \overline{A}\,\overline{B}$. Рассуждая в обратном порядке, мы получим включение $\overline{A}\,\overline{B} \subset \overline{A+B}$. Это доказывает равенство 11).

Доказательство свойства 12). Применим свойство 11) к событиям $\overline{A}$, $\overline{B}$ вместо A, B. Имеем $\overline{\overline{A}+\overline{B}} = \overline{\overline{A}}\,\overline{\overline{B}} = AB$. Во втором равенстве мы воспользовались свойством 7). Взяв дополнение к левому и правому событиям этого равенства и снова используя 7), получим свойство 12).

Доказательство свойства 13). Воспользуемся свойствами 8), 12) и 10). Тогда имеем цепочку равенств

$$AC - BC = AC\overline{BC} = AC(\overline{B}+\overline{C}) = AC\overline{B} = (A\overline{B})C = (A-B)C.$$

Операции над событиями можно последовательно осуществлять несколько раз. Так, например, по определению, $A+B+C = (A+B)+C$ и $ABC = (AB)C$. В силу свойств 1) и 2) порядок осуществления операций сложения и умножения безразличен. Событие $A_1 + \cdots + A_n = \sum\limits_{k=1}^{n} A_k$ заключается в том, что происходит по крайней мере одно из событий $A_1, A_2, \ldots, A_{n-1}$ или A_n. Событие $A_1A_2\ldots A_{n-1}A_n = \prod\limits_{k=1}^{n} A_k$ состоит в том, что все события $A_1, A_2, \ldots, A_{n-1}, A_n$ наступают одновременно.

Следствиями свойств 10) – 12) являются следующие свойства операций над событиями:

$$10')\ B\sum_{k=1}^{n} A_k = \sum_{k=1}^{n} BA_k; \quad 11')\ \overline{\sum_{k=1}^{n} A_k} = \prod_{k=1}^{n} \overline{A_k};$$

$$12')\ \overline{\prod_{k=1}^{n} A_k} = \sum_{k=1}^{n} \overline{A_k}.$$

Эти равенства верны и для $n = \infty$.

Задача 1. Монета подбрасывается три раза подряд. Исходом каждого бросания служит выпадение "герба" – Г или "решетки" – Р. Описать пространство элементарных событий данного эксперимента и описать событие A, состоящее в том, что выпало не менее двух "гербов".

Решение. Исходом данного эксперимента является появление либо трех "гербов", либо трех "решеток", либо комбинаций "гербов" и "решеток". Эти исходы будем записывать следующими наборами символов: ГГГ, РРР, ГРГ и т. д. Используя эти обозначения, пространство элементарных событий Ω можно записать следующим образом: $\Omega = \{$ГГГ, ГГР, ГРГ, РГГ, РРР, РРГ, РГР, ГРР$\}$. Событие A состоит из четырех исходов: $A = \{$ГГГ, ГГР, ГРГ, РГГ$\}$.

Задача 2. Двое играют в шахматы. Событие A означает, что выиграл первый игрок, событие B – что выиграл второй игрок. Что означают события

а) $\overline{B}$; б) $\overline{A}\Delta B$; в) $\overline{A}+\overline{B}$; г) $\overline{B}-A$; д) $\overline{A}-B$?

Решение. Нужно учесть, что $AB = \varnothing$, поскольку одновременное наступление событий "выиграл первый" и "выиграл второй" невозможно. Тогда указанные события означают, что

а) первый игрок выиграл или сыграл вничью;

б) партия закончилась вничью;

в) возможен любой исход партии;

г) партия закончилась вничью;

д) партия закончилась вничью.

Задачи

Задача 2.1. Из таблицы чисел взято одно число. Событие A – выбранное число делится на 5; событие B – данное число оканчивается нулем. Что означают события $A-B$ и $A\overline{B}$?

Задача 2.2. Событие A – хотя бы одно из имеющихся четырех изделий бракованное, событие B – бракованных изделий среди них не менее двух. Что означают противоположные события $\overline{A}$ и $\overline{B}$?

Задача 2.3. Когда возможны равенства: а) $A+B=\overline{A}$; б) $AB=\overline{A}$; в) $A+B=AB$?

Задача 2.4. Совместны ли события A и $\overline{A+B}$?

Задача 2.5. Бросают две игральные кости. Пусть A – событие, состоящее в том, что сумма очков равна пяти, а B – событие, заключающееся в том, что хотя бы на одной из костей выпала единица. Описать события AB и $A\overline{B}$.

Задача 2.6. Бросают две игральные кости. Событие A – выпала четная сумма очков, а B – хотя бы на одной кости выпала шестерка. Что означает событие $\overline{A+B}$?

Задача 2.7. Из колоды вытаскивают две карты. Событие A – обе карты пиковой масти, а B – нет тузов. Что означает событие $A - B$?

Задача 2.8. Рабочий изготовил n деталей. Пусть событие A_i, $i = 1, 2, \ldots, n$, заключается в том, что i-я изготовленная им деталь имеет дефект. Записать событие, заключающееся в том, что: а) ни одна из деталей не имеет дефектов; б) хотя бы одна деталь имеет дефект.

Задача 2.9. Событие A – спортсмен прыгнул дальше 7 метров, событие B – мужчина прыгнул дальше женщины, событие C – женщина прыгнула дальше 7 метров. Что означают события ABC, $A - AB$ и $A\overline{B}C$?

Задача 2.10. Пусть A, B, C – три произвольных события. Найти выражения для событий, состоящих в том, что из A, B, C

а) произошло только A;

б) произошли A и B, но C не произошло;

в) все три события произошли;

г) произошло по крайней мере одно из этих событий;

д) произошли по крайней мере два события;

е) произошло ровно одно из этих событий;

ж) произошло ровно два из этих событий;

з) ни одно событие не произошло;

и) произошло не больше двух событий.

§ 3. Классическое определение вероятности

Вероятность события характеризует возможность (шанс) осуществления события в ходе случайного эксперимента. Смысл вероятности раскрывается в следующих требованиях, налагаемых на вероятности событий. Вероятность события выражается долей от целого, т. е. является числом от нуля до единицы. При этом вероятность наступления достоверного события, т. е. события, которое обязательно происходит в ходе эксперимента, принимается равной единице. Вероятности невозможных событий считаются равными нулю. Если взять произвольный конечный или счетный набор событий, таких что никакие два из них не могут произойти одновременно, то вероятность наступления хотя бы одного из этих событий должна быть равна сумме вероятностей этих событий. Иными словами, шанс наступления суммы попарно несовместных событий равен сумме шансов каждого из событий. Указанные требования налагают жесткие ограничения на численные значения вероятностей.

Вычисление значений вероятностей событий в различных случайных экспериментах является предметом теории вероятностей. Оказывается, что численное значение вероятности того или иного события проявляется следующим образом. Если случайный эксперимент многократно повторить при одних и тех же условиях и вычислить частоту появлений конкретного события среди всех проведенных экспериментов (**частота появлений события** есть отношение числа появлений события к числу экспериментов), то при неограниченном возрастании числа экспериментов эта частота в пределе будет совпадать с вероятностью события. При определенных исходных предположениях теории вероятностей это – строго доказанное утверждение (см. § 16). Основной задачей теории вероятностей является разработка правил вычисления вероятностей событий, появляющихся в сложных случайных экспериментах, по известным вероятностям событий более простых экспериментов. Изучение основ теории вероятностей мы начнем с рассмотрения простых случайных экспериментов, имеющих конечное число исходов.

Пусть пространство элементарных событий конечно:

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

Будем говорить, что на пространстве Ω **заданы вероятности**, если каждому исходу ω_l сопоставлено неотрицательное число $p(\omega_l)$ так, что выполнено условие нормировки

$$\sum_{l=1}^{n} p(\omega_l) = 1.$$

Определение 1. **Вероятностью события** $A = \{\omega_{l_1}, \omega_{l_2}, \ldots, \omega_{l_k}\}$ называется число $\mathbf{P}(A)$, равное сумме вероятностей элементарных исходов, составляющих событие A, т. е.

$$\mathbf{P}(A) = \sum_{\omega_l \in A} p(\omega_l).$$

Рассмотрим простейшую модель теории вероятностей, которую часто называют **"классической схемой"**. Пусть в силу тех или иных свойств, обычно связанных с симметрией, каждый элементарный исход эксперимента имеет одинаковую возможность осуществиться (исходы равновозможны). В такой ситуации этим исходам естественно сопоставить одинаковые вероятности. Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, и все исходы $\omega_1, \omega_2, \ldots, \omega_n$ равновозможны. Тогда полагаем $p(\omega_1) = p(\omega_2) = = \cdots = p(\omega_n) = p$. В силу условия нормировки $np = 1$ или $p(\omega_l) = p = 1/n$, где n – число исходов эксперимента. Для любого множества A символом $\operatorname{card} A$ (**кардинальное число множества** A) будем обозначать количество элементов этого множества. Исходы, составляющие событие A, называются благоприятными для этого события. Тогда в "классической схеме" мы имеем $\operatorname{card}\Omega = n$, $p(\omega_l) = \dfrac{1}{\operatorname{card}\Omega}$, $l = 1, 2, \ldots, n$, и

$$\mathbf{P}(A) = \sum_{\omega_l \in A} p(\omega_l) = p \operatorname{card} A = \frac{\operatorname{card} A}{\operatorname{card}\Omega}.$$

Таким образом, в случае, когда все исходы равновозможны, вероятность события A равна отношению числа благоприятных исходов для события A к общему числу исходов эксперимента. Это – классическое определение вероятности.

Примеры. При бросании правильной монеты $\mathbf{P}(\Gamma) = \mathbf{P}(\mathrm{P}) = 1/2$. При бросании игральной кости $\mathbf{P}(1) = \mathbf{P}(2) = \cdots = \mathbf{P}(6) = 1/6$, $\mathbf{P}$(выпало четное число) $= 3/6 = 1/2$.

В формулировках многих задач при случайном выборе чего-либо будет употребляться слово **"наудачу"**. Это слово означает беспристрастный выбор, при котором все комбинации элементов, которые могут быть выбраны, равновозможны. Иногда слово "наудачу" будет заменяться эквивалентным выражением **"случайным образом"**.

Задача 1. Для участия в лотерее, на карточке, содержащей 49 чисел, нужно отметить 6 чисел. Затем эти числа сверяются с 6 числами, отобранными случайным образом. В зависимости от числа совпавших номеров выплачивается выигрыш. Какова вероятность угадать в лотерее 6 чисел из 49?

Решение. В этой задаче Ω – совокупность всех сочетаний из 49 чисел по 6, и $\operatorname{card}\Omega = C_{49}^{6}$. Столько существует различных вариантов заполнения карточки, и каждый из них имеет одинаковый шанс стать выигрышным. Благоприятствует выигрышу только одно событие: номер на карточке совпал с отобранным. Поэтому $\mathbf{P}$(угадать 6 из 49) $= \dfrac{1}{C_{49}^{6}} = \dfrac{1}{13983816} \approx \dfrac{1}{14000000}$.

Задача 2. В ящике m белых и n черных шаров. Шары тщательно перемешаны. Наудачу вынимаются сразу два шара. Какова вероятность того, что оба шара белые?

Решение. Шары можно для удобства пронумеровать числами от 1 до $m+n$. Поскольку нам не важно, какой шар первый, а какой второй, то возможными исходами эксперимента будут различные сочетания из $m+n$ чисел по 2. Тогда $\operatorname{card}\Omega = C_{m+n}^{2}$. Событию $A = \{$оба белые$\}$ отвечают лишь сочетания из m чисел по 2. Следовательно, $\operatorname{card}A = C_m^2$, и

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

Задача 3 (де Мере). Еще в **XVII** веке француз Шевалье де Мере задался вопросом: какая сумма очков имеет больше шансов выпасть при бросании двух игральных костей – 11 или 12? Сумму 11 могут составить лишь два числа – 5 и 6, а сумму 12 тоже два числа – 6 и 6. На первый взгляд шансы у этих событий равны. Так ли это?

Решение. Будем различать игральные кости. Пусть мысленно одна будет красной, а другая – белой. Различных комбинаций очков на разноцветных костях согласно теореме 4 о числе комбинаций (§ 1) будет $6\cdot 6 = 36$. Следовательно, $\operatorname{card}\Omega = 36$. Сумма 12 выпадает лишь при одной комбинации, когда на красной кости 6 и на белой 6. Таким образом, $\mathbf{P}(12) = 1/36$. Сумма 11 выпадает при двух комбинациях, когда на красной 6, а на белой 5, и наоборот. Тогда

$\mathbf{P}(11) = 2/36$. Следовательно, сумма очков 11 имеет в два раза больше шансов выпасть, чем 12.

Обратимся еще раз к обсуждению понятия вероятности, используя уже рассмотренные нами примеры экспериментов с конечным числом исходов. С математической точки зрения вероятность – неотрицательное число, которое сопоставляется случайному исходу эксперимента, при этом сумма всех вероятностей, отвечающих различным исходам эксперимента, равна единице.

Абстрактная теория занимается в основном разработкой правил пересчета одних вероятностей в другие, предполагая, что исходные вероятности даны и не нуждаются ни в каких обоснованиях их действительных численных значений.

Выяснению численных значений вероятностей в различных конкретных экспериментах посвящена другая математическая дисциплина – статистика. Существуют также эксперименты, в которых априори ясно, чему равны значения вероятностей. Таковы, в частности, эксперименты с равновозможными исходами, в которых исходам сопоставляются равные вероятности. По существу это заложено в самом понятии "равновозможности". Чаще всего вывод о том, какие исходы считать равновозможными, основан на соображениях, связанных с симметрией и однородностью. Если, например, монета симметрична и сделана из однородного материала, то шансы выпадения герба и решетки совпадают, т. е. вероятности равны $1/2$.

Какой же смысл вкладываем мы подсознательно во фразу "при бросании правильной монеты вероятность выпадения герба равна $1/2$", кроме того, что монета симметрична? Оказывается, что вероятность проявляет себя, когда один и тот же случайный эксперимент проводится много раз, причем так, что результаты уже осуществленных экспериментов никак не влияют на последующие. При этих условиях частота наступления события при неограниченном возрастании числа экспериментов стремится к вероятности события.

Пусть правильная монета бросается n раз, и пусть число выпадений герба при этом равно $n_{\text{Г}}$. Тогда при неограниченном увеличении числа бросаний монеты частота выпадений герба будет стремиться к $\frac{1}{2}$, т. е. $\frac{n_{\text{Г}}}{n} \to \frac{1}{2}$ при $n \to \infty$. Этому факту в рамках определенной математической модели будет дано строгое обоснование в § 16.

Эксперименты не обязательно проводить последовательно, можно одновременно провести много одинаковых экспериментов. Например, в лотерее из задачи 1 участвуют сразу много людей. Считаем, что один человек заполняет одну карточку, в которой отмечает 6 чисел из 49 возможных. Процесс заполнения карточки отдельным человеком – случайный эксперимент. В результате доля людей, угадавших 6 из 49, будет близка к вероятности, которую мы вычислили в задаче 1. Так, если в лотерее участвовало 56 млн. человек, то угадавших 6 чисел будет скорее всего 4, поскольку $\frac{4}{56000000} = \frac{1}{14000000}$.

Свойства вероятностей.

1) $\mathbf{P}(\Omega) = 1$ в силу условия нормировки;

2) $\mathbf{P}(\varnothing) = 0$, поскольку сумма, в которой нет слагаемых, равна нулю;

3) если $A \subset B$, то $\mathbf{P}(A) \leqslant \mathbf{P}(B)$, так как выражение для вероятности $\mathbf{P}(B)$ по сравнению с $\mathbf{P}(A)$ содержит дополнительные неотрицательные слагаемые;

4) $0 \leqslant \mathbf{P}(A) \leqslant 1$ для любого события A, что следует из свойств 3) и 1);

5) $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$, так как

$$\mathbf{P}(\overline{A}) + \mathbf{P}(A) = \sum_{\omega_l \in \overline{A}} p(\omega_l) + \sum_{\omega_l \in A} p(\omega_l) = \sum_{\omega_l \in \Omega} p(\omega_l) = 1;$$

6) **(формула сложения вероятностей)** для любых событий A и B

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

Действительно,

$$\mathbf{P}(A + B) = \sum_{\omega_l \in A+B} p(\omega_l) =$$

$$= \sum_{\omega_l \in A} p(\omega_l) + \sum_{\omega_l \in B} p(\omega_l) - \sum_{\omega_l \in AB} p(\omega_l) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(AB).$$

Следствие 1. *Если $AB = \varnothing$, т. е. события A и B несовместны, то*

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

Следствие 2. *Если события $A_l, l = 1, 2, \ldots, n$, попарно несовместны, т. е. $A_l A_j = \varnothing$ для любых $l \neq j$, то*

$$\mathbf{P}\Big(\sum_{l=1}^{n} A_l\Big) = \sum_{l=1}^{n} \mathbf{P}(A_l)$$

Это равенство несложно доказать по индукции, используя следствие 1 и свойство 10′) § 2, из которого следует несовместность событий A_{k+1} и $\sum\limits_{l=1}^{k} A_l$ для любого k.

Следствие 3. *Если $C \subset D$ то*

$$\mathbf{P}(D-C) = \mathbf{P}(D) - \mathbf{P}(C).$$

Действительно, события $D-C$ и C несовместны, а их сумма составляет событие D. Тогда по следствию 1

$$\mathbf{P}(D) = \mathbf{P}(D-C) + \mathbf{P}(C).$$

Задачи

Задача 3.1. Найти вероятность того, что наудачу взятое двузначное число окажется кратным 2, либо 5, либо тому и другому одновременно.

Задача 3.2. Какова вероятность того, что при бросании двух кубиков сумма очков будет кратна четырем?

Задача 3.3. В выпуклом двадцатиугольнике случайным образом берут две вершины и соединяют отрезком. Чему равна вероятность того, что отрезок является диагональю двадцатиугольника?

Задача 3.4. Наудачу взят телефонный номер, состоящий из пяти цифр. Чему равна вероятность того, что все цифры различные?

Задача 3.5. На полке стоят 15 книг, 5 из них в переплете. Берут наудачу три книги. Какова вероятность того, что все три книги в переплете?

Задача 3.6. В ящике 20 шаров с номерами $1, 2, \ldots, 20$. Наудачу выбираются шесть шаров. Найти вероятность того, что среди них есть шары с номерами 1 и 2.

Задача 3.7. Среди 100 фотокарточек есть одна фотокарточка знаменитого артиста. Взяли наудачу 10 фотокарточек. Какова вероятность того, что среди них есть фото артиста?

Задача 3.8. Из колоды в 36 карт вытаскивают две карты. Какова вероятность, что будет хотя бы одна карта пиковой масти?

Задача 3.9. Из 12 лотерейных билетов, среди которых есть 4 выигрышных, наудачу берут 6. Какова вероятность того, что хотя бы один из них выигрышный?

Задача 3.10. Бросают две игральные кости. Какова вероятность того, что сумма очков не более трех?

Задача 3.11. Из 60 вопросов, входящих в экзаменационные билеты, студент подготовил 50. Какова вероятность того, что вытянутый студентом билет, содержащий два вопроса, будет включать в себя подготовленные вопросы?

Задача 3.12. Для уменьшения общего количества игр на соревнованиях 16 волейбольных команд разбиты на две подгруппы по 8 в каждой. Найти вероятность того, что две наиболее сильные команды окажутся: а) в разных подгруппах; б) в одной подгруппе.

Задача 3.13. При наборе телефонного номера абонент забыл две последние цифры и набрал их наудачу, помня только, что эти цифры нечетные и разные. Найти вероятность того, что номер набран правильно.

Задача 3.14. Найти вероятность того, что дни рождения 12 человек придутся на разные месяцы года.

Задача 3.15. В ящике 10 красных и 6 белых шаров. Вынимаются наудачу два шара. Какова вероятность того, что шары будут одноцветными?

Задача 3.16. Бросаются 4 игральные кости. Найти вероятность того, что на них выпадет по одинаковому числу очков.

Задача 3.17. Сколько раз нужно бросить игральную кость, чтобы с вероятностью не меньше а) 0.5; б) 0.9 хотя бы один раз выпала шестерка (шесть очков)?

Задача 3.18. Из колоды в 36 карт вытаскивают 4. Какова вероятность того, что окажется два короля и две дамы?

Какова вероятность того, что окажется три туза и одна шестерка?

Задача 3.19. Из колоды в 36 карт вытаскивают 3. Какова вероятность того, что среди них окажется ровно две карты червовой масти?

Задача 3.20. В ящике 5 белых, 6 красных и 3 черных шара. Какова вероятность того, что вынутые два шара окажутся разного цвета.

Задача 3.21. Из колоды в 36 карт вытаскивают 3. Какова вероятность, что нет тузов?

Задача 3.22. В книжной лотерее разыгрывается 20 билетов, среди которых 4 выигрышных. Наудачу покупают три билета. Какова вероятность того, что среди них ровно два выигрышных?

Задача 3.23. Какова вероятность получить менее 11 очков, бросая два игральных кубика?

Задача 3.24. Лифт в девятиэтажном доме отправляется с четырьмя пассажирами. Найти вероятность того, что на каждом из этажей выйдет не более одного пассажира, предполагая равновероятность всех возможных способов распределения пассажиров по этажам.

Задача 3.25. В ящике 2 белых и 4 черных шара. Один за другим вынимаются все имеющиеся в нем шары. Найти вероятность того, что последний шар будет черным.

Задача 3.26. В течение пяти дней случайным образом поступают сообщения о банкротстве каждого из пяти банков, назовем их условно A, B, C, D, E. Чему равна вероятность того, что сообщения о банкротстве банков A и B не следуют друг за другом?

Задача 3.27. Из последовательности чисел $1, 2, \ldots, n$ наудачу выбираются два числа. Какова вероятность, что одно из них меньше k, а другое больше k, где $1 < k < n$ – произвольное целое число?

Задача 3.28. На шахматную доску из 64 клеток ставятся наудачу две ладьи белого и черного цвета. С какой вероятностью они не будут "бить" друг друга?

§ 4. Геометрические вероятности

Пусть в пространстве задана некоторая область, например, отрезок, если пространством служит прямая, квадрат, если пространством служит плоскость или куб, если пространство трехмерно. Пусть результатом эксперимента является случайный выбор точки из этой области. Предположим, что по каким-то соображениям, обычно связанным с симметрией либо однородностью, выбор любой точки области равновозможен. Чему равны вероятности событий в таком эксперименте? Ответ на этот вопрос связан с понятием геометрических вероятностей.

Пример. Однородная проволока растягивается за концы. Результатом эксперимента является разрыв проволоки в какой-то точке. Всю область исходов эксперимента можно отождествить с отрезком, равным длине проволоки.

Заданную в пространстве область обозначим Ω. В эксперименте, связанном со случайным выбором только одной точки из Ω, множество Ω служит пространством элементарных событий. Символом mes обозначим меру Лебега в пространстве: для прямой это длина, для плоскости – площадь, для трехмерного пространства – объем. Случайными событиями в этом эксперименте можно считать различные подмножества Ω, для которых определена мера Лебега. Будем говорить, что событие A наступило, если случайно выбранная точка принадлежит множеству $A \subset \Omega$. Возникает вопрос: какая вероятность соответствует событию A? В фразу "выбор любой точки области равновозможен" вкладывается следующий смысл: шансы выбрать точки из любых двух множеств одинаковой меры Лебега равны между собой. Разобьем Ω на n частей одинаковой меры Лебега. Пусть событие A таково, что оно целиком состоит из k таких частей. Тогда эксперимент случайного выбора точки из Ω может быть описан с помощью классической схемы. Исходом здесь служит выбор точки из той или иной выделенной части. Всего n равновозможных исходов. При этом вероятность события A равна отношению числа частей, составляющих A, к общему числу частей, т. е.

$$\mathbf{P}(A) = \frac{k}{n} = \frac{\operatorname{mes} A}{\operatorname{mes} \Omega}.$$

Считая, если необходимо, n произвольным и применяя предельный переход при $n \to \infty$, можно показать, что любому событию A, для которого определена мера Лебега, соответствует вероятность

$$\mathbf{P}(A) = \frac{\operatorname{mes} A}{\operatorname{mes} \Omega}.$$

Вероятности, заданные по этой формуле, называют **геометрическими**.

Свойства геометрических вероятностей точно такие же, как и для вероятностей в классическом определении.

В примере о разрыве однородной проволоки можно рассмотреть событие $A = \{$проволока разорвалась ближе к центру, чем к концам$\}$. Здесь, чтобы вычислить вероятность события A, можно разбить проволоку на четыре равные части. Событию A отвечает точка разрыва, появляющаяся в одной из двух средних частей. Следовательно, $\mathbf{P}(A) = \frac{2}{4} = \frac{1}{2}$.

Задача 1. Наудачу выбираются два числа из промежутка $[0, 1]$. Какова вероятность, что их произведение меньше $1/2$?

Решение. Слово "наудачу" здесь означает, что появление любой пары чисел (x, y) из квадрата $[0,1]\times[0,1] \equiv \Omega$ равновозможно. Множество $A = \{(x, y) \in [0,1] \times [0,1] : xy \leqslant 1/2\}$ составляет случайное событие, вероятность которого нужно найти. Очевидно, $\operatorname{mes}\Omega = 1$. Множество A образуют те точки квадрата Ω, которые лежат под кривой $y = \frac{1}{2x}$. Следовательно,

$$\operatorname{mes} A = \frac{1}{2} + \int\limits_{1/2}^{1} \frac{1}{2x} dx = \frac{1}{2}(1 + \ln 2).$$

Это и есть искомая вероятность.

Задача 2 (о встрече). Два человека в течение промежутка времени $[0, T]$ случайным образом приходят к месту встречи и ждут время $\tau < T$. Какова вероятность, что они встретятся?

Решение. Пусть x – время прихода первого человека, а y – второго. Тогда $(x, y) \in [0, T] \times [0, T] \equiv \Omega$. Поскольку они приходят случайным образом, то все исходы Ω равновозможны. Событие $A = \{$они встретятся$\}$ можно задать так:

$A = \{(x, y) \in [0,T] \times [0,T] : |y - x| \leqslant \tau\}$. Множество A образуют те точки квадрата Ω, которые лежат между прямыми $y = x - \tau$ и $y = x + \tau$. Поэтому $\operatorname{mes} A = T^2 - (T - \tau)^2$. Поскольку $\operatorname{mes} \Omega = T^2$, то

$$\mathbf{P}(A) = \frac{\operatorname{mes} A}{\operatorname{mes} \Omega} = 1 - \left(1 - \frac{\tau}{T}\right)^2.$$

Задача 3. Какой толщины должна быть монета радиуса R, чтобы вероятность падения на ребро была равна $1/3$?

Решение. Укажем лишь на основные моменты решения, дающего ответ $R/\sqrt{2}$. Монету рассматриваем как вписанную в сферу. Если радиус, проведенный из центра сферы в направлении, выбранном наудачу, пересекает поверхность монеты, составляющую ребро, то считаем, что монета упала на ребро. Вероятность падения монеты на ребро равна отношению площади сферического пояса, заключенного между круглыми сторонами монеты, к площади сферы.

Задачи

Задача 4.1. На плоскость с нанесенной квадратной сеткой со стороной 4 см бросают монету радиуса 1 см. Найти вероятность того, что монета не пересечет линии.

Задача 4.2. В круг вписан квадрат. Какова вероятность того, что точка, брошенная в круг наудачу, окажется внутри квадрата?

Задача 4.3. Двое условились встретиться в определенном месте. Договорились, что каждый из них будет на месте встречи между 13 и 14 часами, и пришедший, не застав другого, подождет его в течение $1/4$ часа. Вычислить вероятность того, что встреча произойдет?

Задача 4.4. Наудачу выбирают два числа из промежутка $[0, 1]$. Какова вероятность того, что одно число более чем вдвое меньше другого?

Задача 4.5. Наудачу выбирают два числа из промежутка $[0, 1]$. Какова вероятность того, что их сумма заключена между $\frac{1}{2}$ и 1?

Задача 4.6. Из промежутка $[0, 1]$ выбрали наугад два числа. Какова вероятность, что их сумма больше либо равна 1, а их разность меньше либо равна 0?

Задача 4.7. На отрезке $[A, B]$ длиной l поставили наугад две точки L и M. Найти вероятность того, что L будет ближе к точке M, чем к A.

Задача 4.8. Какова вероятность того, что сумма двух наугад взятых положительных чисел, каждое из которых меньше либо равно 1, будет меньше либо равна 1, а их произведение будет не больше $\frac{2}{9}$?

Задача 4.9. На отрезке длиной l наугад ставятся две точки, которые разбивают его на три отрезка. Найти вероятность того, что из этих отрезков можно составить треугольник.

Задача 4.10. Электрон вылетает из случайной точки нити накала и движется по перпендикуляру к нити. С какой вероятностью он свободно пройдет через сетку, окружающую нить и имеющую вид винтовой линии радиуса R, толщиной δ и шага H?

Задача 4.11. Спутник Земли движется по орбите, которая заключена между 60° северной и 60° южной широты. Считая падение спутника в любую точку поверхности Земли между указанными параллелями равновозможным, найти вероятность того, что спутник упадет выше 30° северной широты.

Задача 4.12. Сторожевой катер курсирует по проливу от одного берега до другого, пересекая пролив за час. Какова вероятность того, что идущее вдоль пролива судно замечено, если с катера обнаруживают судно в случае, когда катер пересекает его курс не ранее чем за 20 минут до пересечения судном курса катера, и не познее чем через 20 минут после пересечения. Любой момент и любое место пересечения судном курса равновозможны.

§ 5. Условные вероятности. Независимость событий

Рассмотрим вопрос о том, как определить вероятность какого-либо события A при условии, что уже произошло другое событие B.

Начнем с примера, в котором возникает условная вероятность. Пусть брошена игральная кость и нам неизвестен результат, но известно, что выпало четное число. Мы же хотим, зная эту информацию, подсчитать вероятность того, что выпало число больше трех. Тогда речь идет об условной вероятности события $A = \{$выпало число больше трех$\}$ при условии, что произошло событие $B = \{$выпало четное число$\}$. Нам уже известно, что выпало либо 2, либо 4, либо 6 очков, и все эти исходы равновозможны. Среди этих исходов событию A удовлетворяют лишь исходы 4 и 6. Поэтому условной вероятностью естественно считать отношение $\frac{2}{3}$.

Вероятность события A при условии, что произошло событие B, называется **условной вероятностью** и обозначается $\mathbf{P}(A|B)$.

Условную вероятность мы будем рассматривать лишь для таких событий B, вероятность наступления которых отлична от нуля. Вопрос об условных вероятностях для событий, имеющих нулевую вероятность тоже имеет смысл, но его изучение выходит за рамки элементарного курса.

Определим условные вероятности $\mathbf{P}(A|B)$ для схемы, когда все исходы равновозможны. Пусть Ω – пространство элементарных событий. Поскольку известно, что событие B произошло, то будем рассматривать только те элементарные исходы, которые составляют событие B. Рассмотрим новое пространство элементарных событий $\Omega_1 = B$. Выберем множество исходов из A, которые входят в B, и обозначим его A_1. Очевидно $A_1 = AB$. За условную вероятность $\mathbf{P}(A|B)$ естественно взять вероятность события A_1 при условии, что рассматриваются только события, содержащиеся в B. Чему равна эта вероятность? Мы снова оказались в рамках "классической схемы", только для нового пространства элементарных событий Ω_1. Поэтому эта вероятность равна $\operatorname{card} A_1 / \operatorname{card} \Omega_1$. Следовательно,

$$\mathbf{P}(A|B) = \frac{\operatorname{card} A_1}{\operatorname{card} \Omega_1} = \frac{\operatorname{card} A_1 / \operatorname{card} \Omega}{\operatorname{card} \Omega_1 / \operatorname{card} \Omega} = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Это соотношение оправдывает следующее формальное определение условной вероятности.

Определение 1. **Условной вероятностью** события A при условии, что произошло событие B с $\mathbf{P}(B) \neq 0$, называется число $\mathbf{P}(A|B)$, которое определяется формулой

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

К такому же определению мы придем, если будем рассматривать геометрические вероятности. Это и не удивительно, поскольку при определении геометрических вероятностей мы использовали подход, основанный на “классической схеме”.

В силу того, что условная вероятность – это обычная вероятность, но лишь на более узком пространстве элементарных событий, то для нее справедливы все свойства обычной вероятности.

Свойства условных вероятностей.

1) $\mathbf{P}(\Omega|B) = 1$; 2) $\mathbf{P}(\varnothing|B) = 0$; 3) $0 \leqslant \mathbf{P}(A|B) \leqslant 1$;

4) если $A \subset C$, то $\mathbf{P}(A|B) \leqslant \mathbf{P}(C|B)$;

5)$\mathbf{P}(\overline{A}|B) = 1 - \mathbf{P}(A|B)$;

6) **(формула сложения условных вероятностей)** для любых событий A и C

$$\mathbf{P}(A + C|B) = \mathbf{P}(A|B) + \mathbf{P}(C|B) - \mathbf{P}(AC|B);$$

7) **(формула умножения вероятностей)** для любых событий A и B

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

Эта формула является непосредственным следствием определения условной вероятности.

8) **(общая формула умножения вероятностей)** для любых событий $A_1, A_2, \ldots, A_n$

$$\mathbf{P}(A_1A_2\ldots A_n) =$$
$$= \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1A_2)\ldots\mathbf{P}(A_n|A_1A_2\ldots A_{n-1}).$$

Эту формулу можно получить, применяя последовательно $n-1$ раз формулу умножения вероятностей 7). Действительно, полагая $A = A_n, B = A_1A_2\ldots A_{n-1}$ и применяя 7), получаем

$$\mathbf{P}(A_1A_2\ldots A_n) = \mathbf{P}(A_1A_2\ldots A_{n-1})\mathbf{P}(A_n|A_1A_2\ldots A_{n-1}).$$

Это равенство справедливо для любых n, поэтому можно продолжить цепочку равенств

$$\mathbf{P}(A_1A_2\ldots A_n) =$$
$$= \mathbf{P}(A_1A_2\ldots A_{n-2})\mathbf{P}(A_{n-1}|A_1A_2\ldots A_{n-2})\mathbf{P}(A_n|A_1A_2\ldots A_{n-1}) =$$
$$= \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\ldots\mathbf{P}(A_{n-1}|A_1A_2\ldots A_{n-2})\mathbf{P}(A_n|A_1A_2\ldots A_{n-1}).$$

Задача 1. В ящике m белых и n черных шаров. Шары тщательно перемешаны. Наудачу вынимаются сразу два шара. Какова вероятность того, что оба шара белые?

Решение (ранее (§ 3) данная задача решалась другим способом). Взять два шара сразу или сначала взять один шар, а потом другой – это одно и то же, ведь руку можно не вынимать из ящика, когда берется один шар, а затем другой. События {оба шара белые} и {1-й шар белый и 2-й шар белый} совпадают. Поэтому достаточно вычислить вероятность второго события, которое представимо в виде произведения событий A = {1-й шар белый}, B = {2-й шар белый}. Очевидно, $\mathbf{P}(A) = \frac{m}{m+n}$. Кроме того, $\mathbf{P}(B|A) = \frac{m-1}{m+n-1}$, поскольку, если событие A произошло, то в ящике уже остался $m+n-1$ шар, среди которых $m-1$ белых, и вероятность снова вынуть белый шар будет равна $\frac{m-1}{m+n-1}$. Теперь воспользуемся формулой произведения вероятностей:

$$\mathbf{P}(\text{оба белые}) = \mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B|A) = \frac{m(m-1)}{(m+n)(m+n-1)}.$$

Задача 2. Бросают три игральные кости. Какова вероятность того, что хотя бы на одной из них выпадет шестерка, если на всех трех костях выпали разные числа?

Решение. Определим события A = {хотя бы на одной из трех костей выпала шестерка} и B = {на трех костях выпали разные числа}. Нам нужно вычислить вероятность $\mathbf{P}(A|B)$. Для этого воспользуемся свойством 5), поскольку вероятность $\mathbf{P}(\overline{A}|B)$ вычислить легче. Будем различать кости, например, считая их окрашенными в разные цвета. По теореме 4 § 1 число различных комбинаций чисел на трех костях равно $6 \cdot 6 \cdot 6 = 216$. Число комбинаций разных чисел на трех костях равно числу размещений из 6 элементов по 3, т. е. A_6^3. Мы берем A_6^3, а не C_6^3, поскольку мы различаем кости. Тогда $\mathbf{P}(B) = A_6^3/216$. Событие $\overline{A}B$ означает, что на трех костях выпали разные числа и нет ни одной шестерки. Аналогично предыдущему, имеем $\mathbf{P}(\overline{A}B) = A_5^3/216$. По определению условной вероятности $\mathbf{P}(\overline{A}|B) = \mathbf{P}(\overline{A}B)/\mathbf{P}(B) = A_5^3/A_6^3 = \frac{5\cdot4\cdot3}{6\cdot5\cdot4} = \frac{1}{2}$. По свойству 5) окончательно получаем $\mathbf{P}(A|B) = 1 - \frac{1}{2} = \frac{1}{2}$.

Определение 2. Событие A называется **независимым** от события B с $\mathbf{P}(B) \neq 0$, если $\mathbf{P}(A|B) = \mathbf{P}(A)$, т. е. вероятность наступления события A не зависит от того, произошло событие B или нет.

Пример. Пусть при бросании игральной кости $A = \{$выпало число меньше трех$\}, B = \{$выпало четное число$\}$. Поскольку $A = \{1,2\}$, а $B = \{2,4,6\}$, то $\mathbf{P}(A) = \frac{1}{3}$ и $\mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)} = \frac{1/6}{1/2} = \frac{1}{3}$. Следовательно, событие A не зависит от события B.

Лемма 1 (о взаимной независимости событий). *Если событие A не зависит от B при $\mathbf{P}(A) \neq 0, \mathbf{P}(B) \neq 0$, то и событие B не зависит от A.*

Доказательство. Используя свойство 7), получим

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

и, следовательно, событие B не зависит от A.

Таким образом, события A и B не зависят друг от друга. При этом согласно свойству 7)

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

Это равенство позволяет дать следующее определение независимости событий, симметричное по отношению к событиям A и B и применимое к событиям нулевой вероятности.

Определение 3. События A и B называются **независимыми**, если

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

Это определение удобно тем, что его легко можно распространить на совокупность нескольких событий.

Определение 4. События $A_1, A_2, \ldots, A_n$ называются **независимыми в совокупности**, если для любого набора индексов $1 \leqslant l_1 < l_2 < \cdots < l_k \leqslant n$ выполняется равенство

$$\mathbf{P}(A_{l_1}A_{l_2}\ldots A_{l_k}) = \mathbf{P}(A_{l_1})\mathbf{P}(A_{l_2})\ldots\mathbf{P}(A_{l_k}).$$

Определение 5. События $A_1, A_2, \ldots, A_n$ называются **попарно независимыми**, если $\mathbf{P}(A_i A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$ для любых пар $i, j, 1 \leqslant i < j \leqslant n$.

Замечание 1. Из попарной независимости событий $A_1, A_2, \ldots A_{n-1}, A_n$ не следует независимость этих событий в совокупности.

Убедимся в справедливости этого замечания на примере бросания двух игральных костей. Определим три события следующим образом: $A=\{$нечетное число на первой кости$\}$; $B = \{$нечетное число на второй кости$\}$; $C = \{$нечетная сумма$\}$. Применяя теорему 4 § 1 о числе комбинаций для двух групп, имеем:

$$\mathbf{P}(A) = \ (B) = \frac{3\cdot 6}{6\cdot 6} = \frac{1}{2}; \quad \mathbf{P}(C) = \frac{3\cdot 3 + 3\cdot 3}{6\cdot 6} = \frac{1}{2};$$

$$\mathbf{P}(AB) = \frac{3\cdot 3}{6\cdot 6} = \frac{1}{4}; \quad \mathbf{P}(AC) = \mathbf{P}(BC) = \frac{3\cdot 3}{6\cdot 6} = \frac{1}{4}.$$

Таким образом, эти три события попарно независимы. Однако $ABC = \varnothing$ и, следовательно, $\mathbf{P}(ABC) = 0$, а это означает, что вероятность произведения ABC не равна произведению вероятностей $\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \frac{1}{8}$.

Задача 3. Доказать, что если события A, B и C независимы в совокупности, то события $A, \overline{B}, \overline{C}$ также независимы в совокупности, и независимы в совокупности события $\overline{A}, B$, C.

Решение. Проверим сначала, что вероятность произведения событий $A, \overline{B}, \overline{C}$ распадается в произведение вероятностей. Применяя следствие 3 § 3, получим

$$\begin{aligned}
\mathbf{P}(A\overline{B}\,\overline{C}) &= \mathbf{P}(A\overline{B}(\Omega - C)) = \mathbf{P}(A\overline{B} - A\overline{B}C) = \\
&= \mathbf{P}(A\overline{B}) - \mathbf{P}(A\overline{B}C) = \mathbf{P}(A(\Omega - B)) - \mathbf{P}(A(\Omega - B)C) = \\
&= \mathbf{P}(A) - \mathbf{P}(AB) - \mathbf{P}(AC) + \mathbf{P}(ABC) = \\
&= \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) - \mathbf{P}(A)\mathbf{P}(C) + \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \\
&= \mathbf{P}(A)(1 - \mathbf{P}(B))(1 - \mathbf{P}(C)) = \mathbf{P}(A)\mathbf{P}(\overline{B})\mathbf{P}(\overline{C}).
\end{aligned}$$

Теперь нужно убедиться, что для произведений любых пар событий, т. е. для $A\overline{B}$, $A\overline{C}$ и $\overline{B}\,\overline{C}$ вероятность тоже распадается в произведение вероятностей. Действительно

$$\mathbf{P}(A\overline{B}) = \mathbf{P}(A) - \mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(\overline{B}),$$

$$\mathbf{P}(A\overline{C}) = \mathbf{P}(A) - \mathbf{P}(AC) = \mathbf{P}(A)\mathbf{P}(\overline{C}).$$

Применяя 11-ое свойство операций над событиями и формулу сложения вероятностей, получим

$$\mathbf{P}(\overline{B}\,\overline{C}) = \mathbf{P}(\overline{B+C}) = 1 - \mathbf{P}(B+C) =$$
$$= 1 - \mathbf{P}(B) - \mathbf{P}(C) + \mathbf{P}(BC) =$$
$$= (1 - \mathbf{P}(B))(1 - \mathbf{P}(C)) = \mathbf{P}(\overline{B})\mathbf{P}(\overline{C}).$$

Аналогично можно доказать независимость в совокупности событий $\overline{A}, B, C$.

Задача 4. Доказать, что если события A, B и C независимы в совокупности, то события $A - B$, C независимы, а также независимы события $A + B$, C.

Р е ш е н и е. Воспользуемся предыдущими результатами и восьмым свойством операций над событиями. Тогда

$$\mathbf{P}((A-B)C) = \mathbf{P}(A\overline{B}C) = \mathbf{P}(A)\mathbf{P}(\overline{B})\mathbf{P}(C) =$$

$$= \mathbf{P}(A\overline{B})\mathbf{P}(C) = \mathbf{P}(A-B)\mathbf{P}(C).$$

Независимость событий $A + B$, C также несложно доказать.

Приведенные в задачах утверждения свидетельствуют в пользу следующего факта, который мы сформулируем без доказательства.

Пусть n независимых в совокупности событий разбиты на k групп, не содержащих общих событий, $k \leqslant n$. Над событиями внутри каждой группы произведены любые операции: сложения, умножения, вычитания и т. д. Тогда получившиеся в результате k новых событий будут тоже независимыми в совокупности.

Этот факт отвечает нашим интуитивным представлениям о независимости.

Задачи

Задача 5.1. Из колоды в 52 карты вытаскивается одна. Зависимы или нет события $A = \{\text{вытащен туз}\}$ и $B = \{\text{вытащена карта красной масти}\}$?

Решить ту же задачу, но с колодой в 54 карты (в колоду добавлены два джокера, каждый из которых не имеет цвета и не является тузом).

Задача 5.2. Пусть события A и B_1 независимы и независимы также события A и B_2, при этом $B_1B_2 = \varnothing$. Доказать, что события A и $B_1 + B_2$ независимы.

Задача 5.3. Брошены последовательно три монеты. Определить, зависимы или независимы события:

$A = \{$выпадение "герба" на первой монете$\}$;

$B = \{$выпадение хотя бы одной "решки"$\}$.

Задача 5.4. Одновременно бросают два кубика (цифры 1, 2,...,6). Известно, что выпала сумма очков, равная восьми. Какова вероятность того, что выпали 3 и 5?

Задача 5.5. Какова вероятность при бросании двух кубиков получить четную сумму при условии, что хотя бы на одном выпало 6 очков.

Задача 5.6. В урне 9 белых шаров и 1 черный шар. Вынули сразу три шара. Какова вероятность того, что все шары белые?

Задача 5.7. Из колоды в 36 карт одну за другой вытаскивают 2 карты. Какова вероятность, что среди них два туза?

Задача 5.8. Из колоды в 36 карт одну за другой вытаскивают три карты. Какова вероятность того, что первой картой будет туз, второй – король, третьей – дама?

Задача 5.9. Буквы слова ПОКОЛЕНИЕ выписаны на карточках. Наудачу вынимают одну карточку за другой и укладывают по порядку. Найти вероятность того, что получится слово ПОЛЕ.

Задача 5.10. В двух ящиках находятся шары, отличающиеся только цветом, причем в первой урне 5 белых шаров, 11 черных и 8 красных, а во второй соответственно 10, 8 и 6. Из каждого ящика наудачу извлекается по одному шару. Какова вероятность, что оба шара одного цвета?

Задача 5.11. Два стрелка поочередно стреляют по мишени до первого попадания каждым стрелком. При каждом выстреле вероятность попадания для первого стрелка равна 0.2, а для второго равна 0.3. Найти вероятность того, что первый стрелок сделает больше выстрелов, чем второй.

Задача 5.12. Партия из ста деталей подвергается выборочному контролю. Условием непригодности всей партии является наличие хотя бы одной бракованной детали среди пяти проверяемых. Какова вероятность для данной партии быть непринятой, если она содержит 5% бракованных деталей?

§ 6. Общее определение вероятности

До сих пор мы рассматривали такие примеры случайных экспериментов, в которых все исходы были равновозможны. Однако для многих экспериментов исходы не являются равновозможными. Приведем простой пример.

Пример. Пусть есть однородная проволока длиной 1 метр с несколькими узлами, например, с тремя. Тогда мы не можем считать, что при растяжении проволоки за концы вероятность разрыва в узле будет такой же, как и в остальных точках проволоки. Вероятнее всего проволока будет рваться в узлах. Пусть все узлы одинаковы, и вероятность разрыва в каждом из них равна положительному числу p, такому, что $p < 1/3$, поскольку вероятность всех исходов должна быть равной единице. Разрывы в остальных точках ввиду однородности проволоки равновозможны, поэтому вероятность разрыва в A – множестве точек проволоки, не содержащем узлы, должна быть пропорциональна его длине, т. е. пропорциональна $\operatorname{mes} A$. Отсюда легко понять, что если множество A не содержит узлы, то естественно положить $\mathbf{P}(A) = (1-3p)\operatorname{mes} A$, если A содержит один узел, то $\mathbf{P}(A) = (1-3p)\operatorname{mes} A + p$, если A содержит два узла, то $\mathbf{P}(A) = (1-3p)\operatorname{mes} A + 2p$, и если A содержит три узла, то $\mathbf{P}(A) = (1-3p)\operatorname{mes} A + 3p$.

В этом примере мы встретились со смешанной ситуацией, в которой участвуют как геометрические вероятности, так и вероятности, отвечающие классической схеме. Это достаточно простой пример. В сложных примерах невозможно выписать явные формулы для вероятностей произвольных событий, однако можно сформулировать общие свойства, присущие вероятностям любых событий.

Перейдем к изучению основ теории вероятностей в общем случае Естественно ожидать, что они будут учитывать все известные нам результаты для частных случаев.

Пусть Ω – произвольное множество, природа элементов которого нам совершенно безразлична. Оно может быть как счетным, так и несчетным. Предположим, что Ω является пространством всех элементарных исходов для какого-нибудь случайного эксперимента, каждому результату которого соответствует ровно одна точка из Ω, а разным результатам отвечают разные точки. Выделим некоторую совокупность подмножеств Ω, обозначим ее $\mathcal{A}$. Множества из $\mathcal{A}$ будем трактовать как случайные события. Естественно потребовать, чтобы события, полученные из $\mathcal{A}$ в результате рассмотренных нами выше (в § 2) операций над событиями, снова принадлежали бы $\mathcal{A}$. Это, конечно, налагает на совокупность множеств $\mathcal{A}$ некоторые условия, что приводит нас к следующим ниже определениям алгебры и σ-алгебры событий.

Определение 1. Совокупность $\mathcal{A}$ подмножеств множества Ω называется **алгеброй**, если выполнены следующие условия: 1) $\Omega \in \mathcal{A}$; 2) если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, то $A + B \in \mathcal{A}$; 3) если $A \in \mathcal{A}$, то $\overline{A} \in \mathcal{A}$.

Пример. Пусть $\Omega = [0, 1)$. Тогда на отрезке $[0, 1)$ все множества, состоящие из конечного числа интервалов вида $[a, b)$, образуют алгебру.

Оказывается, что условий 1) – 3) достаточно для того, чтобы другие операции над событиями из $\mathcal{A}$ не выводили бы нас за пределы алгебры $\mathcal{A}$. Это вытекает из следующего утверждения.

Предложение 1. *Пусть $\mathcal{A}$ – алгебра. Тогда а) $\varnothing \in \mathcal{A}$; б) если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, то $AB \in \mathcal{A}$; в) если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, то $A - B \in \mathcal{A}$; г) если $A_l \in \mathcal{A}, l = 1, 2, \ldots, n$, то $\sum\limits_{l=1}^{n} A_l \in \mathcal{A}$; и $\prod\limits_{l=1}^{n} A_l \in \mathcal{A}$.*

Доказательство. а) В силу условий 1), 3) имеем $\varnothing = \overline{\Omega} \in \mathcal{A}$.

б) По свойству 10) § 2 $\overline{AB} = \overline{A} + \overline{B}$, и, следовательно, $AB = \overline{\overline{AB}} = \overline{\overline{A} + \overline{B}}$. Согласно условию 3) $\overline{A} \in \mathcal{A}, \overline{B} \in \mathcal{A}$, а тогда по условию 2) $\overline{A} + \overline{B} \in \mathcal{A}$. Снова воспользовавшись условием 3), имеем $\overline{\overline{A} + \overline{B}} \in \mathcal{A}$, а это и означает, что $AB \in \mathcal{A}$.

в) В силу определения разности событий $A - B = A\overline{B}$. Согласно условию 3) $\overline{B} \in \mathcal{A}$ и по свойству б) $A\overline{B} \in \mathcal{A}$.

г) Первое включение устанавливается с помощью многократного применения условия 2) Если $A_1 + A_2 \in \mathcal{A}$ и $A_3 \in \mathcal{A}$, то по условию 2) выполняется $(A_1 + A_2) + A_3 \in \mathcal{A}$ и т. д. Второе включение устанавливается аналогично с помощью многократного применения свойства б). Если $A_1A_2 \in \mathcal{A}$ и $A_3 \in \mathcal{A}$, то по свойству б) $(A_1A_2)A_3 \in \mathcal{A}$ и т. д.

Из предложения 1 следует, что любое конечное число известных нам операций над событиями из $\mathcal{A}$ не выводит за пределы алгебры. Однако при рассмотрении многих задач теории вероятностей приходится иметь дело и с бесконечным числом операций.

Пример. Пусть эксперимент состоит в последовательном бросании монеты. Нас интересует событие $A = \{$хотя бы один раз выпал герб$\}$. Определим события $A_k = \{$в первые $k-1$ бросаний выпала решетка, а при k-м бросании выпал герб$\}$, $k = 1, 2, \ldots$. Тогда ясно, что $A = \sum\limits_{k=1}^{\infty} A_k$.

Для того, чтобы можно было рассматривать бесконечное число операций над событиями, необходимо усилить ограничения, налагаемые на алгебру $\mathcal{A}$.

Определение 2. Совокупность множеств $\mathcal{A}$ называется **σ-алгеброй**, если выполнены условия 1), 3) из определения алгебры и 2') если $A_l \in \mathcal{A}, l = 1, 2, \ldots$, то $\sum\limits_{l=1}^{\infty} A_l \in \mathcal{A}$.

Условие 2') сильнее, чем 2), поскольку из 2') вытекает 2). Действительно, пусть $A_1 \in \mathcal{A}$ и $A_2 \in \mathcal{A}$. Положим $A_l = \varnothing$ при $l \geqslant 3$. Тогда $A_1 + A_2 = \sum\limits_{l=1}^{\infty} A_l$, поскольку добавляя в сумму невозможные события $\varnothing$, мы ничего не меняем. Теперь из условия 2') очевидно следует 2).

Используя условие 3) и равенство $\overline{\prod\limits_{l=1}^{\infty} A_l} = \sum\limits_{l=1}^{\infty} \overline{A_l}$, легко убедиться в справедливости следующего утверждения.

Предложение 2. *Пусть $\mathcal{A}$ – σ-алгебра. Тогда, если $A_l \in \mathcal{A}, l = 1, 2, \ldots$, то $\prod\limits_{l=1}^{\infty} A_l \in \mathcal{A}$.*

Таким образом, счетное число операций суммирования или перемножения событий не выводит за пределы σ-алгебры.

Событиям из σ-алгебры $\mathcal{A}$ сопоставляются их вероятности. Вид явных формул, конечно, зависит от конкретной задачи, однако и в общей ситуации можно сформулировать естественные аксиомы и вывести некоторые свойства вероятностей, которые будут выполняться для всех частных случаев.

Определение 3. Если каждому множеству из $\mathcal{A}$ сопоставлено некоторое число, то будем говорить, что на $\mathcal{A}$ задана **функция множеств**.

Определение 4. **Вероятностью** называется функция множеств, заданная на σ-алгебре $\mathcal{A}$ подмножеств из Ω и удовлетворяющая следующим условиям:

1) $\mathbf{P}(A) \geqslant 0$ для любого $A \in \mathcal{A}$;

2) $\mathbf{P}(\Omega) = 1$;

3) если события $A_1, A_2, \ldots$ попарно несовместны, т. е. $A_i A_j = \varnothing$ для любых $i \neq j$, то

$$\mathbf{P}\Big(\sum_{l=1}^{\infty} A_l\Big) = \sum_{l=1}^{\infty} \mathbf{P}(A_l).$$

Условия 1), 2′), 3) в определении σ-алгебры $\mathcal{A}$ и условия 1), 2), 3) в определении вероятности составляют **аксиомы теории вероятностей**. Тройка $\big(\Omega, \mathcal{A}, \mathbf{P}\big)$, удовлетворяющая этим аксиомам, называется **вероятностным пространством**. Остальные, уже хорошо известные нам по §3, свойства вероятности выводятся из аксиом. В §3 мы вывели эти свойства, опираясь на классическое определение вероятности для случая конечного числа исходов эксперимента. Здесь мы выведем эти свойства, исходя из аксиом теории вероятностей.

Свойства вероятностей.

1) $\mathbf{P}(\Omega) = 1$.

Это свойство входит в определение вероятности.

2) $\mathbf{P}(\varnothing) = 0$.

Действительно, по аксиоме 3) $\mathbf{P}(\varnothing + \sum\limits_{l=1}^{\infty} A_l) = \mathbf{P}(\varnothing) + \sum\limits_{l=1}^{\infty} \mathbf{P}(A_l)$ и $\mathbf{P}(\sum\limits_{l=1}^{\infty} A_l) = \sum\limits_{l=1}^{\infty} \mathbf{P}(A_l)$. Кроме того,

$\varnothing + \sum\limits_{l=1}^{\infty} A_l = \sum\limits_{l=1}^{\infty} A_l$, поэтому левые части этих равенств совпадают. Отсюда следует совпадение правых частей и выполнение требуемого свойства.

Следствие 1. *Если события A_1 и A_2 несовместны, то*

$$\mathbf{P}(A_1 + A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2).$$

Это утверждение вытекает из свойства 2) и аксиомы 3), если в нем положить $A_3 = \varnothing, A_4 = \varnothing, \ldots$.

Следствие 2. *Если события $A_l, l = 1, \ldots, n$, – попарно несовместны, т. е. $A_iA_j = \varnothing$ для любых $i \neq j$, то*

$$\mathbf{P}\Big(\sum_{l=1}^{n} A_l\Big) = \sum_{l=1}^{n} \mathbf{P}(A_l).$$

3) если $A \subset B$, то $\mathbf{P}(A) \leqslant \mathbf{P}(B)$.

Воспользуемся тем, что $B = (B - A) + A$ при $A \subset B$. Поскольку события $B - A$ и A несовместны, то, применяя следствие 1 и аксиому 1), получим $\mathbf{P}(B) = \mathbf{P}(B - A) + \mathbf{P}(A) \geqslant \mathbf{P}(A)$.

4) $0 \leqslant \mathbf{P}(A) \leqslant 1$ для любого события A.

Это свойство выполняется в силу свойства 3) и аксиом 1), 2).

5) $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.

Поскольку события A и $\overline{A}$ несовместны и $\Omega = A + \overline{A}$, то $1 = \mathbf{P}(\Omega) = \mathbf{P}(A + \overline{A}) = \mathbf{P}(A) + \mathbf{P}(\overline{A})$.

6) (**формула сложения вероятностей**) для любых событий A и B

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

Доказательство. Справедливы следующие представления: $A+B = A+(B-A)$, где события A и $B-A$ несовместны; $B = AB + (B - A)$, где события AB и $B - A$ несовместны. Поэтому в силу следствия 1 имеем $\mathbf{P}(A+B) = \mathbf{P}(A)+\mathbf{P}(B-A)$ и $\mathbf{P}(B) = \mathbf{P}(AB)+\mathbf{P}(B-A)$. Исключая из этих двух равенств $\mathbf{P}(B - A)$, получим формулу сложения вероятностей.

Следствие 3. *Для любых событий* $A_l, l = 1, 2, \ldots, n,$

$$\mathbf{P}\Big(\sum_{l=1}^{n} A_l\Big) \leqslant \sum_{l=1}^{n} \mathbf{P}(A_l).$$

Для $n = 2$ эта оценка следует из свойства 6). Тогда для произвольного $n \geqslant 2$ можно написать

$$\mathbf{P}\Big(\sum_{l=1}^{n} A_l\Big) \leqslant \mathbf{P}\Big(\sum_{l=1}^{n-1} A_l\Big) + \mathbf{P}(A_n).$$

Для оценки вероятности суммы $n-1$ событий снова можно воспользоваться такой же оценкой, и т. д. В результате получим требуемую оценку.

6′) **(общая формула сложения вероятностей)** для любых событий $A_l, l = 1, 2, \ldots, n,$

$$\mathbf{P}\Big(\sum_{l=1}^{n} A_l\Big) = \sum_{l=1}^{n} \mathbf{P}(A_l) - \sum_{l<j} \mathbf{P}(A_l A_j) +$$
$$+ \sum_{l<j<k} \mathbf{P}(A_l A_j A_k) - \cdots + (-1)^{n-1} \mathbf{P}\Big(\prod_{l=1}^{n} A_l\Big).$$

Эту формулу можно вывести из формулы 6) с помощью метода математической индукции. Нужно положить $A_1 + A_2 + \cdots + A_{n-1} = A, A_n = B$, воспользоваться формулой 6) и для $n-1$ слагаемых применить общую формулу сложения вероятностей, предположив, что она верна. Подробно проделаем это лишь для $n = 3$. Дважды применяя формулу 6), имеем

$$\mathbf{P}(A_1 + A_2 + A_3) = \mathbf{P}(A_1 + A_2) + \mathbf{P}(A_3) - \mathbf{P}((A_1 + A_2)A_3) =$$
$$= \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 A_2) + \mathbf{P}(A_3) - \mathbf{P}(A_1 A_3 + A_2 A_3).$$

Для последнего слагаемого, снова применяя формулу 6), получим

$$\mathbf{P}(A_1 A_3 + A_2 A_3) = \mathbf{P}(A_1 A_3) + \mathbf{P}(A_2 A_3) - \mathbf{P}(A_1 A_3 A_2 A_3).$$

Учитывая, что $A_1 A_3 A_2 A_3 = A_1 A_2 A_3$ и подставляя это выражение в предыдущую формулу, окончательно будем иметь

$$\mathbf{P}(A_1 + A_2 + A_3) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) -$$
$$- \mathbf{P}(A_1 A_2) - \mathbf{P}(A_1 A_3) - \mathbf{P}(A_2 A_3) + \mathbf{P}(A_1 A_2 A_3).$$

7) свойства непрерывности:

если $A_1 \subset A_2 \subset A_3 \subset \ldots$, то $\lim\limits_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}\Big(\sum\limits_{k=1}^{\infty} A_k\Big)$;

если $A_1 \supset A_2 \supset A_3 \supset \ldots$, то $\lim\limits_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}\Big(\prod\limits_{k=1}^{\infty} A_k\Big)$.

Доказательство. Поскольку $A_{k-1} \subset A_k$, то

$$\sum_{k=1}^{\infty} A_k = \sum_{k=1}^{\infty}(A_k - A_{k-1}),$$

где $A_0 = \varnothing$, и события $A_k - A_{k-1}$, $k = 1, 2, \ldots$ попарно несовместны. Тогда по третьей аксиоме вероятности

$$\mathbf{P}\Big(\sum_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mathbf{P}\big(A_k - A_{k-1}\big) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{P}\big(A_k - A_{k-1}\big) =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n}\big(\mathbf{P}(A_k) - \mathbf{P}(A_{k-1})\big) = \lim_{n\to\infty}\big(\mathbf{P}(A_n) - \mathbf{P}(\varnothing)\big) = \lim_{n\to\infty} \mathbf{P}(A_n).$$

Докажем второе свойство непрерывности. Поскольку $A_{k-1} \supset A_k$, то $\overline{A}_{k-1} \subset \overline{A}_k$, и по первому свойству непрерывности

$$\lim_{n\to\infty} \mathbf{P}(\overline{A}_n) = \mathbf{P}\Big(\sum_{k=1}^{\infty} \overline{A}_k\Big) = \mathbf{P}\Big(\overline{\prod_{k=1}^{\infty} A_k}\Big).$$

Последнее равенство справедливо в силу свойства 12′) § 2 ($n = \infty$). Переходя к дополнительным вероятностям, получаем

$$\lim_{n\to\infty}\big(1 - \mathbf{P}(A_n)\big) = 1 - \mathbf{P}\Big(\prod_{k=1}^{\infty} A_k\Big),$$

а это и есть второе свойство непрерывности.

Лемма 1. *Пусть события Ω_m таковы, что $\mathbf{P}(\Omega_m) = 1$ для любого натурального m. Тогда*

$$\mathbf{P}\Big(\prod_{m=1}^{\infty} \Omega_m\Big) = 1.$$

Доказательство. Применяя свойство 12′) § 2 и следствие 3, получим

$$0 \leqslant \mathbf{P}\left(\overline{\prod_{m=1}^{n} \Omega_m}\right) = \mathbf{P}\left(\sum_{m=1}^{n} \overline{\Omega}_m\right) \leqslant \sum_{m=1}^{n} \mathbf{P}(\overline{\Omega}_m) = 0,$$

так как $\mathbf{P}(\overline{\Omega}_m) = 1 - \mathbf{P}(\Omega_m) = 0$. По пятому свойству вероятностей $\mathbf{P}\left(\prod\limits_{m=1}^{n} \Omega_m\right) = 1$. Поскольку при $k = 1,2,\ldots$ $\prod\limits_{m=1}^{k-1} \Omega_m \supset \prod\limits_{m=1}^{k} \Omega_m$, то по второму свойству непрерывности вероятности окончательно имеем

$$\mathbf{P}\left(\prod_{m=1}^{\infty} \Omega_m\right) = \lim_{n\to\infty} \mathbf{P}\left(\prod_{m=1}^{n} \Omega_m\right) = 1.$$

По аналогии с классическим случаем дадим следующее определение условной вероятности.

Определение 5. **Условной вероятностью** события A при условии, что произошло событие B с $\mathbf{P}(B) \neq 0$, называется число $\mathbf{P}(A|B)$, которое определяется формулой

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

Используя это определение и свойства вероятности, несложно убедиться, что условная вероятность обладает всеми свойствами вероятности.

Свойства условных вероятностей.

1) $\mathbf{P}(\Omega|B) = 1$; 2) $\mathbf{P}(\varnothing|B) = 0$; 3) $0 \leqslant \mathbf{P}(A|B) \leqslant 1$.

Эти свойства следуют непосредственно из определения.

4) если $A \subset C$, то $\mathbf{P}(A|B) \leqslant \mathbf{P}(C|B)$.

Действительно, из $A \subset C$ следует $AB \subset CB$, и по третьему свойству вероятностей $\mathbf{P}(AB) \leqslant \mathbf{P}(CB)$. Поделив это неравенство на $\mathbf{P}(B)$ и применив определение условной вероятности, получим требуемое утверждение.

5) $\mathbf{P}(\overline{A}|B) = 1 - \mathbf{P}(A|B)$.

Это свойство является следствием цепочки равенств:

$$1 = \frac{\mathbf{P}(\Omega B)}{\mathbf{P}(B)} = \frac{\mathbf{P}((\overline{A} + A)B)}{\mathbf{P}(B)} =$$

$$= \frac{\mathbf{P}(\overline{A}B + AB)}{\mathbf{P}(B)} = \frac{\mathbf{P}(\overline{A}B) + \mathbf{P}(AB)}{\mathbf{P}(B)} = \mathbf{P}(\overline{A}|B) + \mathbf{P}(A|B).$$

6) (**формула сложения условных вероятностей**) для любых событий A и C

$$\mathbf{P}(A+C|B)=\mathbf{P}(A|B)+\mathbf{P}(C|B)-\mathbf{P}(AC|B).$$

Доказательство. Применяя формулу сложения обычных вероятностей, имеем

$$\mathbf{P}((A+C)B)=\mathbf{P}(AB+CB)=\mathbf{P}(AB)+\mathbf{P}(CB)-\mathbf{P}(ACB).$$

Поделив левую и правую части этого равенства на $\mathbf{P}(B)$, получим формулу 6).

7) (**формула умножения вероятностей**) для любых событий A и B

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

Формула непосредственно следует из определения условной вероятности.

8) (**общая формула умножения вероятностей**) для любых событий $A_1, A_2, \ldots, A_n$

$$\mathbf{P}(A_1A_2\ldots A_n)=$$
$$=\mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1A_2)\ldots\mathbf{P}(A_n|A_1A_2\ldots A_{n-1}).$$

Вывод этой формулы из формулы 7) был осуществлен в § 5.

Сохраняются в силе все связанные с независимостью событий определения, сформулированные в § 5.

Большинство задач, которые возникают в теории вероятностей, связано, по-существу, с пересчетом вероятностей. Нужно подсчитать вероятность, возникающую в какой-нибудь сложной модели, когда вероятности, отвечающие простым компонентам этой модели, известны. При этом нас не интересует природа возникновения этих элементарных вероятностей. Их величина может быть получена из соображений, связанных с симметрией и однородностью, или их приближенное значение может быть вычислено с помощью методов математической статистики.

Рассмотрим одну задачу, решение которой опирается на полученные выше общие свойства вероятностей.

