Е. А. Бунимович
В. А. Булычев

МАТЕМАТИКА

ВЕРОЯТНОСТЬ И СТАТИСТИКА

10 класс

БАЗОВЫЙ И УГЛУБЛЁННЫЙ УРОВНИ

Учебное пособие

Москва
«Просвещение»
2023

УДК 373.167.1:519.2+519.2(075.3)
ББК 22.17я721.6
Б91

12+

Бунимович, Евгений Абрамович.

Б91 Математика. Вероятность и статистика : 10-й класс : базовый и углублённый уровни : учебное пособие / Е. А. Бунимович, В. А. Булычев. — Москва : Просвещение, 2023. — 223, [1] с. : ил.

ISBN 978-5-09-110022-8.

Учебное пособие содержит информацию об элементах теории графов, случайных событиях и их вероятностях, сложении и умножении вероятностей, элементах комбинаторики, испытаниях Бернулли, случайных величинах и распределениях.

В учебном пособии содержатся лабораторные работы, выполнение которых предполагает использование электронных таблиц.

Данное учебное пособие разработано в соответствии с требованиями Федерального государственного образовательного стандарта среднего общего образования в редакции Приказа Министерства просвещения № 732 от 12.08.2022 г. и Федеральной образовательной программы среднего общего образования, утверждённой Приказом Министерства просвещения № 371 от 18.05.2023 г.

УДК 373.167.1:519.2+519.2(075.3)
ББК 22.17я721.6

Учебное издание

Бунимович Евгений Абрамович
Булычев Владимир Александрович

МАТЕМАТИКА

Вероятность и статистика

10 класс

Базовый и углублённый уровни

Учебное пособие

Центр математики, физики и астрономии
Ответственный за выпуск *Ю. С. Голубева*
Редактор *Ю. С. Голубева*
Художественный редактор *А. А. Шувалова*
Компьютерная вёрстка *О. В. Сиротиной*
Корректоры *О. Н. Леонова, М. А. Павлушкина*

Подписано в печать 17.08.2023.
Формат 84×108/16. Гарнитура NewtonSanPin.
Усл. печ. л. 23,52. Тираж экз. Заказ № .

Акционерное общество «Издательство «Просвещение».
Российская Федерация, 127473, г. Москва, ул. Краснопролетарская, д. 16, стр. 3, помещение 1Н.

Адрес электронной почты «Горячей линии» — **vopros@prosv.ru**.

ISBN 978-5-09-110022-8

ГЛАВА 1

ПРЕДСТАВЛЕНИЕ ДАННЫХ И ОПИСАТЕЛЬНАЯ СТАТИСТИКА

Статистика — это наука, которая изучает способы сбора, обработки и анализа количественных данных о самых разнообразных массовых явлениях окружающей нас жизни. Термин «статистика» происходит от латинского слова *status* — состояние. Правильно собранные и обработанные статистические данные дают достоверное представление о состоянии экономики, образования, медицины и многих других сфер нашей жизни. Поскольку статистика всегда имеет дело с большими объёмами данных, одна из главных её задач — удобное представление и описание этих данных, замена большого массива наглядной таблицей, диаграммой, одним или несколькими числовыми показателями. В этой главе мы поговорим о простейших методах обработки и анализа статистических данных.

§ 1. Представление данных

1 Таблицы

Одним из самых частых и привычных способов представления информации являются ***таблицы***: расписание уроков, таблица умножения, страницы школьного дневника и классного журнала, календари, программы телепередач, расписания движения поездов и автобусов, турнирные таблицы — этот ряд вы легко продолжите сами.

Простейшая таблица представляет собой совокупность ***строк*** и ***столбцов***, на пересечении которых образуются прямоугольные клетки или ячейки, куда и заносятся данные. Начальная строка часто используется для того, чтобы подписать каждый столбец. Иногда несколько соседних ячеек таблицы объединяют в одну.

Пример 1. Рассмотрим страницу журнала с отметками десятиклассников по математике за февраль (табл. 1).

Строки и столбцы этой таблицы содержат много полезной информации, которую из них несложно извлечь. Например:

- 3, 11 и 24 февраля проводились самостоятельные или контрольные работы — отметки стоят у всех присутствующих;
- во второй половине февраля, скорее всего, началась эпидемия гриппа — было очень много пропусков занятий;
- Борис Боровик успевает по математике очень хорошо, а Олег Орлов — плохо;
- Ерёмин Егор учится стабильно, а Игнатюк Ирина — очень неровно;
- 7, 14, и 21 февраля пришлись на воскресенье, так как занятий в эти дни не было.

Бумажный классный журнал постепенно уходит в прошлое и уступает место электронному. Об электронных таблицах мы ещё поговорим, а сейчас рассмотрим пример из другой области.

Таблица 1

№ п/п	ФИО	Февраль																						
		1	2	3	4	5	6	8	9	10	11	12	13	15	16	17	18	19	20	22	24	25	26	27
1	Анисимов Артём	4		3		2					4										5			3
2	Боровик Борис		5	4				5			5				5		4		н	н	5		5	
3	Васнецова Виктория	4	3	4						4	3			н							4			2
4	Галиева Гюзель			3	3			2			4					н	н	н			4			
5	Дорош Дмитрий	5		5				3			3			4							5			4
6	Ерёмин Егор			4					3		3				4				н	н	4	3		
7	Жукова Евгения			5				4			3		5					5			3			
8	Зыкова Зоя			5			3				4					3					н	н	н	н
9	Игнатюк Ирина		5	н			2				5					н	н	н	2		2		5	
10	Кулаева Касима			3		4				4	4		н	н							4			
11	Лисицын Лев			5							4				3			2		4	н			
12	Магомедов Мурад			4	5						н			н			5				4			
13	Ногина Анастасия			4				2	3		3		5								4			
14	Орлов Олег	3		3				3			2				2	4					3			3
15	Палий Павел			3							4							н	н	н	н			

Пример 2. Перед вами итоговая таблица чемпионата России по футболу 2022/23 г. (табл. 2).

Таблица 2

Место	Команда	И	В	Н	П	М	О
1	Зенит	30	21	7	2	74—20	70
2	ЦСКА	30	17	7	6	56—27	58
3	Спартак	30	15	9	6	60—38	54
4	Ростов	30	15	8	7	48—44	53
5	Ахмат	30	15	5	10	51—39	50
6	Краснодар	30	13	9	8	62—46	48
7	Оренбург	30	14	4	12	58—55	46
8	Локомотив	30	13	6	11	54—46	45
9	Динамо	30	13	6	11	49—45	45
10	Сочи	30	11	5	14	37—54	38
11	Урал	30	10	6	14	33—45	36
12	Крылья Советов	30	8	8	14	32—45	32
13	Пари НН	30	8	6	16	33—50	30
14	Факел	30	6	12	12	36—48	30
15	Химки	30	4	6	20	25—67	18
16	Торпедо М	30	3	4	23	22—61	13

Чтобы разобраться в этой таблице людям, далёким от спорта, придётся поломать голову. Для них нужно, прежде всего, расшифровать названия столбцов: И — игры, В — выигрыши, Н — ничьи, П — проигрыши, М — забитые и пропущенные мячи, О — набранные очки.

После такой расшифровки многое прояснится. Станет понятно, например, почему в столбце «И» у всех команд стоит число 30: поскольку в турнире участвовало 16 команд и он проводился в два круга, то каждая команда сыграла по 30 игр.

Несложно сообразить также, что за каждую победу в каждой игре давалось 3 очка, за ничью — 1 очко, за поражение — 0 очков. Именно поэтому сумма очков, набранная «Зенитом», равна 70:

$$21 \cdot 3 + 7 \cdot 1 + 2 \cdot 0 = 70.$$

Вообще, в спорте таблицы очень популярны. Именно с их помощью удобнее всего представить итоги турнира, рейтинги спортсменов, расписание матчей предстоящего чемпионата и т. д.

Пример 3. Очень популярны таблицы в экономике. Так называемые ***бухгалтерские книги***, в которых ведётся учёт хозяйственной деятельности, есть на любом предприятии (правда теперь, как и классные журналы, они тоже стали электронными).

Рассмотрим фрагмент таблицы, в которой учитывалось движение товаров в магазине (табл. 3).

Таблица 3

Дата	Артикул	Количество упаковок, шт.	Тип операции
01.06.2021	4	180	Поступление
01.06.2021	13	170	Поступление
01.06.2021	1	170	Поступление
01.06.2021	15	140	Поступление
01.06.2021	1	150	Продажа
01.06.2021	15	80	Продажа
01.06.2021	4	130	Продажа
01.06.2021	13	120	Продажа

Загадочное слово ***артикул*** означает код товара. Без его расшифровки понять, чем торговал магазин невозможно. Для этого нужна вторая таблица, где вместе с артикулом приводятся название товара, его цена и другие данные (табл. 4).

Таблица 4

Артикул	Наименование товара	Единица измерения	Количество в упаковке	Цена, р./упак.
1	Молоко	л	1	68
4	Кефир	л	0,5	54
13	Творог	кг	0,2	120
15	Яйцо диетическое	шт.	10	110

Таблицы 3 и 4, связанные между собой по столбцу «Артикул», образуют то, что в информатике называется ***базой данных***. Сопоставляя эти две таблицы, можно ответить на многие вопросы. Используя количество проданных упаковок каждого товара из таблицы 3 и его цену из таблицы 4, можно узнать, например, дневную выручку магазина.

? ВОПРОСЫ

1. Сколько человек в классе из примера 1 не пропустили ни одного занятия в феврале?
2. Как вы думаете, почему при равенстве очков «Локомотив» стоит в таблице 2 на строчку выше, чем «Динамо»?
3. Сколько упаковок творога было продано в примере 3? Сколько килограммов?
4. Приведите примеры таблиц, с которыми вы сталкивались в повседневной жизни.
5. Приведите примеры таблиц, с которыми вы сталкивались на уроках математики, физики, химии, географии.

2 Диаграммы

Перейдём теперь от таблиц к ***диаграммам***, с которыми вы наверняка сталкивались при изучении других предметов, в учебниках или в Интернете. Само это слово происходит от древнегреческого *diagramma* — рисунок, чертёж.

При использовании диаграмм, как правило, теряется точность представления информации, зато значительно повышается её наглядность. Диаграммы удобно использовать для сравнения величин или чтобы показать изменение какой-то величины во времени.

Диаграммы бывают разных видов: диаграммы-линии, столбиковые диаграммы, круговые, лепестковые и многие другие.

Диаграммы-линии часто называют просто графиками. Их удобно использовать для того, чтобы **показать изменение какой-то величины во времени**. При этом по горизонтальной оси откладываются моменты времени (часы, дни, годы), а по вертикальной — значения интересующей нас величины. Полученные точки соединяются ломаной линией.

Пример 1. На рисунке 1 показано изменение цены на нефть марки Brent в течение июня 2023 г. Каждому рабочему дню июня (в нерабочие дни цена не рассчитывалась) соответствует точка на диаграмме.

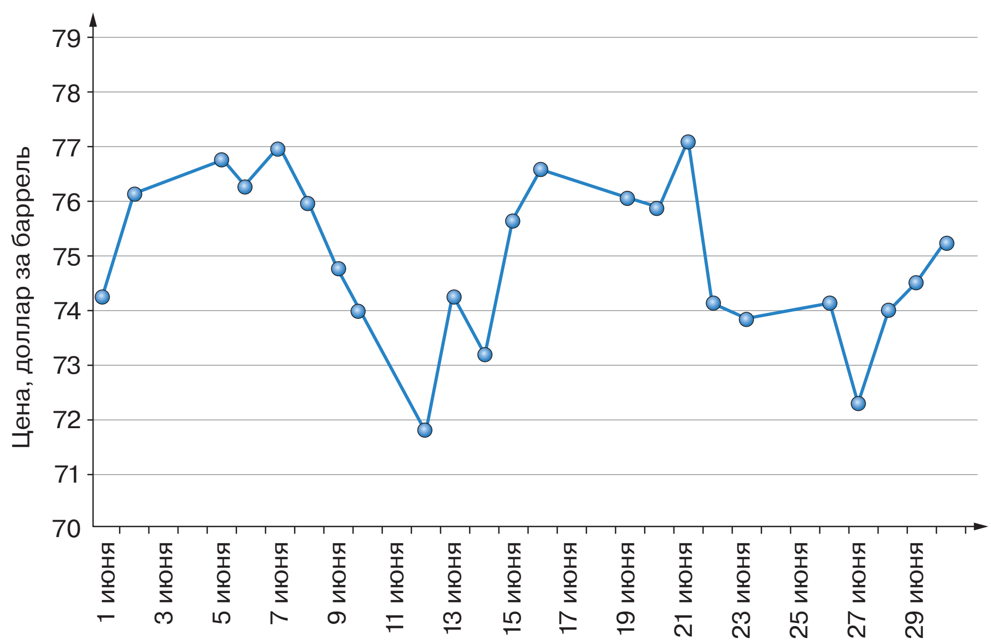


Рис. 1. Цена на нефть марки Brent в июне 2023 г.

Отметим, что в отличие от обычных графиков, которые вы строили на уроках алгебры, отсчёт по горизонтальной и вертикальной осям ведётся чаще всего не от нуля — это касается и диаграммы на рисунке 1.

Следующий популярный тип диаграмм — ***столбиковые*** или ***столбчатые***. Их используют для наглядного **сравнения величин между собой**. Для этого каждую величину изображают в виде столбика, высота которого равна её значению. Все столбики выстраивают в один ряд и подписывают.

Пример 2. На рисунке 2 вы видите столбиковую диаграмму, построенную по данным таблицы 2, в которой приводились итоги чемпионата России по футболу. Высота столбиков показывает число мячей, забитых каждой из команд.

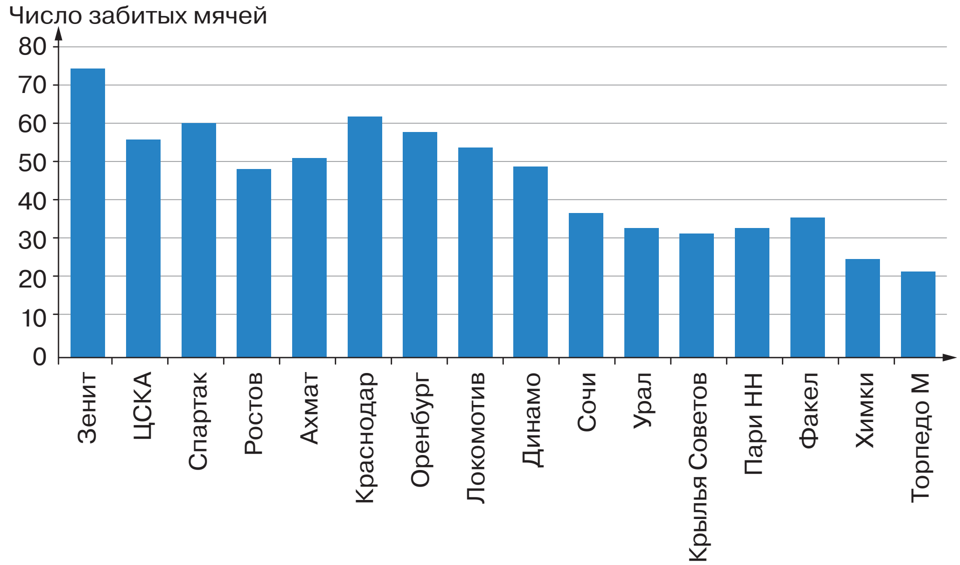


Рис. 2. Число забитых мячей

При построении столбиковых диаграмм можно выбрать любую ширину столбиков и любое расстояние между ними (рис. 3). На диаграмме на рисунке 3, *б* столбики превратились в отрезки. Такие диаграммы иногда называют ***линейчатыми***.

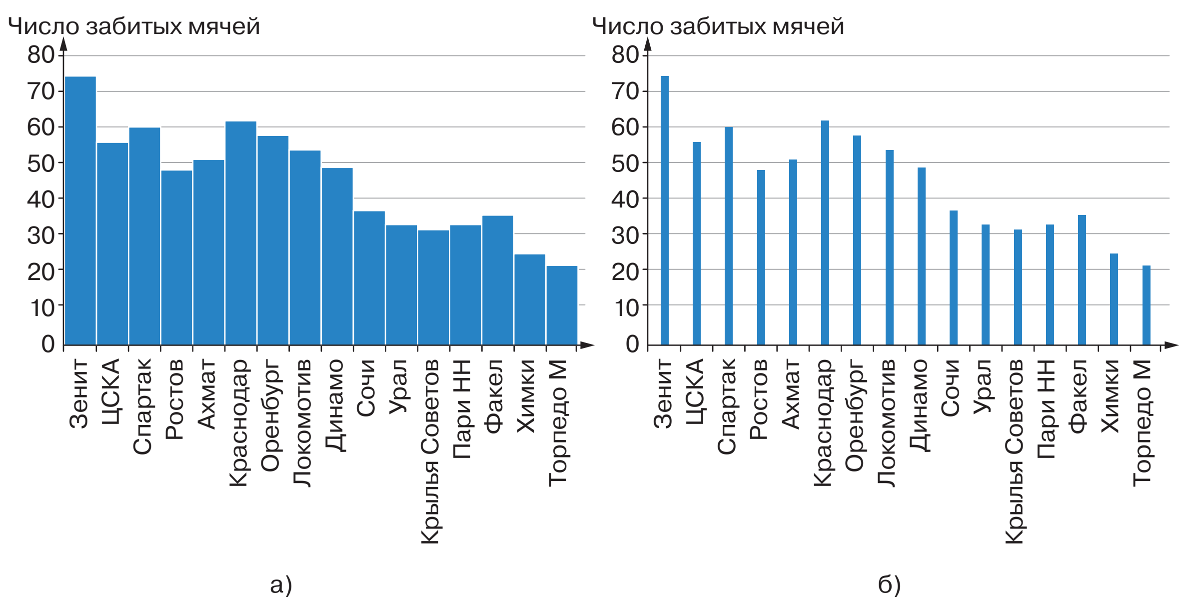


Рис. 3. Варианты оформления столбиковой диаграммы

Иногда на одной диаграмме приходится изображать сразу два числовых набора. В нашем примере интересно сравнить между собой число забитых и число пропущенных разными командами мячей.

На рисунке 4, *а* приведена столбиковая диаграмма, на которой синим цветом показаны столбики, соответствующие числу забитых мячей, а оранжевым цветом — числу пропущенных.

На рисунке 4, *б* те же данные представлены в виде диаграммы-линии. Какой из двух вариантов выбрать, часто определяется по личным предпочтениям. Но если числовых наборов несколько и количество значений в каждом из них велико, то диаграмма-линия будет более наглядной.

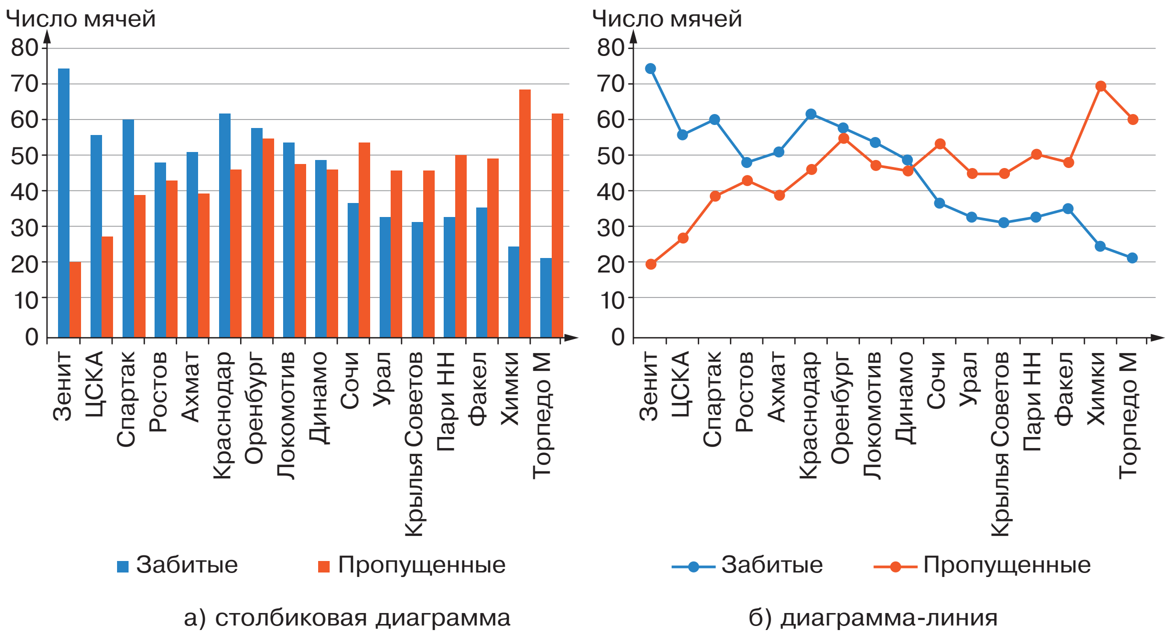


Рис. 4. Число забитых и число пропущенных мячей

Когда на диаграмме нужно показать **соотношение целого и его частей**, то удобнее всего использовать диаграмму, которая называется ***круговой*.** Для построения такой диаграммы рисуют круг и делят его на секторы, угловая величина которых пропорциональна каждому из представленных значений. Визуально такая диаграмма напоминает разрезанный на кусочки пирог. Не случайно в английском языке её называют *pie chart* (*pie* — пирог, *chart* — диаграмма). Иногда эти кусочки раздвигают, чтобы сделать диаграмму нагляднее.

Пример 3. На рисунке 5 с помощью круговой диаграммы представлены результаты контрольной работы, приведённые выше в таблице 1 (попробуйте сами определить, какой именно из трёх контрольных работ соответствует эта диаграмма).

По диаграмме сразу видно, что наибольший процент учащихся написал работу на «4». Бросается в глаза также большая доля отсутствующих учеников.

Заметим, что при использовании круговой диаграммы количество значений в числовом наборе не должно быть слишком большим — иначе диаграмма теряет свою наглядность.

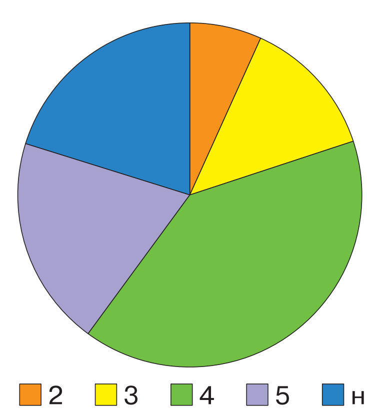


Рис. 5. Результаты контрольной работы

На рисунке 6 показана одна из таких неудачных диаграмм, на которой отражены результаты тестирования по русскому языку, проведённого среди 296 учащихся. Тест содержал 25 вопросов; на диаграмме показано, какая доля учеников правильно ответила на все вопросы, на 24 вопроса, 23 вопроса и т. д. По ней трудно определить, насколько сложным оказался тест для учащихся, поскольку секторов много и они близки по размеру.

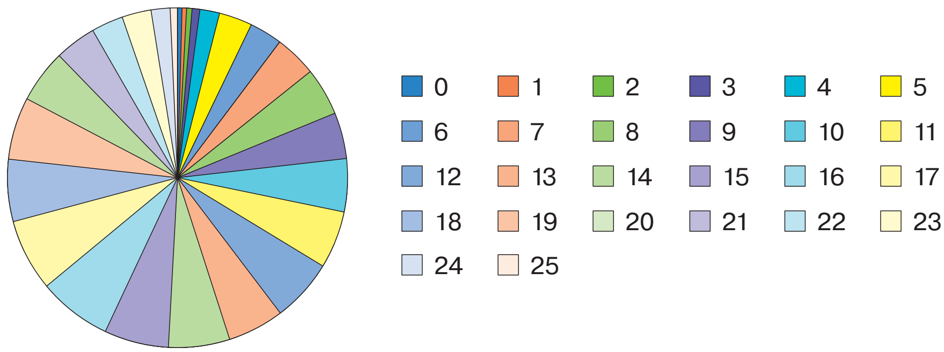


Рис. 6. Неудачная круговая диаграмма

Вот те же результаты, представленные на столбиковой диаграмме (рис. 7). Согласитесь, теперь они выглядят гораздо нагляднее. Сразу видно, что наиболее часто встречались учащиеся, правильно ответившие на 15—17 вопросов, очень немного тех, кто ответил меньше, чем на 5 вопросов, и т. д.

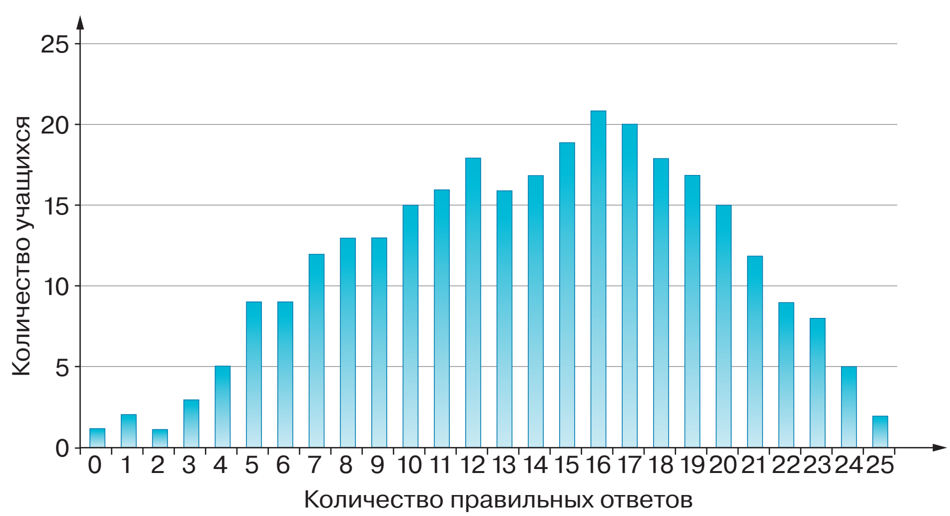


Рис. 7. Результаты тестирования

? ВОПРОСЫ

1. Какую диаграмму лучше использовать, чтобы показать изменение величины во времени?
2. Для чего используют столбиковые диаграммы?
3. Как построить круговую диаграмму? Для чего их используют?
5. По каким причинам может снижаться наглядность круговой диаграммы?
6. Приведите свои примеры таких данных, для которых хорошо подходит диаграмма-линия, столбиковая диаграмма, круговая диаграмма.

3 Таблица частот и полигон

Можно сказать, что первичным результатом любого статистического исследования является достаточно большой массив (набор) чисел. Чтобы увидеть в нём какие-то интересные закономерности и сделать практические выводы, эти данные подвергают специальной обработке, для которой теперь всё чаще используется компьютер. Но прежде чем говорить о компьютерных технологиях, разберёмся, как это делается «вручную», с помощью рассмотренных выше таблиц и диаграмм.

Пример 1. Десятиклассник получил в течение четверти следующие отметки по математике:

5, 2, 4, 5, 3, 2, 3, 2, 4, 3, 4, 5, 4, 3, 4, 5, 5, 5, 4, 5.

В этом числовом наборе не так много чисел — всего 20, — но даже для такого небольшого количества учителю не так просто определить итоговую отметку. Что для этого можно предпринять? Один из вариантов — составить следующую таблицу:

Таблица 5

Отметка	2	3	4	5
Число повторений	3	4	6	7

Мы посчитали, сколько раз встречалась в этом наборе каждая из отметок 2, 3, 4, 5, и записали полученное количество повторений во вторую строку таблицы. Теперь хорошо видно, что самой редкой отметкой была двойка, а самой частой — пятёрка. Чтобы проверить, не ошиблись ли мы при подсчёте, можно сложить все числа во второй строке, должно получиться общее количество выставленных отметок, т. е. 20.

Добавим к этой таблице ещё одну строку — ***частоту***, с которой встречалась каждая отметка. Частоту выражают в виде доли (числа от 0 до 1) или процентного содержания (от 0% до 100%) каждого значения в числовом наборе (табл. 6):

Таблица 6

Отметка	2	3	4	5
Число повторений	3	4	6	7
Частота	0,15	0,20	0,30	0,35

Частота двоек равна 0,15 и получена как отношение количества двоек к общему числу оценок: $\frac{3}{20} = 0{,}15$. Аналогично получены и все остальные частоты. Их сумма, как легко догадаться, всегда равна 1 — это свойство также можно использовать для самопроверки:

$$0{,}15 + 0{,}20 + 0{,}30 + 0{,}35 = 1.$$

Полученная таблица показывает, как частоты распределены между всеми возможными значениями, которые встречались в числовом наборе, поэтому её называют ***распределением частот***. Заметим, что по такой таблице можно восстановить и исходный числовой набор: нужно выписать друг за другом три двойки, четыре тройки, шесть четвёрок и семь пятёрок. Правда, установить, в какой именно последовательности получались отметки, мы уже не сможем. И ещё: если в таблице даны только частоты, а строка с числом повторений каждого значения отсутствует, то для восстановления ряда нужно знать ещё и общее количество значений всего ряда.

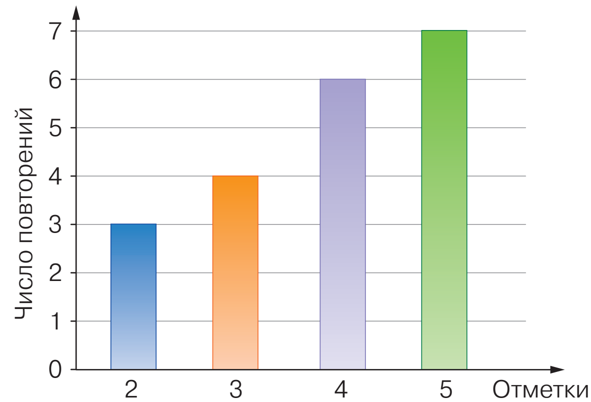


Рис. 8

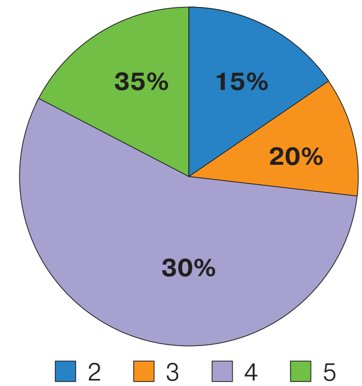


Рис. 9

Распределение частот станет ещё нагляднее, если представить его графически. Это можно сделать с помощью уже знакомых нам столбиковой или круговой диаграмм (рис. 8 и 9).

Однако для распределения частот принято использовать диаграмму-линию, которая в этом случае называется полигоном частот (от латинского *poligon* — многоугольник). Чтобы построить такой полигон, вернёмся к таблице частот.

По горизонтальной оси отметим все числа из первой строки, т. е. полученные нашим учеником отметки. По вертикальной оси укажем для каждой отметки её частоту, записанную в третьей строке. Получили набор из четырёх точек с координатами

$$(2;\ 0{,}15),\ (3;\ 0{,}20),\ (4;\ 0{,}30),\ (5;\ 0{,}35).$$

А теперь соединим их ломаной линией (рис. 10).

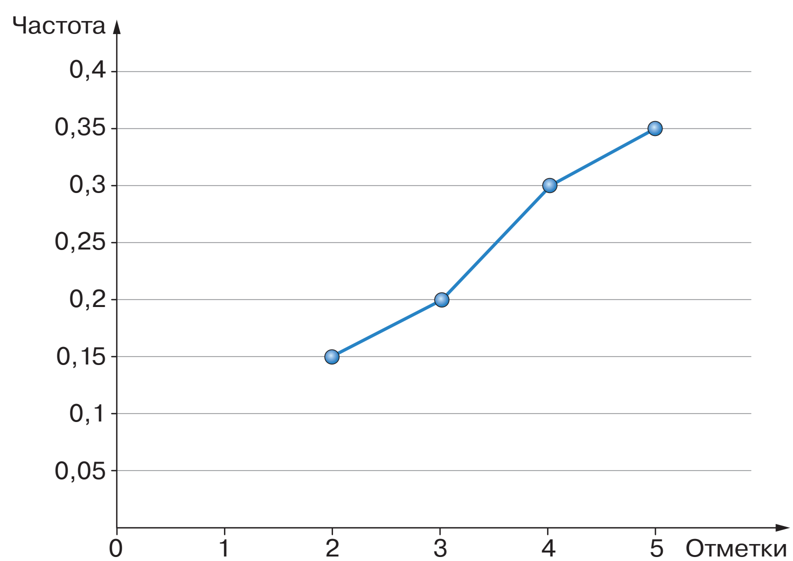


Рис. 10

При сравнении двух и более распределений частот рисуют несколько полигонов в одной системе координат. При этом может оказаться, что удобно отмечать и некоторые значения, которые в одном из рядов отсутствовали (т. е. их частота равнялась 0).

Пример 2. Продолжим анализировать успеваемость десятиклассников и к ученику из примера 1 добавим двух его одноклассников.

Вот отметки второго ученика:

$$3,\ 3,\ 2,\ 3,\ 3,\ 3,\ 2,\ 3,\ 5,\ 3,\ 3,\ 4,\ 4,\ 2,\ 4,\ 3,\ 4,\ 5,\ 3,\ 4,\ 4,\ 5,\ 2.$$

Отметки третьего ученика:

$$2,\ 3,\ 2,\ 4,\ 5,\ 2,\ 3,\ 2,\ 3,\ 4.$$

Построим для каждого из учеников таблицу частот (табл. 7 и 8):

Таблица 7

2-й ученик

Отметка	2	3	4	5
Число повторений	4	10	6	3
Частота	0,17	0,43	0,26	0,13

Таблица 8

3-й ученик

Отметка	2	3	4	5
Число повторений	4	3	2	1
Частота	0,4	0,3	0,2	0,1

Заметим, что каждый из этих учеников получил уже другое количество отметок: у второго их 23, а у третьего только 10. Поэтому при вычислении частот в таблице 7 мы делили каждое число повторений на 23, а в таблице 8 — на 10. Для второго ученика полученные значения частот являются лишь приближёнными: при их записи в виде десятичных дробей мы округлили каждую частоту до 0,01. Заметим, что при этом накопилась погрешность, из-за которой сумма округлённых частот оказалась равна не 1, а 0,99. В статистике такая погрешность вполне допустима.

Теперь изобразим все три распределения в виде полигонов частот (рис. 11).

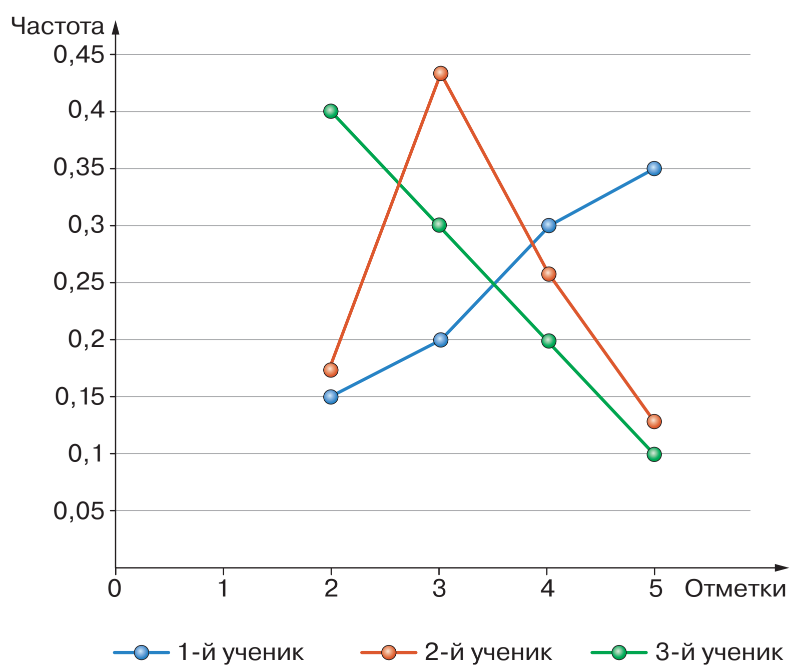


Рис. 11

По построенным графикам хорошо видно различие: отметки первого ученика значительно лучше, чем третьего. Это выражается в том, что частота хороших и отличных отметок у него выше, а плохих (троек и двоек) — ниже. Второй ученик занимает промежуточное положение. Чаще всего у него встречаются тройки и четвёрки, а реже — двойки и пятёрки.

? ВОПРОСЫ

1. Как вычислить частоту данного значения в числовом наборе?
2. В каких границах всегда лежит частота?
3. Чему равна сумма всех частот в распределении?
4. Как построить полигон частот?
5. В каком случае полигон частот превращается в горизонтальный отрезок?

УПРАЖНЕНИЯ

1. По таблице 1 из примера 1 ответьте на следующие вопросы:
а) Сколько учеников получили в феврале только хорошие и отличные отметки?
б) Сколько учеников получили в феврале хотя бы одну двойку?
в) Кто получил наибольшее количество отметок?
г) Сколько выходных дней было в феврале?
д) Кто из учеников, по вашему мнению, лучше всего успевает по математике?

2. Используя итоговую таблицу 2 чемпионата России по футболу, ответьте на следующие вопросы:
а) Сколько всего матчей было сыграно в чемпионате?
б) Сколько матчей закончилось в ничью?
в) Сколько всего голов было забито в чемпионате?
г) Сколько голов забивалось в среднем за одну игру?
д) Какая из команд имеет лучшую разность забитых и пропущенных мячей? худшую разность?

3. По таблицам 3 и 4 из примера 3 ответьте на следующие вопросы:
а) Сколько всего литров кефира было продано?
б) Сколько упаковок яиц осталось в магазине к концу дня?
в) Какую выручку получил магазин от продажи данных товаров за день?
г) Какова общая стоимость всех непроданных в этот день товаров?

4. Представьте, что вы полетите в Сочи рейсом номер 710. Вы прибыли в аэропорт в 15:15. При входе в аэропорт висит табло, на котором даются сведения о вылете самолётов (табл. 9).

Таблица 9

Номер рейса	Пункт назначения	Время вылета	Стойка регистрации	Состояние
212	Казань	15:55	5	Регистрация закончена
357	Сочи	16:10	3	Регистрация закончена
415	Ростов	17:15	7	Идёт регистрация

Окончание

Номер рейса	Пункт назначения	Время вылета	Стойка регистрации	Состояние
512	Мурманск	17:20	1	Вылет задержан до 19:00
140	Сочи	17:30	2	Идёт регистрация
144	Омск	18:00	6	
710	Сочи	18:30	10	

Используя таблицу 9, ответьте на вопросы:
а) Сколько времени осталось до начала регистрации, если регистрация на рейс начинается за 2 ч 30 мин до вылета самолёта?
б) У какой стойки будет проводиться регистрация на этот рейс?
в) Сколько времени осталось до вылета ближайшего рейса в Сочи?
г) На какие рейсы идёт регистрация?

5. Представьте, что вы путешествуете на поезде по маршруту Москва — Волоколамск — Ржев. Ниже даны два расписания движения электропоездов (табл. 10 и 11).

Таблица 10

Москва — Волоколамск

Время отправления	Время прибытия
5:45	7:50
6:32	8:35
8:05	10:00
10:15	12:30
13:20	15:35
15:35	17:40
17:55	20:15
19:00	21:10
21:10	23:25

Таблица 11

Волоколамск — Ржев

Время отправления	Время прибытия
9:30	11:10
11:30	13:10
13:45	15:25
21:00	22:40

а) Сколько времени займёт поездка, если выехать из Москвы в 5:45 и в Волоколамске сесть на ближайшую электричку?
б) Успеете ли вы на последнюю электричку из Волоколамска, если выедете из Москвы в 19:00?
в) Сколько времени вам придётся ждать в Волоколамске, если вы выезжаете из Москвы в 15:35?
г) Подберите два электропоезда — в Москве и в Волоколамске — так, чтобы время на пересадку было наименьшим.

6. В турнирной таблице указаны результаты всех матчей, сыгранных участниками хоккейного турнира (табл. 12). Составьте по ней итоговую таблицу турнира по образцу таблицы 2.
Замечание. Запись 9 : 7 на пересечении строки 1 и столбца 2 означает, что команда «Барсы» выиграла у «Вымпела» со счётом 9 : 7.

Таблица 12

№	Команда	1	2	3	4	5	6	7	8
1	Барсы		9 : 7	6 : 9	3 : 5	2 : 2	6 : 3	9 : 3	5 : 2
2	Вымпел	7 : 9		4 : 4	3 : 6	8 : 1	0 : 2	3 : 3	4 : 1
3	Искра	9 : 6	4 : 4		4 : 1	0 : 7	7 : 3	4 : 1	0 : 9
4	Крылья	5 : 3	6 : 3	1 : 4		1 : 2	7 : 4	3 : 6	6 : 9
5	Ледокол	2 : 2	1 : 8	7 : 0	2 : 1		1 : 8	7 : 2	0 : 4
6	Ракета	3 : 6	2 : 0	3 : 7	4 : 7	8 : 1		2 : 4	2 : 7
7	Ротор	3 : 9	3 : 3	1 : 4	6 : 3	2 : 7	4 : 2		7 : 3
8	Статор	2 : 5	1 : 4	9 : 0	9 : 6	4 : 0	7 : 2	3 : 7	

7. В школе есть возможность организовать занятия по пяти видам спорта. Чтобы определить, какие секции хотели бы посещать десятиклассники, их попросили ответить на вопрос: «Каким видом спорта Вы хотели бы заниматься?» При ответе на вопрос можно было выбрать не более двух из предложенных видов спорта.

