В большинстве стран мира элементы теории вероятностей и статистики изучаются в средней школе и считаются даже более важным математическим предметом (из-за многочисленных применений в физике, биологии, экономике,...), чем многие привычные российским школьникам разделы. В нашей стране, к сожалению, вероятность пока изучают только в очень немногих школах. Цель этой статьи — дать первое представление о понятиях, задачах и результатах теории вероятностей, с надеждой, что отдельные фрагменты этой теории будут все чаще появляться на страницах «Кванта».

Посчитаем вероятности

Н. ВАСИЛЬЕВ, А. СПИВАК

Что такое вероятность? Дать точное определение, которое годилось бы во всех многочисленных применениях, непросто.

Спросите студента-математика, и он отчеканит:

— Вероятностным пространством называется σ-алгебра множеств, на которой задана σ-аддитивная неотрицательная функция (мера), значение которой на всем пространстве равно 1. Вероятность события (множества) — его мера.

— А что такое случайная величина?

— Это измеримая функция.

— А что такое среднее значение случайной величины?

— Это ее интеграл Лебега.

И студент прав. Именно такие определения даны в книге А.Н.Колмогорова «Основные понятия теории вероятностей», вышедшей в свет в 1933 году. Но для первого знакомства этот подход явно не годится: не знает большинство наших читателей, что такое интеграл Лебега!

А как ответит на этот вопрос нематематик? Скорее всего, он скажет, что вероятность события — это его частота. Точнее, если вообразить, что в похожих, многократно воспроизводимых условиях событие иногда происходит, а иногда нет, то вероятность — это доля случаев, в которых оно происходит.

Это объяснение понятнее, но его трудно превратить в точное математическое определение. (Например, если мы 100 раз подбрасываем монету, то естественно ожидать, что примерно в 50 случаях выпадет герб. Но герб не обязательно выпадет именно в 50 случаях. Может быть и 51, и 45, и даже 100 гербов. Вот только вероятность последнего чудовищно мала — скоро мы научимся считать такие вероятности и узнаем, что она равна $1/2^{100} < 1/10^{30}$.)

На самом деле, познакомиться с основными приемами подсчета вероятности может даже шестиклассник. Мы начнем с так называемого *классического (комбинаторного)* определения вероятности и лишь в конце статьи придем, через неравенство Чебышёва, к *закону больших чисел*, отчасти объясняющему связь между вероятностью события и частотой его появления.

Большая часть текста статьи — задачи, одни из которых решены в тексте, а другие оставлены в качестве упражнений (все они занумерованы по порядку, к некоторым даны указания).

Комбинаторное определение

Пусть у нас имеется конечное множество E, состоящее из N элементов. Элементы множества E называются *элементарными событиями* (возможными исходами испытаний). *Событием* называется любое подмножество A множества E. *Вероятность* $\mathbf{P}(A)$ события A задается формулой

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

где $|A|$ — число элементов множества A.

Это значит, что вероятность каждого элементарного события (отдельного элемента множества E) равна $1/N$, а вероятность любого события A складывается из них как из атомов.

Такое определение вероятности годится, например, для задач о бросаниях монеты или игрального кубика и вообще для любой ситуации, где есть несколько равноправных (обычно — из соображений симметрии) возможных исходов.

1. *Игральный кубик бросают два раза. Какова вероятность, что*

а) в первый раз выпало меньше 4 очков, а во второй — больше 4;

б) в первый раз выпало меньше очков, чем во второй;

в) в сумме за два броска выпало k очков, где $k = 1, 2, \ldots, 12$?

Решение. Бросить кубик два раза — все равно что независимо друг от друга бросить два кубика. Поэтому всего здесь $6 \cdot 6 = 36$ возможных исходов. Это — пары (x, y), где x — число очков, выпавших на первом кубике, а y — на втором ($1 \le x \le 6$ и $1 \le y \le 6$). Их удобно изображать 36 точками, расположенными в виде квадрата 6×6 (рис. 1).

(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)

Рис. 1

Осталось отметить точки (x, y), удовлетворяющие соответствующему условию, посчитать их количество и разделить на 36 (рис. 2).

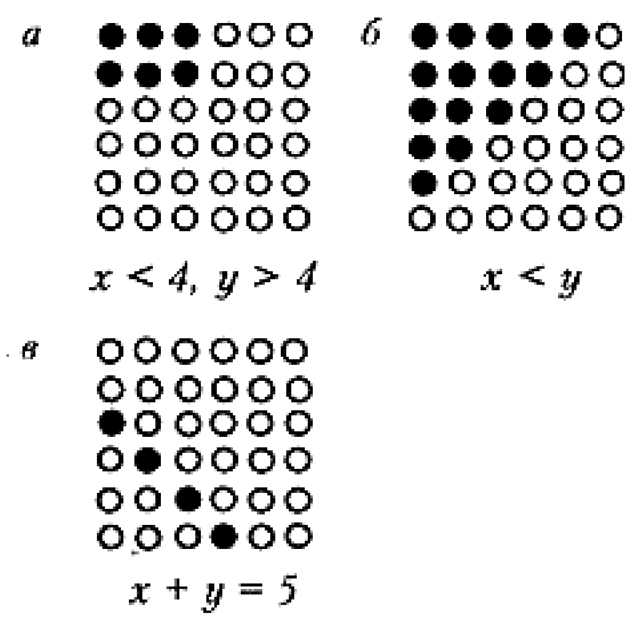

Рис. 2

Получаем ответы: а) $3 \cdot 2/36 = 1/6$; б) $15/36 = 5/12$. А ответ в пункте в) изображен в виде *гистограммы*, где высоты столбиков соответствуют вероятностям отдельных событий (рис. 3).

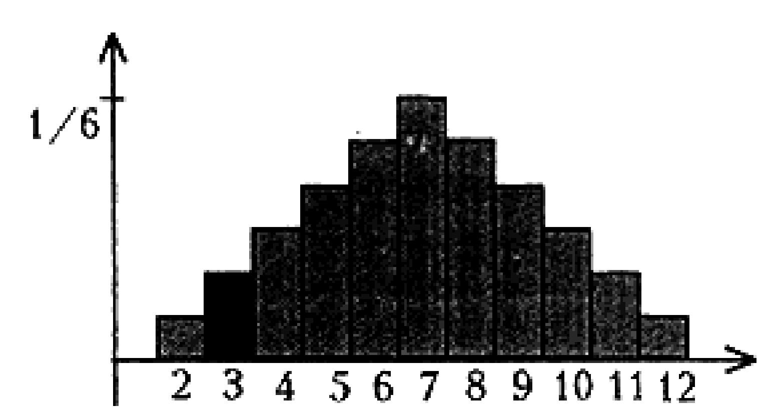

Рис. 3

Теперь разберем три задачи посложнее.