Задача 1 (о расчете надежности электрических цепей). Пусть есть некоторая цепь элементов, соединенных последовательно и параллельно. Например, цепь вида

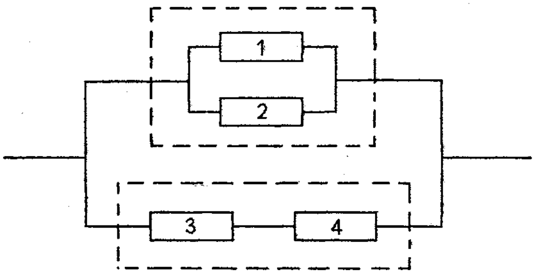

Рис. 3

Пусть известно, что элемент l выходит из строя (происходит разрыв цепи в этом элементе) с вероятностью p_l. Естественно также предположить, что все элементы выходят из строя независимо друг от друга. Какова вероятность того, что вся цепь выйдет из строя, т. е. через нее не пойдет ток?

Решение. Сначала нужно рассчитать вероятности отказа элементарных цепей, т. е. цепей вида

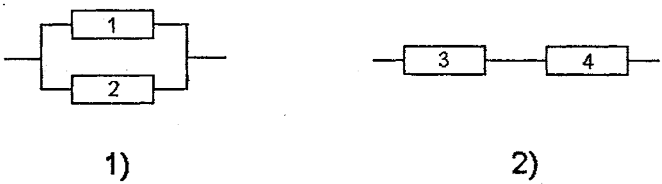

Рис. 4

Введем события $A_l = \{$выход из строя элемента $l\}$. Тогда $\mathbf{P}(A_l) = p_l$. Элементарная цепь вида 1) выходит из строя, когда события A_1 и A_2 наступают одновременно. Следовательно, $\mathbf{P}$(цепь 1) выходит из строя) $= \mathbf{P}(A_1A_2) =$

$\mathbf{P}(A_1)\mathbf{P}(A_2) = p_1p_2$. Элементарная цепь вида 2) выходит из строя, когда наступает либо событие A_3, либо A_4. Поэтому $\mathbf{P}$(цепь 2) выходит из строя) $= \mathbf{P}(A_3+A_4) = \mathbf{P}(A_3)+\mathbf{P}(A_4)-\mathbf{P}(A_3A_4) = p_3+p_4-p_3p_4$. Нашу сложную цепь можно упростить с помощью укрупнения. Это означает, что мы элементарные цепи рассматриваем как новые элементы (они выделены на рис. 4 пунктирной линией), вероятности отказа которых мы уже знаем: это $P_1 = p_1p_2$ и $P_2 = p_3+p_4-p_3p_4$. Но упрощенная цепь будет снова вида 1), и вероятность ее отказа будет равна $P_1P_2 = p_1p_2(p_3+p_4-p_3p_4)$.

Методом последовательного укрупнения схем можно рассчитывать вероятности отказа любых цепей, если вероятности отказов элементов известны.

Задача 2 (о разорении игрока). Рассмотрим игру в “орлянку”, когда игрок выбирает “герб”или “решетку”, после чего бросается монета. Если выпадает сторона монеты, названная игроком, то он выигрывает, получая 1 руб.; в противном случае столько же проигрывает. Пусть начальный капитал игрока составляет x руб. и игрок ставит себе целью довести его до некоторой суммы s руб., большей чем x руб. Игра продолжается до тех пор, пока либо игрок наберет заранее определенную сумму s, либо разорится, проиграв весь имеющийся у него капитал. Какова вероятность того, что игрок разорится?

Решение. Введем события:
$H_1 = \{$игрок выиграл на первом шаге$\}$,
$H_2 = \{$игрок проиграл на первом шаге$\}$,
$A = \{$игрок разорится, имея начальный капитал x руб.$\}$.

Обозначим через $p_x = \mathbf{P}(A)$ искомую вероятность. Очевидно, эта вероятность определена для любого целого x, $0 \leqslant x \leqslant s$ и очевидно $p_0 = 1, p_s = 0$.

Поскольку монета симметрична, то $\mathbf{P}(H_1) = \mathbf{P}(H_2) = \frac{1}{2}$. Если наступает событие H_1, то капитал игрока стал $x+1$, а если наступает событие H_2, то капитал игрока стал $x-1$. После этого можно считать, что игра начинается снова, и согласно принятому обозначению $\mathbf{P}(A|H_1) = p_{x+1}$, $\mathbf{P}(A|H_2) = p_{x-1}$, где x – любое число $1 \leqslant x \leqslant s-1$.

Поскольку $A = A\Omega = A(H_1+H_2) = AH_1 + AH_2$ и события H_1 и H_2 несовместны, то по формуле сложения вероятностей

$\mathbf{P}(A) = \mathbf{P}(AH_1) + \mathbf{P}(AH_2)$. Применяя формулу умножения вероятностей, получим

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

Это в силу принятых обозначений дает следующее рекуррентное уравнение для вероятности p_x: $p_x = \frac{1}{2}p_{x+1} + \frac{1}{2}p_{x-1}$, $1 \leqslant x \leqslant s-1$. Известно, что решением этого уравнения является только линейная функция $p_x = c_1 + c_2 x$, где c_1, c_2 – произвольные коэффициенты. Эти коэффициенты можно определить, подставляя в выражение для p_x значения $x = 0$ и $x = s$ и пользуясь граничными условиями $p_0 = 1, p_s = 0$. Имеем $c_1 = 1, c_1 + c_2 s = 0$. Откуда следует, что $c_1 = 1, c_2 = -\frac{1}{s}$ и, значит, $p_x = 1 - \frac{x}{s}$, $0 \leqslant x \leqslant s$.

Задачи

Задача 6.1. Двое бросают монету. Выигрывает тот, у кого первого выпадет герб. Найти вероятность выигрыша для каждого.

Задача 6.2. Прибор содержит четыре узла: A_1, A_2, A_3, A_4. Узел A_2 дублирует A_1, а A_4 дублирует A_3. При отказе происходит автоматическое переключение. Надежность переключающего устройства равна p. Надежность в течение заданного времени каждого из узлов равна $p_i, i = 1, 2, 3, 4$. Вычислить надежность прибора.

Задача 6.3. Вероятность хотя бы одного попадания в мишень при трех выстрелах равна 0.875. Найти вероятность попадания при одном выстреле.

Задача 6.4. Для того, чтобы сбить самолет достаточно одного попадания. Было сделано три выстрела с вероятностями попадания 0.1, 0.2 и 0.4 соответственно. Какова вероятность того, что самолет сбит?

Задача 6.5. Для того, чтобы разрушить мост, нужно попадание не менее двух бомб. Независимо сбросили три бомбы с вероятностями попадания 0.1, 0.3 и 0.4. Какова вероятность, что мост разрушен?

Задача 6.6. Рассчитать вероятность отказа цепи (не идет ток),

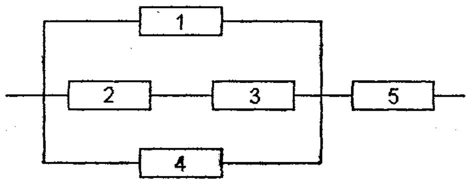

Рис. 5

где p_i – вероятность отказа i-го элемента.

Задача 6.7. Рассчитать вероятность отказа цепи (не идет ток),

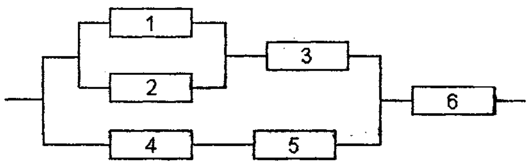

Рис. 6

где p_i – вероятность отказа i-го элемента.

Задача 6.8. Какова вероятность того, что при многократном бросании игральной кости шестерка впервые выпадет на четвертом броске?

Задача 6.9. Какова вероятность того, что при многократном бросании правильной монеты герб впервые выпадет на четном броске?

§ 7. Формула полной вероятности и формула Байеса

При решении вероятностных задач часто возникает следующая ситуация. Все пространство элементарных событий можно разбить на события, не содержащие общих исходов, тем самым выделяется группа несовместных событий, которые в совокупности составляют достоверное событие. Пусть известно, каким образом вычислить условную вероятность при условии, что то или иное событие из этой группы произошло. Пусть, кроме того, ясно как вычислить вероятности самих событий, входящих в эту группу. В таком случае очень полезным оказывается результат, получивший название "формула полной вероятности".

Определение 1. Набор событий $H_1, H_2, \ldots, H_n$ называется **полной группой событий**, если они попарно несовместны и их сумма составляет достоверное событие:

$$H_1 + H_2 + \cdots + H_n = \Omega.$$

Теорема 1 (формула полной вероятности). *Пусть* $H_l, l = 1, 2, \ldots, n$, *– полная группа событий, и* $\mathbf{P}(H_l) > 0$. *Тогда для любого события* A

$$\mathbf{P}(A) = \sum_{l=1}^{n} \mathbf{P}(H_l)\mathbf{P}(A|H_l).$$

Доказательство. В силу свойств операций над событиями (§ 2)

$$A = A\Omega = A(H_1 + H_2 + \cdots + H_n) = AH_1 + AH_2 + \cdots + AH_n.$$

Поскольку события H_l попарно несовместны, то и события AH_l тоже попарно несовместны. Тогда в силу следствия 2 § 6

$$\mathbf{P}(A) = \mathbf{P}(AH_1 + AH_2 + \cdots + AH_n) = \sum_{l=1}^{n} \mathbf{P}(AH_l).$$

Теперь, применяя к слагаемым $\mathbf{P}(AH_l)$ формулу умножения вероятностей (§ 6),

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

получим утверждение теоремы.

Задача 1. Пусть в коробке есть 3 новых и 3 уже использованных теннисных мяча. Для первой игры наудачу берут из коробки 2 мяча и затем их возвращают в коробку. Какова вероятность для второй игры из этой коробки наудачу вынуть два новых мяча?

Решение. Введем событие $A = \{$вынуть два новых мяча для второй игры$\}$. Ситуация перед второй игрой описывается следующими взаимоисключающими возможностями:

$H_3 = \{$в коробке три новых мяча$\}$, если в первый раз играли двумя старыми мячами, $H_2 = \{$в коробке 2 новых мяча$\}$, если играли одним старым и одним новым, $H_1 = \{$в коробке 1 новый мяч$\}$, если играли двумя новыми мячами. События H_1, H_2, H_3 составляют полную группу событий, так как они несовместны и в сумме составляют все возможные исходы. Используя комбинаторные формулы, получим $\mathbf{P}(H_3) = C_3^2/C_6^2 = 1/5, \mathbf{P}(H_2) = 3 \cdot 3/C_6^2 = 3/5, \mathbf{P}(H_1) = C_3^2/C_6^2 = 1/5$. Далее имеем $\mathbf{P}(A|H_3) = C_3^2/C_6^2 = 1/5, \mathbf{P}(A|H_2) = 1/C_6^2 = 1/15, \mathbf{P}(A|H_1) = 0$. Теперь, используя формулу полной вероятности, найдем

$$\mathbf{P}(A) = \frac{1}{5} \cdot 0 + \frac{3}{5} \cdot \frac{1}{15} + \frac{1}{5} \cdot \frac{1}{5} = \frac{2}{25}.$$

В рамках ситуации, которая была описана в начале параграфа, можно ответить и на следующий вопрос. Пусть известно, что произошло некоторое событие. Какова вероятность того, что при этом реализовалось то или иное событие из выделенной полной группы событий? Ответ на этот вопрос дает следующий результат.

Теорема 2 (формула Байеса). *Пусть даны полная группа событий $H_1, H_2, \ldots, H_n$ и некоторое событие A. Тогда для любого $k = 1, 2, \ldots, n$ условная вероятность события H_k при условии, что событие A произошло, задается формулой*

$$\mathbf{P}(H_k|A) = \frac{\mathbf{P}(H_k)\mathbf{P}(A|H_k)}{\sum_{l=1}^{n} \mathbf{P}(H_l)\mathbf{P}(A|H_l)}.$$

Доказательство. В силу определения условной вероятности $\mathbf{P}(H_k|A) = \mathbf{P}(H_k A)/\mathbf{P}(A)$. Согласно формуле умножения вероятностей (§ 6) имеем $\mathbf{P}(H_k A) = \mathbf{P}(H_k)\mathbf{P}(A|H_k)$. Следовательно,

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

Подставляя сюда выражение для $\mathbf{P}(A)$ из формулы полной вероятности, получим формулу Байеса.

Задача 2. Два охотника одновременно и независимо стреляют в кабана. Известно, что первый попадает с вероятностью 0.8, а второй – 0.4. Кабан убит, и в нем обнаружена одна пуля. Как делить кабана?

Решение. Наиболее естественным является следующий принцип деления. Делить следует пропорционально условным вероятностям попадания каждого при условии, что в кабане имеется одна пуля. Дело в том, что при неограниченном (хотя бы и мысленном) повторении описанной ситуации частота, с которой будет добиваться успеха каждый из стрелков, будет сближаться с соответствующей условной вероятностью. Пусть $A = \{$в кабане имеется одна пуля$\}$. Какую выбрать полную группу событий? Выберем такие события: $H_{00} = \{$не попал ни первый, ни второй$\}, H_{10} = \{$попал первый, не попал второй$\}, H_{01} = \{$не попал первый, попал второй$\}, H_{11} = \{$попал первый, попал второй$\}$. Нам нужно найти вероятности $\mathbf{P}(H_{10}|A)$ и $\mathbf{P}(H_{01}|A)$. Очевидно, что $\mathbf{P}(A|H_{00}) = 0, \mathbf{P}(A|H_{10}) = 1, \mathbf{P}(A|H_{01}) = 1, \mathbf{P}(A|H_{11}) = 0$. Поскольку охотники стреляют независимо друг от друга, то по формуле умножения вероятностей

$$\mathbf{P}(H_{10}) = \mathbf{P}(\text{попал первый})\mathbf{P}(\text{не попал второй}) = 0.8 \cdot 0.6 = 0.48,$$

$$\mathbf{P}(H_{01}) = \mathbf{P}(\text{не попал первый})\mathbf{P}(\text{попал второй}) = 0.2 \cdot 0.4 = 0.08.$$

Применяя формулу Байеса, получим

$$\mathbf{P}(H_{10}|A) = \frac{0.48}{0.48 + 0.08} = \frac{6}{7}, \quad \mathbf{P}(H_{01}|A) = \frac{0.08}{0.48 + 0.08} = \frac{1}{7}.$$

Таким образом, первому следует отдать 6 долей из 7, а второму – одну.

Задачи

Задача 7.1. В первом ящике 3 белых и 8 черных шаров, во втором – 6 белых и 5 черных. Из первого во второй наудачу переложили один шар. Какова теперь вероятность вынуть из первого ящика черный шар?

Задача 7.2. В ящике 3 белых и 7 черных шаров. Один шар вынули наудачу и отложили в сторону. Следующий наугад вынутый шар оказался белым. Какова вероятность того, что отложенный шар был белым?

Задача 7.3. Есть четыре кубика с цифрами $1, 2, \ldots, 6$ на гранях и одна правильная пирамидка с цифрами 1,2,3,4 на гранях. Наугад выбрали предмет и бросили. Выпала цифра 4. Какова вероятность того, что взяли кубик?

Задача 7.4. Есть 10 симметричных монет, 8 нормальных, а на двух герб находится с обеих сторон. Наудачу взятая монета бросается три раза. Найти вероятность того, что выпадут три герба.

Задача 7.5. Из двенадцати лотерейных билетов четыре выигрышных. Какова вероятность вытянуть выигрышный билет, если перед этим наудачу вытянули два билета?

Задача 7.6. Из двенадцати лотерейных билетов пять выигрышных. Билеты вытягивают по одному без возвращения. Какова вероятность того, что во второй раз вытянут выигрышный билет?

Задача 7.7. Из двенадцати лотерейных билетов пять выигрышных. Билеты вытягивают по одному без возвращения. Во второй раз был вытянут выигрышный билет. Какова вероятность того, что и в первый раз был вытянут выигрышный билет?

Задача 7.8. Число грузовых машин, проезжающих мимо бензоколонки, относится к числу легковых машин как 3:2. Вероятность того, что грузовая машина будет заправляться, равна 0.1, а того, что будет заправляться легковая – 0.2. У бензоколонки заправляется машина. Найти вероятность того, что это грузовая машина.

Задача 7.9. На рис. 7 изображена схема дорог.

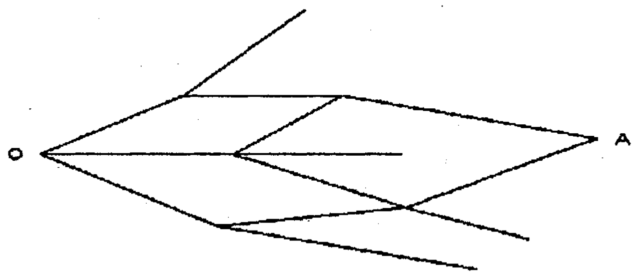

Рис. 7

Туристы вышли из пункта O, выбирая наугад на разветвлении дорог один из возможных путей (изменять направление более чем на 90° нельзя). Какова вероятность того, что они попадут в пункт A?

Задача 7.10. Три охотника одновременно и независимо стреляют в кабана. Известно, что первый попадает с вероятностью 0.8, второй – 0.4, а третий – 0.2. Кабан убит, и в нем обнаружены две пули. Как делить кабана?

§ 8. Последовательные испытания (схема Бернулли)

Схема Бернулли описывает эксперименты со случайным исходом, заключающиеся в следующем. Проводятся n последовательных независимых одинаковых экспериментов (испытаний), в каждом из которых выделяется одно и то же событие A, которое может наступить или не наступить в ходе эксперимента. Под **независимыми** понимаются такие эксперименты, в которых любые события, возникающие в разных экспериментах, являются независимыми в совокупности. Так как испытания одинаковы, то в любом из них событие A наступает с одинаковой вероятностью, обозначим ее $p = \mathbf{P}(A)$. Вероятность дополнительного события обозначим q. Тогда $q = \mathbf{P}(\overline{A}) = 1-p$. Наступление события A обычно называют успехом, а ненаступление – неудачей.

Пример 1. Симметричная однородная монета бросается 10 раз. В каждом из бросаний нас интересует событие $A = \{$выпал герб$\}$. Здесь 10 испытаний, и так как монета симметрична и однородна, то $p = \mathbf{P}(A) = 1/2$, $q = \mathbf{P}(\overline{A}) = 1/2$.

Пример 2. Игральная кость бросается три раза. При каждом бросании нас интересует событие $A = \{$выпала шестерка$\}$. Здесь три испытания, и поскольку кубик симметричен и однороден, то $p = \mathbf{P}(A) = 1/6$, $q = \mathbf{P}(\overline{A}) = 5/6$.

Теорема 1 (формула Бернулли). *Обозначим* $P_n(m) = \mathbf{P}($*событие* A *наступило* m *раз в* n *испытаниях*$)$. *Тогда*

$$P_n(m) = C_n^m p^m q^{n-m}. \tag{8.1}$$

Доказательство. Пронумеруем испытания числами от 1 до n. Обозначим $A_l = \{$наступление A в l-м испытании$\}$. Ясно, что $\mathbf{P}(A_l) = p, \mathbf{P}(\overline{A_l}) = 1 - p = q$, и события $A_l, l = 1, 2, \ldots, n$, независимы в совокупности. Для того, чтобы прояснить суть доказательства, рассмотрим сначала конкретный случай, когда $n = 3, m = 2$. Можно написать следующее представление для интересующего нас события: $\{A$ наступило 2 раза в трех испытаниях $\} = A_1A_2\overline{A_3} + A_1\overline{A_2}A_3 + \overline{A_1}A_2A_3$. Число слагаемых здесь равно $C_3^2 = 3$, что равно числу возможностей выбрать два числа среди трех чисел 1,2,3 без учета порядка. Это соответствует выбору пар номеров испытаний $\{1,2\},\{1,3\},\{2,3\}$, в которых наступает событие A. Оставшийся номер отводится событию $\overline{A}$. Слагаемые $A_1A_2\overline{A_3}$, $A_1\overline{A_2}A_3$, $\overline{A_1}A_2A_3$ несовместны, так как не может, например, наступать и одновременно не наступать событие A в третьем испытании. Каждое слагаемое является произведением независимых событий. Поэтому, применяя формулу сложения для несовместных событий и формулу умножения для независимых событий, получаем

$$\begin{aligned}
&\mathbf{P}(A) = \mathbf{P}(A_1A_2\overline{A_3}) + \mathbf{P}(A_1\overline{A_2}A_3) + \mathbf{P}(\overline{A_1}A_2A_3) \\
&= \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(\overline{A_3}) + \mathbf{P}(A_1)\mathbf{P}(\overline{A_2})\mathbf{P}(A_3) + \mathbf{P}(\overline{A_1})\mathbf{P}(A_2)\mathbf{P}(A_3) \\
&= ppq + pqp + qpp = 3p^2q = C_3^2p^2q.
\end{aligned}$$

Тем самым мы установили формулу Бернулли для $n = 3, m = 2$. Формула для общего случая доказывается аналогично. Существует ровно C_n^m различных возможностей выбрать m номеров испытаний, в которых наступает событие A. Таким образом, имеется C_n^m несовместных событий типа $A_1A_2\overline{A_3}\ldots A_k\overline{A_{k+1}}\ldots A_n$, среди множителей которых m событий вида A_l и $n-m$ событий вида $\overline{A_j}$. В сумме эти C_n^m несовместных событий составляют событие $\{A$ наступило m раз в n испытаниях$\}$. Каждое из слагаемых имеет вероятность

$$\begin{aligned}
&\mathbf{P}(A_1A_2\overline{A}_3\ldots A_k\overline{A}_{k+1}\ldots A_n) \\
&= \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(\overline{A}_3)\ldots\mathbf{P}(A_k)\mathbf{P}(\overline{A}_{k+1})\ldots\mathbf{P}(A_n) = p^mq^{n-m}.
\end{aligned}$$

Поскольку эти вероятности одинаковы, то суммарная вероятность имеет требуемый вид $C_n^mp^mq^{n-m}$.

Теорема 2. *Пусть m_1, m_2 – целые числа, $0 \leqslant m_1 \leqslant m_2 \leqslant n$. Обозначим через $P_n(m_1, m_2)$ вероятность того, что событие A наступило не менее m_1 и не более m_2 раз в n испытаниях. Тогда*

$$P_n(m_1, m_2) = \sum_{k=m_1}^{m_2} C_n^kp^kq^{n-k}, \tag{8.2}$$

или

$$P_n(m_1, m_2) = 1 - \sum_{k=0}^{m_1-1} C_n^k p^k q^{n-k} - \sum_{k=m_2+1}^{n} C_n^k p^k q^{n-k}. \quad (8.3)$$

Замечание 1. Следует использовать ту формулу, в которой число слагаемых меньше.

Доказательство теоремы 2. Для краткости фразу "в n испытаниях" будем опускать. Справедливо представление

$$\{A \text{ наступило не менее } m_1 \text{ и не более } m_2 \text{ раз}\} =$$
$$= \sum_{k=m_1}^{m_2} \{A \text{ наступило ровно } k \text{ раз}\}.$$

Все слагаемые являются несовместными событиями, так как A не может наступать различное число раз в n испытаниях одновременно. Используя формулу сложения вероятностей и формулу Бернулли, получим

$$P_n(m_1, m_2) = \mathbf{P}(A \text{ наступило не менее } m_1 \text{ и не более } m_2 \text{ раз})$$
$$= \sum_{k=m_1}^{m_2} \mathbf{P}(A \text{ наступило ровно } k \text{ раз}) = \sum_{k=m_1}^{m_2} C_n^k p^k q^{n-k}.$$

Второе выражение для $P_n(m_1, m_2)$ справедливо в силу формулы бинома Ньютона:

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

Определение 1. Число наступлений события A называется **наивероятнейшим**, если оно имеет наибольшую вероятность по сравнению с вероятностями наступления A любое другое количество раз.

Теорема 3. *Наивероятнейшее число наступлений события A в n испытаниях заключено между числами $np - q$ и $np + p$.*

Замечание 2. Если $np - q$ – целое число, то наивероятнейших чисел два: $np - q$ и $np + p$.

Доказательство. По формуле Бернулли при $m = 1, 2, \ldots, n$

$$\frac{P_n(m)}{P_n(m-1)} = \frac{C_n^m p^m q^{n-m}}{C_n^{m-1} p^{m-1} q^{n-m+1}} = \frac{n!(m-1)!(n-m+1)!p}{m!(n-m)!n!q} = \frac{(n+1-m)p}{m(1-p)}.$$

Следовательно, вероятность $P_n(m)$ будет больше, равна или меньше вероятности $P_n(m-1)$ в зависимости от того, какое из следующих трех соотношений будет выполняться

$$\frac{(n+1-m)p}{m(1-p)} > 1, \quad \frac{(n+1-m)p}{m(1-p)} = 1, \quad \frac{(n+1-m)p}{m(1-p)} < 1.$$

Если переписать эти соотношения в более простом виде

$$(n+1)p > m, \quad (n+1)p = m, \quad (n+1)p < m,$$

то мы приходим к выводу, что

$$P_n(m) > P_n(m-1), \quad \text{если } m < (n+1)p, \tag{8.4}$$

$$P_n(m) = P_n(m-1), \quad \text{если } m = (n+1)p, \tag{8.5}$$

$$P_n(m) < P_n(m-1), \quad \text{если } m > (n+1)p. \tag{8.6}$$

Следовательно, вероятность $P_n(m)$ сначала возрастает, когда $m < (n+1)p$, а затем убывает, когда $m > (n+1)p$. В случае, когда $(n+1)p$ не является целым числом, для наивероятнейшего числа наступлений события A (обозначим его m_0) должно выполняться неравенство $P_n(m_0+1) < P_n(m_0)$, что согласно (8.6) возможно при $m_0 + 1 > (n+1)p$, т. е. при $m_0 > np - q$, а также должно выполняться неравенство $P_n(m_0 - 1) < P_n(m_0)$, что в силу (8.4) возможно при $m_0 < np + p$. Таким образом, $np - q < m_0 < np + p$. Это и утверждается теоремой. Заметим, что разность между $np+p$ и $np-q$ равна единице, и значит число m_0 единственно. В случае, когда $np+p$ является целым числом, то $m_0 = np + p$ будет наивероятнейшим числом наступлений события A, однако $m_0 - 1$ тоже будет таковым, поскольку в силу (8.5) $P_n(m_0) = P_n(m_0-1)$. Поэтому таких чисел будет два, а именно $m_0^{(1)} = np - q$, $m_0^{(2)} = np + p$.

Задача 1. Вероятность попадания в цель при каждом выстреле из лука равна 1/3. Производится шесть выстрелов. Какова вероятность ровно двух попаданий? Какова вероятность не менее двух попаданий? Каково наивероятнейшее число попаданий?

Решение. Обозначим $A = \{$попадание при одном выстреле$\}$, $p = \mathbf{P}(A) = 1/3$, $q = 1 - p = 2/3$. Число выстрелов $n = 6$. Естественно предположить, что выстрелы не зависят друг от друга. Тогда ответ на первый вопрос находим по формуле Бернулли

$$P_6(2) = C_6^2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4 = \frac{15\cdot 16}{3^6} = \frac{80}{243} \approx \frac{1}{3}.$$

Ответ на второй вопрос следующий:

$$P_6(2,6) = 1 - C_6^0\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^6 - C_6^1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^5 = 1 - \frac{2^6}{3^6} - \frac{6\cdot 2^5}{3^6} = \frac{473}{729} \approx \frac{2}{3}.$$

Наивероятнейшее число попаданий лежит в пределах от $6\cdot\frac{1}{3} - \frac{2}{3}$ до $6\cdot\frac{1}{3} + \frac{1}{3}$, т. е. от $1\frac{1}{3}$ до $2\frac{1}{3}$. Следовательно, оно равно двум.

Задача 2 (Банаха). Некий курящий математик носит с собой две коробки спичек. Каждый раз, когда он хочет достать спичку, он выбирает наугад одну из коробок. Найти вероятность того, что когда математик вынет в первый раз пустую коробку, в другой коробке окажется ровно r спичек, $0 < r \leqslant n$, где n – число спичек, бывших первоначально в каждой из коробок.

Решение. Пусть A – событие, которое состоит в том, что вынимается спичка из коробки, которая в конце оказалась пустой. Если вынутая коробка пуста, а другая коробка содержит r спичек, то это означает, что спички брались всего $2n - r$ раз. При этом событие A наступило ровно n раз, так как коробка стала пустой. Поскольку каждый раз коробка выбирается наугад, то $\mathbf{P}(A) = 1/2$. По формуле Бернулли событие A наступает n раз в $2n - r$ испытаниях с вероятностью

$$P_{2n-r}(n) = C_{2n-r}^n\left(\frac{1}{2}\right)^{2n-r} = \frac{(2n-r)!}{n!(n-r)!2^{2n-r}}.$$

Задачи

Задача 8.1. В семье 10 детей. Считая вероятности рождения мальчика и девочки равными 1/2, определить вероятность того, что в данной семье: а) пять мальчиков; б) мальчиков не менее трех, но и не более восьми.

Задача 8.2. Что вероятнее выиграть у равносильного шахматиста (ничейный исход партии исключен): больше одной партии из четырех или больше двух партий из пяти?

Задача 8.3. В кошельке лежат 8 монет достоинством 5 копеек и 2 монеты достоинством 3 копейки. Наудачу выбирается монета и бросается 5 раз. Какова вероятность того, что в сумме будет 15 очков, если герб принимается за 0.

Задача 8.4. Какова вероятность выпадения хотя бы двух шестерок при трех бросаниях игральной кости?

Задача 8.5. В помещении четыре лампы. Вероятность работы в течение года для каждой лампы 0.8. Найти вероятность того, что к концу года горят три лампы. Чему равно наивероятнейшее число ламп, которые будут работать в течение года?

Задача 8.6. Пирамидка (цифры 1,2,3,4) подбрасывается 4 раза. Какова вероятность, что не выпадет ни одной четверки?

Задача 8.7. Вероятность попадания в мишень при одном выстреле равна $\frac{1}{4}$. Какова вероятность не менее трех попаданий при пяти выстрелах?

Задача 8.8. Какова вероятность того, что при четырех бросаниях пирамидки единица выпадет не более трех раз?

Задача 8.9. Два баскетболиста делают по три броска в корзину. Вероятность попадания мяча при каждом броске равна соответственно 0.6 и 0.7. Найти вероятность того, что у обоих будет равное количество попаданий.

Задача 8.10. Партия изделий содержит 1% брака. Каков должен быть объем контрольной выборки, чтобы вероятность обнаружить в ней хотя бы одно бракованное изделие была не меньше 0.95.

Задача 8.11. Вероятность того, что при одном выстреле стрелок попадает в цель, равна 0.4. Сколько выстрелов должен произвести стрелок, чтобы с вероятностью не менее 0.9

он попал в цель хотя бы один раз?

Задача 8.12. (Проблема Джона Смита) В 1693 г. Джоном Смитом был поставлен следующий вопрос: одинаковы ли шансы на успех у трех человек, если первому надо получить хотя бы одну шестерку при бросании игральной кости 6 раз, второму – не менее двух шестерок при 12 бросаниях, а третьему – не менее трех шестерок при 18 бросаниях.

§ 9. Предельные теоремы для схемы Бернулли

Сохраним все обозначения предыдущего параграфа. В случае, когда число испытаний велико, формулу Бернулли применять неудобно. Для больших n существуют приближенные формулы. Эти формулы тем точнее, чем n больше.

Рассмотрим сначала случай, когда с ростом n вероятность p уменьшается обратно пропорционально n. При малых p, речь идет о появлении очень редких событий, так как вероятность их наступления в отдельном испытании мала. Однако вероятность появления одного или нескольких редких событий в длинной серии испытаний уже не будет малой величиной.

В этом параграфе и далее соотношение $a_n \sim b_n$, где $\{a_n\}$ и $\{b_n\}$ – две числовые последовательности, означает, что $a_n/b_n \to 1$ при $n \to \infty$, или, что то же самое, $a_n = b_n(1+$ $+o(1))$, где $o(1)$ обозначает некоторую величину, стремящуюся к нулю. Соотношение $a_n \approx b_n$, означает, что $a_n - b_n \to 0$ при $n \to \infty$

Теорема 1 (Пуассона). *Предположим, что произведение np является постоянной величиной, когда n неограниченно возрастает. Обозначим $\lambda = np$. Тогда для любого фиксированного m и любого постоянного λ*

$$\lim_{\substack{n\to\infty \\ np=\lambda}} P_n(m) = \frac{\lambda^m}{m!}e^{-\lambda}.$$

Доказательство. Применяя формулу Бернулли, получим

$$P_n(m) = C_n^m p^m (1-p)^{n-m} = \frac{n!}{m!(n-m)!}\left(\frac{\lambda}{n}\right)^m \left(1-\frac{\lambda}{n}\right)^{n-m} =$$
$$= \frac{1}{m!} n(n-1)\dots(n-m+1)\frac{\lambda^m}{n^m}\left(1-\frac{\lambda}{n}\right)^{n-m} =$$
$$= \frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-m}.$$

Поскольку при любом фиксированном s имеет место сходимость $1-\frac{s}{n} \to 1$ при $n \to \infty$, то

$$P_n(m) \sim \frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^n.$$

Теперь, используя **"замечательный предел"**

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda},$$

несложно убедиться в справедливости утверждения теоремы.

Задача 1. Радиоаппаратура состоит из 2000 элементов. Вероятность отказа одного элемента в течение года равна 0.001. Какова вероятность отказа двух элементов за год? Какова вероятность отказа не менее двух элементов за год?

Решение. Работу каждого элемента рассматриваем как отдельное испытание. Обозначим $A = \{$отказ элемента за год$\}$. Имеем $p = \mathbf{P}(A) = 0.001, \lambda = np = 2000 \cdot 0.001 = 2$. По формуле Пуассона

$$P_{2000}(2) \approx \frac{2^2}{2!}e^{-2} = 2e^{-2} \approx 0.2707.$$

Ответ на второй вопрос дается формулой

$$P_{2000}(2, 2000) = 1 - P_{2000}(0) - P_{2000}(1) \approx$$
$$\approx 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} = 1 - 3e^{-2} \approx 0.594.$$

Рассмотрим еще одну приближенную формулу для вероятности $P_n(m)$, когда n велико. В отличие от предыдущего результата число успехов m в этом случае тоже растет с ростом n, а вероятность успеха постоянна.

Теорема 2 (Муавра–Лапласа (локальная)). *Положим* $x_n = \dfrac{m - np}{\sqrt{npq}}$. *Предположим, что* $m \to \infty, n \to \infty$ *и величины* x_n *являются ограниченными. Тогда*

$$\sqrt{npq}P_n(m) \sim \frac{1}{\sqrt{2\pi}}e^{-x_n^2/2}.$$

В частности, если $x_n \to x$, *то*

$$\sqrt{npq}P_n(m) \to \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

Доказательство. Поскольку $n-m = nq - x_n\sqrt{npq}$ (см. например, далее (9.2)), то в силу ограниченности величин x_n разность $n - m$ стремится к ∞ вместе с n и m. Воспользуемся формулой Стирлинга из § 1

$$k! \sim \sqrt{2\pi k}\left(\frac{k}{e}\right)^k$$

для $k = n, k = m$ и $k = n - m$. Тогда получим

$$\sqrt{npq}P_n(m) = \sqrt{npq}C_n^m p^m q^{n-m} = \frac{\sqrt{npq}n!}{m!(n-m)!}p^m q^{n-m} \sim$$
$$\sim \frac{\sqrt{npq}\sqrt{n}n^n}{\sqrt{2\pi}\sqrt{m}m^m\sqrt{n-m}(n-m)^{n-m}}p^m q^{n-m} =$$
$$= \frac{1}{\sqrt{2\pi}}\left(\frac{np}{m}\right)^m\left(\frac{nq}{n-m}\right)^{n-m}\sqrt{\frac{np}{m}}\sqrt{\frac{nq}{n-m}}.$$

Поскольку в силу определения x_n имеем

$$m = np + x_n\sqrt{npq}, \tag{9.1}$$

$$n - m = n - np - x_n\sqrt{npq} = nq - x_n\sqrt{npq}, \tag{9.2}$$

то

$$\frac{m}{np} = 1 + \frac{x_n\sqrt{q}}{\sqrt{np}} \sim 1, \tag{9.3}$$

и

$$\frac{n-m}{nq} = 1 - \frac{x_n\sqrt{p}}{\sqrt{nq}} \sim 1. \tag{9.4}$$

Поэтому, при достаточно большом n,

$$\sqrt{npq}P_n(m) \sim \frac{1}{\sqrt{2\pi}}\left(1+\frac{x_n\sqrt{q}}{\sqrt{np}}\right)^{-m}\left(1-\frac{x_n\sqrt{p}}{\sqrt{nq}}\right)^{-(n-m)} =$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-m\ln\left(1+\frac{x_n\sqrt{q}}{\sqrt{np}}\right)-(n-m)\ln\left(1-\frac{x_n\sqrt{p}}{\sqrt{nq}}\right)\right).$$

Далее воспользуемся следующим асимптотическим выражением:

$$\ln(1+z) = z - \frac{z^2}{2}(1+o(1)) \qquad \text{при } z \to 0.$$

Тогда

$$\sqrt{npq}P_n(m) \sim \frac{1}{\sqrt{2\pi}}\exp\left(-m\left\{\frac{x_n\sqrt{q}}{\sqrt{np}}-\frac{x_n^2 q}{2np}(1+o(1))\right\}-\right.$$
$$\left.-(n-m)\left\{-\frac{x_n\sqrt{p}}{\sqrt{nq}}-\frac{x_n^2 p}{2nq}(1+o(1))\right\}\right). \qquad (9.5)$$

В силу (9.1), (9.2),

$$x_n\left(\frac{(n-m)\sqrt{p}}{\sqrt{nq}}-\frac{m\sqrt{q}}{\sqrt{np}}\right) = \frac{x_n}{\sqrt{npq}}\left(nqp - x_n\sqrt{npq}p - npq - x_n\sqrt{npq}q\right) =$$
$$= -x_n^2(p+q) = -x_n^2.$$

Применяя (9.3), (9.4), получим, что

$$\frac{x_n^2}{2}\left\{\frac{(n-m)p}{nq}(1+o(1))+\frac{mq}{np}(1+o(1))\right\} \sim$$
$$\sim \frac{x_n^2}{2}\left\{p(1+o(1))+q(1+o(1))\right\} \sim \frac{x_n^2}{2}(p+q) = \frac{x_n^2}{2}.$$

Подставляя полученные выражения в (9.5), имеем

$$\sqrt{npq}P_n(m) \sim \frac{1}{\sqrt{2\pi}}\exp\left(-x_n^2+\frac{x_n^2}{2}(1+o(1))\right) \sim \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x_n^2}{2}\right).$$

Теорема доказана.

Рассмотрим приближенную формулу для вероятности $P_n(m_1, m_2)$ того, что событие A наступило не менее m_1 и не более m_2 раз в n испытаниях, когда n велико. Предположим, что числа m_1 и m_2 растут с ростом n, а вероятность успеха постоянна.

Теорема 3 (Муавра–Лапласа (интегральная)). *Положим*

$$a_n = \frac{m_1 - np}{\sqrt{npq}}, \qquad b_n = \frac{m_2 - np}{\sqrt{npq}}.$$

Предположим, что $m_1 \to \infty, n \to \infty$, и величины a_n и b_n являются ограниченными. Тогда

$$\lim_{n\to\infty} \left| P_n(m_1, m_2) - \frac{1}{\sqrt{2\pi}} \int_{a_n}^{b_n} e^{-x^2/2} dx \right| = 0.$$

Доказательство. По локальной теореме Муавра–Лапласа

$$P_n(m) = \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-x_n^2(m)/2}(1 + o_m(1)),$$

где $x_n(m) = \dfrac{m - np}{\sqrt{npq}}$, $m_1 \leqslant m \leqslant m_2$ и $o_m(1)$ – величина, стремящаяся к нулю при $n \to \infty$. Величина $o_m(1)$ зависит от m. При доказательстве теоремы 2 можно было бы установить, что справедливо равномерное предельное соотношение:

$$\lim_{n\to\infty} \sup_{m_1 \leqslant m \leqslant m_2} |o_m(1)| = 0. \tag{9.6}$$

Это привело бы к значительному усложнению выкладок, поэтому мы принимаем это соотношение без доказательства. Согласно теореме 2, § 8

$$P_n(m_1, m_2) = \sum_{m=m_1}^{m_2} P_n(m) = \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-x_n^2(m)/2}(1 + o_m(1)).$$

Обозначим $\Delta x_n(m) = x_n(m+1) - x_n(m)$. Тогда $\Delta x_n(m) = \dfrac{1}{\sqrt{npq}}$, и вероятность $P_n(m_1, m_2)$ представима в виде

$$P_n(m_1, m_2) = \frac{1}{\sqrt{2\pi}} \sum_{m=m_1}^{m_2} e^{-x_n^2(m)/2} \Delta x_n(m)(1 + o_m(1)).$$

Отсюда следует, что

$$\left| P_n(m_1, m_2) - \frac{1}{\sqrt{2\pi}} \sum_{m=m_1}^{m_2} e^{-x_n^2(m)/2} \Delta x_n(m) \right| \leqslant$$

$$\leqslant \sup_{m_1 \leqslant m \leqslant m_1} |o_m(1)| \frac{1}{\sqrt{2\pi}} \sum_{m=m_1}^{m_2} e^{-x_n^2(m)/2} \Delta x_n(m).$$

По определению интеграла

$$\frac{1}{\sqrt{2\pi}} \sum_{m=m_1}^{m_2} e^{-x_n^2(m)/2} \Delta x_n(m) = \frac{1}{\sqrt{2\pi}} \int\limits_{a_n}^{b_n} e^{-x^2/2} dx + o(1).$$

Подставляя это выражение в предыдущее неравенство и учитывая соотношение (9.6), найдем

$$\left| P_n(m_1, m_2) - \frac{1}{\sqrt{2\pi}} \int\limits_{a_n}^{b_n} e^{-x^2/2} dx - o(1) \right| \leqslant o(1) \left(\frac{1}{\sqrt{2\pi}} \int\limits_{a_n}^{b_n} e^{-x^2/2} dx + o(1) \right).$$

Теорема доказана.

Положим

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad \Phi(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int\limits_0^x e^{-t^2/2} dt.$$

Отметим одно очевидное свойство функции $\Phi(x)$: она является нечетной. Кроме того, $\Phi(\infty) = 1$. Это свойство не является элементарным. Значения функций $\phi(x)$ и $\Phi(x)$ находятся из таблиц, при этом таблицы даны лишь для неотрицательных значений x. Функция $\phi(x)$ называется **функцией Гаусса,** а функция $\Phi(x)$ называется **функцией Лапласа**. Используя функции $\phi(x)$ и $\Phi(x)$, сформулированные выше результаты можно кратко выразить так: локальная теорема Муавра–Лапласа утверждает, что

$$P_n(m) \sim \frac{1}{\sqrt{npq}} \phi(x_n), \qquad \text{где} \quad x_n = \frac{m - np}{\sqrt{npq}}, \tag{9.7}$$

а интегральная теорема Муавра–Лапласа – что

$$P_n(m_1, m_2) \approx \frac{1}{2} \big(\Phi(b_n) - \Phi(a_n) \big), \tag{9.8}$$

где $a_n = \dfrac{m_1 - np}{\sqrt{npq}}$, $b_n = \dfrac{m_2 - np}{\sqrt{npq}}$. Требование о том, что величины a_n, b_n предполагаются ограниченными, не является существенным. Это будет следовать из теоремы 2 § 17 (см. пример 1).

Задача 2. Найти вероятность того, что при 150 выстрелах мишень будет поражена ровно 70 раз, если вероятность попадания при одном выстреле равна 0.4.

Решение. Данная задача решается с использованием схемы Бернулли: $n = 150$, $A = \{$попадание при одном выстреле$\}$, $p = \mathbf{P}(A) = 0.4$, $q = 1 - p = 0.6$, $m = 70$. Применим локальную теорему Муавра–Лапласа. Имеем $\sqrt{npq} = \sqrt{150 \cdot 0.4 \cdot 0.6} = 6$, $\quad x_n = \dfrac{70 - 150 \cdot 0.4}{6} = \dfrac{10}{6} \approx 1.67$. Тогда

$$P_{150}(70) \sim \frac{1}{6}\phi(1.67) \approx \frac{1}{6} \cdot 0.0989 \approx 0.0165.$$

Задача 3. Небольшой город ежедневно посещают 100 туристов, которые днем идут обедать. Каждый из них выбирает для обеда один из двух городских ресторанов с равными вероятностями и независимо друг от друга. Владелец одного из ресторанов желает, чтобы с вероятностью приблизительно 0.99 все пришедшие в его ресторан туристы могли там одновременно пообедать. Сколько мест должно для этого быть в его ресторане?

Решение. Пусть $A = \{$турист пообедал у заинтересованного владельца$\}$. Наступление события A будем считать "успехом", $p = \mathbf{P}(A) = 0.5$, $n = 100$. Нас интересует такое наименьшее число m, что вероятность наступления не менее чем m "успехов"в последовательности из $n = 100$ независимых испытаний с вероятностью успеха $p = 0.5$ приблизительно равна $1 - 0.99 = 0.01$. Это как раз вероятность переполнения ресторана. Таким образом, нас интересует такое наименьшее число m, что $P_{100}(m, 100) \approx 0.01$. Применим интегральную теорему Муавра–Лапласа. Имеем $\sqrt{npq} = 5$, $b_n = \dfrac{100 - 50}{5} = 10$, $a_n = \dfrac{m - 50}{5} = \dfrac{m}{5} - 10$, m неизвестно. Тогда

$$0.01 \approx P_{100}(m, 100) \approx \frac{1}{2}\Big(\Phi(10) - \Phi\Big(\frac{m}{5} - 10\Big)\Big) \approx$$

$$\approx \frac{1}{2}\Big(1 - \Phi\Big(\frac{m}{5} - 10\Big)\Big).$$

Откуда следует, что

$$\Phi\Big(\frac{m}{5} - 10\Big) \approx 0.98.$$

Используя таблицы для функции $\Phi(x)$ находим, $\frac{m}{5} - 10 \approx 2.33$, и, значит, $m \approx 2.33 \cdot 5 + 50 \approx 61.65$. Следовательно, в ресторане должно быть 62 места.

Задачи

Задача 9.1. Посеяли 10000 семян. Вероятность для семени не прорасти – 0.0003. Какова вероятность того, что все семена прорастут?

Задача 9.2. Вероятность прерывания телефонного соединения равна 0.03. Какова вероятность того, что среди 200 соединений будет не более двух прерываний?

Задача 9.3. Газета содержит 20000 букв. Каждая буква может быть неправильно напечатана с вероятностью 0.0004. Какова вероятность, что в газете будет не менее двух опечаток?

Задача 9.4. Телефонная станция обслуживает 600 абонентов. Вероятность любого позвонить в течение часа равна 0.005. Какова вероятность того, что в течение часа позвонит один или два человека?

Задача 9.5. Игральную кость бросают 80 раз. Найти приближенно границы, в которых число выпаданий шестерки будет заключено с вероятностью 0.9973.

Задача 9.6. В партии из 768 арбузов каждый арбуз оказывается неспелым с вероятностью 1/4. Найти вероятность того, что количество спелых арбузов будет в пределах от 564 до 600.

Задача 9.7. Вероятность найти белый гриб среди прочих равна 1/4. Какова вероятность того, что среди 300 грибов будет 75 белых?

Задача 9.8. Завод отправил в магазин 5000 лампочек. Вероятность того, что лампочка разобьется при транспортировке равна 0.0002. Найти вероятность того, что в магазин привезли не более трех разбитых лампочек.

Задача 9.9. При социологических опросах граждан каждый человек независимо от других может дать неискренний ответ с вероятностью 0.2. Найти вероятность того, что из 22500 опросов число неискренних ответов будет не более 4620.

Задача 9.10. Правильную монету бросают 400 раз. Какова вероятность того, что герб выпадет не менее 180 раз?

Задача 9.11. Игральный кубик бросают 500 раз. Какова вероятность того, что шестерка выпадет не более 90 раз?

Задача 9.12. Среди 240000 городских ламп каждая будет гореть в течение года с вероятностью 0.6. Какова вероятность того, что к концу года будут гореть от 14160 до 14880 ламп?

Задача 9.13. Вероятность выхода из строя одного элемента в течение года равна 0.2. Какова вероятность того, что из 100 элементов в течение года выйдут из строя 15?

Задача 9.14. В поселке 1000 домов, каждый из которых застрахован на год от пожара в одной страховой компании на сумму 100000 рублей. Страховой взнос за год составляет 300 рублей. Для данного поселка вероятность пожара в доме в течение года оценивается величиной 0.002. Какова вероятность того, что в течение года страховая компания понесет убытки?

§ 10. Случайные величины и функции распределения

Рассмотрим сначала случайный эксперимент с дискретным пространством исходов, т. е. с таким пространством элементарных событий $\Omega = \{\omega_1, \omega_2, \dots\}$, которое состоит из конечного или счетного числа исходов. Пусть есть величина, которая в результате случайного эксперимента принимает различные числовые значения в зависимости от наступления того или иного исхода, при этом каждому исходу соответствует только одно число. Иными словами, на пространстве элементарных событий задана функция.

Определение 1. Функция, заданная на пространстве элементарных событий $\Omega = \{\omega_1, \omega_2, \dots\}$, называется **случайной величиной**.

Для любого исхода ω значение $x = X(\omega)$ – это **реализация случайной величины** X при данном исходе.

В ходе случайного эксперимента реализуется лишь один какой-то исход, это означает, что в результате эксперимента наблюдается лишь какое-то одно значение случайной величины (одна реализация) из всех возможных.

Обозначать случайные величины будем заглавными латинскими буквами $X, Y, \dots$, а конкретные их значения соответствующими строчными буквами $x, y, \dots$.

Пример 1. Пусть бросается игральная кость. Величина X, равная числу выпавших очков, является случайной величиной. Величина X принимает значения 1, 2, 3, 4, 5, 6.

В этом примере элементарный исход – это выпадение той или иной грани игрального кубика. В стандартной ситуации (игральная кость) на гранях кубика написаны цифры 1, 2, 3, 4, 5, 6. Однако можно этим граням сопоставить и другие числа. Тогда мы получим другую случайную величину, хотя пространство элементарных исходов и соответствующие им вероятности $1/6$ останутся прежними.

Для дискретных пространств $\Omega = \{\omega_1, \omega_2, \dots\}$ будем обозначать $x_l = X(\omega_l)$, $l = 1, 2, \dots$. Величина X принимает при этом не более чем счетное число значений $\{x_1, x_2, \dots\}$ и называется **дискретной**. Значение x_l случайной величины X наступает с некоторой вероятностью, обозначим ее p_l. Поскольку одному и тому же значению может соответствовать несколько исходов, то $p_l = \sum\limits_{k: X(\omega_k) = x_l} p(\omega_k)$. Кратко это будем записывать так: $p_l = \mathbf{P}(X = x_l)$. В силу аксиом 3) и 2) из определения вероятности (§ 6) следует, что

$$\sum_{l=1}^{\infty} p_l = \sum_{l=1}^{\infty} \mathbf{P}(X = x_l) = \mathbf{P}\Big(\sum_{l=1}^{\infty} \{X = x_l\}\Big) = 1.$$

Определение 2. Соответствие, которое каждому значению x_l дискретной случайной величины X сопоставляет его вероятность p_l, называется **законом распределения** случайной величины X.

Закон распределения величины X удобно задавать в виде таблицы:

X	x_1	x_2	$\dots$	x_n
$\mathbf{P}$	p_1	p_2	$\dots$	p_n

Пример 2. Пусть X – число гербов, выпавших при четырех бросаниях правильной монеты. Величина X может

принимать значения 0, 1, 2, 3, 4, поэтому в вышеприведенной таблице полагаем $x_k = k - 1$, $k = 1, 2, \ldots, 5$. По формуле Бернулли

$$p_k = \mathbf{P}(X = x_k) = C_4^{(k-1)}\left(\frac{1}{2}\right)^{(k-1)}\left(\frac{1}{2}\right)^{4-(k-1)} = \frac{C_4^{(k-1)}}{16},$$

и, следовательно, закон распределения имеет вид

X	0	1	2	3	4
$\mathbf{P}$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Замечание 1. Если X – дискретная случайная величина, то для любой функции g величина $Y = g(X)$ тоже является дискретной случайной величиной. Эта величина принимает значения $y_l = g(x_l)$ с вероятностями $p_l = \mathbf{P}(X = x_l)$, $l = 1, 2, \ldots, n$, если функция g взаимно однозначна. Если же значения $g(x_l)$ совпадают для различных x_l с величиной y_m, то $Y = g(X)$ принимает общее значение y_m с вероятностью, равной сумме вероятностей p_l, отвечающих всем таким x_l, для которых $g(x_l) = y_m$.

Задача 1. *Дискретная случайная величина* X имеет закон распределения

X	-2	-1	0	1	2
$\mathbf{P}$	0.1	0.2	0.3	0.3	0.1

Построить закон распределения случайной величины $Y = X^2 + 1$.

Решение. Значение 1 величина Y принимает только тогда, когда величина X принимает значение 0, с вероятностью 0.3. Значение 2 величина Y принимает, если величина X принимает значение -1 или 1 с вероятностями 0.2 и 0.3 соответственно. Тогда эти вероятности нужно сложить, что даст вероятность события $Y = 2$. Аналогично вероятность того, что $Y = 5$ будет равна $0.1 + 0.1 = 0.2$. Следовательно, закон распределения случайной величины $Y = X^2 + 1$ имеет вид

Y	1	2	5
$\mathbf{P}$	0.3	0.5	0.2

Рассмотрим случайные величины, которые могут уже не быть дискретными. При определении таких случайных величин множество Ω уже не является счетным, и мы должны, опираться на аксиомы теории вероятностей § 6. Будем считать, что из подмножеств пространства элементарных событий Ω выделено семейство множеств $\mathcal{A}$, которое является σ-алгеброй. Множества из $\mathcal{A}$ составляют всевозможные события случайного эксперимента. Предположим также, что на $\mathcal{A}$ задана вероятность $\mathbf{P}$, удовлетворяющая аксиомам теории вероятностей.

Определение 3. **Случайной величиной** называется функция $X = X(\omega)$, заданная на пространстве элементарных событий Ω, для которой событие $\{X < x\} = \{\omega : X(\omega) < x\}$ принадлежит σ-алгебре $\mathcal{A}$ для любого вещественного x.

Условие $\{X < x\} \in \mathcal{A}$ дает возможность рассматривать вероятности событий $\{X < x\}$, поскольку вероятности определены только на множествах из $\mathcal{A}$. Кроме того, через события $\{X < x\}$, $x \in (-\infty, \infty)$, с помощью известных операций над событиями можно выразить сколь угодно сложное событие, связанное со случайной величиной X. Такое событие будет также принадлежать σ-алгебре $\mathcal{A}$, и, следовательно, для него определена вероятность.

Вся совокупность вероятностей $\mathbf{P}(X < x)$, $x \in (-\infty, \infty)$ задает **закон распределения** случайной величины X в общем случае. Часто для краткости закон распределения называют просто **распределением** случайной величины X.

Определение 4. Функция $F(x) = \mathbf{P}(X < x)$, $x \in (-\infty, \infty)$, называется **функцией распределения** случайной величины X.

Замечание 2. Для дискретных случайных величин $F(x) = \mathbf{P}(X < x) = \mathbf{P}\left(\sum\limits_{x_l < x} \{X = x_l\}\right) = \sum\limits_{x_l < x} \mathbf{P}(X = x_l) = \sum\limits_{x_l < x} p_l$. Ясно, что функция $F(x)$ однозначно определяет все вероятности p_l.

Пример 3. Функция распределения величины X, равной числу гербов, выпавших при четырех бросаниях симметричной монеты, имеет вид

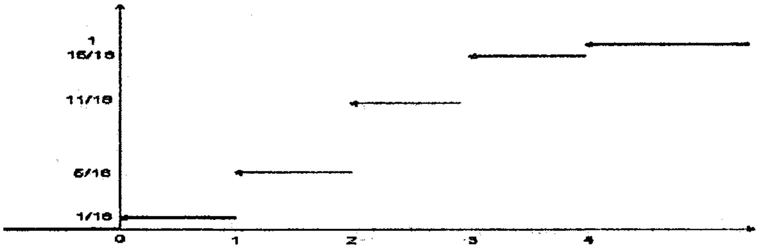

Рис. 8

Свойства функции распределения.

1) если $x_1 < x_2$, то $F(x_1) \leqslant F(x_2)$, т. е. $F(x)$ – неубывающая функция;

2) $F(-\infty) = 0, F(\infty) = 1$;

Выражение $F(\infty)$ служит краткой записью для предела $\lim\limits_{x\uparrow\infty} F(x)$, а $F(-\infty)$ – для предела $\lim\limits_{x\downarrow-\infty} F(x)$

3) функция $F(x)$ непрерывна слева: $\lim\limits_{x\uparrow y} F(x) = F(y)$.

Доказательство этих свойств целиком опирается на свойства вероятностей (§ 6). Положим $A = \{X < x_1\}, B = \{X < x_2\}$. При $x_1 < x_2$ выполняется включение $A \subset B$, т. к. если $X(\omega) < x_1$, то и $X(\omega) < x_2$. По третьему свойству вероятностей имеем $\mathbf{P}(A) \leqslant \mathbf{P}(B)$, а это по определению функции распределения и означает, что $F(x_1) \leqslant F(x_2)$.

Поскольку $\{X < -\infty\} = \varnothing$, а $\{X < \infty\} = \Omega$, т. е. эти события являются соответственно невозможным и достоверным, то по первому и второму свойству вероятностей $\mathbf{P}(X < -\infty) = 0$, а $\mathbf{P}(X < \infty) = 1$. Это составляет второе свойство функции распределения.

Выберем произвольную монотонно возрастающую последовательность x_n, стремящуюся к точке y. Тогда для $k = 2, 3, \ldots$ выполняется $\{X < x_{k-1}\} \subset \{X < x_k\}$, и $\{X < y\} = \sum\limits_{k=1}^{\infty} \{X < x_k\}$. По седьмому свойству вероятностей при $n \to \infty$

$$\mathbf{P}(X < x_n) \to \mathbf{P}\Big(\sum_{k=1}^{\infty} \{X < x_k\}\Big) = \mathbf{P}(X < y).$$

В силу определения функции распределения это соотношение можно переписать следующим образом: $F(x_n) \to F(y)$ при $x_n \uparrow y$. Таким образом, справедливо третье свойство функций распределения.

Лемма 1. *Для любых $a < b$ выполняется равенство*

$$\mathbf{P}(a \leqslant X < b) = F(b) - F(a). \tag{10.1}$$

Доказательство. Очевидно, что

$$\{X < b\} = \{X < a\} + \{a \leqslant X < b\},$$

при этом события $\{X < a\}$ и $\{a \leqslant X < b\}$ несовместны. Используя формулу сложения вероятностей, получим

$$\mathbf{P}(X < b) = \mathbf{P}(X < a) + \mathbf{P}(a \leqslant X < b),$$

и, следовательно,

$$\mathbf{P}(a \leqslant X < b) = \mathbf{P}(X < b) - \mathbf{P}(X < a) = F(b) - F(a).$$

Определение 5. Если существует такая неотрицательная функция $f(y)$, что функция распределения $F(x)$ для каждого $x \in (-\infty, \infty)$ представима в виде

$$F(x) = \int\limits_{-\infty}^{x} f(y)dy, \tag{10.2}$$

то $f(x)$ называется **плотностью распределения** случайной величины X.

Если случайная величина X имеет плотность распределения, то ее функция распределения непрерывна, поскольку интеграл – непрерывная функция верхнего предела.

В случае существования плотности формулу (10.1) можно записать следующим образом:

$$\mathbf{P}(a \leqslant X < b) = \int\limits_{-\infty}^{b} f(y)dy - \int\limits_{-\infty}^{a} f(y)dy = \int\limits_{a}^{b} f(y)dy. \tag{10.3}$$

Для непрерывной плотности распределения из формулы (10.3) и теоремы о среднем для интегралов следует, что при $\Delta \to 0$ справедливо соотношение

$$\mathbf{P}(X \in [x, x+\Delta)) \sim f(x)\Delta. \tag{10.4}$$

Это соотношение раскрывает суть плотности распределения: вероятность попасть в произвольно малый интервал, содержащий данную точку, близка с точностью до белее высокого порядка малости, к произведению плотности распределения в точке на длину интервала.

Свойства плотности распределения.

1) $\int\limits_{-\infty}^{\infty} f(y)dy = 1$;

2) $f(x) = F'(x)$, если функция распределения дифференцируема.

Для проверки первого свойства достаточно в (10.3) положить $a = -\infty, b = \infty$ и заметить, что тогда слева в (10.3) стоит вероятность достоверного события. Дифференцируя равенство (10.2), получим второе свойство.

Пример 4. Однородная проволока длиной 1 метр растягивается за концы и разрывается. Пусть X – случайная величина, равная расстоянию от точки разрыва до левого конца проволоки. Используя геометрические вероятности, найдем, что

$$\mathbf{P}(x_1 \leqslant X < x_2) = \frac{\text{mes}(x_1, x_2)}{\text{mes}(0, 1)} = x_2 - x_1$$

для любых $0 \leqslant x_1 < x_2 \leqslant 1$. Следовательно, функция распределения и плотность распределения этой случайной величины имеют вид

$$F(x) = \begin{cases} 0, \text{ при } x \leqslant 0, \\ x, \text{ при } 0 \leqslant x \leqslant 1, \\ 1, \text{ при } 1 \leqslant x, \end{cases} \qquad f(x) = \begin{cases} 0, \text{ при } x < 0, \\ 1, \text{ при } 0 \leqslant x < 1, \\ 0, \text{ при } 1 \leqslant x. \end{cases}$$

Пусть $y = g(x)$ – монотонно возрастающая функция, $x = g^{-1}(y)$ – обратная функция. Если X – случайная величина, то $Y = g(X)$ тоже является случайной величиной, поскольку $\{Y < y\} = \{g(X) < y\} = \{X < g^{-1}(y)\}$, и, следовательно, событие $\{Y < y\}$ принадлежит σ-алгебре $\mathcal{A}$ для любого вещественного y, поскольку $\{X < g^{-1}(y)\}$ принадлежит $\mathcal{A}$.

Для функций распределения величин X и Y справедливо соотношение

$$F_Y(y) = \mathbf{P}(Y < y) = \mathbf{P}(X < g^{-1}(y)) = F_X(g^{-1}(y)). \qquad (10.5)$$

Если случайная величина X имеет плотность распределения $f_X(x)$, а функция $g(x)$ дифференцируема, то случайная величина Y тоже имеет плотность распределения $f_Y(y)$ и

$$f_Y(y) = f_X(g^{-1}(y))(g^{-1}(y))'. \qquad (10.6)$$

Это равенство, в силу второго свойства плотности распределения, является следствием результата дифференцирования соотношения (10.5). При использовании равенства (10.6) бывает полезной формула

$$(g^{-1}(y))' = 1/g'(g^{-1}(y)).$$

В случае, когда $y = g(x)$ – монотонно убывающая дифференцируемая функция, можно показать, что

$$f_Y(y) = f_X(g^{-1}(y))/|g'(g^{-1}(y))|. \qquad (10.7)$$

Эта формула верна и для взаимно однозначных (существует обратная функция) кусочно монотонных функций $g(x)$. Тот факт, что возможно для счетного числа точек (концов интервалов монотонности) этой формулой значения $f_Y(y)$ не определяются, не является принципиальным. Плотности на выделенном счетном множестве можно придать любое значение, при этом функция распределения не изменится в силу свойств интеграла.

Существует широкий класс функций $g(x)$, не обязательно монотонных, для которых $Y = g(X)$ будет случайной величиной. К нему относятся, например, все непрерывные функции. Обсуждение этой проблемы выходит за рамки элементарного курса, так как требует дополнительных знаний из теории функций.

Задача 2. Случайная величина X имеет **распределение Коши** с плотностью распределения

$$f(x) = \frac{1}{\pi(1+x^2)}, \qquad \text{при } -\infty < x < \infty.$$

Вычислить плотность распределения обратной случайной величины $Y = 1/X$.

Решение. Функция $y = 1/x$ не определена в нуле, убывает на интервалах $(-\infty, 0)$, $(0, \infty)$ и имеет однозначную обратную функцию $x = 1/y$. Применяя формулу (10.7), получим

$$f_Y(y) = \frac{1}{\pi(1+y^{-2})}\frac{1}{y^2} = \frac{1}{\pi(1+y^2)}, \qquad \text{при } -\infty < y < \infty.$$

Следовательно, величина, обратная величине, распределенной по закону Коши, также имеет распределение Коши.

Задачи

Задача 10.1. В урне 5 белых и 25 черных шаров. Вынули 2 шара. Случайная величина X – число вынутых белых шаров. Построить закон распределения и функцию распределения величины X.

Задача 10.2. Построить закон распределения и функцию распределения числа попаданий мячом в корзину при трех бросках, если вероятность попадания каждый раз равна 0.8.

Задача 10.3. Случайная величина X задается плотностью распределения

$$f(x) = \begin{cases} 0, & \text{при } x < 3, \\ \dfrac{7-x}{8}, & \text{при } 3 \leqslant x \leqslant 7, \\ 0, & \text{при } 7 < x. \end{cases}$$

Вычислить функцию распределения $F(x)$ и вероятность $\mathbf{P}(2 \leqslant X \leqslant 4)$.

Задача 10.4. Случайная величина X задается плотностью распределения

$$f(x) = \begin{cases} 0 & \text{при } x < 2, \\ 2x-4 & \text{при } 2 \leqslant x \leqslant 3, \\ 0 & \text{при } 3 < x. \end{cases}$$

Вычислить функцию распределения $F(x)$ и вероятность $\mathbf{P}(2.5 < X < 3.5)$.

Задача 10.5. Случайная величина X имеет плотность распределения

$$f(x) = \begin{cases} 0, & \text{при } \pi < x, \\ \frac{1}{2}\sin x, & \text{при } 0 < x \leqslant \pi, \\ 0, & \text{при } x \leqslant 0. \end{cases}$$

а) Определить функцию распределения $F(x)$; б) найти вероятность того, что величина X примет значение, заключенное в интервале $(0, \pi/4)$.

Задача 10.6. Случайная величина X имеет плотность распределения

$$f(x) = \begin{cases} 0, & \text{при } |x| \geqslant 3, \\ \frac{1}{\pi\sqrt{9 - x^2}}, & \text{при } |x| < 3. \end{cases}$$

а) Определить функцию распределения $F(x)$; б) найти вероятность того, что величина X примет значение, заключенное в интервале $(1.5, 4)$.

Задача 10.7. Дискретная случайная величина X имеет закон распределения

X	0	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{5\pi}{6}$	π
$\mathbf{P}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$

Построить закон распределения случайной величины $Y = \sin X$.

Задача 10.8. Случайная величина X имеет плотность распределения $f_X(x) = e^{-x}, x \geqslant 0$. Найти функцию распределения случайной величины $Y = e^{-X}$.

Задача 10.9. Случайная величина X распределена с плотностью распределения $f_X(x)$. Найти плотность распределения случайной величины $Y = X^2$.

Задача 10.10. Построить закон распределения для величины X, равной числу выпадений очков кратных трем при четырех бросаниях игральной кости.