Таблица 13

Вид спорта	Класс							
	10А		10Б		10В		10Г	
	Д	М	Д	М	Д	М	Д	М
Баскетбол	2	—	2	4	1	1	2	1
Волейбол	—	3	2	—	1	4	3	2
Лыжи	1	3	—	5	2	4	3	7
Футбол	—	5	—	2	—	4	—	8
Художественная гимнастика	6	—	1	—	2	—	5	—

Ответьте на следующие вопросы:
а) Сколько всего мальчиков выбрали волейбол?
б) Сколько всего учащихся выбрали лыжи?
в) Какой вид спорта оказался наиболее популярным у девочек?
г)* Какое наименьшее число десятиклассников может быть в этой школе?
д)* Какое наименьшее и какое наибольшее число мальчиков могло участвовать в опросе?

Здесь и далее * будет отмечен материал углублённого уровня.

8*. Между населёнными пунктами A, B, C, D, E, F построены дороги, протяжённость которых (в км) указана в таблице 14. Отсутствие числа в таблице означает, что прямой дороги между пунктами нет. Найдите кратчайший путь из пункта A в пункт F.

Таблица 14

	A	B	C	D	E	F
A		2	4			
B	2		1		7	
C	4	1		3	4	
D			3		3	
E		7	4	3		2
F					2	

9*. Многие фирмы при отборе персонала проверяют кандидатов на психологическую совместимость друг с другом. Психологу компании поручено сформировать одну или несколько команд для работы над очередным проектом. Для этого он составил таблицу 15 с именами сотрудников и на пересечении каждой строки и столбца поставил 1, если эти два сотрудника психологически совместимы между собой, и оставил пустую клетку в противном случае.

Таблица 15

	Артём	Виктория	Давид	Дмитрий	Ева	Ксения	Кирилл	Мария	Павел	Софья	Элина	Яна
Артём		1										1
Виктория	1											1
Давид					1		1				1	
Дмитрий						1			1			
Ева			1				1				1	
Ксения				1					1			
Кирилл			1		1						1	
Мария									1	1		
Павел				1		1		1		1		
Софья								1	1			
Элина			1		1		1					
Яна	1	1										

Ответьте с помощью таблицы 15 на следующие вопросы:

а) Какую самую большую команду можно выбрать, чтобы все её участники были психологически совместимы друг с другом?

б) На какое наименьшее число команд можно разбить всех сотрудников, чтобы внутри каждой команды не было психологически несовместимых друг с другом людей?

10*. При регистрации участников ток-шоу «Моя семья» их данные заносятся в таблицу 16, и каждому участнику присваивается уникальный номер ID (*идентификатор*). В таблице 17 указываются их родственные связи.

Таблица 16

ID	Фамилия И. О.	Пол
100	Семашко А. И.	М
101	Казакевич И. Э.	Ж
102	Чиковани Л. А.	Ж
103	Лернер М. Г.	М
104	Семашко Т. А.	Ж
105	Акопян А. М.	М
106	Лернер Г. М.	Ж
107	Лернер Ж. М.	Ж
108	Белых К. Ю.	М
109	Акопян Т. А.	Ж
110	Акопян В. А.	М
111	Акопян Г. А.	Ж
112	Ким А. Г.	Ж
113	Ким В. Г.	М
114	Белых Ю. К.	М
115	Белых Э. К.	Ж

Таблица 17

ID родителя	ID ребёнка
100	104
101	104
105	109
105	110
105	111
104	109
104	110
104	111
102	106
102	107
103	106
103	107
108	114
108	115
107	114
107	115
106	112
106	113

На основании этих данных ответьте на следующие вопросы:
а) Какая фамилия у мамы Г. А. Акопян?
б) Сколько детей у Ж. М. Лернер?
в) Сколько внуков и внучек у Л. А. Чиковани?
г) Какая фамилия у бабушки Э. К. Белых?

11*. *Латинским квадратом* называется таблица $n \times n$, заполненная числами от 1 до n так, что в каждой строке и в каждом столбце стоят все числа от 1 до n. Вот один из латинских квадратов размера 3×3:

1	2	3
3	1	2
2	3	1

Постройте какой-нибудь латинский квадрат размера 4×4.

12*. *Магическим квадратом* называется таблица $n \times n$, заполненная различными числами так, что в каждой строке, каждом столбце и на двух диагоналях сумма чисел одна и та же. Вот один из магических квадратов размера 3×3:

2	7	6
9	5	1
4	3	8

Постройте какой-нибудь магический квадрат размера 4×4.

13. На столбиковой диаграмме (рис. 12) показана рождаемость в городе в течение года.

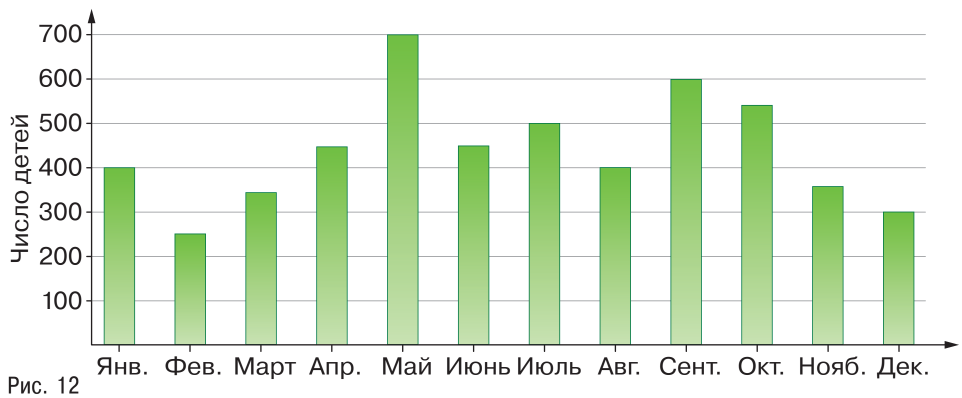


Рис. 12

а) Сколько детей родилось в городе в январе? в мае?
б) В каком месяце родилось 600 детей?
в) В какие месяцы родилось по 400 детей?
г) Сколько детей родилось зимой?
д) В какие месяцы родилось меньше 400 детей? больше 500 детей?

14. В таблице 18 указаны расходы семьи на различные коммунальные услуги (в р.) за первые 6 месяцев года.

Таблица 18

Коммунальные услуги	Январь	Февраль	Март	Апрель	Май	Июнь	Итого
Вода	160	180	220	210	215	170	
Газ	50	50	56	56	56	56	
Электроэнергия	360	414	345	345	310	276	
Телефон	380	380	380	405	405	405	
					Всего		

а) Постройте диаграмму-линию расходов за воду.
б) Постройте столбиковую диаграмму расходов семьи за март.
в) Постройте круговую диаграмму общих расходов семьи за полгода.

15. На круговой диаграмме (рис. 13) изобразили данные о площадях пяти мировых океанов, но забыли подписать секторы. Определите, какой сектор, какому океану соответствует. Найдите необходимые для этого данные в учебнике географии или в Интернете. Для каких секторов определить соответствующий им океан оказалось затруднительно и почему?

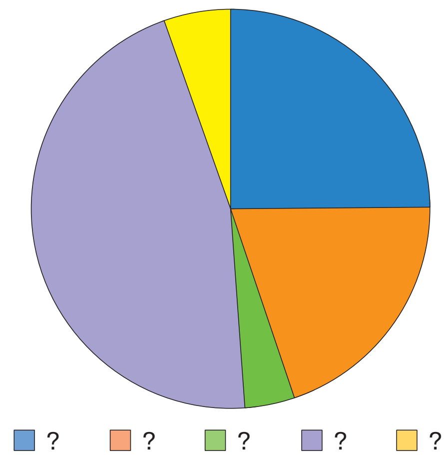


Рис. 13

16. По данным на 1 января 2022 г. вся территория России разделена на 8 федеральных округов. Найдите данные об их площади и населении. Данные о площади изобразите на круговой диаграмме, а о населении на столбиковой.

17. Найдите в Интернете примеры диаграмм каждого из трёх типов: диаграмму-линию, столбиковую и круговую. Объясните, почему для этих данных выбран именно такой тип диаграммы.

18. Вспомните свой вчерашний день (точнее, сутки) и запишите, сколько времени вы провели дома, в школе, на улице, в транспорте, каких-то других местах. Изобразите полученные данные на круговой диаграмме.

19. На последней самостоятельной работе по вероятности и статистике в 10 «А» классе были получены следующие отметки: две двойки, пять троек, восемь четвёрок и три пятёрки. Сколько всего человек писало самостоятельную работу? Составьте по этим данным таблицу частот и нарисуйте полигон частот.

20. В соревнованиях по стрельбе участвовало 12 человек, каждый из которых сделал 10 выстрелов. В таблице 19 указано число попаданий каждого из спортсменов.

Таблица 19

Номер участника	1	2	3	4	5	6	7	8	9	10	11	12
Число попаданий	8	6	7	8	8	5	6	9	8	8	5	9

Постройте по этим данным таблицу частот для числа попаданий. Нарисуйте соответствующий полигон частот.

21. В отделе мужской обуви универмага в течение дня производился учёт размеров купленной обуви. Были получены следующие результаты:

44, 40, 43, 39, 42, 42, 42, 45, 41, 43, 43, 41, 42, 46, 40, 41, 42, 39, 42, 45, 42, 43, 44, 44, 41, 42.

а) Представьте эти результаты в виде таблицы частот (табл. 20), перечертив её в тетрадь. Постройте по таблице полигон частот.

Таблица 20

Размер	39	40	41	42	...
Количество пар					
Частота					

б) Проведите блиц-опрос и заполните аналогичную таблицу с размерами обуви для своего класса. Постройте полигон частот. Сравните его с полигоном из предыдущего пункта. Есть ли между ними какая-то существенная разница? Как её можно объяснить?

22. Завуч школы представила результаты контрольной работы по математике в 10 классах на круговой диаграмме (рис. 14). Опираясь на данные диаграммы, постройте в тетради полигон частот для полученных отметок.

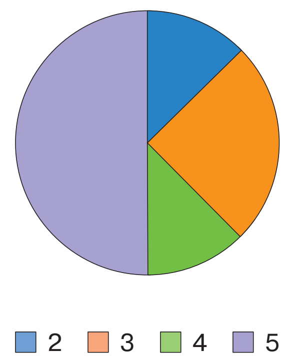


Рис. 14

23. Найдите в Интернете результаты всех матчей XXI чемпионата мира по футболу, который состоялся в 2018 г. в России. Выясните, сколько всего матчей было проведено. Выпишите количество голов, забитых в каждом матче (без учёта послематчевых пенальти). Составьте по этим данным таблицу частот.

24. Возьмите какую-нибудь книгу и откройте её на любой странице.
а) Посчитайте, сколько раз на этой странице встречалась каждая из 33 букв русского алфавита (прописные и строчные буквы различать не нужно).
б) Составьте по полученным данным таблицу частот.
в) Нарисуйте по полученным частотам столбиковую диаграмму.
г) Какая из букв встречалась чаще всего?
д) Сравните свой результат с результатами товарищей.

4 Электронные таблицы

До середины прошлого века основным носителем информации была бумага. Данные о самых разных сферах и явлениях жизни печатались, заносились в таблицы, подшивались в папки и хранились в архивах. Понятно, что многое из этого беспощадно разрушалось временем, а то, что оставалось, становилось всё труднее обрабатывать.

С появлением в XX в. компьютера ситуация начала быстро меняться. Всё чаще стали использоваться электронные носители информации, оказавшиеся намного более долговечными и компактными. Появились программы, которые позволили автоматизировать обработку накопленных данных и сделали их доступными для огромной массы людей. Сегодня никого не удивишь тем, что почти все важные данные хранятся и представляются в электронном виде. Даже традиционный классный журнал с отметками стал сегодня электронным, а основное хранилище информации переместилось в мировую сеть Интернет. Таблицы, с которых мы начали эту главу, тоже стали электронными. О них и пришло время поговорить подробнее.

Электронная таблица — это не цифровая копия обычной таблицы, нарисованной на бумаге. Так называется специальная разновидность компьютерных программ, предназначенных для хранения и автоматической обработки данных, представленных в виде таблицы. В дальнейшем мы будем называть их сокращённо ЭТ.

	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						
7						

Рис. 15

ЭТ содержит, как правило, большое количество строк и столбцов, поэтому на экране компьютера обычно не видно её нижней и правой границы. При этом все строки ЭТ заранее пронумерованы числами: 1, 2, 3, ..., а все столбцы обозначены латинскими буквами: A, B, C, ... (рис. 15). Если букв не хватает, используются двухбуквенные обозначения: AA, AB и т. д. Таким образом, каждая ячейка таблицы имеет свой индивидуальный адрес: B3, C2, D5 (эти ячейки выделены на рисунке 15 цветом).

В ячейках ЭТ могут храниться ***данные*** и ***формулы***. Данные бывают разных типов, чаще всего используются числа (целые и вещественные), текстовые строки и даты (обычно в формате ДД.ММ.ГГГГ, т. е. день, месяц и год, разделённые точками). На рисунке 16 изображена таблица со сведениями о погоде: столбец A содержит даты (первая неделя сентября), столбец B — максимальную температуру в течение суток, столбец C — минимальную (это числа), а в столбце D содержатся сведения о состоянии облачности в виде текстовых строк.

Главное преимущество ЭТ — это использование формул. Чтобы отличить формулу от обычных данных, её ввод в ячейку начинают со знака «=». После этого записывается сама формула, которая состоит из адресов ячеек, арифметических операций и различных функций, которые умеет вычислять электронная таблица. Сразу после ввода формулы в этой ячейке появляется результат вычисления. Если в формуле допущена ошибка, то программа предложит её исправить. Замечательное свойство ЭТ состоит в том, что при изменении любой ячейки с данными все зависящие от неё формулы моментально пересчитываются.

С помощью формулы в столбце E на рисунке 16 посчитана разность максимальной и минимальной температур в течение дня. Для этого в ячейку E1 введена формула

=B1–C1.

	A	B	C	D	E	F
1	01.09.2022	15	8	облачно	7	
2	02.09.2022	18	7	пасмурно	11	
3	03.09.2022	12	8	облачно	4	
4	04.09.2022	22	11	ясно	11	
5	05.09.2022	23	10	ясно	13	
6	06.09.2022	15	6	пасмурно	9	
7	07.09.2022	14	5	облачно	9	
8						

Рис. 16

Поначалу может показаться, что процесс внесения формул очень трудоёмкий, ведь похожую формулу нужно внести ещё в 6 ячеек. Но оказывается, в этом нет необходимости: формулу из ячейки E1 можно скопировать

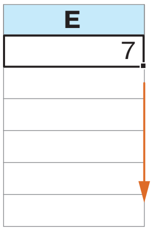


Рис. 17

или «протянуть» в ячейки E2:E7. Этот приём называется *автозаполнением* и заключается в том, что вы наводите курсор мыши на маркер автозаполнения, находящийся в правом нижнем углу активной ячейки, нажимаете на кнопку мыши и «тащите» маркер в нужном направлении по вертикали или горизонтали (рис. 17). При этом в формуле автоматически изменяются все адреса: в ячейке E2 появится формула B2–C2, в ячейке E3 — формула B3–C3 и т. д.

С помощью этого приёма можно ускорить не только процесс ввода формул, но и выполнить автозаполнение ряда последовательных ячеек числами или датами: именно так получены семь последовательных дат в столбце А (см. рис. 16).

ЭТ содержит большое количество встроенных функций, которые используются в вычислениях. Аргументами функций могут выступать как отдельные ячейки, так и целые их диапазоны. Например, чтобы вычислить максимальную разницу дневных и ночных температур в первую неделю сентября, достаточно ввести в любую свободную ячейку нашей таблицы такую формулу:

=МАКС(E1:E7)

МАКС — это название функции «Максимальное значение», а E1:E7 обозначает диапазон из семи ячеек в столбце E.

В электронных таблицах можно не только хранить и вычислять данные, но и строить диаграммы. Столбиковая и круговая диаграммы, о которых мы говорили выше, только маленькая часть тех графических возможностей, которые предоставляет ЭТ для наглядного отображения данных. На рисунке 18 изображена столбиковая диаграмма, построенная в электронной таблице по столбцам A, B и C из нашего примера.

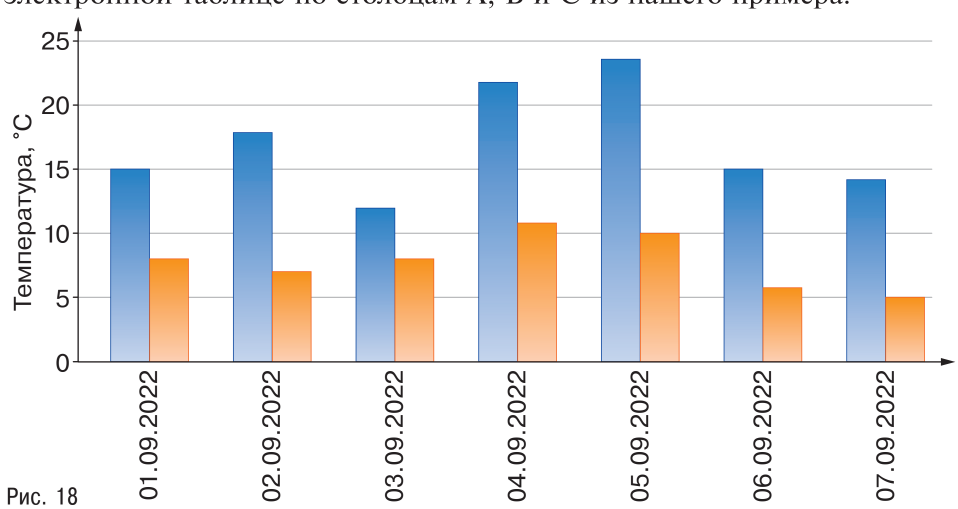


Рис. 18

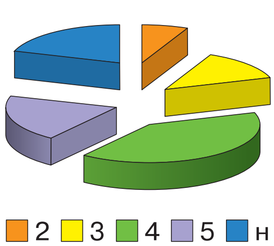


Рис. 19

На рисунке 19 показаны «изобразительные» возможности ЭТ: круговая диаграмма из предыдущего пункта, на которой были представлены результаты контрольной работы по математике, с помощью электронной таблицы действительно превратилась в разрезанный на кусочки пирог.

В электронной таблице частоту каждого значения числового набора можно получить автоматически. Для этого используют специальную функцию

СЧЁТЕСЛИ(диапазон; значение),

с помощью которой можно посчитать, сколько раз данное значение встречается в данном диапазоне. Покажем, как это делается с отметками первого ученика из примера 1 пункта 3.

Прежде всего, нужно внести сами отметки в таблицу, например записать их в столбец A (рис. 20). Запишем в столбец C все различные значения, которые встречались в этом наборе (2, 3, 4, 5), и посчитаем в столбце D число повторений каждого из них. Для этого введём в ячейку D2 формулу

=СЧЁТЕСЛИ(A:A;C2)

и протянем её на ячейки D3:D5. Результат можно увидеть на экране.

Наконец, последнее, что следует сказать об электронных таблицах: в любой момент все данные, включая формулы, графики и всё остальное, можно сохранить на электронном носителе или в облачном хранилище данных. А если к вашему компьютеру подключен принтер, то и распечатать всю таблицу или какую-то её часть.

В дальнейшем мы будем не раз возвращаться к электронным таблицам, указывая их средства и функции, которые оказываются полезными при решении тех или иных задач. ЭТ помогут вам и при выполнении лабораторных работ, которые будут сопровождать весь наш курс в дальнейшем.

Для проверки полученных чисел можно найти в ячейке D6 их сумму (функция СУММ()) — она должна равняться 20. Остаётся поделить каждое число повторений на 20. Для этого введём в ячейку E2 формулу

=D2/D$6

и протянем её вниз. Все частоты получены. Их сумма, конечно, равна 1.

Отметим один важный момент. При указании адреса ячейки D6 мы поставили в её адрес специальный знак «$». С его помощью этот адрес можно «закрепить»: теперь он не будет изменяться при копировании этой формулы вниз.

Чтобы построить полигон частот, воспользуемся специальным типом диаграммы, которая называется «Точечная диаграмма». Она рисует на координатной плоскости набор точек с указанными координатами $(x_1; y_1)$, ..., $(x_n; y_n)$, которые для наглядности можно соединить ломаной линией. Перед построением диаграммы выделим диапазон C2:C5 (это будут значения x_i) и диапазон E2:E5 (это значения y_i). После этого нажмём кнопку с изображением соответствующей диаграммы, и полигон частот готов (рисунок 21 был получен именно таким образом).

	A	B	C	D	E
1	Отметки ученика		Отметка	Число повторений	Частота
2	5		2	3	0,15
3	2		3	4	0,2
4	4		4	6	0,3
5	5		5	7	0,35
6	3			20	1
7	2				
8	....				

Рис. 20

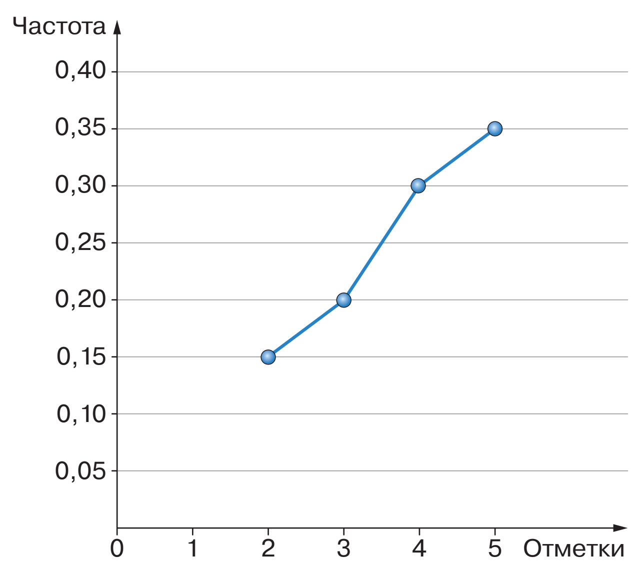


Рис. 21

Конечно, пока числовых данных немного, может показаться, что обработать их вручную даже проще. Но представьте себе, что такое же распределение частот нужно получить для отметок за весь учебный год, и вы сразу почувствуете преимущества компьютерной обработки данных.

? ВОПРОСЫ

1. В чём главное отличие электронной таблицы от бумажной?
2. Сколько ячеек содержат диапазоны A2:A9, B3:H3, B2:F5?
3. Если формулу A3+C4 скопировать на одну ячейку вниз, как она изменится? А если на одну ячейку вправо?
4. Какая функция используется для того, чтобы посчитать количество повторений заданного значения в наборе данных? Какие у неё аргументы?
5. Что означает знак «$» при указании адреса в электронной таблице?
6. Какой тип диаграммы используется для построения полигона частот?

ЛАБОРАТОРНАЯ РАБОТА № 1

Электронные таблицы и диграммы

Задание 1. Использование функций МИН(), МАКС(), СУММ()

Запишите в ячейки столбца A даты последних семи прошедших дней.
Занесите в столбец B время, которое в этот день вы потратили на разговоры по телефону.
С помощью функции СУММ() найдите в ячейке B8 суммарное время, потраченное на разговоры.
С помощью функций МИН() и МАКС() найдите минимальное и максимальное значения в диапазоне B1:B8.
Отразите данные о потраченном времени на столбиковой диаграмме; на круговой диаграмме.

Задание 2. Анализ вашей успеваемости

Выпишите в столбец A все ваши отметки из дневника за последний месяц.
Запишите в столбец C числа 2, 3, 4, 5.
Найдите в столбце D количество повторений каждого из этих чисел в столбце C. Используйте для этого функцию СЧЁТЕСЛИ().
Посчитайте в столбце E частоту каждой отметки.
Постройте по этим данным полигон частот и сделайте выводы о вашей общей успеваемости.

Задание 3*. Последняя цифра полного квадрата

Запишите в столбец A натуральные числа от 1 до 100. Для этого впишите в первые две ячейки числа 1 и 2, а затем «протяните» их вниз.
Вычислите в столбце B квадраты этих чисел: введите в ячейку B1 формулу =A1^2 и «протяните» её вниз.
Найдите в столбце C последнюю цифру каждого квадрата. Используйте для этого функцию ОСТАТ(A;B), которая вычисляет остаток от деления A на B.
Теперь с помощью функции СЧЁТЕСЛИ() определите, с какой частотой в столбце C встречается каждая цифра от 0 до 9. По полученным данным постройте полигон частот. Можно ли было заранее предсказать полученные результаты?

§ 2. Описательная статистика

Поскольку статистика всегда имеет дело с большими массивами данных, то одна из главных её задач — удобное представление и описание этих данных, замена большого массива наглядной таблицей, диаграммой, одним или несколькими числовыми показателями. Отсюда и название вводного раздела в эту науку — «Описательная статистика».

С таблицами и диаграммами мы уже познакомились. Перейдём теперь к числовым показателям, которые могут многое рассказать о тех данных, по которым они посчитаны.

1 Мода

Вернёмся к примерам 1 и 2 пункта 3 «Таблица частот и полигон», в которых речь шла об отметках десятиклассников по математике? Напомним их отметки:

1-й ученик: 5, 2, 4, 5, 3, 2, 3, 2, 4, 3, 4, 5, 4, 3, 4, 5, 5, 5, 4, 5;
2-й ученик: 3, 3, 2, 3, 3, 3, 2, 3, 5, 3, 3, 4, 4, 2, 4, 3, 4, 5, 3, 4, 4, 5, 2;
3-й ученик: 2, 3, 2, 4, 5, 2, 3, 2, 3, 4.

Сведём эти данные в одну общую таблицу частот (табл. 21):

Таблица 21

Отметка	2	3	4	5
Частота для 1-го ученика	0,15	0,2	0,3	0,35
Частота для 2-го ученика	0,17	0,43	0,26	0,13
Частота для 3-го ученика	0,4	0,3	0,2	0,1

Какую же итоговую оценку должен поставить учитель каждому из учеников? Если смотреть на **самую частую** отметку, то для первого ученика это пятёрка, для второго — тройка, а для третьего — двойка.

Число, которое встречается в числовом наборе наиболее часто, называют его *модой*.

Итак, в качестве итоговой оценки можно взять моду. Однако справедливо ли будет не учитывать другие отметки? Ведь двоек у третьего ученика хотя и больше всего, но только 40% от общего количества отметок. А пятёрок у первого и того меньше — всего 35%.

И потом, если мы будем выбирать в качестве итоговой оценки моду, то рискуем попасть в безвыходное положение: ведь вполне может быть так, что две (или даже три) самые популярные отметки будут встречаться одинаково часто. В этом случае говорят, что у числового набора **нет моды**. Таким образом, выбор моды в качестве критерия для выставления итоговой оценки далеко не лучший вариант.

Тем не менее мода довольно часто используется в статистических опросах, таких, например, как выборы.

Это объясняется двумя причинами:

- во-первых, моду можно найти даже для ***не*** числового набора данных (например, можно определить, за какого кандидата чаще всего голосовали избиратели);
- во-вторых, если набор данных большой по объёму, а различных значений немного, то совпадение частот у двух кандидатов очень маловероятно.

? ВОПРОСЫ

1. Что называют модой ряда чисел?
2. Найдите моду ряда чисел:

   а) 5, 5, 5, 5, 5, 5, 5; б) 4, 3, 4, 3, 4; в) –1, 0, 1, 0, –2.

   Приведите пример ряда, у которого нет моды.
3. Витя получил за четверть 7 четвёрок и 3 пятёрки. Чему равна мода его отметок?
4. Проведите блицопрос и узнайте, какой школьный предмет нравится учащимся вашего класса больше всего, определив моду полученного ряда.

2 Среднее арифметическое

Самый популярный способ определить итоговую оценку — вычислить ***среднее арифметическое*** полученных отметок.

Средним арифметическим набора чисел называется частное от деления суммы этих чисел на их количество.

Для набора отметок первого ученика из предыдущего пункта это будет выглядеть так:

$$\frac{5+2+4+5+3+2+3+2+4+3+4+5+4+3+4+5+5+5+4+5}{20} = \frac{77}{20} = 3{,}85.$$

Как видите, результат оказался гораздо меньше моды. И выглядит он справедливее, поскольку учитывает ***каждую*** отметку, полученную учеником. Если мы изменим только одну из отметок всего на один балл, среднее арифметическое обязательно изменится. Правда, если оценок много, то не существенно. Заменим, например, в нашем числовом наборе одну из троек на четвёрку. Легко сообразить, что среднее арифметическое при этом увеличится на $\frac{1}{20} = 0{,}05$ и станет равно 3,9.

В общем случае, если ряд содержит n чисел x_1, x_2, ..., x_n, его среднее арифметическое обозначается $\overline{x}$ и находится по формуле:

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}.$$

Проверьте, что итоговые оценки для второго и третьего учеников, посчитанные по этой формуле, будут равны соответственно 3,35 и 3. Третий ученик, у которого мода отметок равна 2, спасён. По этому принципу выставлять итоговую оценку оказалось гораздо справедливее (по крайней мере, с точки зрения ученика). Заметим, что среднее арифметическое для второго ученика нам тоже пришлось округлить до второго знака после запятой (точное значение равно $\frac{77}{23}$).

Если для набора чисел построена таблица частот, то вычисление среднего арифметического можно упростить: вместо того чтобы складывать много раз одно и то же значение, можно умножить его на количество повторений. В примере с отметками первого ученика это будет выглядеть так:

$$\frac{2 \cdot 3 + 3 \cdot 4 + 4 \cdot 6 + 5 \cdot 7}{20} = \frac{77}{20} = 3{,}85.$$

Наконец, можно обойтись и без деления на 20, если использовать не количество повторений, а частоту каждого значения:

$$2 \cdot 0{,}15 + 3 \cdot 0{,}20 + 4 \cdot 0{,}30 + 5 \cdot 0{,}35 = 3{,}85.$$

Правда, последняя формула, хотя и содержит наименьшее количество арифметических операций, удобна скорее для компьютера, а не для ручного подсчёта.

? ВОПРОСЫ

1. Что называют средним арифметическим ряда чисел?
2. Найдите среднее арифметическое ряда чисел:
   а) 5, 5, 5, 5, 5, 5, 5; б) 4, 3, 4, 3; в) 4, 3, 4, 3, 4.
3. Витя получил за четверть 7 четвёрок и 3 пятёрки. Чему равно среднее арифметическое его отметок?

3 Медиана

Ещё одна числовая характеристика, подходящая для вычисления итоговой оценки, *медиана* (от латинского *mediana* — середина). Медиана — это действительно то число, которое окажется ровно посередине числового набора, **если записать все его числа по возрастанию**.

Правда, в случае, когда количество чисел в наборе чётное (как в примере с отметками), посередине окажутся сразу два числа: десятое (перед ним девять чисел) и одиннадцатое (после него тоже девять чисел). Если их значения совпадают, то это число и будет являться медианой. А если они различны? В этом случае медианой может служить любое значение между этими двумя числами. Обычно берут их среднее арифметическое. Итак, мы пришли к следующему алгоритму вычисления медианы.

Чтобы найти медиану произвольного ряда чисел, их нужно записать по возрастанию и воспользоваться следующим правилом:
- **если ряд состоит из нечётного количества чисел, то медианой будет число, которое находится посередине;**
- **если ряд состоит из чётного количества чисел, то медианой будет среднее арифметическое двух чисел, находящихся посередине.**

Посчитаем медиану отметок для каждого ученика из нашего примера:

1-й ученик: 5, 2, 4, 5, 3, 2, 3, 2, 4, 3, 4, 5, 4, 3, 4, 5, 5, 5, 4, 5;

2-й ученик: 3, 3, 2, 3, 3, 3, 2, 3, 5, 3, 3, 4, 4, 2, 4, 3, 4, 5, 3, 4, 4, 5, 2;

3-й ученик: 2, 3, 2, 4, 5, 2, 3, 2, 3, 4.

Выпишем все отметки первого ученика по возрастанию:

$$2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ \boxed{4,\ 4,}\ 4,\ 4,\ 5,\ 5,\ 5,\ 5,\ 5,\ 5,\ 5.$$

Посередине на 10 и 11 местах стоят две четвёрки: значит, медианой будет число 4. Для второго ученика упорядоченный ряд будет выглядеть так:

$$2,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ \boxed{3,}\ 3,\ 3,\ 4,\ 4,\ 4,\ 4,\ 4,\ 4,\ 5,\ 5,\ 5,$$

а для третьего ученика —

$$2,\ 2,\ 2,\ 2,\ \boxed{3,\ 3,}\ 3,\ 4,\ 4,\ 5.$$

Как видим, с точки зрения медианной оценки успеваемость второго и третьего учеников оказалась одинаковой — медиана и там, и там равна 3.

Сложнее найти медиану, если вам дан не числовой набор, а построенная по нему таблица частот. Конечно, можно сначала выписать все числа по порядку в нужном количестве, а потом найти по ним медиану. Но можно действовать рациональнее: достаточно двигаться по возрастающим значениям ряда, записанным в таблице, и суммировать количество повторений, пока их не станет больше половины (или суммировать частоты, пока сумма не превысит 0,5). То значение, на котором это произошло, и будет медианой.

Для этого можно добавить к нашей таблице ещё одну строку — строку ***накопленных*** (т. е. просуммированных) частот. Вот как такая строка будет выглядеть для первого ученика (табл. 22):

Таблица 22

Отметка	2	3	4	5
Частота	0,15	0,20	0,30	0,35
Накопленная частота	0,15	0,35	0,65	1,00

«Перевал» накопленной частоты через 0,5 произошёл на отметке «4», значит, это число и будет медианой.

Подытоживая, можно дать следующее определение медианы.

Медианой числового набора называется такое число m, что хотя бы 50% исходных чисел $\leqslant m$, и хотя бы 50% исходных чисел $\geqslant m$.

Казалось бы, можно сказать проще: 50% исходных чисел $< m$ и 50% исходных чисел $> m$. Но это будет справедливо только при чётном количестве чисел. Скажем, в наборе

$$1,\ 2,\ 3,\ 4$$

медиана $m = 2{,}5$ и при этом ровно 50% исходных чисел меньше m, и ровно 50% — больше. Возьмём теперь нечётное количество чисел, например:

$$1,\ 2,\ 3,\ 4,\ 5.$$

Медиана $m = 3$. Строго меньше m здесь два числа из пяти (т. е. меньше 50%), и строго больше m — также два. Меньше либо равных медиане чисел — три из пяти (т. е. больше 50%), и столько же чисел больше либо равных.

А вот условие, которое указано в определении, выполняется как для чётного, так и для нечётного количества чисел (проверьте это самостоятельно).

Все три рассмотренные величины называют *средними характеристиками* числового набора. Конечно, они не несут всей полноты информации, полученной в статистическом исследовании (ведь целый массив данных характеризуется всего одним числом!), но содержат ключевую информацию, на основе которой часто принимаются важные решения (выставляется итоговая оценка, определяется победитель соревнований, выбирается президент и т. д.).

В электронной таблице для вычисления каждой из средних характеристик есть специальная функция:

МОДА(диапазон) — мода;

СРЗНАЧ(диапазон) — среднее арифметическое;

МЕДИАНА(диапазон) — медиана.

В качестве аргумента для каждой из этих функций указывается диапазон ячеек, в котором содержится числовой набор.

? ВОПРОСЫ

1. Что называют медианой ряда чисел?
2. Найдите медиану ряда чисел:
   а) 5, 5, 5, 5, 5, 5, 5; б) 3, 4, 4, 3; в) 3, 4, 3, 4, 3.
3. Витя получил за четверть 7 четвёрок и 3 пятёрки. Чему равна медиана его отметок?

4 Средние характеристики: какая лучше?*

Получается, что три разных подхода к выставлению итоговой оценки одних и тех же учеников дают разные результаты. Представим их в виде таблицы (табл. 23):

Таблица 23

Средние характеристики	1-й ученик	2-й ученик	3-й ученик
Мода	5	3	2
Среднее арифметическое	3,85	3,35	3
Медиана	4	3	3

Правда, среднее арифметическое всё равно придётся округлить, и тогда, скорее всего, учитель выставит первому ученику 4, а второму и третьему — 3.

Какая же из этих характеристик лучше? Ответить на этот вопрос невозможно, так как выбор подходящей средней характеристики зависит от конкретной ситуации. Рассмотрим достоинства и недостатки средних характеристик.

МОДА

Мы уже говорили о том, что моды у числового ряда может не быть вовсе. Кроме того, даже когда она есть, её значение может быть малоинформативным, давать искажённое представление о поведении числового набора.

Например, для последовательности чисел

$$1,\ 1,\ 3,\ 4,\ 6,\ 8,\ 9,\ 12$$

мода равна 1, но вряд ли это можно считать какой-то важной характеристикой.

В то же время, когда из всего набора данных (не обязательно числового) нужно по условию задачи выбрать ровно одно значение, например определить лучшую книгу года или выбрать победителя в каком-то конкурсе, мода оказывается самой подходящей характеристикой.

СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ

Среднее арифметическое — наиболее популярная из этих характеристик. Средний балл по какому-то предмету, среднесуточная температура, средняя выработка на производстве — всё это примеры величин, полученных по формуле среднего арифметического.

Для приведённого выше набора чисел среднее арифметическое равно

$$\frac{1+1+3+4+6+8+9+12}{8}=5{,}5,$$

и это гораздо больше согласуется с интуитивным понятием среднего.

Среднее арифметическое имеет важный физический смысл: в этой точке находится «центр тяжести» всего числового набора. Поясним, что это значит. Отметим на числовой оси точки из нашего набора данных и «подвесим» в каждой из этих точек одинаковые грузики (если одно и то же число повторяется в наборе несколько раз, то подвесим в этой точке столько грузиков, сколько раз оно повторяется). Если теперь вычислить среднее арифметическое и поставить в этой точке опору (рис. 22), то вся система будет находиться в равновесии.

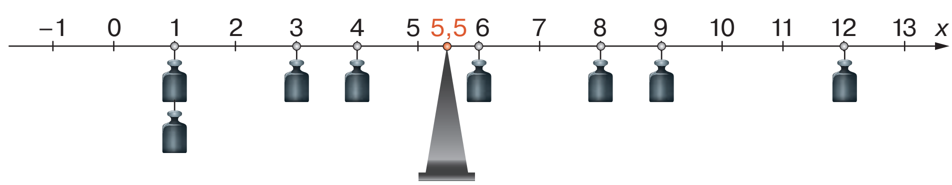


Рис. 22

Но у среднего арифметического есть и свои недостатки. Оно чувствительно к так называемым ***выбросам*** — слишком маленьким или слишком большим значениям числового набора. Например, среднее арифметическое ряда чисел

$$10,\ 20,\ 20,\ 30,\ 30,\ 30,\ 30,\ 30,\ 30,\ 900$$

равно 110 (проверьте), но при этом девять из десяти чисел этого набора меньше этого среднего и только одно — больше.

Этот недостаток среднего арифметического зачастую может искажать весьма важную информацию. Самый известный пример такого рода — **средняя зарплата**. Представим себе, что приведённые выше числа — это зарплата сотрудников какого-то предприятия (выраженная, например, в тысячах рублей). Тогда фраза директора (который получает 900 000 р.) «на нашем предприятии средняя зарплата превышает 100 000 р.» будет чистой правдой, но кому такая «правда» нужна? Ведь *все* его сотрудники получают зарплату, которая намного меньше 100 000 р.!

Не слишком объективным будет и выставление итоговой оценки ученику по среднему арифметическому: если в начале четверти он получал пятёрки, а потом «скатился» на двойки и тройки, то среднее арифметическое вполне может быть близким к четвёрке.

Например, для отметок

$$5,\ 5,\ 5,\ 5,\ 5,\ 3,\ 2,\ 2,\ 3,\ 2$$

среднее арифметическое равно числу 3,7, которое округляется до 4. Но вряд ли этот ученик достоин такой отметки: к моменту подведения итогов четверти он в лучшем случае может считаться «троечником».

МЕДИАНА

Медиана, в отличие от среднего арифметического, не чувствительна к выбросам. Если расположить все значения набора данных на числовой оси, а потом двигать крайнюю правую точку вправо, то среднее арифметическое (центр тяжести) тоже будет смещаться вправо, а вот медиана будет стоять на месте! То же самое будет, если двигать минимальное из чисел влево.

Такое свойство называют ***устойчивостью медианы***. В примере с зарплатами медиана будет равна 30 000 р., и это значение даёт гораздо более объективное представление о материальном положении сотрудников на данном предприятии. Неслучайно в современных статистических сводках и отчётах средний доход и средняя зарплата всё чаще уступают место медианным показателям.

Если директору предприятия захочется сказать «на нашем предприятии медианная зарплата превышает 100 000 р.», то ему придётся повысить до этого уровня заработную плату хотя бы половины своих сотрудников.

Так, может быть, медиана и есть идеальный средний показатель? К сожалению, это не так. Как часто бывает, достоинства медианы являются одновременно и её слабым местом. Обратимся к числовой оси. Если как угодно двигать точки слева от медианы, не перескакивая через неё, то она не изменится. То же самое для точек справа. Так ли это хорошо? Пока речь шла о крайних значениях, устойчивость играла положительную роль, но когда медиана не реагирует и на другие изменения, описанные выше, это становится недостатком.