(Продолжение см. на с. 34)

(Начало см. на с. 31)

Пары, тройки, четверки...

2. *Федя знает ответы на 10 вопросов из 30 возможных. В билет включаются два случайно выбранных вопроса из этих 30. Каковы шансы Феди благополучно ответить на оба вопроса?*

Решение. Выясним, сколько всего разных билетов может составить экзаменатор. Их столько, сколько пар можно составить из 30 элементов, т.е. $30 \cdot 29 = 870$.

Как мы нашли это число? Первый вопрос — это любой из 30 возможных. Если он уже выбран, то вторым вопросом может быть любой из 29 оставшихся.

Точно так же посчитаем количество билетов, «благоприятных» для Феди: $10 \cdot 9 = 90$.

Значит, вероятность сдать экзамен равна $90/870 = 3/29 \approx 0{,}103$.

Мы могли рассуждать и чуть иначе: не различать билеты, отличающиеся только порядком вопросов. Тогда разных билетов будет вдвое меньше, т.е. $30 \cdot 29/2 = 435$. Благоприятных для Феди станет тоже в два раза меньше: $10 \cdot 9/2 = 45$. Ответ, конечно, получится тот же самый.

Число неупорядоченных пар из n элементов обозначается C_n^2. Это число встречается во многих ситуациях и вычисляется, как мы уже поняли, по формуле

$$C_n^2 = \frac{n(n-1)}{2}.$$

3. *Придя на экзамен, Федя узнал, что в билете не 2 вопроса, а 3. Каковы шансы Феди благополучно ответить на все вопросы?*

Решение. На этот раз можно составить $30 \cdot 29 \cdot 28 = 24360$ разных билетов. В самом деле, первый вопрос можно выбрать 30 способами. Как бы он ни был выбран, вторым может оказаться любой из 29 остальных, а после выбора первых двух вопросов третьим может оказаться любой из 28 остальных.

Точно так же, благоприятных для Феди билетов $10 \cdot 9 \cdot 8 = 720$, так что вероятность сдать экзамен равна теперь лишь $720/24360 = 6/203 < 0{,}03$.

Конечно, и в этой задаче можно было не обращать внимания на порядок вопросов в билете, т.е. считать количество трехэлементных подмножеств $\{P, Q, R\}$ множества всех вопросов. Это число в 6 раз меньше, чем $30 \cdot 29 \cdot 28$, поскольку три элемента P, Q и R можно переставить 6 способами: PQR, PRQ, QPR, QRP, RPQ и RQP.

Итак, Федя мог получить любой из $\frac{30 \cdot 29 \cdot 28}{6} = 4060$ разных по содержанию билетов. Благоприятных из них только $10 \cdot 9 \cdot 8/6 = 120$. Разделив 120 на 4060, мы получим ту же вероятность.

Число неупорядоченных троек из n элементов обозначается C_n^3. Оно равно

$$C_n^3 = \frac{n(n-1)(n-2)}{6}.$$

4. *Конечно, Феде не повезло, и экзамен он не сдал. А когда Федя пришел в следующий раз, экзаменатор изменил правила: он задает четыре вопроса (случайно выбирая их из 30 возможных). Если Федя ответит на все 4 вопроса, то получит пятерку, если на 3 вопроса — то четверку, если на 2 — тройку, а в противном случае — двойку.*

С какой вероятностью Федя получит оценку а) 5; б) 4; в) 3; г) 2?

Решение. а) Как мы уже видели, всего можно составить $30 \cdot 29 \cdot 28 \cdot 27$ разных билетов. Из них $10 \cdot 9 \cdot 8 \cdot 7$ билетов состоят только из известных Феде вопросов, так что вероятность сдать экзамен на «5» равна

$$\frac{10 \cdot 9 \cdot 8 \cdot 7}{30 \cdot 29 \cdot 28 \cdot 27} = \frac{2}{261} \approx 0{,}0076.$$

В решении пункта а) мы считали разными билеты, отличающиеся лишь порядком вопросов. В следующих пунктах мы не будем их различать, т.е. будем рассматривать не упорядоченные наборы, а подмножества, состоящие из 4 вопросов. Их количество в 24 раза меньше, чем $30 \cdot 29 \cdot 28 \cdot 27$, поскольку 4 элемента можно переставить 24 способами:

$PQRS, PQSR, PRQS, PRSQ, PSQR, PSRQ,$
$QPRS, QPSR, QRPS, QRSP, QSPR, QSRP,$
$RPQS, RPSQ, RQPS, RQSP, RSPQ, RSQP,$
$SPQR, SPRQ, SQPR, SQRP, SRPQ, SRQP.$

Итак, количество 4-элементных подмножеств 30-элементного множества равно

$$C_{30}^4 = \frac{30 \cdot 29 \cdot 28 \cdot 27}{24} = 27405.$$

б) Федя получит «4», если ему достанется билет, в котором 1 незнакомый вопрос (любой из 20) и 3 знакомых вопроса. Эти 3 вопроса из 10 можно выбрать $C_{10}^3 = \frac{10 \cdot 9 \cdot 8}{6} = 120$ способами. Поэтому вероятность получить «4» равна

$$\frac{20 \cdot 120}{27405} \approx 0{,}0876.$$

в) Для получения тройки нужен билет, в котором 2 знакомых и 2 незнакомых вопроса. Эта вероятность равна

$$\frac{C_{10}^2 C_{20}^2}{C_{30}^4} = \frac{(10 \cdot 9/2) \cdot (20 \cdot 19/2)}{27405} \approx 0{,}312.$$

г) Двойку можно получить двояко: ответив на один вопрос билета или не ответив ни на один. Соответствующие вероятности равны

$$\frac{10 \cdot C_{20}^3}{27405} = \frac{10 \cdot (20 \cdot 19 \cdot 18/6)}{27405} \approx 0{,}416$$

и

$$\frac{C_{20}^4}{27405} = \frac{20 \cdot 19 \cdot 18 \cdot 17/24}{27405} \approx 0{,}177.$$

Значит, вероятность получить двойку примерно равна

$$0{,}416 + 0{,}177 = 0{,}593.$$

Ответ можно записать в виде таблички:

Отметка	2	3	4	5
Вероятность	0,593	0,312	0,0876	0,0076

Наглядно это распределение можно изобразить гистограммой (рис. 4).