Задача 10.11. Случайная величина X имеет плотность распределения вида

$$f(x) = \begin{cases} ax^2, & \text{при } 0 \leqslant x \leqslant 2, \\ 0, & \text{при } x < 0, \quad 2 < x. \end{cases}$$

Вычислить константу a и определить вероятность того, что $X > 1$.

Задача 10.12. Случайная величина X имеет плотность распределения вида

$$f(x) = \begin{cases} a(x-1)^2, & \text{при } 1 \leqslant x \leqslant 5, \\ 0, & \text{при } x < 1, \ 5 < x. \end{cases}$$

Вычислить константу a и определить вероятность того, что $3 \leqslant X < 4$.

Задача 10.13. Вычислить плотность распределения величины $Y = \sqrt{X}$, где X – величина из задачи 10.11.

Задача 10.14. Вычислить плотность распределения величины $Y = (X+1)/2$, где X – величина из задачи 10.12.

§ 11. Совместные функции распределения нескольких случайных величин

Рассмотрим сначала две дискретные случайные величины X и Y. Пусть случайная величина X принимает значения $x_1, x_2, \ldots, x_n$, а случайная величина Y принимает значения $y_1, y_2, \ldots, y_m$. Одновременное наступление событий $\{X = x_i\}$ и $\{Y = y_j\}$ будем обозначать $\{X = x_i, Y = y_j\}$, т. е.

$$\{X = x_i, Y = y_j\} = \{X = x_i\}\{Y = y_j\}.$$

Обозначим $p_{ij} = \mathbf{P}(X = x_i, Y = y_j)$.

Определение 1. Соответствие, которое каждой паре значений (x_i, y_j) дискретных случайных величин X и Y сопоставляет ее вероятность p_{ij}, называется **совместным законом распределения** случайных величин X и Y.

Совместный закон распределения величин (X, Y) можно задавать таблицей

(X,Y)	(x_1,y_1)	$\ldots$	(x_1,y_m)	(x_2,y_1)	$\ldots$	(x_n,y_1)	$\ldots$	(x_n,y_m)
$\mathbf{P}$	$p_{1,1}$	$\ldots$	$p_{1,m}$	$p_{2,1}$	$\ldots$	$p_{n,1}$	$\ldots$	$p_{n,m}$

Задача 1. В ящике два шара, на каждом из которых написана цифра 1, и три шара, на каждом из которых написана цифра 2. Один за другим наудачу вынимают два шара. Пусть величина X – это номер на первом шаре, а Y – номер на втором. Найти совместный закон распределения величин (X, Y).

Решение. Используя формулу умножения вероятностей, найдем $\mathbf{P}((X,Y) = (1,1)) = \mathbf{P}(X = 1)\mathbf{P}(Y = 1|X = 1) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$. Поступая аналогично для вычисления других вероятностей, получим

(X,Y)	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
$\mathbf{P}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$

Свойства вероятностей p_{ij}.

1) $\sum\limits_{i=1}^{n} p_{ij} = p_j^{(2)}$, где $p_j^{(2)} = \mathbf{P}(Y = y_j)$;

2) $\sum\limits_{j=1}^{m} p_{ij} = p_i^{(1)}$, где $p_i^{(1)} = \mathbf{P}(X = x_i)$;

3) $\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} p_{ij} = 1$.

Докажем, например, второе свойство. Поскольку $\sum\limits_{j=1}^{m} \{Y = y_j\} = \Omega$, то используя свойства 10′) и 4) § 2, получим

$$\sum_{j=1}^{m}\{X = x_i, Y = y_j\} = \{X = x_i\}\sum_{j=1}^{m}\{Y = y_j\} =$$

$$= \{X = x_i\}\Omega = \{X = x_i\}.$$

Так как события $\{X = x_i, Y = y_j\}$ при разных j несовместны, то применяя формулу сложения вероятностей, получим

$$p_i^{(1)} = \mathbf{P}(X = x_i) = \mathbf{P}\Big(\sum_{j=1}^{m}\{X = x_i, Y = y_j\}\Big) =$$

$$= \sum_{j=1}^{m}\mathbf{P}\big(X = x_i, Y = y_j\big) = \sum_{j=1}^{m} p_{ij}.$$

Определение 2. Дискретные случайные величины X и Y называются **независимыми**, если для всех пар (i, j) выполняются соотношения

$$\mathbf{P}\big(X = x_i, Y = y_j\big) = \mathbf{P}\big(X = x_i\big)\mathbf{P}\big(Y = y_j\big),$$

т. е. события $\{X = x_i\}$ и $\{Y = y_i\}$ являются независимыми.

Замечание 1. Если X и Y две дискретные случайные величины, то для любой функции двух переменных $h(x, y)$ величина $Z = h(X, Y)$ тоже является дискретной случайной величиной, которая принимает значения $z_{i,j} = h(x_i, y_j)$ с вероятностями $p_{ij} = \mathbf{P}\big(X = x_i, Y = y_j\big), i = 1, 2, \ldots, n, j = 1, 2, \ldots, m$, если $h(x, y)$ является взаимно однозначной функцией. Если же значения $h(x_i, y_j)$ совпадают для различных пар (x_i, x_j) с величиной $z_{l,m}$, то $Z = h(X, Y)$ принимает общее значение $z_{l,m}$ с вероятностью, равной сумме вероятностей $p_{i,j}$, отвечающих всем таким (x_i, x_j), для которых $h(x_i, x_j) = z_{l,m}$.

Рассмотрим теперь общую ситуацию, когда имеется $k \geqslant 2$ произвольных случайных величин $X_1, X_2, \ldots, X_k$. Поскольку для описания этого случая отправной точкой служат аксиомы теории вероятностей (§ 6), то мы предполагаем, что задано Ω – пространство элементарных событий, из подмножеств которого образована σ-алгебра $\mathcal{A}$ и на ней задана вероятность $\mathbf{P}$. При этом предполагается, что при любых $l = 1, 2, \ldots, k$ и $x \in (-\infty, \infty)$ события $\{X_l < x\}$ принадлежат $\mathcal{A}$. Это условие позволяет рассматривать вероятности событий

$$\{X_1 < x_1, X_2 < x_2, \ldots, X_k < x_k\} = \prod_{l=1}^{k} \{X_l < x_l\}.$$

Определение 3. Функция

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

$x_k \in (-\infty, \infty)$ называется **совместной функцией распределения** величин $X_1, X_2, \ldots, X_k$.

Свойства совместной функции распределения.

1) функция $F(x_1, x_2, \ldots, x_k)$ является неубывающей функцией каждого аргумента, при условии, что другие фиксированы;

2) если $x_l = -\infty$ при некотором l, то $F(x_1, x_2, \ldots, x_k) = 0$, а если положить $x_l = \infty$, то $F(x_1, \ldots, x_{l-1}, \infty, x_{l+1}, \ldots, x_k)$ как функция $k-1$ переменной будет являться совместной функцией распределения $k-1$ случайных величин $X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_k$, среди которых отсутствует величина X_l;

3) по каждой переменной x_l функция $F(x_1, x_2, \ldots, x_k)$ непрерывна слева:

$$\lim_{x_l \uparrow y} F(x_1, \ldots, x_l, \ldots, x_k) = F(x_1, \ldots, y, \ldots, x_k).$$

Доказательство этих свойств точно такое же, как для функции распределения одной случайной величины. Достаточно выделить один из сомножителей в произведении множеств $\prod_{l=1}^{k} \{X_l < x_l\}$, отвечающий выбранной координате функции распределения, и для него почти дословно повторить то, что делалось для доказательства свойств функции распределения.

Набор случайных величин $\vec{X} = (X_1, X_2, \ldots, X_k)$ называют **случайным вектором**, функцию $F(\vec{x}) = F(x_1, x_2, \ldots, x_k)$ – **функцией распределения случайного вектора** $\vec{X}$.

Определение 4. Случайные величины $X_1, X_2, \ldots, X_k$ называют **независимыми**, если для любых $(x_1, x_2, \ldots, x_k) \in \mathbf{R}^k$

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

Это соотношение можно выразить в терминах функций распределения

$$F(x_1, x_2, \ldots, x_k) = F_1(x_1)F_2(x_2) \ldots F_k(x_k), \qquad (11.1)$$

где $F_l(x)$ – функция распределения случайной величины X_l.

Лемма 1. *Если величины $X_1, X_2, \ldots, X_k$ независимы, а $g_l(x), l = 1, 2, \ldots, k$, – некоторые функции, то и величины $Y_1 = g_1(X_1), Y_2 = g_2(X_2), \ldots, Y_k = g_k(X_k)$ являются независимыми.*

Справедливость этого утверждения мы проверим лишь для монотонно возрастающих функций. Оно следует из цепочки равенств

$$\begin{aligned}
&\mathbf{P}(Y_1 < y_1, Y_2 < y_2, \ldots, Y_k < y_k) = \\
&= \mathbf{P}(X_1 < g_1^{-1}(y_1), X_2 < g_2^{-1}(y_2), \ldots, X_k < g_k^{-1}(y_k)) = \\
&= \mathbf{P}(X_1 < g_1^{-1}(y_1))\mathbf{P}(X_2 < g_2^{-1}(y_2)) \ldots \mathbf{P}(X_k < g_k^{-1}(y_k)) = \\
&= \mathbf{P}(Y_1 < y_1)\mathbf{P}(Y_2 < y_2) \ldots \mathbf{P}(Y_k < y_k).
\end{aligned}$$

Замечание 2. Утверждение леммы 1 верно для широкого класса функций $g_l(x)$. К этому классу относятся, например, непрерывные функции.

Определение 5. Если существует такая неотрицательная функция $f(y_1, y_2, \ldots, y_k)$, что функция распределения $F(x_1, x_2, \ldots, x_k)$ для всех $x_1, x_2, \ldots, x_k$ представима в виде

$$F(x_1, x_2, \ldots, x_k) = \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} \cdots \int\limits_{-\infty}^{x_k} f(y_1, y_2, \ldots, y_k) dy_1 dy_2 \ldots dy_k,$$

то $f(x_1, x_2, \ldots, x_k)$ называется **совместной плотностью распределения** случайных величин $X_1, X_2, \ldots, X_k$.

Свойства совместной плотности распределения.

1) $\int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_k) dx_1 dx_2 \ldots dx_k = 1;$

2) $f(x_1, x_2, \ldots, x_k) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \ldots \frac{\partial}{\partial x_k} F(x_1, x_2, \ldots, x_k)$, если функция $F(x_1, x_2, \ldots, x_k)$ является дифференцируемой;

3) для любого l функция

$$f_l(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_k) = \int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_l, \ldots, x_k) dx_l$$

является совместной плотностью распределения $k-1$ случайных величин $X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_k$, среди которых отсутствует величина X_l.

Свойства 1) и 2) устанавливаются аналогично тому, как это было сделано для плотности одной случайной величины в § 10. Убедимся в справедливости свойства 3). Воспользуемся определением плотности и свойством, что можно менять порядок интегрирования. Полагая $x_l = \infty$ получим

$$F(x_1, \ldots, x_{l-1}, \infty, x_{l+1}, \ldots, x_k) =$$
$$= \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_k} \left(\int_{-\infty}^{\infty} f(u_1, \ldots, u_l, \ldots, u_k) du_l \right) du_1, \ldots, du_{l+1}, \ldots, du_k.$$

Согласно второму свойству совместной функции распределения слева стоит совместная функция распределения величин $X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_k$. Снова используя определение плотности, получаем, что выражение в скобках является совместной плотностью распределения этих величин.

Аналогом соотношения (10.3) является следующее:

$$\mathbf{P}(a_1 \leqslant X_1 < b_1, a_2 \leqslant X_2 < b_2, \ldots, a_k \leqslant X_k < b_k)$$
$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_k}^{b_k} f(x_1, x_2, \ldots, x_k) dx_1 dx_2 \ldots dx_k. \qquad (11.2)$$

Если у случайных величин $X_1, X_2, \ldots, X_k$ существует совместная плотность распределения, то в силу свойства 3) существуют плотности распределения у каждой из величин X_l в отдельности. Продифференцировав равенство (11.1) по каждой из переменных x_l, $l = 1, 2, \ldots, k$, легко убедиться в том, что свойство независимости случайных величин $X_1, X_2, \ldots, X_k$ через плотность распределения выражается так: для любых $(x_1, x_2, \ldots, x_k) \in \mathbf{R}^k$

$$f(x_1, x_2, \ldots, x_k) = f_1(x_1) f_2(x_2) \ldots f_k(x_k), \qquad (11.3)$$

где $f_l(x)$ – плотность распределения случайной величины X_l.

Задача 2. Некто получил два кредита, каждый из которых он должен вернуть по первому требованию. Требование вернуть кредит $l (l = 1, 2)$ приходит в случайный момент времени τ_l, причем τ_l имеет плотность распределения $p_l(x)$, определяемую формулой $p_l(x) = \alpha_l \exp(-\alpha_l x)$ при $x \geqslant 0$ и $p_l(x) = 0$ при $x < 0$. Требования приходят независимо. Найти распределение длительности случайного промежутка времени, в течение которого в распоряжении заемщика будут находиться оба кредита.

Решение. Вычислим функцию распределения интересующей нас случайной величины $\min(\tau_1, \tau_2)$, которая как раз и равна времени, в течение которого заемщик может пользоваться обоими кредитами. Имеем

$$P(\min(\tau_1, \tau_2) < x) =$$
$$= 1 - P(\min(\tau_1, \tau_2) \geqslant x) = 1 - P(\{\tau_1 \geqslant x\}\{\tau_2 \geqslant x\}) =$$
$$= 1 - P(\tau_1 \geqslant x)P(\tau_2 \geqslant x) = 1 - (1 - P(\tau_1 < x))(1 - P(\tau_2 < x)).$$

Учитывая вид плотности распределения, получаем, что при $x \geqslant 0$ и $l = 1, 2$,

$$P(\tau_l < x) = \int_0^x \alpha_l \exp(-\alpha_l x) dx = 1 - \exp(-\alpha_l x),$$

и поэтому $P(\min(\tau_1, \tau_2) < x) = 1 - \exp\big(-(\alpha_1 + \alpha_2)x\big)$.

Задачи

Задача 11.1. В кошельке лежат 8 пятикопеечных монет и 6 двухкопеечных. Наудачу вынимают одну за другой 2 монеты. Пусть случайная величина X – достоинство первой монеты, Y – второй. Составить совместный закон распределения величин (X, Y).

Задача 11.2. Из 12 лотерейных билетов 4 выигрышных. Двое вытягивают по билету, сначала первый тянет, затем второй. Пусть случайная величина X – число выигрышных билетов у первого, Y – у второго. Составить совместный закон распределения величин (X, Y).

Задача 11.3. Двое стрелков проиводят по два выстрела. Стреляют они независимо друг от друга, и каждый выстрел не зависит от предыдущего. Вероятность попадания первого α_1, втрого – α_2. Пусть случайная величина X – число попаданий первого стрелка, Y – второго. Составить совместный закон распределения величин (X, Y).

Задача 11.4. Бросают две игральные кости. Пусть $X = 1$, если сумма очков четна, и $X = 0$, если сумма нечетна, а

$Y = 1$, если произведение очков четно, и $Y = 0$, если произведение нечетно. Составить совместный закон распределения величин (X, Y).

Задача 11.5. Совместная функция распределения случайных величин X и Y имеет вид $F(x, y) = 1 - 2^{-x^2} - 3^{-y^2} + 2^{-x^2}3^{-y^2}$. Определить совместную плотность распределения этих величин и вычислить вероятность события $\{1 \leqslant X \leqslant 2, 0 \leqslant Y \leqslant 1\}$.

Задача 11.6. Определить совместную плотность распределения двух положительных случайных величин X и Y по заданной функции распределения $F(x, y) = (1-e^{-ax})(1-e^{-by})$.

Задача 11.7. Определить функцию распределения величины, равной максимуму из двух независимых случайных величин X и Y с функциями распределения $F_X(x)$ и $F_Y(x)$ соответственно.

Задача 11.8. Определить функцию распределения величины, равной минимуму из двух независимых случайных величин X и Y с функциями распределения $F_X(x)$ и $F_Y(x)$ соответственно.

Задача 11.9. Плотность распределения двух случайных величин (X, Y) задается формулой: $f(x, y) = \frac{1}{2}\sin(x + y)$, при $0 \leqslant x \leqslant \pi/2$, $0 \leqslant y \leqslant \pi/2$. Для других значений аргументов полагаем ее равной нулю. Определить совместную функцию распределения величин X и Y.

§ 12. Числовые характеристики случайных величин

1. Дискретные случайные величины.

Пусть X – дискретная случайная величина, принимающая значения $x_l, l = 1, 2, \ldots, n$, с вероятностями $p_l = \mathbf{P}(X = x_l)$.

Определение 1. **Математическим ожиданием** дискретной случайной величины X называется число

$$\mathbf{E}(X) = \sum_{l=1}^{n} x_l p_l.$$

Замечание 1. Если случайная величина X имеет счетное число значений, т. е. $n = \infty$, то говорят, что математическое ожидание существует, если ряд $\sum\limits_{l=1}^{\infty} |x_l| p_l$ сходится, в противном случае говорят, что математического ожидания не существует.

Величину $\mathbf{E}(|X|) = \sum\limits_{l=1}^{\infty} |x_l| p_l$ называют **абсолютным моментом** дискретной случайной величины X.

Смысл математического ожидания наиболее ясно проявляется при рассмотрении схемы равновозможных исходов. Это среднее арифметическое значение всех возможных реализаций случайной величины, которые могут появиться в ходе случайного эксперимента. Действительно, пусть N – число всех равновозможных исходов эксперимента. Пусть значение x_l соответствует m_l исходам, $l = 1, 2, \ldots, n$, и $m_1 + m_2 + \cdots + m_n = N$. Поскольку исходы равновозможны, то $p_l = \mathbf{P}(X = x_l) = \frac{m_l}{N}$. Тогда

$$\mathbf{E}(X) = \sum_{l=1}^{n} x_l p_l = \sum_{l=1}^{n} x_l \frac{m_l}{N} =$$

$$= \frac{1}{N}(\underbrace{x_1 + x_1 + \ldots + x_1}_{m_1} + \underbrace{x_2 + x_2 + \ldots + x_2}_{m_2} + \ldots + \underbrace{x_n + x_n + \ldots + x_n}_{m_n}).$$

Правая часть этого равенства и является средним арифметическим всех возможных реализаций случайной величины X. Поэтому $\mathbf{E}(X)$ часто называют **средним значением** величины X.

Замечание 2. Если есть дискретная случайная величина X, принимающая n различных значений, и некоторая взаимно однозначная функция $g(x)$, то случайная величина $g(X)$ принимает значения $g(x_l)$ с вероятностями $p_l = \mathbf{P}(X = x_l)$, $l = 1, \ldots, n$ (см. замечание 1 § 10). Поэтому $\mathbf{E}(g(X)) = \sum\limits_{l=1}^{n} g(x_l) p_l$. Это равенство выполняется и для функций g, не имеющих обратную, поскольку согласно замечанию 1 § 10 $g(X)$ принимает совпадающее для разных x_l значение с соответствующей суммарной вероятностью.

Отметим важный частный случай. Определим

$$\mathbb{I}_A(x) = \begin{cases} 1, & \text{при } x \in A, \\ 0, & \text{при } x \notin A, \end{cases}$$

– **индикатор множества** A.

Пусть A – счетное множество, $g(x) = \mathbb{I}_A(x)$. Тогда справедливо соотношение

$$\mathbf{E}\big(\mathbb{I}_A(X)\big) = \sum_{l=1}^{n} \mathbb{I}_A(x_l)p_l = \sum_{x_l \in A} p_l = \mathbf{P}(X \in A).$$

Замечание 3. Если даны две дискретные случайные величины X и Y с совместными вероятностями $p_{ij} = \mathbf{P}(X = x_i, Y = y_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ и $h(x, y)$ – некоторая функция двух переменных, то в силу замечания 1 § 11 математическое ожидание случайной величины $h(X, Y)$ вычисляется по формуле

$$\mathbf{E}(h(X, Y)) = \sum_{i=1}^{n} \sum_{j=1}^{m} h(x_i, y_j)p_{ij}.$$

Свойства математического ожидания.

1) (свойство линейности) для любых постоянных a и b

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b;$$

2) (свойство аддитивности) для любых двух случайных величин X и Y

$$\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y);$$

3) если случайные величины X и Y независимы, то

$$\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y),$$

и, более того, для широкого класса функций $g(x)$ и $h(y)$

$$\mathbf{E}(g(X)h(Y)) = \mathbf{E}(g(X))\mathbf{E}(h(Y));$$

4) если $X \geqslant 0$, то $\mathbf{E}(X) \geqslant 0$, если $X \geqslant Y$, то $\mathbf{E}(X) \geqslant \mathbf{E}(Y)$.

Доказательство. Первое свойство следует из свойства линейности операции суммирования. Действительно,

$$\mathbf{E}(aX + b) = \sum_{l=1}^{n}(ax_l + b)p_l = a\sum_{l=1}^{n} x_l p_l + b\sum_{l=1}^{n} p_l = a\mathbf{E}(X) + b.$$

Согласно замечанию 1 § 11 величина $X+Y$ является случайной величиной, принимающей значения $x_i + y_j$ с вероятностями $p_{ij} = \mathbf{P}(X = x_i, Y = y_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, и ее математическое ожидание вычисляется согласно замечанию 3. Используя свойства 1) и 2) для вероятностей p_{ij}, имеем

$$\mathbf{E}(X+Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i + y_j)p_{ij} = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} p_{ij} + \sum_{j=1}^{m} y_j \sum_{i=1}^{n} p_{ij} =$$

$$= \sum_{i=1}^{n} x_i p_i^{(1)} + \sum_{j=1}^{m} y_j p_j^{(2)} = \mathbf{E}(X) + \mathbf{E}(Y).$$

Убедимся в справедливости третьего свойства. Поскольку величины X и Y независимы, то
$p_{ij} = \mathbf{P}(X = x_i, Y = y_j) = \mathbf{P}(X = x_i)\mathbf{P}(Y = y_j) = p_i^{(1)} p_j^{(2)}$.

Тогда

$$\mathbf{E}(g(X)h(Y)) = \sum_{i=1}^{n}\sum_{j=1}^{m} g(x_i)h(y_j)p_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} g(x_i)h(y_j)p_i^{(1)}p_j^{(2)} =$$

$$= \sum_{i=1}^{n} g(x_i)p_i^{(1)} \sum_{j=1}^{m} h(y_j)p_j^{(2)} = \mathbf{E}(g(X))\mathbf{E}(h(Y)).$$

Докажем четвертое свойство. У неотрицательной случайной величины математическое ожидание будет неотрицательным, поскольку это сумма неотрицательных слагаемых. Если $X \geqslant Y$, то $X - Y \geqslant 0$, и $\mathbf{E}(X - Y) \geqslant 0$. По второму и первому свойствам $\mathbf{E}(X - Y) = \mathbf{E}(X) - \mathbf{E}(Y)$. Следовательно, $\mathbf{E}(X) - \mathbf{E}(Y) \geqslant 0$, а значит $\mathbf{E}(X) \geqslant \mathbf{E}(Y)$.

Следствие 1. *Для любой случайной величины X*

$$|\mathbf{E}(X)| \leqslant \mathbf{E}(|X|). \tag{12.1}$$

Действительно, поскольку $X \leqslant |X|$ и $-X \leqslant |X|$, то по свойствам 4) и 1) имеем $\mathbf{E}(X) \leqslant \mathbf{E}(|X|)$ и $-\mathbf{E}(X) \leqslant \mathbf{E}(|X|)$. Что и требовалось доказать.

Следствие 2. *Для любых случайных величин* $X_1, X_2, \ldots, X_k$

$$\mathbf{E}\Big(\sum_{l=1}^{k} X_l\Big) = \sum_{l=1}^{k} \mathbf{E}(X_l).$$

Это утверждение доказывается с помощью свойства 2) по индукции.

Определение 2. **Дисперсией** дискретной случайной величины X называется число $\mathbf{D}(X) = \sum\limits_{l=1}^{n}(x_l - \mathbf{E}(X))^2 p_l$.

В силу замечания 2 дисперсией случайной величины X служит математическое ожидание (среднее) случайной величины $(X - \mathbf{E}(X))^2$, т. е. $\mathbf{D}(X) = \mathbf{E}(X - \mathbf{E}(X))^2$. Смысл дисперсии заключается в том, что она характеризует средний квадратичный разброс случайной величины вокруг своего математического ожидания.

Замечание 4. Если $\mathbf{D}(X) = 0$, то в силу определения дисперсии $\mathbf{P}(X = \mathbf{E}(X)) = 1$, т. е. фактически X – **детерминированная** (не случайная) величина.

Пример 1. Пусть X – число гербов, выпавших при четырех бросаниях правильной монеты. Тогда, используя результат примера 2 из § 10, получим

$$\mathbf{E}(X) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{6}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} = 2,$$

$$\mathbf{D}(X) = (0-2)^2 \cdot \frac{1}{16} + (1-2)^2 \cdot \frac{4}{16} +$$
$$+ (2-2)^2 \cdot \frac{6}{16} + (3-2)^2 \cdot \frac{4}{16} + (4-2)^2 \cdot \frac{1}{16} = 1.$$

Свойства дисперсии.

1) для любой случайной величины X

$$\mathbf{D}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X);$$

2) для любых постоянных a и b

$$\mathbf{D}(aX + b) = a^2\mathbf{D}(X);$$

3) если случайные величины X и Y независимы, то

$$\mathbf{D}(X + Y) = \mathbf{D}(X) + \mathbf{D}(Y).$$

Доказательство. Первое свойство устанавливается с помощью следующей цепочки равенств:

$$\mathbf{D}(X) = \sum_{l=1}^{n}(x_l - \mathbf{E}(X))^2 p_l = \sum_{l=1}^{n}(x_l^2 - 2x_l\mathbf{E}(X) + \mathbf{E}^2(X))p_l =$$
$$= \sum_{l=1}^{n} x_l^2 p_l - 2\mathbf{E}(X)\sum_{l=1}^{n} x_l p_l + \mathbf{E}^2(X)\sum_{l=1}^{n} p_l =$$
$$= \mathbf{E}(X^2) - 2\mathbf{E}(X)\mathbf{E}(X) + \mathbf{E}^2(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X).$$

Используя определение дисперсии и свойство линейности математического ожидания, получим

$$\mathbf{D}(aX + b) = \mathbf{E}(aX + b - \mathbf{E}(aX + b))^2 =$$
$$= \mathbf{E}(aX - a\mathbf{E}(X))^2 = a^2\mathbf{E}(X - \mathbf{E}(X))^2 = a^2\mathbf{D}(X).$$

Для доказательства третьего свойства воспользуемся уже доказанным первым свойством дисперсии и вторым и третьим свойствами математического ожидания. Имеем

$$\mathbf{D}(X + Y) = \mathbf{E}(X + Y)^2 - \mathbf{E}^2(X + Y) =$$
$$= \mathbf{E}(X^2 + 2XY + Y^2) - (\mathbf{E}(X) + \mathbf{E}(Y))^2 =$$
$$= \mathbf{E}(X^2) + 2\mathbf{E}(XY) + \mathbf{E}(Y^2) - \mathbf{E}^2(X) - 2\mathbf{E}(X)\mathbf{E}(Y) - \mathbf{E}^2(Y) =$$
$$= \mathbf{E}(X^2) - \mathbf{E}^2(X) + \mathbf{E}(Y^2) - \mathbf{E}^2(Y) = \mathbf{D}(X) + \mathbf{D}(Y).$$

Следствие 3. *Для независимых случайных величин* X_1, $X_2, \ldots, X_k$

$$\mathbf{D}\Big(\sum_{l=1}^{k} X_l\Big) = \sum_{l=1}^{k}\mathbf{D}(X_l).$$

Это утверждение доказывается с помощью свойства 3 по индукции.

Задача 1 (о сборе коллекции). С каждой единицей товара покупатель приобретает наудачу некоторый элемент коллекции, состоящей из n различных элементов (наудачу означает, что при каждой покупке покупатель равновероятно получает любой из этих n элементов независимо от придыдущих приобретений). Какое среднее число покупок надо сделать, чтобы собрать полную коллекцию?

Р е ш е н и е. Пусть случайная величина X – число покупок, которое надо сделать, чтобы собрать полную коллекцию, а величины Y_s, $s = 2, 3, \dots, n$, – число (случайное) покупок, до приобретения нового элемента после того, как уже собран $s-1$ элемент. Очевидно, что $X = 1 + \sum\limits_{s=2}^{n} Y_s$, и, следовательно, $\mathbf{E}X = 1 + \sum\limits_{s=2}^{n} \mathbf{E}Y_s$. Событие {потребуется l покупок до приобретения нового элемента после того, как уже собран $s-1$ элемент } распадается в произведение l событий. Каждое из первых $l-1$ событий состоит в том, что приобретается уже имеющийся элемент. Вероятность такого события равна $\frac{s-1}{n}$. Последнее событие состоит в том, что приобретается новый элемент. Его вероятность $\frac{n-s+1}{n}$. Используя формулу произведения вероятностей для независимых событий, получаем

$$\mathbf{P}(Y_s = l) = \frac{(s-1)^{l-1}(n-s+1)}{n^l}, \qquad l = 1, 2, 3, \dots.$$

Таким образом,

$$\mathbf{E}Y_s = (n-s+1)\sum_{l=1}^{\infty} \frac{l(s-1)^{l-1}}{n^l} = (n-s+1)\Big(\sum_{l=1}^{\infty} \frac{(s-1)^{l-1}}{n^l}\Big)'_s =$$

$$= (n-s+1)\Big(\frac{s-1}{n(1-\frac{s-1}{n})}\Big)'_s = (n-s+1)\Big(\frac{1}{1-\frac{s-1}{n}} - 1\Big)'_s =$$

$$= \frac{(n-s+1)}{n(1-\frac{s-1}{n})^2} = \frac{n}{n-s+1}.$$

Следовательно,

$$\mathbf{E}X = 1 + \sum_{s=2}^{n} \frac{n}{n-s+1} = 1 + n\sum_{k=1}^{n-1} \frac{1}{k} = n\sum_{k=1}^{n} \frac{1}{k}.$$

Итак, для того, чтобы собрать всю коллекцию в среднем надо сделать $n \sum\limits_{k=1}^{n} \frac{1}{k}$ покупок, что при больших n эквивалентно $n \ln n$ покупкам.

2. Случайные величины, имеющие плотность распределения.

Пусть функция распределения $F(x)$ случайной величины X имеет плотность распределения $f(x)$.

Определение 3. **Математическим ожиданием** случайной величины X с плотностью распределения $f(x)$ называется число $\mathbf{E}(X) = \int\limits_{-\infty}^{\infty} xf(x)dx$.

Говорят, что математическое ожидание у случайной величины X существует, если у нее конечен **абсолютный момент** $\mathbf{E}(|X|) = \int\limits_{-\infty}^{\infty} |x|f(x)dx$.

Замечание 5. Если есть случайная величина X с плотностью распределения $f(x)$ и некоторая непрерывная функция $g(x)$, то математическое ожидание случайной величин $g(X)$ вычисляется по формуле

$$\mathbf{E}(g(X)) = \int\limits_{-\infty}^{\infty} g(x)f(x)dx.$$

Если $g(x)$ – монотонно возрастающая функция, то эта формула является следствием формулы (10.6) и правила замены переменной интегрирования. Для важного частного случая, когда $g(x) = \mathbb{I}_A(x)$, где A – интервал $[a, b)$ или объединение интервалов такого вида, справедливо соотношение

$$\mathbf{E}(\mathbb{I}_A(X)) = \int\limits_{-\infty}^{\infty} \mathbb{I}_A(x)f(x)dx = \int\limits_{A} f(x)dx = \mathbf{P}(X \in A).$$

Замечание 6. Если даны две случайные величины X и Y с совместной плотностью распределения $f(x, y)$ и случайная величина $h(X, Y)$, где $h(x, y)$ – непрерывная функция двух переменных, то

$$\mathbf{E}(h(X,Y)) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h(x,y)f(x,y)dxdy.$$

Определение 4. **Дисперсией** случайной величины X с плотностью распределения $f(x)$ называется число

$$\mathbf{D}(X) = \int\limits_{-\infty}^{\infty} (x - \mathbf{E}(X))^2 f(x)dx.$$

Справедливы все свойства математического ожидания и дисперсии, сформулированные для дискретных величин. Принципиальных отличий в доказательстве этих свойств от доказательств соответствующих свойств для дискретных величин нет. Доказательство второго и третьего свойства дисперсии вообще не использует дискретность X, а использует лишь свойства математического ожидания.

Докажем свойство аддитивности математического ожидания. Пусть $f(x, y)$ совместная плотность распределения величин X и Y. Воспользуемся свойством 3) совместной плотности двух случайных величин, согласно которому функция $f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy$ является плотностью распределения величины X, а функция $f_2(y) = \int_{-\infty}^{\infty} f(x, y)dx$ является плотностью распределения величины Y. Для вычисления математического ожидания суммы $X + Y$ применим замечание 5. Тогда получим

$$\mathbf{E}(X+Y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x+y)f(x,y)dxdy =$$

$$= \int\limits_{-\infty}^{\infty} x \int\limits_{-\infty}^{\infty} f(x,y)dydx + \int\limits_{-\infty}^{\infty} y \int\limits_{-\infty}^{\infty} f(x,y)dxdy =$$

$$= \int\limits_{-\infty}^{\infty} xf_1(x)dx + \int\limits_{-\infty}^{\infty} yf_2(y)dy = \mathbf{E}(X) + \mathbf{E}(Y).$$

Свойство аддитивности доказано. Для того, чтобы убедиться в справедливости третьего свойства математического ожидания, воспользуемся соотношением (11.3), которое для двух случайных независимых величин X и Y имеет вид:

$f(x, y) = f_1(x)f_2(y)$. Для вычисления $\mathbf{E}(g(X)h(Y))$ снова применим замечание 5. Тогда будем иметь

$$\mathbf{E}(g(X)h(Y)) =$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_1(x)f_2(y)dxdy =$$
$$= \int_{-\infty}^{\infty} g(x)f_1(x)dx \int_{-\infty}^{\infty} h(y)f_2(y)dy = \mathbf{E}(g(X))\mathbf{E}(g(Y)).$$

Пример 2. Пусть имеется однородная проволока, растягиваемая за концы. Любой точке проволоки сопоставим координату, равную расстоянию от этой точки до левого конца плюс некоторое число a. Тогда левый конец проволоки имеет координату a. Пусть b – координата правого конца проволоки. Очевидно, что $a < b$. Пусть X – случайная величина, равная координате точки разрыва. Ясно, что разрыв происходит в промежутке $[a, b]$. Используя геометрические вероятности, аналогично примеру 4 § 10 найдем, что функция распределения и плотность распределения X имеют вид

$$F(x) = \begin{cases} 0, & \text{при } x \leqslant a, \\ \dfrac{x-a}{b-a}, & \text{при } a < x < b, \\ 1, & \text{при } b \leqslant x, \end{cases} \quad f(x) = \begin{cases} 0, & \text{при } x \leqslant a, \\ \dfrac{1}{b-a}, & \text{при } a < x < b, \\ 0, & \text{при } b \leqslant x. \end{cases}$$

Случайная величина с таким распределением называется **равномерно распределенной** на отрезке $[a, b]$. Ее математическое ожидание и дисперсия имеют следующие значения

$$\mathbf{E}(X) = \frac{1}{b-a}\int_a^b x dx = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{b+a}{2},$$

$$\mathbf{D}(X) = \frac{1}{b-a}\int_a^b \left(x - \frac{b+a}{2}\right)^2 dx = \frac{1}{b-a}\int_{(a-b)/2}^{(b-a)/2} v^2 dv = \frac{(b-a)^2}{12}.$$

Лемма 1 (неравенство Коши–Буняковского). *Для любых случайных величин X и Y*

$$\mathbf{E}(|XY|) \leqslant \big(\mathbf{E}(X^2)\big)^{1/2}\big(\mathbf{E}(Y^2)\big)^{1/2} \tag{12.2}$$

Доказательство. Для любых вещественных чисел z_1, z_2 и α справедливо неравенство

$$z_1 z_2 \leqslant \frac{1}{2}\Big(\frac{1}{\alpha^2}z_1^2 + \alpha^2 z_2^2\Big),$$

которое вытекает из очевидного неравенства $(\alpha^{-1}z_1 - \alpha z_2)^2 \geqslant 0$. Используя это неравенство и свойства 1), 2) и 4) математического ожидания, несложно убедиться, что

$$\mathbf{E}(|XY|) \leqslant \mathbf{E}\Big(\frac{1}{2}\Big(\frac{1}{\alpha^2}X^2 + \alpha^2 Y^2\Big)\Big) \leqslant \frac{1}{2}\Big(\frac{1}{\alpha^2}\mathbf{E}(X^2) + \alpha^2\mathbf{E}(Y^2)\Big).$$

Полагая здесь $\alpha^2 = \big(\mathbf{E}(X^2)\big)^{1/2}/\big(\mathbf{E}(Y^2)\big)^{1/2}$, получаем (12.2).

Следствие 4. *Для любой случайной величины X*

$$\mathbf{E}(|X|) \leqslant \big(\mathbf{E}(X^2)\big)^{1/2}. \tag{12.3}$$

Убедиться в справедливости этого неравенства можно, положив в (12.2) $Y = 1$.

Пусть k – некоторое натуральное число.

Определение 5. **Моментом порядка k** случайной величины X называется число $m_k = \mathbf{E}(X^k)$.

Определение 6. **Абсолютным моментом порядка k** случайной величины X называется число $\mathbf{E}(|X|^k)$.

Определение 7. **Центральным моментом порядка k** случайной величины X называется число $\mu_k = \mathbf{E}(X - \mathbf{E}(X))^k$.

Определение 8. **Ковариацией (корреляционным моментом)** двух случайных величин X и Y называется число

$$\operatorname{cov}(X,Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))).$$

Свойства ковариации.

1) для любых случайных величин X и Y

$$\operatorname{cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y);$$

2) (аддитивность) для любых X, Y и Z

$$\mathrm{cov}(X+Y,Z)=\mathrm{cov}(X,Z)+\mathrm{cov}(Y,Z),$$
$$\mathrm{cov}(X,Y+Z)=\mathrm{cov}(X,Y)+\mathrm{cov}(X,Z).$$

Это свойство легко распространяется на сумму произвольного числа слагаемых.

Первое свойство в силу линейности математического ожидания вытекает из равенств

$$\mathrm{cov}(X,Y)=\mathbf{E}(XY-Y\mathbf{E}(X)-X\mathbf{E}(Y)+\mathbf{E}(X)\mathbf{E}(Y))=$$
$$=\mathbf{E}(XY)-\mathbf{E}(X)\mathbf{E}(Y)-\mathbf{E}(X)\mathbf{E}(Y)+\mathbf{E}(X)\mathbf{E}(Y).$$

Второе свойство столь же очевидно.

Определение 9. **Коэффициентом корреляции** двух случайных величин X и Y называется число $r(X,Y)=$ $=\dfrac{\mathrm{cov}(X,Y)}{\sqrt{\mathbf{D}(X)\mathbf{D}(Y)}}$.

Моменты и коэффициент корреляции определены не всегда. Например, для существования обычных и центральных моментов m_k, μ_k необходимо, чтобы был конечен абсолютный момент: $\mathbf{E}|X|^k<\infty$.

Корреляционный момент и коэффициент корреляции характеризуют степень линейной зависимости случайных величин. Чем больше по модулю коэффициент корреляции, тем сильнее линейная зависимость между величинами.

Определение 10. Если ковариация или коэффициент корреляции равны нулю, то такие величины называются **некоррелированными.**

Для любых случайных величин X и Y

$$\mathbf{D}(X\pm Y)=\mathbf{E}(X\pm Y-(\mathbf{E}(X)\pm\mathbf{E}(Y)))^2=$$
$$=\mathbf{E}((X-\mathbf{E}(X))^2\pm 2(X-\mathbf{E}(X))(Y-\mathbf{E}(Y))+(Y-\mathbf{E}(Y))^2)=$$
$$=\mathbf{E}((X-\mathbf{E}(X)^2)\pm 2\mathbf{E}((X-\mathbf{E}(X)(Y-\mathbf{E}(Y))+\mathbf{E}((Y-\mathbf{E}(Y))^2=$$
$$=\mathbf{D}(X)\pm 2\,\mathrm{cov}(X,Y)+\mathbf{D}(Y). \qquad (12.4)$$

Отметим вытекающее из (12.4) важное свойство некоррелированных величин. Если случайные величины X и Y некоррелированы, то

$$\mathbf{D}(X+Y)=\mathbf{D}(X)+\mathbf{D}(Y).$$

Следствие 3. *Для попарно некоррелированных случайных величин* $X_1, \ldots, X_k$

$$\mathbf{D}\Big(\sum_{l=1}^{k} X_l\Big) = \sum_{l=1}^{k} \mathbf{D}(X_l).$$

Это утверждение доказывается по индукции с применением свойства аддитивности ковариации и вышеприведенного свойства некоррелированных величин.

Свойства коэффициента корреляции.

1) Для любых постоянных a, b, c, d

$$r(aX + b, cY + d) = \operatorname{sign}(ac) r(X, Y),$$

где $\operatorname{sign} x = x/|x|$– функция знака.

Это свойство можно прокомментировать так: линейные преобразования случайных величин не изменяют степени их линейной зависимости. При этом меняется лишь знак зависимости, если a и c разных знаков.

2) $|r(X,Y)| \leqslant 1$, и $|r(X,Y)| = 1$ тогда и только тогда, когда существуют такие постоянные a и b, что $Y = aX + b$ с вероятностью единица.

3) Если величины X и Y независимы, то $r(X,Y) = 0$.

Доказательство. В силу свойства линейности математического ожидания имеем

$$\begin{aligned}
&\operatorname{cov}(aX + b, cY + d) = \\
&= \mathbf{E}((aX + b - \mathbf{E}(aX + b))(cY + d - \mathbf{E}(cY + d))) = \\
&= \mathbf{E}((aX - a\mathbf{E}(X))(cY - c\mathbf{E}(Y))) = ac \operatorname{cov}(X, Y).
\end{aligned}$$

Поскольку $\mathbf{D}(aX + b) = a^2\mathbf{D}(X)$, $\mathbf{D}(cY + d) = c^2\mathbf{D}(Y)$, то

$$\frac{\operatorname{cov}(aX + b, cY + d)}{\sqrt{\mathbf{D}(aX + b)\mathbf{D}(cY + d)}} = \operatorname{sign}(ac)\frac{\operatorname{cov}(X, Y)}{\sqrt{\mathbf{D}(X)\mathbf{D}(Y)}}.$$

Таким образом, первое свойство установлено.

Докажем свойство 2). Положим $\widetilde{X} = \dfrac{X}{\sqrt{\mathbf{D}(X)}}$, $\widetilde{Y} = \dfrac{Y}{\sqrt{\mathbf{D}(Y)}}$. Согласно (12.4)

$$\mathbf{D}(\widetilde{X} \pm \widetilde{Y}) = \mathbf{D}(\widetilde{X}) \pm 2\operatorname{cov}(\widetilde{X}, \widetilde{Y}) + \mathbf{D}(\widetilde{Y}) =$$

$$= \frac{\mathbf{D}(X)}{\mathbf{D}(X)} \pm 2\frac{\operatorname{cov}(X, Y)}{\sqrt{\mathbf{D}(X)\mathbf{D}(Y)}} + \frac{\mathbf{D}(Y)}{\mathbf{D}(Y)} = 2(1 \pm r(X, Y)). \qquad (12.5)$$

Поскольку левая часть этого равенства неотрицательна, то $1 + r(X,Y) \geqslant 0$ и $1 - r(X,Y) \geqslant 0$, что эквивалентно $-1 \leqslant r(X,Y) \leqslant 1$ или $|r(X,Y)| \leqslant 1$.

При $r(X,Y) = \pm 1$ из (12.5) следует, что $\mathbf{D}(\widetilde{X} \mp \widetilde{Y}) = 0$. Согласно замечанию 4 это влечет $\widetilde{X} \mp \widetilde{Y} = \mathbf{E}(\widetilde{X}) \mp \mathbf{E}(\widetilde{Y})$ с вероятностью единица, и, следовательно,

$$Y = \pm\frac{\sqrt{\mathbf{D}(Y)}}{\sqrt{\mathbf{D}(X)}}X + \left(\mathbf{E}(Y) \mp \frac{\sqrt{\mathbf{D}(Y)}}{\sqrt{\mathbf{D}(X)}}\mathbf{E}(X)\right).$$

Пусть теперь $Y = aX + b$. Согласно первому свойству имеем

$$r(X,Y) = \operatorname{sign}(a) r(X,X) = \operatorname{sign}(a)\frac{\mathbf{E}((X-\mathbf{E}(X)(X-\mathbf{E}(X))}{\sqrt{\mathbf{D}(X)\mathbf{D}(X)}} = \operatorname{sign}(a).$$

Убедимся в справедливости третьего свойства. Если величины X и Y независимы, то, согласно определению, величины $X - \mathbf{E}(X)$ и $Y - \mathbf{E}(Y)$ тоже будут независимы. Поэтому по третьему свойству математического ожидания

$$\operatorname{cov}(X,Y) = \mathbf{E}(X - \mathbf{E}(X))\mathbf{E}(Y - \mathbf{E}(Y)) = 0 \cdot 0 = 0.$$

Следовательно, коэффициент корреляции независимых величин равен нулю.

Обратное утверждение неверно, т. е. из равенства нулю коэффициента корреляции не следует независимость величин. Подтверждением этому служит следующий пример.

Пример. Пусть случайная величина X принимает значения ± 1, ± 2 с вероятностями $1/4$, и пусть $Y = X^2$. Тогда совместный закон распределения величин (X,Y) следующий:

(X,Y)	$(-1,1)$	$(1,1)$	$(2,4)$	$(-2,4)$
$\mathbf{P}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

и нетрудно убедиться в том, что величины X и Y зависимы, а их коэффициент корреляции равен нулю.

Задача 1. Вычислить коэффициент корреляции случайных величин X и Y, определенных в задаче 1 § 11.

Решение. По первому свойству коэффициента корреляции он совпадает с коэффициентом для величин $X_1 = X - 1$ и $Y_1 = Y - 1$, совместный закон распределения которых имеет вид

(X_1, Y_1)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$\mathbf{P}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{3}{10}$

По свойству вероятностей $p_{i,j}$ из § 11 закон распределения X_1 следующий: $\mathbf{P}(X_1 = 0) = \frac{2}{5}$, $\mathbf{P}(X_1 = 1) = \frac{3}{5}$. Аналогичное распределение имеет Y_1. Тогда $\mathbf{E}(X_1) = \frac{3}{5}$, $\mathbf{D}(X_1) = \frac{3}{5} - \frac{3^2}{5^2} = \frac{6}{25}$, $\mathbf{E}(X_1Y_1) = \frac{3}{10}$. Окончательно имеем

$$r(X_1, Y_1) = \frac{\mathbf{E}(X_1Y_1) - (\mathbf{E}(X_1))^2}{\mathbf{D}(X_1)} = \frac{\frac{3}{10} - \frac{9}{25}}{\frac{6}{25}} = -\frac{1}{4}.$$

Задачи

Задача 12.1. Вычислить математическое ожидание и дисперсию величины X, равной числу выпавших очков на игральной кости.

Задача 12.2. Вычислить математическое ожидание и дисперсию величины из задачи 10.11.

Задача 12.3. Вычислить математическое ожидание и дисперсию величины из задачи 10.12.

Задача 12.4. Закон распределения случайной величины X имеет вид

X	-2	-1	0	2
$\mathbf{P}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{4}$

Вычислить $\mathbf{E}X$ и $\mathbf{D}X$.

Задача 12.5. Есть правильный жетон, у которого на одной стороне стоит цифра 2, а на другой – 0, и есть правильный кубик, у которого на противоположных гранях написаны цифры 1,2 и 3 соответственно. Жетон и кубик бросаются на стол. Пусть X – случайная величина, равная сумме очков на жетоне и кубике. Построить закон распределения величины X, вычислить математическое ожидание и дисперсию.

Задача 12.6. Есть два правильных жетона, у одного из них на одной стороне стоит цифра 3, на другой – 5, а у другого на одной стороне стоит цифра 1, на другой – 5. Жетоны бросаются на стол. Пусть X – случайная величина, равная сумме очков на жетонах. Вычислить математическое ожидание и дисперсию.

Задача 12.7. Есть правильный кубик, у которого на противоположных гранях написаны цифры 1,2 и 3 соответственно. Пусть X – число единиц, выпавших при трех бросаниях кубика. Найти закон распределения, вычислить математическое ожидание и дисперсию.

Задача 12.8. Случайная величина X задается плотностью распределения

$$f(x) = \begin{cases} 0, & \text{при} \quad x < 3, \\ \frac{7-x}{8}, & \text{при} \quad 3 \leqslant x \leqslant 7, \\ 0, & \text{при} \quad 7 < x. \end{cases}$$

Вычислить $\mathbf{E}X$.

Задача 12.9. Случайная величина X задается плотностью распределения

$$f(x) = \begin{cases} 0 & \text{при} \quad x < 1, \\ x - \frac{1}{2} & \text{при} \quad 1 \leqslant x \leqslant 2, \\ 0 & \text{при} \quad 2 < x. \end{cases}$$

Вычислить $\mathbf{E}X$ и $\mathbf{D}X$.

Задача 12.10. Вычислить коэффициент корреляции величин X и Y, определенных в задаче 11.2.

Задача 12.11. Вычислить коэффициент корреляции величин X и Y, определенных в задаче 11.4.

Задача 12.12. Вычислить коэффициент корреляции величин X и Y, определенных в задаче 11.1.

Задача 12.13. Вычислить коэффициент корреляции величин X и Y, определенных в задаче 11.9.

§ 13. Производящие и характеристические функции

Пусть X – неотрицательная целочисленнозначная случайная величина, т. е. X принимает значения 0,1,2,... с вероятностями $p_k = \mathbf{P}(X = k)$.

Определение 1. **Производящей функцией** целочисленнозначной неотрицательной случайной величины X называется функция $\psi(z) = \sum\limits_{k=0}^{\infty} z^k p_k = \mathbf{E}(z^X)$, определенная для комплексных z таких, что $|z| \leqslant 1$. Для $|z| \leqslant 1$ ряд, определяющий производящую функцию, равномерно сходится, так как

$$|\psi(z)| = \left|\sum_{k=0}^{\infty} z^k p_k\right| \leqslant \sum_{k=0}^{\infty} |z|^k p_k \leqslant \sum_{k=0}^{\infty} p_k = 1.$$

Замечание 1. Производящая функция однозначно определяет распределение случайной величины, поскольку $p_n = \frac{1}{n!}\frac{d^n}{dz^n}\psi(z)\Big|_{z=0}$.
Действительно, дифференцируя под знаком суммы, получим

$$\frac{d^n}{dz^n}\psi(z) = \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1)z^{k-n}p_k. \qquad (13.1)$$

При $z = 0$ все слагаемые в правой части (13.1), кроме первого, обращаются в нуль. Первое же слагаемое равно $n!p_n$.

Лемма 1 (о вычислении моментов). *Пусть неотрицательная целочисленнозначная случайная величина X имеет момент n-го порядка. Тогда*

$$\frac{d^n}{dz^n}\psi(z)\Big|_{z=1} = \mathbf{E}(X(X-1)\dots(X-n+1)).$$

Доказательство леммы непосредственно следует из формулы (13.1), если в ней положить $z = 1$.

Поскольку

$$\mathbf{D}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}^2(X),$$

то из леммы 1 вытекает следующий результат.

Следствие 1. Математическое ожидание и дисперсия выражаются формулами

$$\mathbf{E}(X) = \psi'(1), \qquad \mathbf{D}(X) = \psi''(1) + \psi'(1) - (\psi\ (1))^2.$$

Свойства производящих функций.

1) $|\psi(z)| \leqslant 1$ при $|z| \leqslant 1$.

Действительно, так как X неотрицательна

$$|\psi(z)| \leqslant \mathbf{E}(|z^X|) = \mathbf{E}(|z|^X) \leqslant 1.$$

Для того, чтобы различать производящие функции разных случайных величин, будем случайную величину приписывать к производящей функции в качестве индекса.

2) Для любых вещественных a и b

$$\psi_{aX+b}(z) = z^b \psi_X(z^a).$$

Доказательство опирается на свойство линейности математического ожидания:

$$\psi_{aX+b}(z) = \mathbf{E}\big(z^{(aX+b)}\big) = z^b \mathbf{E}\big((z^a)^X\big) = z^b \psi_X(z^a).$$

3) Если X и Y независимые случайные величины, то

$$\psi_{X+Y}(z) = \psi_X(z)\psi_Y(z).$$

Доказательство. Поскольку величины X и Y независимы, то, используя третье свойство математического ожидания, получаем

$$\psi_{X+Y}(z) = \mathbf{E}\big(z^{(X+Y)}\big) = \mathbf{E}\big(z^X z^Y\big) = \mathbf{E}\big(z^X\big)\mathbf{E}\big(z^Y\big) = \psi_X(z)\psi_Y(z).$$

Задача 1. Найти закон распределения случайной величины X с производящей функцией $\psi(z) = \frac{1}{4}(1+z)^2$.

Решение. Поскольку

$$\psi(z) = \frac{1}{4}(1+z)^2 = \frac{1}{4} + \frac{1}{2}z + \frac{1}{4}z^2,$$

то сравнивая это выражение с определением производящей функции $\psi(z) = \sum\limits_{k=0}^{\infty} z^k p_k$, где $p_k = \mathbf{P}(X = k)$, получаем, что

$$\mathbf{P}(X=0) = \frac{1}{4}, \quad \mathbf{P}(X=1) = \frac{1}{2}, \quad \mathbf{P}(X=2) = \frac{1}{4},$$

$$\mathbf{P}(X=k) = 0, \qquad \text{при } k > 2.$$

Определение 2. **Характеристической функцией** случайной величины X называется комплекснозначная функция $\varphi(t) = \mathbf{E}(e^{itX})$, определенная для всех действительных значений t.

Для дискретной случайной величины с вероятностями $p_k = \mathbf{P}(X = x_k), k = 1, 2, \ldots$, характеристическая функция будет определяться формулой

$$\varphi(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k, \tag{13.2}$$

а для величины, имеющей плотность распределения $f(x)$, – формулой

$$\varphi(t) = \int\limits_{-\infty}^{\infty} e^{itx} f(x) dx. \tag{13.3}$$

Для случайной величины с произвольной функцией распределения $F(x)$ математическое ожидание $\mathbf{E}(e^{itX})$ определяется с помощью интеграла Стилтьеса, т. е. характеристическая функция определяется формулой

$$\varphi(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF(x). \tag{13.4}$$

Для непрерывной ограниченной функции $g(x)$ и неубывающей ограниченной непрерывной слева функции $F(x)$ интеграл Стилтьеса $\int\limits_{-\infty}^{\infty} g(x) dF(x)$ существует и его можно определить равенством

$$\int\limits_{-\infty}^{\infty} g(x) dF(x) = \lim_{n\to\infty} \sum_{k=-n^2}^{n^2} g\left(\frac{k}{n}\right)\left(F\left(\frac{k+1}{n}\right) - F\left(\frac{k}{n}\right)\right).$$

Формула (13.4) является более общей, чем (13.2) и (13.3). Так, согласно определению интеграла Стилтьеса, при

$$F(x) = \sum_{k=1}^{\infty} p_k \mathbb{I}_{(x_k,\infty)}(x)$$

из (13.4) следует (13.2), а в случае, когда у $F(x)$ существует плотность ($dF(x) = f(x)dx$), (13.4) превращается в (13.3).

Замечание 2. Для целочисленных случайных величин $\varphi(t) = \psi(e^{it})$.

Свойства характеристических функций.

1) $\varphi(0) = 1$ и $|\varphi(t)| \leqslant 1$ для всех t.

Действительно: $\varphi(0) = \mathbf{E}(e^0) = \mathbf{E}(1) = 1$; $|\varphi(t)| \leqslant \mathbf{E}|e^{itX}| = \mathbf{E}(1) = 1$.

Для того, чтобы различать характеристические функции разных случайных величин, будем случайную величину приписывать к характеристической функции в качестве индекса.

2) Для любых вещественных a и b

$$\varphi_{aX+b}(t) = e^{itb}\varphi_X(at).$$

3) Если X и Y независимые случайные величины, то

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t).$$

Свойства 2) и 3) следуют, в силу замечания 2, из аналогичных свойств для производящих функций. То обстоятельство, что замечание 2 сформулировано для целочисленных величин, принципиального значения для вывода свойств 2) и 3) не имеет.

Следствие 2. *Если $X_1, X_2, \ldots, X_k$ – независимые случайные величины, то*

$$\varphi_{\sum_{l=1}^k X_l}(t) = \prod_{l=1}^{k} \varphi_{X_l}(t),$$

т. е. характеристическая функция суммы независимых случайных величин распадается в произведение характеристических функций слагаемых.

4) Характеристическая функция $\varphi(t)$ является равномерно непрерывной функцией.

Доказательство. Так как $F(-\infty) = 0$, $F(\infty) = 1$, а функция $F(x)$ не убывает, то для любого $\varepsilon > 0$ можно выбрать столь большое $c > 0$, что будут выполняться неравенства $F(-c) \leqslant \varepsilon, 1 - F(c) \leqslant \varepsilon$. Тогда для любого $h > 0$ получим

$$|\varphi(t+h) - \varphi(t)| \leqslant \int\limits_{-\infty}^{\infty} |e^{i(t+h)x} - e^{itx}| dF(x) \leqslant \int\limits_{-c}^{c} |e^{ihx} - 1| dF(x) +$$

$$+2 \int\limits_{|x|>c} dF(x) \leqslant h \int\limits_{-c}^{c} |x| dF(x) + 2\big(F(-c) + 1 - F(c)\big) \leqslant hc + 4\varepsilon.$$

В силу того, что ε и h выбирались произвольными, из этой оценки следует равномерная непрерывность функции $\varphi(t)$.

Лемма 2 (о вычислении моментов). *Пусть случайная величина X имеет абсолютный момент n-го порядка. Тогда характеристическая функция случайной величины X дифференцируема n раз и при $0 \leqslant k \leqslant n$*

$$\mathbf{E}(X^k) = (-i)^k \frac{d^k}{dt^k} \varphi(t)\Big|_{t=0}.$$

Доказательство. Используя формулу дифференцирования показательной функции, получим

$$\frac{d^k}{dt^k}\varphi(t) = \frac{d^k}{dt^k}\mathbf{E}(e^{itX}) = \mathbf{E}\Big(\frac{d^k}{dt^k}e^{itX}\Big) = \mathbf{E}\big(i^k X^k e^{itX}\big) = i^k \mathbf{E}\big(X^k e^{itX}\big).$$

Деля это равенство на i^k и полагая в нем $t = 0$, получим требуемое утверждение. Внесение дифференцирования под знак математического ожидания корректно, так как математическое ожидание – это либо сумма, либо интеграл, а вносить дифференцирование под знак суммы и интеграла в указанных предположениях можно.

Поскольку $\mathbf{D}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X)$, то из леммы 2 вытекает следующий результат.

Следствие 3. *Математическое ожидание и дисперсия выражаются формулами*

$$\mathbf{E}(X) = -i\varphi'(0), \qquad \mathbf{D}(X) = -\varphi''(0) + (\varphi'(0))^2.$$

Теорема 1 (формулы обращения). *Справедливы следующие утверждения:*

1. *Для целочисленнозначной случайной величины X*

$$p_k \equiv \mathbf{P}(X = k) = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{-itk}\varphi(t)dt, \qquad k = 0, \pm1, \pm2, \ldots .$$

2. *Если характеристическая функция $\varphi(t)$ случайной величины X абсолютно интегрируема, то существует плотность распределения $f(x)$, определяемая формулой*

$$f(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx}\varphi(t)dt. \tag{13.5}$$

Доказательство. Для любого не равного нулю цело го числа m имеем

$$\int\limits_{-\pi}^{\pi} e^{itm}dt = \frac{1}{im}\left(e^{i\pi m} - e^{-i\pi m}\right) = 0.$$

Тогда можно написать следующее равенство, справедливое для всех целых m:

$$\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{itm}dt = \mathbb{I}_{\{0\}}(m) = \begin{cases} 0, & m \neq 0, \\ 1, & m = 0. \end{cases}$$

Используя это равенство и определение характеристической функции целочисленнозначной случайной величины, получим

$$\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{-itk}\varphi(t)dt = \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{-itk}\sum_{l=-\infty}^{\infty} e^{itl}p_l dt =$$

$$= \sum_{l=-\infty}^{\infty} p_l \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} e^{it(l-k)}dt = \sum_{l=-\infty}^{\infty} p_l \mathbb{I}_{\{0\}}(l-k) = p_k.$$

Первый пункт теоремы 1 доказан. Перейдем к доказательству второго пункта. Для этого нам нужно убедиться, что функция $f(x)$, определенная равенством (13.5), является плотностью функции распределения $F(x)$, для которой $\varphi(t)$ определяется равенством (13.4). Так как функция $\varphi(t)$ абсолютно интегрируема, то для произвольных $a < b$ в силу (13.5) имеем

$$\int_a^b f(x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(t)\int_a^b e^{-itx}dxdt = \lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c}\varphi(t)\int_a^b e^{-itx}dxdt. \tag{13.6}$$

Используя формулу (13.4), получим

$$\int_{-c}^{c}\varphi(t)\int_a^b e^{-itx}dxdt = \int_{-c}^{c}\int_{-\infty}^{\infty} e^{ity}dF(y)\int_a^b e^{-itx}dxdt =$$

$$= \int_{-\infty}^{\infty} dF(y)\int_a^b dx\frac{e^{ic(y-x)} - e^{ic(x-y)}}{i(y-x)} =$$

$$= 2\int_{-\infty}^{\infty} dF(y)\int_a^b dx\frac{\sin(c(x-y))}{(x-y)} = 2\int_{-\infty}^{\infty} dF(y)\int_{c(a-y)}^{c(b-y)}\frac{\sin v}{v}dv. \tag{13.7}$$

Известно, что

$$\int_0^{\infty}\frac{\sin v}{v}dv = \frac{\pi}{2}, \qquad \int_{-\infty}^{0}\frac{\sin v}{v}dv = \frac{\pi}{2}, \qquad \int_{-\infty}^{\infty}\frac{\sin v}{v}dv = \pi.$$

Поэтому, отдельно рассматривая случаи $0 \leqslant a - y < b - y$, $a - y < b - y \leqslant 0$ и $a - y < 0 < b - y$, получим

$$\lim_{c\to\infty}\int_{c(a-y)}^{c(b-y)}\frac{\sin v}{v}dv = \begin{cases} \pi, & y \in (a,b), \\ 0, & y \notin [a,b], \\ \frac{\pi}{2}, & y = a, \quad y = b. \end{cases}$$

Пусть a и b точки непрерывности функции $F(y)$. Тогда при $c \to \infty$ правая часть (13.7) стремится к

$$2\pi \int_a^b dF(y) = 2\pi(F(b) - F(a)).$$

Учитывая это и используя (13.6) и (13.7), имеем

$$\int_a^b f(x)dx = F(b) - F(a). \qquad (13.8)$$

Поскольку функция распределения непрерывна слева, то с помощью предельного перехода при $a_n \uparrow a, b_n \uparrow b$, где a_n и b_n – точки непрерывности функции $F(y)$, равенство (13.8) распространяется на произвольные точки a и b. Теперь из (13.8) по определению следует, что $f(x)$ является плотностью функции распределения $F(x)$.

Теорема 2 (о единственности). *Функция распределения однозначно определяется своей характеристической функцией.*

Справедливость этой теоремы в случае, когда функция распределения отвечает целочисленнозначной случайной величине или имеет плотность распределения, следует из формул обращения (теорема 1). В общем случае теорема 2 тоже следует из общей формулы обращения, которую мы не рассматриваем.

Задача 2. Найти закон распределения случайной величины X с характеристической функцией $\varphi(t) = \cos t$.

Решение. Характеристическая функция $\varphi(t) = \cos t$ не является абсолютно интегрируемой на всей прямой, поэтому естественно предположить, что X – дискретная случайная величина. Тогда характеристическая функция имеет вид

$$\varphi(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k = \sum_{k=1}^{\infty} \cos(tx_k)p_k + i \sum_{k=1}^{\infty} \sin(tx_k)p_k, \qquad (13.9)$$

где мы воспользовались **формулой Эйлера**

$$e^{iz} = \cos z + i \sin z.$$

Сравнивая выражение (13.9) с функцией $\cos t$ получаем, что величина X должна принимать лишь два значения $+1$ и -1 с равными вероятностями, т. е.

$$p_1 = \mathbf{P}(X = 1) = \frac{1}{2}, \qquad p_2 = \mathbf{P}(X = -1) = \frac{1}{2}.$$

Задачи

Задача 13.1. Написать производящую функцию величины X, равной числу выпавших очков на игральной кости.

Задача 13.2. Найти законы распределения, которым соответствуют следующие производящие функции:

$$1)\ \frac{1}{2}\left(1 - \frac{1}{2}z\right)^{-1}; \quad 2)\ e^{3\lambda(z-1)}; \quad 3)\ \left(\frac{1}{3} + \frac{2}{3}z\right)^n.$$

Задача 13.3. Найти законы распределения, которым соответствуют следующие характеристические функции:

$$1)\ \cos^2 t; \quad 2)\ \frac{1}{2}(\cos t + \cos 2t); \quad 3)\ \frac{2}{3} + \frac{1}{3}\cos 3t.$$

Задача 13.4. Пусть X – неотрицательная целочисленнозначная величина с производящей функцией $\psi(z)$. Выразить через $\psi(z)$ производящие функции величин $X + 1$ и $2X$.

Задача 13.5. Найти характеристическую функцию **треугольного** распределения с плотностью

$$p_\alpha(x) = \begin{cases} \alpha(1 - \alpha|x|), & \text{при } |x| \leqslant \alpha^{-1}, \\ 0, & \text{при } |x| > \alpha^{-1}, \end{cases} \qquad \alpha > 0.$$

Задача 13.6. Найти математическое ожидание и дисперсию случайной величины X, характеристическая функция которой равна одной из следующих функций

$$1)\ \frac{1}{at}\sin(at) \quad (a \neq 0); \quad 2)\ \frac{4}{t^2}\cos t \sin^2\left(\frac{t}{2}\right); \quad 3)\ \frac{4}{t^2}e^{it}\sin^2\left(\frac{t}{2}\right).$$

Задача 13.7. Вычислить характеристическую функцию случайной величины с плотностью

$$f(x) = \begin{cases} 2x, & \text{при } 0 \leqslant x \leqslant 1, \\ 0, & \text{при } x < 0, \quad 1 < x. \end{cases}$$

Задача 13.8. Найти моменты случайной величины X, характеристическая функция которой $\varphi(t) = \dfrac{1}{1+t^2}$.

Задача 13.9. Найти характеристическую функцию случайной величины X, плотность распределения которой имеет вид

$$f(x) = \frac{1}{2}e^{-|x|}.$$

Задача 13.10. Случайная величина X имеет плотность распределения

$$f(x) = 2h^2 x e^{-h^2 x^2}, \qquad \text{при } x \geqslant 0,$$

и $f(x) = 0$ при отрицательных x. Найти характеристическую функцию случайной величины X.