Перед вами два числовых набора:

$$2,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3;$$

$$3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 5,\ 5,\ 5,\ 5.$$

Медианы обоих наборов равны 3. Но если рассматривать эти наборы, как оценки двух учеников, то у второго они явно лучше. Об этом говорят и средние арифметические, равные соответственно 2,6 и 3,8. Очевидно, первый ученик еле дотягивает до тройки, а второй, возможно, достоин четвёрки.

Таким образом, каждая из трёх рассмотренных средних характеристик — мода, среднее арифметическое и медиана — имеет свои достоинства и недостатки. Зная их, можно выбрать в каждой конкретной задаче наиболее информативную или использовать несколько из них.

? ВОПРОСЫ

1. Приведите пример ряда, у которого мода, медиана и среднее арифметическое совпадают.
2. Приведите пример ряда, у которого все три характеристики разные.
3. Какую среднюю характеристику вы бы использовали, чтобы определить лучшего ученика по математике в своём классе?

5 Среднее гармоническое

Есть и другие способы понимать среднее заданного набора чисел. В некоторых задачах удобно использовать ***среднее гармоническое***.

Средним гармоническим набора из n положительных чисел $x_1, \ldots, x_n$ называется такое число H, для которого выполняется равенство:

$$\frac{n}{H} = \frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}.$$

Отсюда можно получить явную формулу для среднего гармонического:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}.$$

Например, средним гармоническим двух чисел a, b будет:

$$H = \frac{2}{\frac{1}{a} + \frac{1}{b}} = \frac{2ab}{a + b},$$

а средним гармоническим трёх чисел a, b, c будет:

$$H = \frac{3}{\frac{1}{a} + \frac{1}{b} + \frac{1}{c}} = \frac{3abc}{ab + bc + ac}.$$

Такое определение среднего может показаться непривычным, но иногда оказывается полезным.

Пример 1. Автобус проехал расстояние от Калуги до Москвы со средней скоростью 50 км/ч, а возвратился со средней скоростью 70 км/ч. Какова была средняя скорость автобуса на всём пути туда и обратно?

Может показаться, что это будет среднее арифметическое двух данных в задаче скоростей, но это не так. Обозначим неизвестное нам расстояние между городами через s. Тогда время в пути составит:

$$t = \frac{s}{50} + \frac{s}{70}.$$

Чтобы найти среднюю скорость, нужно весь пройденный путь поделить на время:

$$\overline{v} = \frac{2s}{\frac{s}{50} + \frac{s}{70}} = \frac{2}{\frac{1}{50} + \frac{1}{70}} = \frac{2 \cdot 50 \cdot 70}{50 + 70} = \frac{175}{3} \approx 58{,}33.$$

Средняя скорость оказалась средним гармоническим двух данных скоростей.

Термин «гармоническое» происходит из музыки. Для струны музыкального инструмента частоты основного тона и его обертонов представляют собой последовательность аликвотных дробей: $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$.

А в этой последовательности, как легко проверить, каждая дробь является средним гармоническим двух соседних.

Среднее гармоническое обладает некоторыми свойствами среднего:

- среднее гармоническое набора из одинаковых чисел равно этому числу;
- среднее гармоническое всегда лежит в промежутке между наименьшим и наибольшим числами набора.

Ещё одним свойством среднего гармонического является то, что оно всегда меньше или равно среднему арифметическому. Причём равенство достигается только в том случае, когда все числа в наборе одинаковые.

ВОПРОСЫ

1. Что называют средним гармоническим ряда чисел?
2. Какие свойства среднего гармонического вы знаете?
3. Найдите среднее гармоническое ряда чисел:

   а) 2, 2, 2; б) $\frac{1}{2}, \frac{1}{4}$; в) 2, 4.

4. Автомобиль ехал первую половину пути со скоростью 60 км/ч, а вторую — со скоростью 40 км/ч. Чему равна его средняя скорость?

УПРАЖНЕНИЯ

25. Найдите моду ряда чисел:
а) 13; 15; 13; 12; 12; 12; 13; 14; 13; 15; 13; б) 39; 54; 33; 36; 20; 29; 35; 50; 21.

26. Найдите среднее арифметическое ряда чисел:
а) 31; 36; 69; 24; 20; 48; б) 1,6; 4,9; 12,4; 3,1.

27. Найдите медиану ряда чисел:
а) 12; 16; 19; 25; 30; 32; 33; 38; 40;
б) 15; 17; 19; 21; 23; 25; 27; 29; 31; 33;
в) 12; 8; 7; 14; 25; 6; 19; 18;
г) 2,5; 1,3; 1,5; 0,9; 1,7; 2,1.

28. Измерив рост (в см) 12 футболистов, получили следующие данные:

178, 169, 191, 182, 171, 173, 174, 180, 179, 164, 178, 185.

Найдите средний рост футболистов (среднее арифметическое) и число футболистов выше среднего роста. Найдите медиану и число футболистов, рост которых больше медианы.

29. Найдите моду по таблице частот:
а)

Таблица 24

Значения	0	1	2	3	4
Частота	0,1	0,1	0,3	0,4	0,1

б)

Таблица 25

Значения	–3	3	5	8
Частота	0,6	0,3	0,05	0,05

30. Найдите среднее арифметическое по таблице частот:
а)

Таблица 26

Значения	0	10	20	30	100
Частота	0,5	0,1	0,1	0,1	0,2

б)

Таблица 27

Значения	1	2	3	4
Частота	0,1	0,2	0,3	0,4

31. Найдите медиану по таблице частот:

а)

Таблица 28

Значения	0	10	20	30	100
Частота	0,2	0,4	0,2	0,1	0,1

б)

Таблица 29

Значения	1	2	3	4
Частота	0,1	0,4	0,3	0,2

в)

Таблица 30

Значения	–3	3	5	8
Частота	0,6	0,3	0,05	0,05

32. В таблице 31 представлены результаты контрольной работы по математике в 10 классе. Найдите все средние характеристики полученного ряда отметок. Сравните их между собой.

Таблица 31

Отметка	2	3	4	5
Число учеников	4	8	12	6

33. Среди учеников 10 «А» класса провели опрос, сколько примерно часов в день они тратят на выполнение домашних заданий. Ответы школьников представлены на диаграмме, изображённой на рисунке 23.

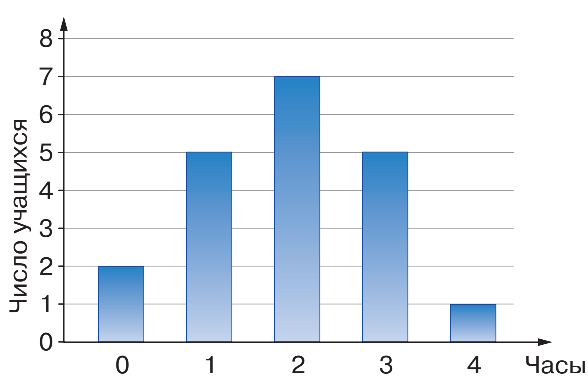


Рис. 23

а) Сколько времени в день в среднем тратят ученики 10 «А» класса на выполнение домашних заданий? (Найдите среднее арифметическое этого ряда данных.)
б) Сколько времени тратит средний ученик на выполнение домашних заданий? (Найдите медиану этих данных.)
в) Сколько времени тратит на выполнение домашних заданий большинство из учеников 10 «А» класса? (Найдите моду этих данных.)

34. Школа подвела итоги Всероссийской проверочной работы по математике в седьмых классах и представила результаты на круговой диаграмме (рис. 24). Найдите моду и медиану этого ряда отметок.

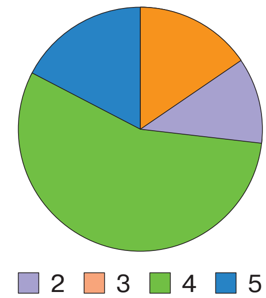


Рис. 24

35. Средняя масса волнистых попугайчиков школьного живого уголка равна 42 г. Масса попугайчика Кеши равна 43 г. Верны ли следующие утверждения?
а) Все попугайчики, кроме Кеши, имеют массу 42 г.
б) Масса каждого попугайчика, кроме Кеши, меньше 42 г.
в) В живом уголке есть попугайчик, масса которого меньше 42 г.
г) В живом уголке есть попугайчик, масса которого равна 41 г.

36. Президент компании получает 1 000 000 р. в год, четверо его заместителей получают по 200 000 р. в год, а 20 служащих компании получают по 100 000 р. в год.
а) Найдите все средние (среднее арифметическое, моду, медиану) зарплат в компании.
б) Компания должна предоставить в муниципальную статистическую службу информацию о средней зарплате служащих компании и о зарплате среднего служащего. Какие из найденных вами данных надо предоставить в каждом случае?
в) Какой показатель — медиана зарплат или их среднее арифметическое — представляется вам более объективным?

37. Директор фирмы решил начать борьбу за здоровый образ жизни и провёл анализ заболеваемости своих сотрудников. Он выписал число рабочих дней, пропущенных в течение года по болезни каждым сотрудником, предварительно разбив сотрудников на две группы — посещающих спортзал и не посещающих. Получились такие результаты:
Посещающие спортзал: 3, 3, 6, 0, 3, 6, 2, 2, 4, 5, 13, 4, 3, 2, 2, 3, 4, 5, 6, 4.
Не посещающие спортзал: 7, 5, 2, 6, 4, 4, 6, 7, 9, 7, 0, 8, 11, 8.
Директор сделал по этим результатам убедительные выводы о пользе занятий спортом. Сделайте и вы то же самое.

38. В таблице 32 представлены данные о количестве детей в семьях города. Если вас попросят охарактеризовать по этим данным среднестатистическую городскую семью, как вы это сделаете?

Таблица 32

Число детей в семье	0	1	2	3	4	5	6
Число семей	255	320	210	80	18	6	1

39. Маша, Саша, Катя, Лена, Ваня и Миша пошли в пиццерию. Ваня съел 5 кусков пиццы, Миша, Саша и Лена — по 3 куска, Катя — 2 куска, Маша — 1 кусок. Найдите все известные вам средние характеристики этих данных. Если бы Ваня съел не 5, а 7 кусков пиццы, как бы изменились эти величины?

40. Выполните задания:
1) Вычислите все средние характеристики ряда 2, 8, 16, 24, 30, 40. Используя полученный результат, попробуйте догадаться, чему равны эти характеристики для следующих рядов:
а) 12, 18, 26, 34, 40, 50;   б) 20, 80, 160, 240, 300, 400.
Проверьте себя с помощью вычислений.
2) Как изменится каждая из средних характеристик, если:
а) ко всем членам ряда прибавить одно и то же число;
б) все члены ряда умножить на одно и то же число?

41*. Постройте ряд из 4 или более чисел (не все из которых равны между собой), у которого:
а) среднее арифметическое равно медиане, но не равно моде;
б) среднее арифметическое равно моде, но не равно медиане;
в) среднее арифметическое, медиана и мода равны между собой.

42*. Числовой набор содержит 9 нулей и некоторое число x:

$$0, 0, 0, 0, 0, 0, 0, 0, 0, x.$$

Найдите для этого ряда среднее арифметическое, моду, медиану. Какие из этих характеристик зависят от x? Нарисуйте графики этих зависимостей.

43*. Числовой набор содержит 5 чисел:

$$1, 2, 3, 4, x.$$

Найдите для этого ряда среднее арифметическое, моду, медиану. Какие из этих характеристик зависят от x? Нарисуйте графики этих зависимостей.

6 Наибольшее и наименьшее значения. Размах

Кроме средних характеристик, используются и другие числа, связанные с набором данных и содержащие иногда не менее важную информацию.

Например, если мы просматриваем в интернет-магазинах цены на интересующий нас товар, то нас, скорее всего, будет интересовать не средняя цена, а наименьшая. А если вы ищете работу, то будете начинать с тех вакансий, где заработная плата наибольшая.

На спортивных соревнованиях мы часто слышим выражение «лучший результат сезона в мире». В этом случае речь тоже идёт о наибольшем или наименьшем значении из всех результатов, показанных на всех соревнованиях в данном виде спорта в этом сезоне. Что именно понимается под «лучшим результатом» — наибольший или наименьший — зависит от вида спорта. Для прыжков в длину — это наибольшее значение, а для бега — наименьшее.

Когда учёные и инженеры планируют запуск летательного аппарата на какую-то планету, их интересует в первую очередь не среднесуточная температура на её поверхности, а минимальная и максимальная температуры, именно от них зависит, из каких материалов должно быть изготовлено оборудование.

Помимо наибольшего и наименьшего значений, представляет интерес и их разность — ***размах***.

Размахом числового набора называют разность между наибольшим и наименьшим значениями этого набора.

Если вернуться к примеру со спортивными соревнованиями, то для конкретного спортсмена размах его результатов в течение сезона несёт очень важную информацию об их стабильности. При прочих равных условиях этот фактор зачастую становится решающим при отборе спортсменов на важные соревнования.

Размах цен на рынке говорит о том, насколько этот рынок уже сложился, а размах доходов граждан — о степени социального расслоения общества.

Очень часто размах дополняет, а иногда заставляет полностью пересмотреть представление о числовых данных, полученное по средним характеристикам. Например, известно, что на экваторе Меркурия среднесуточная температура составляет $+15°$. Исходя из этого статистического показателя можно подумать, что на Меркурии умеренный климат, удобный для жизни людей. Но на самом деле это не так. Температура на Меркурии колеблется от наименьшего значения $-150°$ до наибольшего значения $+350°$. Размах равен $350° - (-150°) = 500°$. Конечно, такого перепада температур человек выдержать не может.

В электронной таблице наибольшее и наименьшее значения числового набора вычисляются с помощью функций

МИН(диапазон) — наименьшее значение;

МАКС(диапазон) — наибольшее значение.

Размах можно найти как разность двух этих значений.

? ВОПРОСЫ

1. Каков размах всех рядов из примеров 1 и 2?
2. Может ли размах равняться 0? Если да, то приведите пример такого ряда.
3. Витя получил за четверть 7 четвёрок и 3 пятёрки. Чему равен размах его отметок?

7 Дисперсия и стандартное отклонение

Размах является простейшей числовой мерой ***разброса данных***. У размаха есть два недостатка. Во-первых, он очень чувствителен к так называемым ***выбросам*** — слишком большим или слишком маленьким значениям, которые могли попасть в числовой набор случайно или по ошибке. Во-вторых, он не учитывает частоту, с которой в выборке встречаются значения. Так, например, для всех трёх учеников, о которых шла речь выше, размах отметок будет одинаковым: $5 - 2 = 3$.

Поэтому в статистике используется другой подход к определению разброса. Попробуем вычислить, насколько **в среднем** значения в наборе данных отклоняются от его среднего арифметического.

Пусть имеется числовой набор $x_1, x_2, ..., x_n$. Обозначим его среднее арифметическое через $\overline{x}$:

$$\overline{x} = \frac{x_1 + ... + x_n}{n}.$$

Тогда ***отклонения*** исходных чисел от их среднего будут равны $x_1 - \overline{x}, x_2 - \overline{x}, ..., x_n - \overline{x}$. Например, для набора из семи чисел 1, 3, 4, 6, 7, 10, 11 их среднее арифметическое будет

$$\frac{1 + 3 + 4 + 6 + 7 + 10 + 11}{7} = 6,$$

а отклонения от него будут равны –5; –3; –2; 0; 1; 4; 5.

Как видите, отклонения могут быть и положительными, и отрицательными, и равными нулю. При этом **сумма всех отклонений от среднего всегда равна нулю**. В данном примере это легко проверить (проверьте самостоятельно), но можно доказать этот факт и для произвольного набора чисел. Действительно:

$$(x_1 - \overline{x}) + (x_2 - \overline{x}) + ... + (x_n - \overline{x}) = (x_1 + x_2 + ... + x_n) - n \cdot \overline{x} =$$

$$= n \cdot \overline{x} - n \cdot \overline{x} = 0.$$

Это значит, что рассматривать в качестве числовой характеристики разброса среднее отклонение от $\overline{x}$ бессмысленно — оно будет всегда равно нулю:

$$\frac{(x_1 - \overline{x}) + (x_2 - \overline{x}) + ... + (x_n - \overline{x})}{n} = 0.$$

Чтобы положительные отклонения от среднего не компенсировались отрицательными отклонениями, можно, например, не учитывать их знаки, т. е. находить сумму их модулей. Но в статистике нашли другой выход: возводить каждое отклонение от среднего в квадрат. Эту меру разброса называют ***дисперсией*** (от латинского слова *dispersio*, означающего «рассеяние») и обозначают S^2:

$$S^2 = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n}$$

(обозначение S^2 лишний раз подчёркивает, что дисперсия всегда неотрицательна).

Дисперсией числового набора называют среднее арифметическое квадратов отклонений от его среднего значения.

На величину дисперсии влияют все отклонения, причём чем больше абсолютное значение отклонения, тем сильнее это влияние (благодаря возведению в квадрат). В отличие от среднего отклонения, которое всегда равно 0, дисперсия всегда положительна, кроме одного случая: если все числа в наборе данных одинаковые, то дисперсия равна 0.

Найдём дисперсию для приведённого выше ряда чисел. Отклонения от среднего мы уже нашли, остаётся возвести каждое из них в квадрат и найти среднее арифметическое:

$$D = \frac{25 + 9 + 4 + 0 + 1 + 16 + 25}{7} = \frac{80}{7} \approx 11{,}43.$$

Рассмотрим теперь реальную ситуацию, в которой полезно учитывать не только средние характеристики, но и разброс данных.

Пример. Представьте, что вы хотите купить себе электронную книгу и решили изучить предложения на одну ту же модель от разных продавцов. Получились такие цены (в рублях):

6400, 5300, 6800, 5900, 6600, 6400, 5600, 6600, 4400.

На первый взгляд кажется, что нужно выбрать самый дешёвый гаджет за 4400 р. Но не торопитесь с выводами — вспомните поговорку «скупой платит дважды». За низкой ценой могут стоять скрытые дефекты, отсутствие гарантии и т. д. Разумнее ориентироваться на среднюю цену. Вычислим её:

$$\frac{6400 + 5300 + \ldots + 4400}{9} = 6000 \text{ р.}$$

Электронной книги с такой ценой нет, но можно взять ближайшую за 5900 р. Если упорядочить все цены по возрастанию, то легко будет найти медиану этого набора чисел:

4400, 5300, 5600, 5900, 6400, 6400, 6600, 6600, 6800.

Медиана равна 6400 р.

Для чего в данном примере может понадобиться разброс? В экономике по разбросу можно судить о стабильности рынка: если разброс небольшой, то цены установились и в ближайшее время вряд ли будут сильно меняться. Увеличение разброса означает, что рынок находится в движении, цены могут начать падать или расти. Так происходит, когда какое-то изделие только появилось на рынке, или наоборот, его снимают с производства. В любом случае информация о разбросе цен позволяет принять решение о том, стоит ли торопиться с этой покупкой.

Размах цен в нашем примере составляет 6800 – 4400 = 2400 р. Вычислим теперь дисперсию. Поскольку среднее арифметическое этого набора равно 6000, то отклонения от среднего будут равны соответственно:

$$400,\ -700,\ 800,\ -100,\ 600,\ 400,\ -400,\ 600,\ -1600.$$

Теперь нужно возвести каждое из них в квадрат и вычислить среднее арифметическое:

$$\frac{160\,000 + 490\,000 + \ldots + 2\,560\,000}{9} \approx 544\,444.$$

Да, получилось именно такое большое число, больше 500 000. Но самое удивительное — это единицы измерения, в которых выражается дисперсия. Они «квадратные»! Если исходная величина измерялась в метрах, то дисперсия будет выражена в квадратных метрах, а если она измерялась в рублях, то будет выражена в «квадратных» рублях. Ведь мы возводили в квадрат каждое отклонение, выраженное в рублях.

Чтобы избавиться от таких странных единиц измерения, рассматривают другую характеристику разброса — ***стандартное отклонение***, которое связано с дисперсией очень простым соотношением.

Стандартным (или средним квадратичным) отклонением числового ряда называется квадратный корень из дисперсии.

Поскольку для дисперсии мы уже ввели обозначение S^2, то естественно обозначить стандартное отклонение буквой S. В рассмотренном примере стандартное отклонение $S = \sqrt{544\ 444} \approx 737{,}86$ уже обыкновенных (не квадратных) рублей. Теперь и полученная величина соответствует разбросу цен: 737 р. гораздо ближе к реальным колебаниям цен, чем 544 000.

Если говорить о том, как в общем случае оценить с помощью стандартного отклонения разброс чисел вокруг среднего значения, то обычно в промежуток $[\bar{x} - S; \bar{x} + S]$ попадает более половины значений числового набора, а в промежуток $[\bar{x} - 3S; \bar{x} + 3S]$ — практически все значения. В нашем примере это можно проверить, если нанести все исходные данные на числовую ось (рис. 25). Более точно сформулировать эти и другие свойства стандартного отклонения мы сможем позже, когда начнём изучать ***случайные величины***.

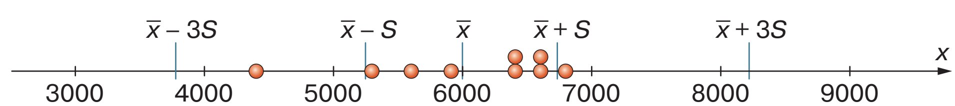


Рис. 25. Среднее арифметическое и стандартное отклонение

Для вычисления дисперсии числового набора в электронной таблице служит функция

ДИСПР(диапазон),

а для стандартного отклонения —

СТАНДОТКЛОНП(диапазон).

? ВОПРОСЫ

1. Чему равно среднее отклонение от среднего?
2. Что называется дисперсией числового набора?
3. Что можно сказать о наборе чисел, в котором размах равен 0? дисперсия равна 0?
4. Как связаны единицы измерения дисперсии и числового набора?

8 Формула для вычисления дисперсии*

Приведённые выше определения позволяют найти дисперсию и стандартное отклонение для любого набора данных. Правда, если чисел много, то для этого может потребоваться большой объём вычислений. Его можно немного сократить, если использовать другую формулу для вычисления дисперсии.

Обозначим через $\overline{x^2}$ среднее арифметическое квадратов всех чисел из нашего набора:

$$\overline{x^2} = \frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}.$$

Тогда для дисперсии числового набора S^2 будет справедлива следующая формула:

$$S^2 = \overline{x^2} - (\overline{x})^2.$$

Заметьте, что, несмотря на схожесть обозначений, числа $\overline{x^2}$ и $(\overline{x})^2$ разные: при вычислении $\overline{x^2}$ мы возводим каждое из исходных чисел в квадрат, а потом находим среднее арифметическое этих квадратов; при вычислении $(\overline{x})^2$ мы находим среднее арифметическое исходных чисел, а потом возводим его в квадрат.

Доказать эту формулу несложно, мы оставим это в качестве самостоятельного упражнения 63*.

Особенно явно преимущество нового способа вычисления дисперсии сказывается для наборов из целых чисел, у которых среднее арифметическое является дробным. В этом случае отклонения от среднего тоже будут дробными, и нам придётся возводить в квадрат дробные числа. По новой формуле мы будем делать то же самое с целыми числами. Покажем это на примере.

Пример 1. Найдём дисперсию числового набора 1, 2, 3, 4, 5, 6 (этот набор вам хорошо знаком: его можно увидеть на гранях любого игрального кубика).

Сначала посчитаем дисперсию по определению. Среднее арифметическое этого набора равно 3,5. Найдём отклонение каждого из чисел от среднего:

$$-2{,}5;\ -1{,}5;\ -0{,}5;\ 0{,}5;\ 1{,}5;\ 2{,}5.$$

Для вычисления дисперсии возведём каждое из отклонений в квадрат и посчитаем их среднее арифметическое:

$$S^2 = \frac{6{,}25 + 2{,}25 + 0{,}25 + 0{,}25 + 2{,}25 + 6{,}25}{6} = \frac{17{,}5}{6} = 2{,}917.$$

Как видите, пришлось выполнить довольно много действий с дробными числами.

Теперь посчитаем дисперсию по новой формуле. Для этого найдём средний квадрат нашего числового набора:

$$\overline{x^2} = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6} = 15{,}167.$$

А теперь вычтем из него квадрат среднего:

$$S^2 = 15{,}167 - 3{,}5^2 = 2{,}917.$$

Здесь дробные числа возникли только на последнем шаге вычислений. Понятно, что если вычисления проводятся на компьютере, то это не имеет значения, а вот человеку оперировать с целыми числами гораздо удобнее.

Если числовой набор представлен в виде таблицы частот, то дисперсию, как и среднее арифметическое, нужно вычислять с учётом числа повторений каждого значения или его частоты. Покажем это на примере.

Пример 2. Найдём дисперсию и стандартное отклонение отметок ученика, заданных таблицей частот (табл. 33).

Таблица 33

Отметка	2	4	5
Число повторений	1	3	6
Частота	0,1	0,3	0,5

Из таблицы видно, что всего у ученика десять отметок: одна двойка, три четвёрки и шесть пятёрок. Размах ряда отметок вычислить просто — он равен разности последнего и первого значений в первой строке: $5 - 2 = 3$.

Найдём среднее арифметическое отметок:

$$\bar{x} = \frac{2 \cdot 1 + 4 \cdot 3 + 5 \cdot 6}{10} = 4{,}4.$$

Среднее арифметическое квадратов, как и среднее арифметическое самих оценок, можно вычислять либо с использованием числа повторений:

$$\overline{x^2} = \frac{2^2 \cdot 1 + 4^2 \cdot 3 + 5^2 \cdot 6}{10} = \frac{4 + 48 + 150}{10} = 20{,}2,$$

либо с использованием частот: $\overline{x^2} = 2^2 \cdot 0{,}1 + 4^2 \cdot 0{,}3 + 5^2 \cdot 0{,}6 = 20{,}2.$

Теперь вычислим дисперсию:

$$S^2 = \overline{x^2} - (x)^2 = 20{,}2 - 4{,}4^2 = 20{,}2 - 19{,}36 = 0{,}84,$$

и стандартное отклонение: $S = \sqrt{0{,}84} = 0{,}92.$

Обратите внимание: почти все отметки ученика отличаются от среднего меньше чем на S, т. е. он учится достаточно стабильно. Одна двойка, которая выпадает из этого диапазона, по-видимому, для него случайная.

? ВОПРОСЫ

1. Запишите две формулы для вычисления дисперсии.
2. Чему равна дисперсия числового набора 1, 2, 3, 4? Попробуйте посчитать её в уме, воспользовавшись приведённой в этом пункте формулой.
3. Какое неравенство следует из новой формулы для вычисления дисперсии? Когда оно превращается в равенство?
4. Что такое стандартное отклонение?

УПРАЖНЕНИЯ

44. Найдите размах ряда чисел:
а) 6; 3; 6; 2; 4; 6; 2; 4; 5; 5; б) 2,7; 4,9; 11,4; 3,5; 0,1.

45. Столбиковая диаграмма на рисунке 26 показывает, сколько книг прочитал каждый из десятиклассников за летние каникулы. Найдите: а) среднее арифметическое этого ряда данных; б) медиану этого ряда данных; в) размах этих данных.

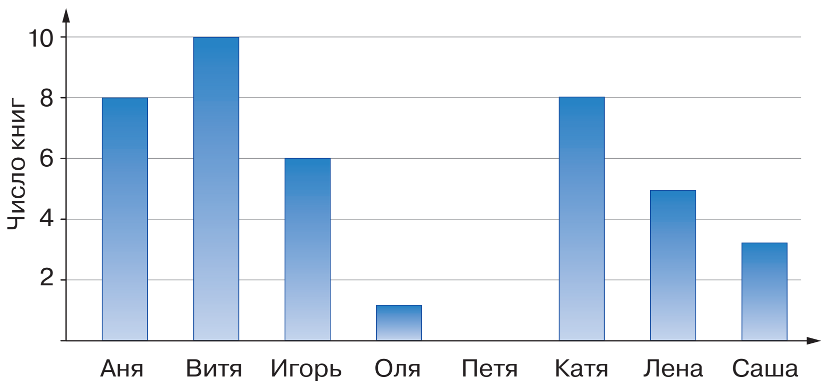


Рис. 26

46. В таблице 34 приведены расходы студента за шесть учебных дней недели.

Таблица 34

День недели	Пн	Вт	Ср	Чт	Пт	Сб
Расходы, р.	380	400	350	400	270	240

Определите, какая статистическая характеристика находится в каждом случае:
а) 380 + 400 + 350 + 400 + 270 + 240 = 2040; 2040 : 6 = 340. Ответ: 340 р.;
б) 240, 270, 350, 380, 400, 400; (350 + 380) : 2 = 365. Ответ: 365 р.;
в) 380, 400, 350, 400, 270, 240. Ответ: 400 р.;
г) 400 – 240 = 160. Ответ: 160 р.

47. Найдите дисперсию числовых наборов:
а) 2; 2; 2; 2; б) –1; 0; 1; в) –1; 0; 10; г) 1; 2; 3.

48. Найдите размах, дисперсию и стандартное отклонение числового набора:

1, 2, 3, 4, 5, 6, 7, 8.

Для каких значений ряда отклонение от среднего не превышает стандартного отклонения?

49. Для каждого из двух наборов чисел вычислите среднее арифметическое, дисперсию, стандартное отклонение и сравните их:
а) 3, 7, 10, 11, 19 и 10, 11, 15, 17, 22; б) 1, 3, 5, 7, 9 и 2, 4, 6, 8, 10.

50. В таблице 35 представлены результаты контрольной работы по математике в 10 классе. Найдите размах, дисперсию и стандартное отклонение набора отметок.

Таблица 35

Отметка	2	3	4	5
Число учеников	2	8	10	9

51. В таблице 36 указано число книг, прочитанных несколькими ребятами за летние каникулы.

Таблица 36

Имя	Аня	Витя	Игорь	Оля	Петя	Катя	Лена	Саша
Число книг	8	10	6	1	0	7	5	3

а) Для данного ряда вычислите среднее арифметическое и стандартное отклонение.

б) Назовите имена тех ребят, для которых модуль разности между количеством прочитанных ими книг и средним арифметическим превышает стандартное отклонение.

52. В таблице 37 приведены данные о доле учеников, получивших соответствующие итоговые отметки по математике в каждом из трёх 10 классов.
Определите, какой класс учится стабильнее.

Таблица 37

Класс	**Частота отметки**			
	2	**3**	**4**	**5**
10 «А»	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{10}$
10 «Б»	$\frac{1}{14}$	$\frac{2}{7}$	$\frac{1}{4}$	$\frac{11}{28}$
10 «В»	$\frac{1}{6}$	$\frac{5}{12}$	$\frac{1}{4}$	$\frac{1}{6}$

53. В таблице 38 приведены данные о возрастном составе участников школьного хора. Найдите процент участников, чей возраст отклоняется от среднего арифметического больше, чем на стандартное отклонение.

Таблица 38

Возраст	7	8	9	10	11	12	13	14	15
Число участников	3	6	5	1	2	3	2	2	1

Задания 54—56 нужно сделать с помощью электронной таблицы.

54. Внесите в электронную таблицу данные классного журнала из таблицы 1 пункта 1 «Таблицы».

1) Вычислите с помощью функции СРЗНАЧ() средний балл каждого ученика.

2) Вычислите с помощью функции МЕДИАНА() медиану отметок каждого ученика.

3) Кто из учеников этого класса лучше всех успевает по математике? Кто — хуже всех?

4) Вычислите с помощью функций МАКС() и МИН() размах отметок каждого ученика.

5) Вычислите с помощью функции СТАНДОТКЛОНП() стандартное отклонение отметок каждого ученика.

6) Кто из учеников этого класса показывает наиболее стабильные результаты? наименее стабильные?

7) Отразите полученные данные на диаграммах.

55. На стройку с кирпичного завода привезли 20 упаковок кирпича. Чтобы проверить качество партии, из каждой упаковки вытащили случайным образом по кирпичу и замерили длину каждого. Ниже представлены полученные величины (в см):

20,5; 20,1; 21,3; 20,3; 19,8; 19,2; 20,1; 19,6; 20,2; 20;

20,5; 19,7; 19,9; 20,5; 19,6; 20,1; 19,4; 19,8; 19,1; 20,3.

а) Определите среднюю длину кирпича.

б) Найдите величину стандартного отклонения длины кирпича от средней.

в) Каков процент кирпичей, длина которых отличается от средней больше чем на 0,2 см?

г) Каков процент кирпичей, длина которых отличается от средней больше чем на величину стандартного отклонения?

56. В таблице 39 приведены данные о росте участников легкоатлетических соревнований.

Таблица 39

Рост (см)	[160;165)	[165;170)	[170;175)	[175;180)	[180;185)	[185;190)	[190;195]
Число участников	5	12	19	25	10	7	2

Определите процент участников, у которых рост попадает в промежуток

а) $[\overline{X} - S, \overline{X} + S]$; б) $[\overline{X} - 2S, \overline{X} + 2S]$; в) $[\overline{X} - 3S, \overline{X} + 3S]$;

где $\overline{X}$ — среднее арифметическое, а $\overline{S}$ — стандартное отклонение.

57. Президент компании «Альфа» получает 100 000 р. в год, четверо его заместителей получают по 20 000 р. в год, а 20 служащих — по 10 000 р. в год.

Президент компании «Бета» получает 200 000 р. в месяц, пятеро его заместителей получают по 20 000 р. в месяц, а 20 служащих — по 9000 р. в месяц.

а) Не используя расчётов на бумаге, попробуйте определить, на каком предприятии больше средний уровень зарплаты, а на каком — разброс зарплаты.

б) Проверьте свои предположения, вычислив соответствующие числовые характеристики числовых рядов.

в) В какой из этих двух компаний вы бы предпочли работать?

58. На предприятии прошло повышение зарплаты.

а) Каждому сотруднику предприятия увеличили зарплату на 1000 р. Как изменилась средняя зарплата? размах? дисперсия? стандартное отклонение?

б) Каждому сотруднику предприятия увеличили зарплату в 10 раз. Как изменилась средняя зарплата? размах? дисперсия? стандартное отклонение?

59. Докажите, что если каждое число в заданном наборе увеличить на одну и ту же константу, то дисперсия не изменится.

60. Докажите, что если каждое число в заданном наборе умножить на одну и ту же константу, то дисперсия умножится на квадрат этой константы.

61. Ребятам было поручено провести статистическое исследование роста одноклассников. Коля записал рост ребят в сантиметрах: 162; 181; 179; ..., а Оля — в метрах: 1,62; 1,81; 1,79; ... . Затем они подсчитали средний рост, дисперсию и стандартное отклонение. Коля получил соответственно 172, 16 и 4. Какие результаты получила Оля?

62. Соберите данные о стоимости одного из основных продуктов питания в магазинах вашего микрорайона (например, о стоимости 1 л молока). Вычислите все известные вам статистические характеристики полученного вами ряда чисел.

63*. Докажите, что дисперсия числового набора может быть вычислена по формуле

$$S^2 = \overline{x^2} - (\overline{x})^2,$$

где $\overline{x}$ — среднее арифметическое всех заданных чисел, а $\overline{x^2}$ — среднее арифметическое квадратов этих чисел.

64*. Числовой набор содержит девять нулей и некоторое число x:

$$0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ x$$

Найдите для этого ряда дисперсию и стандартное отклонение, выразив их через x. Нарисуйте графики этих зависимостей. При каком значении x они достигают минимума?

65*. Числовой набор содержит пять чисел:

$$1,\ 2,\ 3,\ 4,\ x.$$

Найдите для этого ряда дисперсию и стандартное отклонение, выразив их через x. Нарисуйте графики этих зависимостей. При каком значении x они достигают минимума?

66*. Докажите, что стандартное отклонение любого числового набора не больше, чем его размах.

ЛАБОРАТОРНАЯ РАБОТА № 2

Вычисление характеристик числового набора в электронной таблице

Задание 1. Анализ вашей успеваемости (продолжение)

Выберите предмет, по которому у вас больше всего текущих отметок, и выпишите их в столбец A.
Найдите в столбце B средние характеристики этого набора: моду, среднее арифметическое, медиану.
Вычислите в столбце C все характеристики разброса: размах, дисперсию, стандартное отклонение.
Какую итоговую отметку можно выставить по полученным результатам?

Задание 2. Числовые характеристики возраста

Какой средний возраст учеников вашего класса? А его разброс? Электронная таблица поможет найти ответ на этот вопрос с точностью до одного дня.
Запишите в столбец A день рождения каждого ученика в формате даты (например, день рождения 10 февраля 2006 г. записывается как 10.02.2006).
Посчитайте в столбце B возраст каждого ученика в прожитых днях. Для этого можно использовать функцию СЕГОДНЯ(), которая всегда возвращает текущую дату. Если ввести в ячейку B1 формулу

=СЕГОДНЯ()–A1,

то вы получите количество дней, которое прожил первый ученик.
Найдите моду, среднее арифметическое и медиану всех полученных возрастов.
Вычислите размах и межквартильный размах.
Проанализируйте полученные данные.

ГЛАВА 2

ЭЛЕМЕНТЫ ТЕОРИИ ГРАФОВ*

Теория графов может считаться одновременно как одной из самых старых, так и самых молодых математических дисциплин. Первой задачей, с которой началось её развитие, принято считать ***задачу о кёнигсбергских мостах***, сформулированную и решённую Леонардом Эйлером ещё в 1736 г. С тех пор графы стали применяться для решения самых разных задач как внутри математики, так и в смежных дисциплинах.
Но по-настоящему самостоятельным разделом математики теория графов стала только спустя 200 лет благодаря первой научной книге по теории графов, изданной венгерским математиком Денешем Кёнигом. Появление в середине прошлого века компьютеров вызвало новый интерес к теории графов и превратило её в одну из самых активно развивающихся научных дисциплин, находящихся на стыке математики и информатики.

§ 3. Граф и способы его задания

1 Определение графа

Графом в математике называется представление объектов и связей между ними с помощью множества точек, некоторые из которых попарно соединены между собой линиями. Точки называются ***вершинами*** графа, а соединяющие их линии — ***рёбрами***. При необходимости вершины и рёбра обозначают буквами или числами.

Вместо точек иногда удобно использовать круги или другие геометрические фигуры. Рёбрами могут быть любые линии, но чаще всего используют отрезки. На рисунке 27 показано изображение графа, состоящего из 11 вершин и 12 рёбер.

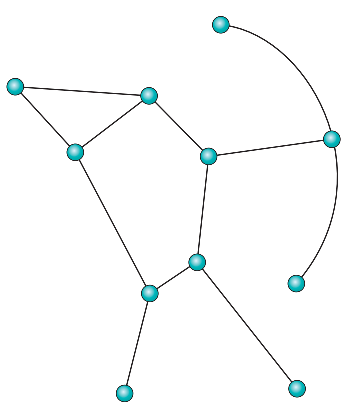
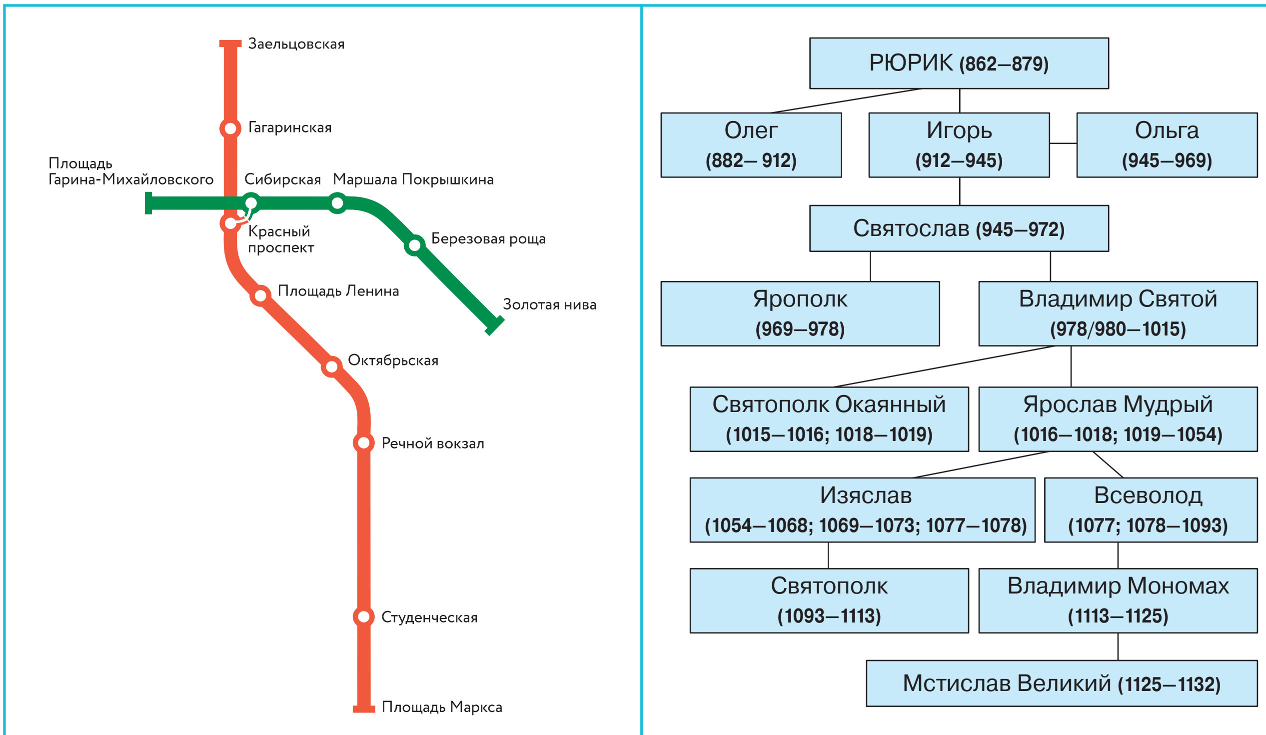


Рис. 27. Граф

Благодаря наглядности графы нашли применение в самых разных областях. С их помощью можно описывать транспортные сети, родственные связи между людьми, строение молекул, соединения деталей и т. д. На рисунке 28 можно увидеть некоторые из перечисленных примеров.