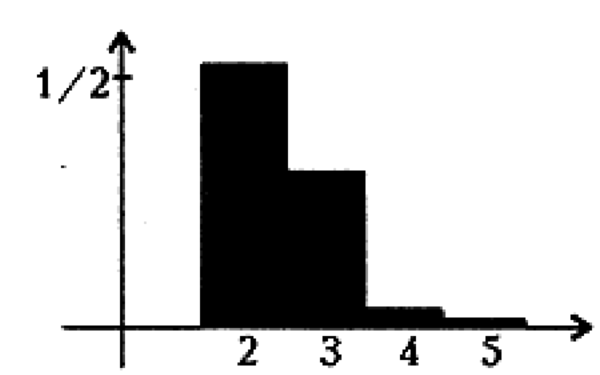

Рис. 4

Заметим, что сумма

$$0{,}0076 + 0{,}0876 + 0{,}312 + 0{,}593 = 1{,}002 \approx 1.$$

На самом деле, конечно, сумма этих вероятностей в точности равна 1, ибо какую-то одну из оценок 2, 3, 4 или 5 Федя получит.

Вообще, если все пространство E элементарных событий разбито на несколько (непересекающихся) множеств $A_1, A_2, \ldots, A_r$, то общее число их элементов равно $N = |E|$:

$$|A_1| + |A_2| + \ldots + |A_r| = |E|$$

и, следовательно, сумма вероятностей событий $A_1, A_2, \ldots, A_r$ равна 1:

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \ldots + \mathbf{P}(A_r) = \frac{|A_1|}{N} + \frac{|A_2|}{N} + \ldots + \frac{|A_r|}{N} = \frac{N}{N} = 1. \quad (1)$$

Про такие события $A_1, A_2, \ldots, A_r$ говорят, что они образуют *полную систему событий*.

Здесь очень важно, что никакие два из этих событий не пересекаются. Вообще, если два множества A и B (содержащиеся в E) не пересекаются, то говорят, что события A и B *несовместны*. При этом вероятность того, что происходит хотя бы одно из них — вероятность попасть в A или в B, — равна сумме вероятностей событий A и B:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B). \qquad (2)$$

Мы пользовались этим выше в решении пункта г) задачи 4.

Упражнения

5. Предположим, что Федя знал ответы не на 10, а на 20 вопросов из 30. Каковы тогда были бы вероятности получить оценки 2, 3, 4, 5? Нарисуйте соответствующую гистограмму.

6. 20 человек написали письма друг другу. Сколько всего писем было послано? (Каждый написал письма всем остальным, по одному письму каждому.)

7. 15 гроссмейстеров провели турнир, в котором каждые двое сыграли между собой одну партию. Сколько партий было сыграно?

8. Среди двузначных чисел случайным образом выбираем одно число. С какой вероятностью а) его цифра десятков больше цифры единиц; б) обе цифры равны; в) это число делится на 9?

9. Из 25 учеников класса, в котором учатся Денис, Данила и Димитрий, случайным образом выбрали 10 учеников, которые должны прийти на экзамен к 9 утра, и 10 человек, которые должны прийти к 12 часам. Остальные 5 учеников должны прийти к 14 часам. Какова вероятность, что Денис, Данила и Димитрий а) втроем попадут в первую группу; б) попадут в одну и ту же группу; в) попадут в три разные группы?

Правило произведения. Перестановки и сочетания

Решая задачи про Федю, мы подсчитывали количества разных комбинаций: перестановок, подмножеств конечного множества и тому подобное. Как мы это делали? Основной прием, которым мы пользовались, — правило произведения. Каждый понимает, что если в лесу 100 елок, у каждой елки 20 веток, на каждой ветке 40 лапок, а на каждой лапке 70 иголок, то всего в лесу $100 \cdot 20 \cdot 40 \cdot 70$ иголок.

Вообще, *если объект x можно выбрать m способами и если после каждого такого выбора объект y можно выбрать n способами, то выбор пары (x,y) можно осуществить mn способами.*

Особенно простой вид правило произведения приобретает, если выбор второго элемента y пары (x,y) совсем не зависит от x. Тогда все mn пар удобно записать в таблицу (рис.5).

$$\begin{array}{cccc}
(x_1,y_1) & (x_1,y_2) & \ldots & (x_1,y_n) \\
(x_2,y_1) & (x_2,y_2) & \ldots & (x_2,y_n) \\
\ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
(x_m,y_1) & (x_m,y_2) & \ldots & (x_m,y_n)
\end{array}$$

Рис. 5

10. *В магазине «Все для чая» есть 5 разных чашек и 3 разных блюдца. Сколькими способами можно купить чашку и блюдце?*

Решение. Выберем чашку. В комплект к ней можно выбрать любое из трех блюдец. Поскольку чашек всего 5, то число различных комплектов равно 15 ($15 = 3 \cdot 5$).

11. *В магазине «Все для чая» есть еще 4 чайные ложки. Сколькими способами можно купить комплект из чашки блюдца и ложки?*

Решение. Выберем любой из 15 комплектов предыдущей задачи. Его можно дополнить ложкой четырьмя различными способами. Поэтому общее число возможных комплектов равно $60 = 15 \cdot 4 = 5 \cdot 3 \cdot 4$.

12. *Выбирают случайным образом шестизначное число. С какой вероятностью вторая его цифра равна 3, а четвертая четна?*

Решение. Всего шестизначных чисел 900000. Из них удовлетворяют условию $9 \cdot 1 \cdot 10 \cdot 5 \cdot 10 \cdot 10$ чисел, так что вероятность равна

$$\frac{9 \cdot 1 \cdot 10 \cdot 5 \cdot 10 \cdot 10}{9 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10} = \frac{1}{10} \cdot \frac{5}{10} = 0{,}05 .$$

Заметьте, что ответ равен произведению $\frac{1}{10}$ (это вероятность того, что вторая цифра равна 3) и $\frac{5}{10}$ (это вероятность того, что четвертая цифра четна). С этим мы уже встречались в задаче 1, а), где вероятность оказалась равной произведению $\frac{1}{2}$ и $\frac{1}{3}$.