§ 14. Законы распределения случайных величин

1. *Дискретные распределения.*

1. **Биномиальное распределение.** Пусть X – случайная величина, равная числу успехов в n испытаниях Бернулли, и пусть вероятность успеха в каждом испытании равна p. Тогда величина X имеет **биномиальное распределение** с параметрами n и p:

$$\mathbf{P}(X = m) = C_n^m p^m (1-p)^{n-m}, \quad m = 0, 1, 2, \ldots, n.$$

Вычислим производящую функцию биномиального распределения. Учитывая, что $\mathbf{P}(X = m) = 0$ при $m > n$, и, используя фомулу бинома Ньютона (§ 1), получим

$$\psi(z) = \sum_{m=0}^{n} z^m \mathbf{P}(X = m) = \\ = \sum_{m=0}^{n} C_n^m (zp)^m (1-p)^{n-m} = (zp + 1 - p)^n.$$

Для вычисления математического ожидания и дисперсии воспользуемся следствием 1 § 13:

$$\mathbf{E}(X) = \psi'(z)|_{z=1} = np(zp + 1 - p)^{n-1}|_{z=1} = np,$$
$$\mathbf{D}(X) = \psi''(z)|_{z=1} + \psi'(1) - (\psi'(1))^2 = \\ = np(n-1)p(zp + 1 - p)^{n-2}|_{z=1} + np - (np)^2 = \\ = n(n-1)p^2 - n^2p^2 + np = np(1-p).$$

2. **Распределение Пуассона.** Это распределение является предельным для биномиального распределения, когда $n \to \infty$, а $\lambda = np$ остается постоянным (см. § 9). Говорят, что случайная величина X имеет **распределение Пуассона** с параметром $\lambda > 0$, если

$$\mathbf{P}(X = m) = \frac{\lambda^m}{m!}e^{-\lambda}, \quad m = 0, 1, 2, \ldots .$$

Производящая функция распределения Пуассона вычисляется следующим образом:

$$\psi(z) = \sum_{m=0}^{\infty} z^m \mathbf{P}(X = m) = e^{-\lambda} \sum_{m=0}^{\infty} \frac{(\lambda z)^m}{m!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}.$$

Применяя следствие 1 предыдущего параграфа, для математического ожидания и дисперсии получим выражения:

$$\mathbf{E}(X) = \psi'(z)|_{z=1} = \lambda, \ \mathbf{D}(X) = \psi''(z)|_{z=1} + \psi'(1) - (\psi'(1))^2 = \lambda.$$

Задача 1. Число космических частиц, попадающих в аппаратный отсек ракеты за время ее полета распределено по закону Пуассона с параметром λ. При этом условная вероятность для каждой из этих частиц попасть в уязвимый блок равна p. Найти закон распределения числа частиц, попадающих в уязвимый блок.

Решение. Обозначим A_k ={в уязвимый блок попало k частиц}, H_l ={в аппаратный отсек попало l частиц}. Тогда события H_l, $l = 0, 1, 2, \ldots$, составляют полную группу событий, и по формуле полной вероятности ($n = \infty$) (§ 7)

$$\mathbf{P}(A_k) = \sum_{l=0}^{\infty} \mathbf{P}(H_l)\mathbf{P}(A_k|H_l) = \sum_{l=k}^{\infty} \mathbf{P}(H_l)\mathbf{P}(A_k|H_l),$$

поскольку при $l < k$ очевидно $\mathbf{P}(A_k|H_l) = 0$. Согласно условию $\mathbf{P}(H_l) = \frac{\lambda^l}{l!}e^{-\lambda}$, $l = 0, 1, 2, \ldots$. Используя схему Бернулли (§ 8), несложно понять, что вероятности $\mathbf{P}(A_k|H_l)$ при $0 \leqslant k \leqslant l$ имеют биномиальное распределение с вероятностью "успеха" (частица попала в уязвимый блок) p. Поэтому $\mathbf{P}(A_k|H_l) = C_l^k p^k (1-p)^{l-k}$, $0 \leqslant k \leqslant l$. Окончательно имеем

$$\mathbf{P}(A_k) = \sum_{l=k}^{\infty} \frac{\lambda^l}{l!} e^{-\lambda} C_l^k p^k (1-p)^{l-k} = \sum_{l=k}^{\infty} \frac{\lambda^l e^{-\lambda}}{k!(l-k)!} p^k (1-p)^{l-k} =$$

$$= \frac{\lambda^k p^k e^{-\lambda}}{k!} \sum_{l=k}^{\infty} \frac{(\lambda(1-p))^{l-k}}{(l-k)!} = \frac{(\lambda p)^k e^{-\lambda}}{k!} e^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!} e^{-\lambda p}.$$

Таким образом, число частиц, попадающих в уязвимый блок, тоже распределено по закону Пуассона, но с параметром, равным произведению λp.

3. **Геометрическое распределение.** Проводится бесконечная последовательность независимых одинаковых испытаний, в каждом из которых событие A наступает с вероятностью $p = \mathbf{P}(A) > 0$. Пусть X – случайная величина, равная числу испытаний до момента первого наступления события A. Тогда

$$\mathbf{P}(X = k) = (1-p)^k p, \quad k = 0, 1, 2, \ldots .$$

Действительно, если A_l = {наступление события A в l-м испытании}, то $\{X = k\} = \overline{A}_1 \overline{A}_2 \ldots \overline{A}_k A_{k+1}$. Принимая во внимание независимость событий A_l, получим

$$\mathbf{P}(X = k) = \mathbf{P}(\overline{A}_1 \overline{A}_2 \ldots \overline{A}_k A_{k+1}) =$$
$$= \mathbf{P}(\overline{A}_1)\mathbf{P}(\overline{A}_2) \ldots \mathbf{P}(\overline{A}_k)\mathbf{P}(A_{k+1}) = (1-p)^k p.$$

Это распределение называется **геометрическим**, так как вероятности $\mathbf{P}(X = k)$ образуют геометрическую прогрессию.

Вероятность того, что событие A наступит не раньше момента m, задается формулой

$$\mathbf{P}(X \geqslant m) = \sum_{k=m}^{\infty} (1-p)^k p = (1-p)^m p \frac{1}{1-(1-p)} = (1-p)^m.$$

Производящая функция геометрического распределения определяется формулой

$$\psi(z) = \sum_{k=0}^{\infty} z^k (1-p)^k p = p \sum_{k=0}^{\infty} (z(1-p))^k = \frac{p}{1 - z(1-p)}.$$

Для математического ожидания и дисперсии имеем

$$\mathbf{E}(X) = \psi'(z)|_{z=1} = \frac{p(1-p)}{(1-z(1-p))^2}\bigg|_{z=1} = \frac{1-p}{p},$$
$$\mathbf{D}(X) = \psi''(z)|_{z=1} + \psi'(1) - (\psi'(1))^2 =$$
$$= \frac{2p(1-p)^2}{(1-z(1-p))^3}\bigg|_{z=1} + \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 =$$
$$= \frac{2(1-p)^2}{p^2} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}.$$

Пусть есть две независимые случайные величины X_1 и X_2, имеющие геометрические распределения с параметрами p_1 и p_2 соответственно. Тогда величина $X = \min(X_1, X_2)$, равная минимуму из этих величин, тоже имеет геометрическое распределение с параметром p, определяемым формулой $1 - p = (1 - p_1)(1 - p_2)$.

Действительно, обозначим $q = 1 - p$, $q_1 = 1 - p_1$, $q_2 = 1 - p_2$. Тогда, используя независимость случайных величин X_1 и X_2, получим

$$\begin{aligned}\mathbf{P}(X = m) &= \mathbf{P}(X_1 = m, X_2 \geqslant m) + \mathbf{P}(X_2 = m, X_1 > m) = \\ &= \mathbf{P}(X_1 = m)\mathbf{P}(X_2 \geqslant m) + \mathbf{P}(X_2 = m)\mathbf{P}(X_1 > m) = \\ &= q_1^m(1 - q_1)\sum_{k=m}^{\infty} q_2^k(1 - q_2) + q_2^m(1 - q_2)\sum_{k=m+1}^{\infty} q_1^k(1 - q_1) = \\ &= (1 - q_1)(1 - q_2)\left(\frac{q_1^m q_2^m}{1 - q_2} + \frac{q_2^m q_1^{m+1}}{1 - q_1}\right) = (q_1 q_2)^m(1 - q_1 q_2).\end{aligned}$$

В справедливости этого утверждения можно убедиться и не проводя вычислений, рассмотрев следующий пример.

Пример 1. Одновременно бросаются игральная кость (цифры 1, 2, 3, 4, 5, 6) и правильная пирамидка (цифры 1, 2, 3, 4). Нас интересует случайная величина X, равная числу бросаний до момента первого наступления события $A =$ {выпала хотя бы одна четверка}. Ясно, что X имеет геометрическое распределение с $\mathbf{P}(\overline{A}) = \frac{5}{6} \cdot \frac{3}{4}$. При этом, игральную кость и пирамидку можно рассматривать поотдельности. Обозначим X_1, число бросаний до момента первого наступления события $A_1 =$ {на кости выпала четверка}, X_2, число бросаний до момента первого наступления события $A_2 =$ {на пирамидке выпала четверка}. Ясно, что эти величины независимы и имеют геометрические распределения с $\mathbf{P}(\overline{A}_1) = \frac{5}{6}$ и $\mathbf{P}(\overline{A}_2) = \frac{3}{4}$ соответственно. Очевидно, также, что $X = \min(X_1, X_2)$.

2. Распределения, имеющие плотность.

1. **Равномерное распределение** для любых $a < b$ задается плотностью

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

Характеристическая функция равномерного распределения вычисляется следующим образом:

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = \frac{1}{b-a}\int_a^b e^{itx}dx = \frac{e^{itb}-e^{ita}}{it(b-a)}.$$

В § 12, пример 2, были вычислены математическое ожидание и дисперсия равномерного распределения: $\mathbf{E}(X) = \frac{b+a}{2}$, $\mathbf{D}(X) = \frac{(b-a)^2}{12}$.

2. **Показательное распределение** возникает как предельное для геометрического в следуюшей схеме. Рассмотрим уплотняющиеся при $\Delta \downarrow 0$ моменты времени $k\Delta$, $k = 0, 1, 2, \ldots$. В моменты времени $k\Delta$ осуществляются независимые одинаковые испытания, в каждом из которых событие A наступает с вероятностью $p_\Delta = \lambda\Delta$, где $\lambda > 0$ – постоянная величина. Обозначим X_Δ – время до момента первого наступления события A. Тогда величина $\frac{X_\Delta}{\Delta}$ равна числу испытаний до первого наступления события A и имеет геометрическое распределение. Поскольку любому фиксированному моменту времени x соответствует $m \approx \frac{x}{\Delta}$ испытаний, то согласно п. 3

$$\mathbf{P}(X_\Delta \geqslant x) = \mathbf{P}\left(\frac{X_\Delta}{\Delta} \geqslant \frac{x}{\Delta}\right) \approx \mathbf{P}\left(\frac{X_\Delta}{\Delta} \geqslant m\right) = \\ = (1-p_\Delta)^m \approx (1-\lambda\Delta)^{x/\Delta} \approx e^{-\lambda x}, \qquad \Delta \downarrow 0.$$

В последнем приближенном равенстве использован "замечательный предел" (см. § 9). Переходя к дополнительным вероятностям, получаем, что предельное распределение для X_Δ имеет вид

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Это – **показательное распределение**, имеющее плотность

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Для характеристической функции показательного распределения имеем выражение:

$$\varphi(t) = \lambda \int_0^\infty e^{itx} e^{-\lambda x} dx = \frac{\lambda}{\lambda - it}.$$

Математическое ожидание и дисперсию вычислим с помощью следствия 3 § 13:

$$\mathbf{E}(X) = -i\varphi'(0) = -i\frac{i\lambda}{(\lambda - it)^2}\Big|_{t=0} = \frac{1}{\lambda}.$$

$$\mathbf{D}(X) = -\varphi''(0) + (\varphi'(0))^2 = -\frac{2i^2\lambda}{(\lambda - it)^3}\Big|_{t=0} + \left(\frac{i}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Задача 2. Пусть время ожидания регистрации прибором солнечного нейтрино распределено по показательному закону с параметром λ. К моменту времени s частица не зарегистрирована. Какова вероятность того, что начиная с момента времени s ожидать регистрации придется не меньше чем время h?

Решение. Обозначим X – время ожидания регистрации с момента включения прибора. Поскольку величина X распределена по показательному закону, то при $t \geqslant 0$ выполняется равенство $\mathbf{P}(X \geqslant t) = e^{-\lambda t}$. Нас интересует условная вероятность $\mathbf{P}(X \geqslant s + h | X \geqslant s)$. По определению условной вероятности

$$\mathbf{P}(X \geqslant s + h | X \geqslant s) = \frac{\mathbf{P}(X \geqslant s + h)}{\mathbf{P}(X \geqslant s)} = \frac{e^{-\lambda(s+h)}}{e^{-\lambda s}} = e^{-\lambda h}.$$

Таким образом, если время ожидания события распределено по показательному закону, то информация о том, что событие не наступило к данному моменту не улучшает шансы на его наступление в дальнейшем.

Это свойство называется **отсутствием последействия**, и им обладает только показательное распределение.

3. **Нормальное распределение** задается плотностью

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \qquad -\infty < x < \infty,$$

где $\sigma > 0$ и $a \in (-\infty, \infty)$ – некоторые параметры.

Это распределение часто также называют **гауссовским распределением**.

Если $a = 0$ и $\sigma = 1$, то нормальное распределение с такими параметрами называется **стандартным**. Функция распределения в этом случае имеет вид

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt =$$

$$= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-t^2/2} dt = \frac{1}{2}\big(1 + \Phi(x)\big), \tag{14.1}$$

где $\Phi(x)$ – функция Лапласа, уже встречавшаяся нам в § 9.

Легко убедиться, что $\Phi(x)$ является при $x \in [0, \infty)$ функцией распределения величины $|X|$, где X – случайная величина со стандартным нормальным распределением. Действительно

$$\mathbf{P}(|X| < x) = \mathbf{P}(-x < X < x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-x}^{x} e^{-t^2/2} dt = \Phi(x). \tag{14.2}$$

В силу этого свойства $\Phi(x)$, $x \in (0, \infty)$, иногда называют функцией **отраженного нормального распределения**. Для отрицательных x функцию распределения следует полагать равной нулю.

Вычислим характеристическую функцию нормального распределения. Используя определение характеристической функции и делая замены переменных в интегралах, получим

$$\varphi(t) = \int\limits_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(x-a)^2}{2\sigma^2}\Big) dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{ita} \int\limits_{-\infty}^{\infty} \exp\Big(ity - \frac{y^2}{2\sigma^2}\Big) dy =$$

$$= \frac{e^{ita - \sigma^2 t^2/2}}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\infty} \exp\Big(-\Big(\frac{y}{\sqrt{2}\sigma} - \frac{it\sigma}{\sqrt{2}}\Big)^2\Big) dy.$$

Теперь, используя метод контурного интегрирования, можно последний интеграл преобразовать

$$\frac{1}{\sigma}\int_{-\infty}^{\infty}\exp\left(-\left(\frac{y}{\sqrt{2}\sigma}-\frac{it\sigma}{\sqrt{2}}\right)^2\right)dy=\int_{-\infty}^{\infty}e^{-v^2/2}dv=\sqrt{2\pi}.$$

Следовательно,

$$\varphi(t)=e^{ita-\sigma^2t^2/2}.$$

Для вычисления математического ожидания и дисперсии нормального распределения применим следствие 3 § 13:

$$\begin{aligned}\mathbf{E}(X)&=-i\varphi'(0)=-i(ia-\sigma^2t)e^{ita-\sigma^2t^2/2}|_{t=0}=-iia=a,\\ \mathbf{D}(X)&=-\varphi''(0)+(\varphi'(0))^2=\\ &=-(-\sigma^2+(ia-\sigma^2t)^2)e^{ita-\sigma^2t^2/2}|_{t=0}+(ia)^2=\\ &=\sigma^2+a^2-a^2=\sigma^2.\end{aligned}$$

Таким образом, параметры a и σ имеют для нормального распределения следующий смысл: a – математическое ожидание, σ – квадратный корень из дисперсии.

Для нормального распределения со средним ноль несложно с помощью леммы 2 § 13 вычислить все моменты. При $a=0$ характеристическую функцию нормального распределения можно представить в виде следующего ряда

$$\varphi(t)=e^{-\sigma^2t^2/2}=\sum_{l=0}^{\infty}(-1)^l\frac{\sigma^{2l}t^{2l}}{2^l l!}.$$

Поэтому

$$\begin{aligned}\frac{d^k}{dt^k}\varphi(t)&=\sum_{l=0}^{\infty}(-1)^l\frac{\sigma^{2l}}{2^l l!}\frac{d^k}{dt^k}t^{2l}=\\ &=\sum_{l\geqslant k/2}(-1)^l\frac{\sigma^{2l}}{2^l l!}2l(2l-1)\dots(2l-k+1)t^{2l-k},\end{aligned}$$

и, следовательно,

$$\left.\frac{d^k}{dt^k}\varphi(t)\right|_{t=0}=\begin{cases}(-1)^{k/2}\dfrac{\sigma^k}{2^{k/2}(k/2)!}k!, & \text{при } k \text{ – четном},\\ 0, & \text{при } k \text{ – нечетном}.\end{cases}$$

Согласно лемме 2 § 13 имеем, что все нечетные моменты равны нулю, а для четных справедлива формула

$$\mathbf{E}(X^{2l}) = \frac{\sigma^{2l}(2l)!}{2^l l!} = \sigma^{2l}(2l-1)!!,$$

где принято обозначение $(2l-1)!! = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2l-1)$.

Используя функцию Лапласа $\Phi(x)$, можно выразить вероятность того, что случайная величина X, распределенная по нормальному закону с параметрами a и σ, лежит в заданных пределах. Справедлива следующая формула

$$\mathbf{P}(x_1 \leqslant X < x_2) = \frac{1}{2}\left(\Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right)\right). \qquad (14.3)$$

Действительно, величина $\frac{X-a}{\sigma}$ имеет стандартное нормальное распределение, поскольку в силу формулы (10.6)

$$f_{\frac{X-a}{\sigma}}(y) = \frac{\sigma}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\sigma y + a - a)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}}e^{-y^2/2}.$$

Применяя формулы (10.1) и (14.1), получим

$$\mathbf{P}(x_1 \leqslant X < x_2) = \mathbf{P}\left(\frac{x_1 - a}{\sigma} \leqslant \frac{X-a}{\sigma} < \frac{x_2 - a}{\sigma}\right) =$$
$$= \mathcal{N}\left(\frac{x_2 - a}{\sigma}\right) - \mathcal{N}\left(\frac{x_1 - a}{\sigma}\right) = \frac{1}{2}\left(\Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right)\right).$$

Следствие 1 (правило “3σ”). *Для нормально распределенной случайной величины X со средним a и дисперсией σ^2*

$$\mathbf{P}(|X - a| < 3\sigma) = \Phi(3) \approx 0.9973.$$

Правило “3σ” означает, что с большой вероятностью, равной 0,9973, значения нормально распределенной случайной величины X лежат в интервале $(a - 3\sigma, a + 3\sigma)$.

Доказательство правила “3σ” основано на формуле (14.3). Применяя (14.3) с $x_1 = -3\sigma + a, x_2 = 3\sigma + a$, получим

$$\mathbf{P}(|X - a| < 3\sigma) = \mathbf{P}(-3\sigma + a < X < 3\sigma + a) =$$
$$= \frac{1}{2}\left(\Phi\left(\frac{3\sigma}{\sigma}\right) - \Phi\left(-\frac{3\sigma}{\sigma}\right)\right) = \frac{1}{2}\Big(\Phi(3) - \Phi(-3)\Big) = \Phi(3).$$

С нормальным распределением мы уже встречались в интегральной теореме Муавра–Лапласа (§ 9). Формула (9.8) при $a_n = \dfrac{m_1 - np}{\sqrt{npq}}, b_n = \dfrac{m_2 - np}{\sqrt{npq}}$ утверждает, что вероятность наступления события A не менее m_1 и не более m_2 раз при большом количестве испытаний ($n \to \infty$) приближенно равна вероятности того, что случайная величина X со стандартным нормальным распределением принимает значения в пределах от a_n до b_n.

4. **Двумерное нормальное распределение.** Пусть X и Y – независимые нормально распределенные случайные величины с параметрами a_1, σ_1 и a_2, σ_2 соответственно. Тогда по формуле (11.3) совместная плотность распределения равна произведению

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(x - a_1)^2}{2\sigma_1} - \frac{(y - a_2)^2}{2\sigma_2}\right).$$

В общем случае совместная плотность **двумерного нормального распределения** имеет вид

$$f(x, y) =$$
$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - r^2}} \exp\left(-\frac{(x - a_1)^2}{2(1 - r^2)\sigma_1^2} + \frac{r(x - a_1)(y - a_2)}{(1 - r^2)\sigma_1\sigma_2} - \frac{(y - a_2)^2}{2(1 - r^2)\sigma_2^2}\right).$$

Смысл параметров, входящих в эту формулу, следующий: $a_1 = \mathbf{E}(X), a_2 = \mathbf{E}(Y), \sigma_1^2 = \mathbf{D}(X), \sigma_2^2 = \mathbf{D}(Y), r = r(X, Y)$ – коэффициент корреляции.

Замечание 1. При $r = 0$ плотность двумерного нормального распределения величин (X, Y) распадается в произведение плотностей величин X и Y, а это означает, что величины X и Y независимы. Таким образом, свойство независимости компонент двумерных нормально распределенных величин полностью характеризуется условием равенства нулю коэффициента корреляции.

Задачи

Задача 14.1. Случайная величина X распределена по нормальному закону с математическим ожиданием $m = 40$ и дисперсией $\sigma^2 = 400$. Вычислить вероятность попадания случайной величины в интервал (30,80).

Задача 14.2. Две игральные кости бросают до выпадения числа 6 хотя бы на одной из них. Найти вероятность того, что впервые число 6 появится при k-ом бросании, $k = 1, 2,$

Задача 14.3. Пусть $X_1, X_2, \ldots, X_n$ – независимые случайные величины, распределенные равномерно на отрезке $[0, 1]$. Найти функции распределения случайных величин $Y = \max\limits_{1 \leqslant k \leqslant n} X_k$ и $Z = \min\limits_{1 \leqslant k \leqslant n} X_k$.

Задача 14.4. Случайная величина X распределена равномерно с математическим ожиданием $\mathbf{E}(X) = 4$ и дисперсией $\mathbf{D}(X) = 3$. Найти плотность распределения величины X.

Задача 14.5. Однородная проволока длиной 1 метр растягивается за концы и разрывается. Случайная величина X равна длине большего куска проволоки. Вычислить функцию распределения величины X.

Задача 14.6. Пусть X_1, X_2 – независимые случайные величины, имеющие геометрическое распределение: $\mathbf{P}(X_l = k) = q^k(1-q), k = 0, 1, 2, \ldots,\ l = 1, 2$. Найти распределение величины $Z = \max(X_1, X_2)$.

Задача 14.7. Изделия испытываются на перегрузку. Число изделий не ограничено. Для каждого изделия вероятность выдержать перегрузку равна 0.8. После того, как первое изделие не выдержит перегрузку, испытания прекращаются. Найти закон распределения величины X равной числу испытанных изделий.

Задача 14.8. Случайная величина X распределена по нормальному закону. Известно, что с вероятностью 0.15 она принимает значения меньше 1.06, а с вероятностью 0.1 больше 3.38. Определить математическое ожидание и дисперсию величины X.

§ 15. Распределения сумм независимых случайных величин. Свертки распределений

Рассмотрим сначала две независимые целочисленнозначные неотрицательные случайные величичы X и Y с распределениями $p_k = \mathbf{P}(X = k), k = 0, 1, 2, \ldots,$ и $q_l = \mathbf{P}(Y = l), l = 0, 1, 2, \ldots$. В силу того, что величины независимы, событие $\{X = k, Y = l\}$ имеет вероятность $p_k q_l$. Сумма $S = X + Y$ есть случайная величина, и событие $\{S = r\}$ равно сумме событий $\{X = k, Y = r - k\}, k = 0, 1, \ldots, r$. Поскольку слагаемые несовместны, то вероятность $u_r = \mathbf{P}(S = r)$ равна сумме вероятностей событий и, следовательно, задается формулой

$$u_r = \sum_{k=0}^{r} p_k q_{r-k}, \quad r = 0, 1, 2, \ldots . \tag{15.1}$$

Распределение $U = \{u_r\}_{r=0}^{\infty}$, определенное этими равенствами по распределениям $P = \{p_k\}_{k=0}^{\infty}$ и $Q = \{q_l\}_{l=0}^{\infty}$, называется **сверткой распределений** P и Q и обозначается $U = P * Q$.

Свертка $U = P * Q$ является распределением суммы $S = X + Y$ независимых величин X и Y с распределениями P и Q соответственно.

Теорема 1. *Пусть распределениям* $P = \{p_k\}_{k=0}^{\infty}$, $Q = \{q_l\}_{l=0}^{\infty}$, $U = \{u_r\}_{r=0}^{\infty}$ *отвечают производящие функции*

$$\psi_P(z) = \sum_{k=0}^{\infty} z^k p_k, \quad \psi_Q(z) = \sum_{l=0}^{\infty} z^l q_l, \quad \psi_U(z) = \sum_{r=0}^{\infty} z^r u_r.$$

Тогда, если $U = P * Q$, *то* $\psi_U(z) = \psi_P(z)\psi_Q(z)$ *и наоборот, если* $\psi_U(z) = \psi_P(z)\psi_Q(z)$, *то* $U = P * Q$.

Доказательство. Перемножив функции $\psi_P(z), \psi_Q(z)$ так, чтобы в результате получился ряд по степеням z, получим

$$\psi_P(z)\psi_Q(z) = \sum_{r=0}^{\infty} \Bigl(\sum_{k=0}^{r} p_k q_{r-k}\Bigr) z^r. \tag{15.2}$$

Из этого равенства следует утверждение теоремы. Действительно, из (15.2) и (15.1) следует, что произведение функций ψ_P и ψ_Q является производящей функцией свертки распределений U, и наоборот, производящая функция свертки распределений распадается в произведение производящих функций компонент.

Пример 1. Пусть независимые величины X и Y имеют распределения Пуассона с параметрами λ_1 и λ_2 соответственно. Тогда сумма $S = X + Y$ снова распределена по закону Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$. Действительно,

$\psi_X(z) = e^{\lambda_1(z-1)}$, $\psi_Y(z) = e^{\lambda_2(z-1)}$, а тогда, согласно теореме 1, $\psi_S(z) = e^{(\lambda_1+\lambda_2)(z-1)}$. Это снова производящая функция распределения Пуассона, но с параметром $\lambda = \lambda_1 + \lambda_2$.

Пусть имеются независимые целочисленнозначные неотрицательные величины $X_l, l = 1, 2, \ldots, n$, с распределениями $P_l = \{p_k^{(l)}\}_{k=0}^{\infty}$. Каким будет распределение суммы $S_n = \sum\limits_{l=1}^{n} X_l$? Составляя сумму последовательно, т. е. каждый раз прибавляя по одному слагаемому, несложно получить, что распределение величины S_n будет иметь вид $P_1 * P_2 * \cdots * P_n$.

Свертку нескольких распределений определяем рекуррентным способом: сначала сворачиваем первые два распределения, затем результат сворачиваем с третьим и т. д. Из теоремы 1 следует, что свертка распределений обладает свойствами коммутативности $P_1 * P_2 = P_2 * P_1$ и ассоциативности $(P_1 * P_2) * P_3 = P_1 * (P_2 * P_3)$, поэтому в какой последовательности сворачивать распределения, безразлично.

Задача 1. Бросаются независимо друг от друга две правильные пирамидки (на гранях цифры $1, 2, 3, 4$). Найти распределение суммы цифр, выпавших на обеих пирамидках. Вычислить функцию распределения этой суммы.

Решение. Возможные значения суммы выпавших цифр составляют множество $\{2,3, \ldots ,8\}$. Обозначим p_k вероятность того, что величина X, равная сумме цифр, принимает значение k. Прямым подсчетом можно проверить, что для $k = 2, 3, \ldots, 8$ выполняется $p_k = (4 - |5 - k|)(1/4)^2$. Так, например, $p_2 = (1/4)^2$, а $p_6 = 3(1/4)^2$, поскольку сумму 2 дает только один набор $(1, 1)$, а сумму 6 составляют 3 таких набора: (2,4), (3,3), (4,2). Тогда функция распределения равна

$$\begin{array}{ll}
0, & \text{при } x \leqslant 2,\\
p_2 = 1/16, & \text{при } 2 < x \leqslant 3,\\
p_2 + p_3 = 3/16, & \text{при } 3 < x \leqslant 4,\\
p_2 + p_3 + p_4 = 6/16, & \text{при } 4 < x \leqslant 5,\\
p_2 + p_3 + p_4 + p_5 = 10/16, & \text{при } 5 < x \leqslant 6,\\
p_2 + p_3 + p_4 + p_5 + p_6 = 13/16, & \text{при } 6 < x \leqslant 7,\\
p_2 + p_3 + p_4 + p_5 + p_6 + p_7 = 15/16, & \text{при } 7 < x \leqslant 8,\\
p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 = 1, & \text{при } 8 < x < \infty.
\end{array}$$

Рассмотрим теперь независимые случайные величины X и Y, имеющие плотности распределения $f(x)$ и $g(y)$ соответственно. Согласно (11.2) и (11.3) для любых $a < b$ и $c < d$ вероятность того, что случайный вектор (X, Y) принадлежит прямоугольнику $[a, b) \times [c, d)$, равна

$$\mathbf{P}(a \leqslant X < b, c \leqslant Y < d) = \int_a^b \int_c^d f(x)g(y)dxdy.$$

Поскольку любое достаточно хорошее множество A на плоскости можно представить в виде суммы непересекающихся прямоугольников, то

$$\mathbf{P}((X, Y) \in A) = \iint_A f(x)g(y)dxdy.$$

Пусть $S = X + Y$. Обозначим $A = \{(x, y) : x + y < s\}$. Тогда $\{S < s\} = \{(X, Y) \in A\}$ и

$$\mathbf{P}(X + Y < s) = \iint_{x+y<s} f(x)g(y)dxdy = \int_{-\infty}^{s} \left(\int_{-\infty}^{\infty} f(x)g(v - x)dx \right) dv.$$

В силу определения плотности из этого представления следует, что функция

$$h(v) = \int_{-\infty}^{\infty} f(x)g(v - x)dx \tag{15.3}$$

является плотностью распределения суммы $X + Y$. Определенную этой формулой функцию h называют **сверткой функций** f и g и обозначают $h = f * g$. Таким образом, плотность распределения суммы двух независимых случайных величин, имеющих плотности распределения, равна свертке плотностей распределений слагаемых.

Если величины X и Y неотрицательны, т. е. $f(x) = 0$ при $x < 0$, и $g(y) = 0$ при $y < 0$, то формула свертки (15.3) превращается в следующую: при $v \geqslant 0$

$$h(v) = \int_0^v f(x)g(v - x)dx, \tag{15.4}$$

а при $v < 0$ имеем $h(v) = 0$. Равенство нулю плотности при отрицательных значениях аргумента отражает тот факт, что сумма двух неотрицательных величин неотрицательна.

Теорема 2. *Пусть f, g и h – плотности распределения и*

$$\varphi_f(t) = \int\limits_{-\infty}^{\infty} e^{itx} f(x)dx, \qquad \varphi_g(t) = \int\limits_{-\infty}^{\infty} e^{itx} g(x)dx,$$

$$\varphi_h(t) = \int\limits_{-\infty}^{\infty} e^{itx} h(x)dx$$

– соответствующие им характеристические функции. Тогда, если $h = f * g$, то $\varphi_h(t) = \varphi_f(t)\varphi_g(t)$, и наоборот, если $\varphi_h(t) = \varphi_f(t)\varphi_g(t)$, то $h = f * g$.

Доказательство. С помощью замены переменной интегрирования легко убедиться, что произведение характеристических функций $\varphi_f(t)$ и $\varphi_g(t)$ можно представить в следующем виде:

$$\varphi_f(t)\varphi_g(t) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{itx} f(x) e^{ity} g(y)dxdy =$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{itx} f(x) e^{it(v-x)} g(v-x)dxdv =$$

$$= \int\limits_{-\infty}^{\infty} e^{itv} \left(\int\limits_{-\infty}^{\infty} f(x)g(v-x)dx \right) dv.$$

Из этого равенства следует, что произведение $\varphi_f(t)\varphi_g(t)$ является характеристической функцией свертки $h = f * g$, и наоборот, характеристическая функция свертки плотностей распределения распадается в произведение характеристических функций компонент.

Свертка нескольких плотностей распределения определяется рекуррентным способом: сначала сворачиваются первые две плотности, затем результат сворачивается с третьей и т. д. Из теоремы 2 следует, что свертка плотностей распределений обладает свойством коммутативности

$f_1 * f_2 = f_2 * f_1$ и ассоциативности $(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3)$, поэтому не имеет значения, в каком порядке сворачивать плотности распределения. Ясно, что плотность распределения суммы $S_n = \sum_{l=1}^{n} X_l$, где X_l – независимые случайные величины с плотностями распределения $f_l(x)$, равна свертке плотностей слагаемых, т. е. $h(s) = f_1 * f_2 * \cdots * f_n(s)$.

Пример 2. Пусть $S_n = X_1 + X_2 + \cdots + X_n$, где X_l – независимые случайные величины, распределенные по показательному закону с параметром λ, т. е. X_l имеют общую плотность

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Тогда плотность распределения суммы S_n имеет вид

$$g_n(x) = \begin{cases} \dfrac{\lambda(\lambda s)^{n-1}}{(n-1)!} e^{-\lambda s}, & \text{при } s \geqslant 0, \\ 0, & \text{при } s < 0. \end{cases}$$

Убедимся в справедливости этой формулы с помощью индукции. Для $n = 1$ эта формула верна, поскольку $g_1(s) = f(s)$. Предположим, что она верна для $n = k$ и докажем ее для $n = k + 1$. Поскольку величины неотрицательны, то применяя равенство (15.4), имеем

$$g_{k+1}(s) = g_k * f(s) = \int_0^s \lambda e^{-\lambda(s-x)} \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} dx =$$

$$= \frac{\lambda^{k+1} e^{-\lambda s}}{(k-1)!} \int_0^s x^{k-1} dx = \frac{\lambda(\lambda s)^k}{k!} e^{-\lambda s}.$$

Задача 2. Найти плотность распределения суммы двух независимых случайных величин $Z = X + Y$, где величина X равномерно распределена в интервале $[a, b]$, а Y имеет нормальное распределение с параметрами m и σ^2.

Решение. Обозначим $h(z)$ – плотность распределения суммы Z. По формуле (15.3)

$$h(z) = \int_a^b \frac{1}{b-a} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{((z-x)-m)^2}{2\sigma^2}\right) dx,$$

или после замены переменных

$$h(z) = \frac{1}{b-a}\frac{1}{\sqrt{2\pi}} \int\limits_{(a-(z-m))/\sigma}^{(b-(z-m))/\sigma} e^{-t^2/2}dt.$$

Используя определение функции Лапласа § 9, окончательно получим

$$h(z) = \frac{1}{2(b-a)}\left(\Phi\left(\frac{b-(z-m)}{\sigma}\right) - \Phi\left(\frac{a-(z-m)}{\sigma}\right)\right).$$

Задачи

Задача 15.1. Брошены две игральные кости. Какова вероятность выпадения на двух костях в сумме 6 очков?

Задача 15.2. Брошены игральная кость и правильная пирамидка (цифры 1, 2, 3, 4). Какова вероятность выпадения в сумме 8 очков?

Задача 15.3. Смешаны две группы деталей, содержащих n_1 и n_2 деталей каждая. Число бракованных деталей в каждой группе X и Y имеют биномиальные распределения

$$P(X = m) = C_{n_1}^m p^m (1-p)^{n_1-m}, \quad m = 0, 1, 2, \dots, n_1,$$

$$P(Y = m) = C_{n_2}^m p^m (1-p)^{n_2-m}, \quad m = 0, 1, 2, \dots, n_2,$$

соответственно. Найти закон распределения числа бракованных деталей $(X+Y)$ в смешанной группе.

Задача 15.4. Пусть X_1 распределена с показательной плотностью

$$f_1(x) = \begin{cases} 4e^{-4x}, & 0 \leqslant x, \\ 0, & x < 0, \end{cases}$$

а X_2 не зависит от X_1 и распределена с плотностью

$$f_2(x) = \begin{cases} \frac{1}{2}, & 2 \leqslant x \leqslant 4, \\ 0, & x < 2, \quad 4 < x. \end{cases}$$

Найти плотность распределения суммы $Z = X_1 + X_2$.

Задача 15.5. Случайная величина X равномерно распределена в интервале [0,2], случайная величина Y равномерно распределена в интервале $[-1,1]$, X и Y независимы. Найти плотность распределения величины $Z = X + Y$.

Задача 15.6. Пусть случайная величина X имеет показательную плотность с параметром α, а величина Y распределена равномерно на интервале $[0,b]$ и не зависит от X. Найти плотности распределения величин $X+Y$ и $X-Y$.

Задача 15.7. Вычислить плотность распределения суммы $Z = X_1 + X_2$, в которой X_1 распределена с плотностью

$$f_1(x) = \begin{cases} \frac{x}{6}, & 2 \leqslant x \leqslant 4, \\ 0, & x < 2, \quad 4 < x, \end{cases}$$

а X_2 не зависит от X_1 и распределена с плотностью

$$f_2(x) = \begin{cases} \frac{1}{2}, & 1 \leqslant x \leqslant 3, \\ 0, & x < 1, \quad 3 < x. \end{cases}$$

§ 16. Неравенство Чебышева. Закон больших чисел

В теории вероятностей часто оценивают вероятности событий, порожденных случайной величиной, через моменты случайной величины.

Лемма 1. *Пусть X – неотрицательная случайная величина, т. е. $X \geqslant 0$. Тогда для любого $\varepsilon > 0$*

$$\mathbf{P}(X \geqslant \varepsilon) \leqslant \frac{\mathbf{E}(X)}{\varepsilon}.$$

Доказательство. Очевидно, что для неотрицательной случайной величины X справедливо равенство

$$\mathbb{1}_{[0,\varepsilon)}(X) + \mathbb{1}_{[\varepsilon,\infty)}(X) = 1,$$

где $\mathbb{1}_A(x)$ – индикатор множества A.

Согласно четвертому свойству математического ожидания (§ 12) из очевидного неравенства $X\mathbb{1}_{[\varepsilon,\infty)}(X) \geqslant \varepsilon\mathbb{1}_{[\varepsilon,\infty)}(X)$ следует, что $\mathbf{E}\big(X\mathbb{1}_{[\varepsilon,\infty)}(X)\big) \geqslant \varepsilon\mathbf{E}\big(\mathbb{1}_{[\varepsilon,\infty)}(X)\big)$. Используя эту оценку и тот факт, что математическое ожидание неотрицательной случайной величины неотрицательно, получим

$$\begin{aligned} \mathbf{E}(X) &= \mathbf{E}\Big(X\big(\mathbb{1}_{[0,\varepsilon)}(X) + \mathbb{1}_{[\varepsilon,\infty)}(X)\big)\Big) = \\ &= \mathbf{E}\big(X\mathbb{1}_{[0,\varepsilon)}(X)\big) + \mathbf{E}\big(X\mathbb{1}_{[\varepsilon,\infty)}(X)\big) \geqslant \\ &\geqslant \mathbf{E}\big(X\mathbb{1}_{[\varepsilon,\infty)}(X)\big) \geqslant \varepsilon\mathbf{E}\big(\mathbb{1}_{[\varepsilon,\infty)}(X)\big) = \varepsilon\mathbf{P}(X \geqslant \varepsilon). \end{aligned}$$

Поделив правую и левую части этого неравенства на ε, получим требуемую оценку.

Лемма 2 (неравенство Чебышева). *Для любой случайной величины Y с конечным вторым моментом и любого $\delta > 0$*

$$\mathbf{P}(|Y - \mathbf{E}(Y)| \geqslant \delta) \leqslant \frac{\mathbf{D}(Y)}{\delta^2}. \tag{16.1}$$

Доказательство. В силу неравенства (12.3) из конечности второго момента следует, что математическое ожидание и дисперсия существуют. Применим лемму 1 для случайной величины $X = (Y - \mathbf{E}(Y))^2$ и $\varepsilon = \delta^2$. Тогда имеем

$$\mathbf{P}(|Y - \mathbf{E}(Y)| \geqslant \delta) = \mathbf{P}((Y - \mathbf{E}(Y))^2 \geqslant \delta^2) \leqslant \leqslant \frac{\mathbf{E}(Y - \mathbf{E}(Y))^2}{\delta^2} = \frac{\mathbf{D}(Y)}{\delta^2}.$$

Следствие 1. *Для любой случайной величины Y с конечным вторым моментом и любого $\delta > 0$*

$$\mathbf{P}(|Y - \mathbf{E}(Y)| < \delta) \geqslant 1 - \frac{\mathbf{D}(Y)}{\delta^2}. \tag{16.2}$$

Поскольку сумма вероятности события и вероятности его дополнения равна единице, то оценка (16.2) непосредственно вытекает из (16.1).

Следствие 2. *Пусть $X_1, X_2, \ldots, X_n$ – независимые случайные величины с конечными вторыми моментами. Тогда для любого $\delta > 0$*

$$\mathbf{P}\left(\left|\sum_{l=1}^{n} X_l - \sum_{l=1}^{n} \mathbf{E}(X_l)\right| \geqslant \delta\right) \leqslant \frac{1}{\delta^2}\sum_{l=1}^{n} \mathbf{D}(X_l), \tag{16.3}$$

$$\mathbf{P}\left(\left|\sum_{l=1}^{n} X_l - \sum_{l=1}^{n} \mathbf{E}(X_l)\right| < \delta\right) \geqslant 1 - \frac{1}{\delta^2}\sum_{l=1}^{n} \mathbf{D}(X_l). \tag{16.4}$$

Доказательство. Положим $Y = \sum\limits_{l=1}^{n} X_l$. В силу аддитивности математического ожидания $\mathbf{E}(Y) = \sum\limits_{l=1}^{n} \mathbf{E}(X_l)$. По следствию 3 § 12 имеем $\mathbf{D}(Y) = \sum\limits_{l=1}^{n} \mathbf{D}(X_l)$. Подставляя эти выражения в (16.1) и (16.2), получаем (16.3) и (16.4) соответственно

Задача 1. Изнашивание орудия при стрельбе ведет к тому, что каждый выстрел уменьшает вероятность попадания в цель на 1 %. При первом выстреле вероятность попадания равна 0.8. Производится 100 выстрелов. Найти границы, в которых с вероятностью не меньшей 0.85 будет заключено число попаданий.

Решение. Поскольку каждый выстрел уменьшает вероятность попадания на 1 %, то при втором выстреле она равна $p_2 = 0.8 \cdot 0.99$, а при l-м равна $p_l = 0.8 \cdot (0.99)^{l-1}$. Пусть величина X_l принимает значение 1 в случае попадания при l-м выстреле и 0 в случае промаха. Сумма $S_n = \sum\limits_{l=1}^{n} X_l$ будет числом попаданий при n выстрелах. Поскольку

$$\mathbf{E}(X_l) = 1 \cdot p_l = 0.8(0.99)^{l-1},$$
$$\mathbf{D}(X_l) = \mathbf{E}(X_l^2) - \mathbf{E}^2(X_l) = 0.8(0.99)^{l-1} - (0.8(0.99)^{l-1})^2,$$

то, используя известную формулу для суммы членов геометрической прогрессии, получим

$$\sum_{l=1}^{n} \mathbf{E}(X_l) = 0.8 \cdot \frac{1-(0.99)^n}{0.01}\bigg|_{n=100} \approx 50.72,$$
$$\sum_{l=1}^{n} \mathbf{D}(X_l) = 0.8 \cdot \frac{1-(0.99)^n}{0.01} - (0.8)^2 \cdot \frac{1-(0.99)^{2n}}{1-(0.99)^2} =$$
$$= 0.8\frac{1-(0.99)^n}{0.01}\left(1 - 0.8\frac{1+(0.99)^n}{1.99}\right)\bigg|_{n=100} \approx 50.72 \cdot 0.45 \approx 22.87.$$

При $n = 100$ из (16.4) следует, что

$$\mathbf{P}\left(\sum_{l=1}^{100} \mathbf{E}(X_l) - \delta < S_{100} < \sum_{l=1}^{100} \mathbf{E}(X_l) + \delta\right) \geqslant$$
$$\geqslant 1 - \frac{1}{\delta^2}\sum_{l=1}^{100} \mathbf{D}(X_l) = 0.85,$$

где δ выбирается так, чтобы выполнялось последнее равенство, т. е. $\delta = (\sum\limits_{l=1}^{100} \mathbf{D}(X_l)/0.15)^{1/2} \approx 12.35$. Таким образом,

$\sum\limits_{l=1}^{100} \mathbf{E}(X_l) - \delta \approx 38.37$, $\quad \sum\limits_{l=1}^{100} \mathbf{E}(X_l) + \delta \approx 63.07$. Следовательно, с гарантированной вероятностью 0.85 число попаданий при ста выстрелах будет заключено в пределах от 38 до 63.

Рассмотрим различные виды сходимости последовательностей случайных величин.

Определение 1. Говорят, что последовательность случайных величин X_n сходится при $n \to \infty$ к величине X **в среднем квадратичном**, если $\mathbf{E}(X_n - X)^2 \to 0$.

Определение 2. Говорят, что последовательность случайных величин X_n сходится при $n \to \infty$ к величине X **по вероятности**, если $\mathbf{P}(|X_n - X| \geqslant \varepsilon) \to 0$ для любого $\varepsilon > 0$.

Определение 3. Говорят что последовательность случайных величин X_n сходится при $n \to \infty$ к величине X **с вероятностью единица**, если вероятность множества всех исходов из Ω, для которых имеет место сходимость числовых последовательностей $X_n(\omega) \to X(\omega)$, имеет единичную вероятность, т. е. $\mathbf{P}(\omega : X_n(\omega) \to X(\omega)) = 1$.

Следующий фундаментальный результат теории вероятностей имеет название "закон больших чисел". Пусть проводится большое количество независимых одинаковых экспериментов, в каждом из которых наблюдается случайная величина одной и той же природы. С математической точки зрения это означает, что наблюдается последовательность независимых случайных величин с одинаковыми распределениями. Предполагается, что существует математическое ожидание отдельно взятой величины. Обозначим его m. Заметим, что математические ожидания у одинаково распределенных случайных величин одинаковы. Тогда среднее арифметическое всех значений случайных величин, полученных в результате экспериментов, приближается при возрастании числа экспериментов к неслучайному числу m.

Теорема 1 (закон больших чисел). *Пусть $\{X_l\}_{l=1}^{\infty}$ – последовательность независимых одинаково распределенных случайных величин с математическим ожиданием $m = \mathbf{E}(X_l)$ и конечными дисперсиями $\mathbf{D}(X_l) < \infty$. Тогда величина $\overline{X}_n = \frac{1}{n}\sum\limits_{l=1}^{n} X_l$, равная среднему арифметическому первых n величин из последовательности $\{X_l\}_{l=1}^{\infty}$, сходится при $n \to \infty$*

в среднем квадратичном, по вероятности и с вероятностью единица к математическому ожиданию m.

Доказательство. Мы докажем лишь сходимость в среднем квадратичном и по вероятности. Утверждение о сходимости с вероятностью единица оставим без доказательства, поскольку доказательство такой сходимости потребовало бы значительных дополнительных сведений из теории вероятностей.

Докажем сходимость в среднем квадратичном. Для этого вычислим квадратичное отклонение $\mathbf{E}(\overline{X}_n - m)^2$. Используя свойства линейности и аддитивности математического ожидания, несложно убедиться в том, что

$$\mathbf{E}(\overline{X}_n) = \mathbf{E}\Big(\frac{1}{n}\sum_{l=1}^{n} X_l\Big) = \frac{1}{n}\sum_{l=1}^{n}\mathbf{E}(X_l) = \frac{1}{n}\cdot nm = m.$$

Тогда по определению дисперсии $\mathbf{E}(\overline{X}_n - m)^2 = \mathbf{D}(\overline{X}_n)$. Для вычисления дисперсии воспользуемся свойством 2) и следствием 3 § 12. В результате получим

$$\mathbf{E}(\overline{X}_n - m)^2 = \mathbf{D}(\overline{X}_n) = \mathbf{D}\Big(\frac{1}{n}\sum_{l=1}^{n} X_l\Big) =$$
$$= \frac{1}{n^2}\mathbf{D}\Big(\sum_{l=1}^{n} X_l\Big) = \frac{1}{n^2}\sum_{l=1}^{n}\mathbf{D}(X_l) = \frac{1}{n}\mathbf{D}(X_1). \tag{16.5}$$

В последнем равенстве мы воспользовались тем, что случайные величины X_l одинаково распределены и, следовательно, имеют одинаковые дисперсии. Правая часть в (16.5) стремится к нулю при $n \to \infty$, а это и означает, что $\overline{X}_n \to m$ в среднем квадратичном.

Докажем сходимость по вероятности. Воспользуемся оценкой (16.3) с $\delta = n\varepsilon$. Для любого $\varepsilon > 0$ по неравенству Чебышева имеем

$$\mathbf{P}(|\overline{X}_n - m| \geqslant \varepsilon) = \mathbf{P}\Big(\Big|\sum_{l=1}^{n} X_l - \sum_{l=1}^{n}\mathbf{E}(X_l)\Big| \geqslant n\varepsilon\Big) \leqslant$$
$$\leqslant \frac{\sum_{l=1}^{n}\mathbf{D}(X_l)}{n^2\varepsilon^2} = \frac{\mathbf{D}(X_1)}{n\varepsilon^2}.$$

Правая часть этого соотношении стремится к нулю при $n \to \infty$ для любого фиксированного ε. Согласно определению это означает, что $\overline{X}_n \to m$ по вероятности.

Пример 1. Пусть n раз бросается симметричная однородная монета. Выпадению герба будем приписывать значение 1, а выпадению решетки – 0. Иными словами, проводятся n случайных экспериментов, в каждом из которых наблюдается случайная величина X_l, $l = 1, 2, \ldots, n$, со значениями 1 и 0 в зависимости от выпадения герба или решетки, где l – номер бросания. Величины X_l независимы, поскольку бросания никак между собой не связаны. Поскольку монета симметричная, то вероятности выпадения герба и решетки совпадают и равны 1/2. Тогда $m = \mathbf{E}(X_l) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$, $\mathbf{D}(X_l) = \left(1 - \frac{1}{2}\right)^2 \frac{1}{2} + \left(0 - \frac{1}{2}\right)^2 \frac{1}{2} = \frac{1}{4}$. Величина $\overline{X}_n = \frac{1}{n} \sum\limits_{l=1}^{n} X_l$ выражает долю выпавших гербов при n бросаниях монеты. Согласно закону больших чисел $\overline{X}_n$ стремится к 1/2, т. е. к вероятности выпадения герба. Если бы монета была несимметричной, т. е. вероятность выпадения герба была бы равна p, $p \neq 1/2$, то $m = \mathbf{E}(X_l) = 1 \cdot p + 0 \cdot (1 - p) = p$, и по закону больших чисел доля выпадений герба будет стремиться к p. Таким образом, бросая монету большое число раз и вычисляя величину $\overline{X}_n$, можно с большой точностью вычислить p, т. е. установить, симметричная монета или нет.

В более общей ситуации, если проводятся n независимых одинаковых случайных экспериментов, в каждом из которых может наступить или не наступить событие A, то частота появления этого события, в силу закона больших чисел, будет стремиться к $\mathbf{P}(A)$. Действительно, как и в примере 1, положим величину X_l, равной 1, если A наступило в l-м эксперименте, и равной 0, если не наступило. Величины X_l, $l = 1, 2, \ldots, n$, независимы, так как независимы эксперименты. Тогда $\overline{X}_n = \frac{1}{n} \sum\limits_{l=1}^{n} X_l$ – частота появления события A. Очевидно $\mathbf{E}(X_l) = \mathbf{P}(A)$, $\mathbf{D}(X_l) = \mathbf{P}(A) - \mathbf{P}^2(A) = \mathbf{P}(A)\mathbf{P}(\bar{A})$. Можно применить закон больших чисел, из которого следует, что $\overline{X}_n \to \mathbf{P}(A)$.

Задача 2. На перрон станции метро каждые 5 минут приходит случайное число пассажиров, распределенное по закону Пуассона с параметром $\lambda = 250$. За то же время с перрона отправляются проходящие поезда, которые могут увезти количество пассажиров, имеющее равномерное распределение в промежутке [195,205]. Можно ли рассчитывать, что перрон не переполнится, если такой режим поддерживается постоянно?

Решение. Пусть X_k – число пассажиров, пришедших на перрон в течение k-го пятиминутного отрезка (отсчет можно вести от произвольного момента), а Y_k – число свободных мест в проходивших за эти 5 минут поездах. Заметим, что $\mathbf{E}(X_k) = \lambda = 250$ и $\mathbf{E}(Y_k) = (195 + 205)/2 = 200$. Пусть N_n – число пассажиров, добавившихся за n пятиминуток к имевшимся на перроне, а c – произвольное положительное число. Поскольку $N_n \geqslant \sum\limits_{k=1}^{n}(X_k - Y_k)$ и $\mathbf{E}(X_k) - \mathbf{E}(Y_k) = 50$, то при $n > c/50$ справедливо

$$\mathbf{P}(N_n < c) \leqslant \mathbf{P}\Big(\sum_{k=1}^{n}(X_k - Y_k) < c\Big) =$$

$$= \mathbf{P}\Big(\sum_{k=1}^{n}(X_k - Y_k - \mathbf{E}(X_k) + \mathbf{E}(Y_k)) < c - 50n\Big) \leqslant$$

$$\leqslant \mathbf{P}\Big(\Big|\sum_{k=1}^{n}(X_k - Y_k - \mathbf{E}(X_k) + \mathbf{E}(Y_k))\Big| > 50n - c\Big) \leqslant$$

$$\leqslant \frac{\sum\limits_{k=1}^{n}\mathbf{D}(X_k - Y_k)}{(50n - c)^2} = \frac{\mathbf{D}(X_1 - Y_1)}{n(50 - c/n)^2} \to 0$$

при $n \to \infty$. Последнее неравенство выполнено благодаря (16.3). Таким образом, переполнение перрона неизбежно, поскольку $\mathbf{P}(N_n \geqslant c) = 1 - \mathbf{P}(N_n < c) \to 1$ для любого c.

Задачи

Задача 16.1. Устройство состоит из 60 независимо работающих элементов. Вероятность отказа каждого элемента за время t равна 0.05. Оценить снизу вероятность того, что число отказавших за время t элементов будет не больше четырех.

Задача 16.2. Игральная кость подбрасывается 100 раз. Оценить вероятность того, что суммарное число очков будет отличаться от 350 не менее чем на 30 очков.

Задача 16.3. Монета бросается 1000 раз. Оценить снизу вероятность отклонения частоты появления герба от 1/2 меньше, чем на 0.1.

Задача 16.4. Вероятность наступления некоторого события A в каждом из 1500 испытаний равна 0.2. Используя неравенство Чебышева, оценить сверху вероятность того, что отклонение числа наступлений события A от математического ожидания будет более 40.

Задача 16.5. Вероятность того, что изделие является качественным, равна 0.9. Сколько следует проверить изделий, чтобы с вероятностью не меньшей 0.95 можно было утверждать, что абсолютная величина отклонения доли качественных изделий от 0.9 не превысит 0.01?

§ 17. Центральная предельная теорема

Многие из случайных явлений возникают в результате взаимодействия большого числа малых случайных возмущений. Примерами могут служить помехи в радиотехнике, диффузии в жидкостях и газах, рост некоторых микроорганизмов и т. д. При определенных условиях действие таких возмущений приводит к неожиданному феномену: величина результирующего (суммарного) воздействия становится мало отличной от нормально распределенной случайной величины. Цель настоящего параграфа – дать математическое обоснование этому феномену. Сам феномен составляет суть предельной теоремы, которая в силу особой важности называется центральной.

При доказательстве предельных теорем, т. е. теорем о сходимости при $n \to \infty$ функций распределения некоторых случайных величин, зависящих от индекса n, к распределению предельной величины, удобно иметь дело не с распределениями величин, а с характеристическими функциями этих величин. Это связано с тем, что характеристические функции более удобны для аналитического изучения, и с тем, что сходимость распределений и сходимость характеристических функций эквивалентны, о чем свидетельствует следующий результат.

Теорема 1. *Пусть $\varphi_n(t)$ – характеристические функции случайных величин Z_n, а $F_n(x)$ – функции распределения Z_n, т. е. $\varphi_n(t) = \mathbf{E}(e^{itZ_n})$, $F_n(x) = \mathbf{P}(Z_n < x)$. Пусть $\varphi(t)$ – характеристическая функция случайной величины Z, а $F(x)$ – функция распределения Z, т. е. $\varphi(t) = \mathbf{E}(e^{itZ})$, $F(x) = \mathbf{P}(Z < x)$. Тогда следующие утверждения эквивалентны:*

1) *в каждой точке непрерывности функции $F(x)$ имеет место сходимость $F_n(x) \to F(x)$ при $n \to \infty$;*

2) *$\varphi_n(t) \to \varphi(t)$ для любого $t \in \mathbf{R}^1$ при $n \to \infty$.*

Для доказательства этого результата используются сложные математические рассуждения, поэтому его можно опустить.

Доказательство. Убедимся, что из 1) следует 2). Пусть ε – произвольное сколь угодно малое число. Так как $F(-\infty) = 0$ и $F(\infty) = 1$, то можно выбрать такое большое число c, что выполняются неравенства $F(-c) \leqslant \varepsilon/2$, $1 - F(c) \leqslant \varepsilon/2$, и в точках $-c$ и c функция $F(x)$ является непрерывной. Используя сходимости $F_n(-c) \to F(-c)$ и $F_n(c) \to F(c)$, можно выбрать столь большой номер N, что при всех $n \geqslant N$ выполняются неравенства $F_n(-c) \leqslant \varepsilon$, $1 - F_n(c) \leqslant \varepsilon$. На интервале $(-c, c)$ комплекснозначную функцию e^{itx} можно равномерно с точностью до любого малого ε приблизить некоторой ступенчатой функцией $h(x) = \sum\limits_{k=0}^{m-1} h_k \mathbb{I}_{[c_k, c_{k+1})}(x)$. Можно считать, что $c_0 = -c, c_m = c$. Оценим разность между характеристическими функциями $\varphi_n(t)$ и $\varphi(t)$. При всех $n \geqslant N$ имеем

$$|\varphi_n(t) - \varphi(t)| = \left| \int\limits_{-\infty}^{\infty} e^{itx} dF_n(x) - \int\limits_{-\infty}^{\infty} e^{itx} dF(x) \right| \leqslant$$

$$\leqslant \left| \int\limits_{-c}^{c} (e^{itx} - h(x)) dF_n(x) \right| + \left| \int\limits_{-c}^{c} h(x) dF_n(x) - \int\limits_{-c}^{c} h(x) dF(x) \right| +$$

$$+ \left| \int\limits_{-c}^{c} (h(x) - e^{itx}) dF(x) \right| + \left| \int\limits_{c}^{\infty} e^{itx} dF_n(x) \right| + \left| \int\limits_{-\infty}^{-c} e^{itx} dF_n(x) \right| +$$

$$+\left|\int\limits_{c}^{\infty} e^{itx}dF(x)\right| + \left|\int\limits_{-\infty}^{-c} e^{itx}dF(x)\right| \leqslant \varepsilon(F_n(c) - F_n(-c)) +$$

$$+\left|\sum_{k=0}^{m-1} h_k\big((F_n(c_{k+1}) - F_n(c_k)) - (F(c_{k+1}) - F(c_k))\big)\right| +$$

$$+\varepsilon(F(c) - F(-c)) + (1 - F_n(c)) + F_n(-c) + (1 - F(c)) + F(-c).$$

В силу выбора c последние четыре слагаемых в правой части этого соотношения не превосходят 3ε. В точках c_k имеет место сходимость $F_n(c_k) \to F(c_k)$. Поэтому

$$\limsup_{n\to\infty} |\varphi_n(t) - \varphi(t)| \leqslant 5\varepsilon.$$

В силу произвольности ε это означает, что $\varphi_n(t) \to \varphi(t)$ при $n \to \infty$.

Докажем теперь, что утверждение 2) влечет 1). Пусть D – какое-нибудь счетное всюду плотное множество точек на прямой, например, множество рациональных чисел. Упорядочим это множество и обозначим его элементы $x_1, x_2, \ldots, x_n, \ldots$. Поскольку последовательность неотрицательных чисел $\{F_n(x_1)\}_{n=1}^{\infty}$ ограничена единицей, то она содержит по крайней мере одну подпоследовательность $F_{m_{11}}(x_1), F_{m_{12}}(x_1), \ldots, F_{m_{1n}}(x_1), \ldots$, сходящуюся к некоторому предельному значению, которое мы обозначим $G(x_1)$. Из последовательности чисел $F_{m_{11}}(x_2), F_{m_{12}}(x_2), \ldots, F_{m_{1n}}(x_2), \ldots$ по аналогичным причинам можно выделить подпоследовательность $F_{m_{21}}(x_2), F_{m_{22}}(x_2), \ldots, F_{m_{2n}}(x_2), \ldots$, сходящуюся к некоторому пределу $G(x_2)$. Продолжая такой выбор, можно для любого целого k выделить такую последовательность $F_{m_{k1}}(x_k), F_{m_{k2}}(x_k), \ldots, F_{m_{kn}}(x_k), \ldots$, для которой одновременно при всех $r \leqslant k$ выполняются предельные соотношения $\lim\limits_{n\to\infty} F_{m_{kn}}(x_r) = G(x_r)$. Составим последовательность функций $\{F_{m_{nn}}(x)\}_{n=1}^{\infty}$. Эта последовательность образована из первоначальной последовательности функций $\{F_n(x)\}_{n=1}^{\infty}$ и обладает следующим свойством: для любой точки x_k из счетного всюду плотного множества D имеет место

$$\lim_{n\to\infty} F_{m_{nn}}(x_k) = G(x_k). \tag{17.1}$$

Это справедливо, так как в силу нашего выбора при $n \geqslant k$ последовательность чисел $F_{m_{nn}}(x_k)$ является подпоследовательностью последовательности чисел $F_{m_{k1}}(x_k), F_{m_{k2}}(x_k), \ldots,$ $F_{m_{kn}}(x_k), \ldots,$ которая сходится к пределу $G(x_k)$.

Поскольку функции $F_{m_{nn}}(x)$ не убывают, неотрицательны и равномерно ограничены единицей, то функция $G(x)$, определенная на множестве D, обладает теми же свойствами. Ее можно доопределить на всей вещественной оси так, чтобы она была непрерывной слева и неубывающей. При этом возможно придется изменить ее значения в точках разрыва. В силу того, что $\{F_{m_{nn}}(x)\}_{n=1}^{\infty}$ – последовательность неубывающих функций, из сходимости на всюду плотном множестве следует сходимость во всех точках непрерывности функции $G(x)$.

Убедимся теперь, что $G(x)$ является функцией распределения, т.е. что $G(-\infty) = 0$ и $G(\infty) = 1$. Пусть это неверно, тогда мы имеем, что

$$\delta = G(\infty) - G(-\infty) < 1. \tag{17.2}$$

Для того, чтобы исходя из (17.2) прийти к противоречию, воспользуемся утверждением, что $\varphi_n(t) \to \varphi(t)$ для любого $t \in \mathbf{R}^1$. Так как $\varphi(0) = 1$ и функция $\varphi(t)$ непрерывна, то для любого ε, $0 < \varepsilon < 1 - \delta$, можно выбрать достаточно малое положительное число τ такое, что будет выполняться неравенство

$$\frac{1}{2\tau}\left|\int\limits_{-\tau}^{\tau} \varphi(t)dt\right| > 1 - \frac{\varepsilon}{2} > \delta + \frac{\varepsilon}{2}. \tag{17.3}$$

Пусть $c > 4/(\tau\varepsilon)$ таково, что точки $-c$ и c являются точками непрерывности функции $G(x)$. Тогда ввиду того, что $F_{m_{nn}}(-c) \to G(-c), F_{m_{nn}}(c) \to G(c)$ и выполняется (17.2), можно выбрать столь большой номер N, что при всех $n \geqslant N$ справедливы оценки

$$F_{m_{nn}}(c) - F_{m_{nn}}(-c) < \delta + \frac{\varepsilon}{4}.$$

Применяя формулу (13.4) для характеристической функции $\varphi_{m_{nn}}(t)$, при $n \geqslant N$ получим следующую оценку

$$\left|\int\limits_{-\tau}^{\tau} \varphi_{m_{nn}}(t)dt\right| = \left|\int\limits_{|x|\leqslant c}\int\limits_{-\tau}^{\tau} e^{itx}dtdF_{m_{nn}}(x) + \int\limits_{|x|>c}\int\limits_{-\tau}^{\tau} e^{itx}dtdF_{m_{nn}}(x)\right| \leqslant$$

$$\leqslant \int\limits_{|x|\leqslant c}\int\limits_{-\tau}^{\tau}|e^{itx}|\,dt\,dF_{m_{nn}}(x)+\Big|\int\limits_{|x|>c}\frac{2\sin(\tau x)}{x}dF_{m_{nn}}(x)\Big|\leqslant$$

$$\leqslant 2\tau(F_{m_{nn}}(c)-F_{m_{nn}}(-c))+\frac{2}{c}\leqslant 2\tau(\delta+\frac{\varepsilon}{4})+\frac{\tau\varepsilon}{2}=2\tau(\delta+\frac{\varepsilon}{2}).$$

В силу равномерной ограниченности характеристических функций, в левой части этого соотношения можно перейти к пределу под знаком интеграла. В результате получим неравенство

$$\Big|\int\limits_{-\tau}^{\tau}\varphi(t)dt\Big|\leqslant 2\tau(\delta+\frac{\varepsilon}{2}),$$

которое противоречит неравенству (17.3). Тем самым наше предположение (17.2) является неверным, и, следовательно, $G(x)$ является функцией распределения.

В силу уже доказанной первой части теоремы из сходимости функций распределения $F_{m_{nn}}(x)$ к функции распределения $G(x)$ следует сходимость характеристических функций $\varphi_{m_{nn}}(t)$ к характеристической функции распределения $G(x)$. Поскольку мы предполагаем, что $\varphi_n(t)\to\varphi(t)$, то $\varphi(t)$ является характеристической функцией распределения $G(x)$. Согласно теореме 2 §13 характеристическая функция однозначно определяет функцию распределения. Поэтому $G(x)=F(x)$.

Чтобы завершить доказательство теоремы, нам осталось показать, что вся последовательность $\{F_n(x)\}_{n=1}^{\infty}$ сходится к $F(x)$. Предположим, что это не так. Тогда с помощью описанного выше метода диагонального выбора можно будет указать подпоследовательность n_k, для которой последовательность $\{F_{n_k}(x)\}_{k=1}^{\infty}$ сходится к некоторой фунции распределения $G^*(x)$, отличной от $F(x)$ по крайней мере в одной из точек непрерывности. По уже доказанному, характеристической функцией для $G^*(x)$ будет функция $\varphi(t)$. По теореме единственности должно быть $G^*(x)=F(x)$. Это противоречит сделанному предположению.