а) схема Новосибирского метрополитена

б) родословная великих князей династии Рюриковичей

в) структурная схема молекулы пропана

г) электрическая схема

Рис. 28. Примеры использования графов

Граф можно задать рисунком, а можно специальной таблицей — ***матрицей смежности***. Она показывает, какие пары вершин в графе соединены рёбрами. На рисунке 29 изображён граф из 5 вершин *a*, *b*, *c*, *d*, *e*, в котором 7 рёбер. Рядом с ним вы видите таблицу 40, которая содержит 5 строк и 5 столбцов, соответствующих вершинам графа. Если вершина *x* соединена с вершиной *y* ребром, то на пересечении строки *x* и столбца *y* в этой таблице стоит 1, если соединения нет, то 0.

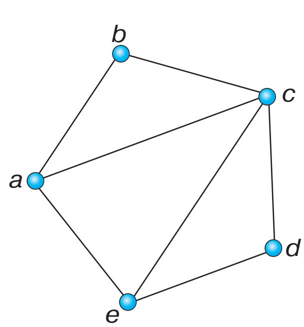


Рис. 29. Граф и его матрица смежности

Таблица 40

	a	***b***	***c***	***d***	***e***
a	0	1	1	0	1
b	1	0	1	0	0
c	1	1	0	1	1
d	0	0	1	0	1
e	1	0	1	1	0

Изображённый на рисунке 29 граф называется ***неориентированным***, так как на его рёбрах не указаны направления. Если нарисовать на каждом ребре стрелку (т. е. указать направление), то он станет ***ориентированным*** (рис. 30).

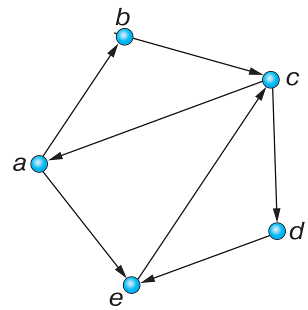


Рис. 30. Ориентированный граф

Таблица 41

	a	**b**	**c**	**d**	**e**
a	0	1	0	0	1
b	0	0	1	0	0
c	1	0	0	1	0
d	0	0	0	0	1
e	0	0	1	0	0

Обратите внимание, что матрица неориентированного графа симметричная: в ней клетки, симметричные относительно диагонали, содержат одинаковые значения. Для ориентированного графа это уже не так. Например, в таблице 41 на пересечении строки *a* и столбца *b* стоит 1, а в симметричной ей клетке — 0. Если рассматривать рёбра графа как дороги, а вершины как перекрёстки, то неориентированный граф будет соответствовать дорогам с двусторонним движением, а ориентированный — с односторонним.

Когда речь заходит о дорогах, то появляется необходимость указывать не только наличие дороги между двумя заданными пунктами, но и её протяжённость (а если дорога платная — то стоимость проезда). В этих случаях каждому ребру сопоставляется соответствующее число (рис. 31). Оно называется ***весом ребра***, а сам граф — ***взвешенным***. Вместо единиц в матрицу смежности взвешенного графа записываются веса рёбер (табл. 42).

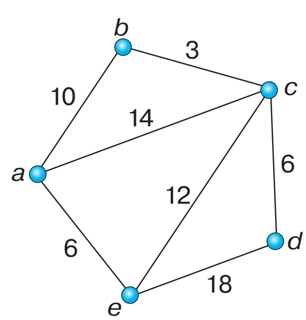


Рис. 31. Взвешенный граф

Таблица 42

	a	**b**	**c**	**d**	**e**
a	0	10	14	0	6
b	10	0	3	0	0
c	14	3	0	6	12
d	0	0	6	0	18
e	6	0	12	18	0

Если допускается наличие нескольких рёбер между одной и той же парой вершин, то такой граф называется ***мультиграфом***, а несколько рёбер — ***кратными*** рёбрами. Сейчас между многими городами и населёнными пунктами есть два варианта автомобильных дорог: платные и бесплатные. Их протяжённость (а тем более стоимость) может быть разной. Для представления такой информации понадобится взвешенный мультиграф.

Обычный (невзвешенный) мультиграф изображён на рисунке 32. Вершины *a* и *b* соединены в нём тремя, а вершины *b* и *c* — двумя рёбрами.

Наконец, на графе могут быть рёбра особого вида, которые соединяют вершину с ней же самой. Такие рёбра называют ***петлями***. На рисунке 33 показан граф, вершины которого пронумерованы цифрами от 0 до 9. Граф показывает, как преобразуется последняя цифра числа после его возведения в квадрат. Петли нарисованы около вершин 0, 1, 5 и 6: если число оканчивается на одну из этих цифр, то после возведения числа в квадрат эта цифра не изменится. Цифра 7 после возведения в квадрат перейдёт в цифру 9, а та в свою очередь — в цифру 1: именно поэтому на графе нарисованы стрелки из 7 в 9, а также из 9 в 1.

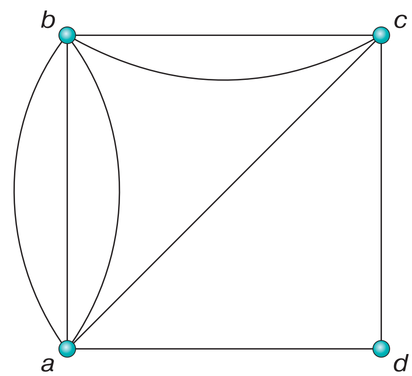


Рис. 32. Мультиграф

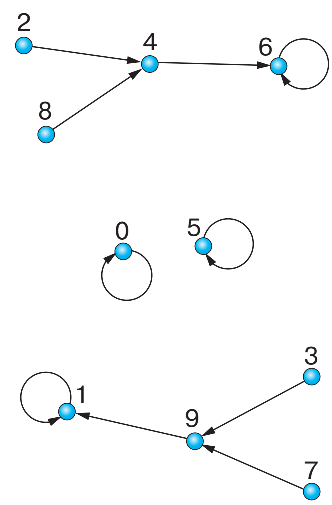
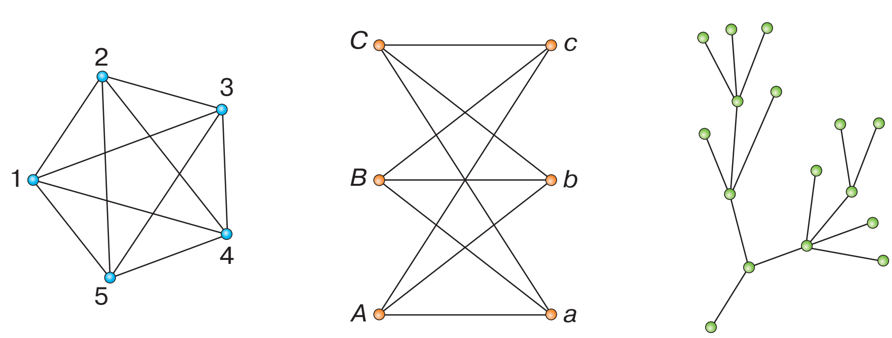


Рис. 33. Граф возведения в квадрат

В дальнейшем мы будем рассматривать в основном **неориентированные, невзвешенные графы без петель и кратных рёбер.** Такие графы называются ***простыми***. Три простых графа изображены на рисунке 34. В первом из них (рис. 34, *а*) вершины обозначены числами от 1 до 5, во втором (рис. 34, *б*) — большими и маленькими латинскими буквами, а в третьем (рис. 34, *в*) — вершины остались без обозначений.



а) б) в)

Рис. 34. Простые графы

? ВОПРОСЫ

1. Что такое граф?
2. Приведите пример графа, с которым вы сталкивались в реальной жизни. Что служило вершинами, а что рёбрами этого графа?
3. Нарисуйте какой-нибудь граф с четырьмя вершинами и запишите его матрицу смежности.
4. Чем ориентированный граф отличается от неориентированного?
5. Какой граф называется взвешенным?
6. Чем мультиграф отличается от графа?
7. Что такое простой граф?

2 Степени вершин

Главной характеристикой вершины графа (или мультиграфа) является её ***степень***.

Степенью вершины называется количество рёбер, проведённых из этой вершины.

Если вершина имеет степень 0 (т. е. не соединена ни с какой другой вершиной), то она называется ***изолированной***.

Чтобы получить максимально возможную степень вершины в простом графе, нужно соединить её со всеми другими вершинами, кроме неё самой. Значит, если в графе n вершин, то максимально возможная степень вершины равна $(n - 1)$. Это утверждение касается только простых графов: в мультиграфах мы можем провести сколько угодно рёбер между любой парой вершин, поэтому там степень вершины ничем не ограничена.

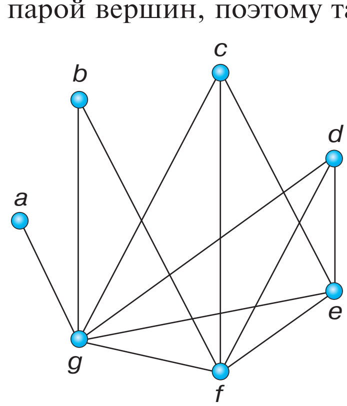


Рис. 35. Степени вершин

Таблица 43

Вершина	*a*	*b*	*c*	*d*	*e*	*f*	*g*
Степень	1	2	3	3	4	5	6

На рисунке 35 изображён простой граф с 7 вершинами. Их степени равны 1, 2, 3, 3, 4, 5, 6 (табл. 43). Мы уже поняли, что диапазон, в котором могут лежать степени вершин такого графа, — это числа от 0 до 6. Но всякий ли набор чисел из этого диапазона может служить степенями вершин? Скажем, можно ли построить простой граф с 7 вершинами, все степени которых равны 2? равны 3? Можно ли построить граф со степенями вершин 0, 1, 2, 3, 4, 5, 6? со степенями 1, 1, 2, 2, 3, 3, 4?

Прежде чем читать параграф дальше, попробуйте построить такие графы (ещё раз напомним, что речь идёт о простых графах, в которых нет петель и кратных рёбер).

Теперь мы рассмотрим две интересные теоремы, которые касаются степеней вершин и часто используются при решении задач. В частности, они объясняют, почему не все графы, описанные в предыдущем абзаце, вам удалось построить.

Теорема 1. Если простой граф содержит больше одной вершины, то в нём обязательно найдутся хотя бы две вершины одинаковой степени.

Доказательство. Доказательство проведём методом от противного. Пусть в графе n вершин. Тогда минимально возможная степень вершины — 0 (когда вершина изолирована), а максимально возможная — $(n - 1)$ (когда она соединена со всеми остальными вершинами). Следовательно, единственный вариант, когда степени всех n вершин различны, — это когда они равняются 0, 1, ..., $(n - 1)$. Но если в графе есть вершина степени 0, то в нём не может быть вершины степени $(n - 1)$: ведь такая вершина степени $(n - 1)$ должна быть соединена со всеми вершинами, а с одной вершиной (у которой степень 0) она точно не соединена. Получили противоречие — значит, все n степеней не могут быть попарно различными. Теперь понятно, почему не получается нарисовать граф со степенями вершин 0, 1, 2, 3, 4, 5, 6.

Теорема 2. Сумма степеней всех вершин графа (или мультиграфа) равна удвоенному числу его рёбер.

Доказательство. Каждое ребро графа имеет два конца, а значит, добавляет по единице к каждой из степеней этих двух вершин. Следовательно, общую сумму всех степеней каждое ребро увеличивает на 2, поэтому эта сумма и равна удвоенному числу рёбер.
Из этой теоремы получаем важное следствие.

Следствие. В любом графе (или мультиграфе) количество вершин нечётной степени всегда чётно.

Доказательство. Обозначим сумму всех степеней через S, сумму нечётных степеней — A, сумму чётных степеней — B. По теореме 2 число S чётно. Число B тоже чётно (ведь это сумма чётных чисел). Но тогда A тоже должно быть чётно. При этом, согласно нашим обозначениям, сумма A содержит только нечётные слагаемые. Значит, количество этих слагаемых чётно — что и требовалось доказать.

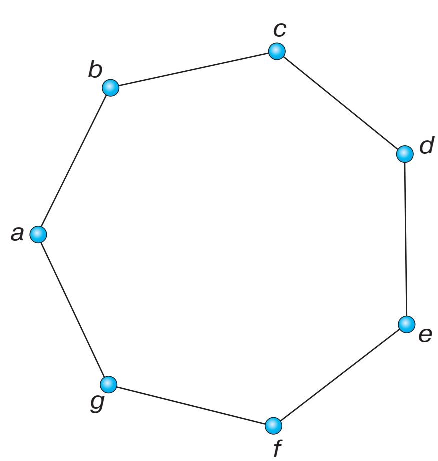


Рис. 36. Простой цикл

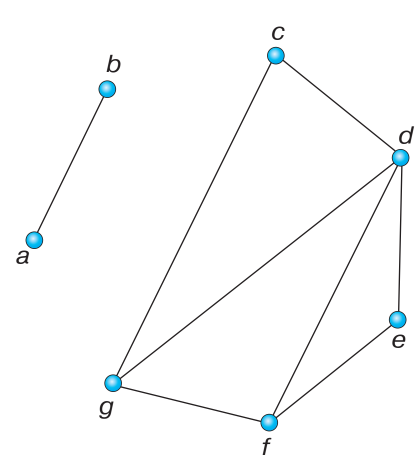


Рис. 37. Несвязный граф

Это следствие объясняет, почему нельзя построить граф с 7 вершинами, все степени которых равны 3.

Рассмотрим ещё два графа с 7 вершинами, описанные в начале этого пункта, которые построить несложно (возможно, вам это уже удалось). Чтобы получить граф, все вершины которого имеют степень 2, достаточно соединить вершины «по кругу»: первую — со второй, вторую — с третьей и так далее до последней. Последнюю вершину нужно соединить с первой. Такой граф называется ***простым циклом*** и показан на рисунке 36 (мы ещё поговорим о циклах в следующем пункте).

Граф со степенями 1, 1, 2, 2, 3, 3, 4 построить чуть сложнее — он показан на рисунке 37.

Кажется, что здесь нарисован не один, а два графа, но на самом деле бывают и такие графы, которые состоят из нескольких частей. Они называются ***несвязными***.

? ВОПРОСЫ

1. Что называется степенью вершины?
2. Если у простого графа 12 вершин, в каких границах могут изменяться их степени?
3. Существует ли граф с тремя вершинами, степени которых равны 0, 1 и 2?
4. Может ли сумма степеней всех вершин графа равняться 13?
5. Если у графа 10 рёбер, чему равна сумма степеней всех его вершин?

3 Пути, цепи и циклы

Если граф представляет собой транспортную сеть, то одна из главных задач, которую можно решить с его помощью, — найти ***путь*** от одной вершины графа к другой. Именно для этих целей в каждом вагоне метро висит схема метрополитена.

Понятие пути между вершинами важно и для других графов.

Путём или маршрутом в графе (или мультиграфе) называется такая последовательность вершин, в которой каждая вершина соединена со следующей за ней вершиной ребром.

Длиной пути называется количество рёбер (оно на 1 меньше количества вершин в пути).

Чтобы описать путь, достаточно просто по порядку перечислить его вершины. Точно так же любой туристический маршрут задаётся перечислением населённых пунктов, через которые он проходит, например: Москва — Ростов — Ярославль — Суздаль — Владимир.

Для графа на рисунке 38 последовательность *abcdeg* будет маршрутом, а *abcdef* нет, поскольку вершины *e* и *f* не соединены ребром.

Обратите внимание, что вершины (и даже рёбра), образующие путь, могут повторяться: например, в маршруте *bcfba* мы дважды заходим в вершину *b*, а в маршруте *abaf* дважды проходим по ребру *ab*.

Цепью в графе (или мультиграфе) называется любой путь без повторяющихся рёбер.

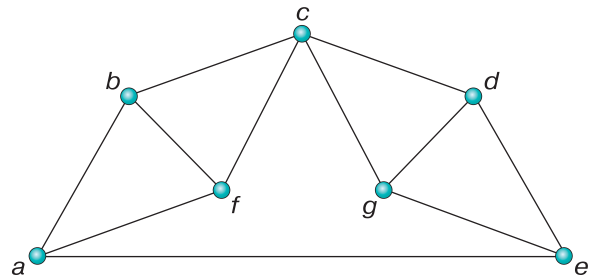


Рис. 38. Пути на графе

Из приведённых выше примеров пути *abcdeg* и *bcfba* — это цепи, а путь *abaf* нет.

Несложно доказать, что если между какими-то двумя вершинами графа есть путь, то есть и цепь, которая их соединяет (это задание будет в упражнении 78).

Ещё одна важная разновидность пути играет в теории графов особую роль.

Циклом называется цепь, в которой начальная и конечная вершина совпадают.

Приведём примеры циклов на рисунке 38:

abfa, *abcfa*, *bcdgcfb*.

Первый из них имеет длину 3, второй — 4, третий — 6 (напомним, что длиной пути называется количество составляющих его рёбер). Обратите внимание на важную деталь: рёбра в цикле повторяться не могут — ведь тогда он не будет цепью. Поэтому, скажем, путь *aba* хотя и замкнутый, но циклом считаться не может.

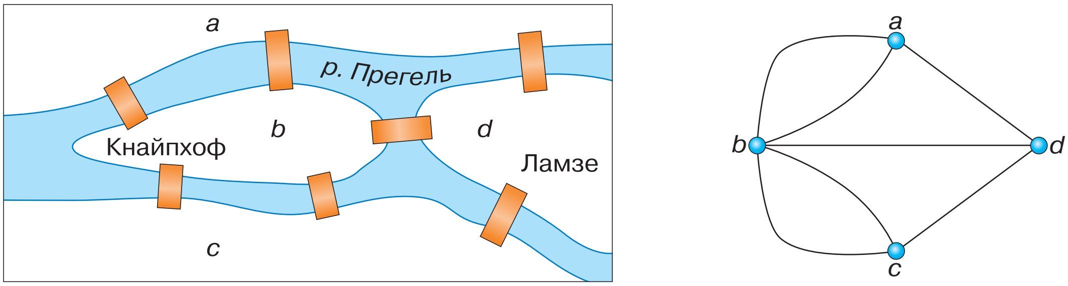
Повторение вершин в циклах допускается, причём не только начальной и конечной. Например, в цикле *bcdgcfb* мы дважды проходим через вершину *c*. Если всё-таки все вершины цикла различны (кроме начальной и конечной), то такой цикл называется ***простым***.

В начале этой главы мы уже упоминали о задаче, с которой началась теория графов, — ***задаче о кёнигсбергских мостах***. Теперь мы можем сформулировать, в чём она состоит.

Город Кёнигсберг (современный Калининград), расположенный на берегах реки Прегель, с раннего Средневековья славился своими мостами — уже к XVII в. их насчитывалось семь. На рисунке 39, *а* вы видите карту с их изображением. Жителей города давно интересовал вопрос: можно ли, выйдя из дома в любом районе Кёнигсберга, прогуляться по каждому из семи мостов ровно один раз (в одном из двух направлений) и вернуться домой?

Если сопоставить каждому из четырёх районов Кёнигсберга вершину графа, а каждому из семи мостов — соответствующее ребро, то получится мультиграф, изображённый на рисунке 39, *б*. С использованием введённых выше понятий задачу о кёнигсбергских мостах можно сформулировать так: существует ли в этом мультиграфе цикл, содержащий все его рёбра?

Такие циклы называют теперь ***эйлеровыми*** в честь великого швейцарского, прусского и российского математика Леонарда Эйлера (1707—1783),



а) карта Кёнигсберга (XVIII в.)　　б) мультиграф мостов

Рис. 39

опубликовавшего решение задачи о мостах в 1736 г. В своей статье «Решение вопроса, связанного с геометрией положения» он доказал, что такого цикла в этом графе не существует.

Как часто бывало в истории математики, Эйлер не просто решил интересную, но частную задачу, а доказал более общий результат, который стал первой теоремой в теории графов.

Теорема Эйлера. Эйлеров цикл в связном графе (или мультиграфе) существует тогда и только тогда, когда все вершины графа (или мультиграфа) имеют чётные степени.

В нашей задаче все четыре вершины имеют нечётные степени, поэтому условие теоремы Эйлера не выполняется, и, следовательно, эйлерова цикла в этом графе не существует.

Эти теоремы вы докажете в упражнениях 79, 80 в конце этого раздела.

Заметим ещё, что если в графе ровно две вершины имеют нечётную степень, а все остальные степени чётные, то эйлерова цикла в нём нет, но зато есть эйлеров путь. Так называется незамкнутый маршрут, который проходит по каждому ребру графа ровно один раз. Начинается и заканчивается он как раз в тех двух вершинах, которые имеют нечётную степень.

? ВОПРОСЫ

1. Дайте определение пути в графе.
2. Что такое цепь?
3. Что такое цикл?
4. Какой цикл называется эйлеровым?
5. Сформулируйте теорему Эйлера.

УПРАЖНЕНИЯ*

67. Определите, существует ли описанный ниже граф. Если да, то постройте его:
а) граф из семи вершин, в котором все вершины имеют степень 2;
б) граф из восьми вершин, в котором все вершины имеют степень 2;
в) граф из восьми вершин, в котором все вершины имеют степень 1;
г) граф из семи вершин, в котором все вершины имеют степень 1;
д) граф из семи вершин, в котором все вершины имеют степень 3;
е) граф из восьми вершин, в котором все вершины имеют степень 3.

68. Петя вбил в землю 5 колышков и соединил некоторые из них верёвками. Могло ли так получиться, что к каждому колышку привязано ровно по 3 верёвки?

69. На олимпиаде по математике каждый из 30 участников решил по 4 задачи, а каждую задачу решило ровно 10 человек. Сколько задач было на олимпиаде?

70. В чемпионате города по футболу участвует 12 команд. Чемпионат проводится в один круг. Докажите, что в любой момент проведения чемпионата всегда найдутся хотя бы две команды, сыгравшие одинаковое число матчей.

71. На лавке сидят семеро ребят. Каждый имеет среди остальных не менее трёх братьев. Докажите, что все семеро — братья.

72. Какие из графов, изображённых на рисунке 40, содержат эйлеровы циклы? Найдите эти циклы.

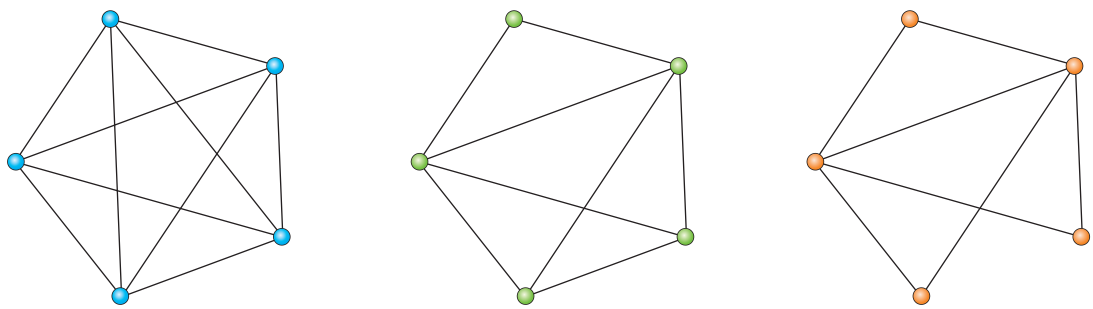


Рис. 40 а) б) в)

73. Какие из фигур, изображённых на рисунке 41, можно нарисовать, не отрывая карандаш от бумаги и не проводя никакую линию дважды? Если возможно, покажите, как это сделать.

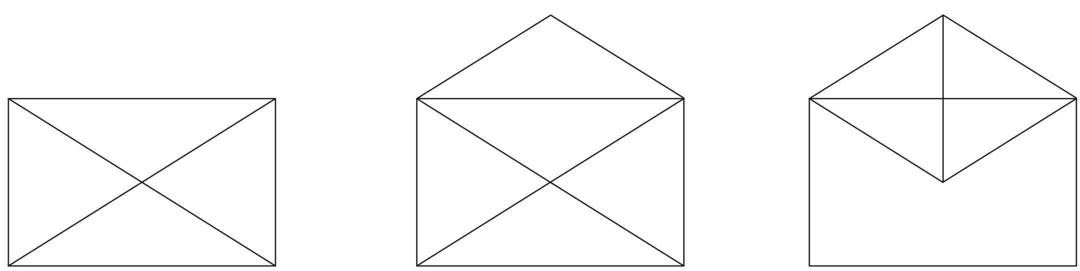


Рис. 41 а) б) в)

74. Какое максимальное число кёнигсбергских мостов можно пройти по одному разу и вернуться в исходную точку? А если не требовать возвращения в исходную точку?

75. Можно ли из куска проволоки длиной 120 см изготовить каркас куба с ребром 10 см? Проволоку разрешается сгибать, но не разламывать.

76. Можно ли раскрасить рёбра куба в красный и зелёный цвета так, чтобы муравей мог пройти из любой вершину в любую только по красным рёбрам, а жук — только по зелёным?

77. **Задача Рамсея.** Докажите, что в любой компании из 6 человек найдутся либо трое попарно незнакомых друг с другом, либо трое попарно знакомых.

78. Докажите, что если между двумя различными вершинами графа есть путь, то существует и цепь, которая их соединяет.

79. Докажите, что если мультиграф содержит эйлеров цикл, то степени всех его вершин чётные (необходимое условие в теореме Эйлера).

80. Докажите, что если степени всех вершин мультиграфа чётные, то в нём есть эйлеров цикл (достаточное условие в теореме Эйлера).

§ 4. Виды графов

1 Связные графы

Вернёмся к одному из графов, который мы рассмотрели в начале этой главы, — схеме новосибирского метро. Этот граф обладает одним замечательным свойством: от любой его вершины можно пройти по рёбрам до любой другой вершины (от любой станции можно доехать до любой другой — возможно, с пересадкой, но при этом не выходя из метро!). Такой граф называется ***связным***.

Граф называется связным, если для любых двух его вершин найдётся хотя бы один путь, который их соединяет.

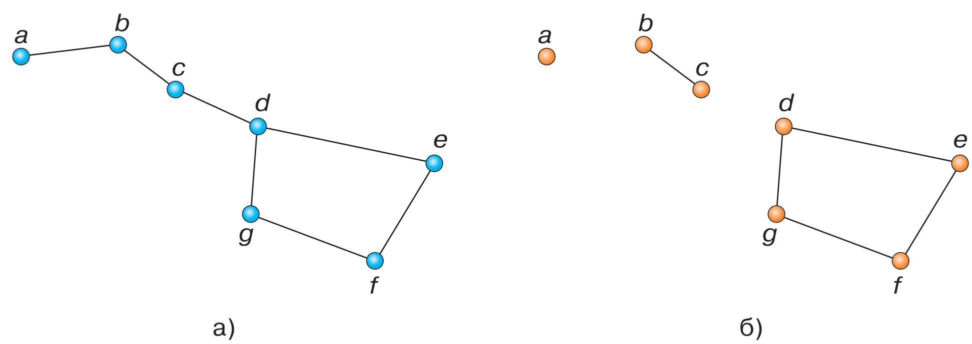


Рис. 42. Примеры связного и несвязного графов

Граф на рисунке 42, *а* — связный, а на рисунке 42, *б* — несвязный: он состоит из трёх ***связных компонент*** (говорят ещё, что у него три компоненты связности).

Примером несвязного графа может служить граф новосибирских трамвайных путей: он состоит из двух связных компонент, расположенных на разных берегах реки Оби (рис. 43).

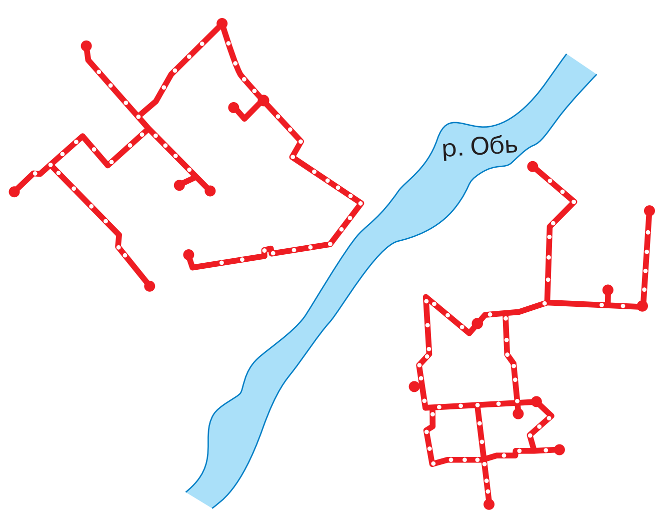


Рис. 43. Схема трамвайных путей в Новосибирске

Понятно, что на связность влияет количество соединений, т. е. рёбер: чем больше рёбер, тем «более связным» бывает обычно граф. Если у графа n вершин и вовсе нет рёбер (и такое бывает!), то каждая его вершина — это отдельная компонента связности. Это пример «самого несвязного» графа.

Противоположный случай: представьте себе, что мы соединили ребром каждую пару вершин. Тогда, конечно, такой граф будет связным. Он имеет специальное название — ***полный граф***.

На рисунке 44 изображён полный граф, состоящий из пяти вершин. Этот граф содержит 10 рёбер.

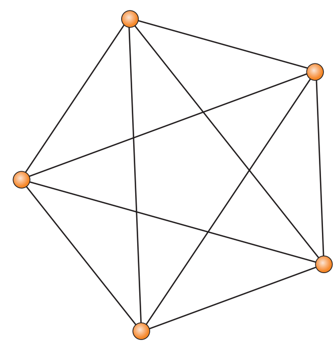


Рис. 44. Полный граф из пяти вершин

Можно доказать (см. упражнение 89 после этого раздела), что у полного графа с n вершинами будет $\dfrac{n(n-1)}{2}$ рёбер.

Свойство связности графа выделяет среди всех рёбер некие «критические» рёбра, которые обеспечивают эту связность. Они называются ***мостами***.

Ребро связного графа называется мостом, если после его удаления граф становится несвязным (т. е. распадается на две компоненты связности).

Как при разрушении моста через реку становится невозможным попасть с одного берега реки на другой, так и при удалении моста в графе оказываются не связанными между собой две его части. На рисунке 45 приведены три графа: у первого из них (рис. 45, *а*) три ребра являются мостами (они выделены оранжевым цветом), у второго графа (рис. 45, *б*) все рёбра — мосты, а у третьего (рис. 45, *в*) вообще нет ни одного моста.

Мосты в графе связаны с циклами: **ребро является мостом тогда и только тогда, когда оно не принадлежит ни одному циклу.** В самом деле, если удалить ребро, которое связывает какие-то вершины x и y и входит в цикл, то связь между этими вершинами не нарушится: от x к y можно будет пройти по другой части цикла.

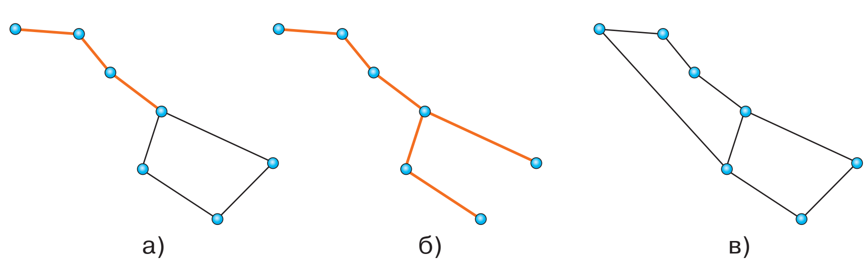


Рис. 45. Мосты в графах

? ВОПРОСЫ

1. Какой граф называется связным?
2. Что такое полный граф?
3. Сколько рёбер содержит полный граф с 4 вершинами?
4. Какие рёбра называются мостами?

2 Деревья

На рисунке 46 изображена сеть авиалиний некоторого воображаемого государства.

Представим себе, что это государство охватил экономический кризис и часть авиалиний решено было закрыть, но так, чтобы из любого города можно было по-прежнему долететь до любого другого хотя бы с пересадками. На языке графов это означает, что граф авиалиний должен остаться связным. При этом в нём не должно остаться «лишних» авиалиний. Какие же линии можно считать лишними?

В исходном графе есть циклы. Возьмём один из них — например, $ALNMA$. Очевидно, что из него можно выбросить любое из четырёх рёбер — граф всё равно останется связным. Выбросим, например, ребро MN. В полученном графе снова есть цикл — $AHLA$. Выбросим какое-нибудь из его рёбер — например, HL. Поступим так же ещё с двумя циклами.

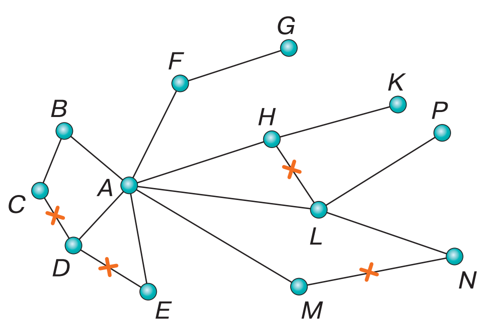


Рис. 46. Сеть авиалиний

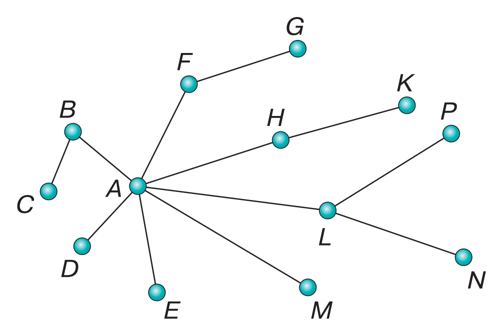


Рис. 47. Дерево

Получился граф, в котором циклов уже нет (рис. 47). Если вы попробуете убрать любое из его рёбер, граф потеряет связность и распадётся на две компоненты — тогда перелететь на самолёте из одной компоненты в другую будет уже невозможно. Такие графы, которые не имеют лишних рёбер, называются ***деревьями***, и возникают в самых разных задачах.

Деревом называется связный граф без циклов.

Если дерево, как в нашем примере, получено из некоторого графа вычёркиванием «лишних» рёбер, то оно называется ***остовным деревом***, ***остовом*** или ***каркасом*** данного графа. Можно сказать, что дерево — это связный граф, у которого нет «лишних» рёбер: если выкинуть любое из его рёбер, нарушится связность. Это утверждение можно сформулировать в виде следующей теоремы.

■ **Теорема 1.** В дереве каждое ребро является мостом.

Доказательство. Доказательство проведём ***методом от противного***. Предположим, что в дереве есть какое-то ребро xy, которое не является мостом. Это значит, что после его удаления дерево остаётся связным, а значит, найдётся цепь, которая связывает вершины x и y. Добавим к этой цепи ребро xy — получим цикл. Но это противоречит условию того, что граф является деревом. Полученное противоречие показывает, что все рёбра дерева являются мостами.

Со связностью и отсутствием циклов связано ещё одно важное свойство дерева.

■ **Теорема 2.** Между любыми двумя вершинами дерева существует единственная цепь, которая их соединяет.

Существование такой цепи следует из того, что любое дерево является связным графом.

Для доказательства единственности снова воспользуемся методом от противного. Предположим, что найдутся такие вершины x и y, что между ними есть две разные цепи (рис. 47). Пусть эти цепи расходятся в вершине p, а потом сходятся в вершине q (возможно, $p = x$ и $q = y$).

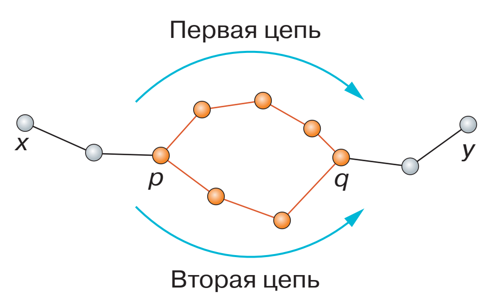


Рис. 48

Но тогда мы можем составить из них цикл, который выделен на рисунке 48. Это противоречит условию теоремы, поскольку в дереве не может быть циклов. Значит, двух разных цепей между вершинами x и y тоже быть не может.

Дерево, которое получилось из сети авиалиний на рисунке 47, содержит 13 вершин и 12 рёбер. Оказывается, это соотношение между количеством вершин и рёбер не случайное.

■ **Теорема 3.** Дерево из n вершин содержит $(n - 1)$ рёбер.

Если вы ещё не доказывали эту теорему в основной школе, обязательно выполните упражнение 90 после этого раздела.

? ВОПРОСЫ

1. Какой граф называется деревом?
2. Сформулируйте известные вам свойства деревьев.
3. Сколько рёбер в дереве, у которого 2024 вершины?
4. Сколько вершин в дереве, у которого 1000 рёбер?

3 Дерево случайного эксперимента

У вас, наверное, уже возник вопрос: почему связные графы без циклов стали называть деревьями — ведь граф, изображённый на рисунке 49, мало похож на дерево. Сейчас мы в этом разберёмся.

Дело в том, что во многих задачах на дереве часто выделяют одну, особую вершину, которую называют ***корнем***. И тогда дерево рисуют обычно так: выше всех остальных вершин изображается вершина-корень, на ярус ниже — её соседи, ещё ниже — соседи этих соседей и т. д. Получается, что мы подняли дерево за корень и слегка его «встряхнули».

Дерево, в котором одна из вершин выбрана в качестве корня, называется ***корневым деревом***. На рисунке 50 в качестве корня выбрана вершина a.

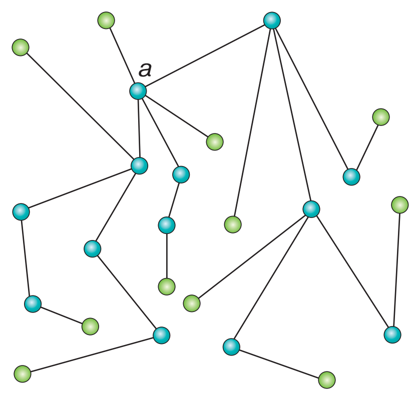


Рис. 49. Пример дерева

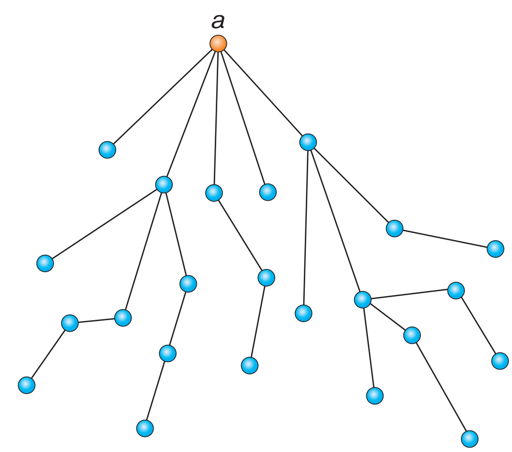


Рис. 50. Дерево с выделенным корнем a

В соответствии с этой терминологией конечные вершины дерева, имеющие степень 1, называют ***листьями***. На рисунке 49 дерево имеет 11 листьев (все они закрашены в зелёный цвет).

Примером дерева с выделенным корнем может служить дерево вложенных друг в друга папок на диске вашего компьютера. Самая верхняя папка на диске так и называется — корневая.

Ещё один пример корневого дерева — это ***дерево случайного эксперимента***. Такое дерево удобно использовать для решения задач, в которых случайный эксперимент состоит из нескольких этапов. Например, несколько раз бросают монету или кубик, вытаскивают друг за другом несколько шаров из коробки и т. д.

Корню дерева соответствует начало эксперимента. Вершинами служат состояния, в которые мы переходим на каждом следующем этапе, а рёбрами — соответствующие переходы из одного состояния в другое.

Листья дерева — это конечные состояния, которые представляют собой все возможные исходы данного эксперимента. Посмотрим, как выглядит такое дерево на конкретных примерах.

Пример 1. Рассмотрим опыт, в котором три раза подряд бросают монету. Он состоит из трёх шагов, на каждом из которых проводится очередное испытание с монетой. У каждого такого испытания два возможных исхода. Если взять в качестве корня начальное состояние, когда монету ещё ни разу не бросали, то получим дерево, изображённое на рисунке 51.