Объясним, почему так получается. Представим себе, что мы случайным образом выбираем одну из $N = mn$ клеток таблицы $m \times n$ (см. рис.5), причем событие A состоит в том, что клетка попадает в одну из заданных k строк, а B — что она попадает в один из заданных l столбцов. Тогда пересечение $A \cap B$ состоит из kl клеток и поэтому

$$\mathbf{P}(A \cap B) = \frac{kl}{N} = \frac{k}{m} \cdot \frac{l}{n} = \mathbf{P}(A)\mathbf{P}(B).$$

Вообще, два события A и B называются *независимыми*, если

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\ \mathbf{P}(B). \qquad (3)$$

13. *В ряд расположены n фонарей. Сколькими способами можно осветить улицу, считая и тот «способ освещения», когда ни один фонарь не горит? (Говоря математическим языком, надо подсчитать количество подмножеств множества из n элементов.)*

Решение. Первый фонарь может находиться в двух состояниях: гореть или не гореть. Независимо от него второй фонарь может гореть или не гореть. Значит, первые два фонаря могут находиться в $2 \cdot 2 = 4$ состояниях. Добавляя третий фонарь, мы удвоим число возможных состояний. Точно так же, добавление каждого следующего фонаря будет удваивать число состояний. Таким образом, имеется 2^n способов освещения n фонарями.

Заодно мы посчитали количество последовательностей длины n из нулей и единиц. Оно равно 2^n. В само деле, зажженный фонарь можно обозначать единицей, а негорящий — нулем. (Тогда последовательность из n нулей означает, что ни один фонарь не горит, т.е. соответствует пустому множеству. А последовательность из n единиц означает, что все фонари горят.)

Используя правило произведения, легко посчитать общее количество *перестановок*, которые можно составить из n элементов, т.е. количество способов, которыми можно все эти элементы расположить в ряд. Мы уже видели, что оно равно 6 для $n = 3$ и 24 для $n = 4$. В общем случае оно равно

$$n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1.$$

В самом деле, если нам нужно расположить в строку n элементов, то первый можно выбрать n способами, после чего второй можно выбрать $(n-1)$ способами, после каждого выбора первого и второго третий — $(n-2)$ способами,..., предпоследний элемент выбирается из двух оставшихся к этому моменту, а последний определен однозначно.

Займемся теперь не менее важными для комбинаторики и теории вероятностей числами сочетаний C_n^k. Мы получим их, решая следующую задачу.

14. *В ряд расположены 5 лампочек. Сколькими способами можно зажечь k из них (для каждого k = 0, 1, 2, 3, 4 и 5)? А если лампочек n?*

Решение. 0 ламп можно «зажечь» только одним способом — ни одну не зажигать. 1 лампу из 5 можно выбрать 5 способами, 2 лампы из 5 — 10 способами (действительно, первую можно зажечь 5 способами, после этого остается 4 возможности зажечь вторую лампу; но число $5 \cdot 4$ надо разделить пополам, ибо когда лампы горят, не ясно, какая была зажжена раньше).

Для $k = 3$ ламп получим ответ $5 \cdot 4 \cdot 3/3! = 10$, для $k = 4$ получим

$5 \cdot 4 \cdot 3 \cdot 2/4! = 5$, а для $k = 5 - 1$ способ. (Впрочем, сразу можно было сообразить, что строчка ответов 1, 5, 10, 10, 5, 1 получится симметричной: ведь случаев, когда k ламп горят и остальные $5 - k$ не горят, столько же, сколько случаев, когда все наоборот.)

Вы, конечно, заметили, что, решая задачи 2 и 3 про Федю, мы именно так получали формулы для C_n^2 и C_n^3.

Точно так же выводится и общая формула для C_n^k — числа способов зажечь k ламп из n. Первую мы выбираем из n возможностей, вторую — из $(n-1)$, третью — $(n-2)$, и так k раз (при этом заметьте, что k-ю лампу мы выбираем из $n-k+1$ варианта, а вовсе не из $n-k$, как можно подумать сгоряча). А делить надо на $k!$ — число перестановок k элементов:

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

Числа сочетаний C_n^k в англоязычных учебниках обозначают иначе: $\binom{n}{k}$. Они появляются в разных ситуациях. C_n^k — это

число k-элементных подмножеств множества из n элементов;

число последовательностей длины n, состоящих из k единиц и $(n-k)$ нулей;

число путей длины n из точки $(0,0)$ в точку $(k, n-k)$ по сторонам единичных клеток;

коэффициент при x^k после раскрытия скобок в выражении $(x+1)^n$.

Биномиальные коэффициенты удобно расположить в виде треугольника Паскаля (рис.6), где каждое число

$$\begin{array}{c}
1\\
1\quad 1\\
1\quad 2\quad 1\\
1\quad 3\quad 3\quad 1\\
1\quad 4\quad 6\quad 4\quad 1\\
1\quad 5\quad 10\quad 10\quad 10\quad 1\\
\cdots\cdots\cdots\cdots\\
C_n^0\; C_n^1\; C_n^2\;\ \ldots\ \; C_n^{n-1}\; C_n^n\\
\cdots\cdots\cdots\cdots\cdots
\end{array}$$

Рис. 6

равно сумме двух стоящих над ним, что соответствует формуле $C_{n+1}^{k+1} = C_n^k + C_n^{k+1}$.

Числа сочетаний еще понадобятся нам ниже.

Упражнения

15. Из города A в город B ведут 4 дороги, из B в C — 6 дорог. Сколькими способами можно проехать из A в C?

16. Сколько существует трехзначных чисел, у которых все цифры нечетны и различны?

17. Сколькими способами можно отметить 8 полей шахматной доски, чтобы никакие два отмеченных поля не лежали ни на одной вертикали, ни на одной горизонтали?

18. Сколько всего шестизначных чисел, все цифры которых нечетные?

19. Выбирают случайным образом трехзначное нечетное число. С какой вероятностью

а) все три цифры разные;

б) первая цифра равна 1;

в) первая цифра равна последней?

20. Пусть в ряд расположены n фонарей, каждый из которых горит или не горит с вероятностью $1/2$. Найдите вероятность того, что горят ровно k фонарей.

21. Из юго-западного угла квадратного города $n \times n$, разбитого улицами на единичные квадраты, выходят 2^n человек. Половина идет на север, половина — на восток. Из всех, кто дошел до очередного перекрестка, половина снова идет на север, половина — на восток. Сколько человек придет на k-й перекресток диагонали?