Лемма 1. *Если функция распределения $F(x)$ непрерывна на всей прямой, то сходимость $F_n(x)\to F(x)$ для любого $x\in\mathbf{R}^1$*

эквивалентна равномерной сходимости: при $n \to \infty$

$$\sup_{x \in \mathbf{R}^1} |F_n(x) - F(x)| \to 0. \tag{17.4}$$

Доказательство. Для произвольного $\varepsilon > 0$ разобьем интервал $[0,1]$ на интервалы $[y_k, y_{k+1}), k = 0, 1, \ldots, m$, так, чтобы $0 < y_{k+1} - y_k \leqslant \varepsilon$. Считаем, что $y_0 = 0, y_{m+1} = 1$. Поскольку функция $F(x)$ непрерывна и не убывает, то существуют такие точки x_k, что $F(x_k) = y_k$ и $x_k < x_{k+1}, k = 0, 1, \ldots, m$. Поскольку $F_n(x_k) \to F(x_k)$ при $n \to \infty$, то можно указать столь большое N, что при $n \geqslant N$ будут выполнены неравенства

$$|F_n(x_k) - F(x_k)| < \varepsilon, \quad k = 0, 1, \ldots, m+1.$$

В силу этих оценок для любого $n \geqslant N$ и $k = 0, 1, \ldots, m$ будем иметь

$$0 \leqslant F_n(x_{k+1}) - F_n(x_k) \leqslant |F_n(x_{k+1}) - F(x_{k+1})| + \\ + F(x_{k+1}) - F(x_k) + |F(x_k) - F_n(x_k)| \leqslant 3\varepsilon.$$

Используя теперь монотонность функций $F_n(x)$ и $F(x)$, для $n \geqslant N$ получим

$$\sup_{x \in \mathbf{R}^1} |F_n(x) - F(x)| \leqslant \max_{0 \leqslant k \leqslant m} \Big\{ \sup_{x \in [x_k, x_{k+1})} (F_n(x) - F_n(x_k)) + \\ + |F_n(x_k) - F(x_k)| + \sup_{x \in [x_k, x_{k+1})} (F(x) - F(x_k)) \Big\} \leqslant \\ \leqslant \max_{0 \leqslant k \leqslant m} \{F_n(x_{k+1}) - F_n(x_k) + |F_n(x_k) - F(x_k)| + \\ + F(x_{k+1}) - F(x_k)\} \leqslant 3\varepsilon + \varepsilon + \varepsilon = 5\varepsilon.$$

Так как ε выбиралось произвольным, из этой оценки следует утверждение леммы.

Теорема 1 об эквивалентности утверждений о сходимости характеристических функций и функций распределения играет принципиальную роль при доказательстве следующего основополагающего результата теории вероятностей.

Теорема 2 (центральная предельная теорема). *Пусть $\{X_l\}_{l=1}^{\infty}$ – последовательность независимых одинаково распределенных случайных величин с математическими ожиданиями $\mathbf{E}(X_l) = a$ и конечными дисперсиями $\mathbf{D}(X_l) = \sigma^2$. Положим $S_n = \sum\limits_{l=1}^{n} X_l$. Пусть $F_n(x) = \mathbf{P}(Z_n < x)$ – функция распределения случайной величины $Z_n = \dfrac{S_n - na}{\sigma\sqrt{n}}$, а $\mathcal{N}(x)$ – функция распределения нормального закона со средним 0 и дисперсией 1. Тогда при $n \to \infty$*

$$\sup_{x \in \mathbf{R}^1} |F_n(x) - \mathcal{N}(x)| \to 0. \tag{17.5}$$

Доказательство. Характеристическая функция нормального распределения со средним 0 и дисперсией 1 имеет вид $e^{-t^2/2}$ (см. § 14). Пусть $\varphi_n(t) = \mathbf{E}(e^{itZ_n})$ – характеристическая функция величины Z_n. Согласно теореме 1 и лемме 1 достаточно доказать, что для любого $t \in \mathbf{R}^1$

$$\varphi_n(t) \to e^{-t^2/2}. \tag{17.6}$$

Пусть $\varphi(t) = \varphi_{Y_l}(t) = \mathbf{E}(e^{itY_l})$ – характеристическая функция величины $Y_l = \dfrac{X_l - a}{\sigma}$. Величину Z_n представим в виде суммы величин Y_l, деленной на $\sqrt{n}$:

$$Z_n = \frac{S_n - an}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} \frac{(X_l - a)}{\sigma} = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} Y_l.$$

Величины Y_l независимы и одинаково распределены, поскольку таким свойством обладают величины X_l. Используя свойство 2 и следствие 2 для характеристических функций (§ 13), получим

$$\varphi_n(t) = \varphi_{\sum_{l=1}^{n} Y_l}\left(\frac{t}{\sqrt{n}}\right) = \prod_{l=1}^{n} \varphi_{Y_l}\left(\frac{t}{\sqrt{n}}\right) = \varphi^n\left(\frac{t}{\sqrt{n}}\right). \tag{17.7}$$

По формуле Тейлора, примененной к функции $\varphi(t)$, имеем

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + \alpha(t)t^2, \tag{17.8}$$

где $\alpha(t) \to 0$ при $t \to 0$. Согласно первому свойству характеристической функции $\varphi(0) = 1$. По следствию 3 § 13

$$\mathbf{E}(Y_l) = -i\varphi'(0), \quad \mathbf{D}(Y_l) = -\varphi''(0) + (\varphi'(0))^2.$$

Используя известные нам свойства, вычислим математическое ожидание $\mathbf{E}(Y_l)$ и дисперсию $\mathbf{D}(Y_l)$:

$$\mathbf{E}(Y_l) = \mathbf{E}\left(\frac{X_l - a}{\sigma}\right) = \frac{1}{\sigma}(\mathbf{E}(X_l) - a) = 0, \quad \mathbf{D}(Y_l) = \frac{1}{\sigma^2}\mathbf{D}(X_l) = 1.$$

В результате имеем $\varphi'(0) = 0$, $\varphi''(0) = -1$. Подставляя эти значения в (17.8), получим

$$\varphi(t) = 1 - \frac{t^2}{2} + \alpha(t)t^2.$$

Поэтому формулу (17.7) можно переписать в виде

$$\varphi_n(t) = \left(1 - \frac{t^2}{2n} + \alpha\left(\frac{t}{\sqrt{n}}\right)\frac{t^2}{n}\right)^n.$$

Обозначим $\beta_n(t) = -\frac{t^2}{2n} + \alpha\left(\frac{t}{\sqrt{n}}\right)\frac{t^2}{n}$. Тогда $n\beta_n(t) \to -t^2/2$, $\beta_n(t) \to 0$, и, принимая во внимание "замечательный предел"

$$\lim_{\beta \to 0}(1 + \beta)^{1/\beta} = e,$$

получим

$$\varphi_n(t) = \left((1 + \beta_n(t))^{1/\beta_n(t)}\right)^{n\beta_n(t)} \to e^{-t^2/2}.$$

Требуемое соотношение (17.6), а, следовательно, и теорема 2 доказаны.

Замечание 1. Как следует из доказательства, величину Z_n можно представить в виде $Z_n = \sum\limits_{l=1}^{n} Y_{n,l}$, где слагаемые $Y_{n,l} = \frac{Y_l}{\sqrt{n}}$ независимы и одинаково распределены со средним $\mathbf{E}(Y_{n,l}) = 0$ и дисперсией $\mathbf{E}(Y_{n,l}^2) = \frac{1}{n}$ – обратно пропорциональной числу слагаемых. Часто оказывается, что зависимость слагаемых $Y_{n,l}$ от n является более сложной, чем зависимость вида $Y_l/\sqrt{n}$. Тем не менее, если эти слагаемые при каждом n независимы, имеют среднее 0 и дисперсию, обратно пропорциональную n, то при некотором дополнительном ограничении справедливо соотношение (17.5). Очень важным обстоятельством является то, что центральная предельная теорема справедлива для случайных величин X_l с любыми распределениями, для которых $\mathbf{E}X_l = 0, \mathbf{D}(X_l) = \sigma^2 < \infty$.

Замечание 2. Из формул (17.5), (10.1) и (14.1) следует, что для $s_1 < s_2$

$$\begin{aligned}\mathbf{P}(s_1 \leqslant S_n < s_2) = \mathbf{P}\left(\frac{s_2 - na}{\sigma\sqrt{n}} \leqslant \frac{S_n - na}{\sigma\sqrt{n}} < \frac{s_1 - na}{\sigma\sqrt{n}}\right) &= \\ = F_n\left(\frac{s_2 - na}{\sigma\sqrt{n}}\right) - F_n\left(\frac{s_1 - na}{\sigma\sqrt{n}}\right) &\approx \\ \approx \mathcal{N}\left(\frac{s_2 - na}{\sigma\sqrt{n}}\right) - \mathcal{N}\left(\frac{s_1 - na}{\sigma\sqrt{n}}\right) &= \\ = \frac{1}{2}\left(\Phi\left(\frac{s_2 - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{s_2 - na}{\sigma\sqrt{n}}\right)\right).& \qquad (17.9)\end{aligned}$$

Пример 1. Приведем вывод интегральной теоремы Муавра–Лапласа (§ 9) из центральной предельной теоремы. Проводятся n испытаний Бернулли, и нас интересует число наступлений события A. Пусть случайная величина X_l принимает значение 1, если в l-м испытании событие A наступило, и 0 в противоположном случае. Тогда $\mathbf{E}(X_l) = p$, $\mathbf{D}(X_l) = pq$, где $p = \mathbf{P}(A)$ и $q = 1-p$. Обозначим $S_n = \sum\limits_{l=1}^{n} X_l$. Величина S_n равна числу наступлений события A в n испытаниях. Для $m_1 < m_2$ положим $a_n = \dfrac{m_1 - np}{\sqrt{npq}}$, $b_n = \dfrac{m_2 - np}{\sqrt{npq}}$. Используя введенные обозначения и (17.6), получим, что для $P_n(m_1, m_2)$ (определение см. § 8) справедливо соотношение

$$\begin{aligned}P_n(m_1, m_2) = \mathbf{P}(m_1 \leqslant S_n \leqslant m_2) = \mathbf{P}(m_1 \leqslant S_n < m_2 + 1) \approx \\ \approx \frac{1}{2}\left(\Phi\left(b_n + \frac{1}{\sqrt{npq}}\right) - \Phi(a_n)\right) \approx \frac{1}{2}\left(\Phi(b_n) - \Phi(a_n)\right).\end{aligned}$$

В последнем приближенном равенстве мы воспользовались равномерной непрерывностью функции $\Phi(x)$. Полученное соотношение и составляет утверждение интегральной теоремы Муавра–Лапласа.

Задача 1. Производится стрельба с большого расстояния по круглой мишени так, что при каждом выстреле попадание равновозможно в любую точку мишени. Промахов нет.

Мишень разделена концентрическими окружностями, равноотстоящими друг от друга на 10 областей. Центральная область является кругом с радиусом, равным расстоянию между окружностями, и попадание в нее оценивается в 10 очков. Попадание в прилегающее к ней кольцо оценивается в 9 очков, в последующее – 8 очков и так далее – до 1 очка для последнего кольца. Вычислить приближенное значение вероятности того, что при 100 независимых выстрелах будет набрано от 360 до 430 очков.

Решение. Поскольку все точки мишени равновозможны для попадания, то шансы попадания в каждое из колец описываются геометрическими вероятностями. Площадь круга равна πr^2, где r – радиус. Обозначим расстояние между окружностями через d. Легко подсчитать, что площадь k-го кольца равна $\pi d^2((k+1)^2 - k^2) = \pi d^2(2k+1)$. За нулевое кольцо принимается центр круга, а за 9-е – последнее кольцо. Площадь всей мишени равна $100\pi d^2$. Следовательно, вероятность попадания в k-е кольцо равна $(2k+1)/100$, $k = 0, 1, 2, \ldots, 9$. Пусть X – случайная величина, равная числу очков, набранных при одном выстреле. Тогда закон распределения X задается следующей таблицей

X	10	9	8	7	6	5	4	3	2	1
$\mathbf{P}$	$\frac{1}{100}$	$\frac{3}{100}$	$\frac{5}{100}$	$\frac{7}{100}$	$\frac{9}{100}$	$\frac{11}{100}$	$\frac{13}{100}$	$\frac{15}{100}$	$\frac{17}{100}$	$\frac{19}{100}$

Вычисляя математическое ожидание и дисперсию величины X, получим $\mathbf{E}(X) = 3.85$, $\mathbf{D}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X) = 5.5275$. Следовательно, $\sqrt{\mathbf{D}(X)} \approx 2.351$.

Пусть X_l – случайная величина, равная числу очков при l-м выстреле. Тогда $S_n = \sum\limits_{l=1}^{n} X_l$ – суммарное число очков при n выстрелах. Применяя центральную предельную теорему (точнее формулу 17.6), для любых $s_1 < s_2$ получим

$$\mathbf{P}\left(s_1 \leqslant S_n < s_2\right) \approx \frac{1}{2}\left(\Phi\left(\frac{s_2 - n\mathbf{E}(X)}{\sqrt{\mathbf{D}(X)}\sqrt{n}}\right) - \Phi\left(\frac{s_1 - n\mathbf{E}(X)}{\sqrt{\mathbf{D}(X)}\sqrt{n}}\right)\right).$$

Подставляя в правую часть этого соотношения значения математического ожидания, дисперсии и параметров $n = 100$,

$s_1 = 360$, $s_2 = 430$, получим, что вероятность набрать при стрельбе от 360 до 430 очков приблизительно равна

$$\frac{1}{2}(\Phi(1.91) - \Phi(-1.06)) = \frac{1}{2}(\Phi(1.91) + \Phi(1.06)) \approx 0.825.$$

Задачи

Задача 17.1. Игральная кость подбрасывается 420 раз. Какова вероятность того, что суммарное число очков будет находиться в пределах от 1400 до 1505?

Задача 17.2. При выстреле по мишени стрелок попадает в десятку с вероятностью 0.5, в девятку – 0.3, в восьмерку – 0.1, в семерку – 0.05, в шестерку – 0.05. Стрелок сделал 100 выстрелов. Какова вероятность того, что он набрал не менее 900 очков?

Задача 17.3. Предположим, что на станцию скорой помощи вызовы в течение суток поступают по закону Пуассона с параметром $\lambda = 73$ и в разные сутки их количество не зависит друг от друга. Определить вероятность того, что в течение года (365 дней) общее число вызовов будет в пределах от 26500 до 26800.

Часть 2
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Математическая статистика – это дисциплина, изучающая методы оценивания и сравнения распределений случайных величин и их характеристик по наблюдениям случайных величин.

§ 18. Случайная выборка. Эмпирическая функция распределения

Пусть X – случайная величина с функцией распределения $F(x) = \mathbf{P}(X < x)$. Если в ходе случайного эксперимента наблюдается одно значение случайной величины X, то существенной информации о распределении X или ее характеристиках получить нельзя. Однако, если проводится большое количество не зависимых друг от друга одинаковых случайных экспериментов, в каждом из которых наблюдается значение случайной величины X, то при достаточно большом количестве экспериментов можно получить хорошие оценки функции распределения величины X и ее характеристик. Построение таких оценок и является одной из основных задач статистики.

Определение 1. **Случайной выборкой объема** n, отвечающей случайной величине X с функцией распределения $F(x)$, называется набор n независимых случайных величин $X_1, X_2, \ldots, X_n$, каждая из которых имеет распределение $F(x)$.

Определение 2. Случайная величина, являющаяся функцией случайной выборки, называется **статистикой**.

Таким образом, для любой достаточно хорошей (например, кусочно-непрерывной) функции $g(\vec{x})$, $\vec{x} = (x_1, x_2, \ldots, x_n)$, величина $g(X_1, X_2, \ldots, X_n)$ – статистика.

Величина $X_{(k)} = g_k(X_1, X_2, \ldots, X_n)$, где g_k – отображение из R^n в R^1, сопоставляющее каждому вектору $\vec{x} \in R^n$ ту из его координат, которая занимает k-е по порядку значение в упорядоченном по возрастанию наборе, составленном из координат вектора, называется k-й **порядковой статистикой**. В частности, $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ и

$X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$. Набор упорядоченных случайных величин $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$ называется **вариационным рядом**.

Часто для краткости случайную выборку $X_1, X_2, \ldots, X_n$ объединяют в один случайный n-мерный вектор $\vec{X} = (X_1, X_2, \ldots, X_n)$. Экспериментатор (статистик), как правило, располагает одной реализацией этой случайной выборки, т. е. набором чисел $\vec{x} = (x_1, x_2, \ldots, x_n)$, полученных в результате наблюдений величины X при n независимых повторениях случайного эксперимента в одинаковых условиях. Согласно теории, оценки параметров величины X или ее распределения, построенные по случайной выборке $(X_1, X_2, \ldots, X_n)$, будут сходиться при $n \to \infty$ к истинному значению параметра или распределению либо по вероятности, либо с вероятностью единица, т. е. для почти всех реализаций выборки. Таким образом, уже по одной реализации $(x_1, x_2, \ldots, x_n)$ случайной выборки, когда n велико, можно построить оценку, с большой вероятностью хорошо аппроксимирующую истинный параметр или распределение.

Пример 1. Рассмотрим задачу о приближенном вычислении вероятности выпадения герба для несимметричной монеты. Пусть p – вероятность выпадения герба, а $1-p$ – вероятность выпадения решетки. Для того, чтобы приближенно вычислить параметр p, производятся независимые бросания монеты, и каждый раз выпавшему гербу сопоставляется единица, а решетке – ноль. Таким образом, мы имеем последовательность независимых одинаково распределенных случайных величин

$$X_l = \begin{cases} 1, & \text{если при } l\text{-м бросании выпадает герб,} \\ 0, & \text{если при } l\text{-м бросании выпадает решетка.} \end{cases}$$

Набор величин $(X_1, X_2, \ldots, X_n)$ является случайной выборкой для величины X с $\mathbf{P}(X=1) = p, \mathbf{P}(X=0) = 1-p$. Поскольку $\mathbf{E}(X_l) = 1 \cdot p + 0 \cdot (1-p) = p$, то из закона больших чисел (см. часть 1, § 16) следует, что случайная величина $\overline{X}_n = \frac{1}{n}\sum\limits_{l=1}^{n} X_l$, равная доле выпавших гербов, сходится с вероятностью единица при $n \to \infty$ к величине p. Следовательно, статистика $\overline{X}_n$ является хорошей оценкой для p, если n достаточно велико. При многократном бросании монеты наблюдатель имеет лишь набор n чисел вида $(1,0,1,1,0,\ldots,1) = (x_1, x_2, \ldots, x_n)$, т. е. одну конкретную реализацию случайной выборки $(X_1, X_2, \ldots, X_n)$. Вместо случайной величины $\overline{X}_n$ он тоже имеет ее реализацию $\bar{x}_n = \frac{1}{n}\sum\limits_{l=1}^{n} x_l$. С математической точки зрения реализация случайной выборки $(x_1, \ldots, x_n)$ представляет собой значение случайного вектора $\vec{X}$ при каком-то элементарном исходе, реализованном в целой последовательности экспериментов. Поскольку сходимость $\overline{X}_n \to p$ имеет место для почти всех исходов (с вероятностью единица), то и $\bar{x}_n$ при достаточно больших n будет близка к p.

Замечание 1. Одна из основных задач математической статистики состоит в построении по случайным выборкам $\vec{X}$ других случайных величин, которые служат оценками параметров или распределений. Эти оценки при неограниченном увеличении числа компонент выборки сходятся с вероятностью единица или по вероятности к истинному значению параметра наблюдаемой величины или ее распределению. Прикладная статистика имеет дело с конкретными реализациями случайных величин, полученными в ходе случайных экспериментов. Она использует оценки и формулы математической статистики, при этом вместо случайных величин X_l подставляются их конкретные реализации x_l.

Рассмотрим задачу о построении оценки для функции распределения $F(x)$ случайной величины X по случайной выборке $\vec{X} = (X_1, X_2, \ldots, X_n)$. Такой оценкой будет служить так называемая эмпирическая функция распределения.

О п р е д е л е н и е 3. **Эмпирической функцией распределения**, построенной по случайной выборке $(X_1, X_2, \ldots, X_n)$, называется случайная функция

$$F_n(x) = \frac{1}{n}\sum_{l=1}^{n} \mathbb{I}_{(-\infty,x)}(X_l),$$

где $\mathbb{I}_A(y)$ – индикатор множества A.

Замечание 2. *Для конкретной реализации* $(x_1, x_2, \ldots, x_n)$ выборки $(X_1, X_2, \ldots, X_n)$ и фиксированного числа x величина $F_n(x)$ равна доле тех значений x_l, которые меньше x.

Замечание 3. Свойства эмпирической функции распределения $F_n(x)$ аналогичны свойствам обычной функции распределения: $F_n(x)$ – неубывающая функция по x, $0 \leqslant F_n(x) \leqslant 1$ для любого x, и $F_n(-\infty) = 0, F_n(\infty) = 1$.

Замечание 4. Эмпирическую функцию распределения можно описать следующим образом. От минус бесконечности вплоть до первой порядковой статистики $X_{(1)}$ значение $F_n(x)$ равно нулю. Далее значение $F_n(x)$ увеличивается на $\frac{1}{n}$ в каждой из точек $X_{(k)}$, т. е. $F_n(x) = \frac{k}{n}$ при $x \in (X_{(k)}, X_{(k+1)}]$. Если несколько значений $X_{(k)}$ совпадают, то значение $F_n(x)$ увеличивается на $\frac{1}{n}$, домноженное на число совпадений. При x большем чем n-я порядковая статистика $X_{(n)}$ значение $F_n(x)$ равно 1.

Для любого фиксированного x случайная величина $F_n(x)$ имеет следующее распределение

$$\mathbf{P}\big(F_n(x) = \frac{k}{n}\big) = C_n^k(F(x))^k(1 - F(x))^{n-k}, \quad k = 1, 2, \ldots, n.$$

Действительно, событие $\{F_n(x) = k/n\}$ состоит в том, что произошло ровно k из событий $\{X_l < x\}$, $l = 1, 2, \ldots, n$, и, следовательно, произошло $n - k$ противоположных событий. Воспользуемся схемой Бернулли (§ 8). Будем говорить, что в l-м испытании наступил успех, если $\{X_l < x\}$, и неудача, если $\{X_l \geqslant x\}$. Тогда вероятность успеха $p = \mathbf{P}(X_l < x) = F(x)$, а вероятность неудачи $q = \mathbf{P}(X_l \geqslant x) = 1 - F(x)$. По формуле Бернулли (8.1) вероятность наступления ровно k успехов равна $C_n^k p^k q^{n-k}$, и эта вероятность равна вероятности события $\{F_n(x) = k/n\}$. Что и требовалось доказать.

В рамках приведенной схемы легко вычислить функцию распределения порядковой статистики $X_{(m)}$. Событие $\{X_{(m)} < x\}$ означает, что наступило не менее чем m событий вида $\{X_l < x\}$, $l = 1, 2, \ldots, n$. Вероятность такого события по формулам (8.2), (8.3) равна

$$\mathbf{P}(X_{(m)} < x) = P_n(m, n) = \sum_{k=m}^{n} C_n^k(F(x))^k(1 - F(x))^{n-k} =$$

$$= 1 - \sum_{k=0}^{m-1} C_n^k(F(x))^k(1 - F(x))^{n-k}.$$

Для крайних порядковых статистик эти формулы принимают простой вид:

$$\mathbf{P}(X_{(1)} < x) = 1 - (1 - F(x))^n, \quad \mathbf{P}(X_{(n)} < x) = (F(x))^n.$$

Пример 2. Пусть наблюдается следующая реализация выборки: 0.2; −1.7; −3.6; 2.1; 5.2; −3.4; 4.1; 1.8; −1.3; 2.6.

График эмпирической функции распределения для этой реализации имеет вид

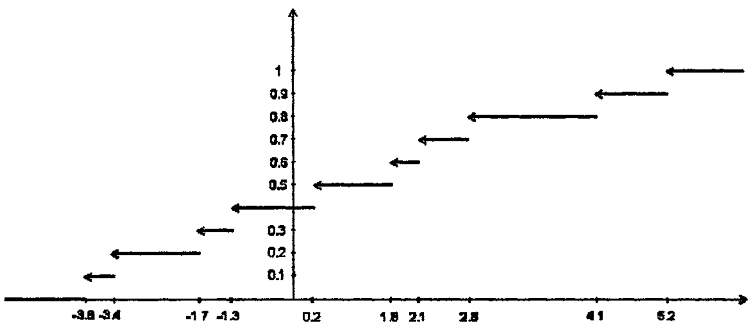

Рис. 9

Это ступенчатая функция, имеющая в точках выборки скачки размера 0.1.

Следующий результат утверждает, что эмпирическая функция распределения $F_n(x)$ является хорошей оценкой для функции распределения $F(x)$.

Теорема 1 (Гливенко). *С вероятностью единица при* $n \to \infty$

$$\sup_{x \in \mathbf{R}^1} |F_n(x) - F(x)| \to 0. \tag{18.1}$$

Доказательство. Мы докажем это утверждение лишь для случая, когда функция распределения $F(x)$ непрерывна. Для произвольной функции $F(x)$ доказательство осуществляется аналогично, хотя и требует дополнительных рассуждений. Пусть $\varepsilon > 0$ – произвольное малое число, для которого число $m = 1/\varepsilon$ целое. Так как функция $F(x)$ непрерывна и не убывает, то можно выбрать такую возрастающую последовательность чисел $x_0 = -\infty, x_1, \ldots, x_{m-1}, x_m = \infty$, что

$$F(x_0) = 0,\ F(x_1) = \varepsilon,\ \ldots,\ F(x_k) = \varepsilon k,\ \ldots, F(x_m) = 1.$$

Поскольку фунции $F_n(x)$ и $F(x)$ являются неубывающими, то для любого k и $x \in [x_k, x_{k+1})$ справедливы соотношения

$$F_n(x) - F(x) \leqslant F_n(x_{k+1}) - F(x_k) = F_n(x_{k+1}) - F(x_{k+1}) + \varepsilon,$$

$$F_n(x) - F(x) \geqslant F_n(x_k) - F(x_{k+1}) = F_n(x_k) - F(x_k) - \varepsilon.$$

Отсюда следует, что

$$\sup_{x \in \mathbf{R}^1} |F_n(x) - F(x)| \leqslant \max_{0 \leqslant k \leqslant m} |F_n(x_k) - F(x_k)| + \varepsilon. \qquad (18.2)$$

Положим $Y_l = \mathbb{1}_{(-\infty,x)}(X_l)$. Поскольку величины X_l независимы и одинаково распределены, то и величины Y_l независимы и одинаково распределены. В силу определения математического ожидания и первого свойства дисперсии имеем

$$\mathbf{E}(Y_l) = \mathbf{E}(\mathbb{1}_{(-\infty,x)}(X_l)) = \mathbf{P}(X_l \in (-\infty, x)) = \mathbf{P}(X_l < x) = F(x),$$
$$\mathbf{D}(Y_l) = \mathbf{E}(\mathbb{1}^2_{(-\infty,x)}(X_l)) - \mathbf{E}^2(Y_l) = F(x)(1 - F(x)) \leqslant 1.$$

Тогда, согласно закону больших чисел (теорема 1, § 16), для любого фиксированного x

$$F_n(x) = \frac{1}{n}\sum_{l=1}^{n} Y_l \to \mathbf{E}(Y_l) = F(x)$$

с вероятностью единица. Поэтому для конечного набора точек $x_0, x_1, \ldots, x_m$ каждое из предельных соотношений $F_n(x_k) \to F(x_k)$, $k = 0, 1, \ldots, m$, будет выполняться с вероятностью единица. Теперь, принимая во внимание оценку (18.2), получим, что с вероятностью единица

$$\limsup_{n\to\infty} \sup_{x\in\mathbf{R}^1} |F_n(x) - F(x)| \leqslant \varepsilon = \frac{1}{m}.$$

Событие, состоящее из тех исходов, для которых выполняется это неравенство, обозначим Ω_m. Это неравенство справедливо для всех натуральных m. Следовательно, для тех исходов, которые входят во все события Ω_m одновременно, т. е. входят в произведение $\prod_{m=1}^{\infty} \Omega_m$, выполняется соотношение (18.1). Согласно Лемме 1 § 6 $\mathbf{P}\left(\prod_{m=1}^{\infty} \Omega_m\right) = 1$. Таким образом, множество исходов, для которых выполняется соотношение (18.1), имеет вероятность единица.

Сформулируем без доказательства еще один результат, который характеризует качество оценки $F_n(x)$. Из него следует, что оценка $F_n(x)$ является достаточно хорошей оценкой функции распределения $F(x)$.

Теорема 2 (Колмогорова). *Пусть*

$$\mathcal{K}(z) = \begin{cases} \sum\limits_{l=-\infty}^{\infty} (-1)^l \exp(-2l^2 z^2), & \textit{при } 0 < z, \\ 0, & \textit{при } z \leqslant 0, \end{cases}$$

– **функция распределения Колмогорова**. *Тогда, если функция распределения $F(x)$ непрерывна, то для любого z*

$$\mathbf{P}(\sqrt{n} \max_{x \in \mathbf{R}^1} |F_n(x) - F(x)| < z) \to \mathcal{K}(z), \qquad \textit{при } n \to \infty.$$

Этот результат можно переформулировать следующим образом: случайная величина $Z_n = \sqrt{n} \max\limits_{x \in \mathbf{R}^1} |F_n(x) - F(x)|$ при больших n имеет функцию распределения, мало отличающуюся от $\mathcal{K}(z)$. В силу этого, для любого наперед заданного малого числа α, выбирая $z_{1-\alpha}$ так, чтобы $\mathcal{K}(z_{1-\alpha}) = 1 - \alpha$, мы получим, что оценка $\max\limits_{x \in \mathbf{R}^1} |F_n(x) - F(x)| < \frac{z_{1-\alpha}}{\sqrt{n}}$ выполняется с вероятностью, приблизительно равной $1 - \alpha$. Таким образом, с вероятностью, близкой к единице, для всех x оценка $F_n(x)$ отличается от $F(x)$ не более чем на $z_{1-\alpha}/\sqrt{n}$.

Перейдем к оцениванию плотности распределения случайной величины X. Пусть $f(x)$ – плотность распределения величины X, и пусть $(X_1, X_2, \ldots, X_n)$ – случайная выборка, отвечающая величине X. Выберем произвольное число $h > 0$. Положим $x_h = [x/h]h$, где $[a]$ – обозначает целую часть числа a, т. е. $[a]$ – наибольшее целое число, не превосходящее a (например: $[2.31] = 2, [-2.11] = -3$).

Определение 3. **Эмпирической плотностью распределения**, построенной по случайной выборке $(X_1, X_2, \ldots, X_n)$. называется случайная функция

$$f_{n,h}(x) = \frac{1}{nh} \sum_{l=1}^{n} \mathbb{I}_{[x_h, x_h+h)}(X_l).$$

Рассмотрим эту функцию подробнее. Когда x пробегает всю вещественную ось от $-\infty$ до ∞, число $k = [x/h]$ принимает все целые значения $\ldots, -2, -1, 0, 1, 2, \ldots$. Обозначим ν_k – число тех значений величин X_l, которые попадают в интервал $[kh, (k+1)h)$, т. е. $\nu_k = \sum\limits_{l=1}^{n} \mathbb{I}_{[kh,(k+1)h)}(X_l)$. При $kh \leqslant x < (k+1)h$ имеем $k = [x/h], x_h = kh$, и, следовательно, $f_{n,h}(x) = \frac{\nu_k}{nh}$. Теперь ясно, как построить график функции $f_{n,h}(x)$, который называется **гистограммой** распределения случайной величины X. Нужно разбить всю вещественную ось на интервалы длины h и каждому интервалу $[kh, (k+1)h)$ сопоставить число $\frac{\nu_k}{nh}$.

Пример 3. Пусть дана следующая реализация выборки, состоящей из 20 компонент:

0.78; −0.12; −0.61; 0.92; 0.55; 1.63; −1.16; 0.01; −0.45; 0.20; −0.38; 0.62; 0.46; −0.22; 1.11; −0.77; 0.37; 0.72; −0.24; 0.42.

Гистограмма с шагом $h = 0.4$ для этой реализации имеет следующий вид:

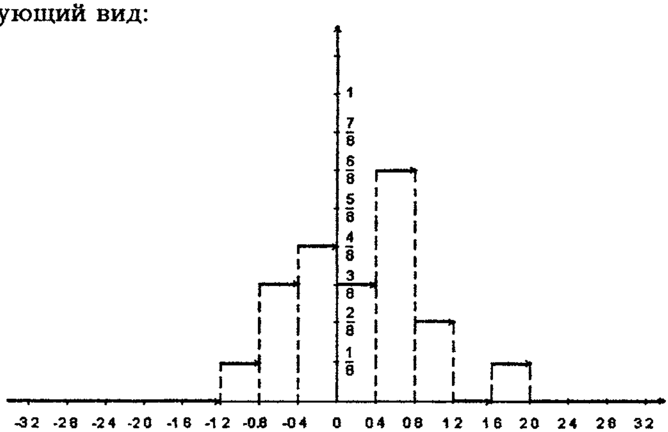

Рис. 10

Эмпирическая плотность распределения $f_{n,h}(x)$ используется в качестве оценки для $f(x)$ – плотности распределения случайной величины X.

Теорема 3 (о сходимости эмпирических плотностей распределения). *Пусть $n \to \infty$ и $h \to 0$ так, что $nh \to \infty$. Тогда в каждой точке x, в которой плотность $f(x)$ непрерывна, $f_{n,h}(x) \to f(x)$ по вероятности.*

Доказательство. В целях упрощения доказательства мы не будем переходить к пределу одновременно при $n \to \infty, h \to 0$ и $nh \to \infty$. Предельный переход будет осуществлен последовательно. Сначала перейдем к пределу при $n \to \infty$ при фиксированном h, а затем – при $h \to 0$. Предложенное ниже доказательство хорошо отражает существо дела. Доказательство утверждения теоремы 3, в котором осуществляется одновременный предельный переход, не намного сложнее, но оно использует вариант закона больших чисел, который мы не рассматривали.

Положим $Y_l = \mathbb{I}_{[x_h, x_h+h)}(X_l)$. Как и X_l, величины Y_l являются независимыми и одинаково распределенными. Математическое ожидание и дисперсия величины Y_l имеют вид

$$\begin{aligned}
&\mathbf{E}(Y_l) = \mathbf{P}(X \in [x_h, x_h + h)),\\
&\mathbf{D}(Y_l) = \mathbf{P}(X \in [x_h, x_h + h)) - \mathbf{P}^2(X \in [x_h, x_h + h)) \leqslant 1.
\end{aligned}$$

Применяя закон больших чисел (теорема 1, § 16), получим, что при $n \to \infty$

$$f_{n,h}(x) = \frac{1}{nh}\sum_{l=1}^{n} Y_l \to \frac{1}{h}\mathbf{E}(Y_1) = \frac{1}{h}\mathbf{P}(X \in [x_h, x_h + h))$$

по вероятности. Далее, поскольку $f(x)$ непрерывна, а $x_h \to x$ при $h \to 0$, имеем

$$\frac{1}{h}\mathbf{P}(X \in [x_h, x_h + h)) = \frac{1}{h}\int_{x_h}^{x_h+h} f(y)dy \to f(x).$$

Таким образом установлено, что по вероятности

$$\lim_{h\to 0}\lim_{n\to\infty} f_{n,h}(x) = f(x).$$

Именно это мы и планировали доказать.

Замечание 5. Чтобы обеспечить выполнение условий теоремы 3, следует при построении эмпирической плотности распределения выбирать h в зависимости от n так, чтобы при больших n и малых h произведение nh тоже было большим, например, $h = 1/\sqrt{n}$.

В дальнейшем нам потребуется следующая характеристика функции распределения.

Определение 2. **Квантилью порядка** p функции распределения $F(x)$ называется такое число z_p, для которого $F(z_p) \leqslant p$, а $\lim\limits_{x \downarrow z_p} F(x) \geqslant p$, если оно единственно, а если множество таких чисел составляет целый интервал, то полагаем z_p – средняя точка интервала.

Для непрерывной строго возрастающей функции $F(x)$ при любом $0 < p < 1$ квантиль z_p единственна и $F(z_p) = p$. В данном случае это равенство служит определением квантили.

Квантиль порядка 1/2 функции распределения $F(x)$ называют **медианой** распределения.

Для случайной величины с непрерывной строго возрастающей функцией распределения вероятность того, что величина примет значение меньше медианы, равна вероятности того, что она примет значение больше медианы и равна 1/2.

Многие численные характеристики эмпирического распределения служат в качестве **выборочных характеристик**.

Определим **выборочную квантиль порядка** p, как квантиль порядка p для эмпирического распределения $F_n(x)$, т. е. определим ее равенством

$$\hat{z}_p = \begin{cases} X_{([np]+1)}, & \text{если } np \text{ – не целое число,} \\ \frac{1}{2}\left(X_{([np])} + X_{([np]+1)}\right), & \text{если } np \text{ – целое число,} \end{cases}$$

где $[a]$ – целая часть числа a.

Выборочная медиана – это медиана эмпирического распределения. Она равна величине $X_{((n+1)/2)}$, если n – нечетное, и $\frac{1}{2}\left(X_{(n/2)} + X_{((n+2)/2)}\right)$, если n – четное.

В следующем параграфе речь пойдет о других выборочных характеристиках.

Задача 1. Вычислить выборочную квантиль порядка $p = 0.35$ для реализации выборки из примера 2. Определить выборочную медиану.

Решение. В примере 2 $x_{(1)} = -3.6$, $x_{(2)} = -3.4$, $x_{(3)} = -1.7$, $x_{(4)} = -1.3$, $x_{(5)} = 0.2$, $x_{(6)} = 1.8$, $x_{(7)} = 2.1$, $x_{(8)} = 2.6$, $x_{(9)} = 4.1$, $x_{(10)} = 5.2$. Так как $n = 10$, то $[np] = [3.5] = 3$. Следовательно, $\hat{z}_{0.35} = x_{(4)} = -1.3$. Поскольку 10 – четное число, то выборочная медиана равна $\frac{1}{2}(x_{(5)} + x_{(6)}) = \frac{1}{2}(0.2 + 1.8) = 1$.

Задачи

Задача 18.1. В течение дня ежечасные измерения напряжения тока в электросети дали следующие значения:

227.3; 219.1; 215.3; 229.6; 231.8; 218.7;
222.4; 229.2; 225.8; 219.5; 219.9; 217.1.

Построить график эмпирической функции распределения и вычислить выборочную медиану.

Задача 18.2. Пусть наблюдается следующая реализация выборки, состоящей из 40 компонент:

0.42	−0.44	0.59	0.79	−0.10	0.12	0.48	−1.27
0.59	0.65	1.27	−1.40	−1.57	−1.61	1.45	1.26
0.86	−0.88	−0.14	−0.52	0.53	1.38	0.02	−0.26
0.23	−0.62	−0.02	−0.46	−0.89	−0.39	−0.53	1.70
−0.85	0.40	0.97	0.83	0.41	−0.34	−1.09	0.58

Построить гистограмму с шагом $h = 0.4$ для этой реализации. Написать реализацию вариационного ряда, вычислить квантиль порядка 0.8 и выборочную медиану.

§ 19. Оценки параметров распределения. Выборочные моменты

Пусть $\vec{X} = (X_1, X_2, \ldots, X_n)$ – случайная выборка, отвечающая величине X с функцией распределения $F(x)$. Пусть θ – некоторый параметр, характеризующий распределение случайной величины X, например: $\theta = \mathbf{E}(X)$ – математическое ожидание или $\theta = \mathbf{D}(X)$ – дисперсия, или $\theta = \mathbf{E}(X^k)$ – момент порядка k. Пусть по случайной выборке $\vec{X}$ для параметра θ построена некоторая оценка $\gamma_n(\vec{X})$, которая является случайной величиной.

Определение 1. Оценка $\gamma_n(\vec{X})$ называется **состоятельной** оценкой параметра θ, если $\gamma_n(\vec{X}) \to \theta$ по вероятности при $n \to \infty$.

В этом определении, если говорить абсолютно строго, речь идет о состоятельной последовательности оценок, а термин "состоятельная оценка" служит сокращенным названием для этого понятия.

Определение 2. Оценка $\gamma_n(\vec{X})$ называется **несмещенной** оценкой праметра θ, если $\mathbf{E}(\gamma_n(\vec{X})) = \theta$.

Определение 3. Оценка $\gamma_n(\vec{X})$ называется **инвариантной относительно сдвига** (безразличной к сдвигу), если для любой постоянной c

$$\gamma_n(X_1 + c, X_2 + c, \ldots, X_n + c) = \gamma_n(X_1, X_2, \ldots, X_n)$$

Рассмотрим конкретные оценки параметров распределения случайной величины X. Для того, чтобы в дальнейшем не усложнять формулировки теорем, мы сразу предположим, что выполняется следующее условие: если рассматривается оценка момента порядка k или центрального момента порядка k, то у случайной величины X конечен момент порядка $2k$. Это условие нам потребуется для того, чтобы применять закон больших чисел (§ 16).

1. Выборочное среднее. В качестве оценки для математического ожидания $m = \mathbf{E}(X)$ используется оценка $\overline{X}_n = \frac{1}{n}\sum_{l=1}^{n} X_l$, которая называется **выборочным средним**.

Замечание 1. Выборочное среднее является средним значением (математическим ожиданием) для эмпирической функции распределения.

В этом утверждении присутствует необычное наслоение, связанное с понятием "случайный". Эмпирическая функция распределения сама зависит от случая, т. е. каждому значению случайной выборки отвечает своя реализация эмпирической функции распределения. Рассматривая эту реализацию как обычную функцию распределения, мы можем вычислить среднее той случайной величины, которая ей соответствует. Согласно замечанию 4 § 18 эта величина принимает значения (если их упорядочить) $X_{(k)}$, $k = 1, 2, \ldots, n$, с равными вероятностями $\frac{1}{n}$. Среднее же этой дискретной случайной величины по определению (см. § 12) равно $\sum\limits_{k=1}^{n} X_{(k)}\frac{1}{n} = \overline{X}_n$, т. е. равно выборочному среднему.

Теорема 1. *Выборочное среднее $\overline{X}_n$ является несмещенной состоятельной оценкой математического ожидания* $\mathbf{E}(X)$.

Доказательство. Несмещенность является следствием цепочки равенств

$$\mathbf{E}(\overline{X}_n) = \frac{1}{n}\mathbf{E}\Big(\sum_{l=1}^{n} X_l\Big) = \frac{1}{n}\sum_{l=1}^{n}\mathbf{E}(X_l) = \frac{1}{n}\sum_{l=1}^{n} m = m.$$

Состоятельность $\overline{X}_n$ следует из закона больших чисел (§ 16), согласно которому $\overline{X}_n \to m$ по вероятности.

Для выборочного среднего очевидно следующее равенство

$$\overline{X}_n(X_1 + c, \ldots, X_n + c) =$$
$$= \frac{1}{n}\sum_{l=1}^{n}(X_l + c) = c + \overline{X}_n(X_1, X_2, \ldots, X_n), \qquad (19.1)$$

которое можно охарактеризовать как свойство аддитивности выборочного среднего.

2. Выборочная дисперсия. В качестве оценки для дисперсии $\mathbf{D}(X)$ используется оценка

$$S_n^2 = \frac{1}{n-1}\sum_{l=1}^{n}(X_l - \overline{X}_n)^2,$$

которая называется **выборочной дисперсией**.

Выборочная дисперсия характеризует среднеквадратичное отклонение выборочных величин от выборочного среднего.

В определении выборочной дисперсии множитель $\frac{1}{n-1}$ взят вместо множителя $\frac{1}{n}$ для того, чтобы добиться важного свойства несмещенности оценки S_n^2.

Теорема 2. *Выборочная дисперсия S_n^2 является инвариантной относительно сдвига несмещенной состоятельной оценкой дисперсии* $\mathbf{D}(X)$.

Доказательство. Используя равенство (19.1), получим

$$S_n^2(X_1+c,\ldots,X_n+c) =$$
$$= \frac{1}{n-1}\sum_{l=1}^{n}(X_l+c-\overline{X}_n(X_1+c,\ldots,X_n+c))^2 =$$
$$= \frac{1}{n-1}\sum_{l=1}^{n}(X_l-\overline{X}_n(X_1,\ldots,X_n))^2 = S_n^2(X_1,\ldots,X_n). \qquad (19.2)$$

Это доказывает инвариантность оценки S_n^2 относительно сдвига.

Проверим несмещенность. Пусть $m = \mathbf{E}(X_l), \sigma^2 = \mathbf{D}(X_l)$. Положим $Y_l = X_l - m$. Тогда $\mathbf{E}(Y_l) = \mathbf{E}(X_l) - m = 0, \mathbf{D}(Y_l) = \mathbf{D}(X_l) = \sigma^2$. В силу инвариантности относительно сдвига $S_n^2(\vec{X}) = S_n^2(\vec{Y})$. Далее имеем

$$S_n^2(\vec{Y}) = \frac{1}{n-1}\sum_{l=1}^{n}(Y_l-\overline{Y}_n)^2 = \frac{1}{n-1}\sum_{l=1}^{n}(Y_l^2-2Y_l\overline{Y}_n+\overline{Y}_n^2) =$$

$$= \frac{1}{n-1}\sum_{l=1}^{n}Y_l^2 - \frac{2n}{n-1}\overline{Y}_n^2 + \frac{n}{n-1}\overline{Y}_n^2 =$$
$$= \frac{1}{n-1}\sum_{l=1}^{n}Y_l^2 - \frac{n}{n-1}\overline{Y}_n^2. \qquad (19.3)$$

Используя третье свойство математического ожидания (§ 12) и равенства $\mathbf{E}(Y_l) = 0$, получим

$$\mathbf{E}(\overline{Y}_n^2) = \frac{1}{n^2}\mathbf{E}\Big(\sum_{l=1}^{n}Y_l\Big)^2 = \frac{1}{n^2}\mathbf{E}\Big(\sum_{l=1}^{n}Y_l^2 + 2\sum_{l<j}Y_lY_j\Big) =$$

$$= \frac{1}{n^2}\Big(\sum_{l=1}^{n}\mathbf{E}(Y_l^2) + 2\sum_{l<j}\mathbf{E}(Y_l)\mathbf{E}(Y_j)\Big) = \frac{1}{n^2}\sum_{l=1}^{n}\mathbf{E}(Y_l^2) = \frac{\sigma^2}{n}.$$

Теперь, применяя (19.3), найдем, что

$$\mathbf{E}(S_n^2(\vec{X})) = \mathbf{E}(S_n^2(\vec{Y})) =$$
$$= \frac{1}{n-1}\sum_{l=1}^{n}\mathbf{E}(Y_l^2) - \frac{n}{n-1}\mathbf{E}(\overline{Y}_n^2) = \frac{n\sigma^2}{n-1} - \frac{n}{n-1}\cdot\frac{\sigma^2}{n} = \sigma^2.$$

Следовательно, $S_n^2(\vec{X})$ является несмещенной оценкой дисперсии случайной величины X.

Докажем состоятельность. Аналогично (19.3) имеем

$$S_n^2(\vec{X}) = \frac{1}{n-1}\sum_{l=1}^{n}X_l^2 - \frac{n}{n-1}\overline{X}_n^2 =$$
$$= \frac{n}{n-1}\Big(\frac{1}{n}\sum_{l=1}^{n}X_l^2 - \overline{X}_n^2\Big). \qquad (19.4)$$

В силу состоятельности выборочного среднего $\overline{X}_n \to \mathbf{E}(X)$ по вероятности. Поскольку квадраты независимых одинаково распределенных случайных величин являются независимыми одинаково распределенными величинами, то по закону больших чисел (§ 16) имеем

$$\frac{1}{n}\sum_{l=1}^{n}X_l^2 \to \mathbf{E}(X^2)$$

по вероятности. Так как $\dfrac{n}{n-1} \to 1$, то из (19.4) следует, что

$$S_n^2(\vec{X}) \to \mathbf{E}(X^2) - \mathbf{E}^2(X) = \mathbf{D}(X)$$

по вероятности, а это и означает состоятельность оценки $S_n^2(\vec{X})$.

Замечание 2. Для различных вычислений часто бывает удобней пользоваться следующей формулой для выборочной дисперсии:

$$S_n^2(\vec{X}) = \frac{1}{n-1}\sum_{l=1}^{n}X_l^2 - \frac{n}{n-1}\overline{X}_n^2.$$

Эта формула содержится в (19.4).

3. Выборочные моменты. **Выборочным моментом порядка k**, построенным по выборке $\vec{X}$, называется величина

$$\overline{X}_n^{(k)} = \frac{1}{n}\sum_{l=1}^{n} X_l^k.$$

Аналогично случаю с выборочным средним легко убедиться, что выборочный момент порядка k является моментом порядка k для эмпирической функции распределения.

Теорема 3. *Выборочный момент $\overline{X}_n^{(k)}$ является несмещенной состоятельной оценкой момента m_k.*

Доказательство. Несмещенность следует из равенств

$$\mathbf{E}(\overline{X}_n^{(k)}) = \frac{1}{n}\mathbf{E}\Big(\sum_{l=1}^{n} X_l^k\Big) = \frac{1}{n}\sum_{l=1}^{n}\mathbf{E}(X_l^k) = \mathbf{E}(X^k) = m_k.$$

Состоятельность непосредственно вытекает из закона больших чисел: при $n \to \infty$

$$\overline{X}_n^{(k)} = \frac{1}{n}\sum_{l=1}^{n} X_l^k \to \mathbf{E}(X^k)$$

по вероятности.

Выборочным центральным моментом порядка $k, k \geqslant 2$, построенным по выборке $\vec{X}$, называется величина

$$\overline{S}_n^{(k)} = \frac{1}{n}\sum_{l=1}^{n}(X_l - \overline{X}_n)^k.$$

Теорема 4. *Выборочный центральный момент $\overline{S}_n^{(k)}$ является инвариантной относительно сдвига состоятельной оценкой центрального момента μ_k.*

Доказательство. Инвариантность относительно сдвига оценки $\overline{S}_n^{(k)}(\vec{X})$ устанавливается точно так же, как инвариантность оценки $S_n^2(\vec{X})$ (см. доказательство теоремы 2). Для доказательства состоятельности воспользуемся равенством

$$a^k - b^k = (a - b)\sum_{m=0}^{k-1} a^m b^{k\ \ 1-m}.$$

В силу этого равенства

$$\frac{1}{n}\sum_{l=1}^{n}(X_l-\overline{X}_n)^k-\frac{1}{n}\sum_{l=1}^{n}(X_l-\mathbf{E}(X))^k=$$
$$=(\overline{X}_n-\mathbf{E}(X))\frac{1}{n}\sum_{l=1}^{n}\sum_{m=0}^{k-1}(X_l-\overline{X}_n)^m(X_l-\mathbf{E}(X))^{k-1-m}=$$
$$=(\overline{X}_n-\mathbf{E}(X))\sum_{m=0}^{k-1}\frac{1}{n}\sum_{l=1}^{n}(X_l-\overline{X}_n)^m(X_l-\mathbf{E}(X))^{k-1-m}. \tag{19.5}$$

Применяя оценки (12 1), (12.2) и свойства 1), 2) математического ожидания (§ 12), получим

$$\mathbf{E}\left(\left|\sum_{l=1}^{n}(X_l-\overline{X}_n)^m(X_l-\mathbf{E}(X))^{k-1-m}\right|\right)\leqslant$$
$$\leqslant\sum_{l=1}^{n}\mathbf{E}\left(\left|(X_l-\overline{X}_n)^m(X_l-\mathbf{E}(X))^{k-1-m}\right|\right)\leqslant$$
$$\leqslant n\max_{1\leqslant l\leqslant n}\left\{\left(\mathbf{E}(X_l-\overline{X}_n)^{2m}\right)^{1/2}\left(\mathbf{E}(X_l-\mathbf{E}(X))^{2k-2-2m}\right)^{1/2}\right\}.$$

В силу состоятельности выборочного среднего $\overline{X}_n-\mathbf{E}(X)\to 0$ по вероятности. С помощью прямых вычислений можно проверить, что математические ожидания квадратов случайных величин $(X_l-\overline{X}_n)^m$ и $(X_l-\mathbf{E}(X))^{k-1-m}$ конечны и не зависят от l. Таким образом, в правой части (19.5) стоит величина, равная произведению величины, стремящейся по вероятности к нулю, на величину, абсолютный момент которой равномерно ограничен по n. Можно доказать, что произведение таких величин стремится к нулю по вероятности. Тогда из (19.5) следует, что для того, чтобы доказать состоятельность выборочного центрального момента $\overline{S}_n^{(k)}$, достаточно убедиться, что

$$\frac{1}{n}\sum_{l=1}^{n}(X_l-\mathbf{E}(X))^k\to\mathbf{E}(X-\mathbf{E}(X))^k$$

по вероятности. Этот факт следует из закона больших чисел, так как величины $(X_l-\mathbf{E}(X))^k$ являются независимыми и одинаково распределенными.

4. Выборочный коэффициент корреляции. В § 18 дано определение случайной выборки объема n, отвечающей одной случайной величине. Аналогичное понятие можно ввести для двух и более случайных величин.

Определение 4. **Случайной выборкой объема** n, отвечающей паре случайных величин (X, Y), называется набор n независимых одинаково распределенных пар случайных величин $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, каждая из которых имеет такое же совместное распределение, как и пара величин (X, Y).

Оценкой для $\mathrm{cov}(X, Y)$, построенной по выборке $(\vec{X}, \vec{Y}) = ((X_1, Y_1), \ldots, (X_n, Y_n))$, служит **выборочная ковариация**, определяемая по формуле

$$C_n(\vec{X}, \vec{Y}) = \frac{1}{n-1} \sum_{l=1}^{n} (X_l - \overline{X}_n)(Y_l - \overline{Y}_n). \tag{19.6}$$

В качестве оценки для коэффициента корреляции $r(X, Y)$ используется **выборочный коэффициент корреляции**, определяемый по формуле

$$R_n(\vec{X}, \vec{Y}) = \frac{C_n(\vec{X}, \vec{Y})}{\sqrt{S_n^2(\vec{X}) S_n^2(\vec{Y})}} = \frac{\sum\limits_{l=1}^{n} (X_l - \overline{X}_n)(Y_l - \overline{Y}_n)}{\sqrt{\sum\limits_{l=1}^{n} (X_l - \overline{X}_n)^2 \sum\limits_{l=1}^{n} (Y_l - \overline{Y}_n)^2}}.$$

Для выборочной ковариации и для выборочного коэффициента корреляции несложно получить следующие выражения:

$$C_n(\vec{X}, \vec{Y}) = \frac{1}{n-1} \sum_{l=1}^{n} X_l Y_l - \left(\frac{n}{n-1}\right) \overline{X}_n \overline{Y}_n, \tag{19.7}$$

$$R_n(\vec{X}, \vec{Y}) = \frac{\frac{1}{n} \sum\limits_{l=1}^{n} X_l Y_l - \overline{X}_n \overline{Y}_n}{\sqrt{\left(\frac{1}{n} \sum\limits_{l=1}^{n} X_l^2 - \overline{X}_n^2\right)\left(\frac{1}{n} \sum\limits_{l=1}^{n} Y_l^2 - \overline{Y}_n^2\right)}}. \tag{19.8}$$

Теорема 5. *Выборочная ковариация* $C_n(\vec{X},\vec{Y})$ *является инвариантной относительно сдвига несмещенной состоятельной оценкой ковариации* $\operatorname{cov}(X,Y)$.

Доказательство. Используя формулу (19.6) и выкладки, аналогичные (19.1), (19.2), несложно проверить, что для любых c_1, c_2 справедливо равенство

$$C_n(\vec{X}+\vec{c}_1, \vec{Y}+\vec{c}_2) = C_n(\vec{X},\vec{Y}),$$

где $\vec{c}_1 = (c_1, c_1, \ldots, c_1), \vec{c}_2 = (c_2, c_2, \ldots, c_2)$ — n-мерные вектора с одинаковыми компонентами. Это равенство и означает, что выборочная ковариация $C_n(\vec{X},\vec{Y})$ инвариантна относительно сдвига.

Докажем несмещенность оценки (19.7). В силу инвариантности относительно сдвига можно считать, что X_l и Y_l имеют нулевые средние ($\mathbf{E}X_l = \mathbf{E}Y_l = 0$), иначе из этих величин можно вычесть математические ожидания, не изменив оценку. При нулевых средних

$$\mathbf{E}(\overline{X}_n\overline{Y}_n) = \frac{1}{n^2}\sum_{l=1}^{n}\mathbf{E}(X_lY_l) = \frac{1}{n}\mathbf{E}(XY).$$

Теперь несложно убедиться в несмещенности оценки:

$$\mathbf{E}C_n(\vec{X},\vec{Y}) = \frac{1}{n-1}\sum_{l=1}^{n}\mathbf{E}(X_lY_l) - \left(\frac{n}{n-1}\right)\mathbf{E}(\overline{X}_n\overline{Y}_n) =$$

$$= \frac{n}{n-1}\mathbf{E}(XY) - \frac{1}{n-1}\mathbf{E}(XY) = \mathbf{E}(XY) = \operatorname{cov}(X,Y).$$

Последнее равенство верно, поскольку мы предположили, что величины X и Y имеют нулевые средние.

Для доказательства состоятельности воспользуемся формулой (19.7). Поскольку двумерные величины (X_l, Y_l) являются независимыми и одинаково распределенными, то произведения их координат X_lY_l будут независимыми одинаково распределенными величинами. Применим закон больших чисел (§ 16), согласно которому

$$\frac{1}{n}\sum_{l=1}^{n}X_lY_l \to \mathbf{E}(XY)$$

по вероятности. Так как выборочное среднее является состоятельной оценкой, то имеем $\overline{X}_n \to \mathbf{E}(X), \overline{Y}_n \to \mathbf{E}(Y)$ по вероятности. Используя эти предельные соотношения в формуле (19.7), получим, что

$$C_n(\vec{X},\vec{Y}) = \frac{n}{n-1}\left(\frac{1}{n}\sum_{l=1}^{n} X_l Y_l - \overline{X}_n\overline{Y}_n\right) \to \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$$

по вероятности. Правая часть этого соотношения равна $\mathrm{cov}(X,Y)$. Следовательно, $C_n(\vec{X},\vec{Y})$ является состоятельной оценкой ковариации.

Теорема 6. *Выборочный коэффициент корреляции $R_n(\vec{X},\vec{Y})$ является инвариантной относительно сдвига состоятельной оценкой коэффициента корреляции $r(X,Y)$.*

Доказательство. Поскольку выборочный коэффициент корреляции равен отношению выборочной ковариации к корню из произведения выборочных дисперсий, то инвариантность относительно сдвига и состоятельность этой оценки следует из аналогичных свойств выборочной ковариации и выборочных дисперсий.

Задача 1. При обработке данных 15 испытаний спортивного самолета были получены следующие значения его максимальной скорости: 422.2; 418.7; 425.6; 420.3; 425.8; 423.1; 431.5; 428.2; 438.3; 434.0; 411.3; 417.2; 413.5; 441.3; 420.0 м/сек. Определить несмещенные оценки математического ожидания и дисперсии максимальной скорости самолета.

Решение. Для того чтобы упростить вычисления, воспользуемся свойством инвариантности относительно сдвига для выборочной дисперсии, а для выборочного среднего используем равенство (19.1). Тогда вместо исходной реализации выборки можно рассмотреть, например, реализацию $\vec{x}$: 2.2; −1.3; 5.6; 0.3; 5.8; 3.1; 11.5; 8.2; 18.3; 14.0; −8.7; −2.8; −6.5; 21.3; 0.0, которая отличается от исходной на значение $c = 420$. Выборочное среднее этой реализации равно $\bar{x}_{15} = \frac{1}{15}\sum_{l=1}^{15} x_l = 4.73$. Для вычисления выборочной дисперсии воспользуемся равенством (19.4). Имеем

$s^2_{15} = \frac{1}{14}\sum\limits_{l=1}^{15} x_l^2 - \frac{15}{14}(\bar{x}_{15})^2 = \frac{1}{14}\cdot 1391.08 - \frac{15}{14}(4.73)^2 \approx 75.392$. Следовательно, оценка математического ожидания максимальной скорости спортивного самолета равна 424.73 м/сек. Несмещенная оценка дисперсии приблизительно равна 75.392 $м^2/сек^2$.

Задачи

Задача 19.1. Произведено 16 измерений начальной скорости снаряда. Результаты измерений (в м/с) следующие: 1235.6; 1237.5; 1232.9; 1236.2; 1238.5; 1234.2; 1235.9; 1233.3; 1234.5; 1236.8; 1237.6; 1233.1; 1234.3; 1237.5; 1235.4; 1234.7. Вычислить оценки математического ожидания и дисперсии начальной скорости снаряда.

Задача 19.2. На телефонной станции производились наблюдения за числом неправильных соединений в минуту. Наблюдения в течение часа дали следующие результаты:

3,	1,	3,	4,	2,	1,	1,	3,	2,	7,	2,	0,	1,	2,	1,
2,	4,	0,	3,	0,	2,	0,	1,	3,	3,	1,	2,	0,	3,	4,
2,	0,	2,	1,	4,	3,	4,	2,	0,	2,	3,	1,	1,	2,	2,
3,	1,	4,	2,	2,	1,	2,	5,	1,	1,	0,	1,	1,	1,	5.

Оценить среднее и дисперсию числа неправильных соединений.

Задача 19.3. Данные ежедневных измерений температуры в течение месяца представлены в следующей таблице:

15.3	16.5	15.8	14.7	13.9	12.2	12.6	12.8	12.3	14.5
14.2	16.7	15.4	16.9	17.1	16.6	13.7	13.4	14.3	18.1
17.4	17.7	18.5	17.9	19.4	20.4	21.2	19.8	19.2	20.5

Оценить среднее значение температуры за месяц и среднеквадратичное отклонение.

Задача 19.4. Ошибки 40 измерений приведены в таблице задачи 18.2. Найти оценки математического ожидания и дисперсии случайной величины, характеризующей ошибки измерений.

Задача 19.5. Сырье, поступающее на завод из карьера, содержит два полезных компонента – минералы A и B. Результаты анализов десяти образцов сырья, поступившего в разное время из разных мест карьера, приведены в таблице, где x и y – выборочные значения случайных величин X и Y, выражающих соответственно процентное содержание минералов A и B в образцах.

x	67	54	72	64	39	22	58	43	46	34
y	24	15	23	19	16	11	20	16	17	13

Оценить коэффициент корреляции величин X и Y.

§20. Асимптотические свойства выборочных моментов

Пусть θ – некоторый параметр распределения случайной величины X, значение которого нам неизвестно, и $\gamma_n(\vec{X}) = \gamma_n(X_1, X_2, \ldots, X_n)$ – оценка параметра θ, построенная по выборке $\vec{X}$. Хорошая оценка, естественно, должна быть состоятельной, т. е. $\gamma_n(\vec{X})$ должно сходиться к θ при неограниченном возрастании n. В этом случае разность $\gamma_n(\vec{X}) - \theta$ стремится к нулю. Естественно ожидать, что, домножив ее на некоторый возрастающий множитель, можно получить в пределе случайную величину, неравную постоянной. Оказывается, и это является следствием центральной предельной теоремы, что для большинства оценок таким множителем является $\sqrt{n}$.

Определение 1. Оценка $\gamma_n(\vec{X})$ называется **асимптотически нормальной** с дисперсией Δ^2, если функция распределения случайной величины $\frac{\sqrt{n}}{\Delta}(\gamma_n(\vec{X}) - \theta)$ сходится при $n \to \infty$ к функции распределения стандартного нормального закона.

Исследуем условия асимптотической нормальности конкретных оценок. Как и в предыдущем параграфе мы предположим, что если рассматривается оценка момента порядка k или центрального момента порядка k, то у случайной величины X конечен момент порядка $2k$.

1. Выборочное среднее.

Теорема 1. *Выборочное среднее $\overline{X}_n$ асимптотически нормально с дисперсией $\Delta^2 = \mathbf{D}(X)$.*

Доказательство. Воспользуемся центральной предельной теоремой (§ 17). Обозначим $m = \mathbf{E}(X_l)$, $\Delta^2 = \mathbf{D}(X_l)$,

$S_n = \sum_{l=1}^{n} X_l$. Величина $\overline{X}_n$ служит оценкой для m, и нам нужно убедиться в том, что нормированная разность $\frac{\sqrt{n}}{\Delta}(\overline{X}_n - m)$ распределена в пределе по нормальному закону со средним 0 и дисперсией 1. Имеем

$$\frac{\sqrt{n}}{\Delta}(\overline{X}_n - m) = \frac{\sqrt{n}}{\Lambda}\left(\frac{1}{n}\sum_{l=1}^{n} X_l - m\right) = \frac{S_n - mn}{\Delta\sqrt{n}}.$$

Поскольку величины X_l независимы и одинаково распределены со средним m и дисперсией Δ^2, то, согласно центральной предельной теореме, распределение величины $Z_n = \frac{1}{\Delta\sqrt{n}}(S_n - mn)$ сходится при $n \to \infty$ к стандартному нормальному распределению, что и требовалось доказать.

2. Выборочный момент порядка k. Выборочный момент $\overline{X}_n^{(k)} = \frac{1}{n}\sum_{l=1}^{n} X_l^k$ является оценкой для $m_k = \mathbf{E}(X^k)$.

Теорема 2. *Выборочный момент $\overline{X}_n^{(k)}$ асимптотически нормален с дисперсией $\Delta^2 = m_{2k} - m_k^2$.*

Доказательство. Применим центральную предельную теорему для величин $Y_l = X_l^k$. Тогда $S_n = \sum_{l=1}^{n} Y_l$ и $\mathbf{E}(Y_l) = \mathbf{E}(X_l^k) = m_k$, $\mathbf{D}(Y_l) = \mathbf{E}(Y_l^2) - \mathbf{E}^2(Y_l) = \mathbf{E}(X_l^{2k}) - \mathbf{E}^2(X_l^k) = m_{2k} - m_k^2 = \Delta^2$. Имеем

$$\frac{\sqrt{n}}{\Delta}(\overline{X}_n^{(k)} - m_k) = \frac{\sqrt{n}}{\Delta}\left(\frac{1}{n}\sum_{l=1}^{n} X_l^k - m_k\right) = \frac{S_n - m_k n}{\Delta\sqrt{n}}.$$

Величины Y_l независимы, поскольку независимыми являются величины X_l, и одинаково распределены со средним m_k и дисперсией $\Delta^2 = m_{2k} - m_k^2$. Согласно центральной предельной теореме распределение величины $Z_n = \frac{1}{\Delta\sqrt{n}}(S_n - m_k n)$ сходится при $n \to \infty$ к стандартному нормальному распределению.

3. Выборочная дисперсия. Пусть $\mu_k = \mathbf{E}(X - m)^k$ – k-й центральный момент случайной величины X со средним m. Заметим, что $\mu_2 = \sigma^2 = \mathbf{D}(X)$, т. е. второй центральный момент совпадает с дисперсией. Выборочная дисперсия S_n^2 является оценкой для дисперсии σ^2.