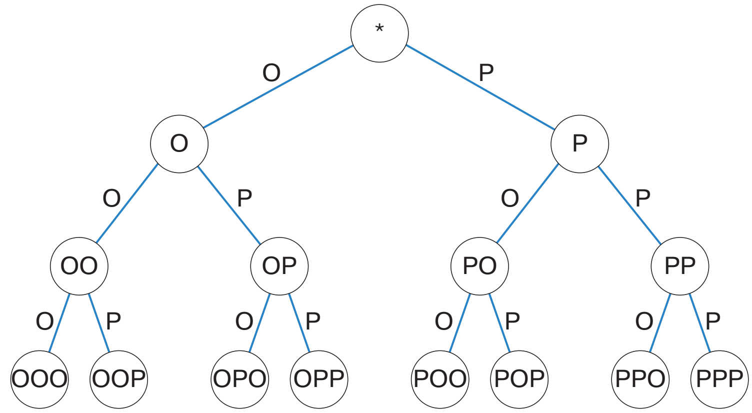


Рис. 51. Три испытания с монетой

Корень дерева отмечен символом «*». На первом уровне находятся две вершины: это возможные состояния после первого испытания. На втором уровне их уже четыре, а на третьем — восемь. От каждой вершины, кроме восьми конечных, отходят вниз по два ребра, которые соответствуют двум возможным исходам очередного испытания с монетой. Нижние восемь вершин-листьев отмечают конечные состояния всего эксперимента — это его возможные исходы. Поскольку в данном опыте все они равновозможны, то вероятность каждого из них равна $\frac{1}{8}$.

При увеличении количества шагов дерево начинает быстро расти, и рисовать его становится всё сложнее. Если монету подбрасывать n раз, то дерево будет содержать 2^n листьев. Такое дерево называют ***двоичным***, поскольку у каждой его вершины два потомка — по числу возможных исходов очередного бросания монеты.

Ещё быстрее будет ветвиться дерево эксперимента с кубиком, поскольку у каждой вершины здесь будет уже по шесть потомков.

Сложнее устроены деревья экспериментов, в которых может быть разное количество шагов. Рассмотрим, например, такую задачу.

Пример 2. Монету подбрасывают до появления орла, но не больше трёх раз. С какой вероятностью орёл при этом появится?

На этот раз эксперимент может закончиться на первом, втором или третьем шаге. Дерево возможных вариантов изображено на рисунке 52.

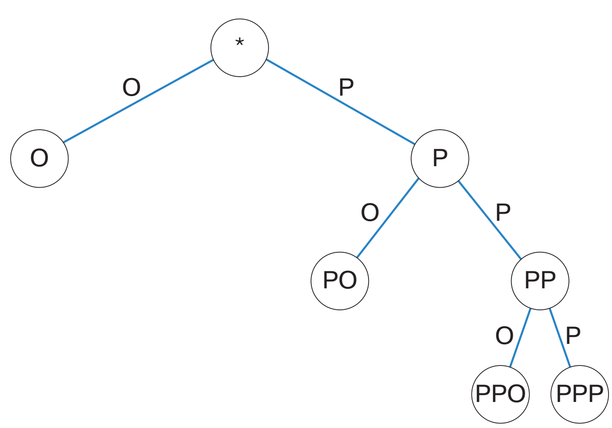


Рис. 52. Три испытания с монетой до первого орла

У этого дерева 4 листа, которые соответствуют четырём возможным исходам эксперимента:

О, РО, РРО, РРР.

Считать эти исходы равновозможными, конечно, уже нельзя. Очень скоро мы научимся вычислять вероятность каждого из этих исходов с помощью специального ***правила умножения***. Пока можно рассуждать так:

– исход О — один из двух равновозможных исходов при подбрасывании монеты, поэтому его вероятность равна $\frac{1}{2}$;

– исход РО — один из четырёх равновозможных исходов при двукратном бросании монеты, поэтому его вероятность равна $\frac{1}{4}$;

– исходы РРО и РРР — это два из восьми равновозможных исходов при троекратном бросании монеты, поэтому вероятность каждого из них равна $\frac{1}{8}$.

Таким образом, мы нашли вероятности каждого из четырёх исходов. Поскольку вероятность любого события равна сумме вероятностей благоприятных для него исходов, то вероятность появления орла равна $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$.

Заметим, что тот же самый ответ можно получить с помощью дерева из примера 1, если рассмотреть событие «в трёх испытаниях появится хотя бы один орёл». Тогда из восьми равновозможных исходов (восемь листов дерева на рисунке 51) семь будут благоприятными для нашего события и его вероятность тоже будет равна $\frac{7}{8}$.

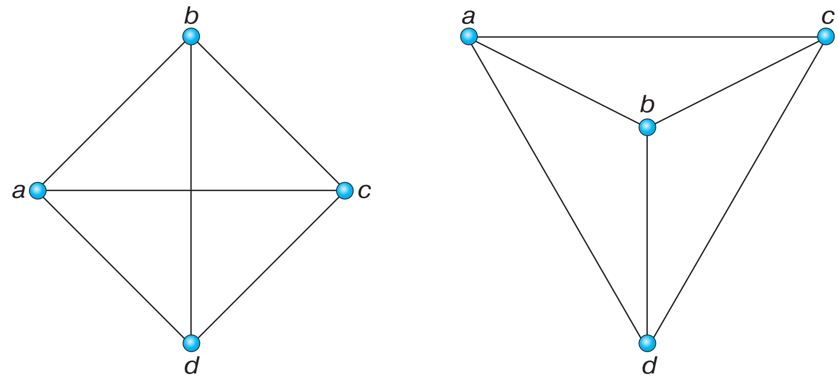
Конечно, в таком решении мы немного изменили условия опыта. Но ведь вероятность события не может измениться, если после появления орла мы бросим монету ещё один или два раза, чтобы довести число испытаний до трёх. В теории вероятностей, да и в математике в целом, при решении задач можно использовать разные модели — главное, чтобы их выбор не приводил к неверным выводам и ответам.

? ВОПРОСЫ

1. Что такое корневое дерево?
2. Какие вершины дерева называются листьями?
3. Как построить дерево случайного эксперимента?
4. Постройте дерево какого-нибудь случайного эксперимента.

4 Планарные графы

Изображая граф на плоскости, его стараются нарисовать так, чтобы рёбра не имели других общих точек, кроме вершин, — тогда легче разобраться в структуре соединений. На рисунках 53, *а* и 53, *б* изображён один и тот же граф: в первом случае с пересечением рёбер, во втором — без пересечения. Так же как у одного и того же человека могут быть разные фотографии, у одного и того же графа могут быть разные изображения: одни — более удачные, другие — менее.

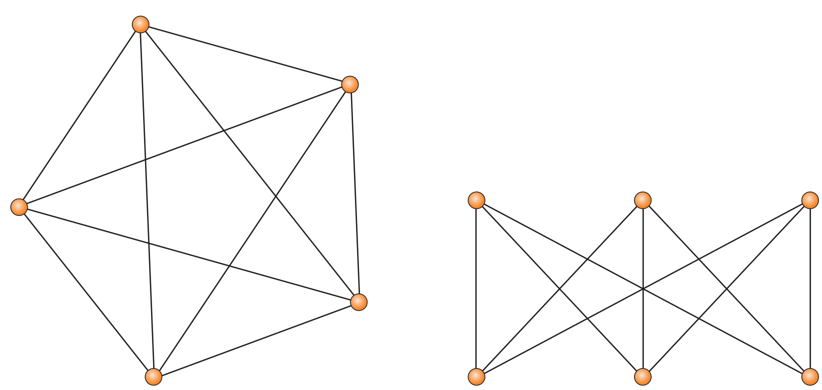


а) с пересечением рёбер б) без пересечения рёбер

Рис. 53. Разные изображения одного и того же графа

Если граф является деревом, то его всегда можно изобразить на плоскости без пересечения рёбер. Для этого достаточно, например, выбрать одну вершину в качестве корня и дальше рисовать вершины по уровням, постепенно опускаясь всё ниже и всё дальше уходя от корня.

Но изобразить произвольный граф без пересечения рёбер не всегда удаётся: например, для графов на рисунке 54 это сделать невозможно (можете попробовать). Второй из этих графов наверняка знаком вам по детской ***головоломке о трёх домиках и трёх колодцах***: нужно проложить от каждого из трёх домиков (верхние три вершины на рисунке 54, *б* тропинки к каждому из трёх колодцев (нижние три вершины) так, чтобы они не пересекались.



а) б)

Рис. 54. Эти графы невозможно нарисовать без пересечения рёбер

Сколько бы вы ни пытались прокладывать тропинки, самая последняя всегда будет пересекать какие-то из уже проложенных. Но из этого ещё не следует, что сделать этого нельзя: возможно, вы просто плохо искали.

Только в первой половине прошлого века (по меркам истории математики — совсем недавно) советский математик Л. С. Понтрягин (1908—1988) и польский математик К. Куратовский (1896—1980) независимо друг от друга доказали замечательную теорему, из которой следовало, что проложить тропинки в этой задаче невозможно. Неожиданно значение этой теоремы оказалось намного важнее, чем решение головоломки. Чтобы её сформулировать, нам понадобится следующее определение.

Граф называется планарным, если его можно нарисовать на плоскости так, чтобы его рёбра попарно не пересекались (разумеется, пересечения в вершинах при этом не считаются).

Изображение планарного графа без пересечения рёбер называется ***плоским*** графом.

Теорема Понтрягина — Куратовского утверждает, что двумя графами, приведёнными на рисунке 54, в некотором смысле исчерпываются все непланарные графы. Более точно это теорема утверждает, что:

– во-первых, графы на рисунках 54, *а* и 54, *б* непланарные;
– во-вторых, любой непланарный граф содержит один из этих двух графов в качестве своей части.

Одновременно с доказательством теоремы были найдены алгоритмы, которые строят изображения планарных графов без пересечения рёбер. Эти алгоритмы нашли неожиданное применение в современной микроэлектронике. При проектировании печатных плат нужно соединять проводниками входы и выходы микроэлементов так, чтобы проводники не пересекались. Если схема соединения удовлетворяет условиям теоремы Понтрягина — Куратовского, то с помощью такого алгоритма печатную плату можно сделать однослойной, если нет, то приходится располагать её элементы в разных слоях.

? ВОПРОСЫ

1. Какие графы называются планарными?
2. Нарисуйте какой-нибудь плоский граф.
3. Приведите пример непланарного графа.
4. Являются ли все деревья планарными графами?

5 Теорема Эйлера

С планарностью графов связан ещё один замечательный результат, полученный всё тем же Леонардом Эйлером совсем в другой области математики. В 1750 г. он доказал теорему, которую вы будете доказывать в курсе геометрии.

Теорема Эйлера о многогранниках. Если B — число вершин, P — число рёбер, а Γ — число граней выпуклого многогранника, то

$$B - P + \Gamma = 2.$$

Доказательство этой замечательной теоремы основано на аналогичном результате для плоских графов. Чтобы сформулировать этот результат, нам понадобится ещё одно понятие, связанное с плоским изображением планарного графа.

Если нарисовать планарный граф на плоскости без внутренних пересечений рёбер, то вся плоскость разделится на отдельные, не связанные друг с другом области, которые называются ***гранями*** плоского графа.

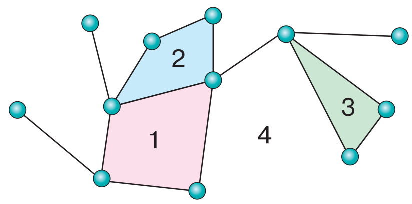


Рис. 55. Грани плоского графа

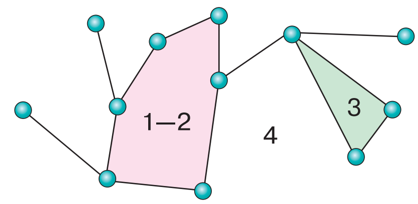


Рис. 56. Слияние двух смежных граней

На рисунке 55 изображён граф, который разбивает плоскость на четыре такие области. Заметим, что одна из граней всегда будет неограниченной (грань с номером 4).

Конечно, фраза «не связанные друг с другом области» нуждается в уточнении, но интуитивно понятно, что это значит: внутри одной грани любые две точки можно соединить друг с другом непрерывной кривой, не пересекающей рёбра, а две точки из разных граней — нельзя.

Докажем следующую теорему.

Теорема Эйлера о плоских графах. Для любого связного плоского графа справедливо равенство

$$В - Р + Г = 2,$$

где $В$ — число вершин, $Р$ — число рёбер, а $Г$ — число граней плоского графа.

Поскольку граф связный, то он либо является деревом, либо содержит цикл.

Для дерева, как мы знаем,

$$Р = В - 1,\ Г = 1,$$

поэтому приведённое в теореме равенство выполнено.

Если в графе есть цикл, то возьмём любое ребро, которое содержится в этом цикле, и удалим его. Связность при этом сохранится. Сохранится и значение выражения $В - Р + Г$: ведь и рёбер, и граней стало на 1 меньше. Мы пользуемся здесь интуитивно понятным фактом, что если стереть ребро, входящее в цикл, то две грани, находящиеся внутри и снаружи цикла, сольются в одну грань (рис. 56).

Продолжая этот процесс «уничтожения циклов», мы рано или поздно придём к дереву, в котором, как уже было сказано, $В - Р + Г = 2$. Поскольку значение этого выражения в течение всего процесса не изменялось, то, значит, и для исходного графа оно было равно 2. Теорема доказана.

Чтобы получить из теоремы о плоских графах теорему о многогранниках, построим для выпуклого многогранника плоский граф. Сделать это не так просто, как кажется. Понятно, что вершинами такого графа будут вершины многогранника, а рёбрами графа — рёбра многогранника. Но как нарисовать такой граф на плоскости?

Представим себе, что мы выбрали одну из граней многогранника (например, заднюю) и вырезали её, а оставшуюся поверхность многогранника растянули на плоскости (возможно, с деформацией рёбер, но без разрывов). Полученный плоский граф и будет искомым. При этом удалённой грани соответствует неограниченная грань плоского графа.

Поскольку для графа формула Эйлера доказана, то она автоматически выполнена и для многогранника.

ВОПРОСЫ

1. Запишите формулу Эйлера для многогранников.
2. Проверьте выполнение формулы Эйлера для известных вам многогранников.
3. Что называется гранью плоского графа?
4. Запишите формулу Эйлера для плоских графов.

УПРАЖНЕНИЯ*

81. В некоторой стране железнодорожная сеть устроена так, что можно попасть из любого города в любой другой (возможно, через другие города). Докажите, что всегда можно выбрать такой город, что если закрыть на ремонт все ведущие в него дороги, то всё равно можно будет попасть из любого города в любой другой (кроме самого выбранного города).

82. Сколько рёбер нужно удалить из графа, изображённого на рисунке 57, чтобы получилось дерево? Удалите их и перерисуйте полученное дерево в тетрадь.

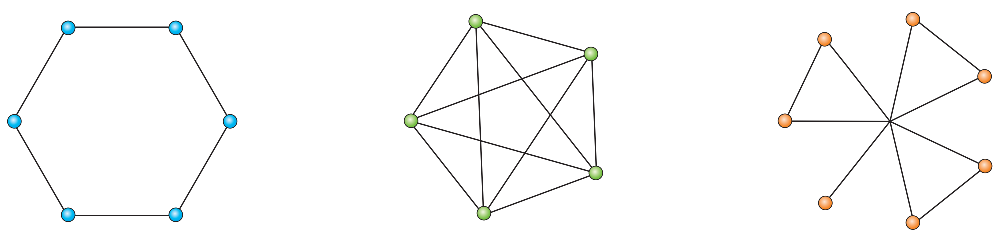


Рис. 57 а) б) в)

83. Сколько рёбер нужно добавить к графу, изображённому на рисунке 58, чтобы получилось дерево? Добавьте их и перерисуйте полученное дерево в тетрадь.

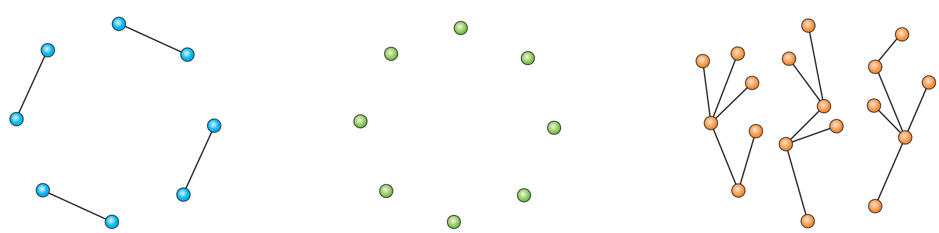


Рис. 58 а) б) в)

84. Если дерево с заданными степенями вершин существует — постройте его, если нет — объясните, почему такого дерева не существует:

а) 1, 1, 2, 2, 2; в) 1, 1, 1, 1, 1;

б) 2, 2, 2, 3; г) 1, 1, 1, 1, 1, 1, 2, 4.

85. В деревне 20 домов, между некоторыми из которых проложены тропинки. По тропинкам можно пройти от любого дома к любому другому и притом единственным образом. Сколько всего тропинок проложено в деревне?

86. Все 30 городов некоторого государства соединены прямыми авиалиниями. Правительство решило сократить число авиалиний. Какое наибольшее их количество можно закрыть, чтобы можно было по-прежнему долететь из любого города в любой другой, но, возможно, с пересадками?

87. В стране 13 городов, каждый из которых соединён авиасообщением с 6 другими. Докажите, что из любого города можно добраться в любой другой (возможно, с пересадками).

88. После экзамена 10 одноклассников обсуждали решение самой трудной задачи, и каждый из них выяснил, что его ответ совпал как минимум с четырьмя одноклассниками. Можно ли быть уверенным, что все они дали одинаковый ответ?

89. Докажите, что в полном графе с n вершинами количество рёбер равно $\frac{n(n-1)}{2}$.

90. Докажите, что в дереве с n вершинами количество рёбер равно $(n-1)$.

91. Какое наименьшее и какое наибольшее количество листьев может иметь дерево, у которого 5 вершин? 50 вершин? 2024 вершины?

92. Сколько рёбер в связном графе с n вершинами, если в нём имеется единственный цикл?

93. В некотором графе все вершины имеют степень 3. Докажите, что в нём есть цикл.

94. Известно, что дерево имеет 3 вершины степени 3 и 4 вершины степени 2. Остальные вершины имеют степень 1. Сколько этих вершин?

95. В дереве имеется 40 вершин степени 4, а все остальные вершины — листья. Сколько этих листьев?

96. Волейбольная сетка имеет вид прямоугольника 50×600 клеток. Какое наибольшее количество верёвочек можно перерезать так, чтобы сетка не распалась на куски?

97. В коробке лежат два белых и два чёрных шара. Из неё один за другим достают два шара. Нарисуйте дерево этого эксперимента. Сколько листьев у этого дерева? Сколько равновозможных исходов у этого эксперимента?

98. Из ящика, в котором лежат две одинаковые пары перчаток, одну за другой достают три перчатки. Нарисуйте дерево этого эксперимента. Сколько листьев у этого дерева? Сколько равновозможных исходов у этого эксперимента?

99. Игра в бадминтон идёт до победы в двух геймах. Перечислите с помощью дерева все возможные варианты изменения счёта в игре. Сколько всего таких вариантов получилось?

100. Перерисуйте графы, изображённые на рисунке 59, так, чтобы их рёбра не пересекались.

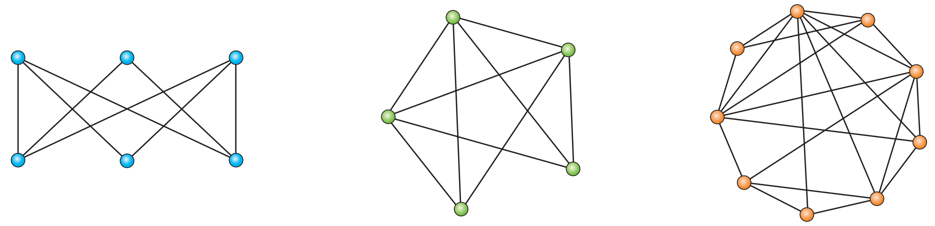


Рис. 59 а) б) в)

101. Какие из графов на рисунке 60 планарные? Свой ответ обоснуйте.

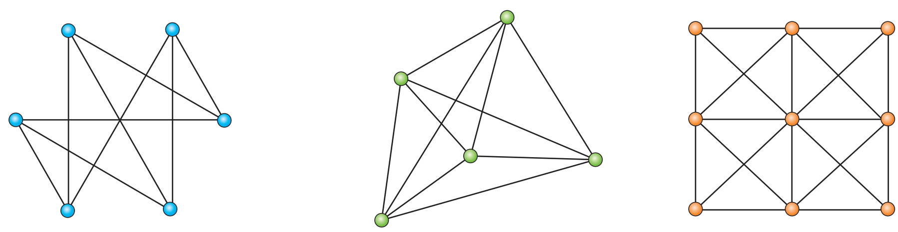


Рис. 60 а) б) в)

102. Какое наибольшее число рёбер может быть у планарного графа с 5 вершинами? с 6 вершинами? Свои ответы обоснуйте.

ГЛАВА 3

СЛУЧАЙНЫЕ СОБЫТИЯ И ИХ ВЕРОЯТНОСТИ

Если подбросить игральный кубик один раз, то вряд ли можно гарантировать, что на нём выпадет шесть очков. Это событие случайное: оно может как произойти, так и не произойти. Но если повторить такой эксперимент 100 раз, то хотя бы одна шестёрка выпадет наверняка. Больше того, при увеличении числа опытов доля шестёрок будет неуклонно приближаться к $\frac{1}{6}$ — числу, которое называют *вероятностью* этого случайного события.

Таким образом, даже в случайных опытах при их многократном повторении начинают проявляться закономерности. Изучает их теория вероятностей. Как и в ситуации с теорией графов, задачи на вычисление вероятностей возникли ещё в XVII в., но в самостоятельную математическую дисциплину теория вероятностей оформилась меньше ста лет тому назад. Произошло это во многом благодаря трудам российских математиков: П. Л. Чебышёва (1821—1894), А. А. Маркова (1856—1922), А. Н. Колмогорова (1903—1987) и других.

В этой главе мы кратко повторим некоторые важные понятия, которые используются в теории вероятностей и, возможно, уже знакомы вам по основной школе.

§ 5. Случайные события

1 Случайный опыт и случайные события

Случайным опытом или ***случайным экспериментом*** называется опыт, результат которого невозможно точно предсказать. Некоторые из таких опытов несложно провести самому: подбросить монету или игральный кубик, вытащить наугад шар из коробки с разноцветными шарами. Ещё больше примеров случайных опытов можно найти в окружающей нас жизни:

- стрельба биатлониста на огневом рубеже;
- жеребьёвка перед проведением спортивного турнира;
- проверка случайно выбранной детали на соответствие стандартам качества;
- выбор билета на экзамене;
- разлив рек во время весеннего половодья;
- измерение кровяного давления;
- взвешивание выбранного на рынке арбуза и т. д.

Важным требованием к любому случайному опыту является возможность его многократного повторения. Это связано с тем, что **закономерности, которые изучает теория вероятностей, начинают проявляться только при большом числе повторных испытаний.**

Случайным событием называется любое событие, связанное со случайным опытом. В результате проведения опыта оно может произойти или не произойти. Например, с подбрасыванием кубика связаны события:

A = {на кубике выпадет чётное число очков};

B = {на кубике выпадет число 4};

C = {на кубике выпадет число больше 4}.

В результате стрельбы биатлониста могут произойти или не произойти события:

D = {все выстрелы попадут в цель};

E = {все мишени будут сбиты}.

Хотя события D и E кажутся одинаковыми, на самом деле они разные. Если вы следите за соревнованиями по биатлону, то знаете, что в эстафете для поражения мишеней биатлонисту даётся несколько запасных патронов, которые он может использовать, если совершил промахи.

С весенним разливом рек связаны случайные события:

F = {будет подтоплен прибрежный район};

G = {уровень подъёма воды не превысит 6 метров};

H = {уровень подъёма воды превысит исторический максимум}.

События F, G и H хотя и являются случайными, но могут быть определённым образом связаны между собой. Если, например, прибрежный район находится на уровне 8 метров выше уровня реки, а исторический максимум равен 10 метрам, то при наступлении события H событие F обязательно наступает, а событие G, наоборот, не наступает.

? ВОПРОСЫ

1. Приведите примеры случайных опытов.
2. Приведите примеры случайных событий, связанных с этими опытами.
3. Какими свойствами обладает случайный опыт?
4. Приведите пример таких случайных событий A и B, что из наступления A обязательно следует наступление B.

2 Исходы и элементарные события

Все возможные результаты случайного опыта называют его ***исходами***. При подбрасывании монеты мы имеем два исхода, которые называют «орёл» и «решка». Для опыта с кубиком этих исходов уже шесть, и обозначать их удобно цифрами от 1 до 6. Если мы извлекаем наугад шар из коробки, то число возможных исходов совпадает с количеством шаров, находящихся в коробке.

Для каждого случайного события все исходы опыта делятся на две категории: при одних исходах это событие наступает, а при других нет. Например, событие A = {на кубике выпадет чётное число очков} наступает, когда выпало 2, 4 или 6 очков, и не наступает при выпадении 1, 3 или 5 очков.

Исходы, при которых событие наступает, называют ***благоприятными*** для этого события. В теории вероятностей принято рассматривать любое случайное событие как **множество благоприятных для него исходов.** При таком подходе событие (как и любое множество) можно или описать словами, или задать перечислением составляющих его элементов, например:

A = {на кубике выпадет чётное число очков} = {2, 4, 6};

B = {на кубике выпадет число 4} = {4};

C = {на кубике выпадет число больше 4} = {5, 6}.

Заметим, что событие B состоит в этом примере из одного исхода. Такие события называются ***элементарными***. Это своеобразные «атомы», мельчайшие события, из которых состоят все остальные события. Опыт всегда завершается наступлением ровно одного из своих элементарных событий.

Формально между исходом опыта и элементарным событием есть некоторая разница: исход — это **элемент множества**, а элементарное событие — **подмножество**, состоящее из одного этого элемента. В приведённом примере с кубиком цифра 4 — исход, а множество {4} — элементарное событие. Однако на практике про это различие часто забывают и считают, что исходы опыта и элементарные события — это одно и то же.

В отличие от опытов с монетой, кубиком или шарами описать все возможные исходы более сложных случайных опытов бывает гораздо труднее. При этом важно сформулировать условия, в которых проводится случайный эксперимент. Например, чтобы говорить о возможных исходах стрельбы биатлониста по мишеням, нужно прежде всего уточнить правила соревнований.

Пример 1. На очередном огневом рубеже биатлонист должен поразить три мишени. Для этого ему даётся 4 патрона, один из которых запасной. Сколько возможных исходов имеет такой случайный опыт?

Если считать исходами число сбитых мишеней, то возможных исходов будет четыре: 0, 1, 2, 3. Но они не совсем «элементарные». Например, сбить 3 мишени можно по-разному: с использованием запасного патрона или без него.

В этом опыте полную информацию об исходе всей стрельбы дают, скорее, результаты каждого выстрела. Обозначим попадание в мишень цифрой 1, а промах — цифрой 0. Тогда исходами опыта будут последовательности из 1 и 0, например:

111 — все три выстрела достигли цели, все мишени сбиты;

0111 — все мишени сбиты, но использован дополнительный патрон;

0000 — все выстрелы прошли мимо и т. д.

Три выстрела будут сделаны только в одном случае: когда все они попали в цель. Значит, остальные исходы будут состоять из четырёх 0 и 1. При этом годятся все такие последовательности, кроме одной: 1110. Ведь, сбив три мишени, биатлонист не будет использовать дополнительный патрон.

Используя полученные исходы, можно выразить через них те случайные события, о которых шла речь выше:

D = {все выстрелы попадут в цель} = {111};

E = {все мишени будут сбиты} = {111, 0111, 1011, 1101}.

Как мы и говорили, события оказались разными: первое содержит только один исход (значит, является элементарным), а второе состоит из 4 исходов.

Пример 2. Регистрируется максимальный уровень подъёма воды во время весеннего разлива реки Оки в районе города Калуги. Что можно считать элементарными исходами этого случайного эксперимента?

Здесь исходом будет число из некоторого промежутка от 0 до M, где M — тот максимальный уровень, выше которого вода точно не поднимется. Чтобы его определить, нужно обратиться к статистике прошлых лет и взять его «с запасом» — например, 20 метров (исторический максимум для Калуги равен 16 м 77 см).

Таким образом, в этом опыте мы имеем дело с ***бесконечным множеством возможных исходов***, что встречается в реальных ситуациях довольно часто.

Как видите, выбор множества элементарных событий является важным этапом в построении ***математической модели*** реального случайного опыта. От того, насколько построенная модель соответствует реальности, во многом будет зависеть правильность полученных в дальнейшем результатов.

? ВОПРОСЫ

1. Опишите множество возможных исходов опыта с подбрасыванием монеты; с подбрасыванием двух монет.
2. Сколько элементарных событий в опыте с подбрасыванием одного кубика? двух кубиков?
3. Приведите пример случайного опыта с бесконечным множеством возможных исходов.
4. Перечислите благоприятные исходы для события A = {выпадет простое число} в опыте с подбрасыванием кубика.

3 Частота и вероятность

Вспомним теперь, что такое ***частота*** и что такое ***вероятность*** случайного события.

Поскольку любой случайный опыт можно проводить (или наблюдать) многократно, представим, что мы провели серию из N таких опытов. Если в этой серии случайное событие A произошло N_A раз, то дробь $\frac{N_A}{N}$ называют ***частотой*** события A в этой серии. Иногда этот термин уточняют: указанную дробь называют ***относительной частотой***, а целое число N_A — ***абсолютной частотой*** события A. Аналогично определяется и частота любого исхода — это доля опытов, завершившихся этим исходом.

Пример 1. Представим, что вы подбросили кубик 20 раз и получили следующие результаты (табл. 44):

Таблица 44

Число на кубике	1	2	3	4	5	6
Количество выпадений	3	3	2	5	3	4

Чтобы посчитать частоту каждого из 6 исходов этого опыта, достаточно поделить количество его выпадений на 20 (табл. 45):

Таблица 45

Число на кубике	1	2	3	4	5	6
Количество выпадений	3	3	2	5	3	4
Частота	0,15	0,15	0,10	0,25	0,15	0,2

Если продолжить эту серию опытов с кубиком, то все частоты будут меняться. Предсказать их значения после 100 или 200 опытов заранее невозможно — в этом и состоит явление ***случайной изменчивости***. Однако и в такой изменчивости есть свои закономерности.

Заметим прежде всего, что поскольку случайное событие не может произойти в N опытах более N раз, то $N_A \leqslant N$, и поэтому $\frac{N_A}{N} \leqslant 1$. Это значит, что частота любого случайного события хотя и колеблется, но всегда лежит в промежутке от 0 до 1. При этом частота невозможного события равна 0, а частота достоверного события равна 1.

Очевидно также, что сумма частот всех исходов в любой серии опытов будет равна 1. В самом деле, ведь каждый опыт завершается ровно одним из этих исходов, поэтому сумма их абсолютных частот равна числу проведённых опытов N, а значит, сумма их относительных частот равна 1.

Поскольку любое случайное событие A наступает только в тех случаях, когда опыт завершается одним из благоприятных для него исходов, то сумма частот всех благоприятных для события A исходов будет равна частоте этого события.

Получаем следующие четыре свойства частот:

1. Частота любого случайного события всегда лежит в промежутке от 0 до 1.
2. Частота невозможного события равна 0, а частота достоверного равна 1.
3. Сумма частот всех возможных исходов опыта равна 1.
4. Частота события A равна сумме частот всех благоприятных для него исходов.

В поведении частот есть ещё одна замечательная закономерность. Она называется ***устойчивостью частот*** и состоит в том, что в длинной серии случайных опытов частота любого случайного события A стабилизируется и приближается к некоторому числу $P(A)$, которое называют ***вероятностью*** случайного события и рассматривают как числовую меру его правдоподобия.

Чтобы сделать это определение вероятности математическим, нужно установить два факта:

– во-первых, что стабилизация частоты события A будет происходить **в любой** достаточно длинной серии опытов;
– во-вторых, что число $P(A)$ для события A в любой такой серии будет **одним и тем же.**

В основе этих утверждений лежит так называемый ***закон больших чисел***, с которым вы уже отчасти знакомы и к доказательству которого мы ещё вернёмся в 11 классе. Пока можно рассматривать его как «закон природы», многократно подтверждённый опытом — в том числе и лично вашим.

Приведённые выше четыре свойства частот выполняются после любого числа проведённых опытов. Поскольку при большом числе опытов частоты почти совпадают с вероятностями, то **аналогичные свойства будут выполнены и для вероятностей**:

1. Вероятность любого случайного события всегда лежит в промежутке от 0 до 1.
2. Вероятность невозможного события равна 0, а вероятность достоверного события равна 1.

3. Сумма вероятностей всех элементарных исходов опыта равна 1.
4. Вероятность события A равна сумме вероятностей всех благоприятных для него исходов.

Используем последнее свойство для вычисления вероятностей случайных событий в следующем примере.

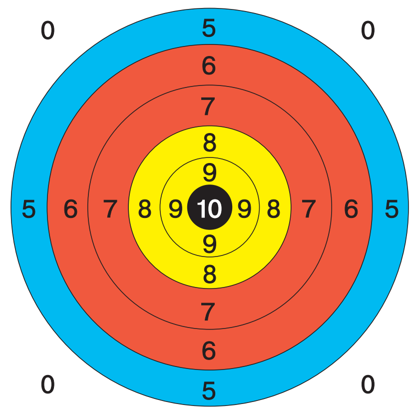


Рис. 61

Пример 2. Спортсмен стреляет из лука в мишень, изображённую на рисунке 61. Вероятности попадания в зоны мишени заданы в таблице 46:

Таблица 46

Зона	0	5	6	7	8	9	10
Вероятность	0,02	0,06	0,12	0,18	0,24	0,22	0,16

Найдём вероятности следующих событий:

A = «спортсмен получит больше семи очков»;
B = «спортсмен попадёт в красную зону»;
C = «спортсмен получит меньше 10 очков».

Для вычисления вероятностей A и B воспользуемся свойством 4 и сложим вероятности благоприятных для этих событий исходов:

$P(A) = 0,24 + 0,22 + 0,16 = 0,62;$

$P(B) = 0,12 + 0,18 = 0,3.$

При вычислении вероятности события C можно воспользоваться свойством 3 и вычесть из 1 вероятность единственного неблагоприятного исхода:

$P(C) = 1 - 0,16 = 0,84.$

Итак, по вероятностям возможных исходов опыта мы можем вычислять вероятности случайных событий. Но откуда взять вероятности самих исходов? Конечно, можно провести длинную серию экспериментов и оценить вероятность каждого исхода по его частоте. Но такой способ может оказаться слишком трудоёмким или просто невозможным.

Оказывается, во многих случаях можно найти вероятности всех исходов *априори*, без проведения реальных экспериментов (латинское выражение *a priori* можно перевести как «знание, полученное до опыта и независимо от него»). Об этом мы и поговорим в следующем разделе.

ВОПРОСЫ

1. В 100 бросаниях монеты орёл выпал 45 раз. Чему равна частота выпадения орлов? частота решек?
2. В каком промежутке лежит вероятность любого случайного события?
3. Чему равна вероятность невозможного события? достоверного события?
4. Чему равна сумма вероятностей всех возможных исходов случайного опыта?
5. Как найти вероятность случайного события, если известны вероятности всех возможных исходов опыта?

УПРАЖНЕНИЯ

103. Подбрасывают игральный кубик. Запишите каждое из событий в виде множества благоприятных исходов:
A = «выпадет число меньше 3»;
B = «выпадет число меньше 10»;
C = «выпадет простое число»;
D = «выпадет совершенное число».

Замечание. Число называется ***совершенным***, если оно равно сумме всех своих делителей, кроме себя самого.

104. Монету подбрасывают два раза. Выпишите все возможные исходы этого опыта. Запишите каждое из следующих событий в виде множества благоприятных исходов:
A = {первым выпал орёл}; C = {орлов выпало больше};
B = {орлов и решек выпало поровну}; D = {выпал хотя бы один орёл}.

105. Из колоды, содержащей 36 карт, вытаскивают одну карту. Сколько благоприятных исходов содержит каждое из следующих событий:
A = {вытянут даму}; C = {вытянут даму пик};
B = {вытянут пику}; D = {вытянут красную масть}?
Какие из этих событий будут элементарными?

106. Одновременно подбрасывают игральный кубик и монету. Опишите все возможные исходы этого опыта. Сколько их?

107. В коробке лежат два красных и один зелёный шар. Из неё вытаскивают одновременно два шара. Опишите все элементарные события в этом опыте. Сколько их?

108. Монету подбрасывают три раза. Выпишите все возможные исходы этого опыта. Посчитайте, сколько среди них благоприятных исходов для каждого из следующих событий:
A = {первым выпал орёл};
B = {орлов и решек выпало поровну};
C = {орлов выпало больше};
D = {не выпало двух орлов подряд}.

109. Три человека пришли в ресторан и сдали свои шляпы в гардероб. Уходя из ресторана, они разобрали свои шляпы наугад. Выпишите все возможные исходы этого опыта. Сколько их?

110. В каких из перечисленных случайных опытов имеется бесконечное число возможных исходов:
— подбрасываем 100 кубиков;
— бросаем монету до появления первого орла;
— раздаём карты перед началом игры в дурака;
— раскрываем книгу на случайной странице;
— бросаем дротик в мишень;
— измеряем температуру раствора на уроке химии?

111. Каждый пятый посетивший сайт книжного интернет-магазина покупает какую-то книгу. С какой вероятностью очередной клиент, зашедший на сайт, сделает заказ?

112. Судья назначил пенальти в ворота «Динамо». С какой примерно вероятностью он будет реализован, если за последние три сезона вратарь «Динамо» отразил 6 пенальти из 25?

113. В среднем из каждых 200 велосипедных насосов 5 оказываются с браком. Вы купили велосипедный насос. С какой вероятностью он исправен?

114. За столетнюю историю наблюдений в городе Морозове температура 1 января 15 раз опускалась ниже −30 °С. С какой вероятностью в следующую новогоднюю ночь температура будет выше −30 °С?

115. Из каждых 20 белых грибов 4 оказываются червивыми. Максим набрал в лесу 50 белых грибов. Сколько червивых грибов среди них он должен ожидать?

116. Сергей может получить на ближайшей контрольной по математике одну из оценок от 2 до 5 с вероятностями соответственно 0,1; 0,2; 0,4 и 0,3. С какой вероятностью он не получит двойку? С какой вероятностью его оценка будет больше 3?

§ 6. Опыты с равновозможными исходами

1 Классическое определение вероятности

Определение вероятности как предельного значения частоты называется статистическим. У него есть два существенных недостатка: во-первых, чтобы им воспользоваться, нужно провести достаточно длинную серию реальных опытов; во-вторых, какой бы длинной ни была эта серия, она даст только приближённое значение вероятности.

Понятно, что провести большое количество опытов возможно далеко не всегда. Например, на экспериментальное вычисление вероятности выигрыша в лотерею может просто не хватить денег. К счастью, во многих ситуациях существуют более «экономичные» способы расчёта вероятностей.

Мы уже приводили примеры простейших случайных опытов, в которых несложно указать все возможные исходы: для подбрасывания монеты — это орёл и решка, кубика — шесть его граней, для случайного выбора одного шара — любой из n шаров, находящихся в коробке. Поскольку монета и кубик симметричны, а шар вытаскивают наугад, то в каждом из этих опытов все исходы ***равновозможны***, т. е. могут произойти с одинаковой вероятностью. Как же найти эту вероятность?

Напомним, что для любого опыта с конечным числом исходов сумма их вероятностей равна 1. Если обозначить количество таких исходов n и предположить, что все они имеют одинаковую вероятность, то эта вероятность будет равна $\frac{1}{n}$.

Например, для монеты $n = 2$, поэтому вероятности выпадения орла и решки равны $\frac{1}{2}$. Для кубика $n = 6$, поэтому вероятность выпадения каждой из шести граней равна $\frac{1}{6}$. Если в коробке находится 12 шаров, то любой из них можно вытащить с одной и той же вероятностью $\frac{1}{12}$.

Таким образом, вычислять вероятности элементарных событий в опытах с равновозможными исходами можно без проведения экспериментов и подсчёта частот.

Пример 1. Для проведения экзамена по географии приготовили 24 билета. Андрей не выучил один билет и очень боится его вытянуть. Какова вероятность того, что Андрею достанется невыученный билет?

Всего у данного эксперимента 24 исхода, и все они равновозможны. Поэтому вероятность того, что Андрею достанется несчастливый для него билет, равна $\frac{1}{24}$.

Посмотрим теперь, как в опыте с n равновозможными исходами вычислить вероятность не только элементарного, но и любого другого случайного события. Если для события A имеется m благоприятных исходов, то его вероятность, как мы уже знаем, равна сумме их вероятностей:

$$P(A) = \underbrace{\frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n}}_{m \text{ слагаемых}} = \frac{m}{n}.$$

Полученную формулу часто называют ***классическим определением вероятности***. Такое название связано с тем, что это определение возникло на начальном этапе развития теории вероятностей. Считается, что окончательная формулировка этого определения принадлежит французскому учёному П. Лапласу (1749—1827). Сформулируем классическое определение вероятности ещё раз.