Сложение и умножение вероятностей

Не всегда вероятности вычисляются по формуле $\mathbf{P}(A) = |A|/N$. Иногда удается одни вероятности вычислять через другие. Например, если даны два множества A и B (содержащиеся в E), то число элементов объединения $A \cup B$ можно посчитать, сложив числа элементов множеств A и B и вычтя число элементов их пересечения $A \cap B$:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Поделив на N, получаем для вероятностей

$$\mathbf{P}(A \cup B) = \mathbf{P}(A)\mathbf{P}(B) - \mathbf{P}(A \cap B). \quad (4)$$

Частный случай этой формулы — для несовместных событий A и B — формула (2).

22. *На дискотеке 80% времени был выключен свет, 90% времени играла музыка и 50% времени шел дождь. Какую наименьшую долю времени все это обязано было происходить одновременно?*

Решение. Перейдем к дополнительным событиям: свет был выключен 20% времени, музыка молчала 10%, а дождь не шел 50% времени, так что дополнительные события не могли занять более 20 + 10 + 50 = 80% времени. Следовательно, музыка под дождем в темноте звучала не меньше 100 − 80 = 20% времени.

В этой задаче мы воспользовались переходом от события A к дополнительному $E \setminus A$,[1] которое для краткости часто обозначают $\overline{A}$ (читается: «дополнение к A» или «не A»). При этом

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

23. Сколько существует шестизначных чисел, в записи которых есть хотя бы одна четная цифра?

Подсказка. Вместо того, чтобы подсчитывать количество требуемых шестизначных чисел, определите количество шестизначных чисел, у которых все цифры нечетны.

24. *В классе 40% мальчиков и 60% девочек. Из мальчиков на роликовых коньках катается каждый второй, а из девочек — 30%. Какая доля учеников этого класса катается на роликовых коньках?*

Решение. Пусть в классе n человек. Тогда мальчиков всего $0{,}4n$, а девочек $0{,}6n$. Значит, катаются $0{,}5 \cdot 0.4n$ мальчиков и $0{,}3 \cdot 0.6n$ девочек, так что искомая доля равна

$$\frac{0{,}5 \cdot 0{,}4n + 0{,}3 \cdot 0{,}6n}{n} =$$

$$= 0{,}5 \cdot 0{,}4 + 0{,}3 \cdot 0{,}6 = 0{,}38.$$

Последняя формула, если бы речь шла не о процентах и долях, а о вероятностях, называлась бы *формулой полной вероятности*. В общем виде она выглядит так: если $A_1, A_2, \ldots, A_r$ — полная система событий, то

$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}_{A_1}(B) +$$

$$+ \mathbf{P}(A_2)\mathbf{P}_{A_2}(B) + \ldots + \mathbf{P}(A_r)\mathbf{P}_{A_r}(B). \quad (5)$$

Объясним обозначения. Пусть A и B — два события (например, A — «быть мальчиком», B — «кататься на коньках»). Тогда *условной вероятностью* $P_A(B)$ *события B при условии A* называется отношение $\mathbf{P}(A \cap B)/\mathbf{P}(A)$. (В нашем примере $\mathbf{P}_A(B) = 0{,}5$, $\mathbf{P}_{\overline{A}}(B) = 0{,}3$.)

Чтобы пояснить это определение, вспомним комбинаторное определение вероятности. Если мы желаем перейти от всего пространства E к меньшему пространству A, то следует учитывать только те элементы из B, которые содержатся в A, а их количество равно $|A \cap B|$. При этом $\mathbf{P}(A \cap B) = |A \cap B|/|E|$ и $\mathbf{P}(A) = |A|/|E|$. Отсюда и получается формула условной вероятности

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

Условную вероятность $\mathbf{P}_A(B)$ обозначают и так: $\mathbf{P}(B|A)$. Некоторым больше нравится одно обозначение, некоторым — другое.

[1] *Вообще, для любых двух множеств X и Y через $X \setminus Y$ обозначают множество тех элементов из X, которые не принадлежат Y.*

Часто используется такая формула, вытекающая из определения условной вероятности:

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

Между прочим, поскольку в левую часть буквы A и B входят симметрично, заодно выполняется и формула

$$\mathbf{P}(B)\ \mathbf{P}_B(A)=\mathbf{P}(A\cap B)=\mathbf{P}(A)\ \mathbf{P}_A(B).$$

Из формулы (7) легко получить формулу полной вероятности:

$$\mathbf{P}(B)=\mathbf{P}(A_1\cap B)+ \\ +\mathbf{P}(A_2\cap B)+\ldots+\mathbf{P}(A_r\cap B)= \\ =\mathbf{P}(A_1)\mathbf{P}_{A_1}(B)+ \\ +\mathbf{P}(A_2)\mathbf{P}_{A_2}(B)+\ldots+\mathbf{P}(A_r)\mathbf{P}_{A_r}(B).\quad(7)$$

Упражнения

25. Мне прислали две посылки. В одной из них 20% груш, 50% яблок и 30% апельсинов, а в другой — 30% груш, 10% яблок и 60% киви. Я, зажмурив глаза, случайным образом выбираю посылку и в ней — фрукт. Какова вероятность, что я выберу апельсин?

26. Известно, что при броске кости выпало четное число. Какова вероятность того, что это число меньше пяти?

27. Докажите, что если события A и B независимы, то и события A и $\overline{B}$ независимы.

28. Двое друзей подкидывают три монеты и хотят вычислить вероятность того, что все три выпадут одной стороной — орлом или решкой. Первый утверждает, что эта вероятность равна $\frac{1}{4}$. Он рассуждает так: вероятность того, что вторая монета ляжет так же, как первая, равна $\frac{1}{2}$, а вероятность того, что третья монета ляжет тем же способом, вдвое меньше — т.е. равна $\frac{1}{4}$.

Второй утверждает, что искомая вероятность равна $\frac{1}{2}$. Он рассуждает так: какие-нибудь две из трех монет обязательно выпадут одной и той же стороной. Вероятность того, что и третья монета ляжет тем же способом, равна $\frac{1}{2}$. Кто прав?

29. Про некоторую семью известно, что там двое детей. Как-то раз мама вывела на прогулку одного (случайно выбранного) ребенка. Оказалось, что это мальчик. Что более вероятно: что второй ребенок является мальчиком или девочкой?

30 (Задача Паскаля). Два одинаково искусных игрока играют в игру, не допускающую ничейного исхода. Они сделали равные ставки и условились, что тот, кто первым наберет 10 выигранных партий, получит все деньги. Игра была прервана при счете 9 : 8 и не могла быть продолжена. Как должны они разделить деньги?