Теорема 3. *Выборочная дисперсия S_n^2 асимптотически нормальна с дисперсией $\Delta^2 = \mu_4 - \mu_2^2$.*

Доказательство. Имеет место следующее представление

$$\sqrt{n}(S_n^2 - \sigma^2) = \sqrt{n}\Big(\frac{1}{n-1}\sum_{l=1}^{n}(X_l - \overline{X}_n)^2 - \sigma^2\Big) =$$
$$= \Big(\frac{n}{n-1}\Big)\frac{1}{\sqrt{n}}\sum_{l=1}^{n}\Big((X_l - \overline{X}_n)^2 - \Big(1 - \frac{1}{n}\Big)\sigma^2\Big).$$

Множитель $n/(n-1)$ на предельное поведение не влияет, поскольку $n/(n-1) \to 1$. Далее, так как $\overline{X}_n \to m$ со скоростью $1/\sqrt{n}$ в том смысле, что среднеквадратичное отклонение имеет порядок $1/n$ (см. (16.5)), то можно доказать, что предельное поведение нормированной суммы в правой части не изменится, если в ней слагаемые заменить на $Y_l = (X_l - m)^2 - \sigma^2$. Это доказательство довольно громоздкое, и мы его опускаем. Таким образом, предельное поведение нормированной разности $\sqrt{n}(S_n^2 - \sigma^2)$ совпадает с предельным поведением величины

$$Z_n = \frac{1}{\sqrt{n}}\sum_{l=1}^{n} Y_l.$$

Случайные величины Y_l независимы, одинаково распределены и

$$\mathbf{E}(Y_l) = \mathbf{E}(X_l - m)^2 - \sigma^2 = 0,$$
$$\mathbf{D}(Y_l) = \mathbf{D}((X_l - m)^2) = \mathbf{E}(X_l - m)^4 - \mathbf{E}^2(X_l - m)^2 = \mu_4 - \mu_2^2.$$

По центральной предельной теореме распределение величины $Z_n/\sqrt{\mu_4 - \mu_2^2}$ сходится при $n \to \infty$ к стандартному нормальному распределению. Аналогичным будет и предельное поведение распределения величины $\sqrt{n}(S_n^2 - \sigma^2)/\sqrt{\mu_4 - \mu_2^2}$. Это и требовалось доказать.

В следующем параграфе свойство асимптотической нормальности будет использовано при построении граничных значений для оцениваемого параметра.

§ 21. Доверительные интервалы

При решении некоторых практических задач вместо оценки неизвестного параметра распределения случайной величины важнее знать границы, в которых этот параметр находится. Границы интервала, содержащего неизвестный параметр, строятся по выборке. При небольшом объеме выборки, как правило, не удается указать верхнюю и нижнюю границы, достаточно близкие к параметру с большой вероятностью. Чем больше компонент в выборке, тем более точные границы с более близкой к единице вероятностью можно построить. Пусть θ – неизвестный параметр распределения случайной величины X, и $\vec{X} = (X_1, X_2, \ldots, X_n)$ – случайная выборка, отвечающая величине X. Пусть α – некоторое число между нулем и единицей, а $\underline{\theta}(\vec{X}), \bar{\theta}(\vec{X})$ – две функции от случайной выборки $\vec{X}$ такие, что $\underline{\theta}(\vec{X}) \leqslant \bar{\theta}(\vec{X})$.

Определение 1. Интервал $(\underline{\theta}(\vec{X}), \bar{\theta}(\vec{X}))$ называется **доверительным интервалом** для оценки параметра θ, отвечающим **доверительной вероятности** α, если

$$\mathbf{P}(\underline{\theta}(\vec{X}) < \theta < \bar{\theta}(\vec{X})) \geqslant \alpha.$$

Величина $\underline{\theta}(\vec{X})$ называется нижней доверительной границей, а $\bar{\theta}(\vec{X})$ – верхней доверительной границей для параметра θ.

Иными словами, доверительным называется такой интервал, который с наперед заданной вероятностью содержит оцениваемый параметр. Границы $\underline{\theta}(\vec{X})$ и $\bar{\theta}(\vec{X})$ кроме наблюдений $\vec{X}$ будут зависеть от α и от числа наблюдений n.

Рассмотрим методы построения доверительных интервалов для конкретных параметров.

1. Доверительные интервалы для неизвестного математического ожидания $m = \mathbf{E}(X)$ при известной дисперсии $\sigma^2 = \mathbf{D}(X)$.

Согласно свойству асимптотической нормальности выборочного среднего $\overline{X}_n$ имеем

$$\mathbf{P}\left(\frac{\sqrt{n}}{\sigma}(\overline{X}_n - m) < x\right) \approx \mathcal{N}(x).$$

Тогда для любого $z > 0$

$$\mathbf{P}\left(-z < \frac{\sqrt{n}}{\sigma}(\overline{X}_n - m) < z\right) \approx \mathcal{N}(z) - \mathcal{N}(-z).$$

Очевидно, что неравенства

$$-z < \frac{\sqrt{n}}{\sigma}(\overline{X}_n - m) < z$$

эквивалентны неравенствам

$$\overline{X}_n - \frac{z\sigma}{\sqrt{n}} < m < \overline{X}_n + \frac{z\sigma}{\sqrt{n}}.$$

Поэтому, используя формулу $\mathcal{N}(x) = \frac{1}{2} + \frac{1}{2}\Phi(x)$, выражающую функцию распределения стандартного нормального закона через функцию Лапласа $\Phi(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int\limits_0^x e^{-v^2/2} dv$, получим

$$\mathbf{P}\left(\overline{X}_n - \frac{z\sigma}{\sqrt{n}} < m < \overline{X}_n + \frac{z\sigma}{\sqrt{n}}\right) \approx \frac{1}{2}(\Phi(z) - \Phi(-z)) = \Phi(z).$$

В последнем равенстве мы использовали нечетность функции Лапласа $\Phi(x)$. Для значений функции Лапласа и ее квантилей существуют таблицы. Поэтому, если в качестве z взять z_α – квантиль порядка α функции $\Phi(x)$, то окончательно будем иметь

$$\mathbf{P}\left(\overline{X}_n - \frac{z_\alpha\sigma}{\sqrt{n}} < m < \overline{X}_n + \frac{z_\alpha\sigma}{\sqrt{n}}\right) \approx \Phi(z_\alpha) = \alpha. \tag{21.1}$$

Сравнивая это с определением доверительного интервала, мы видим, что случайный интервал $\left(\overline{X}_n - \frac{z_\alpha\sigma}{\sqrt{n}}, \overline{X}_n + \frac{z_\alpha\sigma}{\sqrt{n}}\right)$ является доверительным интервалом для математического ожидания m с доверительной вероятностью, приблизительно равной α. Для того, чтобы приближенные равенства были более точными, необходимо, чтобы n было велико. Таблица значений функции Лапласа позволяет по заданной величине α найти значение z_α. Выборочное среднее $\overline{X}_n = \frac{1}{n}\sum\limits_{l=1}^{n} X_l$ вычисляется по наблюдениям выборки $\vec{X}$, σ – известно, n – известно, поэтому полностью определен и доверительный интервал, указанный выше.

2. Доверительные интервалы для неизвестного математического ожидания $m = \mathbf{E}(X)$ при неизвестной дисперсии $\mathbf{D}(X)$.

Можно поступить аналогично предыдущему случаю, когда дисперсия σ^2 была известна. Однако поскольку теперь она неизвестна, то ее следует предварительно оценить. В качестве оценки для σ^2 возьмем выборочную дисперсию

$$S_n^2 = \frac{1}{n-1}\sum_{l=1}^{n}(X_l - \overline{X}_n)^2.$$

Поскольку $S_n^2 \to \sigma^2$ при $n \to \infty$ по вероятности, то можно доказать, что если в приближенном равенстве (21.1) вместо σ использовать $\sqrt{S_n^2}$, то мы получим аналогичное приближенное равенство

$$\mathbf{P}\left(\overline{X}_n - \frac{z_\alpha\sqrt{S_n^2}}{\sqrt{n}} < m < \overline{X}_n + \frac{z_\alpha\sqrt{S_n^2}}{\sqrt{n}}\right) \approx \alpha.$$

Таким образом, случайный интервал

$$\left(\overline{X}_n - \frac{z_\alpha\sqrt{S_n^2}}{\sqrt{n}}, \overline{X}_n + \frac{z_\alpha\sqrt{S_n^2}}{\sqrt{n}}\right), \tag{21.2}$$

где z_α – квантиль порядка α функции $\Phi(x)$, является доверительным интервалом для математического ожидания при неизвестной дисперсии с доверительной вероятностью, приблизительно равной α.

Пусть можно считать, что исходная случайная величина X, математическое ожидание которой мы оцениваем, распределена по нормальному закону с неизвестной дисперсией. Тогда можно построить доверительный интервал для математического ожидания m с доверительной вероятностью, точно равной α. Кроме того, при таком предположении, если оно справедливо, число наблюдений может быть небольшим потому, что при построении доверительного интервала не будут использоваться асимптотические формулы.

Для случайной выборки $\vec{X}$, отвечающей нормально распределенной случайной величине, положим

$$\mathcal{S}_{n-1}(x) = \mathbf{P}\left(\frac{\sqrt{n}}{\sqrt{S_n^2}}(\overline{X}_n - m) < x\right), \qquad x \in (-\infty, \infty).$$

Распределение $\mathcal{S}_k(x)$ – так называемое **распределение Стьюдента** с k степенями свободы, для которого есть таблицы. Плотность этого распределения задается выражением

$$s_k(x) = \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\Gamma(k/2)}\left(1+\frac{x^2}{k}\right)^{-(k+1)/2}, \quad x \in (-\infty, \infty).$$

Поскольку $\mathcal{S}_k(-x) = 1 - \mathcal{S}_k(x)$, то для любого $z > 0$

$$\mathbf{P}\left(-z < \frac{\sqrt{n}}{\sqrt{S_n^2}}(\overline{X}_n - m) < z\right) = \\ = \mathcal{S}_{n-1}(z) - \mathcal{S}_{n-1}(-z) = 2\mathcal{S}_{n-1}(z) - 1,$$

и, следовательно,

$$\mathbf{P}\left(\overline{X}_n - \frac{z\sqrt{S_n^2}}{\sqrt{n}} < m < \overline{X}_n + \frac{z\sqrt{S_n^2}}{\sqrt{n}}\right) = 2\mathcal{S}_{n-1}(z) - 1.$$

Существуют таблицы, с помощью которых по значениям α и $n-1$ можно определить такое число $z_{\alpha,n-1}$, что $2\mathcal{S}_{n-1}(z_{\alpha,n-1}) - 1 = \alpha$. Величина $z_{\alpha,k}$ является квантилью порядка $(1+\alpha)/2$ распределения Стьюдента с k степенями свободы, поскольку $\mathcal{S}_k(z_{\alpha,k}) = (1+\alpha)/2$.

Таким образом, случайный интервал

$$\left(\overline{X}_n - \frac{z_{\alpha,n-1}\sqrt{S_n^2}}{\sqrt{n}}, \overline{X}_n + \frac{z_{\alpha,n-1}\sqrt{S_n^2}}{\sqrt{n}}\right) \tag{21.3}$$

является доверительным интервалом с доверительной вероятностью равной α для математического ожидания m нормально распределенной случайной величины при неизвестной дисперсии.

3. Доверительные интервалы для неизвестной дисперсии $\sigma^2 = D(X)$ при неизвестных других параметрах.

В силу свойства асимптотической нормальности выборочной дисперсии имеем

$$\mathbf{P}\left(\frac{\sqrt{n}}{\sqrt{\mu_4 - \mu_2^2}}(S_n^2 - \sigma^2) < x\right) \approx \mathcal{N}(x).$$

Вместо неизвестных параметров μ_4 и μ_2 возьмем их оценки

$$\overline{S}_n^{(4)} = \frac{1}{n}\sum_{l=1}^{n}(X_l - \overline{X}_n)^4, \qquad S_n^2 = \frac{1}{n-1}\sum_{l=1}^{n}(X_l - \overline{X}_n)^2.$$

Тогда в силу состоятельности этих оценок выполняется приближенное равенство

$$\mathbf{P}\left(\frac{\sqrt{n}(S_n^2 - \sigma^2)}{\sqrt{\overline{S}_n^{(4)} - (S_n^2)^2}} < x\right) \approx \mathcal{N}(x).$$

Отсюда, как и в пункте 1, получим, что

$$\mathbf{P}\left(S_n^2 - \frac{z_\alpha}{\sqrt{n}}\sqrt{\overline{S}_n^{(4)} - (S_n^2)^2} < \sigma^2 < S_n^2 + \frac{z_\alpha}{\sqrt{n}}\sqrt{\overline{S}_n^{(4)} - (S_n^2)^2}\right) \approx$$
$$\approx \Phi(z_\alpha) = \alpha,$$

где z_α – квантиль порядка α функции Лапласа $\Phi(x)$.

Следовательно, в качестве доверительного интервала для дисперсии σ^2 можно использовать

$$\left(S_n^2 - \frac{z_\alpha}{\sqrt{n}}\sqrt{\overline{S}_n^{(4)} - (S_n^2)^2}, S_n^2 + \frac{z_\alpha}{\sqrt{n}}\sqrt{\overline{S}_n^{(4)} - (S_n^2)^2}\right). \qquad (21.4)$$

Задача 1. Пусть наблюдается следующая реализация выборки, состоящей из 40 компонент:

−1.80	−2.01	−1.63	0.54	0.25	−0.16	0.03	0.07
−1.18	1.18	1.11	0.88	1.09	−0.20	0.15	−0.37
0.65	−1.14	1.15	−1.21	−0.92	0.42	0.29	−0.90
−0.43	0.35	−1.93	0.89	−0.22	0.60	0.87	−0.43
−1.39	−0.23	0.38	−0.64	−0.57	0.23	−0.28	0.51

Вычислить выборочное среднее и выборочную дисперсию. Построить доверительный интервал для математического ожидания с доверительной вероятностью, приблизительно равной 0.9.

Решение. Используя формулы

$$\overline{X}_n = \frac{1}{n}\sum_{l=1}^{n} X_l, \qquad S_n^2 = \frac{1}{n-1}\sum_{l=1}^{n} X_l^2 - \left(\frac{n}{n-1}\right)\overline{X}_n^2,$$

получим следующие численные значения выборочного среднего и выборочной дисперсии: $\bar{x}_{40} = \frac{1}{40} \cdot (-6.00) = -0.15$; $s^2_{40} = \frac{1}{39} \cdot 31.1328 - \frac{40}{39} \cdot 0.0225 = \frac{1}{39} \cdot (31.1328 - 0.9) \approx 0.80$ (для обозначения реализаций используются соответствующие строчные буквы).

Так как дисперсия неизвестна, то мы используем следующую формулу для доверительного интервала, основанную на реализации случайной выборки:

$$\left(\bar{x}_n - \frac{z_\alpha\sqrt{s_n^2}}{\sqrt{n}}, \bar{x}_n + \frac{z_\alpha\sqrt{s_n^2}}{\sqrt{n}}\right),$$

где z_α – квантиль порядка $\alpha = 0.9$ функции Лапласа $\Phi(x)$. Из таблиц для функции Лапласа можно найти, что $z_{0.9} \approx 1.684$. Тогда

$$\frac{z_{0.9}}{\sqrt{40}}\sqrt{s^2_{40}} \approx \frac{1.684}{2\sqrt{10}}\sqrt{0.8} \approx 0.24, \quad \bar{x}_{40} \pm \frac{z_{0.9}}{\sqrt{40}}\sqrt{s^2_{40}} \approx -0.15 \pm 0.24.$$

Следовательно, доверительный интервал для математического ожидания имеет вид $(-0.39, 0.09)$. Таким образом, для данной выборки наблюдений с вероятностью, приблизительно равной 0.9 математическое ожидание m удовлетворяет неравенству $-0.39 < m < 0.09$.

Задачи

Задача 21.1. По реализации выборки, приведенной в задаче 18.2, вычислить доверительный интервал для математического ожидания с доверительной вероятностью 0.95.

Задача 21.2. По результатам измерений, приведенных в задаче 19.1, требуется: а) вычислить доверительный интервал для математического ожидания начальной скорости с доверительной вероятностью 0.9; б) вычислить доверительный интервал для дисперсии с доверительной вероятностью 0.92.

Задача 21.3. Произведено пять независимых измерений для определения заряда электрона. Получены следующие значения (в кулонах): $1.594 \cdot 10^{-19}, 1.597 \cdot 10^{-19}, 1.596 \cdot 10^{-19}, 1.593 \cdot 10^{-19}, 1.590 \cdot 10^{-19}$. Предполагая, что ошибки измерения распределены по нормальному закону, оценить величину заряда электрона и найти для нее доверительные границы при доверительной вероятности 0.99.

§ 22. Неравенство Рао–Крамера

Меру качества оценки можно выбирать по-разному. Наиболее часто используемой мерой качества служит среднеквадратичное отклонение оценки от оцениваемого параметра. Чем меньше величина среднеквадратичного отклонения $\mathbf{E}(\gamma_n(\vec{X}) - \theta)^2$, тем оценка $\gamma_n(\vec{X})$ параметра θ лучше. При определенных условиях существует нижняя граница для величины среднеквадратичного отклонения, которую улучшить нельзя. Вычислим эту границу.

Пусть случайная величина X имеет плотность распределения $f_\theta(x), x \in (-\infty, \infty)$, которая является дифференцируемой по θ. Поскольку величины $X_1, X_2, \ldots, X_n$ независимы, то согласно формуле (11.3) совместная плотность распределения случайного вектора $\vec{X} = (X_1, X_2, \ldots, X_n)$ имеет вид

$$f_\theta(\vec{x}) = f_\theta(x_1) f_\theta(x_2) \ldots f_\theta(x_n). \qquad (22.1)$$

Предположим, что оценка $\gamma_n(\vec{X})$ является несмещенной. По определению это означает, что

$$\mathbf{E}(\gamma_n(\vec{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \gamma_n(\vec{x}) f_\theta(\vec{x}) d\vec{x} = \theta.$$

Продифференцировав это равенство по θ, получим

$$1 = \frac{\partial}{\partial\theta} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \gamma_n(\vec{x}) f_\theta(\vec{x}) d\vec{x} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \gamma_n(\vec{x}) \frac{\partial}{\partial\theta} f_\theta(\vec{x}) d\vec{x} =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \gamma_n(\vec{x}) \frac{\partial}{\partial\theta} \big(\ln f_\theta(\vec{x})\big) f_\theta(\vec{x}) d\vec{x} = \mathbf{E}\Big(\gamma_n(\vec{X}) \frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\Big). \qquad (22.2)$$

В силу первого свойства совместной плотности распределения (§ 11)

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_\theta(\vec{x}) d\vec{x} = 1. \qquad (22.3)$$

Продифференцировав по θ это равенство, получим

$$0 = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \frac{\partial}{\partial\theta} f_\theta(\vec{x})d\vec{x} = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \frac{\partial}{\partial\theta}\big(\ln f_\theta(\vec{x})\big) f_\theta(\vec{x})d\vec{x} =$$
$$= \mathbf{E}\Big(\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\Big). \qquad (22.4)$$

Домножая это равенство на θ и вычитая результат из (22.2), будем иметь

$$\mathbf{E}\Big(\big(\gamma_n(\vec{X}) - \theta\big)\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\Big) = 1.$$

Воспользуемся четвертым свойством математического ожидания и неравенством Гельдера (12.2). Тогда

$$1 = \mathbf{E}\Big(\big(\gamma_n(\vec{X}) - \theta\big)\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\Big) \leqslant \mathbf{E}\Big|\big(\gamma_n(\vec{X}) - \theta\big)\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\Big| \leqslant$$
$$\leqslant \Big(\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2\Big)^{1/2} \Big(\mathbf{E}\big(\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\big)^2\Big)^{1/2}.$$

Возведя левую и правую части этого неравенства в квадрат, получим **неравенство Рао–Крамера**

$$\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2 \geqslant \frac{1}{I_n(\theta)}, \qquad (22.5)$$

где

$$I_n(\theta) = \mathbf{E}\big(\frac{\partial}{\partial\theta} \ln f_\theta(\vec{X})\big)^2 = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \Big(\frac{\partial}{\partial\theta} \ln f_\theta(\vec{x})\Big)^2 f_\theta(\vec{x})d\vec{x}.$$

Величина $I_n(\theta)$ называется **информационным количеством Фишера** относительно параметра θ, содержащимся в n наблюдениях $X_1, X_2, \ldots, X_n$.

Согласно неравенству Рао–Крамера среднеквадратичное отклонение оценки от оцениваемого параметра не может быть меньше величины, обратной к информационному количеству Фишера.

Определение 1. Оценка $\gamma_n(\vec{X})$ параметра θ, для которой выполняется равенство

$$\mathbf{E}\big(\gamma_n(\vec{X})-\theta\big)^2=\frac{1}{I_n(\theta)},$$

называется **эффективной оценкой**.

Пусть функция $f_\theta(x)$ дважды непрерывно дифференцируема по θ. В этом случае можно получить следующее выражение для информационного количества Фишера

$$I_n(\theta)=-\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(\vec{X})\Big), \tag{22.6}$$

которое часто является более удобным для вычислений. Оно следует из цепочки равенств

$$\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(\vec{X})\Big)=\mathbf{E}\left(\frac{\frac{\partial^2}{\partial\theta^2}f_\theta(\vec{X})}{f_\theta(\vec{X})}\right)-\mathbf{E}\left(\frac{\frac{\partial}{\partial\theta}f_\theta(\vec{X})}{f_\theta(\vec{X})}\right)^2=$$
$$=\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty}\frac{\partial^2}{\partial\theta^2}f_\theta(\vec{x})d\vec{x}-\mathbf{E}\Big(\frac{\partial}{\partial\theta}\ln f_\theta(\vec{X})\Big)^2=$$
$$=\frac{\partial^2}{\partial\theta^2}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty}f_\theta(\vec{x})d\vec{x}-I_n(\theta)=-I_n(\theta),$$

где для получения последнего равенства использовано соотношение (22.3).

Величина $I_n(\theta)$ возрастает с ростом n. Выразим информационное количество Фишера, содержащееся в n наблюдениях, через информационное количество Фишера, содержащееся в одном наблюдении, т. е. через величину

$$I_1(\theta)=\mathbf{E}\Big(\frac{\partial}{\partial\theta}\ln f_\theta(X_1)\Big)^2=\int\limits_{-\infty}^{\infty}\Big(\frac{\partial}{\partial\theta}\ln f_\theta(x)\Big)^2 f_\theta(x)dx.$$

Для функции $f_\theta(x)$ дважды непрерывно дифференцируемой по θ, используя (22.1), (22.6) и свойство, что логарифм от произведения величин равен сумме логарифмов, получим

$$I_n(\theta)=-\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\sum_{k=1}^{n}\ln f_\theta(X_k)\Big)=-\sum_{k=1}^{n}\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(X_k)\Big)=$$
$$=-n\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(X_1)\Big)=nI_1(\theta).$$

В случае, когда функция $f_\theta(x)$ не имеет второй производной по θ, аналогичное равенство получается следующим образом:

$$I_n(\theta) = \mathbf{E}\Big(\frac{\partial}{\partial\theta}\sum_{k=1}^{n}\ln f_\theta(X_k)\Big)^2 = \mathbf{E}\Big(\sum_{k=1}^{n}\frac{\partial}{\partial\theta}\ln f_\theta(X_k)\Big)^2 =$$

$$= \sum_{k=1}^{n}\mathbf{E}\big(\frac{\partial}{\partial\theta}\ln f_\theta(X_k)\big)^2 + 2\sum_{k=1}^{n}\sum_{l=1}^{k-1}\mathbf{E}\big(\big(\frac{\partial}{\partial\theta}\ln f_\theta(X_k)\big)\big(\frac{\partial}{\partial\theta}\ln f_\theta(X_l)\big)\big).$$

Величины $X_k, k = 1, 2, \ldots, n$ независимы и одинаково распределены. Воспользуемся третьим свойством математического ожидания (§ 12) и равенством (22.4) для $n = 1$. В результате получим

$$I_n(\theta) = nI_1(\theta) + 2\sum_{k=1}^{n}\sum_{l=1}^{k-1}\mathbf{E}\big(\frac{\partial}{\partial\theta}\ln f_\theta(X_k)\big)\mathbf{E}\big(\frac{\partial}{\partial\theta}\ln f_\theta(X_l)\big) = nI_1(\theta).$$

Таким образом, информационное количество Фишера, содержащееся в n наблюдениях, равно информационному количеству Фишера, содержащемуся в одном наблюдении, домноженному на число наблюдений n.

С учетом этого неравенству Рао–Крамера можно придать следующий вид:

$$\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2 \geqslant \frac{1}{nI_1(\theta)}. \tag{22.7}$$

Оценка $\gamma_n(\vec{X})$ будет эффективной, если

$$\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2 = \frac{1}{nI_1(\theta)}.$$

Эффективные оценки можно построить лишь в редких случаях. Чаще бывает так, что оценка лишь с ростом n приближается к неулучшаемой.

О п р е д е л е н и е 2. Оценка $\gamma_n(\vec{X})$ называется **асимптотически эффективной оценкой** параметра θ, если

$$n\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2 \underset{n\to\infty}{\longrightarrow} \frac{1}{I_1(\theta)}. \tag{22.8}$$

Замечание 1. В § 20 мы рассматривали асимптотически нормальные оценки с дисперсией Δ^2. Если $\Delta^2 = \dfrac{1}{I_1(\theta)}$, то асимптотически нормальная оценка будет и асимптотически эффективной.

Действительно, согласно определению асимптотической нормальности с $\Delta^2 = \dfrac{1}{I_1(\theta)}$, функция распределения величины $\sqrt{nI_1(\theta)}\big(\gamma_n(\vec{X}) - \theta\big)$ сходится при $n \to \infty$ к функции распределения стандартного нормального закона, который имеет дисперсию 1. Если, например, у оценки существует абсолютный момент порядка больше, чем два, то сходимость распределений влечет и сходимость дисперсий, т. е.

$$\mathbf{D}\big(\sqrt{nI_1(\theta)}\big(\gamma_n(\vec{X}) - \theta\big)\big) \underset{n\to\infty}{\to} 1. \tag{22.9}$$

Поскольку оценка $\gamma_n(\vec{X})$ предполагается несмещенной, то

$$\mathbf{D}\big(\sqrt{nI_1(\theta)}\big(\gamma_n(\vec{X}) - \theta\big)\big) = nI_1(\theta)\mathbf{E}\big(\gamma_n(\vec{X}) - \theta\big)^2.$$

С учетом этого равенства соотношение (22.9) эквивалентно (22.8).

Пример 1. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает нормальному распределению с плотностью

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\theta)^2/2\sigma^2},$$

где математическое ожидание θ является неизвестным параметром, а дисперсия $\mathbf{D}(X_1) = \sigma^2$ известна. Тогда

$$I_1(\theta) = \mathbf{E}\Big(\frac{\partial}{\partial\theta}\Big(\ln\frac{1}{\sqrt{2\pi}\sigma} - \frac{(X_1-\theta)^2}{2\sigma^2}\Big)\Big)^2 = \mathbf{E}\Big(\frac{(X_1-\theta)^2}{\sigma^4}\Big) = \frac{1}{\sigma^2}.$$

Среднеквадратичное отклонение выборочного среднего $\overline{X}_n$ от среднего θ вычислялось при доказательстве закона больших чисел. Согласно (16.5)

$$\mathbf{E}(\overline{X}_n - \theta)^2 = \frac{\mathbf{D}(X_1)}{n} = \frac{\sigma^2}{n} = \frac{1}{nI_1(\theta)}.$$

Таким образом, выборочное среднее для выборки, отвечающей нормальному распределению с известной дисперсией, является эффективной оценкой для математического ожидания.

Это один из тех редких примеров, когда можно явно указать эффективную оценку параметра. Другой пример связан с показательным распределением.

Пример 2. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает показательному распределению с параметром $\frac{1}{\theta} > 0$, т. е. распределению с плотностью

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Тогда, принимая во внимание следующие выражения для математического ожидания и дисперсии (§ 14, п. 2) $\mathbf{E}(X_1) = \theta$, $\mathbf{D}(X_1) = \theta^2$, получим

$$I_1(\theta) = -\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\Big(\ln f_\theta(X_1)\Big)\Big) = -\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\Big(\ln\frac{1}{\theta} - \frac{X_1}{\theta}\Big)\Big) =$$
$$= -\mathbf{E}\Big(\frac{1}{\theta^2} - \frac{2X_1}{\theta^3}\Big) = -\frac{1}{\theta^2} + \frac{2}{\theta^3}\mathbf{E}(X_1) = \frac{1}{\theta^2} = \frac{1}{\mathbf{D}(X_1)}.$$

В этом примере информационное количество Фишера зависит от параметра θ и не является постоянным, как это было в примере 1, однако тоже является обратно пропорциональным дисперсии. Отсюда, как показано в примере 1, следует, что выборочное среднее $\overline{X}_n$ является эффективной оценкой параметра θ.

§ 23. Проверка статистических гипотез

Предположим, что функция распределения случайной величины X нам неизвестна, но мы располагаем случайной выборкой $\vec{X} = (X_1, X_2, \ldots, X_n)$. По наблюдениям выборки $\vec{X}$ мы хотим дать ответ на вопрос: совпадает функция распределения $F(x)$ с некоторой наперед заданной функцией распределения $F_0(x)$ или нет. При такой постановке задачи говорят, что речь идет о проверке статистической гипотезы согласия.

Используя наблюдения выборки $\vec{X}$, нужно либо принять гипотезу о том, что функция распределения $F(x)$ совпадает с заданной функцией распределения $F_0(x)$, либо ее отвергнуть. Правило принятия одного из этих двух решений называется **статистическим критерием** или просто **критерием**. В качестве функции $F_0(x)$ обычно выбирается одно из известных распределений, например: нормальное, равномерное или показательное распределение с известными параметрами. Рассмотрим две конкретные задачи. Методы решения этих задач предложены А. Н. Колмогоровым и К. Пирсоном.

1. Критерий согласия Колмогорова.

Рассматривается гипотеза о том, что функция распределения $F(x)$ совпадает с непрерывной функцией распределения $F_0(x)$. Обозначим эту гипотезу символом H_0. Символом H_1 обозначим противоположную гипотезу о том, что $F(x) \neq F_0(x)$ хотя бы при одном значении x. Проверка гипотезы о распределении состоит в том, чтобы по наблюдениям выборки $X_1, X_2, \ldots, X_n$ сделать вывод, что функция распределения $F(x)$ совпадает с $F_0(x)$, т. е. справедлива гипотеза H_0, или заключить, что $F(x)$ не совпадает с $F_0(x)$, т. е. справедлива гипотеза H_1. Поскольку наблюдения случайны, то абсолютно достоверно такие утверждения сделать нельзя. При любом естественном подходе есть положительные вероятности того, что мы примем гипотезу H_0, когда она на самом деле не верна, или, что мы примем гипотезу H_1, когда она не является верной. Пусть α – некоторое наперед заданное малое положительное число. Мы хотим указать такое правило (критерий), которое бы по наблюдениям выборки $X_1, X_2, \ldots, X_n$ отвергало гипотезу H_0 при условии, что она на самом деле верна, с вероятностью α или близкой к α для больших n. При этом правило следует выбирать таким, чтобы вероятность принять гипотезу H_0 при условии, что верна гипотеза H_1, была бы как можно меньшей.

Определение 1. Вероятность α отвергнуть гипотезу H_0 при условии, что она верна, называется **вероятностью ошибки первого рода**.

Вероятность ошибки первого рода также называют **уровнем значимости** критерия.

Определение 2. Вероятность принять гипотезу H_0, при условии, что верна гипотеза H_1, называется **вероятностью ошибки второго рода**.

Вероятность ошибки второго рода будем обозначать β.

Обычно за H_0 принимается такая гипотеза, отвержение которой, когда она на самом деле верна, будет иметь наихудшие последствия по сравнению с теми, когда за H_0 выбирается другая гипотеза.

В рассматриваемой задаче гипотеза H_0 характеризуется выбором функции $F_0(x)$. Однако вышеприведенные определения и замечание справедливы безотносительно к конкретному выбору гипотез H_0 и H_1.

Поскольку при фиксированном объеме выборки вероятности ошибок первого и второго рода одновременно сделать достаточно малыми нельзя, то α задают заранее, а β при уже выбранном α стараются сделать как можно меньше.

Все статистические критерии устроены так, что при уменьшении α повышаются шансы на принятие гипотезы H_0, а при $\alpha = 0$ гипотеза H_0 обязательно принимается.

Пусть $F_n(x)$ – эмпирическая функция распределения, построенная по выборке $\vec{X}$, т. е. $F_n(x) = \frac{1}{n}\sum\limits_{l=1}^{n} \mathbb{I}_{(-\infty,x)}(X_l)$. Положим

$$D_n(\vec{X}) = \sqrt{n} \max_{x\in(-\infty,\infty)} |F_n(x) - F_0(x)|.$$

Поскольку $F_n(x)$ при больших n является хорошей оценкой $F(x)$ – истинной функции распределения случайной величины X, то функция $D_n(\vec{X})$ является мерой различия между истинной функцией распределения $F(x)$ и предполагаемой $F_0(x)$. Функция $D_n(\vec{X})$ называется **статистикой Колмогорова**. Согласно теореме Колмогорова (§ 18), если $F(x) = F_0(x)$, т. е. при условии, что верна гипотеза H_0, справедливо

$$\mathbf{P}(D_n(\vec{X}) < z) \approx \mathcal{K}(z),$$

где $\mathcal{K}(z)$ – функция распределения Колмогорова. Для функции $\mathcal{K}(z)$ есть таблицы и по заданному малому числу α можно выбрать $z_{1-\alpha}$ так, чтобы $\mathcal{K}(z_{1-\alpha}) = 1 - \alpha$ ($z_{1-\alpha}$ является квантилью порядка $1-\alpha$ функции распределения $\mathcal{K}(z)$).

Определим следующее правило, которое называется **критерием согласия Колмогорова**: если $D_n(\vec{X}) < z_{1-\alpha}$, то принимаем гипотезу H_0, т. е. считаем, что $F = F_0$, а если $D_n(\vec{X}) \geqslant z_{1-\alpha}$, то принимаем гипотезу H_1, т. е. считаем, что $F \neq F_0$. При таком правиле вероятность отвергнуть гипотезу H_0, при условии, что она верна, есть $\mathbf{P}(D_n(\vec{X}) \geqslant z_{1-\alpha})$, и согласно теореме Колмогорова

$$\mathbf{P}(D_n(\vec{X}) \geqslant z_{1-\alpha}) = 1 - \mathbf{P}(D_n(\vec{X}) < z_{1-\alpha}) \approx 1 - \mathcal{K}(z_{1-\alpha}) = \alpha.$$

Таким образом, первое требование выполнено: вероятность ошибки первого рода для данного правила при больших n приблизительно равна α. Предположим теперь, что $F \neq F_0$, т. е. верна гипотеза H_1. Тогда в силу того, что $F_n(x) \to F(x)$ равномерно по $x \in (-\infty, \infty)$ (теорема 1 § 18), имеем

$$\max_{x\in(-\infty,\infty)} |F_n(x) - F_0(x)| \underset{n\to\infty}{\longrightarrow} \max_{x\in(-\infty,\infty)} |F(x) - F_0(x)| > 0.$$

Поэтому $D_n(\vec{X}) \to \infty$, и, следовательно, для любой константы c

$$\mathbf{P}(D_n(\vec{X}) < c) \to 0.$$

Возьмем в качестве c величину $z_{1-\alpha}$. Тогда имеем

$$\beta = \mathbf{P}(D_n(\vec{X}) < z_{1-\alpha}) \underset{n\to\infty}{\longrightarrow} 0.$$

При $F \neq F_0$, величина β равна вероятности принять гипотезу H_0 при условии, что верна гипотеза H_1, т. е. β – вероятность ошибки второго рода. При больших n величина β становится все меньше и меньше. Тем самым выполнено и второе требование о том, что величина β должна быть как можно меньшей при выбранном α.

Задача 1. Результаты измерений (в мм) длины 20 однотипных деталей дали следующие значения

86.44; 88.37; 88.56; 87.73; 89.04; 88.05; 87.34; 87.88; 88.29; 86.87;
88.77; 89.65; 87.58; 88.83; 87.94; 87.99; 86.32; 87.02; 89.11; 88.44.

Проверить с помощью критерия Колмогорова с уровнем значимости 0.05 гипотезу о том, что величина длины детали X имеет нормальное распределение со средним 88 и дисперсией 1.

Решение. Поскольку в статистике Колмогорова максимум берется по всем значениям аргумента x, то она не зависит от сдвига аргумента у эмпирической и предполагаемой функций распределения на одну и ту же величину. Это означает, что вместо исходной случайной величины X можно рассмотреть величину $Y = X - 88$ и проверить с помощью критерия Колмогорова гипотезу о том, что величина Y имеет нормальное распределение со средним 0 и дисперсией 1.

Сначала нужно упорядочить значения величины $Y = X - 88$, т. е. построить реализацию вариационного ряда $y_{(1)} \leqslant y_{(2)} \leqslant \cdots \leqslant y_{(20)}$. Имеем $-1.68 \leqslant -1.56 \leqslant -1.13 \leqslant -0.98 \leqslant -0.66 \leqslant -0.42 \leqslant -0.27 \leqslant -0.12 \leqslant -0.06 \leqslant -0.01 \leqslant 0.05 \leqslant 0.29 \leqslant 0.37 \leqslant 0.44 \leqslant 0.56 \leqslant 0.77 \leqslant 0.83 \leqslant 1.04 \leqslant 1.11 \leqslant 1.65$. Далее нужно вычислить значения $\mathcal{N}(x)$ – нормального распределения со средним 0 и дисперсией 1 в этих точках. Для этого можно использовать таблицу для значений $\mathcal{N}(x)$, а для отрицательных аргументов применить формулу $\mathcal{N}(-x) = 1 - \mathcal{N}(x)$. Другой вариант – использовать формулу (14.1) с учетом нечетности функции Лапласа ($\Phi(-x) = -\Phi(x)$) и таблицу для значений $\Phi(x)$. В результате получим: $\mathcal{N}(-1.68) \approx 0.04648$; $\mathcal{N}(-1.56) \approx 0.05938$; $\mathcal{N}(-1.13) \approx 0.12924$; $\mathcal{N}(-0.98) \approx 0.16354$; $\mathcal{N}(-0.66) \approx 0.25463$; $\mathcal{N}(-0.42) \approx 0.33724$; $\mathcal{N}(-0.27) \approx 0.39358$; $\mathcal{N}(-0.12) \approx 0.45224$; $\mathcal{N}(-0.06) \approx 0.47608$; $\mathcal{N}(-0.01) \approx 0.49601$; $\mathcal{N}(0.05) \approx 0.51994$; $\mathcal{N}(0.29) \approx 0.61409$; $\mathcal{N}(0.37) \approx 0.64431$; $\mathcal{N}(0.44) \approx 0.67003$; $\mathcal{N}(0.56) \approx 0.71224$; $\mathcal{N}(0.77) \approx 0.77935$; $\mathcal{N}(0.83) \approx 0.79673$; $\mathcal{N}(1.04) \approx 0.85083$; $\mathcal{N}(1.11) \approx 0.86650$; $\mathcal{N}(1.65) \approx 0.95053$.

Поскольку при 20 наблюдениях эмпирическая функция распределения увеличивается в каждой из точек $y_{(k)}$ на $1/20$ то, нарисовав графики функций, несложно понять, что

$$D_{20}(\vec{y}) = \sqrt{20} \max_{1 \leqslant k \leqslant 20} \left\{ \left| \mathcal{N}(y_{(k)}) - \frac{k}{20} \right|, \left| \mathcal{N}(y_{(k)}) - \frac{k-1}{20} \right| \right\}.$$

Подсчет показывает, что этот максимум достигается на значении

$$\sqrt{20} \left| \mathcal{N}(y_{(8)}) - \frac{7}{20} \right| \approx \sqrt{20}(0.45244 - 0.35) \approx 4.4721 \cdot 0.10224 \approx 0.457.$$

Из таблицы для распределения Колмогорова находим $z_{0\,95} \approx 1.36$. Поскольку $D_{20}(\vec{y}) < z_{0\,95}$, то гипотеза принимается.

Перейдем к изложению другого подхода, при котором по-иному ставится и решается вопрос о равенстве функций распределения. Пусть нас интересует не вся функция распределения, а лишь ее приращения на некоторых выделенных интервалах. При этом вопрос о совпадении функции распределения с наперед заданной функцией заменяется, естественно, на вопрос о совпадении приращений функции распределения с соответствующими приращениями заранее выделенной функции. Заметим, что для дискретных распределений в качестве выделенных интервалов можно взять интервалы постоянства функции распределения. Тогда приращения функции распределения на них полностью характеризуют закон распределения. В этом случае задача становится аналогичной предыдущей, однако метод решения будет принципиально отличным.

2. **Критерий согласия Пирсона (хи-квадрат критерий).**

Предположим, что вещественная прямая разбита на r непересекающихся интервалов $\Delta_l, l = 1, 2, \ldots, r$, т. е. $\bigcup_{l=1}^{r} \Delta_l = (-\infty, \infty)$, $\Delta_i \cap \Delta_j = \varnothing$. Нас интересует не вся функция распределения $F(x)$ случайной величины X, а только набор вероятностей $p_l = \mathbf{P}(X \in \Delta_l)$, $l = 1, 2, \ldots, r$. Задача состоит в следующем. Пусть задан набор вероятностей $p_l^\circ, l = 1, 2, \ldots, r$, таких, что $\sum\limits_{l=1}^{r} p_l^\circ = 1$. Обычно $p_l^\circ = \mathbf{P}(X^\circ \in \Delta_l)$, где X° – случайная величина с каким-нибудь известным распределением. Мы хотим знать, совпадают ли вероятности p_l, отвечающие наблюдениям выборки $X_1, X_2, \ldots, X_n$, с вероятностями p_l° или нет. Обозначим символом H_0 гипотезу о том, что $p_l = p_l^\circ$ при всех $l = 1, 2, \ldots, r$. Символом H_1 обозначим противоположную гипотезу о том, что $p_l \neq p_l^\circ$ хотя бы для одного l.

Пусть $X_1, X_2, \ldots, X_n$ – выборка, отвечающая случайной величине X. Определим правило, по которому мы будем принимать или отвергать гипотезу H_0. Обозначим через $\nu_l = \sum\limits_{j=1}^{n} \mathbb{I}_{\Delta_l}(X_j)$ число тех наблюдений X_j, которые принадлежат интервалу Δ_l. Поскольку $\mathbf{E}\big(\mathbb{I}_{\Delta_l}(X_j)\big) = \mathbf{P}(X_j \in \Delta_l) = p_l$, то по закону больших чисел $\frac{\nu_l}{n} \to p_l$ по вероятности.

Поэтому $\frac{\nu_l}{n}$ является хорошей оценкой для p_l. Положим

$$\Delta_n(\vec{X}) = \sum_{l=1}^{r} \frac{n}{p_l^\circ}\Big(\frac{\nu_l}{n} - p_l^\circ\Big)^2 = \sum_{l=1}^{r} \frac{(\nu_l - np_l^\circ)^2}{np_l^\circ}.$$

Функцию $\Delta_n(\vec{X})$ называют **статистикой Пирсона или статистикой χ^2 (хи-квадрат)** с $r-1$ степенью свободы.

То, что мера различия между истинными вероятностями p_l и предполагаемыми p_l° имеет именно такой вид, продиктовано желанием иметь хорошее предельное распределение для величины $\Delta_n(\vec{X})$. Справедлив следующий результат, который приводится без доказательства.

Теорема (Пирсона). *Пусть $p_l = p_l^\circ$ при всех $l = 1, 2, \ldots, r$, т. е. выполнена гипотеза H_0. Тогда справедливо следующее соотношение*

$$\sup_{z\in(0,\infty)} |\mathbf{P}(\Delta_n(\vec{X}) < z) - \mathcal{P}_{r-1}(z)| \to 0$$

при $n \to \infty$, где $\mathcal{P}_{r-1}(z)$ – так называемое **распределение Пирсона** *с $r-1$ степенью свободы. Для любого целого неотрицательного k распределение $\mathcal{P}_k(x)$ определяется плотностью*

$$p_k(x) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k-2)/2} e^{-x/2}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0, \end{cases}$$

где $\Gamma(t) = \int\limits_0^\infty x^{t-1} e^{-x} dx$ – **гамма-функция.**

При целых n для гамма-функции выполняются равенства

$$\Gamma(n+1) = n!, \quad \Gamma\big(n + \tfrac{1}{2}\big) = 1 \cdot 3 \cdot 5 \cdots (2n-3)(2n-1)\sqrt{\pi} 2^{-n}.$$

Для распределения Пирсона существуют таблицы. По заданному значению α можно определить такое число $z_{1-\alpha}$ (квантиль порядка $1-\alpha$), что $\mathcal{P}_{r-1}(z_{1-\alpha}) = 1 - \alpha$.

Определим следующее правило, которое называется **критерием согласия Пирсона** или **критерием хи-квадрат**:

если $\Delta_n(\vec{X}) < z_{1-\alpha}$, то принимаем гипотезу H_0, т. е. считаем, что $p_l = p_l^0$ для всех $l = 1, 2, \ldots, r$, если $\Delta_n(\vec{X}) \geqslant z_{1-\alpha}$, то принимаем гипотезу H_1, т. е. считаем, что $p_l \neq p_l^0$ при некотором l. При таком правиле вероятность отвергнуть гипотезу H_0 при условии, что она верна, равна $\mathbf{P}(\Delta_n(\vec{X}) \geqslant z_{1-\alpha})$, и согласно теореме Пирсона

$$\mathbf{P}(\Delta_n(\vec{X}) \geqslant z_{1-\alpha}) = 1 - \mathbf{P}(\Delta_n(\vec{X}) < z_{1-\alpha}) \approx 1 - \mathcal{P}_{r-1}(z_{1-\alpha}) = \alpha.$$

Таким образом, вероятность ошибки первого рода приблизительно равна α. Предположим теперь, что $p_l \neq p_l^0$ для какого-нибудь l, т. е. предположим, что верна гипотеза H_1. Тогда, поскольку в силу закона больших чисел $\frac{\nu_l}{n} \to p_l$, имеем

$$\left(\frac{\nu_l}{n} - p_l^0\right)^2 \underset{n\to\infty}{\longrightarrow} (p_l - p_l^0)^2 > 0,$$

и, следовательно,

$$\frac{n}{p_l^0}\left(\frac{\nu_l}{n} - p_l^0\right)^2 \to \infty.$$

Тем самым, при выполнении гипотезы H_1 имеет место соотношение $\Delta_n(\vec{X}) \to \infty$. Поэтому

$$\mathbf{P}(\Delta_n(\vec{X}) < z_{1-\alpha}) \to 0 \quad \text{при } n \to \infty.$$

Положим $\beta = \mathbf{P}(\Delta_n(\vec{X}) < z_{1-\alpha})$ – вероятность принять гипотезу H_0 при условии, что верна гипотеза H_1. Тогда с ростом n вероятность β стремится к нулю. В результате оба требования, предъявляемые к критериям согласия, выполнены.

Задача 2. Через равные промежутки времени в препарате регистрируется число микробов, попавших в поле зрения микроскопа. В результате наблюдений получены следующие данные:

l	0	1	2	3	4	5	6	7
ν_l	112	168	130	68	32	5	1	1

В первой строке приведено число микробов, а во второй строке – число моментов времени, соответствующих наблюдению ровно l микробов.

Проверить, используя критерий хи-квадрат, что с вероятностью ошибки первого рода $\alpha = 0.05$ число микробов, попадающих в поле зрения микроскопа в любой момент регистрации, распределено по закону Пуассона с математическим ожиданием $\lambda = 1.5$.

Решение. Пусть X_j – неотрицательная случайная величина, равная числу микробов, регистрируемых за j-й промежуток времени. Считаем величины $X_j, j = 1, 2, \ldots, n$, независимыми и одинаково распределенными. Общий объем выборки равен $n = \sum\limits_{l=0}^{7} \nu_l = 517$. Можно считать, что неотрицательная полуось разбита на 8 интервалов $\Delta_l = [l, l+1), l = 0, 1, \ldots, 6$ и $\Delta_7 = [7, \infty)$. Тогда $\nu_l = \sum\limits_{j=1}^{517} \mathbb{I}_{\Delta_l}(X_j)$ – число моментов времени, когда регистрируется ровно l микробов. Значения ν_l приведены в условии задачи. Пусть X° – случайная величина, распределенная по закону Пуассона с параметром $\lambda = 1.5$:

$$\mathbf{P}(X^\circ = l) = \frac{\lambda^l e^{-\lambda}}{l!}, \quad l = 0, 1, 2, \ldots .$$

Используя эту формулу, найдем теоретические вероятности попадания числа микробов в поле зрения микроскопа:

$$\begin{aligned}
p_0^\circ &= \mathbf{P}(X^\circ = 0) \approx 0.2231; & p_4^\circ &= \mathbf{P}(X^\circ = 4) \approx 0.0471;\\
p_1^\circ &= \mathbf{P}(X^\circ = 1) \approx 0.3347; & p_5^\circ &= \mathbf{P}(X^\circ = 5) \approx 0.0141;\\
p_2^\circ &= \mathbf{P}(X^\circ = 2) \approx 0.2510; & p_6^\circ &= \mathbf{P}(X^\circ = 6) \approx 0.0035;\\
p_3^\circ &= \mathbf{P}(X^\circ = 3) \approx 0.1255; & p_7^\circ &= \mathbf{P}(X^\circ \geqslant 7) \approx 0.0010.
\end{aligned}$$

Вероятность p_7° является вероятностью того, что наблюдается не менее семи микробов, так как теоретически в поле зрения микроскопа за выбранный промежуток времени может наблюдаться любое их количество. Значение Δ_n статистики Пирсона вычисляем по формуле

$$\Delta_n = \sum_{l=0}^{7} \frac{(\nu_l - np_l^\circ)^2}{np_l^\circ} \approx 4.3.$$

По таблицам для квантилей распределения Пирсона $\mathcal{P}_{r-1}(z)$ с $r = 8$ найдем критическое значение $z_{1-\alpha}$ при $\alpha = 0.05$. Имеем $z_{0.95} \approx 14.07$. Так как $\Delta_n < z_{1-\alpha}$, то мы принимаем гипотезу о том, что число микробов, попадающих в поле зрения микроскопа, распределено по закону Пуассона с математическим ожиданием $\lambda = 1.54$.

Задачи

Задача 23.1. В городе 17036 семей имеют двоих детей. В 4529 семьях – два мальчика, в 4019 – две девочки, в 8488 семьях – мальчик и девочка. Можно ли с уровнем значимости 0.05 считать, что количество мальчиков в семьях с двумя детьми имеет биномиальное распределение с вероятностью рождения мальчика 0.515?

Задача 23.2. Датчик случайных чисел выдает независимые значения случайной величины, которая должна иметь равномерное распределение на $[0, 1]$. Получены следующие 20 значений:

$0.100; 0.253; 0.520; 0.863; 0.354; 0.809; 0.911; 0.292; 0.453; 0.204;$
$0.648; 0.429; 0.805; 0.372; 0.610; 0.008; 0.166; 0.422; 0.531; 0.509.$

С помощью критерия Колмогорова с уровнем значимости 0.1 проверить гипотезу о том, что датчик действительно генерирует значения случайной величины с равномерным распределением на $[0, 1]$.

Задача 23.3. Результаты подсчета частот появления цифр $0, 1, \ldots, 9$ в 10002 знаках десятичной записи числа $\pi - 3$ приведены в таблице:

0	1	2	3	4	5	6	7	8	9
968	1026	1021	974	1014	1046	1021	970	948	1014

С помощью критерия Пирсона с вероятностью ошибки первого рода 0.2 проверить гипотезу о равновероятности появления в данной записи каждой из цифр $0, 1, \ldots, 9$.

§ 24. Оценка параметров общей линейной модели (метод наименьших квадратов)

Рассматривается следующая линейная модель. Имеется вектор неизвестных параметров, который нужно определить. Однако наблюдается не сам вектор параметров, а другой вектор, который получается из искомого вектора применением линейного преобразования. Кроме того, наблюдения (измерения) осуществляются с некоторыми случайными ошибками. Предполагается, что эти ошибки не являются систематическими, т. е. среднее этих ошибок равно нулю.

Математически эта модель описывается следующим образом. Обозначим $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ – вектор неизвестных параметров, подлежащих оценке. Пусть $A = \{a_{ij}\}_{i=1,j=1}^{n\ m}$ – матрица известных коэффициентов, а $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ – случайные величины с математическими ожиданиями $\mathbf{E}(\mathcal{E}_i) = 0$, с дисперсиями $\mathbf{D}(\mathcal{E}_i) = \sigma_i^2 < \infty$ и с ковариациями $\mathbf{E}(\mathcal{E}_i\mathcal{E}_j) = 0$ при $i \neq j$. Наблюдаются величины X_i:

$$X_i = \sum_{j=1}^{m} a_{ij}\theta_j + \mathcal{E}_i, \quad i = 1, 2, \ldots, n. \tag{24.1}$$

Если бы случайных ошибок измерения не было, то значения параметров θ_j по известным a_{ij} и X_i определялись бы как решение системы n алгебраических линейных уравнений. В случае, когда есть случайные ошибки $\mathcal{E}_i$, предлагается поступать следующим образом. Поскольку мы не знаем истинных значений параметров, выберем какие-нибудь значения параметров $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n$ и рассмотрим сумму квадратичных отклонений

$$R(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m) = \sum_{i=1}^{n}\Big(X_i - \sum_{j=1}^{m} a_{ij}\hat{\theta}_j\Big)^2.$$

Определим, при каких значениях $\hat{\theta}_j$ математическое ожидание $R(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ будет минимальным. Имеем

$$\mathbf{E}(R(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)) = \sum_{i=1}^{n}\mathbf{E}\Big(X_i - \sum_{j=1}^{m} a_{ij}\hat{\theta}_j\Big)^2 =$$

$$= \sum_{i=1}^{n}\mathbf{E}\Big(\sum_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j) + \mathcal{E}_i\Big)^2 =$$

$$= \sum_{i=1}^{n}\Big\{\Big(\sum_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j)\Big)^2 + 2\sum_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j)\mathbf{E}(\mathcal{E}_i) + \mathbf{E}(\mathcal{E}_i^2)\Big\} =$$

$$= \sum_{i=1}^{n}\Big\{\Big(\sum_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j)\Big)^2 + \sigma_i^2\Big\} = \sum_{i=1}^{n}\Big(\sum_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j)\Big)^2 + \sum_{i=1}^{n}\sigma_i^2.$$

Мы видим, что это математическое ожидание будет минимальным при $\sum\limits_{i=1}^{n}\big(\sum\limits_{j=1}^{m} a_{ij}(\theta_j - \hat{\theta}_j)\big)^2 = 0$, в частности, при $\hat{\theta}_j = \theta_j$, т. е. при истинных значениях параметров. Поэтому в качестве оценки для неизвестных параметров θ_j разумно взять такие значения параметров $\hat{\theta}_j$, при которых функция $R(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ принимает минимальное значение. Такая оценка называется **оценкой по методу наименьших квадратов**. Функция $R(\theta_1, \theta_2, \ldots, \theta_m)$ – функция многих переменных, и минимума она достигает в той точке, в которой все частные производные этой функции по каждой из переменных равны нулю. Частные производные имеют вид

$$\frac{\partial}{\partial\theta_k}R(\theta_1, \theta_2, \ldots, \theta_m) = 2\sum_{i=1}^{n}\Big(X_i - \sum_{j=1}^{m} a_{ij}\theta_j\Big)(-a_{ik}), \quad k = 1, \ldots, m.$$

Найдем такие значения $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$, которые обращают эти производные в ноль. Отсюда получим систему m линейных уравнений относительно неизвестных $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$:

$$\sum_{j=1}^{m}\Big(\sum_{i=1}^{n} a_{ij}a_{ik}\Big)\hat{\theta}_j = \sum_{i=1}^{n} a_{ik}X_i, \quad k = 1, 2, \ldots, m. \qquad (24.2)$$

Матрица этой системы состоит из коэффициентов $b_{jk} = \sum\limits_{i=1}^{n} a_{ij}a_{ik}$, а вектор свободных коэффициентов $\vec{Y}$ состоит из компонент $Y_k = \sum\limits_{i=1}^{n} a_{ik}X_i$. Если определитель матрицы $B = \{b_{jk}\}_{j=1}^{m}{}_{k=1}^{m}$ отличен от нуля, то эта система имеет единственное решение $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$. Решение это можно найти либо методом Гаусса, либо методом определителей, либо вычисляя обратную матрицу к B. Исходная линейная модель (24.1) и система уравнений (24.2) могут быть записаны в краткой форме с использованием матриц и векторов. Обозначим

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix},$$

$$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{pmatrix}, \qquad \vec{\mathcal{E}} = \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \vdots \\ \mathcal{E}_n \end{pmatrix}.$$

Тогда (24.1) переписывается в виде

$$\vec{X} = A\vec{\theta} + \vec{\mathcal{E}}.$$

Далее, полагая

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix}, \qquad \vec{\hat{\theta}} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_m \end{pmatrix},$$

$$A^{\top} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix},$$

где матрица $A^{\top}$ является транспонированной (отраженной) по отношению к матрице A, систему уравнений (24.2) представим в виде

$$B\vec{\hat{\theta}} = A^{\top}\vec{X}.$$

Решение этой системы можно записать так: $\vec{\hat{\theta}} = B^{-1}A^{\top}\vec{X}$, где B^{-1} – матрица, обратная к B. Таким образом, искомая оценка $\vec{\hat{\theta}}$, которая является оценкой для вектора неизвестных параметров $\vec{\theta}$, равна произведению матрицы B^{-1} на вектор $A^{\top}\vec{X}$. В силу определения коэффициентов b_{jk} матрица B представима в виде произведения $B = A^{\top}A$, что позволяет выразить оценку $\vec{\hat{\theta}}$, построенную по методу наименьших квадратов, в терминах исходной матрицы коэффициентов A и вектора наблюдений $\vec{X}$:

$$\vec{\hat{\theta}} = \vec{\hat{\theta}}(X_1, X_2, \dots, X_n) = (A^{\top}A)^{-1}A^{\top}\vec{X}. \qquad (24.3)$$

Пример 1. Пусть есть полином

$$P(y) = \theta_m y^{m-1} + \theta_{m-1}y^{m-2} + \dots + \theta_1$$

степени $m-1$ с неизвестными коэффициентами. Предположим, что у нас есть результаты измерений значений этого полинома в различных точках $y_1, y_2, \ldots, y_n$. Но измерения в каждой точке y_i производятся со случайной ошибкой $\mathcal{E}_i$, т. е. мы в качестве измерений имеем лишь значения величин X_i,

$$X_i = P(y_i) + \mathcal{E}_i, \quad i = 1, 2, \ldots, n. \tag{24.4}$$

Как определить коэффициенты θ_j? Оказывается, можно применить метод наименьших квадратов. Положим $a_{ij} = y_i^{j-1}$, тогда

$$P(y_i) = \sum_{j=1}^{m} y_i^{j-1}\theta_j = \sum_{j=1}^{m} a_{ij}\theta_j,$$

и равенства (24.4) переходят в равенства (24.1). Но для такой линейной модели мы уже умеем строить оценку параметров $\theta_1, \theta_2, \ldots, \theta_m$, и эта оценка определяется по коэффициентам a_{ij} системой уравнений (24.2), которая в данном случае имеет вид

$$\sum_{j=1}^{m}\Bigl(\sum_{i=1}^{n} y_i^{j+k-2}\Bigr)\hat{\theta}_j = \sum_{i=1}^{n} y_i^{k-1} X_i, \quad k = 1, 2, \ldots, m.$$

Решение этой системы не обязательно выписывать формулой (24.3). Его можно получать различными известными методами.

Задача 1. Затраты y на развитие производства и величина годовой прибыли фирмы X в течение 5 лет представлены в условных единицах следующей таблицей

$\vec{y}$	6	3	7	5	10
$\vec{X}$	33	27	32	28	42

На величину прибыли влияют случайные факторы. Предполагается, что всегда имеет место линейная зависимость между затратами y и прибылью X вида

$$X = \theta_2 y + \theta_1 + \mathcal{E} \tag{24.5}$$

с неизменными параметрами θ_1, θ_2 и величиной случайного влияния $\mathcal{E}$ со средним ноль и конечной дисперсией. Каждый год случайное влияние некоррелировано с предыдущими годами. Оценить параметры θ_1 и θ_2.

Решение. Зависимость (24.5), представленная по годам $i = 1, 2, 3, 4, 5$, имеет вид

$$X_i = \theta_2 y_i + \theta_1 + \mathcal{E}_i. \tag{24.6}$$

Таким образом, она описывается моделью типа (24.1). Оценим параметры θ_1 и θ_2 по методу наименьших квадратов. Имеем $n = 5$, $m = 2$, $a_{i1} = 1$, $a_{i2} = y_i$. Система уравнений (24.2) в этом случае следующая

$$\begin{cases} n\hat{\theta}_1 + \sum\limits_{i=1}^{n} y_i\hat{\theta}_2 = \sum\limits_{i=1}^{n} X_i, & k = 1, \\ \sum\limits_{i=1}^{n} y_i\hat{\theta}_1 + \sum\limits_{i=1}^{n} y_i^2\hat{\theta}_2 = \sum\limits_{i=1}^{n} y_i X_i, & k = 2. \end{cases} \tag{24.7}$$

Вычисляя значение θ_2 методом определителей, получим

$$\hat{\theta}_2 = \frac{n\sum\limits_{i=1}^{n} y_i X_i - \sum\limits_{i=1}^{n} y_i \sum\limits_{i=1}^{n} X_i}{n\sum\limits_{i=1}^{n} y_i^2 - (\sum\limits_{i=1}^{n} y_i)^2} = \frac{\overline{yX}_n - \bar{y}_n\overline{X}_n}{\overline{y^2}_n - \bar{y}_n^{\,2}}, \tag{24.8}$$

где $\overline{y^2}_n = \sum\limits_{i=1}^{n} y_i^2/n$, $\overline{yX}_n = \sum\limits_{i=1}^{n} y_i X_i/n$, $\bar{y}_n = \sum\limits_{i=1}^{n} y_i/n$, $\overline{X}_n = \sum\limits_{i=1}^{n} X_i/n$. Теперь из первого уравнения системы (24.6) находим

$$\hat{\theta}_1 = \overline{X}_n - \hat{\theta}_2\bar{y}_n = \frac{\bar{y}_n^{\,2}\overline{X}_n - \bar{y}_n\overline{yX}_n}{\overline{y^2}_n - \bar{y}_n^{\,2}}. \tag{24.9}$$

Для конкретных вычислений вместо случайных величин X_i подставляют их реализации x_i. Для имеющихся данных задачи 1

$$\sum_{i=1}^{5} y_i = 31, \quad \sum_{i=1}^{5} x_i = 162, \quad \sum_{i=1}^{5} y_i x_i = 1063, \quad \sum_{i=1}^{5} y_i^2 = 219.$$

Тогда

$$\hat{\theta}_2 = \frac{5\cdot 1063 - 31\cdot 162}{5\cdot 219 - 31^2} = \frac{293}{134} \approx 2.187,$$

$$\hat{\theta}_1 = \frac{1}{5}\cdot 162 - \frac{293}{134}\cdot\frac{1}{5}\cdot 31 \approx 18.843.$$

В результате получаем, что зависимость величины годовой прибыли фирмы X от затрат на развитие производства y имеет вид

$$X \approx 2.187y + 18.843 + \mathcal{E}.$$

Зависимость (24.6) является частным случаем следующей зависимости, в которой величины y_i могут быть случайными:

$$X_i = \theta_2 Y_i + \theta_1 + \mathcal{E}_i \quad i = 1, 2, \ldots, n. \tag{24.10}$$

Предполагается, что Y_i, $i = 1, 2, \ldots, n$, являются одинаково распределенными независимыми случайными величинами, независящими от величин $\mathcal{E}_i$, которые интерпретируются как ошибки измерения. В свою очередь величины $\mathcal{E}_i$, $i = 1, 2, \ldots, n$, тоже являются независимыми со средним ноль ($\mathbf{E}(\mathcal{E}_i) = 0$) и конечной дисперсией ($\mathbf{E}(\mathcal{E}_i^2) < \infty$). По наблюдениям значений пар случайных величин X_i, Y_i, $i = 1, 2, \ldots, n$, нужно определить параметры θ_1, θ_2. Такая модель называется **простой линейной регрессией**.

По тем же соображениям, которые приведены для вывода оценки наименьших квадратов, в качестве оценок $\hat{\theta}_1$, и $\hat{\theta}_2$ для параметров θ_1, и θ_2 естественно взять те значения, которые минимизируют функцию квадратичного отклонения

$$R(\hat{\theta}_1, \hat{\theta}_2) = \sum_{i=1}^{n}\left(X_i - \hat{\theta}_2 Y_i - \hat{\theta}_1\right)^2.$$

Приравнивая к нулю частные производные этой функции по $\hat{\theta}_1$, и $\hat{\theta}_2$, получим систему уравнений (24.7) с величинами Y_i вместо y_i. Решения системы уравнений будут аналогичны (24.8), (24.9):

$$\hat{\theta}_2 = \frac{n\sum\limits_{i=1}^{n} Y_i X_i - \sum\limits_{i=1}^{n} Y_i \sum\limits_{i=1}^{n} X_i}{n\sum\limits_{i=1}^{n} Y_i^2 - (\sum\limits_{i=1}^{n} Y_i)^2} = \frac{C_n(\vec{X}, \vec{Y})}{S_n^2(\vec{Y})}, \tag{24.11}$$

$$\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i - \hat{\theta}_2 \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{X}_n - \hat{\theta}_2 \overline{Y}_n. \tag{24.12}$$

Соответствие $x = \hat{\theta}_2 y + \hat{\theta}_1$, $y \in (-\infty, \infty)$, называется **уравнением линейной регрессии**. Не удивительно, что для параметров θ_1, и θ_2 получились именно такие оценки. Вычислим ковариацию величин X и Y, связанных линейной зависимостью $X = \theta_2 Y + \theta_1 + \mathcal{E}$. Поскольку величины $\mathcal{E}$ и Y независимы и $\mathbf{E}(\mathcal{E}) = 0$, то $\mathbf{E}(X) = \theta_2 \mathbf{E}(Y) + \theta_1$,

$$\mathrm{cov}(X,Y) = \mathbf{E}((\theta_2 Y + \mathcal{E} - \theta_2 \mathbf{E}(Y))(Y - \mathbf{E}(Y))) =$$
$$= \theta_2 \mathbf{E}(Y - \mathbf{E}(Y))^2 + \mathbf{E}(\mathcal{E})\mathbf{E}(Y - \mathbf{E}(Y)) = \theta_2 D(Y).$$

Отсюда следует, что $\theta_2 = \dfrac{\mathrm{cov}(X,Y)}{\mathbf{D}(Y)}$, $\theta_1 = \mathbf{E}(X) - \theta_2 \mathbf{E}(Y)$. Теперь, если в этих выражениях вместо ковариации, дисперсии и математических ожиданий взять их выборочные характеристики из § 19 (формулы (19.7), (19.3)), то для θ_1, и θ_2 получим оценки (24.11) и (24.12).

Задачи

Задача 24.1. Зависимость между количеством различных видов внесенных удобрений y и урожайностью (в центнерах с гектара) X представлена следующей таблицей:

$\vec{y}$	1	2	3	4	5
$\vec{X}$	14	14.5	15.7	16.5	18

Предполагается линейная зависимость между количеством удобрений y и урожайностью X вида $X = \theta_2 y + \theta_1 + \mathcal{E}$ с неизменными параметрами θ_1, θ_2 и величиной случайного влияния $\mathcal{E}$ со средним ноль и конечной дисперсией. Оценить параметры θ_1 и θ_2 с помощью метода наименьших квадратов.