Если случайный эксперимент может завершиться одним из n равновозможных исходов, из которых m исходов благоприятны для случайного события A, то вероятность $P(A)$ можно вычислить по формуле

$$P(A) = \frac{m}{n}.$$

Применим это определение к вычислению вероятностей некоторых случайных событий.

Пример 2. Никита — одноклассник Андрея из примера 1 — не успел выучить к тому же самому экзамену по географии 6 билетов. С какой вероятностью он не сдаст экзамен?

Эксперимент у нас такой же, как и в примере 1, — выбор одного билета из 24. Все его исходы равновозможны. Но теперь речь идёт о вероятности неэлементарного случайного события A. Благоприятными для этого события (но не для Никиты!) будут 6 исходов. По классическому определению вероятности получаем:

$$P(A) = \frac{6}{24} = 0{,}25.$$

Пример 3. В классе, где учатся Андрей и Никита, 16 девочек и 12 мальчиков. Из них по жребию выбирают одного дежурного. С какой вероятностью дежурить придётся девочке?

Когда говорят «по жребию», то автоматически предполагают, что он проводится так, что все его участники имеют равные шансы быть выбранными (придумайте, как провести такой жребий). Поэтому мы можем считать, что все 16 + 12 = 28 исходов нашего опыта равновозможны. Благоприятными для события A = «дежурить придётся девочке» будут 16 исходов. Отсюда

$$P(A) = \frac{16}{28} \approx 0{,}571.$$

? ВОПРОСЫ

1. Сколько равновозможных исходов в опыте с подбрасыванием монеты? кубика?
2. Из коробки, в которой 10 конфет с белой начинкой и 6 конфет с чёрной начинкой, выбирают одну конфету. Сколько равновозможных исходов у этого опыта?
3. При регистрации на сайте пользователь указывает год своего рождения. Можно ли считать исходы этого опыта равновозможными? Почему?
4. Приведите свой пример опыта с равновозможными исходами. Сколько этих исходов?

2 Равновозможные исходы в сложных опытах

Итак, классическое определение вероятности даёт нам замечательную возможность предсказать, к чему будет приближаться частота случайного события в предстоящей серии испытаний, не проводя самих испытаний. Вычислив значение дроби $\frac{m}{n}$ (отношение числа благоприятных исходов к числу всех возможных исходов опыта), вы можете определить свои шансы сдать экзамен или выиграть в лотерею ещё до их проведения.

Но не нужно забывать и о тех ограничениях, которые накладывает это определение: опыт должен иметь **конечное** число **равновозможных** исходов. К сожалению, это бывает не так уж часто.

Во-первых, далеко не всегда исходы случайного опыта имеют равные вероятности. Если мы подбрасываем не монету, а канцелярскую кнопку, пуговицу или металлическую крышку от бутылки, то уже не можем считать два исхода такого опыта равновозможными: кнопка, пуговица и крышка не симметричны.

Во-вторых, встречаются опыты с бесконечным числом исходов. Представьте, например, что вы закрыли глаза и случайно выбрали точку на глобусе. С какой вероятностью вы попадёте на водную поверхность? Даже если предположить, что выбор всех точек равновозможен, это вряд ли позволит найти ответ: точек на глобусе (как и на его «водной» части) бесконечно много.

В первом случае у нас нет другого выхода, как вернуться к статистическому определению вероятности, а во втором можно использовать ещё одно — ***геометрическое*** — определение вероятности, которое мы рассматривали раньше в курсе математики основной школы.

Но даже если в случайном опыте есть симметрия, которая гарантирует равновозможность всех исходов, правильно определить эти исходы не всегда просто. До сих пор мы бросали одну монету и один кубик, вытаскивали один шар из коробки. Ситуация осложняется, если подбрасывают сразу несколько монет, кубиков или вытаскивают несколько шаров.

Начнём с опыта, в котором подбрасывают две монеты.

Пример 1. Подбрасывают две монеты. С какой вероятностью они выпадут на одну сторону?

Знаменитый французский учёный-энциклопедист Жан Лерон Даламбер (1717—1783) вошёл в историю науки не только со своими замечательными математическими результатами и философскими сочинениями, но и с ошибкой, которую он сделал (и долго на ней настаивал), решая эту задачу. Вот его решение.

«Опыт имеет 3 возможных исхода: обе монеты выпадут на орла, обе монеты выпадут на решку, монеты выпадут на разные стороны. Два из этих трёх исходов благоприятны для нашего события, поэтому его вероятность равна $\frac{2}{3}$».

Только реальный эксперимент смог убедить Даламбера, что частота этого события приближается вовсе не к $\frac{2}{3}$, а к $\frac{1}{2}$. Вы наверняка уже поняли, в чём здесь дело: исходы, предложенные Даламбером, не равновозможны. И связано это с тем, что событие «монеты выпадут на разные стороны» не является элементарным. Оно распадается на два исхода: на первой монете орёл, на второй — решка, и наоборот. Таким образом, у этого опыта не 3, а 4 равновозможных исхода. Если обозначить выпадение орла буквой О, а решки — буквой Р, то эти исходы легко перечислить:

ОО, ОР, РО, РР.

Из этих четырёх исходов два благоприятных для нашего события — это ОО и РР, поэтому вероятность равна $\frac{2}{4} = \frac{1}{2}$.

Теперь приведём более сложный пример, в котором можно привести целых два неправильных решения с разными ответами.

Пример 2. Из коробки, в которой 2 белых и 2 синих шара (рис. 62), вытаскивают два шара. Какова вероятность, что они окажутся одного цвета?

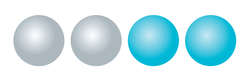
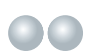
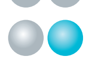


Рис. 62

Первое неправильное решение. Этот опыт имеет три возможных исхода:

1)  — вынули 2 белых шара;

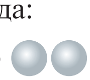
2)  — вынули 1 белый и 1 синий шар;

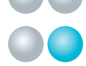
3)  — вынули 2 синих шара.

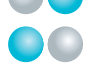
Благоприятными для нашего события будут два исхода: 1 и 3, поэтому вероятность равна $\frac{2}{3}$.

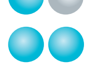
В этом решении мы считали, что шары вынимались одновременно, поэтому не различали, какой из них вынут первым, а какой — вторым. Если считать, что шары вынимают друг за другом, то получим следующее решение.

Второе неправильное решение. Этот опыт имеет четыре возможных исхода:

1)  — и первый, и второй шар белые;

2)  — первый шар белый, второй шар синий;

3)  — первый шар синий, второй шар белый;

4)  — и первый, и второй шар синие.

Благоприятными для нашего события будут два исхода: 1 и 4, поэтому вероятность равна $\frac{2}{4} = \frac{1}{2}$.

Получили другой ответ и, что удивительно, — опять неправильный! В чём же дело? Снова в неравновозможности выбранных исходов. Приведём теперь правильное решение.

Правильное решение. Пронумеруем шары, находящиеся в коробке (рис. 63):

Рис. 63

Этот опыт имеет шесть равновозможных исходов:

12, 13, 14, 23, 24, 34.

Мы специально выписали только номера шаров — по ним вы сами легко догадаетесь, какого цвета оказались вынутые шары. Только в двух исходах из шести шары оказываются одного цвета: 12 (оба белые) и 34 (оба синие). Поэтому искомая вероятность равна $\frac{2}{6} = \frac{1}{3}$. Это и есть правильный ответ, который можно подтвердить экспериментально.

Есть ещё одно правильное решение (разумеется, с тем же ответом). Если мы будем учитывать порядок, в котором вынимаются шары, то количество исходов удвоится и их станет 12:

12, 21, 13, 31, 14, 41, 23, 32, 24, 42, 34, 43.

Но теперь и благоприятных исходов будет не 2, а 4, поэтому вероятность останется прежней.

В чём же была причина сделанных ошибок? В неправильных решениях в качестве возможных исходов рассматривались не элементарные события, которые оказались не равновозможными. Чтобы избежать подобных ошибок, при подсчёте исходов опыта нужно различать сами предметы, из которых делается случайный выбор, а не их цвета, размеры или какие-то другие свойства.

Если возможных исходов опыта немного, можно их просто выписать друг за другом. Иногда существуют и более удобные способы представления — например, таблицы.

Пример 3. Одновременно бросают два кубика — красный и зелёный. Какова вероятность того, что на красном выпадет больше очков, чем на зелёном?

Исходы этого эксперимента удобно представить в виде следующей таблицы (табл. 47):

Таблица 47

к \ з	1	2	3	4	5	6
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

По строкам идут числа для красного кубика, по столбцам — для зелёного. Каждому исходу опыта соответствует клетка таблицы, в которой записано соответствующее двузначное число: первая цифра показывает, сколько очков выпало на красном кубике, а вторая — сколько очков выпало на зелёном. Таким образом, общее количество исходов равно количеству клеток в таблице: $6 \cdot 6 = 36$. Благоприятные исходы можно получить, если посчитать клетки, в которых первая цифра больше второй (они выделены оранжевым фоном). Поскольку этих клеток 15, то вероятность события «на красном кубике выпадет больше очков, чем на зелёном» равна $\frac{15}{36} = \frac{5}{12}$.

Такую таблицу можно использовать для вычисления вероятности любого события, связанного с опытом, в котором бросают два кубика, — достаточно выделить в ней благоприятные исходы, посчитать их количество и поделить на 36.

? ВОПРОСЫ

1. Монету бросают дважды. Сколько равновозможных исходов у этого опыта?
2. Кубик бросают дважды. Сколько равновозможных исходов у этого опыта?
3. Одновременно подбрасывают монету и кубик. Сколько равновозможных исходов у этого опыта?
4. Из пакета, в котором 2 апельсина и 2 грейпфрута, выбирают наугад 2 фрукта. Сколько равновозможных исходов у этого опыта? Перечислите их.
5. Из пакета, в котором 2 апельсина и 2 грейпфрута, выбирают наугад 3 фрукта. Сколько равновозможных исходов у этого опыта? Перечислите их.

УПРАЖНЕНИЯ

117. Для каждого из следующих случайных опытов найдите число всех возможных исходов, число благоприятных исходов и вычислите вероятность указанного события.

а) На столе 12 кусков пирога. В трёх «счастливых» из них запечены призы. Какова вероятность взять «счастливый» кусок пирога?

б) В коробке 15 белых и 25 чёрных шаров. Из неё наугад вынимается один шар. Какова вероятность того, что он будет белым?

в) Для школьной лотереи отпечатали 500 билетов, из них 25 выигрышных. Какова вероятность вытянуть билет без выигрыша?

118. Из колоды, содержащей 36 карт, вытаскивают одну карту. Определите вероятности следующих событий:

A = {вытянут даму};
B = {вытянут пику};
C = {вытянут даму пик};
D = {вытянут туза};
E = {вытянут не короля};
F = {вытянут красную масть}.

119. В коробке лежит 8 красных, 2 жёлтых и 20 зелёных карандашей.

а) Вы наугад вынимаете один карандаш. Какова вероятность того, что этот карандаш будет красным? жёлтым? зелёным?

б) Какое наименьшее количество карандашей нужно вынуть, чтобы с вероятностью, равной 1, среди них оказался красный карандаш? жёлтый? зелёный?

120. Наугад выбрано двузначное число. Какова вероятность того, что оно окажется:

а) чётным;
б) кратным 5;
в) простым;
г) квадратом натурального числа?

121. В группе российских туристов 2 человека владеют английским и французским языками, 1 человек — английским и немецким, 7 человек — только английским и 10 человек не владеют ни одним иностранным языком. Найдите вероятность того, что случайно выбранный гидом турист владеет:

а) французским языком;
б) двумя языками;
в) английским языком.

122. Подбрасывают два игральных кубика — красный и синий. Определите вероятности следующих событий:

A = {на кубиках выпадут одинаковые числа};
B = {на кубиках выпадут разные числа};
C = {числа на кубиках будут одинаковой чётности};
D = {сумма чисел будет равна 6};
E = {выпадет хотя бы одна шестёрка}.

123. Монету подбрасывают три раза. Найдите вероятности следующих событий:

A = {первым выпадет орёл};
B = {орлов и решек выпадет поровну};
C = {орлов выпадет больше};
D = {не выпадет двух орлов подряд}.

124. Одновременно подбрасывают игральный кубик и монету. С какой вероятностью в этом опыте выпадут орёл и шестёрка? решка и чётное число?

125. Наугад выбрано трёхзначное число. Какова вероятность того, что:

а) все цифры в нём разные;
б) в нём нет ни одной 1;
в) в нём есть хотя бы одна 1;
г) в нём есть ровно одна 1?

126. Из коробки с красными и жёлтыми шарами одновременно вынимают два шара. С какой вероятностью они будут одного цвета, если в коробке:

а) два красных и один жёлтый шар;
б) два красных и два жёлтых шара;
в) два красных и три жёлтых шара?

127. Даны отрезки длиной 2, 5, 6 и 10 см. Какова вероятность того, что из выбранных наудачу трёх отрезков можно составить треугольник?

128. На трёхместную скамейку случайным образом садятся двое мужчин и женщина. Какова вероятность того, что мужчины окажутся рядом?

129. Пятеро друзей случайным образом садятся за круглый стол. С какой вероятностью Света и Андрей окажутся рядом?

130. В соревнованиях по прыжкам в длину участвуют 4 спортсмена из России, 3 из Белоруссии, 2 из Казахстана, 1 из Армении. Порядок, в котором они будут выступать, определяется жребием. Найдите вероятность того, что первым будет выступать:

а) спортсмен из России;
б) спортсмен из Казахстана;
в) спортсмен не из России.

131. **Задача Эйлера.** Три человека пришли в ресторан и сдали свои шляпы в гардероб. Уходя из ресторана, они разобрали свои шляпы наугад. С какой вероятностью:

а) каждый уйдёт в своей шляпе;
б) каждый уйдёт в чужой шляпе;
в) один уйдёт в своей шляпе, а двое — в чужих;
г) двое уйдут в своих шляпах, а один — в чужой?

132. Бабушка дала внучке, не умеющей читать, три кубика. Какова вероятность того, что у внучки случайным образом получится из них слово русского языка, если на кубиках написаны буквы:

а) Д, М, Ы;
б) К, О, Т;
в) Г, Д, О;
г) Б, Б, О?

133*. 16 команд по жребию делятся на четыре равные группы. С какой вероятностью «ЦСКА» и «Спартак» окажутся в одной группе? в разных группах?

134*. На рисунке 64 изображены развёртки двух необычных игральных кубиков. Оба они необычны тем, что на их гранях написаны не совсем обычные числа. С какой вероятностью на первом кубике выпадет число больше, чем на втором? меньше, чем на втором? на кубиках выпадут равные числа?

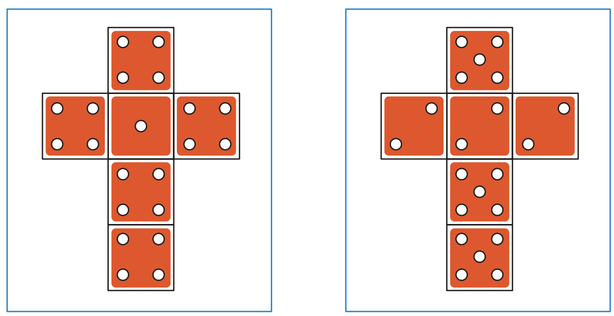


Рис. 64

135*. К двум кубикам из предыдущей задачи добавляют третий кубик (рис. 65) и предлагают сыграть вам в такую игру. Вы выбираете один из трёх необычных кубиков. После этого ваш соперник выбирает себе один из двух оставшихся. Каждый подбрасывает свой кубик. У кого выпало больше очков — выигрывает. Какой кубик вы выберете?

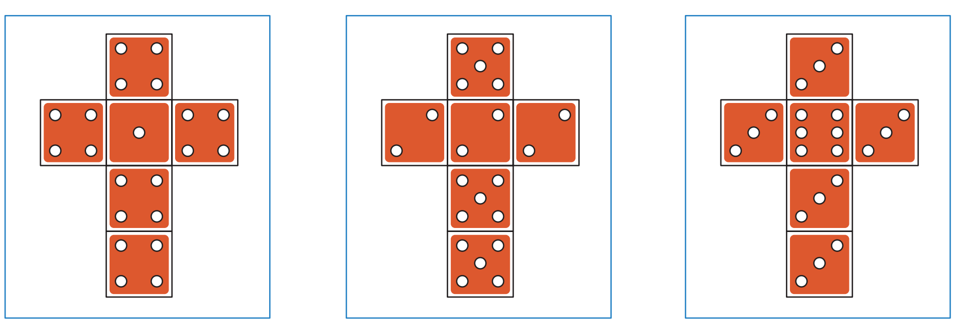


Рис. 65

ГЛАВА 4

СЛОЖЕНИЕ И УМНОЖЕНИЕ ВЕРОЯТНОСТЕЙ

В Большой советской энциклопедии сказано, что теория вероятностей — это «математическая наука, позволяющая по вероятностям одних случайных событий находить вероятности других, связанных каким-либо образом с первыми».
Оказывается, со случайными событиями можно производить такие же операции, как со множествами: их можно объединять, пересекать, находить дополнение. Что происходит при этом с их вероятностями? Как, например, вероятность объединения или пересечения выражается через вероятности исходных событий?
О том, как одни события выражаются через другие и как при этом найти вероятности полученных событий, вы и узнаете в этой главе.

§ 7. Операции над событиями

1 События и множества

В предыдущей главе мы говорили, что **любое случайное событие можно рассматривать как множество благоприятных для него исходов.**

Например, в опыте с подбрасыванием кубика случайное событие A = «на кубике выпадет чётное число очков» можно задать как множество $A = \{2, 4, 6\}$. Ничего удивительного в такой двойственности нет — ведь и другие множества можно описывать словами, а можно задавать перечислением составляющих его элементов.

Роль ***универсума***, содержащего все случайные события, играет множество всех возможных исходов случайного опыта. В теории вероятностей для обозначения этого множества принято использовать заглавную греческую букву «омега» Ω — последнюю букву греческого алфавита. В опыте с кубиком множество возможных исходов состоит из шести элементов: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Так как множество Ω содержит все возможные исходы, то это событие происходит при любом исходе опыта, т. е. является ***достоверным*** событием.

Пустое множество $\varnothing$, наоборот, не содержит ни одного исхода, а значит, никогда не происходит и является ***невозможным*** событием.

Чтобы показать соотношения между множествами, используют диаграммы Эйлера, на которых множества изображают кругами, а их элементы — точками. Такие диаграммы можно использовать и для случайных событий. Покажем, как они выглядят на нескольких примерах.

Пример 1. Бросают игральный кубик. Изобразим на диаграмме Эйлера (рис. 66) случайные события:

A = «на кубике выпадет чётное число очков» = $\{2, 4, 6\}$;
B = «на кубике выпадет шестёрка» = $\{6\}$;
C = «на кубике выпадет простое число» = $\{2, 3, 5\}$.

Для события B благоприятным будет всего один исход — напомним, что такие события называются ***элементарными***. События A и C имеют по три благоприятных исхода. При этом один исход у них общий — это число 2, которое является одновременно и чётным, и простым. Исход 1 не попадает ни в одно из трёх событий A, B, C.

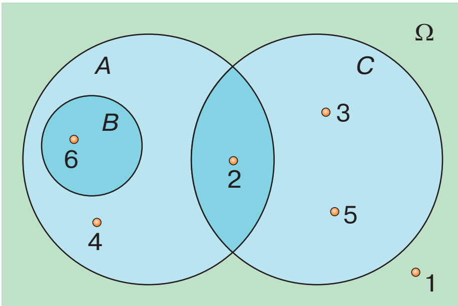


Рис. 66. Исходы и события в опыте с кубиком

Поскольку в этом опыте все 6 исходов ***равновозможны***, мы можем легко найти вероятность каждого из трёх событий A, B, C, используя ***классическое определение вероятности***:

$$P(A) = \frac{3}{6} = \frac{1}{2};\ P(B) = \frac{1}{6};\ P(C) = \frac{3}{6} = \frac{1}{2}.$$

В следующем примере исходы опыта не равновозможны, да и вопрос об их количестве уже не такой простой.

Пример 2. Перед футбольным матчем «Спартак» — «Динамо» болельщики обсуждают шансы событий:

A = «матч закончится вничью»;
B = «Динамо» не забьёт ни одного гола»;
C = «Спартак» выиграет».

Возможным исходом матча здесь можно считать итоговый счёт:

0:0, 1:0, 0:1, ...

Понятно, что таких исходов много и найти какое-то разумное ограничение для возможного счёта непросто (вспомните знаменитый матч из повести «Старик Хоттабыч», в котором команда «Шайба» забила в ворота «Зубила» 24 безответных мяча!). Тем не менее изобразить события A, B, C на диаграмме Эйлера вполне возможно (рис. 67).

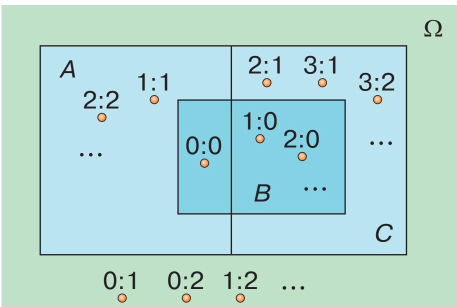


Рис. 67. Исходы и события в опыте с футбольным матчем

При перечислении возможных исходов на этой диаграмме пришлось ограничиться несколькими примерами и дальше поставить многоточие. Отметим ещё, что вместо традиционных кругов Эйлера мы изобразили события A, B, C в виде прямоугольников. Так легче показать, что у событий A и C нет общих благоприятных исходов, а все исходы, благоприятные для B, лежат либо в A, либо в C.

? ВОПРОСЫ

1. Как представить случайное событие в виде множества?
2. Из каких исходов состоит множество A, соответствующее событию A = «на кубике выпадет больше 3 очков»?
3. Изобразите на диаграмме Эйлера все исходы опыта с подбрасыванием кубика и события A = «выпадет больше 3 очков» и B = «выпадет меньше 5 очков».

2 Противоположное событие

Поскольку случайные события можно рассматривать как множества исходов, то над ними можно совершать те же операции, что и над любыми множествами: пересечение, объединение, дополнение. Посмотрим, какие события будут при этом получаться.

Два случайных события естественно назвать противоположными, если в каждом опыте происходит одно и только одно из них. Например, «лотерейный билет выиграет» и «лотерейный билет будет без выигрыша», «на монете выпадет орёл» и «на монете выпадет решка» и т. д. Все исходы опыта разбиваются при этом на два непересекающихся множества: благоприятные для первого события и благоприятные для второго. Таким образом, каждое из двух множеств является дополнением к другому. Это приводит нас к следующему определению.

Дополнением к событию A или противоположным к A событием называется событие $\overline{A}$, которое состоит из всех исходов, не содержащихся в A (т. е. неблагоприятных для A). Оно происходит всякий раз, когда не происходит A.

Результат этой операции показан на рисунке 68.

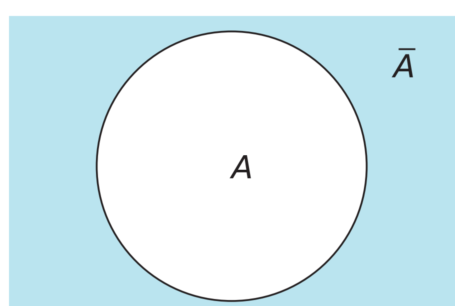


Рис. 68. Противоположное событие

Понятно, что событием, противоположным к событию $\overline{A}$, будет событие A.

Противоположное событие можно строить **как дополнение $\overline{A}$ к множеству** A, а можно с помощью логики, используя для этого **высказывание $\overline{A}$, противоположное для исходного высказывания** A. Покажем, как это делается в рассмотренных ранее примерах.

Пример 1. Найдём события, противоположные к событиям в опыте с кубиком:

A = «выпадет чётное число очков» = {2, 4, 6};
B = «выпадет шесть очков» = {6};
C = «выпадет простое число» = {2, 3, 5}.

В этом примере противоположные события одинаково легко описать как словами, так и перечислением исходов:

$\overline{A}$ = «выпадет нечётное число очков» = {1, 3, 5};
$\overline{B}$ = «выпадет не шесть очков» = {1, 2, 3, 4, 5};
$\overline{C}$ = «выпадет не простое число» = {1, 4, 6}.

Казалось бы, событие $\overline{C}$ можно описать как «выпадет составное число», но это будет неверно: число 1 не является ни простым, ни составным.

Пример 2. Для футбольного матча «Спартак» — «Динамо» найдём события, противоположные к событиям:

A = «матч закончится вничью»;
B = «Динамо» не забьёт ни одного гола»;
C = «Спартак» выиграет».

Здесь, как мы уже говорили, сложно перечислить все возможные исходы, поэтому воспользуемся логикой и построим отрицания к приведённым высказываниям:

$\overline{A}$ = «матч закончится не вничью» = «какая-то из команд выиграет»;
$\overline{B}$ = «Динамо» забьёт хотя бы один гол»;
$\overline{C}$ = «Спартак» не выиграет» = «Динамо» не проиграет» = «либо «Динамо» выиграет, либо матч закончится вничью».

Обратите внимание, что благодаря богатству русского языка сформулировать противоположное высказывание можно по-разному. Лучше выбрать самую простую формулировку, но главное — не делать при этом ошибок. Очень часто при построении противоположного события получают «неполное» дополнение (табл. 48).

Таблица 48

Высказывание	Неверное отрицание
«Динамо» не забьёт ни одного гола	«Динамо» забьёт один гол
«Спартак» выиграет	«Динамо» выиграет

В первом случае при построении отрицания пропущена важная частица «хотя бы», которая подразумевает не ровно один гол, а один и больше. Во втором случае не учтён ещё один возможный вариант неблагоприятного исхода — ничья.

Как видите, логика, теория множеств и теория вероятностей тесно связаны друг с другом — в чём вы ещё не раз сможете убедиться в дальнейшем.

? ВОПРОСЫ

1. Какое событие называется противоположным к событию A? Как оно обозначается?
2. Какое событие будет противоположным к событию $\overline{A}$?
3. Опишите событие, противоположное к событию A = «в трёх бросаниях монеты выпадет только решка». Сколько исходов содержит событие A? событие $\overline{A}$?
4. Опишите событие, противоположное к событию A = «при подбрасывании трёх кубиков ни разу не выпадет шесть очков». Сколько исходов содержит событие A? событие $\overline{A}$?

3 Пересечение событий

Напомним, что в пересечение двух множеств A и B входят элементы, которые содержатся в обоих этих множествах. Если A состоит из исходов, благоприятных для одного события, а B — для другого, то их пересечение состоит из исходов, благоприятных одновременно для A и B. Таким образом, событие $A \cap B$ будет происходить только в том случае, когда произошли оба события A и B.

Получаем следующее определение.

Пересечением двух событий A, B называют событие $A \cap B$, которое состоит из всех исходов, благоприятных для обоих событий A и B. Оно происходит всякий раз, когда происходят сразу оба события A и B.

Результат этой операции показан на рисунке 69.

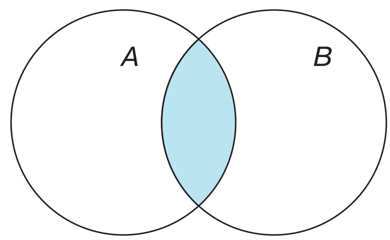


Рис. 69. Пересечение (произведение) событий

Пересечение событий в теории вероятностей часто называют **произведением событий** и обозначают AB. Найдём пересечения событий в примерах 1—2.

Пример 1. Вернёмся к примеру с кубиком и найдём попарные пересечения событий:

A = «выпадет чётное число очков» = $\{2, 4, 6\}$;
B = «выпадет шестёрка» = $\{6\}$;
C = «выпадет простое число» = $\{2, 3, 5\}$.

$A \cap B = \{6\}$, $A \cap C = \{2\}$, $B \cap C = \varnothing$.

Поскольку $B \subseteq A$, т. е. B является частью A, то пересечением A и B оказалось событие B. Пересечение событий B и C является пустым множеством, т. е. невозможным событием. Для непересекающихся событий в теории вероятностей используется специальный термин — ***несовместные***.

События A и B называют несовместными, если их пересечение — пустое множество, т. е. является невозможным событием. Несовместные события не могут произойти одновременно в одном случайном опыте.

Пример 2. Найдём пересечения событий, связанных с футбольным матчем:

A = «матч закончится вничью»;
B = «Динамо» не забьёт ни одного гола»;
C = «Спартак» выиграет».

Снова воспользуемся логикой. Пересечению множеств, как вы помните, соответствует логическая операция «и», которой мы и воспользуемся:

$A \cap B$ = «матч закончится вничью, и «Динамо» не забьёт ни одного гола» = «матч закончится со счётом 0 : 0»;

$A \cap C = \varnothing$;

$B \cap C$ = «Спартак» забьёт хотя бы один гол».

Заметим, что при нахождении $B \cap C$ мы использовали тот факт, что событие B произошло (команда «Динамо» не забила ни одного гола), а значит, «Спартаку» для победы достаточно забить хотя бы один гол.

? ВОПРОСЫ

1. Какое событие называется пересечением событий A и B? Как оно обозначается?
2. Как называются события, для которых $A \cap B = \varnothing$?
3. Монету бросают три раза. Запишите все благоприятные исходы для пересечения событий A = «выпадут ровно две решки» и B = «первый раз выпадет решка».
4. Бросают три кубика. Опишите пересечение событий A = «в сумме выпадет 5 очков» и B = «произведение очков будет равно 6».

4 Объединение событий

В ***объединение*** двух множеств A и B входят элементы, которые содержатся хотя бы в одном из этих множеств. Используя эту операцию, можно определить объединение случайных событий.

Объединением событий A и B называют событие $A \cup B$, которое состоит из всех исходов, которые входят хотя бы в одно из этих событий (т. е. благоприятных хотя бы для одного события). Оно происходит всякий раз, когда происходит хотя бы одно из событий A или B.

Результат этой операции показан на рисунке 70.

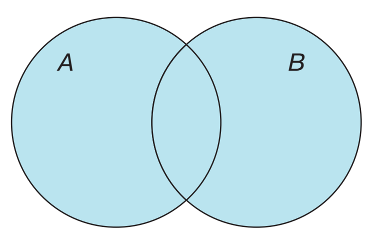


Рис. 70. Объединение (сумма) событий

Объединение событий в теории вероятностей часто называют ***суммой событий*** и обозначают $A + B$.

Отметим, что пересечение всегда содержится в объединении:

$$A \cap B \subseteq (A \cup B).$$

Пример 1. Найдём попарные объединения событий из опыта с кубиком:

A = «выпадет чётное число очков» = {2, 4, 6};
B = «выпадет шестёрка» = {6};
C = «выпадет простое число» = {2, 3, 5}.

Получаем:

$$A \cup B = \{2, 4, 6\},\ A \cup C = \{2, 3, 4, 5, 6\},\ B \cup C = \{2, 3, 5, 6\}.$$

Напомним, что при объединении событий общие элементы множеств не повторяются: элемент 2 входит в объединение $A \cup C$ только один раз. Отсюда вытекает ***формула включения-исключения***, по которой можно посчитать количество элементов в объединении. Для двух множеств она выглядит так: количество элементов в объединении любых двух множеств A и B можно вычислить по формуле

$$n_A + n_B - n_{AB},$$

где n_A, n_B — количество элементов в каждом из исходных множеств, а n_{AB} — количество элементов в их пересечении. Для объединения $A \cup C$ в нашем примере она будет выглядеть так: $3 + 3 - 1 = 5$.

Эта формула пригодится нам в дальнейшем при вычислении вероятности объединения двух случайных событий.

? ВОПРОСЫ

1. Какое событие называется объединением событий A и B? Как оно обозначается?
2. Приведите пример событий, для которых $A \cup B = \Omega$.
3. Монету бросают три раза. Запишите все благоприятные исходы для объединения событий A = «выпадут ровно две решки» и B = «первый раз выпадет решка».
4. Бросают три кубика. Сколько исходов содержит объединение событий A = «в сумме выпадет 5 очков» и B = «произведение очков будет равно 6»?

5 События, формулы и диаграммы

Итак, мы научились выполнять над случайными событиями такие операции, как дополнение, пересечение и объединение, используя для этого язык логики или теории множеств. Но в дальнейшем нам чаще придётся решать обратную задачу: выражать заданное случайное событие как результат каких-то операций над другими, более простыми событиями. Рассмотрим, как это делается на примерах.

Пример 1. Два стрелка делают по выстрелу в мишень. Даны два случайных события:

A = «первый стрелок попадёт в мишень»;
B = «второй стрелок попадёт в мишень».

Выразим через них следующие случайные события:

C = «оба стрелка попадут в мишень»;
D = «хотя бы один из них попадёт в мишень»;
E = «оба стрелка промахнутся»;
F = «хотя бы один из них промахнётся».

Событие C означает, что произошли **оба события** A и B, поэтому $C = A \cap B$.

Событие D происходит, когда происходит **хотя бы одно из двух событий** A или B, поэтому $D = A \cup B$.

Чтобы произошло событие E, должны произойти одновременно $\overline{A}$ и $\overline{B}$, поэтому $E = \overline{A} \cap \overline{B}$. Эту формулу можно преобразовать, если воспользоваться *законом де Моргана*, который знаком вам из основной школы: $E = \overline{A} \cap \overline{B} = \overline{A \cup B}$. Мы видим, что событие E будет противоположным к событию D: $E = \overline{A \cup B} = \overline{D}$.

Аналогично получаем, что $F = \overline{A} \cup \overline{B} = \overline{A \cap B} = \overline{C}$.

Таким образом, все четыре события C, D, E, F нам удалось выразить через исходные события A, B с использованием трёх операций: дополнения, пересечения и объединения. Заметим также, что среди наших событий оказались две пары взаимно противоположных: это пара D, E и пара C, F.

Все эти соотношения помогут нам в дальнейшем, когда нужно будет выражать вероятности одних событий через вероятности других.

Теперь рассмотрим более сложный пример, в котором буду даны три исходных события.

Пример 2. Студенту предстоит сдать в ближайшую сессию три экзамена: по высшей математике, по иностранному языку и по истории. Даны три случайных события:

A = {студент сдаст экзамен по высшей математике};
B = {студент сдаст экзамен по иностранному языку};
C = {студент сдаст экзамен по истории}.

Нужно выразить через них следующие события:

D = {студент сдаст все экзамены};
E = {студент не сдаст все экзамены};
F = {студент сдаст не все экзамены};
G = {студент сдаст хотя бы один экзамен};
H = {студент сдаст хотя бы два экзамена}.

Для наглядности будем изображать каждое из событий на диаграмме Эйлера.

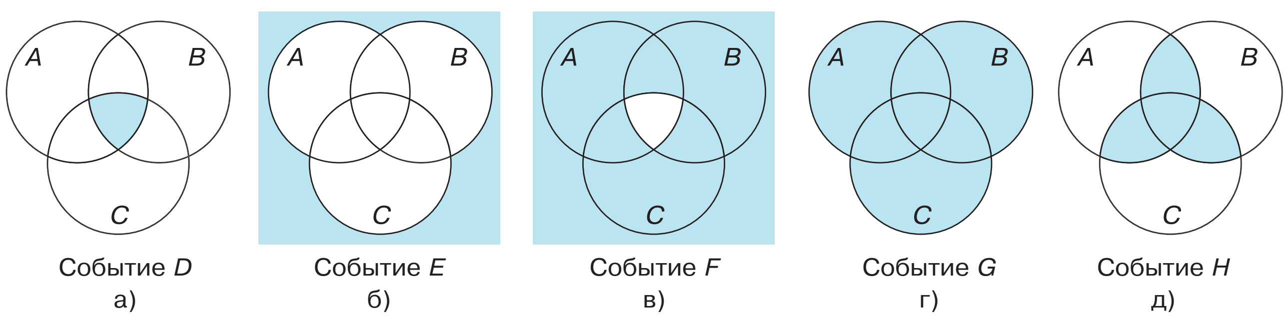


Рис. 71. Случайные события в примере 2

Проще всего найти событие D: чтобы оно произошло, должны произойти все три события A, B, C, поэтому оно равно их пересечению: $D = A \cap B \cap C$ (рис. 71, *а*).

Событие E происходит, когда происходят одновременно события $\overline{A}$, $\overline{B}$, $\overline{C}$, значит, оно равно их пересечению: $E = \overline{A} \cap \overline{B} \cap \overline{C}$ (рис. 71, *б*).

Может показаться, что событие F — то же самое, что E, но это не так! Оно происходит, когда не происходит D, — значит, оно является противоположным к нему: $F = \overline{D} = \overline{A \cap B \cap C} = \overline{A} \cup \overline{B} \cup \overline{C}$ (рис. 71, *в*). В последнем равенстве мы воспользовались законом де Моргана.

Событие G, очевидно, равно объединению событий A, B, C — ведь оно происходит, когда происходит хотя бы одно из них: $G = A \cup B \cup C$ (рис. 71, *г*).

Осталось выразить событие H. Оно состоит в том, что студент сдаст два или все три экзамена, поэтому выражение для него будет самым длинным:

$$H = (A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) \cup (A \cap B \cap C).$$

Если использовать для его записи знаки умножения и сложения вместо пересечения и объединения (мы уже говорили, что в теории вероятностей это допускается), то выражение будет значительно короче:

$$H = AB\overline{C} + A\overline{B}C + \overline{A}BC + ABC.$$

Ещё нагляднее это событие выглядит на диаграмме Эйлера (рис. 71, *д*).

ВОПРОСЫ

1. Из ящика, в котором находятся три красных и три зелёных яблока, наугад вынимают три яблока. События A_1, A_2, A_3 означают, что первое, второе и третье яблоко окажутся красными. Опишите словами события:
$A_1 \cap A_2 \cap A_3$,
$A_1 \cup A_2 \cup A_3$,
$(A_1 \cup A_2) \cap A_3$,
$(A_1 \cup A_2) \cup A_3$.
2. Сегодня два урока математики, на каждом из которых вас могут вызвать к доске. Событие A означает, что вас вызовут на первом уроке, событие B — на втором. Запишите через них следующие события:
— «вас сегодня не вызовут к доске»;
— «вас сегодня хотя бы раз вызовут к доске»;
— «вас вызовут только на втором уроке»;
— «вас вызовут только на одном уроке».

УПРАЖНЕНИЯ

136. Бросают игральный кубик. Изобразите на диаграмме Эйлера все возможные исходы этого опыта и следующие случайные события:
A = «выпадет тройка»;
B = «выпадет нечётное число»;
C = «выпадет больше трёх очков»;
D = «выпадет не меньше четырёх очков».

137. Бросают два кубика. Найдите события, противоположные к следующим событиям:
A = «на первом кубике выпадет больше очков, чем на втором»;
B = «хотя бы на одном из кубиков выпадет чётное число очков»;
C = «одно из выпавших чисел будет делиться на другое»;
D = «оба выпавших числа будут простыми».

138. Четыре раза подряд бросают монету. Найдите события, противоположные к следующим событиям, и посчитайте количество благоприятных для них исходов:
A = «орлов и решек выпадет поровну»;
B = «орлов выпадет больше, чем решек»;
C = «орлов выпадет меньше, чем решек»;
D = «ни разу не выпадет два орла подряд».

139. Баскетболист готовится выполнить три штрафных броска. За каждый точный бросок его команда получает одно очко. События A_1, A_2, A_3 означают, что соответствующий бросок окажется точным. Выразите через эти события следующие события:
B_0 = «команда не получит ни одного очка»;
B_1 = «команда получит одно очко»;
B_2 = «команда получит два очка»;
B_3 = «команда получит три очка».

140. Из ящика, в котором находятся 2 красных и 2 синих носка, одновременно вытаскивают один за другим 2 носка. Событие A_1 состоит в том, что первый носок красный, событие A_2 — что второй носок красный. Выразите через них следующие события:
A = «вынуты два красных носка»;
B = «вынуты два синих носка»;
C = «вынуты носки одного цвета»;
D = «вынуты носки разных цветов».

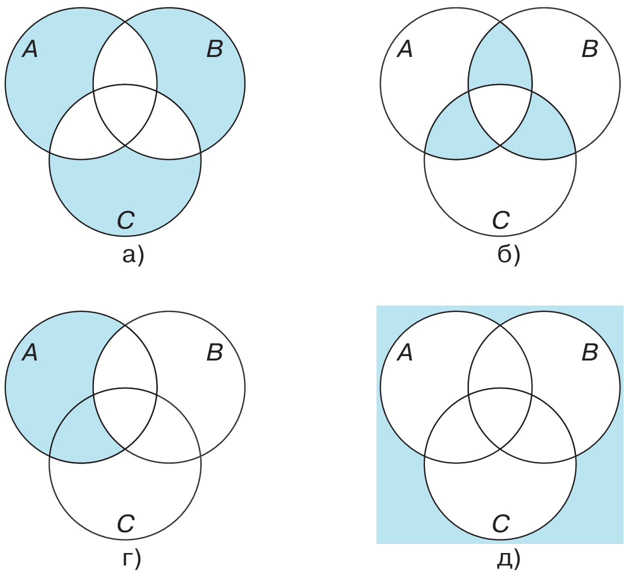


Рис. 72

141*. Команде знатоков в телевизионной игре «Что? Где? Когда?» предстоит сыграть 3 матча с телезрителями. События A, B, C означают, что соответствующий матч закончится победой знатоков. Опишите словами события, изображённые на диаграммах Эйлера — Венна (рис. 72).