Случайная величина и среднее значение

Как мы уже говорили во введении, на практике вероятности находят, многократно повторяя опыт и вычисляя долю случаев («частоту»), в которых произошло интересующее нас событие. Например, если много раз подбросить монету, то она упадет цифрой вверх примерно в половине случаев. Такая «устойчивость частоты» при многократном повторении испытания наблюдается во многих ситуациях. Математическое объяснение этой устойчивости дал Я.Бернулли в книге «Искусство предположения», опубликованной в 1713 году. Он установил *закон больших чисел*. Если в каждом из n независимых испытаний с одной и той же вероятностью p может произойти некоторое событие A, то количество Z появлений события A не обязано в точности равняться np и может сильно отклоняться от этой величины; но вероятности значительных отклонений малы: *для всяких положительных чисел* ε *и* η *вероятность* $\mathbf{P}\left(\left|\frac{Z}{n}-p\right|>\varepsilon\right)$ *будет меньше* η *при всех достаточно больших* n.

Более простое, чем у Бернулли, доказательство закона больших чисел получится в конце следующего параграфа из неравенства Чебышёва.

В этом разделе мы будем заниматься такой задачей. Пусть некоторое событие A происходит с вероятностью p. Какова вероятность, что за n независимых испытаний оно произойдет ровно k раз?

Сначала надо придать точный смысл словам «n независимых испытаний A с одной и той же вероятностью p». Для этого служит очень важное вероятностное пространство — так называемая *схема Бернулли*. Оно состоит из 2^n элементарных событий — строк $(z_1,z_2,\ldots,z_n)$ из n нулей и единиц. Вероятность, приписываемая строке, в которой k единиц и $n-k$ нулей, равна $p^k(1-p)^{n-k}$. При $p=1/2$ получается вероятностное пространство, которое описывает n бросаний симметричной монеты и встречалось в задаче 20. Заметим, что схема Бернулли (при $p\neq 1/2$) не подходит под « комбинаторное» определение вероятностного пространства E, где все «атомы» были равновероятны.

Вот более общее определение, кстати, более подходящее для практических применений.

Определение. *Конечным вероятностным пространством* называется конечное множество $E=\{e_1,e_2,\ldots,e_n\}$, каждому элементу e_i которого приписано неотрицательное число w_i (называемое вероятностью элементарного события e_i), причем их сумма равна единице:

$$w_1+w_2+\ldots+w_n=1.$$

Любому событию A (подмножеству $A\subseteq E$) приписывается вероятность

$$\mathbf{P}(A)=\sum_{e_i\in A}w_i.$$

Последняя формула означает, что $\mathbf{P}(A)$ — это сумма вероятностей всех элементарных событий, из которых состоит A.

Заметим, что основные правила вычисления вероятностей (1) — (7), о которых шла речь выше, при этом сохраняются.

Далее нам потребуются другие новые понятия: случайная величина, ее распределение, ее среднее значение и т.п. Собственно, эти понятия намного старше теории вероятностей и всем хорошо знакомы: они относятся к традиционной статистике, возникшей одновременно с умением записывать числа.

Определение. *Случайной величиной* называется функция X, заданная на множестве E.

Каждому событию $A\subseteq E$ соответствует «характеристическая» случайная величина ξ_A, принимающая значения 0 и 1: $\xi_A(e)=1$, если $e\in A$, и $\xi_A(e)=0$ в противном случае (так что можно считать, что «случайная величина» — некоторое обобщение понятия «событие»).

Поскольку мы рассматриваем только конечные множества E, всякая случайная величина $X:E\to\mathbf{R}$ может быть задана набором чисел $X(e_i)$, где $i=1,2,\ldots,N$, которые можно записать в виде таблички. Например, в задаче 1 о двух бросаниях кубика величина «сумма выпавших очков» может быть представлена такой таблицей:

Элементарное событие	(1,1)	(1,2)	(2,1)	(1,3)	...	(6,6)
Вероятнось	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	...	$\frac{1}{36}$
X	2	3	3	4	...	12

Если вероятностное пространство состоит из большого числа элементов, то таблица становится совершенно необозримой. Между тем, обычно достаточно знать *распределение* случайной величины, т.е. перечень всех возможных ее значений $\{x_1,\ldots,x_m\}$ и вероятность $u_j=\mathbf{P}(X=x_j)$ каждого из них. Здесь $X=x_j$ — это событие «случайная величина X равна x_j»; его вероятность u_j — сумма w_i по всем e_i, для которых $X(e_i)=x_j$. События $X=x_j$ $(j=1,2,\ldots$

..., m) образуют, очевидно, полную систему событий, так что $u_1 + u_2 + \ldots + u_m = 1$. Распределение случайной величины полностью определяет ее основные свойства — среднее значение, величину «разброса», наличие лишь одного или нескольких наиболее вероятных значений... В примере с кубиком распределение таково:

x_j	2	3	4	5	6	7	8	9	10	11	12
u_j	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

А в схеме Бернулли случайная величина «число единиц» — обозначим ее z — имеет *биномиальное распределение*

$$\mathbf{P}(z=k) = C_n^k p^k (1-p)^{n-k}.$$

Распределение случайной величины «оценка Феди на экзамене» из задачи 4 (обозначим ее Φ) задано таблицей или гистограммой рисунка 4.

Определение. *Средним значением* случайной величины X называется сумма

$$\mathbf{M}(X) = \sum_{i=1}^{N} w_i X(e_i) = \\ = w_1 X(e_1) + w_2 X(e_2) + \ldots + w_N X(e_N).$$

Зная распределение $\mathbf{P}(X = x_j) = u_j$ ($j = 1, 2, \ldots, m$), эту формулу можно, собрав вместе слагаемые с одинаковыми значениями $X(e_i)$, переписать в виде

$$\mathbf{M}(X) = \sum_{j=1}^{m} u_j x_j = u_1 x_1 + u_2 x_2 + \ldots + u_m x_m.$$

Среднее значение называют также *математическим ожиданием* случайной величины.