Задача 24.2. В книге "Основы химии" Д. И. Менделеева приводятся данные о растворимости азотнокислого натрия $NaNO_3$ в зависимости от температуры воды. В 100 частях воды растворяется следующее число условных частей $NaNO_3$ при соответствующих температурах:

$\vec{y}$	0	4	10	15	21	29	36	51	68
$\vec{X}$	66.7	71.0	76.3	80.6	85.7	92.9	99.4	113.6	125.1

Через $\vec{y}$ и $\vec{X}$ обозначены соответственно температура раствора и количество $NaNO_3$. На количество растворившегося $NaNO_3$ влияют случайные факторы. Предполагается, что имеет место линейная зависимость между температурой y и количеством растворившегося $NaNO_3$ вида $X = \theta_2 y + \theta_1 + \mathcal{E}$ с неизменными параметрами θ_1, θ_2 и величиной случайного влияния $\mathcal{E}$ со средним ноль и конечной дисперсией. Оценить параметры θ_1 и θ_2 с помощью метода наименьших квадратов.

Задача 24.3. В таблице приведены результаты 14 лабораторных испытаний прочности стальных канатов. Случайная величина Y характеризует сумму разрывных усилий (в кг) отдельных нитей, из которых свит канат, а X – разрывное усилие всего каната.

$\vec{Y}$	1436	1386	1365	1440	1390	1414	1424
$\vec{X}$	1135	1140	1141	1156	1150	1154	1145

$\vec{Y}$	1425	1463	1408	1435	1421	1435	1402
$\vec{X}$	1146	1164	1145	1145	1140	1150	1140

Используя модель линейной регрессии (24.10), оценить параметры θ_1 и θ_2.

Задача 24.4. Данные опыта приведены в таблице:

$\vec{y}$	0	2	4	6	8	10
$\vec{X}$	5	−1	0.5	1.5	4.5	8.5

Полагая, что y и X связаны зависимостью $X = \theta_3 y^2 + \theta_2 y + \theta_1 + \mathcal{E}$, где величина случайного влияния $\mathcal{E}$ имеет среднее ноль и конечную дисперсию, найти коэффициенты θ_1, θ_2 и θ_3 методом наименьших квадратов.

§ 25. Метод максимального правдоподобия

Пусть $\vec{X} = (X_1, X_2, \ldots, X_n)$ – случайная выборка, отвечающая случайной величине X с плотностью распределения $f(x, \theta)$, где θ – неизвестный параметр. Предполагается, что вид плотности известен, а параметр θ неизвестен. Например, $f(x, \theta)$ может быть плотностью нормального распределения с математическим ожиданием θ и дисперсией 1, т. е.

$$f(x, \theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}.$$

Задача состоит в том, чтобы, зная вид плотности, оценить неизвестный параметр θ по наблюдениям $X_1, X_2, \ldots, X_n$.

Определение 1. **Функцией правдоподобия** для случайных наблюдений $X_1, X_2, \ldots, X_n$ с плотностью распределения $f(x, \theta)$ называется функция $L(x_1, x_2, \ldots, x_n, \theta) = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta)$.

Оценкой параметра θ, построенной **методом максимального правдоподобия** по наблюдениям $X_1, X_2, \ldots, X_n$, называется оценка $\hat{\theta}_n(\vec{X})$, определяемая равенством

$$L(X_1, X_2, \ldots, X_n, \hat{\theta}_n(\vec{X})) = \max_{\theta} L(X_1, X_2, \ldots, X_n, \theta). \quad (25.1)$$

Иными словами эта оценка определяется следующим образом. В качестве первых n координат функции правдоподобия нужно взять случайные наблюдения $X_1, X_2, \ldots, X_n$, затем рассмотреть функцию $L(X_1, X_2, \ldots, X_n, \theta)$ как функцию одной переменной θ и определить такое значение переменной $\theta = \hat{\theta}_n$, при котором эта функция принимает максимальное значение. Ясно, что величина переменной $\hat{\theta}_n$ будет зависеть еще и от $\vec{X}$. Формально все это задано равенством (25.1). Почему именно так выбирается оценка $\hat{\theta}_n$? Дело в том, что в силу независимости наблюдений $X_1, X_2, \ldots, X_n$ их совместная плотность равна произведению плотностей величин X_l, которые имеют одинаковый вид и равны $f(x_l, \theta)$. Таким образом, функция $L(\vec{x}, \theta)$ и есть совместная плотность распределения случайной выборки $\vec{X}$. Для тех значений $\vec{x}$, для которых плотность $L(\vec{x}, \theta)$ больше, вероятность появления значений выборки $\vec{X}$ в малой окрестности $\vec{x}$ в ходе эксперимента тоже больше. Поэтому, если мы уже наблюдаем какие-то значения величин $X_1, X_2, \ldots, X_n$, то естественно предположить, что и значение плотности в этих точках, т. е. значение $L(X_1, X_2, \ldots, X_n, \theta)$, должно быть велико. Поэтому мы и выбираем оценку $\hat{\theta}_n(\vec{X})$ для истинного значения параметра θ так, чтобы $L(X_1, X_2, \ldots, X_n, \hat{\theta}_n(\vec{X}))$ было наибольшим возможным значением. Если функция дифференцируема и достигает максимального значения в какой-нибудь точке, то производная в этой точке должна быть равна нулю. Таким образом, для нахождения оценки $\hat{\theta}_n(\vec{X})$ имеем уравнение

$$\frac{\partial}{\partial \theta} L(\vec{X}, \theta) = 0. \quad (25.2)$$

Считаем, что совместная плотность L нигде не обращается в нуль, тогда уравнение (25.2) можно записать $\frac{1}{L}\frac{\partial}{\partial\theta}L = 0$. Пользуясь формулой для производной логарифма, перепишем его следующим образом: $\frac{\partial}{\partial\theta}\ln L(\vec{X},\theta) = 0$. Учитывая определение $L(\vec{X},\theta)$ и свойство логарифма $\ln(ab) = \ln a + \ln b$, окончательно получим

$$\frac{\partial}{\partial\theta}\sum_{l=1}^{n}\ln f(X_l,\theta) = 0. \qquad (25.3)$$

Это уравнение, в котором неизвестным является параметр θ, называется **уравнением максимального правдоподобия**. Если у него существует единственное решение $\hat{\theta}_n(\vec{X})$, то оно называется **оценкой максимального правдоподобия** для параметра θ.

Замечание 1. Метод максимального правдоподобия можно применять, когда $\vec{\theta} = (\theta_1,\ldots,\theta_m)$ – многомерный параметр. Функция максимального правдоподобия будет достигать максимума в той точке $\vec{\theta}$, в которой все частные производные по θ_k, $k = 1,2,\ldots,m$, равны нулю. Поэтому для нахождения оценки $\vec{\hat{\theta}}_n(\vec{X}) = (\hat{\theta}_n^{(1)}(\vec{X}),\hat{\theta}_n^{(2)}(\vec{X}),\ldots,\hat{\theta}_n^{(m)}(\vec{X}))$ вместо (25.2) имеем систему уравнений

$$\frac{\partial}{\partial\theta_k}L(\vec{X},\vec{\theta}) = 0, \qquad k = 1,2,\ldots,m.$$

Свойства оценки максимального правдоподобия.

1) Оценка максимального правдоподобия $\hat{\theta}_n(\vec{X})$ является состоятельной оценкой параметра θ, т. е. $\hat{\theta}_n(\vec{X}) \to \theta$ по вероятности при $n \to \infty$.

2) Оценка $\hat{\theta}_n(\vec{X})$ является асимптотически нормальной оценкой параметра θ с дисперсией $\Delta^2 = I_1^{-1}(\theta)$, где

$$I_1(\theta) = \int_{-\infty}^{\infty}\left(\frac{\partial}{\partial\theta}\ln f(x,\theta)\right)^2 f(x,\theta)dx,$$

– информационное количество Фишера относительно параметра θ, содержащееся в одном наблюдении. В силу замечания 1 § 22 оценка $\hat{\theta}_n(\vec{X})$ будет асимптотически эффективной оценкой параметра θ.

Свойства 1) и 2) справедливы при определенных условиях на плотность распределения $f(x,\theta)$. Доказательство этих свойств выходит за рамки элементарного курса математической статистики.

Пример 1. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает плотности распределения $f(x,\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)$. Найдем оценку для параметра θ методом максимального правдоподобия. Уравнение (25.3) в этом случае имеет вид

$$\frac{\partial}{\partial\theta} \sum_{l=1}^{n} \left(-\frac{(X_l-\theta)^2}{2} + \ln\left(\frac{1}{\sqrt{2\pi}}\right)\right) = 0.$$

Дифференцируя под знаком суммы, получим

$$\sum_{l=1}^{n} (X_l - \theta) = 0.$$

Решением этого уравнения является $\hat{\theta}_n(\vec{X}) = \frac{1}{n} \sum_{l=1} X_l$. Таким образом, оценка максимального правдоподобия для θ, когда $f(x,\theta)$ – плотность нормального распределения, есть выборочное среднее. Это естественно, поскольку θ является математическим ожиданием для нормального закона.

Пример 2. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает показательному распределению с параметром $\theta^{-1}, \theta > 0$, с плотностью

$$f(x,\theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Найдем оценку для параметра θ методом максимального правдоподобия. Уравнение (25.3) в этом случае преобразуется к виду

$$\sum_{l=1}^{n} \left(\frac{X_l}{\theta^2} - \frac{1}{\theta}\right) = 0.$$

Откуда находим решение $\hat{\theta}_n(\vec{X}) = \frac{1}{n} \sum_{l=1}^{n} X_l$, которое, как и в предыдущем примере, есть выборочное среднее.

Замечание 2. Метод максимального правдоподобия в полной мере применим и для случая, когда $\vec{X} = (X_1, \ldots, X_n)$ – случайная выборка, отвечающая случайной величине X, принимающей дискретные значения $x_k, k = 1, 2, \ldots$, с вероятностями $p(x_k, \theta) = \mathbf{P}(X_1 = x_k)$. Предполагается, что вид функции известен, а параметр θ неизвестен. **Функция правдоподобия** для дискретной величины имеет вид $L(x_1, x_2, \ldots, x_n, \theta) = p(x_1, \theta)p(x_2, \theta)\ldots p(x_n, \theta)$. Как и в случае с плотностью, оценка, построенная методом максимального правдоподобия, определяется равенством (25.1). Это равенство, если функция $p(x, \theta)$ дифференцируема по θ, влечет **уравнение максимального правдоподобия**

$$\frac{\partial}{\partial\theta}\sum_{l=1}^{n}\ln p(X_l, \theta) = 0. \tag{25.4}$$

Если существует единственное решение $\hat{\theta}_n(\vec{X})$, то оно называется **оценкой максимального правдоподобия** для параметра θ.

Пример 3. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает распределению Пуассона с параметром θ, т. е.

$$p(m, \theta) = \mathbf{P}(X_1 = m) = \frac{\theta^m}{m!}e^{-\theta}, \qquad m = 0, 1, 2, \ldots.$$

Найдем оценку для параметра θ методом максимального правдоподобия. Уравнение (25.4) в этом случае превращается в следующее

$$\frac{\partial}{\partial\theta}\sum_{l=1}^{n}\big(-\theta + X_l\ln\theta - \ln(X_l!)\big) = 0.$$

После дифференцирования получим

$$-n + \frac{1}{\theta}\sum_{l=1}^{n}X_l = 0.$$

Отсюда для параметра θ находим оценку максимального правдоподобия $\hat{\theta}_n(\vec{X}) = \frac{1}{n}\sum\limits_{l=1}^{n}X_l$, которая, как и в предыдущем примере, есть выборочное среднее.

Пример 4. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает равномерному распределению, сосредоточенному на интервале единичной длины с неизвестным центром θ, т. е. распределению с плотностью

$$f(x,\theta) = \mathbb{I}_{[\theta-\frac{1}{2},\theta+\frac{1}{2}]}(x),$$

где $\mathbb{I}_A(x)$ – индикатор множества A, определенный в п.1 § 12. Равенство (25.1) принимает вид

$$\prod_{l=1}^{n} \mathbb{I}_{[\hat\theta_n-\frac{1}{2},\hat\theta_n+\frac{1}{2}]}(X_l) = \max_\theta \prod_{l=1}^{n} \mathbb{I}_{[\theta-\frac{1}{2},\theta+\frac{1}{2}]}(X_l). \qquad (25.5)$$

Хотя плотность $f(x,\theta)$ не дифференцируема для всех θ, оценку $\hat\theta_n$, удовлетворяющую (25.5), найти можно. Ясно, что максимальное значение правой части в (25.5) равно единице, и это происходит тогда, когда интервал $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ содержит все значения $X_1, X_2, \ldots, X_n$. Пусть $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$ – вариационный ряд. Интервал $[\theta-\frac{1}{2}, \theta+\frac{1}{2}]$ будет содержать все значения X_k, $k = 1, 2, \ldots, n$, если $\theta-\dfrac{1}{2} \leqslant X_{(1)}$ и $X_{(n)} \leqslant \theta+\dfrac{1}{2}$. Таких значений θ – целый интервал $\left[X_{(n)}-\dfrac{1}{2}, X_{(1)}+\dfrac{1}{2}\right]$, и любое из них может служить оценкой максимального правдоподобия. В качестве одного из возможных возьмем $\hat\theta_n = \dfrac{1}{2}\big(X_{(n)} + X_{(1)}\big)$. Оно лежит внутри интервала, поскольку равномерное распределение сосредоточено на интервале единичной длины, что влечет $X_{(n)} - X_{(1)} \leqslant 1$, и

$$\hat\theta_n - \frac{1}{2} = X_{(1)} + \frac{1}{2}\big(X_{(n)} - X_{(1)}\big) - \frac{1}{2} \leqslant X_{(1)},$$

$$\hat\theta_n + \frac{1}{2} = X_{(n)} + \frac{1}{2} - \frac{1}{2}\big(X_{(n)} - X_{(1)}\big) \geqslant X_{(n)}.$$

Следовательно, $\hat\theta_n$ удовлетворяет (25.5), и ее можно считать оценкой, построенной методом максимального правдоподобия.

Задачи

Задача 25.1. Страховая компания за год произвела n выплат по портфелю однотипных договоров страхования, размер выплат по которым имеет распределение с плотностью $f(x) = \theta e^{-\theta x}$, $x > 0$. Найти оценку максимального правдоподобия для θ.

Задача 25.2. На заводе производят шарики для шарикоподшипников. Из продукции, произведенной за час, выбрали 20 шариков. Измерение их диаметров дало следующие результаты: 5.05, 5.01, 5.25, 4.97, 4.99, 5.03, 4.97, 5.13, 5.02, 4.90, 5.01, 4.75, 4.95, 4.98, 5.05, 5.02, 4.88, 5.00, 5.04, 4.96 мм. При предположении, что диаметр имеет нормальное распределение с неизвестным средним и дисперсией, построить оценку максимального правдоподобия для этих параметров.

Задача 25.3. Пусть $X_1, X_2, \ldots, X_n$ независимы и равномерно распределены на интервале $\theta, \theta+1$. Показать, что любая точка интервала $\left[\max\limits_{1\leqslant i\leqslant n} X_i - 1, \min\limits_{1\leqslant i\leqslant n} X_i\right]$ является оценкой максимального правдоподобия для параметра θ.

§ 26. Процедура рекуррентного оценивания

Пусть как в предыдущем параграфе по случайной выборке $\vec{X} = (X_1, X_2, \ldots, X_n)$, отвечающей случайной величине X с известным видом плотности распределения $f(x, \theta)$, нужно оценить неизвестный параметр θ.

В большинстве случаев точное решение уравнения максимального правдоподобия

$$\sum_{l=1}^{n} \frac{\partial}{\partial \theta} \ln f(X_l, \theta) = 0 \tag{26.1}$$

найти не удается. Поэтому для него ищут различные приближенные выражения. Опишем один интересный подход к решению этой проблемы.

Рассмотрим решения уравнения (26.1) при n и при $n+1$, т. е. рассмотрим $\hat{\theta}_n$ и $\hat{\theta}_{n+1}$, удовлетворяющие равенствам

$$\sum_{l=1}^{n} \frac{\partial}{\partial \theta} \ln f(X_l, \hat{\theta}_n) = 0, \qquad \sum_{l=1}^{n+1} \frac{\partial}{\partial \theta} \ln f(X_l, \hat{\theta}_{n+1}) = 0.$$

Вычитая из правого равенства левое, получим

$$\frac{\partial}{\partial\theta}\ln f(X_{n+1},\hat\theta_n)+ \\ +\sum_{l=1}^{n+1}\Big(\frac{\partial}{\partial\theta}\ln f(X_l,\hat\theta_{n+1})-\frac{\partial}{\partial\theta}\ln f(X_l,\hat\theta_n)\Big)=0. \qquad (26.2)$$

Поскольку оценка максимального правдоподобия является состоятельной оценкой параметра θ, т. е. $\hat\theta_n\to\theta$ при $n\to\infty$, то разность $\hat\theta_{n+1}-\hat\theta_n$ стремится к нулю. Предположим, что у функции $f(x,\theta)$ существует непрерывная вторая производная по параметру θ. Тогда функция $g(\theta)=\dfrac{\partial}{\partial\theta}\ln(x,\theta)$ при любом фиксированном x обладает непрерывной производной по θ, и, в силу определения производной, приращение этой функции можно выразить следующим образом $g(\hat\theta_{n+1})-g(\hat\theta_n)\approx g'(\hat\theta_n)(\hat\theta_{n+1}-\hat\theta_n)$. Используя это, преобразуем (26.2) к виду

$$\frac{\partial}{\partial\theta}\ln f(X_{n+1},\hat\theta_n)+\sum_{l=1}^{n+1}\frac{\partial^2}{\partial\theta^2}\ln f(X_l,\hat\theta_n)(\hat\theta_{n+1}-\hat\theta_n)\approx 0$$

или

$$\hat\theta_{n+1}-\hat\theta_n\approx-\frac{1}{\sum\limits_{l=1}^{n+1}\dfrac{\partial^2}{\partial\theta^2}\ln f(X_l,\hat\theta_n)}\frac{\partial}{\partial\theta}\ln f(X_{n+1},\hat\theta_n). \qquad (26.3)$$

Величины $X_l, l=1,2,\ldots,n+1$, независимы и одинаково распределены. Для любого фиксированного θ по закону больших чисел (§ 16)

$$\frac{1}{n+1}\sum_{l=1}^{n+1}\frac{\partial^2}{\partial\theta^2}\ln f(X_l,\theta)\approx\mathbf{E}\Big(\frac{\partial^2}{\partial\theta^2}\ln f(X_1,\theta)\Big)=-I_1(\theta), \qquad (26.4)$$

где $I_1(\theta)$ – информационное количество Фишера. Равенство в правой части следует из (22.6).

При некоторых дополнительных условиях можно доказать, что соотношение (26.4) выполняется равномерно по всем θ, достаточно близким к заданному значению, поскольку функция $\frac{\partial^2}{\partial\theta^2}f(x,\theta)$ непрерывна по θ. С учетом этого, соотношение (26.3) эквивалентно следующему

$$\hat{\theta}_{n+1}-\hat{\theta}_n \approx \frac{1}{(n+1)I_1(\hat{\theta}_n)}\frac{\partial}{\partial\theta}\ln f(X_{n+1},\hat{\theta}_n).$$

Замена приближенного равенства в этом соотношении на строгое приводит к следующей **процедуре рекуррентного оценивания** неизвестного параметра θ:

$$\tilde{\theta}_{n+1}=\tilde{\theta}_n+\frac{1}{(n+1)I_1(\tilde{\theta}_n)}\frac{\partial}{\partial\theta}\ln f(X_{n+1},\tilde{\theta}_n), \qquad \tilde{\theta}_0=c, \qquad (26.5)$$

где c – произвольная постоянная.

Эта процедура замечательна тем, что приближенное значение для параметра θ на $n+1$ шаге выражается только через приближенное значение на n-м шаге и через наблюдение X_{n+1}. Это и означает, что оценивание рекуррентное. При определенных условиях можно доказать, что оценка $\tilde{\theta}_n$ ничем не хуже оценки максимального правдоподобия $\hat{\theta}_n$, поскольку она обладает теми же свойствами 1) и 2). Следовательно, она является асимптотически эффективной оценкой параметра θ.

Пример 1. Пусть, как и в примере 1 § 25,

$$f(x,\theta)=\frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{(x-\theta)^2}{2}\Big).$$

Тогда $\frac{\partial}{\partial\theta}f(x,\theta)/f(x,\theta)=x-\theta$. Из примера 1 § 22 следует, что $I_1(\theta)=1$. Следовательно, процедура (26.5) для этого случая имеет вид

$$\tilde{\theta}_{n+1}=\tilde{\theta}_n+\frac{1}{n+1}(X_{n+1}-\tilde{\theta}_n), \qquad \tilde{\theta}_0=c. \qquad (26.6)$$

Из этой процедуры легко вычислить явное выражение для $\tilde{\theta}_n$:

$$\tilde{\theta}_1=c+1(X_1-c)=X_1, \qquad \tilde{\theta}_n=\frac{1}{n}\sum_{k=1}^{n}X_k=\overline{X}_n.$$

Следовательно, в случае нормальной плотности распределения рекуррентная оценка математического ожидания совпадает с оценкой максимального правдоподобия и совпадает с выборочным средним. Это и не удивительно, поскольку для этой плотности все вышеприведенные приближенные равенства превращаются в строгие равенства.

Пример 2. Пусть выборка $X_1, X_2, \ldots, X_n$ отвечает показательному распределению с параметром $\theta^{-1}, \theta > 0$, т. е. распределению с плотностью

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & \text{при } x \geqslant 0, \\ 0, & \text{при } x < 0. \end{cases}$$

Тогда $\frac{\partial}{\partial\theta} f(x,\theta)/f(x,\theta) = \frac{x}{\theta^2} - \frac{1}{\theta}$. Из примера 2 § 22 следует, что $I_1(\theta) = 1/\theta^2$. Процедура (26.5) в этом случае имеет вид

$$\tilde{\theta}_{n+1} = \tilde{\theta}_n + \frac{1}{(n+1)\frac{1}{\tilde{\theta}_n^2}} \left(\frac{X_{n+1}}{\tilde{\theta}_n^2} - \frac{1}{\tilde{\theta}_n} \right), \qquad \tilde{\theta}_0 = c,$$

т. е. снова имеет вид (26.6), и снова $\tilde{\theta}_n = \hat{\theta}_n = \overline{X}_n$.

Ответы и решения к задачам

Теория вероятностей

1.1. $5 \cdot 4 = 20$.

1.2. $C_{10}^3 = 120$.

1.3. $A_{30}^2 = 30 \cdot 29 = 870$.

1.4. $C_9^3 C_5^2\, 3 = 2520$.

1.5. $C_8^2 - 2 = 26$.

1.6. $A_5^2 = 20$.

1.7. $6^5 = 7776$.

1.8. $(n!)^2$.

1.9. Числа 1, 2, 3 можно разместить так, чтобы они стояли рядом $n-2$ способами. Упорядочить остальные $n-3$ мест можно $(n-3)!$ способами. Ответ $(n-2)(n-3)! = (n-2)!$.

1.10. $2 + 2^2 + 2^3 + 2^4 + 2^5 = 62$.

1.11. $32^3 \cdot 10^4 = 327680000$.

1.12. Поскольку числа не могут начинаться нулем, ответ следующий: $A_{10}^6 - A_9^5 = 9 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 136080$.

1.13. Приведем два решения этой задачи. Число людей, у которых есть k белых зубов, равно C_{32}^k, поскольку порядок зубов не имеет значения. Следовательно, всего людей с учетом свойства 1) для биномиальных коэффициентов будет $\sum\limits_{k=0}^{32} C_{32}^k = 2^{32} = 1024 \cdot 1024 \cdot 1024 \cdot 4 = 4294967296$.

Другое решение. Сопоставим наличию белого зуба на той или иной позиции среди 32 возможных цифру 1, черному – цифру 0. Тогда возникает полное соответствие между зубами и 32 разрядными двоичными числами. Следовательно, число людей будет совпадать с числом 32 разрядных двоичных чисел. Их число очевидно равно 2^{32}.

1.14. Из слова КОЛОКОЛ машинные слова можно составлять следующим образом. Сначала, например, на 7 свободных для букв позиций можно разместить буквы О. Таких возможностей C_7^3, так как буквы О не различаются. На оставшиеся 4 позиции можно разместить 2 буквы К C_4^2 способами. Буквы Л займут два последних места. Таким образом, всего можно составить $C_7^3 C_4^2 = \frac{7!}{3!2!2!} = 5 \cdot 6 \cdot 7 =$

210 машинных слов. Из слова ВОДОРОД можно составить $C_7^3C_4^2C_2^1 = \frac{7!}{3!2!} = 2 \cdot 5 \cdot 6 \cdot 7 = 420$ машинных слов.

1.15. Места можно пронумеровать от 1 до 14. Можно на нечетные места посадить женщин, а на четные мужчин, и наоборот. Женщин можно разместить 7! способами, и столькими же способами можно разместить мужчин. Всего способов $2(7!)^2$.

1.16. Конфеты расположим в ряд. Между конфетами будет 8 промежутков, в которые C_8^4 способами можно поместить 4 разделительных перегородки по одной в каждый промежуток. Эти перегородки разделяют все конфеты на 5 групп, которые соответствуют 5 различным пакетам. Следовательно, на первый вопрос ответ следующий $C_8^4 = \frac{8!}{4!4!} = 70$. В случае, когда пакеты могут быть пустыми, конфеты и перегородки могут располагаться в ряд в любом порядке, в частности несколько перегородок могут следовать друг за другом. Всего таких возможностей C_{13}^4, поскольку среди $9 + 4 = 13$ позиций 4 можно выбрать для перегородок. Ответ на второй вопрос следующий: $C_{13}^4 = \frac{13!}{4!9!} = 715$.

2.1. $A - B = A\overline{B} =${число оканчивается цифрой 5}.

2.2. $\overline{A} =${ни одного бракованного изделия}, $\overline{B} =${либо одно бракованное изделие, либо бракованных нет}.

2.3. а) $A = \varnothing, B = \Omega$; б) никогда; в) $A = B$.

2.4. Нет.

2.5. $AB =${на одной кости выпала 1, а на другой – 4}, $A\overline{B} =${на одной кости выпала 2, а на другой – 3}.

2.6. По формуле 11) $\overline{A+B} = \overline{A}\overline{B}$ – выпала нечетная сумма, и ни на одной кости не выпала шестерка, т.е. на одной кости выпало четное число очков, отличное от шести, а на другой – нечетное число очков.

2.7. $A - B = A\overline{B}$ – одна карта туз пик и другая карта пиковой масти.

2.8. $a)$ $\prod\limits_{i=1}^{n} \overline{A_i}$, $b)$ $\sum\limits_{i=1}^{n} A_i$.

2.9. $ABC =${оба спортсмена прыгнули дальше 7 метров, причем мужчина прыгнул дальше женщины}, $A - AB = A\overline{B} =${женщина прыгнула дальше 7 метров, а мужчина не дальше женщины}, $A\overline{B}C$ ={оба спортсмена прыгнули дальше 7 метров, но мужчина прыгнул не дальше женщины}.

2.10. а) $A\overline{B}\,\overline{C}$; б) $AB\overline{C}$; в) ABC; г) $A+B+C$; д) $AB+AC+BC$; е) $A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C$; ж) $\overline{A}BC+A\overline{B}C+AB\overline{C}$; з) $\overline{A}\,\overline{B}\,\overline{C}$; и) $\overline{ABC}$.

3.1. Число должно оканчиваться цифрами 0, 2, 4, 5, 6, 8 из десяти возможных. Поэтому $P=\frac{3}{5}$.

3.2. Различаем кубики. Тогда при бросании двух кубиков равновозможны $6\cdot 6 = 36$ пар чисел. Сумму очков, кратную четырем, составляют три несовместных события: $A_1 = \{4 \text{ очка}\}$, $A_2 = \{8 \text{ очков}\}$, $A_3 = \{12 \text{ очков}\}$. Событию A_1 соответствуют пары чисел (1,3), (2,2), (3,1), при этом $\mathbf{P}(A_1)=\frac{3}{36}$. Событию A_2 соответствуют пары чисел (2,6), (3,5), (4,4), (5,3), (6,2), при этом $\mathbf{P}(A_2)=\frac{5}{36}$. И, наконец, событию A_3 соответствует пара (6,6), при этом $\mathbf{P}(A_3)=\frac{1}{36}$. Тогда $\mathbf{P}(\text{сумма кратна четырем}) = \frac{3}{36}+\frac{5}{36}+\frac{1}{36}=\frac{1}{4}$.

3.3. $P=\dfrac{C_{20}^2-20}{C_{20}^2}=\dfrac{17}{19}$.

3.4. $P=\dfrac{A_{10}^5}{10^5}=0.3024$.

3.5. $P=\dfrac{C_5^3}{C_{15}^3}=\dfrac{2}{91}$.

3.6. $P=\dfrac{C_{18}^4}{C_{20}^6}=\dfrac{3}{38}$.

3.7. $P=\dfrac{C_{99}^9}{C_{100}^{10}}=\dfrac{1}{10}$.

3.8. Выразим искомую вероятность через вероятность дополнительного события, которое состоит в том, что нет карт пиковой масти. Тогда

$$P=1-\frac{C_{27}^2}{C_{36}^2}=1-\frac{27\cdot 26}{36\cdot 35}=1-\frac{39}{70}=\frac{31}{70}.$$

3.9. $P=1-\dfrac{C_8^6}{C_{12}^6}=1-\dfrac{1}{33}=\dfrac{32}{33}$.

3.10. $\mathbf{P}(\{1,1\}+\{1,2\}+\{2,1\})=\dfrac{3}{36}=\dfrac{1}{12}$.

3.11. $P=\dfrac{C_{50}^2}{C_{60}^2}=\dfrac{245}{354}\approx 0.692$.

3.12. а) $\frac{8\cdot 8}{C_{16}^2}=\frac{8}{15}$; б) $\frac{7}{15}$.

3.13. $P=\frac{1}{A_5^2}=\frac{1}{20}$.

3.14. Ответ $\frac{12!}{12^{12}}$, по формуле Стирлинга, преобразуется к виду $\frac{12!}{12^{12}}\approx\frac{\sqrt{2\pi 12}}{e^{12}}\approx\frac{8.683515}{162754.7914}\approx 0.00005335$.

3.15. $P=1-\frac{10\cdot 6}{C_{16}^2}=\frac{1}{2}$.

3.16. $P=\frac{6}{6^4}=\frac{1}{216}$.

3.17. Вероятность того, что при n бросаниях хотя бы один раз выпала шестерка, равна 1 минус вероятность того, что она ни разу не выпадет при n бросаниях, и очевидно равна $1-\frac{5^n}{6^n}$.

а) Нужно найти такое минимальное n, что $1-\frac{5^n}{6^n}\geqslant 0.5$. Несложно подсчитать, что $n=3$.

б) Минимальное n такое, что $1-\frac{5^n}{6^n}\geqslant 0.9$ равно 13.

3.18. $\mathbf{P}(2\text{К},2\text{Д})=\frac{C_4^2C_4^2}{C_{32}^4}=\frac{4\cdot3\cdot3\cdot2\cdot3\cdot4}{36\cdot35\cdot34\cdot33}=\frac{4}{6545}$.

$\mathbf{P}(3\text{Т},6)=\frac{C_4^3\cdot 4}{C_{32}^4}=\frac{4\cdot4\cdot2\cdot3\cdot4}{36\cdot35\cdot34\cdot33}=\frac{16}{58905}$.

3.19. $\mathbf{P}(2\text{Ч})=\frac{C_9^2\cdot 27}{C_{36}^3}=\frac{3\cdot 27}{35\cdot 17}=\frac{81}{595}$.

3.20. $\mathbf{P}(\text{шары разного цвета})=\mathbf{P}(\text{БК})+\mathbf{P}(\text{БЧ})+\mathbf{P}(\text{КЧ})=$
$=\frac{5\cdot 6}{C_{14}^2}+\frac{5\cdot 3}{C_{14}^2}+\frac{6\cdot 3}{C_{14}^2}=\frac{3\cdot(10+5+6)}{7\cdot 13}=\frac{9}{13}$.

3.21. $P=\frac{C_{32}^3}{C_{36}^3}=\frac{32\cdot31\cdot30}{36\cdot35\cdot34}=\frac{31\cdot 8}{21\cdot 17}=\frac{248}{357}$.

3.22. $P=\frac{C_4^2C_{16}^1}{C_{20}^3}=\frac{144}{1615}$

3.23. $P=1-\frac{2}{36}-\frac{1}{36}=\frac{11}{12}$.

3.24. $P=\frac{A_8^4}{8^4}=\frac{8\cdot7\cdot6\cdot5}{8^4}=\frac{105}{256}$.

3.25. Два белых шара могут быть на любом из шести мест в порядке выемки. Таких возможностей C_6^2. Зарезервировав за черным шаром последнее место, имеем C_5^2 возможностей для белых шаров занять два из первых пяти мест. Поэтому ответ $P=\frac{C_5^2}{C_6^2}=\frac{4}{6}=\frac{2}{3}$.

3.26. $P=1-\frac{4\cdot2\cdot3!}{5!}=1-\frac{2}{5}=\frac{3}{5}$.

3.27. $\dfrac{2(k-1)(n-k)}{n(n-1)}$.

3.28. $P = 1 - \dfrac{14}{63} = \dfrac{7}{9}$.

4.1. $P = (a-2r)^2/a^2$, где $a = 4, r = 1$, т. е. $P = 2^2/16 = 1/4$.

4.2. $P = \dfrac{2r^2}{\pi r^2} = \dfrac{2}{\pi}$.

4.3. Данная задача является частным случаем задачи о встрече. Время 11 часов можно взять за начало отсчета. Тогда $T = 1$, $\tau = \dfrac{1}{4}$, и согласно решению задачи 2 имеем $P = 1 - \left(1 - \frac{1}{4}\right)^2 = \dfrac{7}{16}$.

4.4. Каждой паре чисел отвечает точка в квадрате $[0,1] \times [0,1]$ с координатами (x,y). Все точки квадрата равновозможны. Событиям, связанным со случайным выбором точки из квадрата, отвечают геометрические вероятности. Событию $\{$одно число более чем вдвое меньше другого$\}$ соответствуют два случая: 1) $2x < y$, $x \in [0,1]$, $y \in [0,1]$; 2) $2y < x$, $x \in [0,1]$, $y \in [0,1]$. В первом случае – это точки квадрата, лежащие выше прямой $y = 2x$, а во втором – ниже прямой $y = \dfrac{1}{2}x$. Площади соответствующих треугольников равны по $1/4$. В результате искомая вероятность равна $P = \dfrac{1}{4} + \dfrac{1}{4} = \dfrac{1}{2}$.

4.5. Это задача на геометрические вероятности. Событие $\{$сумма чисел заключена между $\dfrac{1}{2}$ и $1\}$ можно представить множеством $\dfrac{1}{2} \leqslant x + y \leqslant 1$. Это точки, лежащие между двумя прямыми $y = \dfrac{1}{2} - x$, $y = 1 - x$. С учетом того, что пара чисел (x,y) выбирается из квадрата $[0,1]\times[0,1]$, имеем $P = \dfrac{1}{2} - \dfrac{1}{8} = \dfrac{3}{8}$.

4.6. $P = 1/4$.

4.7. Пусть x – длина отрезка $[A, L]$, y – длина отрезка $[A, M]$. Тогда условие задачи запишется в виде $|y - x| < x$. Это влечет $0 < y < 2x$, и, следовательно, искомая вероятность равна

$$P = \frac{l^2 - \frac{l^2}{4}}{l^2} = \frac{3}{4}.$$

4.8. Две кривые $x + y = 1$ и $xy = \dfrac{2}{9}$ пересекаются в двух точках. Координаты абсцисс у них следующие: $x_1 = \dfrac{1}{3}, x_2 =$

$=\frac{2}{3}$. Ответ в задаче $P=\frac{1}{3}+\int\limits_{1/3}^{2/3}\frac{2dx}{9x}=\frac{1}{3}+\frac{2}{9}\ln 2$.

4.9. Обозначим через x, y и $l-x-y$ длины отрезков. Ограничение $0\leqslant x+y\leqslant l$ следует из условия задачи. Следующие ограничения вытекают из свойства треугольника, что сумма двух любых сторон больше третей: $0\leqslant x\leqslant\frac{l}{2}$, $0\leqslant y\leqslant\frac{l}{2}$, $x+y\geqslant\frac{l}{2}$. В результате искомая вероятность равна

$$P=\frac{\operatorname{mes}\left\{x,y:0\leqslant x\leqslant\frac{l}{2},0\leqslant y\leqslant\frac{l}{2},x+y\geqslant\frac{l}{2}\right\}}{\operatorname{mes}\{x,y:0\leqslant x,0\leqslant y,0\leqslant x+y\leqslant l\}}=\frac{1}{4}.$$

4.10. $P=1-\frac{\delta}{H}\sqrt{1+\frac{H^2}{4\pi^2R^2}}$.

4.11. $P=\frac{3-\sqrt{3}}{6}\approx 0.2113$.

4.12. Принимаем 20 минут движения катера за единицу длины. Тогда

$$P=\frac{2}{3}\cdot\frac{1}{3}+2\cdot\frac{1}{3}\int\limits_0^1\frac{x+1}{3}dx=\frac{2}{9}+\frac{1}{3}=\frac{5}{9}.$$

Это решение не является очевидным.

5.1. Для колоды с 52 картами вычислим вероятности:

$$\mathbf{P}(A)=\frac{4}{52}=\frac{1}{13};\quad \mathbf{P}(B)=\frac{26}{52}=\frac{1}{2};\quad \mathbf{P}(AB)=\frac{2}{52}=\frac{1}{26}.$$

Очевидно, что $\mathbf{P}(AB)=\mathbf{P}(A)\mathbf{P}(B)$, и, следовательно, события A и B независимы. Иными словами, сведение о том, что событие B произошло, не дает никакой информации о вероятности события A.

Для колоды с 54 картами имеем:

$$\mathbf{P}(A)=\frac{4}{54}=\frac{2}{27};\quad \mathbf{P}(B)=\frac{26}{54}=\frac{13}{27};\quad \mathbf{P}(AB)=\frac{2}{54}=\frac{1}{27},$$

и $\mathbf{P}(AB)\neq\mathbf{P}(A)\mathbf{P}(B)$. Следовательно, события A и B зависимы. При этом $\mathbf{P}(A|B)=\frac{\mathbf{P}(AB)}{\mathbf{P}(B)}=1/13$.

5.2. События AB_1 и AB_2 несовместны, так как $B_1B_2 = \varnothing$. Тогда

$$\mathbf{P}(A(B_1 + B_2)) = \mathbf{P}(AB_1 + AB_2) = \mathbf{P}(AB_1) + \mathbf{P}(AB_2) =$$

$$= \mathbf{P}(A)\mathbf{P}(B_1) + \mathbf{P}(A)\mathbf{P}(B_2) = \mathbf{P}(A)(\mathbf{P}(B_1) + \mathbf{P}(B_2)) =$$

$$= \mathbf{P}(A)\mathbf{P}(B_1 + B_2).$$

5.3. Зависимы, поскольку $\mathbf{P}(A) = \frac{1}{2}, \mathbf{P}(B) = 1 - \frac{1}{2^3} = \frac{7}{8}$, а $\mathbf{P}(AB) = \frac{1}{2}\left(1 - \frac{1}{2^2}\right) = \frac{3}{8}$.

5.4. Сумму, равную восьми, составляют следующие пары цифр: $(2,6)$, $(3,5)$, $(4,4)$, $(5,3)$, $(6,2)$. Тогда $\mathbf{P}(3,5) = \frac{2}{5}$.

5.5. Обозначим события $A = \{$четная сумма$\}$, $B = \{$шестерка хотя бы на одном из двух кубиков$\}$. Тогда $\{AB\} = \{(6,6),(6,4),(6,2),(2,6),(4,6)\}$ и $\mathbf{P}(AB) = \frac{5}{36}$. Кроме того, $\mathbf{P}(B) = 1 - \frac{5}{36} = \frac{11}{36}$. В результате имеем $\mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)} = \frac{5}{11}$.

5.6. Вынули сразу три шара или поочередно – одно и тоже. Поэтому

$\mathbf{P}(1\text{-й б.}, 2\text{-й б.}, 3\text{-й б.}) =$
$= \mathbf{P}(1\text{-й б.})\mathbf{P}(2\text{-й б.}|1\text{-й б.})\mathbf{P}(3\text{-й б.}|1\text{-й б.}, 2\text{-й б.}) =$
$= \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{7}{8} = \frac{7}{10}$.

5.7. $\mathbf{P}(\text{Т}_1\text{Т}_2) = \mathbf{P}(\text{Т}_1)\mathbf{P}(\text{Т}_2|\text{Т}_1) = \frac{4}{36} \cdot \frac{3}{35} = \frac{1}{9} \cdot \frac{3}{35} = \frac{1}{105}$.

5.8. $\mathbf{P}(\text{ТДК}) = \mathbf{P}(\text{Т})\mathbf{P}(\text{К}|\text{Т})\mathbf{P}(\text{Д}|\text{ТК}) = \frac{1}{9} \cdot \frac{4}{35} \cdot \frac{2}{17} = \frac{8}{5355}$.

5.9. $\mathbf{P}(\text{ПОЛЕ}) = \mathbf{P}(\text{П})\mathbf{P}(\text{О}|\text{П})\mathbf{P}(\text{Л}|\text{ПО})\mathbf{P}(\text{Е}|\text{ПОЛ}) =$
$= \frac{1}{9} \cdot \frac{2}{8} \cdot \frac{1}{7} \cdot \frac{2}{6} = \frac{1}{756}$.

5.10. По формуле сложения вероятностей несовместных событий

$\mathbf{P}(\text{шары одного цвета}) =$
$= \mathbf{P}(\text{белые}) + \mathbf{P}(\text{черные}) + \mathbf{P}(\text{красные}) =$
$= \frac{5}{24} \cdot \frac{10}{24} + \frac{11}{24} \cdot \frac{8}{24} + \frac{8}{24} \cdot \frac{6}{24} = \frac{186}{578} \approx 0.323$.

5.11. Обозначим искомую вероятность P. Положим $A = \{$первый сделал больше выстрелов, чем второй$\}$, $A_1 = \{$первый стрелок попал при первом выстреле$\}$, $B_1 = \{$второй стрелок попал при первом выстреле}. Поскольку после двух выстрелов, если оба промахнулись, все начинается снова, то имеем $A = \overline{A}_1(B_1 + \overline{B}_1 A)$. Следовательно,

$$\mathbf{P}(A) = \mathbf{P}(\overline{A}_1)(\mathbf{P}(B_1) + \mathbf{P}(\overline{B}_1)\mathbf{P}(A)),$$

и $P = 0.8(0.3 + 0.7P)$), т. е. $P = \frac{6}{11}$.

5.12. Событие A является произведением событий $A_k = \{k\text{-я деталь качественная}\}$, $k = 1, 2, \ldots, 5$. Вычислим вероятность q события $A = A_1A_2A_3A_4A_5$ – партия принята. Очевидно $\mathbf{P}(A_1) = \frac{95}{100}$, так как всего деталей 100, а качественных 95. После осуществления события A_1 деталей останется 99, среди которых качественных 94, поэтому $\mathbf{P}(A_2|A_1) = \frac{94}{99}$. Аналогично

$$\mathbf{P}(A_3|A_1A_2) = \frac{93}{98},\ \mathbf{P}(A_4|A_1A_2A_3) = \frac{92}{97},\ \mathbf{P}(A_5|A_1A_2A_3A_4) = \frac{91}{96}.$$

По общей формуле умножения вероятностей находим

$$q = \frac{95}{100} \cdot \frac{94}{99} \cdot \frac{93}{98} \cdot \frac{92}{97} \cdot \frac{91}{96} \approx 0.77.$$

Искомая вероятность равна $p = 1 - q \approx 0.23$.

6.1. Первый бросивший монету игрок выигрывает при первом броске с вероятностью $\frac{1}{2}$. При втором броске он выигрывает, если при первом броске у него и у второго игрока выпало по решетке (вероятность этого $\frac{1}{4}$), а на втором броске у него выпал герб. В силу независимости событий вероятность выиграть на втором броске равна $\left(\frac{1}{4}\right) \cdot \frac{1}{2}$. Легко понять, что вероятность выиграть на k-м броске равна $\left(\frac{1}{4}\right)^k \cdot \frac{1}{2}$. Складывая эти вероятности, получим искомую вероятность:

$$p = \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} + \left(\frac{1}{4}\right)^2 \frac{1}{2} + \cdots + \left(\frac{1}{4}\right)^k \cdot \frac{1}{2} + \cdots = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{2}{3}.$$

Первый бросивший монету игрок выигрывает с вероятностью $p = \frac{2}{3}$, второй – $q = 1 - p = \frac{1}{3}$.

6.2. Решение легко понять из ответа для вероятности надежности прибора: $P = (p_1 + (1 - p_1)pp_2)(p_3 + (1 - p_3)pp_4)$.

6.3. Вероятность хотя бы одного попадания в мишень равна единице минус вероятность всех трех промахов. Пусть p – вероятность попадания, а q – вероятность промаха при одном выстреле. Тогда $0.875 = 1 - q^3$, и, следовательно, $q = 0.5$, $p = 1 - q = 0.5$.

6.4. $P = 1 - 0.9 \cdot 0.8 \cdot 0.6 = 1 - 0.432 = 0.568$.

6.5. Обозначим A_l ={попадание l-й бомбы}. Тогда

$$\mathbf{P}(\text{мост разрушен}) = \mathbf{P}(\overline{A}_1 A_2 A_3 + A_1 \overline{A}_2 A_3 + A_1 A_2 \overline{A}_3 + A_1 A_2 A_3)$$

$$= \mathbf{P}(\overline{A}_1)\mathbf{P}(A_2)\mathbf{P}(A_3) + \mathbf{P}(A_1)\mathbf{P}(\overline{A}_2)\mathbf{P}(A_3) +$$

$$+ \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(\overline{A}_3) + \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3) =$$

$$= 0.9 \cdot 0.3 \cdot 0.4 + 0.1 \cdot 0.7 \cdot 0.4 + 0.1 \cdot 0.3 \cdot 0.6 + 0.1 \cdot 0.3 \cdot 0.4 = 0.166.$$

6.6. $P = p_1(p_2 + p_3 - p_2 p_3)p_4(1 - p_5) + p_5$.

6.7. $P = (p_1 p_2 + p_3 - p_1 p_2 p_3)(p_4 + p_5 - p_4 p_5)(1 - p_6) + p_6$.

6.8. $P = \left(\frac{5}{6}\right)^3 \frac{1}{6} = \frac{5^3}{6^4} = \frac{125}{1296} \approx \frac{1}{10}$.

6.9. $P = \sum\limits_{m=1}^{\infty} q^{2m-1} p \Big|_{p=q=1/2} = \sum\limits_{m=1}^{\infty} \left(\frac{1}{2}\right)^{2m} = \dfrac{1}{4\left(1 - \frac{1}{4}\right)} = \dfrac{1}{3}$.

7.1. Обозначим $Б_1$ ={переложили белый}, $Ч_1$ ={переложили черный}, Ч ={вынуть черный шар}. По формуле полной вероятности

$$\mathbf{P}(Ч) = \mathbf{P}(Б_1)\mathbf{P}(Ч|Б_1) + \mathbf{P}(Ч_1)\mathbf{P}(Ч|Ч_1) =$$

$$= \frac{3}{11} \cdot \frac{8}{10} + \frac{8}{11} \cdot \frac{7}{10} = \frac{8}{10} \cdot \frac{10}{11} = \frac{8}{11}.$$

7.2. Обозначим $Б_1$ ={первый белый}, $Ч_1$ ={первый черный}, A ={следующий за первым – белый}. По формуле Байеса

$$\mathbf{P}(Б_1|A) = \frac{\mathbf{P}(Б_1)\mathbf{P}(A|Б_1)}{\mathbf{P}(Б_1)\mathbf{P}(A|Б_1) + \mathbf{P}(Ч_1)\mathbf{P}(A|Ч_1)} = \frac{\frac{3}{10} \cdot \frac{2}{9}}{\frac{3}{10} \cdot \frac{2}{9} + \frac{7}{10} \cdot \frac{3}{9}} = \frac{2}{9}.$$

7.3. Обозначим К ={выбрали кубик}, П ={выбрали пирамидку}, A ={выпала цифра 4}. По формуле Байеса

$$\mathbf{P}(\text{К}|A) = \frac{\mathbf{P}(\text{К})\mathbf{P}(A|\text{К})}{\mathbf{P}(\text{К})\mathbf{P}(A|\text{К}) + \mathbf{P}(\text{П})\mathbf{P}(A|\text{П})} = \frac{\frac{4}{5}\cdot\frac{1}{6}}{\frac{4}{5}\cdot\frac{1}{6}+\frac{1}{5}\cdot\frac{1}{4}} = \frac{8}{11}.$$

7.4. По формуле полной вероятности $P = \frac{8}{10}\cdot\frac{1}{2^3}+\frac{2}{10} = \frac{3}{10}$.

7.5. По формуле полной вероятности

$$\mathbf{P}(\text{В}) = \mathbf{P}(\text{ВВ})\mathbf{P}(\text{В}|\text{ВВ}) + \mathbf{P}(\text{ВН})\mathbf{P}(\text{В}|\text{ВН}) + \mathbf{P}(\text{НН})\mathbf{P}(\text{В}|\text{НН}) =$$

$$= \frac{C_4^2}{C_{12}^2}\cdot\frac{2}{10}+\frac{4\cdot 8}{C_{12}^2}\cdot\frac{3}{10}+\frac{C_8^2}{C_{12}^2}\cdot\frac{4}{10} = \frac{2}{12\cdot 11\cdot 10}\Big(\frac{4\cdot 3}{2}\cdot 2+4\cdot 8\cdot 3+\frac{8\cdot 7}{2}\cdot 4\Big) = \frac{1}{3}.$$

7.6. По формуле полной вероятности

$$\mathbf{P}(\text{В}_2) = \mathbf{P}(\text{В}_1)\mathbf{P}(\text{В}_2|\text{В}_1) + \mathbf{P}(\text{Н}_1)\mathbf{P}(\text{В}_2|\text{Н}_1) =$$

$$= \frac{5}{12}\cdot\frac{4}{11}+\frac{7}{12}\cdot\frac{5}{11} = \frac{5}{11}\cdot\frac{11}{12} = \frac{5}{12}.$$

7.7. По формуле Байеса

$$\mathbf{P}(\text{В}_1|\text{В}_2) = \frac{\mathbf{P}(\text{В}_1)\mathbf{P}(\text{В}_2|\text{В}_1)}{\mathbf{P}(\text{В}_1)\mathbf{P}(\text{В}_2|\text{В}_1) + \mathbf{P}(\text{Н}_1)\mathbf{P}(\text{В}_2|\text{Н}_1)} =$$

$$= \frac{\frac{5}{12}\cdot\frac{4}{11}}{\frac{5}{12}\cdot\frac{4}{11}+\frac{7}{12}\cdot\frac{5}{11}} = \frac{4}{4+7} = \frac{4}{11}.$$

7.8. $\mathbf{P}$(грузовая машина) $= \frac{3}{7}$.

7.9. $P = \frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\big(\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{2}\big)+\frac{1}{3}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{5}{12}$.

7.10. Начало решения аналогично решению задачи 2 (§ 7). Обозначим A ={в кабане две пули}, H_{110} ={попал первый, попал второй, не попал третий} и т. д. В силу независимости $\mathbf{P}(H_{110}) = 0.8\cdot 0.4\cdot 0.8 = 0.256$, $\mathbf{P}(H_{101}) = 0.8\cdot 0.6\cdot 0.2 = 0.096$, $\mathbf{P}(H_{011}) = 0.2\cdot 0.4\cdot 0.2 = 0.016$. По формуле Байеса $\mathbf{P}(H_{110}|A) = \frac{0.256}{0.368} = \frac{16}{23}$, $\mathbf{P}(H_{101}|A) = \frac{0.096}{0.368} = \frac{6}{23}$, $\mathbf{P}(H_{011}|A) = \frac{0.016}{0.368} = \frac{1}{23}$. Таким образом, первому и второму вместе нужно отдать 16 долей из 23, первому и третьему вместе – 6 долей, а второму и третьему вместе – 1 долю. Между собой они должны делить пополам, поскольку, если оба попали, то независимо от вероятностей попадания их вклад одинаков. Окончательно имеем, что первому нужно отдать $\frac{16}{23}\cdot\frac{1}{2}+\frac{6}{23}\cdot\frac{1}{2}=\frac{22}{46}$ доли добычи, второму – $\frac{16}{23}\cdot\frac{1}{2}+\frac{1}{23}\cdot\frac{1}{2}=\frac{17}{46}$, а третьему – $\frac{6}{23}\cdot\frac{1}{2}+\frac{1}{23}\cdot\frac{1}{2}=\frac{7}{46}$.

8.1. а) $P=P_{10}(5)=C_{10}^{5}\frac{1}{2^{10}}=\frac{63}{256}$;
б) $P=P_{10}(3,8)=1-\frac{1}{2^{10}}(1+C_{10}^{1}+C_{10}^{2}+C_{10}^{9}+1)=\frac{957}{1024}$.

8.2. Поскольку $P_4(2,4)=\frac{11}{16}\geqslant\frac{8}{16}=P_5(3,5)$, то вероятнее выиграть больше одной партии из четырех.

8.3. $P=\frac{8}{10}C_5^3\left(\frac{1}{2}\right)^5+\frac{2}{10}C_5^5\left(\frac{1}{2}\right)^5=\frac{8\cdot 10+2\cdot 1}{10\cdot 2^5}=\frac{41}{160}$.

8.4. $P=C_3^2\cdot\left(\frac{1}{6}\right)^2\cdot\frac{5}{6}+C_3^3\cdot\left(\frac{1}{6}\right)^3=\frac{1}{216}\cdot(3\cdot 5+1)=\frac{16}{216}=\frac{2}{27}$.

8.5. Три лампы горят с вероятностью $P=C_4^3\cdot\left(\frac{8}{10}\right)^3\cdot\frac{2}{10}=$ $=\frac{4\cdot 8^3\cdot 2}{10000}=0.4096$. Наивероятнейшее число лежит в пределах от $3.2-0.2$ до $3.2+0.8$. Следовательно, чисел два: 3 и 4.

8.6. Вероятность события, что при одном бросании не выпадет четверки, равна $p=\frac{3}{4}$. По формуле Бернулли вероятность того, что это событие наступит 4 раза при четырех бросаниях иеет вид:

$$P=C_4^4\left(\frac{3}{4}\right)^4=\left(\frac{3}{4}\right)^4=\frac{81}{256}.$$

8.7. Вероятность не менее трех попаданий при пяти выстрелах определяется формулой

$$C_5^3\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2+C_5^4\left(\frac{1}{4}\right)^4\left(\frac{5}{4}\right)+C_3^5\left(\frac{1}{4}\right)^5=\frac{5\cdot 4}{2}\cdot\frac{9}{45}+5\cdot\frac{3}{45}+\frac{1}{45}=$$

$$=\frac{1}{45}(90+15+1)=\frac{53}{2\cdot 4^4}=\frac{53}{512}.$$

8.8. Единица выпадет не более трех раз с вероятностью

$$P=C_4^0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^4+C_4^1\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^3+C_4^2\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2=$$

$$=\frac{3^4+4\cdot 3^3+6\cdot 3^2}{256}=\frac{243}{256}.$$

Другой подсчет этой вероятности производится по формуле

$$P = 1 - C_4^3\left(\frac{1}{4}\right)^3\frac{3}{4} - C_4^4\left(\frac{1}{4}\right)^4 = 1 - \frac{12}{4^4} - \frac{1}{4^4} = 1 - \frac{13}{256} = \frac{243}{256}.$$

8.9. Искомая вероятность равна $P = \sum\limits_{k=0}^{3} P_3^{(1)}(k)P_3^{(2)}(k) =$

$$= \sum_{k=0}^{3}(C_3^k)^2\left(\frac{6}{10}\right)^k\left(\frac{4}{10}\right)^{3-k}\left(\frac{7}{10}\right)^k\left(\frac{3}{10}\right)^{3-k} = \sum_{k=0}^{3}(C_3^k)^2\left(\frac{42}{100}\right)^k\left(\frac{12}{100}\right)^{3-k}$$

$$= \frac{1}{1000000}(12^3 + 9\cdot 42\cdot 12^2 + 9\cdot 42^2\cdot 12 + 42^3) = 0.32076.$$

8.10 Вероятность того, что изделие бракованное $p = 0.01$. Тогда $\mathbf{P}$(среди n хотя бы одно бракованное) $=$
$= 1 - \mathbf{P}$(среди n нет бракованных) $= 1 - (1 - 0.01)^n \geqslant 0.95$. Решая это неравенство, получим $n \geqslant \dfrac{\ln(0.05)}{\ln(0.99)} \approx 298.073$. Следовательно, наименьший объем должен состоять из 299 изделий.

8.11. Обозначим $A_n = \{$при n выстрелах стрелок попадает в цель хотя бы один раз$\}$. Тогда $\mathbf{P}(A_n) = 1 - (0.6)^n \geqslant 0.9$. Следовательно, $n\ln(0.6) \leqslant \ln(0.1)$, и $n \geqslant \dfrac{\log(0.1)}{\log(0.6)} = 4.5$. Необходимо произвести не менее пяти выстрелов.

8.12. $\mathbf{P}$(хотя бы одна шестерка при 6 бросаниях) $=$

$$= 1 - \left(\frac{5}{6}\right)^6 = 1 - \frac{15625}{46656} \approx 1 - 0.3349 = 0.6651.$$

$\mathbf{P}$(не менее двух шестерок при 12 бросаниях) $=$

$$= 1 - \left(\frac{5}{6}\right)^{12} - 12\cdot\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{11} = 1 - \frac{17\cdot 5^{11}}{6^{12}} \approx 0.6187.$$

$\mathbf{P}$(не менее трех шестерок при 18 бросаниях) $=$

$$= 1 - \left(\frac{5}{6}\right)^{18} - 18\cdot\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{17} - 153\cdot\frac{1}{6^2}\cdot\left(\frac{5}{6}\right)^{16} = 1 - \frac{5^{16}(25 + 90 + 153)}{6^{18}} \approx$$

$$\approx 1 - 0.4027 = 0.5973.$$

Следовательно, шансы убывают.

9.1. Поскольку вероятность не прорасти для семени обратно пропорциональна количеству посеянных семян, то применим формулу Пуассона с параметром $\lambda = 10000 \cdot 0.0003 = 3$. В результате получим

$$\mathbf{P}_{10000}(0) \approx \frac{3^0}{0!}e^{-3} = e^{-3}.$$

9.2. Применим формулу Пуассона с $\lambda = 200 \cdot 0.03 = 6$. Тогда

$$\mathbf{P}(\text{число прерываний} \leqslant 2) \approx \frac{6^0}{0!}e^{-6} + \frac{6}{1!}e^{-6} + \frac{6^2}{2!}e^{-6} = 25e^{-6}.$$

9.3. $P \approx 1 - e^{-8} - 8e^{-8} - \frac{64}{2}e^{-8} = 1 - 41e^{-8} \approx 0.986$.

9.4. $P \approx e^{-3}(3 + \frac{3^2}{2}) = e^{-3}\frac{15}{2} \approx 0.373$.

9.5. Вероятность выпадения шестерки $p = \frac{1}{6}$. Воспользуемся формулой 9.8. Имеем $q = 1-p = \frac{5}{6}$, $np = \frac{80}{6}$, $\sqrt{npq} = \frac{20}{6}$. Естественно выбирать границы, симметричные относительно среднего np, т. е. $m_1 = np - m$, $m_2 = np + m$, где m – неизвестно и его нужно определить. Согласно формуле 9.8

$$0.9973 = P_n(np-m, np+m) \approx \frac{1}{2}\left(\Phi\left(\frac{m}{\sqrt{npq}}\right) - \Phi\left(-\frac{m}{\sqrt{npq}}\right)\right) = \Phi\left(\frac{6m}{20}\right).$$

Из таблицы для функции Лапласа получаем $\frac{6m}{20} = 3$ или $m = 10$. Следовательно, $m_1 = \frac{80}{6} - 10 = 3\frac{2}{3}$, $m_2 = \frac{80}{6} + 10 = 23\frac{2}{3}$, и, значит, при 80 бросаниях шестерка выпадает с вероятностью, близкой к 0.9973 в пределах от 4 до 23 раз.

9.6. Воспользуемся формулой 9.8. Имеем $p = 3/4, q = 1/4$, $\sqrt{npq} = \sqrt{768 \cdot \frac{3}{4} \cdot \frac{1}{4}} = 12$, $np = 768 \cdot \frac{3}{4} = 576$. Тогда $a_n = \frac{564 - 576}{12} = -1, b_n = \frac{600 - 576}{12} = 2$, и искомая вероятность

$$P \approx \frac{1}{2}(\Phi(2) + \Phi(1)) \approx \frac{1}{2}(0.9545 + 0.6827) = 0.8186.$$

9.7. Поскольку число 75 сравнимо по порядку с числом 300, то применим формулу 9.7. Согласно этой формуле, имеем $P_{300}(75) \approx \frac{2}{15}\phi(0) \approx 0.053$.

9.8. $\mathbf{P}$(не более трех разбитых) $\approx e^{-1}(1+1+\frac{1}{2}+\frac{1}{6}) = = \frac{8}{3}e^{-1} \approx 0.981$.

9.9. Применим формулу 9.8. Имеем $n = 22500$, $A =$\{ответ неискренний\}, $p = \mathbf{P}(A) = 0.2$, $q = 1-p = 0.8$, $m_1 = 4380, m_2 = 4560$. Тогда $\sqrt{npq} = \sqrt{22500 \cdot 0.2 \cdot 0.8} = 60$, $a_n = \frac{0 - 22500 \cdot 0.2}{60} = -75$, $b_n = \frac{4620 - 22500 \cdot 0.2}{60} = 2$. Теперь вычислим искомую вероятность

$$P_{22500}(0, 4620) \approx \frac{1}{2}\big(\Phi(2) - \Phi(-75)\big) \approx \frac{1}{2}\big(\Phi(2) + 1\big) \approx 0.9773.$$

9.10. Применим интегральную теорему Муавра–Лапласа. В данной задаче $n = 400, p = q = \frac{1}{2}$, $\sqrt{npq} = \sqrt{400 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 10$. Тогда

$$\mathbf{P}(\text{число выпадений герба} > 180) \approx \frac{1}{2}\left(1 - \Phi\left(\frac{180-200}{10}\right)\right) =$$

$$= \frac{1}{2}(1 + \Phi(2)) \approx 0.9773.$$

9.11. Применим интегральную теорему Муавра–Лапласа. В данной задаче $n = 500, p = \frac{1}{6}, q = \frac{5}{6}$, $\sqrt{npq} = \sqrt{500 \cdot \frac{1}{6} \cdot \frac{5}{6}} = = \frac{50}{6}$. Тогда

$$\mathbf{P}(\text{шестерка выпадет} < 90) \approx \frac{1}{2}\left(\Phi\left(\frac{90 \cdot 6}{50} - 10\right) - \Phi(-10)\right) =$$

$$= \frac{1}{2}(\Phi(0.8) + 1) \approx 0.788145.$$

9.12. $$P \approx \frac{1}{2}\left(\Phi\left(\frac{14880 - 14400}{240}\right) - \Phi\left(\frac{14160 - 14400}{240}\right)\right) =$$

$$= \frac{1}{2}\big(\Phi(2) + \Phi(1)\big) \approx \frac{1}{2}(0.6827 + 0.9545) = 0.8186.$$

9.13. Применим локальную теорему Муавра–Лапласа. В данной задаче $n = 100$, $p = 0.2$, $q = 0.8$. Тогда

$$P \approx \frac{1}{\sqrt{100 \cdot 0.2 \cdot 0.8}}\phi\left(\frac{15-20}{4}\right) = \frac{1}{4}\phi(-1.25) = \frac{1}{4}\phi(1.25) \approx 0.04566.$$

9.14. Страховая компания собирает со всех домов суммарный взнос 300000 рублей. Поэтому, если за год сгорят более чем 3 дома, то компания понесет убытки. Применим теорему Пуассона с параметром $\lambda = 1000 \cdot 0.002 = 2$. Тогда

$$\mathbf{P}(\text{сгорят более чем 3 дома}) \approx$$

$$\approx 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} - \frac{2^2}{2!}e^{-2} - \frac{2^3}{3!}e^{-2} =$$

$$= 1 - \frac{19}{3}e^{-2} \approx 0.358.$$

10.1. Закон распределения имеет вид

X	0	1	2
P	$\frac{60}{87}$	$\frac{25}{87}$	$\frac{2}{87}$

10.2. Закон распределения имеет вид

X	0	1	2	3
P	$\frac{8}{1000}$	$\frac{96}{1000}$	$\frac{384}{1000}$	$\frac{512}{1000}$

10.3.

$$F(x) = \begin{cases} 0, & \text{при} \quad x < 3, \\ \int\limits_3^x \frac{7-y}{8}dy = 1 - \frac{(7-x)^2}{16}, & \text{при} \quad 3 \leqslant x \leqslant 7, \\ 1, & \text{при} \quad 7 < x. \end{cases}$$

$$\mathbf{P}(2 \leqslant X \leqslant 4) = F(4) - F(2) = 1 - \frac{3^2}{16} = \frac{7}{16}.$$

10.4.

$$F(x) = \begin{cases} 0 & \text{при} \quad x < 2, \\ (x-2)^2 & \text{при} \quad 2 \leqslant x \leqslant 3, \\ 1 & \text{при} \quad 3 < x. \end{cases}$$

$$\mathbf{P}(2.5 < X < 3.5) = F(3.5) - F(2.5) = 1 - (0.5)^2 = 0.75 = \frac{3}{4}.$$

10.5. *a*) $F(x) = \begin{cases} 1, & \text{при } x \leqslant 0, \\ \frac{1}{2}(1 - \cos x), & \text{при } 0 < x \leqslant \pi, \\ 0, & \text{при } \pi < x. \end{cases}$

b) $\mathbf{P}(0 \leqslant X \leqslant \frac{\pi}{4}) = F(\frac{\pi}{4}) - F(0) = \frac{1}{2}(1 - \cos\frac{\pi}{4}) \approx 0.146.$

10.6.

$$F(x) = \begin{cases} 0 & \text{при} \quad x < -3, \\ \frac{1}{2} + \frac{1}{\pi}\arcsin\frac{x}{3} & \text{при} \quad -3 \leqslant x \leqslant 3, \\ 1 & \text{при} \quad 3 < x. \end{cases}$$

$\mathbf{P}(1.5 < X < 4) = F(1.5) - F(4) = \frac{1}{3}.$

10.7. Закон распределения имеет вид

Y	0	0.5	1
P	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{1}{10}$

10.8. Величина $Y = e^{-X}$, поскольку X неотрицательна, принимает значения в интервале [0,1]. Для $0 \leqslant y \leqslant 1$

$$F_Y(y) = \mathbf{P}\left(e^{-X} < y\right) = \mathbf{P}(X > -\ln y) = \int\limits_{-\ln y}^{\infty} e^{-v}dv = y.$$

Следовательно, величина Y равномерно распределена на промежутке [0,1].