142*. Бросают два кубика. Пусть A_1 = «на первом кубике выпадет чётное число», A_2 = «на втором кубике выпадет чётное число». Выразите через события A_1, A_2 следующие события:
A = «сумма выпавших очков чётна»;
B = «сумма выпавших очков нечётна»;
C = «произведение выпавших очков чётно»;
D = «произведение выпавших очков нечётно».

143*. На рисунке 73 дана электрическая схема соединения пяти нагревательных элементов. События A_1, A_2, A_3, A_4, A_5 означают, что соответствующий элемент исправен и пропускает электрический ток. Событие A означает, что вся схема пропускает ток. Выразите событие A через A_1, A_2, A_3, A_4, A_5.

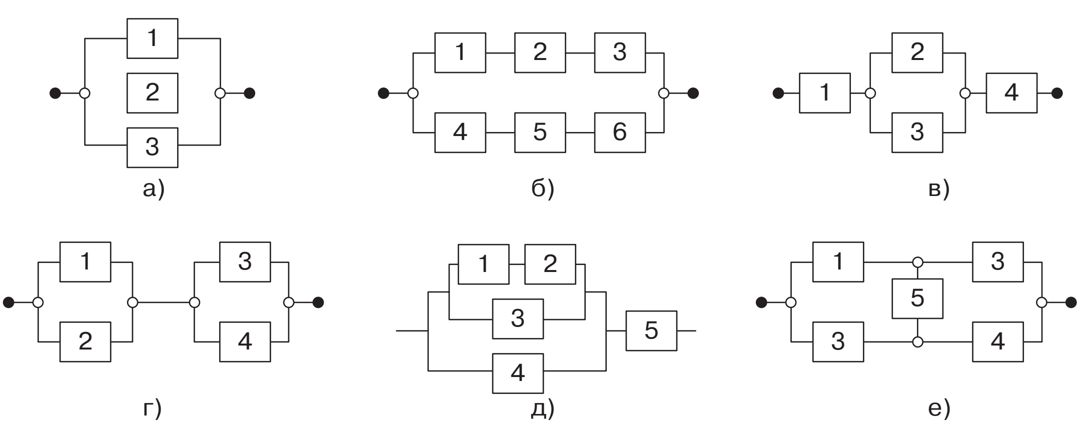


Рис. 73

§ 8. Сложение вероятностей

1 Вероятность противоположного события

Итак, мы научились оперировать со случайными событиями: находить их пересечение и объединение, строить противоположное событие. А что при этом происходит с вероятностями исходных событий? Проще всего ответить на этот вопрос для противоположного события.

Рассмотрим сначала опыт с n **равновозможными** исходами. Если m из этих исходов благоприятны для события A, то неблагоприятных (т. е. таких, при которых A не происходит) будет $(n - m)$. Поэтому по классическому определению вероятности

$$P(\overline{A}) = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - P(A).$$

Мы получили формулу для **вероятности противоположного события**:

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

Чтобы обобщить эту формулу на другие ситуации, в которых исходы не обязательно равновозможны, вспомним определение вероятности через частоту. Если в серии из N опытов событие A наступало M раз, то не наступало оно $(N - M)$ раз. Обозначим через $F(A)$ частоту события A в этой серии опытов. Тогда

$$F(\overline{A}) = \frac{N-M}{N} = 1 - \frac{M}{N} = 1 - F(A).$$

Поскольку частоты с ростом числа опытов приближаются к вероятностям, то полученное соотношение будет выполнено и для вероятностей.

Пример 1. Каждый восьмой выпускник автошколы сдаёт экзамен на получение водительских прав с первого раза. С какой вероятностью выпускник этой автошколы не сдаст экзамен с первого раза?

Пусть событие A = «выпускник сдаст экзамен с первого раза». Тогда $\overline{A}$ = «выпускник не сдаст экзамен с первого раза». Фраза «каждый восьмой сдаёт экзамен с первого раза» означает, что $P(A) = \frac{1}{8}$. Отсюда $P(\overline{A}) = 1 - \frac{1}{8} = \frac{7}{8}$.

Пример 2. Бросают два кубика. С какой вероятностью на них выпадут разные числа?

Обозначим через A событие «на кубиках выпадут одинаковые числа». Тогда $\overline{A}$ = «на кубиках выпадут разные числа». В § 6 мы уже выяснили, что этот опыт может закончиться одним из 36 равновозможных исходов. Благоприятными для события A будут исходы, в которых на кубиках выпадут одинаковые числа. Всего таких исходов шесть: (1; 1), (2; 2), (3; 3), (4; 4), (5; 5), (6; 6).

По классическому определению вероятности $P(A) = \frac{6}{36} = \frac{1}{6}$. Отсюда $P(\overline{A}) = 1 - \frac{1}{6} = \frac{5}{6}$.

? ВОПРОСЫ

1. Если вероятность события A равна 0, то чему равна вероятность события $\overline{A}$?
2. Могут ли обе вероятности $P(A)$ и $P(\overline{A})$ быть больше 0,5? больше 0,3? меньше 0,8?
3. Приведите пример такого события A, для которого $P(A) = P(\overline{A}) = 0{,}5$.

2 Формула суммы для несовместных событий

Теперь перейдём к вычислению вероятности объединения событий. Эту задачу мы разобьём на два случая: простой (но очень важный) и общий, более сложный.

Начнём с простого. Предположим, что наши события A и B не пересекаются, т. е. не имеют общих исходов, а значит, не могут произойти в одном опыте одновременно. Напомним, что такие события в теории вероятностей называются *несовместными*.

Как и для противоположного события, рассмотрим сначала опыт с равновозможными исходами. Пусть n — общее число исходов опыта, из которых n_A благоприятных для события A, n_B — для события B. Поскольку события A и B не имеют общих исходов, то объединение $A \cup B$ будет содержать $(n_A + n_B)$ исходов. Отсюда

$$P(A \cup B) = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = P(A) + P(B).$$

В общем случае эту формулу можно получить из соотношения для частот: если в N опытах событие A произошло N_A раз, а событие B — N_B раз, то событие $(A + B)$ произошло $(N_A + N_B)$ раз.

Мы получили формулу для **вероятности объединения несовместных событий**:

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

Если использовать для объединения событий знак «плюс», то формула становится ещё «красивее»:

$$P(A + B) = P(A) + P(B)$$

и может быть прочитана так: **вероятность суммы несовместных событий равна сумме их вероятностей**. Важно только, чтобы за этой формулировкой вы не забыли два важных условия:

– во-первых, это справедливо только для несовместных событий;
– во-вторых, **сумма событий — это не сумма чисел, а объединение множеств**.

Пример 1. Антон и Тимур ходят в шахматный клуб и часто играют между собой в шахматы. По статистике, примерно в 25% партий выигрывал Антон, в 15% — Тимур и 60% партий закончилась вничью. С какой вероятностью в очередной партии Тимур не проиграет?

Рассмотрим три события:

A = «выиграет Антон»;
B = «выиграет Тимур»;
C = «партия закончится вничью».

По условию задачи $P(A) = 0{,}25$, $P(B) = 0{,}15$, $P(C) = 0{,}6$. Событие «Тимур не проиграет» означает, что Тимур выиграет или будет ничья, т. е. оно равно объединению $B \cup C$. Поскольку события B и C несовместны, то

$$P(B + C) = P(B) + P(C) = 0{,}15 + 0{,}6 = 0{,}75.$$

Пример 2. Марина решила поработать на летних каникулах и отправила заявки в два сетевых супермаркета. Вероятность положительного ответа от первого супермаркета составляет 0,7, от второго — 0,4. С какой вероятностью она получит хотя бы один положительный ответ?

Обозначим A, B два события, о которых идёт речь в задаче:

A = «первый супермаркет ответит положительно»;
B = «второй супермаркет ответит положительно».

Тогда $P(A) = 0{,}7$; $P(B) = 0{,}4$. Хотя бы один положительный ответ Марина получит, если произойдёт хотя бы одно из двух событий A или B, что равносильно объединению $A \cup B$.

Можно подумать, что для вычисления $P(A \cup B)$ нужно сложить вероятности событий A и B:

$P(A + B) = P(A) + P(B) = 0{,}7 + 0{,}4 = 1{,}1$.

Полученная вероятность оказалась больше 1 — но такого не может быть!

Где же допущена ошибка? Было забыто главное условие, при котором можно складывать вероятности: события A и B должны быть **несовместными**! В нашем случае они, конечно, совместны: Марина вполне может получить положительные ответы сразу от двух супермаркетов. Значит, пользоваться формулой для вероятности объединения несовместных событий в этом примере нельзя.

О том, как вычислять вероятность объединения совместных событий, мы узнаем в следующем пункте.

? ВОПРОСЫ

1. Какие случайные события называются несовместными?
2. Приведите примеры двух совместных и двух несовместных случайных событий.
3. Могут ли события, вероятности которых равны 0,4 и 0,7, быть несовместными? Свой ответ объясните.
4. Если $P(A) = 0{,}3$, $P(B) = 0{,}4$ и события A и B несовместны, то чему равна вероятность $P(A \cup B)$?

3 Формула суммы для произвольных событий

Итак, пусть теперь события A и B совместны, т. е. имеют общие исходы, пересекаются. Обозначим через n_{AB} количество исходов в пересечении. По *формуле включения-исключения*, о которой уже говорилось выше, можно записать:

$$n_{A+B} = n_A + n_B - n_{AB}$$

(напомним, что вычитать n_{AB} приходится потому, что при сложении n_A и n_B мы посчитали общие исходы событий A и B дважды). Если все n исходов опыта равновозможны, то

$$P(A \cup B) = \frac{n_{A+B}}{n} = \frac{n_A + n_B - n_{AB}}{n} = \frac{n_A}{n} + \frac{n_B}{n} - \frac{n_{AB}}{n} = P(A) + P(B) - P(A \cap B).$$

Для опыта с неравновозможными исходами можно, как и раньше, доказать аналогичное равенство для частот (запишите его самостоятельно), а затем перейти к вероятностям.

Мы получили формулу для вероятности объединения любых событий:

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

Заметим, что если события несовместны, то $A \cap B = \varnothing$, поэтому $P(A \cap B) = 0$ и наша общая формула переходит в уже доказанную ранее формулу для несовместных событий.

Пример 1. Из колоды, в которой 36 карт, вытаскивают наугад одну карту. С какой вероятностью эта карта будет дамой или пикой?

Чтобы ответить на этот вопрос, рассмотрим два события:

A = «из колоды вытянут даму»;
B = «из колоды вытянут пику».

Вероятности этих событий можно посчитать по классическому определению — как отношение числа благоприятных исходов к числу всех равновозможных исходов опыта. Всего карт в колоде 36, а дам — 4, поэтому $P(A) = \frac{4}{36} = \frac{1}{9}$. Поскольку всего разных мастей в колоде 4, то карт пиковой масти 36 : 4 = 9, поэтому $P(A) = \frac{9}{36} = \frac{1}{4}$.

Нам нужно найти вероятность объединения $A \cup B$. Чтобы воспользоваться для этого доказанной выше формулой, осталось найти $P(A \cap B)$. Но событие $A \cap B$ означает, что вынутая карта является одновременно и дамой, и картой пиковой масти, т. е. это дама пик. Такая карта всего одна, поэтому $P(A \cap B) = \frac{1}{36}$.

Подставим все полученные вероятности в формулу и найдём ответ:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{9} + \frac{1}{4} - \frac{1}{36} = \frac{12}{36} = \frac{1}{3}.$$

Пример 2. В зале ожидания стоят два автомата по разливу кофе. Вероятность того, что к концу дня кофе закончится в первом автомате, — 0,1, во втором — тоже 0,1. Вероятность, что кофе закончится в обоих автоматах, — 0,05. С какой вероятностью к концу дня кофе останется в обоих автоматах?

Рассмотрим два события:

A = «кофе останется в первом автомате»;
B = «кофе останется во втором автомате».

Тогда событиями, о которых идёт речь в задаче, будут:

$\overline{A}$ = «кофе закончится в первом автомате»;
$\overline{B}$ = «кофе закончится во втором автомате».

Нам даны вероятности $P(\overline{A})$, $P(\overline{B})$, $P(\overline{A} \cap \overline{B})$, а найти нужно $P(A \cap B)$. Чтобы воспользоваться формулой суммы, перейдём к противоположному событию $\overline{A \cap B}$ и применим закон де Моргана:

$$P(A \cap B) = 1 - P(\overline{A \cap B}) = 1 - P(\overline{A} \cup \overline{B}).$$

По формуле суммы для совместных событий

$$P(\overline{A} \cup \overline{B}) = P(\overline{A}) + P(\overline{B}) - P(\overline{A} \cap \overline{B}) = 0{,}1 + 0{,}1 - 0{,}05 = 0{,}15.$$

Отсюда $P(A \cap B) = 1 - 0{,}15 = 0{,}85$.

? ВОПРОСЫ

1. Если $P(A) = 0{,}3$, $P(B) = 0{,}4$, $P(A \cap B) = 0{,}2$, то чему равна вероятность $P(A \cup B)$?

2 Если $P(A) = 0{,}6$, $P(B) = 0{,}3$, $P(A \cup B) = 0{,}8$, то чему равна вероятность $P(A \cap B)$?

3. Какое наибольшее значение может иметь вероятность $P(A \cap B)$, если $P(A) = 0{,}5$, $P(B) = 0{,}9$?

4. Какое наибольшее значение может иметь вероятность $P(A \cup B)$, если $P(A) = 0{,}5$, $P(B) = 0{,}9$?

УПРАЖНЕНИЯ

144. Про случайные события A и B известно, что они несовместные и что $P(A) = 0{,}3$, а $P(B) = 0{,}4$. Найдите вероятности событий $P(\overline{A})$, $P(\overline{B})$, $P(A \cap B)$, $P(A \cup B)$.

145. Могут ли события быть несовместными, если:

а) $P(A) = 0{,}5,\ P(B) = 0{,}3$;
б) $P(A) = 0{,}5,\ P(B) = 0{,}5$;
в) $P(A) = 0{,}4,\ P(B) = 0{,}7$;
г) $P(A) = 0{,}1,\ P(B) = 0{,}1$?

146. Нарисуйте события $A,\ B$ на диаграмме Эйлера и найдите $P(A \cup B)$, если:

а) $P(A) = 0{,}5,\ P(B) = 0{,}3,\ P(A \cap B) = 0{,}2$;
б) $P(A) = 0{,}4,\ P(B) = 0{,}7,\ P(A \cap B) = 0{,}4$;
в) $P(A) = 0{,}4,\ P(B) = 0{,}3,\ P(A \cap B) = 0$;
г) $P(A) = 0{,}8,\ P(B) = 0{,}5,\ P(A \cap B) = 0{,}3$.

147. Нарисуйте события A, B на диаграмме Эйлера и найдите $P(A \cap B)$, если:

а) $P(A) = 0{,}7,\ P(B) = 0{,}5,\ P(A \cup B) = 0{,}7$;
б) $P(A) = 0{,}6,\ P(B) = 0{,}3,\ P(A \cup B) = 0{,}7$;
в) $P(A) = 0{,}3,\ P(B) = 0{,}3,\ P(A \cup B) = 0{,}6$;
г) $P(A) = 0{,}9,\ P(B) = 0{,}8,\ P(A \cup B) = 1$.

148. Известно, что $P(A) = 0{,}4,\ P(B) = 0{,}5,\ P(A \cap B) = 0{,}3$. Найдите вероятности событий:

а) $P(A \cup B)$;
б) $P(A \cup B)$;
в) $P(\overline{A} \cup \overline{B})$;
г) $P(\overline{A} \cap \overline{B})$;
д) $P(A \cap \overline{B})$;
е) $P(\overline{A} \cup B)$.

149. Справедливы ли следующие утверждения:

а) $P(A) + P(B) = 1 \Rightarrow$ события A и B противоположные;
б) $P(A \cup B) = P(A) + P(B) \Rightarrow$ события A и B несовместные?

150. В отделении банка стоят два банкомата. Первый может выдавать и принимать купюры, второй — только выдавать. Вероятность, что к концу дня купюры закончатся в первом автомате, равна 0,1, а что закончатся во втором, — 0,3. Вероятность, что они закончатся в двух автоматах, — 0,06. Найдите вероятность того, что к концу дня купюры:

а) останутся в обоих автоматах;
б) останутся хотя бы в одном автомате;
в) закончатся хотя бы в одном автомате;
г) закончатся только во втором банкомате, а в первом останутся;
д) закончатся только в первом банкомате, а во втором останутся.

151. Известно, что $P(A) = 0{,}4,\ P(B) = 0{,}3$. Какое наибольшее и какое наименьшее значение может иметь вероятность события $A \cap B$? вероятность события $A \cup B$?

152. Друзья Петров и Васечкин часто прогуливают лекции. Вероятность их совместного появления на лекции равна 0,1, а вероятность, что оба не придут на лекцию, — 0,8. Петров прогулял 87% лекций. Сколько процентов лекций прогулял Васечкин?

153. Вероятность того, что на тестировании по истории ученик верно ответит больше, чем на 8 вопросов, равна 0,65. Вероятность того, что он верно ответит больше, чем на 7 вопросов, равна 0,74. Найдите вероятность того, что ученик ответит ровно на 8 вопросов.

154. Вероятность, что масса буханки хлеба будет больше 790 г, составляет 0,6, а что масса будет меньше 810 г, — 0,7. С какой вероятностью масса буханки хлеба будет от 790 до 810 г?

155. Вероятность того, что в десяти бросаниях монеты выпадет больше пяти орлов, составляет 0,377. С какой вероятностью выпадет меньше шести решек?

156. Монету бросают 13 раз. С какой вероятностью орлов выпадет больше, чем решек?

157. Допустимое отклонение диаметра CD-диска от номинального составляет 0,01 мм. Если отклонение больше этого значения, то диск бракуется. Вероятность, что диаметр диска будет больше максимально допустимого, составляет 0,02, а меньше минимально допустимого, — 0,03. С какой вероятностью изготовленный диск окажется без брака?

158. По данным статистических наблюдений, вероятность того, что уровень весеннего подъёма воды в реке будет больше 6 м, составляет 0,1, больше 3 м, — 0,5, больше 1 м, — 0,8. Оцените сверху и снизу вероятность того, что уровень подъёма воды превысит 4 м.

159. Вероятность, что в предстоящем матче команда «Спартак» не проиграет команде «Динамо» равна 0,6. Вероятность, что «Динамо» не проиграет «Спартаку», равна 0,7. С какой вероятностью этот матч завершится вничью?

160. Светлана обещала вернуться домой к 19:00. Вероятность, что она опоздает больше, чем на час, равна 0,1. Вероятность, что она не опоздает, равна 0,3. С какой вероятностью она опоздает, но не больше, чем на час?

161. Вероятность, что следующая композиция, которую выберет диджей, будет звучать больше 10 мин, составляет 0,15. Вероятность, что она будет не длиннее 5 мин, равна 0,2, а что не короче 3 мин, — 0,95. С какой вероятностью длительность композиции будет от 3 до 5 мин? от 5 до 10 мин?

§ 9. Умножение вероятностей

1 Условная вероятность

Очень часто мы оказываемся в ситуации, когда в процессе проведения случайного опыта нам приходится пересмотреть наше первоначальное представление о вероятности какого-либо события.

До начала футбольного матча мы оценивали вероятность выигрыша любимой команды примерно как 0,9, а после того, как в первом тайме она пропустила три безответных мяча, наша уверенность в её победе упала до 0,2.

Готовясь к экзамену, Борис выучил только 10 билетов из 20, и его шансы сдать экзамен были равны $\frac{1}{2}$. Пятеро его товарищей уже зашли в аудиторию и вытащили билеты, причём всем попались те, что выучил Борис. Его шансы сдать экзамен упали до $\frac{1}{3}$ (объясните почему).

Перед тем как трижды бросить монету, мы вычислили вероятность того, что все три раза выпадет орёл: она оказалась равной $\frac{1}{8}$. Вы бросили монету, и на ней выпал орёл. Очевидно, вероятность нашего события увеличилась и стала равна $\frac{1}{4}$.

Почему вероятность в приведённых примерах меняется? Ответ простой: происходят события, из-за которых какие-то элементарные исходы вовсе исключаются из рассмотрения, а вероятность каждого из оставшихся исходов увеличивается. Например, в опыте с монетой было изначально 8 равновозможных исходов:

$$\text{ООО, ООР, ОРО, ОРР, РОО, РОР, РРО, РРР.}$$

Из них только один был благоприятным для нашего события A, поэтому $P(A) = \frac{1}{8}$.

После того как в первом испытании выпал орёл, осталось только 4 равновозможных исхода:

$$\text{ООО, ООР, ОРО, ОРР,}$$

поэтому вероятность $P(A)$ стала равняться $\frac{1}{4}$.

В такой ситуации говорят об ***условной вероятности*** события A при условии, что произошло некоторое событие B. Эта вероятность обозначается $P(A \mid B)$. Обычная вероятность $P(A)$, которую мы вычисляем до начала опыта, называется при этом ***безусловной вероятностью***.

Условной вероятностью события A при условии события B называется вероятность события A, вычисленная при условии, что событие B произошло. В отличие от обычной (безусловной) вероятности она обозначается как $P(A \mid B)$.

Таким образом, в приведённых выше примерах мы вычисляли две разных вероятности: сначала безусловную $P(A)$, а потом условную $P(A \mid B)$. Например, в последнем примере с монетой безусловная вероятность $P(A) = \frac{1}{8}$, а условная вероятность $P(A \mid B) = \frac{1}{4}$, где событие B = «при первом подбрасывании выпал орёл».

Рассмотрим ещё два примера на вычисление условной вероятности.

Пример 1. В кармане у Макара лежат 6 орехов, 2 из которых пустые. Он вынимает из кармана и раскалывает один за другим 2 ореха. Рассмотрим события:

A = {первый орех будет полным},
B = {второй орех будет полным}.

Очевидно, что $P(A) = \frac{4}{6} = \frac{2}{3}$. А как вычислить $P(B)$? Если событие A произошло, то вероятность события B снижается до $\frac{3}{5}$, а если не произошло — наоборот, увеличивается до $\frac{4}{5}$. Обе этих вероятности есть не что иное, как условные вероятности события B: первая — при условии, что A произошло, а вторая — при условии, что A не произошло:

$$P(B \mid A) = \frac{3}{5} = 0{,}6, \quad P(B \mid \overline{A}) = \frac{4}{5} = 0{,}8.$$

Конечно, у события B есть и безусловная вероятность. Для её вычисления можно использовать две найденные условные вероятности и ***формулу полной вероятности***, с которой мы познакомимся уже в следующем разделе.

Тем не менее найти $P(B)$ можно и сейчас — достаточно воспользоваться классическим определением вероятности. По комбинаторному правилу умножения наш опыт имеет $6 \cdot 5 = 30$ равновозможных исходов (первый орех можно вытащить 6 способами, после чего второй орех — 5 способами). Чтобы посчитать благоприятные для B исходы, разобьём их на два непересекающихся множества:

- первый орех полный, и второй полный: $4 \cdot 3 = 12$ исходов;
- первый орех пустой, а второй полный: $2 \cdot 4 = 8$ исходов.

Теперь сложим эти два числа: $12 + 8 = 20$. Получаем, что $P(B) = \frac{20}{30} = \frac{2}{3}$.

Вероятности событий A и B оказались одинаковыми. Ничего удивительного в этом нет — ведь мы нашли ***безусловную*** вероятность события B, полученную до того, как мы вынули первый орех. После того как событие A произошло (или не произошло), вероятность события B меняется и $\frac{2}{3} = 0{,}66(6)$ превращается, соответственно, в 0,6 или в 0,8.

В следующем примере вычислить условную вероятность будет уже не так просто.

Пример 2. При подбрасывании двух кубиков в сумме выпало 6 очков. С какой вероятностью оба выпавшие числа чётные?

Разберёмся, о какой вероятности идёт речь. Вот два события, которые рассматриваются в задаче:

A = {в сумме выпало 6 очков},
B = {на кубиках выпали чётные числа}.

Поскольку говорится, что событие A произошло, то нужно найти условную вероятность $P(B \mid A)$.

В отличие от предыдущего примера, в котором опыт состоял из двух шагов, здесь **событие A не предшествует событию B во времени**. Но оказывается, это не мешает вычислению условной вероятности. Посмотрим, как это делается.

Если событие A произошло, то множество возможных исходов опыта сократилось: теперь вместо 36 возможных пар чисел остались только такие, которые дают в сумме 6:

$$(1,\ 5);\ (2,\ 4);\ (3,\ 3);\ (4,\ 2);\ (5,\ 1).$$

В двух из этих пяти исходов оба числа чётные, значит, $P(B \mid A) = \frac{2}{5} = 0{,}4$.

Заметим, что безусловная вероятность события B равнялась 0,25. Её можно найти, если посчитать благоприятные исходы для события B по правилу умножения: $3 \cdot 3 = 9$ (3 варианта выбрать чётное число для первого кубика, и 3 варианта — для второго). Отсюда $P(B) = \frac{9}{36} = \frac{1}{4} = 0{,}25$. Как видите, событие A сильно увеличило шансы события B — с 0,25 до 0,4.

Обобщим рассмотренный пример на произвольный опыт с n равновозможными исходами. Пусть A и B — случайные события, состоящие соответственно из n_A и n_B исходов (рис. 74). В соответствии с классическим определением вероятности $P(A) = \frac{n_A}{n}$.

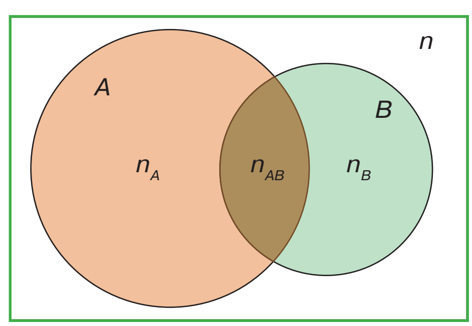


Рис. 74. Условная вероятность для опыта с равновозможными исходами

Предположим теперь, что событие B произошло. Как теперь найти условную вероятность $P(A \mid B)$? Очевидно, нужно исключить из рассмотрения все исходы, которые не попадают в B: теперь они стали невозможными. Таким образом, количество всех равновозможных исходов опыта сократилось до n_B. Благоприятными для A среди них будут те, что попали в пересечение событий $A \cap B$, — обозначим их количество n_{AB}. По классическому определению вероятности:

$$P(A \mid B) = \frac{n_{AB}}{n_B}.$$

Именно эту формулу мы и применили в примере 2 с кубиками.

Чтобы выразить условную вероятность через безусловные, преобразуем полученную дробь, умножив её числитель и знаменатель на n:

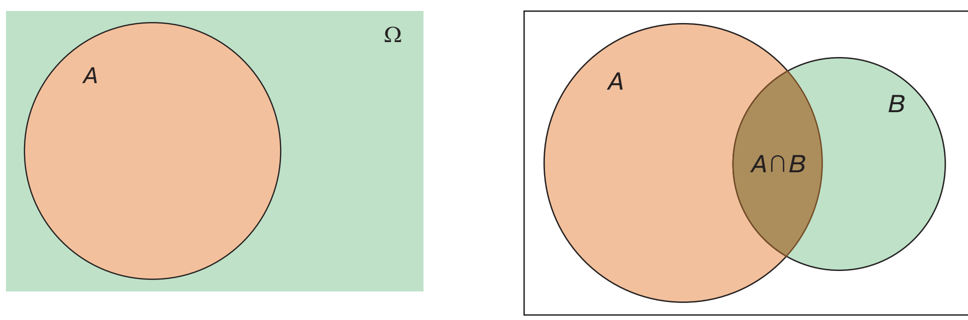
$$P(A \mid B) = \frac{n_{AB}}{n_B} = \frac{n_{AB}}{n} \cdot \frac{n}{n_B} = \frac{P(A \cap B)}{P(B)}.$$

Получили **формулу условной вероятности для произвольных случайных событий**:

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

Чтобы данное отношение имело смысл, должно выполняться $P(B) \neq 0$. Это требование вполне естественно: если B произошло, то вряд ли его вероятность равна 0.

Соотношение условной и безусловной вероятности поясняют две диаграммы Эйлера на рисунке 75.



а) Безусловная вероятность

б) Условная вероятность

Рис. 75

Безусловную вероятность $P(A)$ можно рассматривать как условную вероятность $P(A \mid \Omega)$ относительно множества всех возможных исходов Ω.

? ВОПРОСЫ

1. Что называется условной вероятностью события A при условии события B? Как обозначается эта вероятность?
2. Запишите формулу для вычисления условной вероятности.
3. Для каких случайных событий B не имеет смысла условная вероятность $P(A \mid B)$?
4. Чему равна условная вероятность $P(A \mid \Omega)$?

2 Вероятность пересечения событий

Если из определения условной вероятности выразить $P(A \cap B)$, то получим формулу для вероятности пересечения произвольных событий A и B, которую называют также **правилом умножения вероятностей**:

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

События A и B в этой формуле можно поменять местами:

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

При этом в качестве условия удобно выбрать то из двух событий, которое по времени происходит раньше. Покажем, как пользоваться этой формулой на примерах.

Вернёмся к примеру с орехами.

Пример 1. В кармане у Макара лежат 6 орехов, половина из которых пустые. Он вынимает из кармана и раскалывает один за другим два ореха. С какой вероятностью оба ореха будут полными?

Рассмотрим снова два события:

A = «первый орех будет полным»,
B = «второй орех будет полным».

Мы уже показали, что $P(A) = \frac{3}{6} = \frac{1}{2}$, $P(B \mid A) = \frac{2}{5}$. Отсюда по правилу умножения вероятностей $P(A \cap B) = P(A) \cdot P(B \mid A) = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5} = 0{,}2$.

Правило умножения можно обобщить на три и большее число событий. Продолжим наш пример с орехами. Пусть Макар вынимает теперь три ореха подряд. С какой вероятностью все они будут полными?

Введём третье событие C = «третий орех будет полным». Найдём вероятность события C при условии, что A и B произошли, т. е. условную вероятность $P(C \mid A \cap B)$. Поскольку из кармана вынуты 2 полных ореха, то там осталось 4 ореха, среди которых один полный, поэтому $P(C \mid A \cap B) = \frac{1}{4}$. Теперь применим правило умножения к событиям $A \cap B$ и C:

$$P(A \cap B \cap C) = P(A \cap B) \cdot P(C \mid A \cap B) = \frac{1}{5} \cdot \frac{1}{4} = \frac{1}{20} = 0{,}05.$$

Шансы получились не очень большие.

Решая эту задачу, мы получили **правило умножения для трёх событий**:

$$P(A \cap B \cap C) = P(A) \cdot P(A \mid B) \cdot P(C \mid A \cap B).$$

Аналогично можно записать его для четырёх и большего числа событий.

Пример 2. У Ирины на связке 5 разных ключей, из которых один подходит к двери. Она наугад пытается подобрать нужный ключ. С какой вероятностью она откроет дверь с первой попытки? с последней (т. е. пятой) попытки?

Ответить на первый вопрос не составляет труда: вероятность выбрать один нужный ключ из пяти составляет $\frac{1}{5}$.

Для ответа на второй вопрос рассмотрим пять событий A_1, A_2, A_3, A_4, A_5, каждое из которых означает, что соответствующий ключ (первый, второй, третий и т. д.) подошёл к двери. Мы уже выяснили, что $P(A_1) = \frac{1}{5}$. Чтобы дверь открылась со второй попытки, должны произойти события $\overline{A}_1$ и A_2. Найдём вероятность их пересечения по правилу умножения:

$$P(\overline{A}_1 \cap A_2) = P(\overline{A}_1) \cdot P(A_2 \mid \overline{A}_1) = \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5}.$$

Вероятность получилась такая же. А теперь для третьей попытки:

$$P(\overline{A}_1 \cap \overline{A}_2 \cap A_3) = P(\overline{A}_1) \cdot P(\overline{A}_2 \mid \overline{A}_1) \cdot P(A_3 \mid \overline{A}_1 \cap \overline{A}_2) = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{5}.$$

Аналогично для четвёртой:

$$P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap A_4) = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{5}$$

и, наконец, пятой попытки:

$$P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \overline{A}_4 \cap A_5) = \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{5}.$$

Мы получили интересный результат: для всех попыток — от первой до последней — вероятность оказалась одинаковой!

? ВОПРОСЫ

1. Запишите правило умножения для двух случайных событий A и B.
2. Запишите правило умножения для трёх случайных событий A, B и C.
3. Если вероятность одного из двух событий A, B равна 0, то чему равна вероятность $P(A \cap B)$?
4. Если вероятность одного из трёх событий A, B, C равна 0, то чему равна вероятность $P(A \cap B \cap C)$?

3 Независимые события

Итак, условная вероятность $P(A \mid B)$ отличается от безусловной вероятности $P(A)$ тем, что она вычисляется в новых условиях, когда событие B произошло. При этом она может измениться — стать меньше или больше безусловной, а может остаться прежней.

В случае когда $P(A \mid B) = P(A)$, естественно считать, что событие A **не зависит** от события B, поскольку B никак не влияет на его вероятность. За примерами такой независимости далеко ходить не нужно. Когда вы дважды подбрасываете монету, то результат первого испытания никак не влияет на результат второго. Вероятность того, что во втором испытании выпадет орёл, никак не зависит от того, выпал ли орёл в первом испытании, — она в любом случае равна $\frac{1}{2}$.

Интересно, что из равенства $P(A \mid B) = P(A)$ можно вывести симметричное равенство $P(B \mid A) = P(B)$:

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

Другими словами, если A не зависит от B, то и B не зависит от A. Таким образом, свойство независимости является для событий A и B взаимным.

Для независимых событий A и B правило умножения упрощается: поскольку $P(B \mid A) = P(B)$, то

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

т. е. **вероятность пересечения независимых событий равна произведению их вероятностей**.

Формула $P(A \cap B) = P(A) \cdot P(B)$ в отличие от приведённых выше равенств с условной вероятностью симметрична относительно событий A и B и допускает равенство вероятностей A и B нулю. В теории вероятностей выполнение этого равенства принимают за определение независимых событий.

Два случайных события A и B называются независимыми, если вероятность их пересечения равна произведению их вероятностей:

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

Конечно, в большинстве случаев независимость случайных событий следует из условий самого эксперимента. Так будет в опытах, где проводится несколько независимых испытаний (например, несколько раз подбрасывают монету) или участвует несколько разных, не связанных друг с другом объектов (подбрасывают одновременно несколько кубиков).

Но бывают ситуации, когда независимость можно установить (или опровергнуть), только проверив выполнение приведённого выше равенства $P(A \cap B) = P(A) \cdot P(B)$. Такие примеры мы тоже рассмотрим в дальнейшем.

Пример 1. Бросают два кубика. С какой вероятностью на первом выпадет 3 очка, а на втором 5 очков? Пусть A = {на первом кубике выпадет 3 очка}, B = {на втором кубике выпадет 5 очков}. Нужно найти $P(A \cap B)$. Поскольку события A и B связаны с разными кубиками, то можно считать их независимыми. По формуле умножения для независимых событий получаем:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Заметим, что эту задачу можно было решить и без использования независимости: у нашего опыта 36 равновозможных исходов, из которых только 1 исход благоприятен для события $P(A \cap B)$. А вот в следующем примере без независимости уже не обойтись.

Пример 2. Два стрелка делают по одному выстрелу в мишень. Вероятность попадания для первого — 0,6, для второго — 0,8. Чему равны вероятности следующих событий:

C = «оба стрелка попадут в мишень»;
D = «хотя бы один из них попадёт»;
E = «оба стрелка промахнутся»;
F = «хотя бы один из них промахнётся»?

Рассмотрим два события, вероятности которых нам заданы:

A = «первый стрелок попадёт в мишень»;
B = «второй стрелок попадёт в мишень».

В § 7 «Операции над событиями» мы уже рассматривали этот пример и нашли выражения каждого из событий C, D, E, F через события A, B — напомним их:

$C = A \cap B$;
$D = A \cup B$;
$E = \overline{A} \cap \overline{B} = \overline{A \cup B} = \overline{D}$;
$F = \overline{A} \cup \overline{B} = \overline{A \cap B} = \overline{C}$.

Из условий нашего опыта следует, что события A и B можно считать независимыми: каждый из стрелков не может повлиять своим выстрелом на результат другого. Поэтому вероятность события C можно вычислить по **правилу умножения для независимых событий**:

$$P(C) = P(A \cap B) = P(A) \cdot P(B) = 0{,}6 \cdot 0{,}8 = 0{,}48.$$

Событие D является объединением событий A и B, поэтому для вычисления вероятности $P(D)$ нужно использовать правило суммы. При этом события A и B могут произойти одновременно, т. е. пересекаются, поэтому используем **правило суммы для пересекающихся (совместных) событий**:

$$P(D) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}6 + 0{,}8 - 0{,}48 = 0{,}92.$$

Вероятность события E можно найти двумя способами. С одной стороны, E является пересечением независимых событий $\overline{A}$ и $\overline{B}$, поэтому

$$P(E) = P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B}) = (1 - P(A))(1 - P(B)) = 0{,}4 \cdot 0{,}2 = 0{,}08.$$

С другой стороны, $E = \overline{D}$, поэтому

$$P(E) = P(\overline{D}) = 1 - P(D) = 1 - 0{,}92 = 0{,}08.$$

Как видите, у нас есть разные способы вычислить вероятность одного события через вероятности других в зависимости от того, как мы выразили само это событие. Можно выбрать самый простой способ (в данном случае второй), а можно использовать несколько — это позволит проверить полученные ответы.

Для последнего события F используем самый короткий способ:

$$P(F) = P(\overline{C}) = 1 - P(C) = 1 - 0{,}48 = 0{,}52.$$

Как видите, ключом к решению примера 2 стала именно независимость событий A и B — без неё вычисление всех найденных нами вероятностей было бы невозможным.

Теперь рассмотрим более сложный пример, в котором независимость никак не следует из условий опыта и её нужно установить, применив определение независимых событий, т. е., проверив выполнение равенства $P(A \cap B) = P(A) \cdot P(B)$.

Пример 3. а) Из колоды в 36 карт вытягивают одну карту. Будут ли независимыми события A = «вытянут пику» и B = «вытянут даму»?

Здесь нельзя сразу сделать вывод о независимости: оба события связаны с одним и тем же испытанием и одним и тем же предметом. Будем действовать по определению. Для этого нужно найти три вероятности — $P(A)$, $P(B)$, $P(A \cap B)$ — и проверить выполнение равенства $P(A \cap B) = P(A) \cdot P(B)$.

Всего карт в колоде — 36, дам — 4, пик — 9, пиковая дама — одна. Поэтому по классическому определению вероятности получаем:

$$P(A) = \frac{9}{36} = \frac{1}{4};\ P(B) = \frac{4}{36} = \frac{1}{9};\ P(A \cap B) = \frac{1}{36}.$$

Подставляем найденные вероятности в равенство, которое нужно проверить:

$$\frac{1}{36} = \frac{1}{9} \cdot \frac{1}{4}.$$

Равенство выполнено, поэтому события A и B независимы.

б) В эту же колоду из 36 карт подложили ещё одну карту — тройку пик. Будут ли теперь события A и B независимыми?

Всего карт в колоде — 37, дам — 4, пик — 10, пиковая дама — одна. Поэтому

$$P(A) = \frac{10}{37};\ P(B) = \frac{4}{37};\ P(A \cap B) = \frac{1}{37}.$$

Подставляем найденные вероятности в равенство, которое нужно проверить:

$$\frac{1}{37} \neq \frac{10}{37} \cdot \frac{4}{37}.$$

Равенство не выполнено, поэтому события A и B зависимы!

Как можно объяснить возникшую вдруг зависимость этих событий? Почему добавление одной лишней карты «испортило» наше равенство? Разумное объяснение найти несложно.

В пункте а): если произошло событие A (т. е. вытянутая карта оказалась пикой), это никак не изменяло шансов, что на этой карте дама, поскольку в каждой из четырёх мастей дама — одна из 9 карт.

В пункте б): если вытянутая карта оказалась пикой, то шансы, что на этой карте дама, снижаются: в пиковой масти дама — одна из 10 карт, а в остальных — одна из 9 карт. Это хорошо видно при вычислении условной вероятности:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{37} : \frac{10}{37} = \frac{1}{10} \quad < \quad P(A) = \frac{1}{9}.$$

? ВОПРОСЫ

1. Приведите примеры независимых событий.
2. Приведите примеры зависимых событий.
3. Если события A и B независимы, то чему равна условная вероятность $P(A \mid B)$? условная вероятность $P(B \mid A)$?
4. Запишите правило умножения для независимых событий.

УПРАЖНЕНИЯ

162. События A, B, изображённые на диаграмме Эйлера — Венна (рис. 76), независимы и $P(A) = 0{,}6$, $P(B) = 0{,}7$. Найдите для каждого из четырёх непересекающихся событий, получившихся на диаграмме, соответствующую вероятность p_1, p_2, p_3, p_4.