Среднее значение характеристической функции $\xi = \xi_A$ события A, принимающей только значения 0 и 1, равно вероятности события A:

$$\mathbf{M}(\xi) = \mathbf{P}(\xi = 0)\cdot 0 + \mathbf{P}(\xi = 1)\cdot 1 = \\ = \mathbf{P}(\xi = 1) = \mathbf{P}(A).$$

Среднее значение оценки Феди равно

$$\mathbf{M}(\Phi) \approx 0{,}0076\cdot 5 + 0{,}0876\cdot 4 + \\ + 0{,}312\cdot 3 + 0{,}593\cdot 2 \approx 2{,}51.$$

Среднее значение числа единиц в схеме Бернулли равно по определению

$$\mathbf{M}(Z) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k}\cdot k.$$

Подсчитаем эту сумму. Поскольку при $k \ge 1$

$$C_n^k \cdot k = \frac{n!}{k!(n-k)!}\cdot k = \frac{n!}{(k-1)!(n-k)!} = nC_{n-1}^{k-1}$$

имеем

$$\mathbf{M}(Z) = \sum_{k=1}^{n} nC_{n-1}^{k-1} p^k (1-p)^{n-k} = \\ = np\sum_{k=1}^{n} p^{k-1}(1-p)^{n-k} = \\ = np\sum_{l=0}^{n-1} p^l (1-p)^{n-1-l}.$$

По формуле бинома Ньютона сумма $\sum_{l=0}^{n-1} p^l (1-p)^{n-1-l}$ равна $(p + (1-p))^n = 1$. Значит, среднее значение величины Z равно np.

Но гораздо проще считать это среднее значение при помощи следующего свойства

Теорема 1. *Среднее значение суммы $X+Y$ случайных величин X и Y равно сумме их средних:*

$$\mathbf{M}(X+Y) = \mathbf{M}(X) + \mathbf{M}(Y).$$

Доказательство. По определению,

$$\mathbf{M}(X) = w_1 X(e_1) + w_2 X(e_2) + \ldots + w_N X(e_N), \tag{8}$$

$$\mathbf{M}(Y) = w_1 Y(e_1) + w_2 Y(e_2) + \ldots + w_N Y(e_N). \tag{9}$$

Складывая эти равенства почленно, получаем

$$\mathbf{M}(X) + \mathbf{M}(Y) = w_1(X(e_1) + Y(e_1)) + \\ + w_2(Xe(e_2) + Y(e_2)) + \ldots \\ \ldots + w_N(X(e_N) + Y(e_N)) = \mathbf{M}(X+Y).$$

Разумеется, утверждение теоремы верно и для суммы нескольких случайных величин.

Чтобы применить его к нахождению $\mathbf{M}(Z)$, представим Z в виде суммы n случайных величин

$$Z = Z_1 + Z_2 + \ldots + Z_n,$$

где Z_j — значение j-й координаты строки $(Z_1, Z_2, \ldots, Z_n)$. Среднее значение каждого из n слагаемых равно

$$\mathbf{M}(Z_j) = p, \tag{10}$$

поэтому $\mathbf{M}(Z) = np$.

Для произведения средних не все так просто — не всегда математическое ожидание произведения равно произведению математических ожиданий. Например, если случайная величина X принимает значения 1 и -1 с равными вероятностями $1/2$, то $\mathbf{M}(X) = 0$, а $\mathbf{M}(X^2) = \mathbf{M}(1) = 1$.

Определение. Случайные величины X и Y называются *независимыми*, если для любых значений x_j и y_k события $X = x_j$ и $Y = y_k$ независимы, т.е.

$$\mathbf{P}(X = x_j, Y = y_k) = \mathbf{P}(X = x_j)\mathbf{P}(Y = y_k).$$

Теорема 2. *Среднее значение произведения независимых случайных величин X и Y равно произведению их средних значений:*

$$\mathbf{M}(XY) = \mathbf{M}(X)\mathbf{M}(Y).$$

Доказательство. Пусть величина X принимает значения $x_1, x_2, \ldots, x_m$, а величина Y принимает значения $y_1, y_2, \ldots, y_l$. Обозначим $\mathbf{P}(X = x_j) = u_j$, $\mathbf{P}(Y = y_k) = v_k$. Тогда

$$\mathbf{P}(X = x_j, Y = y_k) = u_j v_k,$$

так что

$$\sum_{j=1}^{m}\sum_{k=1}^{l} \mathbf{P}(X = x_j, Y = y_k) x_j u_k = \\ = \sum_{j=1}^{m}\sum_{k=1}^{l} u_j v_k x_j y_k = \\ = \sum_{j=1}^{m} u_j x_j \cdot \sum_{k=1}^{l} v_k y_k = \mathbf{M}(X)\mathbf{M}(Y).$$

Упражнения

31. Двое играют в такую игру. Если при броске кости выпадает 1 или 2, то первый выигрывает у второго 5 очков; в противном случае второй выигрывает у первого 2 очка. Для кого эта игра выгодна — для первого или для второго игрока?

32. Я доезжаю на работу обычно либо автобусом за 20 минут, либо троллейбусом за полчаса, причем автобусом езжу втрое чаще, чем троллейбусом. В виде исключения я раз в десять дней доезжаю на такси за 10 минут и раз в десять дней хожу пешком за 1 час. Сколько времени в день в среднем я трачу на дорогу?

33. Каждым ходом игрок бросает игральную кость и получает столько очков, сколько выпадет. К тому же, если выпадет шестерка, он бросает кость второй раз за тот же ход и получает дополнительно выпавшее число очков. Сколько в среднем очков игрок получает за ход?

34. Докажите, что если случайные величины X и Y связаны соотношением $Y = aX + b$, где a и b — постоянные числа, то $\mathbf{M}(Y) = a\mathbf{M}(X) + b$.