10.9. Поскольку $Y = X^2$, то $F_Y(y) = 0$ при $y \leqslant 0$, а при $y > 0$ имеем $F_Y(y) = \mathbf{P}(X^2 < y) = \mathbf{P}(-\sqrt{y} < X < \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$. Дифференцируя это равенство по y, получим $F'_Y(y) = \frac{1}{2\sqrt{y}}F'_X(\sqrt{y}) + \frac{1}{2\sqrt{y}}F'_X(-\sqrt{y})$. Отсюда в силу второго свойства плотности распределения (§ 10) следует, что при $y > 0$

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left(f_X(-\sqrt{y}) + f_X(\sqrt{y})\right),$$

и $f_Y(y) = 0$ при $y \leqslant 0$.

10.10. Закон распределения имеет вид

X	0	1	2	3	4
P	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$

10.11. $a = \frac{3}{8}$, $\mathbf{P}(X > 1) = \frac{7}{8}$.

10.12. $a = \frac{3}{64}$, $\mathbf{P}(3 \leqslant X < 4) = \frac{19}{64}$.

10.13. Применяя формулу (10.6) для $g(x) = \sqrt{x}$, получим

$$f_{\sqrt{X}}(y) = \begin{cases} \frac{3}{4}y^5, & 0 \leqslant y \leqslant \sqrt{2}, \\ 0, & y < 0, \quad \sqrt{2} < y. \end{cases}$$

10.14. Применяя формулу (10.6), получим

$$f_Y(y) = \begin{cases} \frac{3}{8}(y-1)^2, & 1 \leqslant y \leqslant 3, \\ 0, & y < 1, \quad 3 < y. \end{cases}$$

11.1. Совместный закон распределения имеет вид

(X, Y)	$(5,5)$	$(5,2)$	$(2,5)$	$(2,2)$
P	$\frac{28}{91}$	$\frac{24}{91}$	$\frac{24}{91}$	$\frac{15}{91}$

11.2. Совместный закон распределения имеет вид

(X, Y)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
P	$\frac{14}{33}$	$\frac{8}{33}$	$\frac{8}{33}$	$\frac{3}{33}$

11.3. Пусть $p_{i,j} = \mathbf{P}((X, Y) = (i, j))$, где $i = 0, 1, 2$, $j = 0, 1, 2$. Обозначим $\beta_1 = 1 - \alpha_1$ – вероятность промаха первым стрелком, $\beta_2 = 1 - \alpha_2$ – вторым. Тогда в силу независимости величин X и Y, и ввиду того, что выстрелы не зависят друг от друга, $p_{0,0} = \beta_1^2\beta_2^2$, $p_{0,1} = 2\beta_1^2\alpha_2\beta_2$, $p_{0,2} = \beta_1^2\alpha_2^2$, $p_{1,0} = 2\alpha_1\beta_1\beta_2^2$,

$p_{1,1} = 4\alpha_1\beta_1\alpha_2\beta_2$, $p_{1,2} = 2\alpha_1\beta_1\alpha_2^2$, $p_{2,0} = \alpha_1^2\beta_2^2$, $p_{2,1} = 2\alpha_1^2\alpha_2\beta_2$, $p_{2,2} = \alpha_1^2\alpha_2^2$,

11.4. Пара величин (X, Y) значение $(0, 0)$ никогда не принимает. Значение $(1, 0)$ она принимает, если на обеих костях нечетные числа, вероятность этого $1/4$. Значение $(1, 1)$ она принимает, если на обеих костях четные числа, вероятность этого $1/4$. И наконец значение $(0, 1)$ она принимает, если на одной кости четное число, а на другой нечетное. Вероятность такого события $1/2$.

11.5. Плотность определяется по свойству 2): $f(x, y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x, y) = 4\ln 2\ln 3xy2^{-x^2}3^{-y^2}$. Для вычисления искомой вероятности применим формулу (11.2). В результате получим

$$\mathbf{P}(1 \leqslant X \leqslant 2, 0 \leqslant Y \leqslant 1) = 2\ln 2\int_1^2 x2^{-x^2}dx\ 2\ln 3\int_1^2 y3^{-y^2}dy =$$

$$= \left(2^{-1} - 2^{-4}\right)\left(1 - 3^{-1}\right) = 7/24.$$

11.6. $f(x, y) = abe^{-ax-by}$.

11.7. Пусть $Z = \max\{X, Y\}$. Тогда $F_Z(x) = F_X(x)F_Y(x)$.

11.8. Пусть $Z = \min\{X, Y\}$. Тогда $F_Z(x) = 1 - (1 - F_X(x))(1 - F_Y(x))$.

11.9. Пользуясь определением совместной плотности распределения, вычислим совместную функцию распределения. При $0 \leqslant x \leqslant \pi/2$, $\quad 0 \leqslant y \leqslant \pi/2$

$$F(x, y) = \int_0^x\int_0^y \frac{1}{2}\sin(u + v)dudv = \frac{1}{2}(\sin x + \sin y - \sin(x + y)).$$

При $0 \leqslant x \leqslant \pi/2$, $\quad \pi/2 \leqslant y$

$$F(x, y) = \int_0^x\int_0^{\pi/2} \frac{1}{2}\sin(u + v)dudv = \frac{1}{2}(\sin x + 1 - \sin(x + \frac{\pi}{2})).$$

При $\pi/2 \leqslant x$, $\quad 0 \leqslant y \leqslant \pi/2$ аналогично

$$F(x, y) = \frac{1}{2}(\sin y + 1 - \sin(y + \frac{\pi}{2})).$$

При $x \leqslant 0$ или $y \leqslant 0$ имеем $F(x, y) = 0$.

12.1. $\mathbf{E}(X) = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$,

$\mathbf{D}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X) = \frac{1}{6}(1+4+9+16+25+36) - \frac{49}{4} = \frac{35}{12}$.

12.2. $\mathbf{E}(X) = 1.5, \mathbf{D}(X) = 0.15$.

12.3. $\mathbf{E}(X) = 4, \mathbf{D}(X) = \frac{3}{5}$.

12.4. $\mathbf{E}X = -\frac{1}{8}, \mathbf{D}X = \frac{119}{64}$.

12.5. Пусть X_1 – очки, выпавшие на жетоне, X_2 – на кубике. Тогда $X = X_1 + X_2$, и закон распределения имеет вид

X	1	2	3	4	5
$\mathbf{P}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

Используя свойства математического ожидания и дисперсии, получим $\mathbf{E}(X) = \mathbf{E}(X_1) + \mathbf{E}(X_2) = 1 + 2 = 3, \mathbf{D}(X) = \mathbf{D}(X_1) + \mathbf{D}(X_2) = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{5}{3}$.

12.6. Пусть X_1 – очки, выпавшие на одном жетоне, X_2 – на другом. Тогда $X = X_1 + X_2$ и $\mathbf{E}(X) = \mathbf{E}(X_1) + \mathbf{E}(X_2) = 4 + 3 = 7, \mathbf{D}(X) = \mathbf{D}(X_1) + \mathbf{D}(X_2) = 1 + 4 = 5$.

12.7. Закон распределения имеет вид

X	0	1	2	3
$\mathbf{P}$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

Математическое ожидание и дисперсия равны $\mathbf{E}(X) = 1$, $\mathbf{D}(X) = \frac{2}{3}$.

12.8.

$$\mathbf{E}X = \int_3^7 x\frac{7-x}{8}dx = \frac{7}{8} \cdot \frac{x^2}{2}\Big|_3^7 - \frac{x^3}{8 \cdot 3}\Big|_3^7 =$$

$$= \frac{1}{8}\Big(\frac{7}{2} \cdot 4 \cdot 10 - \frac{1}{3} \cdot 4(49+21+9)\Big) = 4\frac{1}{3}.$$

12.9.

$$\mathbf{E}X = \int\limits_1^2 x(x - \frac{1}{2})dx = (\frac{x^3}{3} - \frac{x^2}{4})\Big|_1^2 = \frac{8-1}{3} - \frac{4-1}{4} = \frac{19}{12},$$

$$\mathbf{D}X = (\frac{x^4}{4} - \frac{x^3}{6})\Big|_1^2 - \frac{19^2}{12^2} = \frac{15}{4} - \frac{7}{6} - \frac{19^2}{12^2} = \frac{62}{24} - \frac{19^2}{12^2} = \frac{11}{144}.$$

12.10. Используем решение задачи 11.2 и свойства вероятностей p_{ij} из § 11. Закон распределения величины X следующий: $\mathbf{P}(X = 0) = \frac{2}{3}, \mathbf{P}(X = 1) = \frac{1}{3}$. Закон распределения величины Y такой же. Тогда $\mathbf{E}(X) = \frac{1}{3}$, $\mathbf{D}(X) = \frac{2}{9}$, $\mathbf{E}(XY) = \frac{1}{11}$. В результате получим $r(X,Y) = \frac{\mathbf{E}(XY) - \mathbf{E}^2(X)}{\mathbf{D}(X)} = -\frac{1}{11}$.

12.11. Законы распределения величин X и Y имеют вид: $\mathbf{P}(X = 0) = \frac{1}{2}, \mathbf{P}(X = 1) = \frac{1}{2}$ и $\mathbf{P}(Y = 0) = \frac{1}{4}, \mathbf{P}(Y = 1) = \frac{3}{4}$ Тогда $\mathbf{E}(X) = \frac{1}{2}$, $\mathbf{D}(X) = \frac{1}{4}$, $\mathbf{E}(Y) = \frac{3}{4}$, $\mathbf{D}(Y) = \frac{3}{16}$, $\mathbf{E}(XY) = \frac{1}{4}$ В результате получим $r(X,Y) = \frac{\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)}{\sqrt{\mathbf{D}(X)\mathbf{D}(Y)}} = -\frac{1}{\sqrt{3}}$.

12.12. Используем решение задачи 11.1 и свойства вероятностей p_{ij} из § 11. Закон распределения величины X следующий: $\mathbf{P}(X = 2) = \frac{3}{7}, \mathbf{P}(X = 5) = \frac{4}{7}$. Закон распределения величины Y такой же. Тогда $\mathbf{E}(X) = \frac{26}{7}$, $\mathbf{D}(X) = \frac{108}{49}$, $\mathbf{E}(XY) = \frac{1240}{91}$. В результате получим $r(X,Y) = \frac{\mathbf{E}(XY) - \mathbf{E}^2(X)}{\mathbf{D}(X)} = -\frac{1}{13}$.

12.13. Математическое ожидание случайной величины X равно

$$\mathbf{E}(X) = \frac{1}{2}\int\limits_0^{\pi/2}\int\limits_0^{\pi/2} x\sin(x+y)dxdy$$

$$= \frac{1}{2}\int\limits_0^{\pi/2} x(\cos x - \cos(x + \pi/2))dx = \frac{\pi}{4} \approx 0.785$$

Дисперсия случайной величины X равна

$$\mathbf{D}(X) = \frac{1}{2}\int_0^{\pi/2}\int_0^{\pi/2} x^2\sin(x+y)dxdy - \frac{\pi^2}{16} =$$

$$= \frac{1}{2}\int_0^{\pi/2} x^2(\cos x - \cos(x+\pi/2))dx - \frac{\pi^2}{16} = \frac{\pi^2}{16} + \frac{\pi}{2} - 2 \approx 0.188$$

Из симметрии плотности вероятности относительно x и y следует, что $\mathbf{E}(Y) = \mathbf{E}(X)$, $\mathbf{D}(Y) = \mathbf{D}(X)$. Используя первое свойство ковариации, получим

$$\operatorname{cov}(X,Y) = \frac{1}{2}\int_0^{\pi/2}\int_0^{\pi/2} xy\sin(x+y)dxdy - \frac{\pi^2}{16} =$$

$$= \frac{1}{2}\int_0^{\pi/2} x(\cos x + (\frac{\pi}{2} - 1)\sin x)dx - \frac{\pi^2}{16} = \frac{\pi}{2} - \frac{\pi^2}{16} - 1.$$

Таким образом, коэффициент корреляции равен

$$r(X,Y) = \frac{\operatorname{cov}(X,Y)}{\mathbf{D}(X)} = \frac{\frac{\pi}{2} - \frac{\pi^2}{16} - 1}{\frac{\pi^2}{16} + \frac{\pi}{2} - 2} \approx -\frac{0.046}{0.188} \approx 0.245.$$

13.1. $\psi(z) = \frac{1}{6}(z + z^2 + z^3 + z^4 + z^5 + z^6)$.

13.2. 1) Поскольку $\frac{1}{2}\left(1 - \frac{1}{2}z\right)^{-1} = \frac{1}{2} + \frac{1}{4}z + \frac{1}{8}z^2 + \ldots$, то это соответствует дискретному распределению
$\mathbf{P}(X = k) = \frac{1}{2^{k+1}}$, $k = 0, 1, 2, \ldots$.

2) Поскольку $e^{3\lambda(z-1)} = e^{-3\lambda}\sum\limits_{k=0}^{\infty}\frac{(3\lambda)^k z^k}{k!}$, то
$\mathbf{P}(X = k) = \frac{(3\lambda)^k}{k!}e^{-3\lambda}$, $k = 0, 1, 2, \ldots$.

3) Поскольку $(\frac{1}{3} + \frac{2}{3}z)^n = \sum\limits_{k=0}^{n} C_n^k\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{n-k}z^k$, то $\mathbf{P}(X = k) =$
$= C_n^k\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{n-k}$, $k = 0, 1, \ldots, n$, и $\mathbf{P}(X = k) = 0$ при $k > n$.

3.3. 1) Имеем $\cos^2 t = \frac{1}{4}e^{it(-2)} + \frac{1}{2} + \frac{1}{4}e^{it2}$, следовательно,

$$F(x) = \begin{cases} 0, & x \leqslant -2; \\ \frac{1}{4}, & -2 < x \leqslant 0; \\ \frac{3}{4}, & 0 < x \leqslant 2; \\ 1, & 2 < x. \end{cases}$$

2) Имеем $\frac{1}{2}(\cos t + \cos 2t) = \frac{1}{4}e^{-it} + \frac{1}{4}e^{it} + \frac{1}{4}e^{-i2t} + \frac{1}{4}e^{i2t}$. Это – дискретное распределение со скачками $\frac{1}{4}$ в точках $\pm 1, \pm 2$.
3) Имеем $\frac{2}{3} + \frac{1}{3}\cos 3t = \frac{2}{3} + \frac{1}{6}e^{-i3t} + \frac{1}{6}e^{i3t}$, следовательно, $\mathbf{P}(X = 0) = \frac{2}{3}, \mathbf{P}(X = \pm 3) = \frac{1}{6}$.

13.4. Можно записать
$$\psi(z) = \sum_{k=0}^{\infty} z^k \mathbf{P}(X + 1 = k + 1) = \frac{1}{z}\sum_{l=1}^{\infty} z^l \mathbf{P}(X + 1 = l).$$
Используя это равенство и тот факт, что X – неотрицательная величина, получим ответ на первый вопрос:
$$\psi_{X+1}(z) = \sum_{l=0}^{\infty} z^l \mathbf{P}(X + 1 = l) = \mathbf{P}(X + 1 = 0) + z\psi(z) = z\psi(z).$$
Ответ на второй вопрос с учетом неотрицательности и целочисленности X следующий:
$$\psi_{2X}(z) = \sum_{l=0}^{\infty} z^l \mathbf{P}(2X = l) = \sum_{k=0}^{\infty} z^{2k} \mathbf{P}(2X = 2k) = \psi(z^2).$$

13.5. $\varphi(t) = \frac{2\alpha^2}{t^2}(1 - \cos\frac{t}{\alpha})$ при $t \neq 0$, $\varphi(t) = 1$ при $t = 0$.

13.6. 1) 0, $a^2/3$; 2) 0, 7/6; 3) 1, 7/6.

13.7. $\varphi(t) = e^{it}\left(\frac{2}{it} + \frac{2}{t^2}\right) - \frac{2}{t^2}$.

13.8. $\mathbf{E}(X^k) = \begin{cases} k!, & \text{при } k \text{ четном,} \\ 0, & \text{при } k \text{ нечетном.} \end{cases}$

13.9. По определению характеристической функции

$$\varphi(t) = \frac{1}{2}\int_{-\infty}^{\infty} e^{itx-|x|}dx = \frac{1}{2}\int_{0}^{\infty} e^{(it-1)x}dx + \frac{1}{2}\int_{-\infty}^{0} e^{(it+1)x}dx =$$
$$= \frac{1}{2}\left(\frac{1}{1-it} - \frac{1}{1+it}\right) = \frac{1}{1+t^2}.$$

13.10. $\varphi(t) = 1 + e^{-t^2/4h^2}\left(\frac{it\sqrt{\pi}}{2h} - \frac{t}{h}\int_0^{t/2h} e^{v^2}dv\right)$.

14.1. По формуле 14.3

$$\mathbf{P}(30 < X < 80) = \frac{1}{2}\Big(\Phi\Big(\frac{80-40}{\sqrt{400}}\Big) - \Big(\frac{30-40}{\sqrt{400}}\Big)\Big) =$$

$$= \frac{1}{2}\big(\Phi(2) + \Phi(\tfrac{1}{2})\big) \approx \frac{1}{2}(0.9545 + 0.3829) = 0.6687.$$

14.2. Положим $p = 1 - \frac{5^2}{6^2} = \frac{11}{36}$ – вероятность появления хотя бы одной шестерки. Шестерка впервые выпадет при k-ом бросании с вероятностью, задаваемой геометрическим распределением $P = \big(\frac{5^2}{6^2}\big)^{k-1}\frac{11}{36} = \frac{11\cdot 5^{2k-2}}{6^{2k}}, k = 1, 2, \ldots.$

14.3.

$$P(Y < y) = \begin{cases} 1, & 1 \leqslant y, \\ y^n, & 0 \leqslant y \leqslant 1, \\ 0, & y \leqslant 0, \end{cases}$$

$$P(Z < z) = \begin{cases} 1, & 1 \leqslant z, \\ 1-(1-z)^n, & 0 \leqslant z \leqslant 1, \\ 0, & z \leqslant 0. \end{cases}$$

14.4. Поскольку $\mathbf{E}(X) = \frac{a+b}{2} = 4, \mathbf{D}(X) = \frac{(b-a)^2}{12} = 3$, то $a = 1$ и $b = 7$. Следовательно, плотность имеет вид $f(x) = = \frac{1}{6}\mathbb{I}_{[1,7]}(x)$.

14.5. Обозначим Y длину левого куска проволоки, выбирая левый конец за начало координат. Согласно примеру 4 §10 величина Y распределена равномерно на $[0,1]$. По условию задачи $X = \max\{Y, 1-Y\}$. Тогда при $x \in [1/2, 1]$

$$\mathbf{P}(X < x) = \mathbf{P}(Y < x, 1-Y < x) = \mathbf{P}(1-x < Y < x) = 2x - 1.$$

Следовательно, величина X распределена равномерно на интервале $[1/2, 1]$.

14.6. Используя независимость случайных величин X_1 и

X_2, получим

$$\begin{aligned}\mathbf{P}(Z = m) &= \mathbf{P}(X_1 = m, X_2 < m) + \\ &+ \mathbf{P}(X_2 = m, X_1 < m) + \mathbf{P}(X_1 = m, X_2 = m) = \\ &= 2\mathbf{P}(X_1 = m)\mathbf{P}(X_2 < m) + \mathbf{P}^2(X_1 = m) = \\ &= 2q^m(1-q)\sum_{k=0}^{m-1} q^k(1-q) + q^{2m}(1-q)^2 = \\ &= q^m(1-q)(2 - q^m - q^{m+1}).\end{aligned}$$

14.7. Величина X имеет геометрическое распределение с вероятностями

$$\mathbf{P}(X = k) = 0.2 \cdot (0.8)^{k-1}.$$

14.8. Обозначим a неизвестное математическое ожидание и σ^2 неизвестную дисперсию величины X. По формуле 14.3

$$0.15 = \mathbf{P}(X < 1.06) = \frac{1}{2}\Big(\Phi\Big(\frac{1.06-a}{\sigma}\Big) - \Phi(-\infty)\Big) = \frac{1}{2}\Big(\Phi\Big(\frac{1.06-a}{\sigma}\Big) + 1\Big).$$

Следовательно, $\Phi\Big(\frac{1.06-a}{\sigma}\Big) = -0.7$ или $\Phi\Big(\frac{a-1.06}{\sigma}\Big) = 0.7$. Из таблицы для функции Лапласа находим $\frac{a-1.06}{\sigma} \approx 1.04$. Аналогично

$$0.1 = \mathbf{P}(X > 3.38) = \frac{1}{2}\Big(\Phi(\infty) - \Phi\Big(\frac{3.38-a}{\sigma}\Big)\Big) = \frac{1}{2}\Big(1 - \Phi\Big(\frac{3.38-a}{\sigma}\Big)\Big).$$

Следовательно, $\Phi\Big(\frac{3.38-a}{\sigma}\Big) = 0.8$ и $\frac{3.38-a}{\sigma} \approx 1.28$. Два полученных приближенных равенства можно рассматривать как систему двух уравнений с двумя неизвестными. Решая эту систему, получим $a \approx 2.1$, $\sigma \approx 1$.

15.1. $\mathbf{P}(\{1,5\} + \{2,4\} + \{3,3\} + \{4,2\} + \{5,1\}) = \frac{5}{36}$.

15.2. $\mathbf{P}(\{2,6\} + \{3,5\} + \{4,4\}) = \frac{1}{4}\cdot\frac{1}{6} + \frac{1}{4}\cdot\frac{1}{6} + \frac{1}{4}\cdot\frac{1}{6} = \frac{1}{8}$.

15.3. Производящие функции величин X и Y имеют вид $\psi_X(z) = (zp + 1 - p)^{n_1}$ и $\psi_Y(z) = (zp + 1 - p)^{n_2}$. По свойству 3) для производящих функций

$$\psi_{X+Y}(z) = \psi_X(z)\psi_Y(z) = (zp + 1 - p)^{n_1+n_2}.$$

Следовательно, $X+Y$ тоже распределена по биномиально закону $\mathbf{P}(X+Y=m)=C^m_{n_1+n_2}p^m(1-p)^{n_1+n_2-m}$, $m=0,1,2,\ldots,n_1+n_2$.

15.4. По формуле (15.4)

$$f_{X_1+X_2}(z)=\int\limits_0^z 4e^{-4(z-x)}\frac{1}{2}\mathbb{I}_{[2,4]}(x)dx=$$

$$=2\mathbb{I}_{[2,4]}(z)e^{-4z}\int\limits_2^z e^{4x}dx+2\mathbb{I}_{[4,\infty)}(z)e^{-4z}\int\limits_2^4 e^{4x}dx=$$

$$=\frac{1}{2}(1-e^{8-4z})\mathbb{I}_{[2,4]}(z)+\frac{1}{2}e^{8-4z}(e^2-1)\mathbb{I}_{[4,\infty)}(z).$$

15.5. По формуле (15.3)

$$f(z)=\begin{cases}0, & \text{при } z\leqslant -1,\\ \dfrac{1+z}{4}, & \text{при } -1<z\leqslant 1,\\ \dfrac{3-z}{4}, & \text{при } 1<z\leqslant 3,\\ 0, & \text{при } z>3.\end{cases}$$

15.6. По формуле (15.4)

$$f_{X+Y}(z)=\begin{cases}0, & z\leqslant 0,\\ \frac{1}{b}(1-e^{-\alpha z}), & 0<z<b,\\ \frac{1}{b}(e^{\alpha b}-1)e^{-\alpha z}, & b\leqslant z.\end{cases}$$

Величина $-Y$ распределена равномерно на интервале $[-b,0]$. Тогда полагая $X-Y=X+(-Y)$ и применяя формулу (15.3), получим

$$f_{X-Y}(z)=\begin{cases}0, & z\leqslant -b,\\ \frac{1}{b}(1-e^{-\alpha(z+b)}), & -b<z<0,\\ \frac{1}{b}(1-e^{-\alpha b})e^{-\alpha z}, & 0\leqslant z.\end{cases}$$

15.7. Поскольку величины X_1 и X_2 неотрицательны, то по формуле (15.4)

$$f_{X_1+X_2}(z)=\int_0^z \frac{x}{6}\mathbb{I}_{[2,4]}(x)\frac{1}{2}\mathbb{I}_{[1,3]}(z-x)dx=$$
$$=\int_0^z \frac{x}{12}\mathbb{I}_{[2,4]}(x)\mathbb{I}_{[z-3,z-1]}(x)dx$$

Произведение индикаторов $\mathbb{I}_{[2,4]}(x)\mathbb{I}_{[z-3,z-1]}(x)$ равно нулю, если интервал $[z-3,z-1]$ расположен левее или правее интервала $[2,4]$, т.е. при $z<3$ либо при $7<z$. Интервалы $[z-3,z-1]$ и $[2,4]$ имеют пересечение $[2,z-1]$ при $2\leqslant z-1\leqslant 4$ и имеют пересечение $[z-3,4]$ при $2\leqslant z-3\leqslant 4$. В результате получим, что

$$\mathbb{I}_{[2,4]}(x)\mathbb{I}_{[z-3,z-1]}(x)=\begin{cases}0, & z<3,\quad 7<z,\\ \mathbb{I}_{[2,z-1]}(x), & 3\leqslant z\leqslant 5,\\ \mathbb{I}_{[z-3,4]}(x), & 5\leqslant z\leqslant 7.\end{cases}$$

Поэтому

$$f_{X_1+X_2}(z)=\begin{cases}0, & z<3,\quad 7<z,\\ \frac{1}{12}\int\limits_2^{z-1} xdx=\frac{1}{24}(z-3)(z+1), & 3\leqslant z\leqslant 5,\\ \frac{1}{12}\int\limits_{z-3}^{4} xdx=\frac{1}{24}(7-z)(z+1), & 5\leqslant z\leqslant 7.\end{cases}$$

16.1. Пусть величина X_l равна 1, если l-й элемент отказал за время t, и $X_l=0$, если не отказал. Очевидно $\mathbf{E}(X_l)=0.05$. Применим лемму 1 (§ 16) для $X=\sum\limits_{l=1}^{60}X_l$. Тогда получим

$$\mathbf{P}\Big(\sum_{l=1}^{60}X_l\leqslant 4\Big)=1-\mathbf{P}\Big(\sum_{l=1}^{60}X_l>5\Big)\geqslant 1-\frac{1}{5}\mathbf{E}\Big(\sum_{l=1}^{60}X_l\Big)=\frac{2}{5}.$$

16.2. Пусть X_l – число очков на игральной кости при l-ом бросании, а $S_n = \sum\limits_{l=1}^{n} X_l$ – суммарное число очков при n бросаниях, $n = 100$. Согласно решению задачи 12.1 $\mathbf{E}(X_l) = \frac{7}{2}$, $\mathbf{D}(X_l) = \frac{35}{12}$. Тогда $n\mathbf{E}(X_l) = 350$, $n\mathbf{D}(X_l) = \frac{1750}{6}$ и по формуле (16.3)

$$\mathbf{P}\left(\left|S_{100} - 350\right| \geqslant 30\right) \leqslant \frac{1750}{900 \cdot 6} \approx 0.324.$$

16.3. Пусть величина X_l равна 1, если при l-ом бросании выпал герб, и 0, если выпала решетка. Тогда $\mathbf{E}(X_l) = \frac{1}{2}$, $\mathbf{D}(X_l) = \frac{1}{4}$. Применим (16.4) при $n = 1000$ и $\delta = 100$. Полагая ν_{1000} – частота появления герба, получим

$$\mathbf{P}\left(\left|\nu_{1000} - \frac{1}{2}\right| < 0.1\right) = \mathbf{P}\left(\left|\sum_{l=1}^{1000} X_l - 500\right| < 100\right) \geqslant$$

$$\geqslant 1 - \frac{1000}{10000} \cdot \frac{1}{4} = \frac{39}{40}.$$

16.4. Обозначим S_n – число наступлений A в n испытаниях с вероятностью успеха p. Тогда $\mathbf{E}(S_n) = np$, $\mathbf{D}(S_n) = npq$. Следовательно, при $p = 0.2$ математическое ожидание $\mathbf{E}(S_{1500}) = 1500 \cdot 0.2 = 300$ и дисперсия $\mathbf{D}(S_{1500}) = 1500 \cdot 0.2 \cdot 0.8 = 240$. Неравенство Чебышева дает

$$\mathbf{P}(|S_{1500} - 300| > 40) \leqslant \frac{240}{40^2} = 0.15.$$

16.5. Применим (16.2) для величины $Y = \frac{1}{n}S_n$, где S_n та же, что в решении задачи 16.4. Тогда $\mathbf{D}(Y) = \frac{pq}{n}$ и число n следует выбирать так, чтобы $1 - \frac{pq}{n\delta^2} \geqslant 0.95$. Решая это неравенство относительно n, получаем $n \geqslant \frac{pq}{0.05\delta^2}$. При $p = 0.9$, $q = 1 - 0.9 = 0.1$, $\delta = 0.01$ имеем $n \geqslant \frac{0.9 \cdot 0.1}{0.05(0.01)^2} = \frac{0.09}{0.000005} = 18000$, т. е. наименьшее число изделий, которые следует проверить, равно 18000.

17.1. Пусть X_l – число очков на игральной кости при l-ом бросании, а S_n – суммарное число очков при n бросаниях,

$n = 420$. Согласно решению задачи 12.1 $\mathbf{E}(X_l) = \frac{7}{2}$, $\mathbf{D}(X_l) = \frac{35}{12}$. Тогда $n\mathbf{E}(X_l) = 1470$, $\sqrt{n\mathbf{D}(X_l)} = 35$ и по формуле (17.9)

$$\mathbf{P}(1400 \leqslant S_{420} < 1505) \approx \frac{1}{2}\left(\Phi\left(\frac{1505-1470}{35}\right) - \Phi\left(\frac{1400-1470}{35}\right)\right) =$$

$$= \frac{1}{2}\big(\Phi(1) - \Phi(-2)\big) \approx \frac{1}{2}(0.6827 + 0.9545) = 0.8186.$$

17.2. Пусть X_l – число очков при l-ом выстреле, а S_n – суммарное число очков при n выстрелах, $n = 100$. Тогда $\mathbf{E}(X_l) = 10\cdot0.5+9\cdot0.3+8\cdot0.1+7\cdot0.05+6\cdot0.05 = 9.15$, $\mathbf{D}(X_l) = \mathbf{E}(X_l^2) - \mathbf{E}^2(X_l) = 50+24.3+6.4+2.45+1.8-83.7225 = 1.2275$, $\sqrt{\mathbf{D}(X_l)} \approx 1.1079$. По формуле (17.9)

$$\mathbf{P}(900 \leqslant S_{100} < \infty) \approx \frac{1}{2}\left(\Phi(\infty) - \Phi\left(\frac{900-915}{11.079}\right)\right) \approx$$

$$\approx \frac{1}{2}\big(1 + \Phi(1.35)\big) \approx \frac{1}{2}(1 + 0.823) = 0.9115.$$

17.3. Пусть S_n – число вызовов в течение года, $n = 365$. Тогда $\mathbf{E}(X_l) = \lambda = 73$, $\mathbf{D}(X_l) = \lambda = 73$, и $\sqrt{n}\sqrt{\mathbf{D}(X_l)} = \sqrt{365}\sqrt{73} = 73\sqrt{5}$, $l = 1, 2, \dots, 365$, и по формуле (17.9)

$$\mathbf{P}(26500 < S_n < 26800) \approx \frac{1}{2}\left(\Phi\left(\frac{155}{73\sqrt{5}}\right) + \Phi\left(\frac{145}{73\sqrt{5}}\right)\right) \approx$$

$$\approx \frac{1}{2}\big(\Phi(0.95) + \Phi(0.89)\big) \approx \frac{1}{2}(0.6579 + 0.6265) = 0.6422.$$

Математическая статистика

18.1. Значения $x_{(1)} \leqslant x_{(2)} \leqslant \dots \leqslant x_{(12)}$ – реализации вариационного ряда следующие: 215.3; 217.1; 218.7; 219.1; 219.5; 219.9; 222.3; 225.8; 227.4; 229.6; 229.9; 231.8. График эмпирической функции распределения имеет в каждой из этих точек скачок величины $\frac{1}{12}$. При $x \leqslant 215.3$ функция имеет нулевое значение, а при $x > 231.8$ – единичное. Так как 12 – четное число, то выборочная медиана равна $\frac{1}{2}(x_{(6)} + x_{(7)}) = \frac{1}{2}(219.9 + 222.3) = 221.1$.

18.2. Значения $x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(40)}$ – реализации вариационного ряда следующие: -1.61; -1.57; -1.40; -1.27; -1.09; -0.89; -0.88; -0.85; -0.62; -0.53; -0.52; -0.46; -0.44; -0.39; -0.34; -0.26; -0.14; -0.10; -0.02; 0.02; 0.12; 0.23; 0.40; 0.41; 0.42; 0.48; 0.53; 0.58; 0.59; 0.59; 0.65; 0.79; 0.83; 0.86; 0.97; 1.26; 1.27; 1.38; 1.45; 1.70. Поскольку нас интересует квантиль порядка $p = 0.8$, а $n = 40$, то $[np] = np = 32$, и $\hat{z}_{0.8} = \frac{1}{2}(x_{(32)} + x_{(33)}) = \frac{1}{2}(0.79 + 0.83) = 0.81$. Так как 40 – четное число, то выборочная медиана равна $\frac{1}{2}(x_{(20)} + x_{(21)}) = \frac{1}{2}(0.02 + 0.12) = 0.07$.

19.1. Для вычисления оценки математического ожидания начальной скорости снаряда воспользуемся формулой (19.1) при $c = 1235$: $\bar{x}_{16} = 1235 + \frac{1}{16}(16 - 8) = 1235.5$ м/с. Оценку дисперсии начальной скорости вычисляем по формуле (19.2): $s_{16}^2 = \frac{1}{15}(0.01 + 4 + 6.76 + 0.49 + 9 + 1.69 + 0.16 + 4.84 + 1 + 1.69 + 4.41 + 5.76 + 1.44 + 4 + 0.01 + 0.64) = 3.06$ м^2/с^2.

19.2. Значение выборочного среднего числа неправильных соединений равно $\bar{x}_{60} = 2$. Вычисляя значение выборочной дисперсии по формуле (19.4), получим $s_{60}^2 = \frac{1}{59} \cdot 366 - \frac{60}{59} \cdot 2^2 \approx 2.1356$.

19.3. $\bar{x}_{30} = \frac{1}{30} \cdot 489 = 16.3$, $s_{30}^2 = \frac{1}{29} \sum\limits_{l=1}^{30} (x_l - \bar{x}_{30})^2 = \frac{1}{29} \cdot 199.78 \approx 6.888966$.

19.4. $\bar{x}_{40} = \frac{1}{40} \cdot 2.15 = 0.05375$, $s_{40}^2 = \frac{1}{39} \sum\limits_{l=1}^{40} (x_l - \bar{x}_{40})^2 = \frac{1}{39} \cdot 29.32714 \approx 0.751978$.

19.5. Имеем $\bar{x}_{10} = \frac{499}{10} = 49.9$; $\bar{y}_{10} = \frac{174}{10} = 17.4$; $\sum\limits_{l=1}^{10} x_l y_l = 9228$; $\sum\limits_{l=1}^{10} x_l^2 = 27175$; $\sum\limits_{l=1}^{10} y_l^2 = 3182$. Тогда

$$r_{10}(\vec{x}, \vec{y}) = \frac{\sum\limits_{l=1}^{10} x_l y_l - 10 \cdot \bar{x}_{10} \bar{y}_{10}}{\sqrt{\left(\sum\limits_{l=1}^{10} x_l^2 - 10 \cdot \bar{x}_{10}^2\right)\left(\sum\limits_{l=1}^{10} y_l^2 - 10 \cdot \bar{y}_{10}^2\right)}} \approx 0.92.$$

Это значение характеризует сильную линейную зависимость.

21.1. Воспользуемся результатом решения задачи 19.4 и формулой (21.2). Из таблицы для функции Лапласа по $\alpha = 0.95$ определяем $z_\alpha \approx 1.96$. Тогда $z_\alpha \frac{\sqrt{s_{40}^2}}{\sqrt{40}} \approx 0.26874$, и в силу того, что $\bar{x}_{40} = 0.05375$, нижняя доверительная граница будет равной $0.05375 - 0.26874 = -0.21499$, а верхняя будет равной $0.05375 + 0.26874 = 0.32249$. Таким образом, интервал $(-0.21499, 0.32249)$ содержит математическое ожидание с вероятностью, приблизительно равной 0.95.

21.2. а) Воспользуемся доверительным интервалом для математического ожидания с доверительной вероятностью α при неизвестной дисперсии, определенном в п. 2 § 21. По результатам решения задачи 19.1 $\bar{x}_{16} = 1235.5, s_{16}^2 = 3.06$. Из таблицы для функции Лапласа по $\alpha = 0.9$ определяем $z_\alpha \approx 1.65$. Тогда $z_\alpha \frac{\sqrt{s_{16}^2}}{\sqrt{16}} \approx 0.72$. Таким образом, интервал (1234.78, 1236.22) содержит математическое ожидание с вероятностью, приблизительно равной 0.9.

б) Построим доверительный интервал для дисперсии. Воспользуемся формулой (21.4). Из таблицы для функции Лапласа по $\alpha = 0.92$ определяем $z_\alpha \approx 1.75$. Имеем $\bar{s}_{16}^{(4)} = \sum_{l=1}^{16}(x_l - \bar{x}_{16})^4 = \frac{1}{16} \cdot 244.2102 \approx 15.2631$. Следовательно,

$$\frac{z_\alpha}{\sqrt{16}}\sqrt{\bar{s}_{16}^{(4)} - (s_{16}^2)^2} \approx \frac{1.75}{4}\sqrt{15.2631 - 9.3636} \approx 1.06.$$

Поскольку $s_{16}^2 = 3.06$, то нижняя доверительная граница будет равной $3.06 - 1.06 = 2$, а верхняя будет равной $3.06 + 1.06 = 4.12$. Таким образом, интервал (2, 4.12) содержит значение дисперсии с вероятностью, приблизительно равной 0.92.

21.3. Воспользуемся формулой (21.3). Имеем $\bar{x}_5 = 1.594 \cdot 10^{-19}$, $\sqrt{s_5^2} \approx 0.002739 \cdot 10^{-19}$, $z_{0\,99,4} \approx 4.6$. Следовательно, $\frac{z_{0\,99,4} \cdot \sqrt{s_5^2}}{\sqrt{5}} \approx 0.005634 \cdot 10^{-19}$, и доверительный интервал имеет вид $(1.588 \cdot 10^{-19}, 1.600 \cdot 10^{-19})$. Заметим, что современное значение заряда электрона приблизительно равно $1.602 \cdot 10^{-19}$ Это означает, что приведенные в задаче данные измерений имели систематическую ошибку.

23.1. Пусть X – число мальчиков в семье с двумя детьми. Пусть ν_i – количество семей с i мальчиками ($i = 0, 1, 2$) среди семей с двумя детьми. Проверяется гипотеза о том, что величина X имеет биномиальное распределение с параметром $p = 0.515$, т. е. распределение вида

$$\mathbf{P}(X=0) = (1-p)^2,\ \mathbf{P}(X=1) = 2p(1-p),\ \mathbf{P}(X=2) = p^2.$$

Имеем $n = 17036$,

$$\Delta_n(\vec{x}) = \frac{(4019 - 17036\cdot(0.485)^2)^2}{17036\cdot(0.485)^2} + \frac{(8488 - 17036\cdot 2\cdot 0.515\cdot 0.485)^2}{17036\cdot 2\cdot 0.515\cdot 0.485} + \\ + \frac{(4529 - 17036\cdot(0.515)^2)^2}{17036\cdot(0.515)^2} \approx 0.118.$$

По таблице квантилей распределения Пирсона с двумя степенями свободы имеем $z_{0.95} \approx 5.991$. Следовательно, гипотеза принимается.

23.2. Сначала нужно упорядочить значения, т. е. построить реализацию вариационного ряда $x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(20)}$. Имеем $0.008 \leqslant 0.1 \leqslant 0.166 \leqslant 0.204 \leqslant 0.253 \leqslant 0.292 \leqslant 0.354 \leqslant 0.372 \leqslant 0.422 \leqslant 0.429 \leqslant 0.453 \leqslant 0.509 \leqslant 0.520 \leqslant 0.531 \leqslant 0.610 \leqslant 0.648 \leqslant 0.805 \leqslant 0.809 \leqslant 0.863 \leqslant 0.911$. Используя определение эмпирической функции распределения, несложно понять, что в данном примере

$$D_{20}(\vec{x}) = \sqrt{20} \max_{1\leqslant k\leqslant 20}\left\{\left|x_{(k)} - \frac{k}{20}\right|, \left|x_{(k)} - \frac{k-1}{20}\right|\right\} = \\ = \sqrt{20}\left|x_{(14)} - \frac{14}{20}\right| = \sqrt{20}(0.7 - 0.531) \approx 0.7758.$$

Из таблицы для распределения Колмогорова находим $z_{0.9} \approx 1.23$. Поскольку $D_{20}(\vec{x}) < z_{0.9}$, то гипотеза принимается.

23.3. Поскольку все вероятности p_i° равны 0.1, получаем $\Delta_n(\vec{x}) = \frac{(968-1000.2)^2}{1000.2} + \frac{(1026-1000.2)^2}{1000.2} + \cdots + \frac{(1014-1000.2)^2}{1000.2} \approx$ ≈ 9.3677. По таблице квантилей распределения Пирсона с девятью степенями свободы имеем $z_{0.8} \approx 12.24$. Следовательно, критерий Пирсона подтверждает гипотезу о равновероятности появления цифр в записи числа $\pi - 3$.

24.1. Воспользуемся формулами (24.8), (24.9). Для имеющихся данных

$$\sum_{i=1}^{5} y_i = 15,\quad \sum_{i=1}^{5} x_i = 78.7,\quad \sum_{i=1}^{5} y_i x_i = 246.1,\quad \sum_{i=1}^{5} y_i^2 = 55.$$

Тогда

$$\hat{\theta}_2 = \frac{5 \cdot 246.1 - 15 \cdot 78.7}{5 \cdot 55 - 15^2} = \frac{50}{50} = 1,$$

$$\hat{\theta}_1 = \frac{1}{5} \cdot 78.7 - 1 \cdot \frac{1}{5} \cdot 15 = 12.74.$$

В результате получаем, что зависимость величины урожайности X от количества видов удобрений y имеет вид

$$X = y + 12.74 + \mathcal{E}.$$

24.2. Формулы (24.8), (24.9) дают следующий ответ: $\hat{\theta}_1 \approx 67.5, \hat{\theta}_2 \approx 0.87$.

24.3. Воспользуемся формулами (24.11), (24.12), в которые вместо случайных величин подставим значения их реализаций. Для имеющихся данных

$$\sum_{i=1}^{14} y_i = 19844, \quad \sum_{i=1}^{14} x_i = 16051 = 7 \cdot 2293,$$

$$\sum_{i=1}^{14} y_i x_i = 22752428, \quad \sum_{i=1}^{14} y_i^2 = 28135942.$$

Тогда

$$\hat{\theta}_2 = \frac{14(22752428 - 9922 \cdot 2293)}{4(7 \cdot 14067971 - 98446084)} = \frac{4487}{29713} \approx 0.151,$$

$$\hat{\theta}_1 = \frac{16050}{14} - \frac{4487}{29713} \cdot \frac{19844}{14} = \frac{55411905}{59426} \approx 932.4522.$$

В результате получаем, что зависимость величины разрывного усилия всего каната X от суммарной величины разрывных усилий составляющих его нитей Y имеет вид

$$X \approx 0.151\, Y + 932.4522 + \mathcal{E}.$$

24.4. $\hat{\theta}_1 = 4, \hat{\theta}_2 = -2, \hat{\theta}_3 = 0.25$.

25.1. Имеем $L(x, \theta) = \theta^n \exp\left(-\theta \sum\limits_{i=1}^{n} X_i\right)$, и $\dfrac{\partial \ln L(x,\theta)}{\partial \theta} = \dfrac{n}{\theta} - \sum\limits_{i=1}^{n} X_i$. Следовательно, оценка максимального правдоподобия имеет вид $\hat{\theta}_n = 1/\overline{X}_n$.

25.2. Пусть $\vec{\theta} = (\theta_1, \theta_2)$, где θ_1 – математическое ожидание, θ_2 – дисперсия распределения диаметров. Воспользуемся замечанием 1. Тогда

$$L(\vec{X}, \vec{\theta}) = \frac{1}{(2\pi\theta_2)^{n/2}} \exp\left(-\frac{1}{2\theta_2}\sum_{i=1}^{n}(X_i - \theta_1)^2\right).$$

Решая систему

$$\frac{\partial \ln L(\vec{X}, \vec{\theta})}{\partial \theta_1} = -\frac{1}{\theta_2}\sum_{i=1}^{n}(X_i - \theta_1) = 0,$$

$$\frac{\partial \ln L(\vec{X}, \vec{\theta})}{\partial \theta_2} = -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2}\sum_{i=1}^{n}(X_i - \theta_1)^2 = 0,$$

находим оценку максимального правдоподобия $\vec{\hat{\theta}}_n = (\overline{X}_n, \overline{S}_n^{(2)})$, где $\overline{S}_n^{(2)}$ определен в п. 3 § 19. Для имеющихся данных $\bar{x}_{20} = 4.998, \bar{s}_{20}^{(2)} = 0.009056$.

25.3. См. пример 4 § 25.

Приближенные значения функции Колмогорова
$$\mathcal{K}(x) = \sum_{-\infty}^{\infty} (-1)^l \exp(-2l^2 x^2),$$ домноженные на 10^5

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.2	00000	00000	00000	00000	00000	00000	00000	00000	00000	00000
0.3	00001	00002	00005	00009	00017	00030	00051	00083	00129	00193
0.4	00281	00397	00548	00738	00973	01259	01600	02002	02468	03002
0.5	03605	04281	05031	05853	06750	07718	08758	09866	11040	12276
0.6	13573	14923	16323	17775	19268	20799	22364	23958	25578	27219
0.7	28877	30547	32227	33911	35598	37283	38964	40637	42300	43951
0.8	45586	47204	48803	50381	51937	53468	54974	56455	57907	59332
0.9	60727	62093	63429	64734	66008	67252	68464	69645	70794	71913
1.0	73000	74057	75083	76078	77043	77979	78886	79764	80613	81434
1.1	82228	82995	83736	84450	85139	85804	86444	87061	87655	88226
1.2	88775	89303	89810	90297	90765	91213	91643	92056	92451	92829
1.3	93191	93537	93868	94185	94487	94776	95051	95314	95565	95804
1.4	96032	96249	96455	96652	96838	97016	97185	97345	97479	97641
1.5	97778	97908	98031	98148	98258	98362	98461	98554	98643	98726
1.6	98805	98879	98949	99015	99078	99136	99192	99244	99293	99339
1.7	99383	99423	99461	99497	99531	99563	99592	99620	99646	99670
1.8	99693	99715	99735	99753	99771	99787	99802	99815	99830	99842
1.9	99854	99864	99874	99884	99892	99900	99908	99915	99921	99927
2.0	99933	99938	99943	99947	99952	99955	99959	99962	99965	99968
2.1	99971	99972	99975	99977	99979	99981	99982	99984	99985	99986
2.2	99987	99989	99990	99990	99991	99992	99993	99993	99994	99994
2.3	99995	99995	99996	99996	99997	99997	99997	99997	99998	99998
2.4	99998	99998	99998	99999	99999	99999	99999	99999	99999	99999
2.5	99999	99999	99999	99999	99999	10^5	10^5	10^5	10^5	10^5

Приближенные значения стандартного нормального распределения $\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt$, домноженные на 10^5

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0.1	53983	54380	54776	55172	55567	55962	56356	56750	57142	57535
0.2	57926	58317	58706	59095	59484	59871	60257	60642	61026	61409
0.3	61791	62172	62552	62930	63307	63683	64058	64431	64803	65173
0.4	65542	65910	66276	66640	67003	67365	67724	68082	68439	68793
0.5	69146	69497	69847	70194	70540	70884	71226	71566	71904	72241
0.6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0.7	75804	76115	76424	76731	77035	77337	77637	77935	78231	78524
0.8	78815	79103	79389	79673	79955	80234	80511	80785	81057	81327
0.9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1.0	84135	84375	84614	84850	85083	85314	85543	85769	85993	86214
1.1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1.2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90148
1.3	90320	90490	90658	90824	90988	91149	91309	91466	91621	91774
1.4	91924	92073	92220	92364	92507	92647	92786	92922	93056	93189
1.5	93319	93448	93575	93699	93822	93943	94062	94179	94295	94408
1.6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1.7	95544	95637	95728	95819	95907	95994	96080	96164	96246	96327
1.8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1.9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97671
2.0	97725	97778	97831	97882	97933	97982	98030	98077	98124	98169
2.1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2.2	98610	98645	98679	98713	98746	98778	98809	98840	98870	98899
2.3	98928	98956	98983	99010	99036	99061	99086	99111	99134	99158
2.4	99180	99202	99224	99245	99266	99286	99305	99324	99343	99361
2.5	99379	99396	99413	99430	99446	99461	99477	99492	99506	99520
2.6	99534	99547	99560	99573	99586	99598	99609	99621	99632	99643
2.7	99653	99664	99674	99683	99693	99702	99711	99720	99728	99737
2.8	99745	99752	99760	99767	99774	99781	99788	99795	99801	99807
2.9	99813	99819	99825	99831	99836	99841	99846	99851	99856	99861
3.0	99865	99869	99874	99878	99882	99886	99889	99893	99897	99900
3.1	99903	99907	99910	99913	99916	99918	99921	99924	99926	99929
3.2	99931	99934	99936	99938	99940	99942	99944	99946	99948	99950
3.3	99952	99953	99955	99957	99958	99960	99961	99962	99964	99965
3.4	99966	99968	99969	99970	99971	99972	99973	99974	99975	99976
3.5	99977	99978	99978	99979	99980	99981	99982	99982	99983	99984
3.6	99984	99985	99985	99986	99986	99987	99987	99988	99988	99989
3.7	99989	99990	99990	99990	99991	99991	99992	99992	99992	99993
3.8	99993	99993	99993	99994	99994	99994	99994	99995	99995	99995
3.9	99995	99995	99996	99996	99996	99996	99996	99996	99997	99997

Приближенные значения функции Лапласа

$$\Phi(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int\limits_0^x e^{-t^2/2} dt, \text{ домноженные на } 10^5$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	00000	00798	01596	02393	03191	03988	04784	05581	06376	07171
0.1	07966	08759	09552	10348	11134	11924	12712	13499	14285	15069
0.2	15852	16633	17413	18191	18967	19741	20514	21284	22052	22818
0.3	23582	24344	25103	25860	26614	27366	28115	28862	29605	30346
0.4	31084	31819	32552	33280	34006	34729	35448	36164	36877	37587
0.5	38292	38995	39694	40387	41080	41768	42452	43132	43809	44481
0.6	45149	45814	46474	47131	47783	48431	49075	49714	50350	50981
0.7	51607	52230	52848	53461	54070	54675	55275	55870	56461	57047
0.8	57629	58206	58778	59346	59909	60468	61021	61570	62114	62653
0.9	63188	63718	64243	64763	65278	65789	66294	66795	67291	67783
1.0	68269	68750	69227	69699	70166	70628	71086	71538	71986	72429
1.1	72867	73300	73729	74152	74571	74986	75395	75800	76200	76595
1.2	76986	77372	77754	78130	78502	78870	79233	79592	79945	80295
1.3	80640	80980	81316	81648	81975	82298	82617	82931	83241	85547
1.4	83849	84146	84439	84728	85013	85294	85571	85844	86113	86378
1.5	86639	86696	87149	87398	87644	87886	88124	88358	88589	88817
1.6	89040	89260	89477	89690	89899	90106	90309	90508	90704	90897
1.7	91087	91273	91457	91637	91814	91988	92159	92327	92492	92655
1.8	92814	92970	93124	93275	93423	93569	93711	93852	93989	94124
1.9	94257	94387	94514	94639	94762	94882	95000	95116	95230	95341
2.0	95450	95557	95662	95764	95865	95964	96060	96155	96247	96338
2.1	96427	96514	96599	96683	96765	96844	96923	96999	97074	97148
2.2	97219	97289	97358	97425	97491	97555	97618	97679	97739	97798
2.3	97855	97911	97966	98019	98072	98123	98172	98221	98269	98315
2.4	98360	98405	98448	98490	98531	98571	98611	98649	98686	98723
2.5	98758	98793	98826	98859	98891	98923	98953	98983	99012	99040
2.6	99068	99095	99121	99146	99171	99195	99219	99241	99263	99285
2.7	99307	99327	99347	99367	99386	99404	99422	99439	99456	99473
2.8	99489	99505	99520	99535	99549	99563	99576	99590	99602	99615
2.9	99627	99639	99650	99661	99672	99682	99692	99702	99712	99721
3.0	99730	99739	99747	99755	99763	99771	99779	99786	99793	99800
3.1	99806	99813	99819	99825	99831	99837	99842	99848	99853	99858
3.2	99863	99867	99872	99876	99880	99885	99889	99892	99896	99900
3.3	99903	99907	99910	99913	99916	99919	99922	99925	99928	99930
3.4	99933	99935	99937	99940	99942	99944	99946	99948	99950	99952
3.5	99953	99955	99957	99958	99960	99961	99963	99964	99966	99967
3.6	99968	99969	99971	99972	99973	99974	99975	99976	99977	99978
3.7	99978	99979	99980	99981	99982	99982	99983	99984	99984	99985
3.8	99986	99986	99987	99987	99988	99988	99989	99989	99990	99990
3.9	99990	99991	99991	99992	99992	99992	99992	99993	99993	99993
4.0	99994	99994	99994	99994	99995	99995	99995	99995	99996	99996

Приближенные значения функции Гаусса

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx,$$ домноженные на 10^5

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0 07	0.08	0.09
0.0	39894	39892	39886	39876	39862	39844	39823	39797	39767	39733
0.1	39695	39654	39608	39559	39505	39448	39387	39322	39253	39181
0.2	39104	39024	38940	38853	38762	38667	38568	38466	38361	38252
0.3	38139	38023	37903	37780	37654	37524	37391	37255	37115	36973
0.4	36827	36678	36526	36371	36214	36053	35889	35723	35553	35381
0.5	35207	35029	34849	34667	34482	34294	34105	33912	33718	33521
0.6	33323	33122	32918	32713	32506	32297	32086	31874	31659	31443
0.7	31225	31006	30785	30563	30339	30114	29887	29660	29431	29200
0.8	28969	28737	28504	28269	28034	27799	27562	27324	27086	26848
0.9	26609	26369	26129	25888	25647	25406	25164	24923	24681	24439
1.0	24197	23955	23713	23471	23230	22988	22747	22506	22265	22025
1.1	21785	21546	21307	21069	20831	20594	20357	20121	19886	19652
1.2	19419	19186	18954	18723	18494	18265	18037	17810	17585	17360
1.3	17137	16915	16694	16474	16256	16038	15823	15608	15395	15183
1.4	14973	14764	14556	14351	14146	13943	13742	13542	13344	13147
1.5	12952	12758	12567	12376	12188	12001	11816	11632	11451	11270
1.6	11092	10916	10741	10568	10396	10227	10059	09893	09728	09566
1.7	09405	09246	09089	08933	08780	08628	08478	08329	08183	08038
1.8	07895	07754	07614	07477	07341	07206	07074	06943	06814	06687
1.9	06562	06438	06316	06195	06077	05959	05844	05730	05618	05508
2.0	05399	05292	05186	05082	04980	04879	04780	04682	04586	04491
2.1	04398	04307	04217	04128	04041	03955	03871	03788	03706	03626
2.2	03547	03470	03394	03319	03246	03174	03103	03034	02966	02898
2.3	02833	02768	02705	02643	02582	02522	02463	02406	02349	02294
2.4	02239	02186	02134	02083	02033	01984	01936	01888	01842	01797
2.5	01753	01709	01667	01625	01585	01545	01506	01468	01431	01394
2.6	01358	01323	01289	01256	01223	01191	01160	01130	01100	01071
2.7	01042	01014	00987	00961	00935	00909	00885	00861	00837	00814
2.8	00792	00770	00748	00727	00707	00687	00668	00649	00631	00613
2.9	00595	00578	00562	00545	00530	00514	00499	00485	00471	00457
3.0	00443	00430	00417	00405	00393	00381	00370	00358	00348	00337
3.1	00327	00317	00307	00298	00288	00279	00271	00262	00254	00246
3.2	00238	00231	00224	00217	00210	00203	00196	00190	00184	00178
3.3	00172	00167	00161	00156	00151	00146	00141	00136	00132	00128
3.4	00123	00119	00115	00111	00108	00104	00100	00097	00094	00090
3.5	00087	00084	00081	00079	00076	00073	00071	00068	00066	00063
3.6	00061	00059	00057	00055	00053	00051	00049	00047	00046	00044
3.7	00043	00041	00039	00038	00037	00035	00034	00033	00032	00030
3.8	00029	00028	00027	00026	00025	00024	00023	00022	00022	00021
3.9	00020	00019	00018	00018	00017	00016	00016	00015	00015	00014
4.0	00013	00013	00012	00012	00011	00011	00011	00010	00010	00009

Приближенные значения квантилей $z_{\alpha,k}$ порядка $(1+\alpha)/2$ для распределения Стьюдента с k степенями свободы:

$$S_k(z) = \int\limits_{-\infty}^{z} \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\,\Gamma(k/2)}\left(1+\frac{x^2}{k}\right)^{-(k+1)/2} dx$$

$k \setminus \alpha$	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.98	0.99	0.999
1	0.727	1.000	1.376	1.963	3.080	6.310	12.71	31.80	63.70	63.70
2	0.617	0.816	1.061	1.336	1.886	2.920	4.300	6.960	9.920	31.60
3	0.584	0.765	0.978	1.250	1.638	2.350	3.180	4.540	5.840	12.94
4	0.569	0.741	0.941	1.190	1.533	2.130	2.770	3.750	4.600	8.610
5	0.559	0.727	0.920	1.156	1.476	2.020	2.570	3.360	4.030	6.860
6	0.553	0.718	0.906	1.134	1.440	1.943	2.450	3.140	4.710	5.960
7	0.549	0.711	0.896	1.119	1.415	1.895	2.360	3.000	3.500	5.400
8	0.546	0.706	0.889	1.108	1.397	1.860	2.310	2.900	3.360	5.040
9	0.543	0.703	0.883	1.100	1.383	1.833	2.260	2.820	3.250	4.780
10	0.542	0.700	0.879	1.093	1.372	1.812	2.230	2.760	3.170	4.590
11	0.540	0.697	0.876	1.088	1.363	1.796	2.200	2.720	3.110	4.490
12	0.539	0.695	0.873	1.083	1.356	1.782	2.180	2.680	3.060	4.320
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.010	4.220
14	0.537	0.692	0.868	1.076	1.345	1.761	2.140	2.620	2.980	4.140
15	0.536	0.691	0.866	1.074	1.341	1.753	2.130	2.600	2.950	4.070
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.580	2.920	4.020
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.570	2.900	3.960
18	0.534	0.688	0.862	1.067	1.330	1.734	2.100	2.550	2.880	3.920
19	0.533	0.688	0.861	1.066	1.328	1.729	2.090	2.540	2.860	3.880
20	0.533	0.687	0.860	1.064	1.325	1.725	2.090	2.530	2.840	3.850
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.520	2.830	3.820
22	0.532	0.686	0.858	1.061	1.321	1.717	2.070	2.510	2.820	3.790
23	0.532	0.685	0.858	1.060	1.319	1.714	2.070	2.500	2.810	3.770
24	0.531	0.685	0.857	1.059	1.318	1.711	2.060	2.490	2.800	3.740
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.480	2.790	3.720
26	0.531	0.684	0.856	1.058	1.315	1.706	2.060	2.480	2.780	3.710
27	0.531	0.684	0.855	1.057	1.314	1.703	2.050	2.470	2.770	3.690
28	0.530	0.683	0.855	1.056	1.313	1.701	2.050	2.470	2.760	3.670
29	0.530	0.683	0.854	1.055	1.311	1.699	2.040	2.460	2.760	3.660
30	0.530	0.683	0.854	1.055	1.310	1.697	2.040	2.460	2.750	3.650
40	0.529	0.681	0.851	1.050	1.303	1.684	2.020	2.420	2.700	3.550
60	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.360	2.620	3.370
∞	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.330	2.580	3.290

Приближенные значения квантилей $z_{1-\alpha}$ порядка $1-\alpha$ для распределения Пирсона с k степенями свободы:

$$\mathcal{P}_k(z) = \int\limits_0^z \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k-2)/2} e^{-x/2} dx$$

k \ α	0.5	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001
1	0.455	0.708	1.074	1.642	2.706	3.841	5.024	6.635	7.879	10.83
2	1.386	1.833	2.408	3.219	4.605	5.991	7.378	9.210	10.60	13.82
3	2.366	2.946	3.665	4.642	6.251	7.815	9.348	11.35	12.84	16.27
4	3.357	4.045	4.878	5.989	7.779	9.488	11.14	13.28	14.86	18.47
5	4.351	5.132	6.064	7.289	9.236	11.07	12.83	15.09	16.75	20.52
6	5.348	6.211	7.231	8.558	10.65	12.59	14.45	16.81	18.55	22.46
7	6.346	7.283	8.383	9.803	12.02	14.07	16.01	18.48	20.29	24.32
8	7.344	8.351	9.524	11.03	13.36	15.51	17.54	20.09	21.96	26.13
9	8.343	9.414	10.66	12.24	14.68	16.92	19.02	21.67	23.59	27.88
10	9.342	10.47	11.78	13.44	15.99	18.31	20.48	23.21	25.19	29.59
11	10.34	11.53	12.90	14.63	17.28	19.68	21.92	24.73	26.76	31.26
12	11.34	12.58	14.01	15.81	18.55	21.03	23.34	26.22	28.30	32.91
13	12.34	13.64	15.12	16.99	19.81	22.36	24.74	27.69	29.82	34.53
14	13.34	14.69	16.22	18.15	21.06	23.69	26.12	29.14	31.32	36.12
15	14.34	15.73	17.32	19.31	22.31	25.00	27.49	30.58	32.80	37.70
16	15.34	16.78	18.42	20.47	23.54	26.30	28.85	32.00	34.27	39.25
17	16.34	17.82	19.51	21.62	24.77	27.59	30.19	33.41	35.72	40.79
18	17.34	18.87	20.60	22.76	25.99	28.87	31.53	34.81	37.16	42.31
19	18.34	19.91	21.69	23.90	27.20	30.14	32.85	36.19	38.58	43.82
20	19.34	20.95	22.78	25.04	28.41	31.41	34.17	37.57	40.00	45.32
21	20.34	21.99	23.86	26.17	29.62	32.67	35.48	38.93	41.40	46.80
22	21.34	23.03	24.94	27.30	30.81	33.92	36.78	40.29	42.80	48.27
23	22.34	24.07	26.02	28.43	32.01	35.17	38.08	41.64	44.18	49.73
24	23.34	25.11	27.10	29.55	33.20	36.42	39.36	42.98	45.56	51.18
25	24.34	26.14	28.17	30.68	34.38	37.65	40.65	44.31	46.93	52.62
26	25.34	27.18	29.25	31.80	35.56	38.89	41.92	45.64	48.29	54.05
27	26.34	28.21	30.32	32.91	36.74	40.11	43.19	46.96	49.65	55.48
28	27.34	29.25	31.39	34.03	37.92	41.34	44.46	48.28	50.99	56.89
29	28.34	30.28	32.46	35.14	39.09	42.56	45.72	49.59	52.34	58.30
30	29.34	31.32	33.53	36.25	40.26	43.77	46.98	50.89	53.67	59.70
35	34.34	36.48	38.86	41.78	46.06	49.80	53.20	57.34	60.28	66.62
40	39.34	41.62	44.17	47.27	51.81	55.76	59.34	63.69	66.77	73.40
50	49.34	51.89	54.72	58.16	63.17	67.51	71.42	76.15	79.49	86.66
60	59.34	62.14	65.23	68.97	74.40	79.08	83.30	88.38	91.95	99.61
70	69.33	72.36	75.69	79.72	85.53	90.53	95.02	100.4	104.2	112.3
80	79.33	82.57	86.12	90.41	96.58	101.9	106.6	112.3	116.3	124.8
90	89.33	92.76	96.52	101.1	107.6	113.1	118.1	124.1	128.3	137.2
100	99.33	102.9	106.9	111.7	118.5	124.3	129.6	135.8	140.2	149.4

Литература

1. *Агапов Г. И.* Задачник по теории вероятностей. М., 1986.

2. *Большев Л. Н., Смирнов Н.В.* Таблицы математической статистики. М., 1983.

3. *Боровков А. А.* Теория вероятностей. М., 1984.

4. *Ван дер Варден Б. Л.* Математическая статистика. М., 1960.

5. *Виленкин Н. Я.* Комбинаторика. М., 1969.

6. *Вентцель Е. С.* Теория вероятностей. М., 1962.

7. *Гихман И. И., Скороход А. В., Ядренко М. И.* Теория вероятностей и математическая статистика. Киев, 1979.

8. *Гнеденко Б. В.* Курс теории вероятностей. М., 1961.

9. *Гнеденко Б. В., Хинчин А. Я.* Элементарное введение в теорию вероятностей, М., 1982.

10. *Гордин М. И., Фролов А. Н.* Теория вероятностей в задачах и решениях. Математическая статистика в задачах и решениях (учебно-методические пособия). СПбГУ, 1997.

11. *Гусак А. А., Бричикова Е. А.* Теория вероятностей. Справочное пособие к решению задач. Минск, 2000.

12. *Ивченко Г. И., Медведев Ю. И.* Математическая статистика. 2-е изд. М., 1992.

13. *Крамер Г.* Математические методы статистики. 2-е изд. М., 1975.

14. *Колмогоров А. Н.* Основные понятия теории вероятностей. М., 1974.

15. *Мешалкин Л. Д.* Сборник задач по теории вероятностей. М., 1963.

16. *Нейман Ю. В.* Вводный курс теории вероятностей и математической статистики. М., 1968.

17. *Прохоров Ю. В., Розанов Ю. А.* Теория вероятностей. М., 1973.

18. *Свешников А. А.* Сборник задач по теории вероятностей, математической статистике и теории случайных функций. СПб., 2008.

19. *Севастьянов Б. А.* Курс теории вероятностей и математической статистики. М.,1982.

20. *Феллер В.* Введение в теорию вероятностей и ее применения. Т.1, 2. М., 1967.

Предметный указатель

Андрей Николаевич БОРОДИН

ЭЛЕМЕНТАРНЫЙ КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Учебное пособие

Издание восьмое,
стереотипное

Художественный редактор *С. Л. Шапиро*
Подготовка оригинал-макета *А. Н. Бородин*
Выпускающий *С. В. Николаева*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812) 412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 22.07.11.
Бумага офсетная. Гарнитура Школьная. Формат 84×108 1/32.
Печать офсетная. Усл. п. л. 13,44. Тираж 1500 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Издательско-полиграфическое предприятие «Правда Севера».
163002, г. Архангельск, пр. Новгородский, д. 32.
Тел./факс (8182) 64-14-54; www.ippps.ru