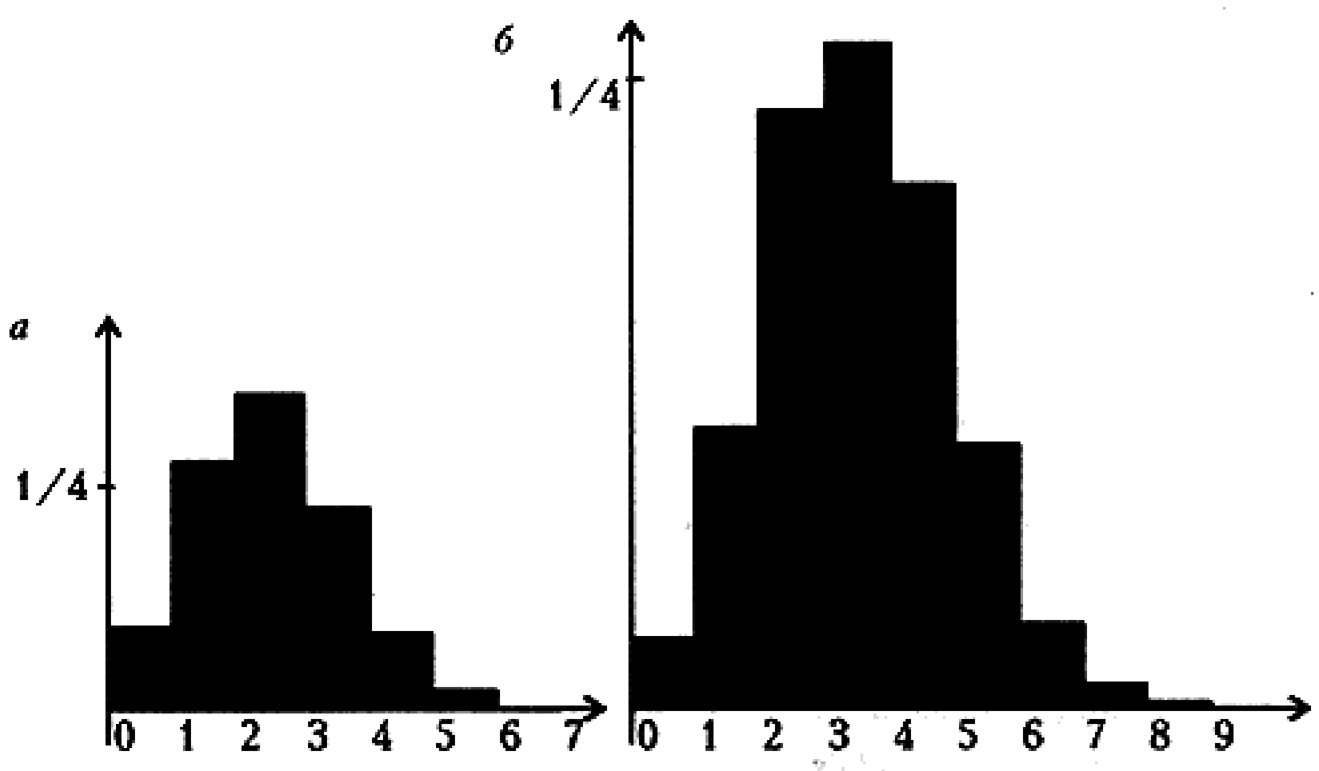

Рис. 7

35. Докажите, что если всегда $X \le Y$, то $\mathbf{M}(X) \le \mathbf{M}(Y)$.

36. Докажите, что если $\mathbf{M}(X^2) = \mathbf{M}(X)^2$, то X — постоянная величина.

37. Феде на экзамене задают а) 6; б) 9 вопросов, на каждый из которых он отвечает правильно с вероятностью $1/3$. Найдите вероятности, что он ответит правильно на k вопросов (соответствующие гистограммы изображены на рис.7, *а,б*).

Дисперсия и неравенство Чебышёва. Закон больших чисел

Кроме среднего значения $\mathbf{M}(X)$, случайная величина имеет другие характеристики. Если мы знаем, что в среднем за час на остановку приходит 10 автобусов, то отсюда еще не следует, что нам не придется ждать автобуса полчаса. Рассмотрим случайную величину $\hat{X} = X - \mathbf{M}(X)$ — отклонение X от математического ожидания (чем большие по модулю значения она имеет, тем больше «разброс» величины X).

Разумеется, математическое ожидание величины $\hat{X}$ равно 0. Математическое ожидание ее модуля может служить мерой разброса значений X. Но для математиков значительно более удобна другая характеристика, которую называют *дисперсией*:

$$\mathbf{D}(X) = \mathbf{M}(\hat{X})^2 = \mathbf{M}(X - \mathbf{M}(X))^2.$$

Дисперсия — это *квадрат* «характерного уклонения» величины X от ее среднего значения. Чем меньше дисперсия, тем острее (уже) гистограмма распределения.

38. Докажите, что

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

А теперь — самое замечательное свойство дисперсии.

Теорема 3. *Дисперсия суммы независимых случайных величин X и Y равна сумме их дисперсий:*

$$\mathbf{D}(X + Y) = \mathbf{D}(X) + \mathbf{D}(Y).$$

39. Докажите эту теорему.

Теорема 3 верна, конечно, и для суммы нескольких независимых случайных величин. Отсюда ясно, что для n одинаково распределенных независимых величин (например, для числа единиц схемы Бернулли) дисперсия их суммы равна $n\mathbf{D}$, где $\mathbf{D}$ — дисперсия одной величины. Значит, характерное уклонение суммы n случайных значений от ее математического ожидания равно $\sqrt{n\mathbf{D}}$ («закон корня из n»), тем самым, среднее арифметическое n значений отличается от вероятности p на величину порядка $\sqrt{\mathbf{D}}/\sqrt{n}$. На этом и строится доказательство закона больших чисел.

Докажите самостоятельно следующие утверждения:

40. Дисперсия величины Z схемы Бернулли равна $np(1-p)$.

Теорема 4. Если случайная величина X принимает только неотрицательные значения, то для любого положительного числа ε выполняется неравенство

$$\mathbf{P}(X \ge \varepsilon) \le \mathbf{M}(X)/\varepsilon.$$

Теорема 5. *Неравенство Чебышёва.* Для любого положительного числа ε и любой случайной величины X выполняется неравенство

$$\mathbf{P}(\hat{X} \ge \varepsilon) \le \mathbf{D}(X)/\varepsilon^2.$$

Теорема 6. *Закон больших чисел.* Если случайная величина Y есть сумма n независимых случайных величин, у каждой из которых среднее значение равно a, а дисперсия D, то для любого положительного числа ε

$$\mathbf{P}\left(\left|\frac{Y}{n} - a\right| > \varepsilon\right) < \frac{D}{n\varepsilon^2}.$$

Последние теоремы типичны для теории вероятностей. Мы не можем наверняка утверждать, что результат опыта (случайная величина) или среднее арифметическое нескольких опытов отличается от истинного среднего — математического ожидания — не более чем на ε, но можем быть уверены, что вероятность большого отклонения мала. Основное содержание теории вероятностей — различные неравенства, позволяющие оценивать вероятности.

Давайте посмотрим, какую оценку дает неравенство Чебышёва для 400 бросаний симметричной монеты. Вероятность $\mathbf{P}$ того, что число выпадений цифры отличается от 200 более чем на 20, оценивается так:

$$\mathbf{P} < \frac{D}{400 \cdot (1/20)^2} = \frac{1/4}{400 \cdot (1/20)^2} = \frac{1}{4},$$

где значение $D = 1/4$.

На самом деле, существуют замечательные теоремы, позволяющие значительно точнее оценивать такие величины. Так, на последнем примере $\mathbf{P} \approx 0{,}05$. Но об этом мы поговорим в следующий раз.