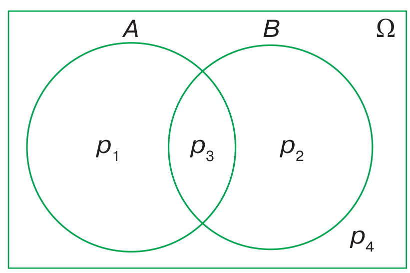


Рис. 76

163. Про случайные события A и B известно, что они независимые и что $P(A) = 0{,}3$, $P(B) = 0{,}4$. Найдите вероятности событий $P(\overline{A})$, $P(\overline{B})$, $P(A \cap B)$, $P(A \cup B)$.

164. В ящике лежит 6 красных и 12 синих шаров. Из него один за другим вынимают 2 шара. С какой вероятностью:

а) оба шара будут красными;
в) шары будут одного цвета;
б) оба шара будут синими;
г) шары будут разных цветов?

165. В корзине 10 красных и 5 зелёных яблок. Из неё наугад извлекают 2 яблока. Какова вероятность, что они разного цвета?

166. В коробке лежат 10 фломастеров, из которых 3 уже закончились, а 7 продолжают писать. Фломастеры вытаскивают из коробки один за другим наугад. С какой вероятностью фломастер, который не пишет, появится первый раз третьим по счёту?

167. Каждый из 4 друзей, которые собрались за город, случайно садится в один из 10 вагонов электрички. Найдите вероятность того, что все они окажутся в разных вагонах.

168. Каждый из 5 учеников случайно записывается на один из 8 предложенных факультативов. С какой вероятностью хотя бы двое из них выберут один и тот же факультатив?

169. Кубик подбрасывают 6 раз. Найдите вероятность того, что выпадет хотя бы одна 3.

170. Шесть школьников случайным образом рассаживаются на шесть свободных мест. С какой вероятностью Женя и Таня будут сидеть рядом, если эти шесть мест расположены:
а) за круглым столом;
б) на скамейке?

171. Семизначные номера телефонов могут изменяться от 100-00-00 до 999-99-99. С какой вероятностью в случайно выбранном номере не будет нулей? не будет девяток?

172. Трёхзначные номера автомобилей могут изменяться от 000 до 999. С какой вероятностью в случайно взятом номере будет хотя бы одна цифра 9?

173. Одновременно бросают 3 кубика. Какова вероятность того, что на всех кубиках выпадут одинаковые числа? все числа на кубиках будут разные? выпадет ровно два одинаковых числа?

174*. Выясните, с какими из перечисленных ниже событий B, C, D независимо событие:
A = «на первом кубике выпало 5 очков»;
B = «максимальное из двух чисел чётно»;
C = «сумма очков чётна»;
D = «произведение очков чётно».

175*. Вероятность того, что первое воскресенье сентября будет дождливым, равна 0,6. Вероятность, что каждое следующее воскресенье сентября будут дождливым при условии, что предыдущее было дождливым, равна 0,8, а при условии, что предыдущее было не дождливым, — 0,4. С какой вероятностью три первых воскресенья сентября будут дождливыми? будут солнечными?

176*. Студенты Петров и Иванов посещают лекции независимо друг от друга, причём Петров чаще, чем Иванов. Установлено, что вероятность их совместного появления на лекции равна 0,02, а вероятность того, что ни один не придёт на лекцию, равна 0,72. Найдите вероятности появления на лекции для каждого из студентов.

177*. Спортсмен-биатлонист должен поразить 3 мишени четырьмя выстрелами. Каждый выстрел попадает в цель с вероятностью 0,8. С какой вероятностью биатлонисту не придётся бежать штрафные круги?

178*. Чтобы пройти в следующий круг соревнований, футбольной команде нужно набрать хотя бы 4 очка в двух играх. Если команда выигрывает, она получает 3 очка, в случае ничьей — 1 очко, если проигрывает — 0 очков. Найдите вероятность того, что команде удастся выйти в следующий круг соревнований, если в каждой игре вероятности выигрыша и проигрыша одинаковые и равны 0,3.

179*. Кубик бросили 3 раза. Известно, что в сумме выпало 6 очков. Какова вероятность того, что хотя бы раз выпало 3 очка?

180*. При подбрасывании двух кубиков в сумме выпало 8 очков. Какова вероятность того, что хотя бы на одном кубике выпало 4 очка?

181*. Олег бросает кубик до тех пор, пока не наберёт в сумме больше 4 очков. Сколько бросков он, скорее всего, сделает?

182*. В викторине участвуют 6 команд. Все команды разной силы, и в каждой встрече выигрывает та команда, которая сильнее. В первом раунде встречаются две случайно выбранные команды. Ничьих не бывает. Проигравшая команда выбывает из викторины, а победившая команда играет со следующим случайно выбранным соперником. Известно, что в первых трёх играх победила команда «Умники». С какой вероятностью эта команда выиграет четвёртый раунд?

183*. Турнир по теннису проводится по олимпийской системе: игроки случайным образом разбиваются на игровые пары; проигравший в каждой паре выбывает из турнира, а победитель выходит в следующий тур, где встречается со следующим противником, который определён жребием. Всего в турнире участвует 16 игроков, все они играют одинаково хорошо, поэтому в каждой встрече вероятность выигрыша и поражения у каждого игрока равна 0,5. Среди игроков есть два друга. Какова вероятность того, что они встретятся между собой в каком-то туре?

§ 10. Полная вероятность и формула Байеса

1 Дерево вероятностей

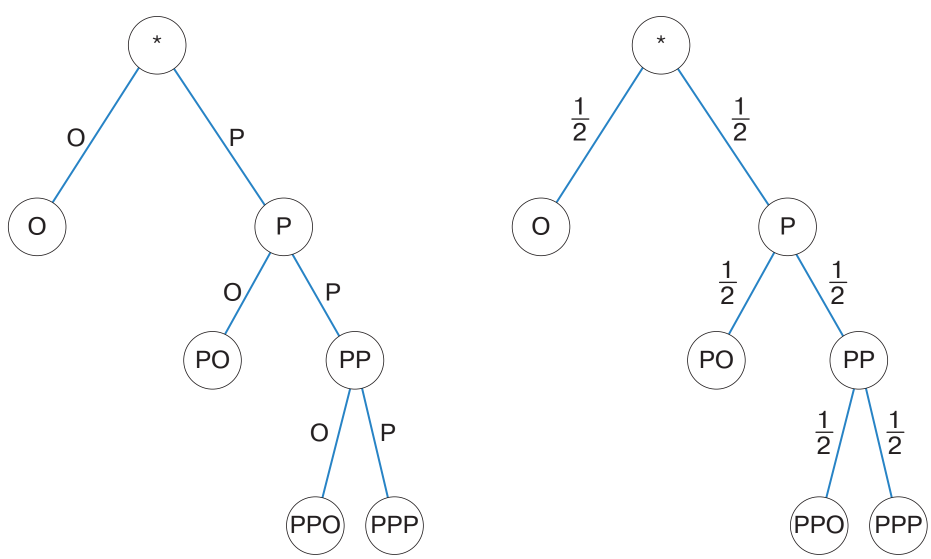
В главе 2 «Элементы теории графов» мы познакомились с важной разновидностью графов, которые называются деревьями, и научились строить ***дерево случайного эксперимента***.

Напомним, что в корне такого дерева находится начальное состояние всего эксперимента, когда никакие действия ещё не производились. Вершинами служат состояния, в которые мы попадаем в процессе проведения эксперимента, а рёбрами — возможные переходы из одного состояния в другое. Построение дерева эксперимента помогает найти и перечислить все его исходы: ими служат конечные вершины, т. е. листья дерева.

Оказывается, дерево эксперимента можно использовать и для вычисления вероятностей любых событий, связанных с этим экспериментом. Вернёмся к эксперименту, для которого мы уже строили дерево.

Пример 1. Монету подбрасывают до появления орла, но не больше трёх раз. С какой вероятностью орёл появится?

Раньше при изображении дерева эксперимента (рис. 77, *а*) мы помечали каждое ребро тем исходом, который переводит нас в следующее состояние, например: если мы находимся в вершине РР (т. е. два раза уже выпала решка), то ребро О переводит нас в состояние РРО, а ребро Р — в состояние РРР. Договоримся теперь писать на каждом ребре вероятность такого перехода (рис. 77, *б*). Для нашего примера все эти вероятности будут равны $\frac{1}{2}$. Полученное дерево называют ***деревом вероятностей***.



а) дерево эксперимента б) дерево вероятностей

Рис. 77

Заметим, что такое дерево всегда обладает одним замечательным свойством: для любой вершины сумма вероятностей на всех выходящих (т. е. направленных вниз) рёбрах равна 1. Исключение составляют только конечные вершины-листья, на которых весь эксперимент завершается и из которых не выходит ни одного ребра.

По свойству дерева, доказанному ранее, **к любой его вершине ведёт единственная цепь из корня**. Она показывает, какая цепочка случайных событий должна произойти, чтобы попасть в данное состояние. Например, чтобы попасть в вершину РРО нужно, чтобы в первом и втором испытаниях выпала решка, а в третьем — орёл.

Поскольку события в любой цепочке относятся к разным испытаниям, то они независимы, и поэтому вероятность их пересечения можно найти как произведение их вероятностей. Таким образом, чтобы посчитать вероятность попадания в любую вершину, нужно перемножить все вероятности на пути от корня к этой вершине. Например, вероятность попадания в вершину РРО будет равна $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$.

Вероятность любого случайного события теперь можно получить, сложив вероятности всех благоприятных для него конечных вершин-исходов. Для события A = «орёл появится» благоприятными являются вершины О, РО, РРО. Найдём их вероятности по правилу умножения и сложим:

$$P(A) = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}.$$

В этом примере все следующие друг за другом испытания были независимы. Однако дерево вероятностей можно использовать и для зависимых событий. В этом случае нужно записывать на каждом ребре **условную вероятность очередного события при условии, что произошли все предыдущие события в цепочке**.

Пример 2. В ящике лежат 5 спелых груш и 3 неспелых. Соня вытаскивает из него по одной груше до тех пор, пока не вытащит спелую. Какое количество груш, вероятнее всего, будет вынуто из ящика?

Чтобы ответить на этот вопрос, нарисуем для данного эксперимента дерево вероятностей (рис. 78). Буква С на нём соответствует спелой груше, а буква Н — неспелой.

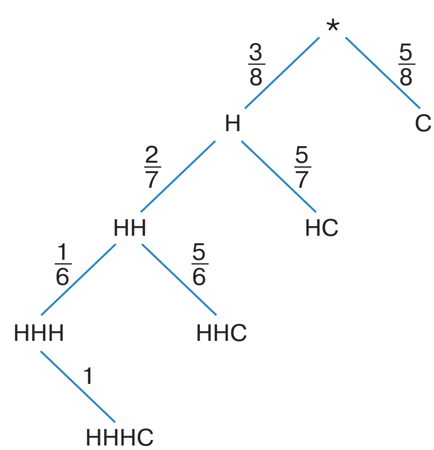


Рис. 78. Дерево вероятностей для опыта с грушами

На рёбрах дерева выписаны **условные вероятности** каждого очередного события в соответствующей цепочке. Например, чтобы попасть в вершину ННС, нужно вытащить друг за другом 2 неспелых, а потом спелую грушу. На рёбрах пути от корня «*» в вершину «ННС» записаны следующие вероятности:

$$P(\text{Н}) = \frac{3}{8};\ P(\text{Н} \mid \text{Н}) = \frac{2}{7};\ P(\text{С} \mid \text{НН}) = \frac{5}{6}.$$

Посмотрим, как они получены.

P(Н) — это вероятность вытащить первую неспелую грушу из корзины. Она равна $\frac{3}{8}$, поскольку в начальный момент в корзине 8 груш, из которых 3 неспелых.

P(Н | Н) — это вероятность вытащить вторую неспелую грушу при условии, что первая оказалась неспелой. Она равна $\frac{2}{7}$, поскольку в этот момент в корзине 7 груш, из которых 2 неспелых (одну неспелую вытащили на первом шаге).

P(С | НН) — это вероятность вытащить на третьем шаге спелую грушу при условии, что первая и вторая оказались неспелыми. Она равна $\frac{5}{6}$, поскольку в этот момент в корзине 6 груш, из которых 5 спелых (пока все спелые груши находятся в корзине).

Аналогично получены и все остальные вероятности. Теперь можем вычислить вероятность каждого из четырёх возможных исходов. Для этого нужно перемножить все вероятности, написанные на рёбрах от корня до соответствующей вершины:

$$P(\text{С}) = \frac{5}{8}; \qquad P(\text{ННС}) = \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{5}{6} = \frac{5}{56};$$

$$P(\text{НС}) = \frac{3}{8}\cdot\frac{5}{7} = \frac{15}{56}; \qquad P(\text{НННС}) = \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{1}{6}\cdot 1 = \frac{1}{56}.$$

Получаем, что максимальную вероятность имеет исход *С*, т. е., наиболее вероятно, что спелая груша попадётся Соне с первого раза. Интересно, что так будет всегда, кроме одного случая, когда спелая груша всего одна. Тогда её с равной вероятностью можно будет вынуть на любом шаге (вспомните пример, в котором девочка подбирала ключи к двери).

В примере 2 нам понадобилось только перемножать вероятности, расставленные на рёбрах. Рассмотрим задачу, в которой полученные произведения нужно будет ещё складывать.

Пример 3. На заводе микроэлектроники 8% микросхем производятся с браком. Система контроля, через которую проходят все микросхемы, выявляет этот брак с вероятностью 0,95. Какой процент микросхем попадает к заказчику? С какой вероятностью схема, поступившая к заказчику, будет бракованной?

Случайный эксперимент в этом примере состоит из двух этапов (рис. 79): сначала микросхему производят, а потом проверяют в системе контроля. На первом этапе может произойти одно из двух событий: либо

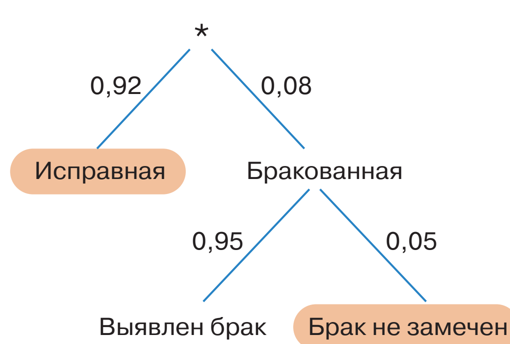


Рис. 79. Дерево вероятностей к примеру 3

схема с вероятностью 0,92 окажется исправной, либо с вероятностью 0,08 окажется бракованной. После этого схема поступает в систему контроля. Если это исправная схема, то она отправляется заказчику, а если она бракованная, то может с вероятностью 0,95 быть отправлена в утиль, а может всё-таки с вероятностью 0,05 отправиться к заказчику.

Таким образом, благоприятными для события A = «микросхема попадёт к заказчику» будут две вершины дерева: «исправная» и «брак не замечен». Они выделены на дереве цветом. Чтобы найти вероятность события A, нужно сложить вероятности этих двух исходов:

$$P(A) = 0{,}92 + 0{,}08 \cdot 0{,}05 = 0{,}924.$$

Таким образом, к заказчику попадает 92,4% произведённых микросхем — все исправные и те бракованные, которые не заметила система контроля.

Рассмотрим событие B = «микросхема будет бракованной». Чтобы ответить на второй вопрос, нам нужно вычислить условную вероятность $P(B \mid A)$. По определению условной вероятности

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0{,}08 \cdot 0{,}05}{0{,}924} = 0{,}004329.$$

Т. е. заказчик получит меньше, чем полпроцента бракованных микросхем, — система контроля качества работает достаточно надёжно.

Рассмотрим более сложный пример, в котором «благоприятных листьев» на дереве эксперимента гораздо больше.

Пример 4. Алексей приехал на каникулах в деревню и пошёл в лес. Он идёт по тропинкам, пока не найдёт гриб — мухомор или белый. Схема тропинок изображена на рисунке 80. На любой развилке он выбирает следующую тропинку случайным образом. Сорвав гриб, он возвращается домой тем же путём, что пришёл. С какой вероятностью Алексей принесёт домой белый гриб?

Здесь дерево вероятностей дано в самом условии, только вероятности на нём не написаны. Они зависят от количества тропинок, на которые разделяется дальнейший путь. При выходе из дома путь разделяется на две тропинки — значит, выбор каждой из них происходит с вероятностью $\frac{1}{2}$. В точке A происходит разветвление на 5 тропинок, поэтому каждая из

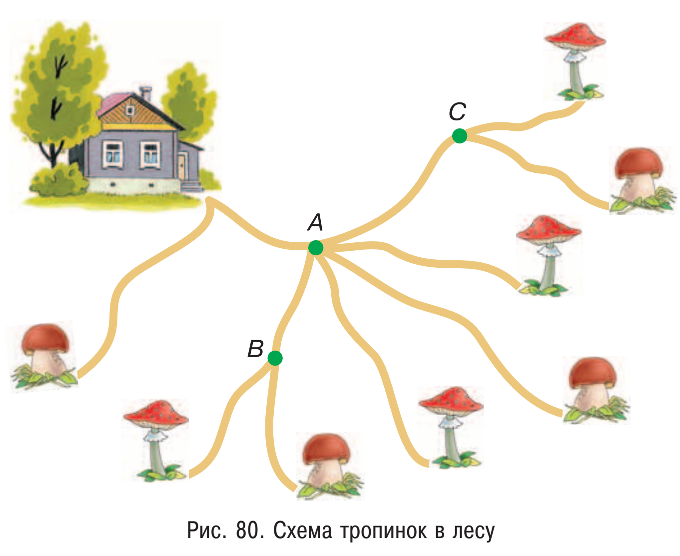


Рис. 80. Схема тропинок в лесу

них может быть выбрана с вероятностью $\frac{1}{5}$. В точках B и C расходятся 2 тропинки, поэтому соответствующие вероятности равны $\frac{1}{2}$.

Чтобы найти вероятность, с которой Алексей вернётся домой с белым грибом, нужно найти вероятности дойти до каждого из четырёх белых грибов, а потом их сложить:

$$\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{5}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{5} + \frac{1}{2}\cdot\frac{1}{5}\cdot\frac{1}{2} = \frac{1}{2} + \frac{1}{20} + \frac{1}{10} + \frac{1}{20} = \frac{7}{10}$$

(мы обходили четыре белых гриба против часовой стрелки, начиная с самого ближнего к дому). Как видите, вероятность получилась значительно больше, чем для мухомора, хотя белых грибов и мухоморов в лесу растёт поровну. Это связано, конечно, с устройством тропинок в лесу.

? ВОПРОСЫ

1. Какой граф называется деревом?
2. Какие вершины на дереве вероятностей соответствуют элементарным событиям?
3. Как с помощью дерева вероятностей вычислить вероятность элементарного события?
4. Как с помощью дерева вероятностей вычислить вероятность произвольного случайного события?

2 Формула полной вероятности

Нам уже встречались примеры, в которых условную вероятность посчитать проще, чем безусловную. Напомним один из них.

Пример 1. В кармане у Макара лежат 6 орехов, 2 из которых пустые. Он вынимает из кармана и раскалывает один за другим 2 ореха. Рассмотрим события:

A = {первый орех будет полным},
B = {второй орех будет полным}.

Чему равна вероятность $P(B)$? Если бы мы знали, произошло или нет событие A, то ответ легко вычислялся бы по классическому определению вероятности:

$$P(B \mid A) = \frac{3}{5},\ P(B \mid \overline{A}) = \frac{4}{5}.$$

Но это **условные** вероятности, а как найти по ним **безусловную** вероятность $P(B)$? Может быть, взять их среднее арифметическое? В нашем примере оно будет равно

$$\frac{1}{2}\left(\frac{3}{5} + \frac{4}{5}\right) = \frac{7}{10} = 0{,}7.$$

Но этот ответ неверный — мы уже выяснили в § 9, что $P(B) = P(A) = \frac{2}{3}$. Тем не менее найденные условные вероятности $P(B \mid A)$ и $P(B \mid \overline{A})$ можно использовать для вычисления безусловной вероятности $P(B)$, если применить ***формулу полной вероятности***, которую мы сейчас докажем.

Введём сначала одно новое понятие.

Набор случайных событий H_1, H_2, ..., H_n называется *полной группой событий*, если все эти события попарно несовместны, а их объединение равно множеству всех исходов Ω.

Можно сказать, что полная группа событий — это разбиение множества исходов Ω на непересекающиеся множества (рис. 81).

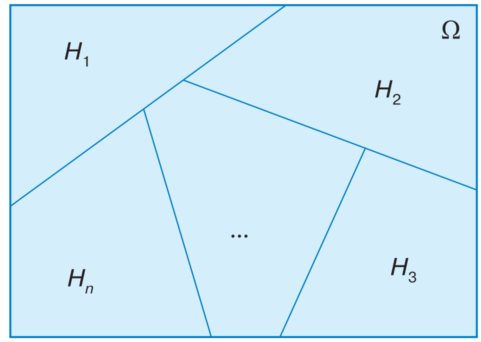


Рис. 81. Полная группа событий

Самым простым (и часто используемым) примером полной группы событий является набор из двух событий: H и $\overline{H}$.

Рассмотрим произвольное случайное событие A. Полная группа событий $H_1, H_2, \ldots, H_n$ разбивает A на непересекающиеся подмножества:

$$A = AH_1 + AH_2 + \ldots + AH_n$$

(чтобы упростить запись, мы заменили знак объединения на знак «плюс», а знак пересечения — на «умножить»). Полученное разбиение можно увидеть на рисунке 82.

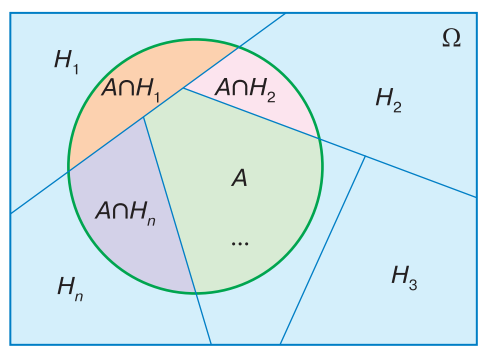


Рис. 82. Разбиение события A

Отсюда по формуле суммы вероятностей для несовместных событий получаем:

$$P(A) = P(AH_1) + P(AH_2) + \ldots + P(AH_n).$$

Теперь применим к каждому слагаемому формулу произведения вероятностей:

$$P(AH_1) = P(H_1)\,P(A \mid H_1),\ P(AH_2) = P(H_2)\,P(A \mid H_2),\ \ldots,$$
$$P(AH_n) = P(H_n)\,P(A \mid H_n).$$

Окончательно получаем следующую формулу, которая называется *формулой полной вероятности* и выражает безусловную вероятность события A через условные вероятности:

$$P(A) = P(H_1)\,P(A \mid H_1) + P(H_2)\,P(A \mid H_2) + \ldots + P(H_n)\,P(A \mid H_n).$$

В частном случае, когда полная группа событий состоит из H и $\overline{H}$, эта формула выглядит так:

$$P(A) = P(H)\,P(A \mid H) + P(\overline{H})\,P(A \mid \overline{H}).$$

Применим эту формулу для вычисления вероятности $P(B)$ в примере 1:

$$P(B) = P(A)\,P(B \mid A) + P(\overline{A})\,P(B \mid \overline{A}) = \frac{4}{6} \cdot \frac{3}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{2}{3}.$$

Получился верный ответ, который мы уже нашли раньше другим способом.

В следующем примере формула полной вероятности помогает ответить на вполне практический вопрос.

Пример 2. Готовясь к экзамену, Борис выучил только 15 билетов из 20. Каким по счёту ему лучше заходить на экзамен — первым или вторым?

Назовём те билеты, которые выучил Борис, «счастливыми». Нам нужно сравнить вероятности двух событий:

A = «первый ученик вытащит «счастливый» билет»;
B = «второй ученик вытащит «счастливый» билет».

С вероятностью первого события всё понятно: $P(A) = \frac{15}{20} = \frac{3}{4}$. Для события B легко найти условные вероятности:

$$P(B \mid A) = \frac{14}{19},\ P(B \mid \overline{A}) = \frac{15}{19}.$$

Отсюда по формуле полной вероятности

$$P(B) = \frac{3}{4} \cdot \frac{14}{19} + \frac{1}{4} \cdot \frac{15}{19} = \frac{57}{76} = \frac{3}{4}.$$

Как видите, ответ не изменился. Такой же результат мы получим для третьего, четвёртого и любого другого ученика. Повысить шансы успешной сдачи экзамена можно только одним способом: выучить побольше билетов.

В следующем примере полная группа будет состоять из трёх событий.

Пример 3. Смартфоны марки «Бетафон» производятся на трёх заводах. При этом 20% всей продукции выпускает первый завод, 30% — второй, 50% — третий. Процент брака на каждом из этих заводов составляет соответственно 4%, 2% и 1%. Какова вероятность, что случайно выбранный смартфон будет бракованным?

Введём событие A = «смартфон будет бракованным».

Для вычисления $P(A)$ рассмотрим полную группу событий:

H_1 = «смартфон был выпущен на первом заводе»;
H_2 = «смартфон был выпущен на втором заводе»;
H_3 = «смартфон был выпущен на третьем заводе».

В задаче даны вероятности каждого из этих событий:

$$P(H_1) = 0{,}2,\ P(H_2) = 0{,}3,\ P(H_3) = 0{,}5,$$

а также условные вероятности:

$$P(A \mid H_1) = 0{,}04,\ P(A \mid H_2) = 0{,}02,\ P(A \mid H_3) = 0{,}01.$$

Чтобы найти $P(A)$, применим формулу полной вероятности:

$$P(A) = 0{,}2 \cdot 0{,}04 + 0{,}3 \cdot 0{,}02 + 0{,}5 \cdot 0{,}01 = 0{,}008 + 0{,}006 + 0{,}005 = 0{,}019.$$

Обратите внимание, что безусловная вероятность оказалась больше минимальной (0,01) и меньше максимальной (0,04) из условных вероятностей. Это свойство будет выполнено всегда. Мы докажем его в упражнении 198*.

Формулу полной вероятности можно изобразить в виде дерева вероятностей. Для полной группы из трёх событий это дерево изображено на рисунке 83.

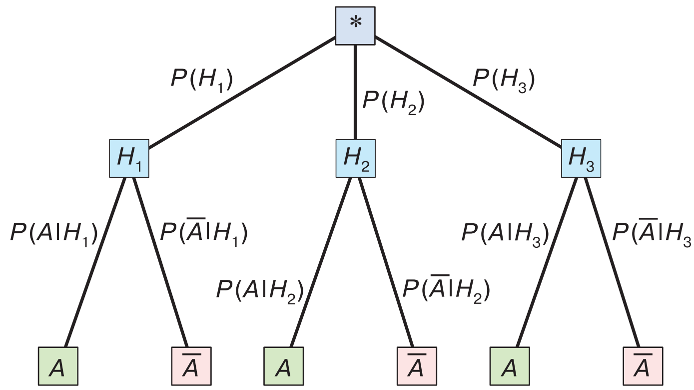


Рис. 83. Дерево для формулы полной вероятности

Чтобы найти с помощью этого дерева вероятность $P(A)$, нужно найти все конечные вершины, благоприятные для события A, вычислить с помощью правила умножения их вероятности и сложить. Результатом и будет формула полной вероятности.

? ВОПРОСЫ

1. Что такое полная группа событий?
2. Приведите пример полной группы событий.
3. Запишите формулу полной вероятности.
4. В каких границах лежит $P(A)$ в формуле полной вероятности?

3 Формула Байеса*

Британский математик и священник Томас Байес (1702—1761) опубликовал при жизни всего две работы — одну богословскую, другую математическую. Тем не менее сегодня без его имени не обходится ни один учебник по теории вероятностей. В его работе, опубликованной в 1763 г. уже после смерти автора, был предложен «метод для корректировки убеждений, основанный на обновлённых данных».

Чтобы пояснить суть этого метода, вернёмся к примеру со смартфонами, рассмотренному в предыдущем пункте.

Пример 1. Напомним, что смартфоны марки «Бетафон» производятся на трёх заводах: 20% всей продукции выпускает первый завод, 30% — второй, 50% — третий. Процент брака на каждом из этих заводов составляет соответственно 4%, 2% и 1%.

Мы уже нашли вероятность того, что случайно выбранный смартфон этой марки будет бракованным: по формуле полной вероятности она равна 0,019.

Представим теперь, что вы купили смартфон этой марки и он оказался бракованным. С какой вероятностью он был сделан на первом заводе? на втором? на третьем?

Вопрос может показаться бессмысленным — ведь достаточно открыть технический паспорт этого изделия и найти в нём адрес производителя. Но допустим, что такого паспорта нет и производитель нигде не указан. Получается, что вам **нужно пересмотреть первоначальные вероятности $P(H_1)$, $P(H_2)$, $P(H_3)$ на основе новых данных**. Эти новые данные состоят в том, что произошло событие A «выбранный смартфон оказался бракованным». Следовательно, в качестве «пересмотренных» нужно взять условные вероятности

$$P(H_1 \mid A),\ P(H_2 \mid A),\ P(H_3 \mid A).$$

Попробуем вычислить первую из них. По определению условной вероятности

$$P(H_1 \mid A) = \frac{P(AH_1)}{P(A)}.$$

Найдём вероятность пересечения событий H_1 и A, используя формулу произведения вероятностей:

$$P(AH_1) = P(H_1)\,P(A \mid H_1) = 0{,}2 \cdot 0{,}04 = 0{,}008.$$

Теперь подставим найденные значения $P(H_1A)$ и $P(A)$ в формулу для условной вероятности:

$$P(H_1 \mid A) = \frac{0{,}008}{0{,}019} = 0{,}421.$$

Как видите, условная вероятность $P(H_1 \mid A)$ оказалась в два с лишним раза больше безусловной $P(H_1)$. Причина здесь вполне объяснима: первый завод производит больше всего брака и поэтому первоначальная вероятность пересмотрена в сторону её увеличения.

Аналогично можно вычислить две другие условные вероятности:

$$P(H_2 \mid A) = \frac{P(H_2) \cdot P(A \mid H_2)}{P(A)} = \frac{0{,}3 \cdot 0{,}02}{0{,}019} = 0{,}316;$$

$$P(H_3 \mid A) = \frac{P(H_3) \cdot P(A \mid H_3)}{P(A)} = \frac{0{,}5 \cdot 0{,}01}{0{,}019} = 0{,}263.$$

Условная вероятность $P(H_1 \mid A)$ почти не изменилась по сравнению с безусловной, а $P(H_1 \mid A)$ почти в два раза уменьшилась, так как третий завод производит наиболее качественную продукцию.

Сведём полученные результаты в таблицу 49.

Таблица 49

До опыта	После опыта
$P(H_1) = 0{,}2$	$P(H_1 \mid A) = 0{,}421$
$P(H_2) = 0{,}3$	$P(H_2 \mid A) = 0{,}316$
$P(H_3) = 0{,}5$	$P(H_3 \mid A) = 0{,}263$

В первом столбце выписаны вероятности событий H_1, H_2, H_3 до проведения опыта — т. е. до выяснения, какой смартфон вам достался. Во втором столбце выписаны вероятности тех же событий после проведения опыта, который закончился результатом A = «выбранный смартфон оказался бракованным».

Вероятности в первом столбце часто называют ***априорными*** от латинского *a priori* — «знание, полученное до опыта». Вероятности во втором столбце называют ***апостериорными*** от латинского *a posteriori* — «знание, полученное после опыта». Сами события $H_1, \ldots, H_n$, которые составляют полную группу, называют при этом ***гипотезами*** (отсюда и обозначение этих событий буквами H — от латинского *hypothesis*).

В общем случае ***формула Байеса*** для полной группы событий $H_1, \ldots, H_n$ и произвольного события A записывается так:

$$P(H_k \mid A) = \frac{P(H_k)\,P(A \mid H_k)}{P(H_1)\,P(A \mid H_1) + \ldots + P(H_n)\,P(A \mid H_n)}.$$

Знаменатель дроби — это выражение для вероятности $P(A)$, полученное по формуле полной вероятности. Оно представляет собой сумму из n слагаемых. В числителе находится одно из этих слагаемых, а именно то, которое соответствует событию H_k.

В каком-то смысле формула Байеса переставляет причину и следствие: по известному свершившемуся факту A вычисляется вероятность того, что он был вызван причиной H_k.

В современном виде формула Байеса была опубликована только в 1812 г. одним из основателей теории вероятностей — французским математиком П. Лапласом (1749—1827). Многие учёные посчитали, что эта формула открывает новый путь к научному познанию, так как позволяет пересмотреть уровень доверия к любым первоначальным гипотезам H_1, ..., H_n. Но, как часто бывает в таких случаях, такое обобщение оказалось сильно преувеличенным.

Новая волна интереса к формуле Байеса поднялась уже в последние десятилетия и была связана с её применением в технологиях искусственного интеллекта. Сегодня так называемый ***байесовский подход*** лежит в основе всего машинного обучения.

Рассмотрим ещё несколько примеров применения этой замечательной формулы.

Пример 2. Известно, что некоторое заболевание встречается в среднем у одного из 1000 человек. УЗИ (ультразвуковое исследование) позволяет выявить это заболевание у больного человека с вероятностью 0,9, но при этом может обнаружить его и у здорового человека — с вероятностью 0,01. Какова вероятность, что человек, получивший заключение УЗИ о наличии заболевания, действительно болен?

Сначала разберёмся, что произошло — т. е. какую информацию мы получили. Произошло событие A = «пациент получил заключение о наличии заболевания».

Нужно найти вероятность события H = «пациент болен» при условии, что событие A произошло, т. е. $P(H \mid A)$. Рассмотрим полную группу событий H и $\overline{H}$. Из условия задачи

$$P(H) = 0{,}001, \quad P(\overline{H}) = 0{,}999.$$

Кроме того, известно, что

$$P(A \mid H) = 0{,}9, \; P(A \mid \overline{H}) = 0{,}01.$$

Отсюда по формуле полной вероятности

$$P(A) = P(H) \cdot P(A \mid H) + P(\overline{H}) \cdot P(A \mid \overline{H}) = 0{,}001 \cdot 0{,}9 + 0{,}999 \cdot 0{,}01 = 0{,}01089.$$

Для получения ответа применим формулу Байеса:

$$P(H \mid A) = \frac{P(H)\,P(A \mid H)}{P(A)} = \frac{0{,}001 \cdot 0{,}9}{0{,}01089} \approx 0{,}083.$$

Несмотря на заключение УЗИ о наличии заболевания, вероятность, что человек действительно болен, получилась совсем небольшой. Как же в этом случае поступить пациенту? Стоит ли сразу начинать лечение — ведь с вероятностью 92% он здоров?

В этом случае лучше всего провести **повторное независимое** обследование. Попробуем рассмотреть его возможные результаты и последующие выводы. Пусть события A_1 и A_2 состоят в том, что пациент получил заключение о наличии заболевания соответственно при первом и при втором обследовании. Эти события можно считать условно независимыми относительно событий H и $\overline{H}$, поэтому

$$P(A_1A_2 \mid H) = 0{,}9^2 = 0{,}81, \; P(A_1A_2 \mid \overline{H}) = 0{,}01^2 = 0{,}0001.$$

Отсюда

$$P(A) = 0{,}001 \cdot 0{,}9^2 + 0{,}999 \cdot 0{,}01^2 = 0{,}0009099,$$

и тогда

$$P(H \mid A_1A_2) = \frac{P(H)\,P(A_1A_2 \mid H)}{P(A_1A_2)} = \frac{0{,}001 \cdot 0{,}9^2}{0{,}0009099} \approx 0{,}89.$$

Вероятность, что пациент болен, выросла на целый порядок! Вот теперь нужно срочно начинать лечение. Как видите, формула Байеса имеет самое непосредственное отношение к нашей повседневной жизни.

Формулу Байеса, как и формулу полной вероятности, можно «увидеть» на дереве вероятностей. На рисунке 84 изображено дерево вероятностей для случайного события A и полной группы из трёх гипотез.

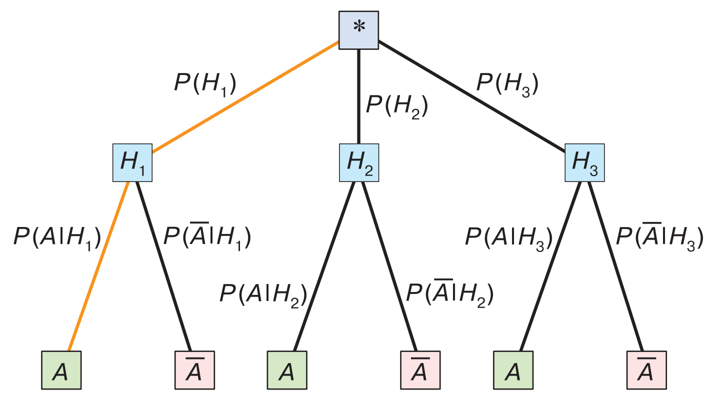


Рис. 84. Формула Байеса на дереве вероятностей

Формула Байеса показывает, какую долю составляет вероятность ветки с гипотезой H_k (на рисунке 84 выделена ветка для $k = 1$) в общей вероятности всех веток, ведущих к событию A.

ВОПРОСЫ

1. В какой ситуации применяется формула Байеса и что позволяет вычислить?
2. Что такое априорные и апостериорные вероятности гипотез?
3. Запишите формулу Байеса для гипотез H и $\overline{H}$.

УПРАЖНЕНИЯ

184. В ящике лежит 6 красных и 12 синих шаров. Из него один за другим вынимают 3 шара. С какой вероятностью:

а) первый шар будет красным; б) второй шар будет красным; в) третий шар будет красным?

185. В долине Стабильности бывают только дождливые и солнечные дни, причём с вероятностью 0,9 на следующий день сохраняется та же погода, которая была в предыдущий. 1 мая был дождь. С какой вероятностью 5 мая будет солнце?

186. В отдел технического контроля поступает партия, содержащая 20 изделий, среди которых имеется 5 бракованных. Контролёр для проверки отбирает 2 изделия, при этом в бракованном изделии он обнаруживает брак с вероятностью 0,9. Партия бракуется, если среди отобранных для проверки изделий обнаружено хотя бы одно бракованное. Найдите вероятность того, что данная партия изделий будут забракована.

187. Автоматическая линия изготавливает батарейки. Вероятность того, что готовая батарейка неисправна, равна 0,03. Перед упаковкой каждая батарейка проходит систему контроля. Вероятность, что система забракует неисправную батарейку, равна 0,95. Вероятность, что система по ошибке забракует исправную батарейку, равна 0,04. Найдите вероятность того, что случайно выбранная батарейка будет забракована системой контроля.

188. Из десяти студентов, пришедших на экзамен, Иванов и Петров знают 20 билетов из 30, Сидоров — только 15, а остальные выучили все 30 билетов. Знание билета гарантирует сдачу экзамена профессору Злобину с вероятностью 0,85, а незнание — только 0,1. С какой вероятностью случайно вызванный студент сдаст экзамен?

189*. В ящике лежат 20 теннисных мячей: 15 новых и 5 игранных. Для игры наудачу берут два мяча и после игры возвращают обратно в ящик. Для второй игры снова извлекают два мяча. С какой вероятностью оба мяча будут новыми в первой игре? во второй игре?

190*. Первый игральный кубик обычный, а на гранях второго написаны две единицы, две тройки и две пятёрки. Из двух кубиков наугад выбрали один и бросили его два раза. Известно, что в каком-то порядке выпали 3 и 5 очков. Какова вероятность, что бросали второй кубик?

191*. У вас имеются две игральные кости: правильная и фальшивая (с дробинкой внутри), на которой вероятность шестёрки — $\frac{1}{4}$, единицы — $\frac{1}{12}$, остальных чисел — по $\frac{1}{6}$. Вы выбрали из двух костей одну и подбросили — выпала шестёрка. С какой вероятностью эта кость является фальшивой?

192*. Студенты сдают в сессию три экзамена A, B, C. Вероятность сдачи по каждому из них составляет соответственно 0,6, 0,5, 0,9. Известно, что студент Разгуляев сдал не все экзамены. С какой вероятностью он сдал экзамен A? экзамен B? экзамен C?

193*. Братья-близнецы A и B иногда прогуливают занятия в школе. Брат A прогуливает около 4% всех занятий, а брат B — около 8%. Примерно 1% занятий они прогуливают одновременно. На очередном занятии присутствует только один из них. С какой вероятностью это A? с какой B?

194*. Жуки некоторого вида делятся на два подвида: обычные (99%) и редкие (1%). В редком подвиде 98% жуков имеют на крыльях узор в виде цветка, а в обычном — только 5%. Энтомолог нашёл жука с узором в виде цветка. С какой вероятностью это жук редкого подвида?

195*. Три стрелка A, B, C попадают по мишени с вероятностями 0,7, 0,8 и 0,9 соответственно, независимо друг от друга. Все стрелки выстрелили по разу, и один выстрел достиг цели. Какова вероятность, что попал стрелок A? Что A промахнулся?

196*. Перед экзаменом по географии в классе из 25 учеников имеется 10 отлично подготовленных (знают все 30 вопросов), 7 — хорошо подготовленных (20 вопросов), 5 — удовлетворительно (10 вопросов), 3 — плохо (2 вопроса). Вызванный ученик ответил на 2 вопроса. С какой вероятностью он подготовлен отлично? хорошо? удовлетворительно? плохо?

197*. При подозрении на наличие некоторого заболевания пациента отправляют на ПЦР-тест. Если заболевание действительно есть, то тест подтверждает его в 86% случаев. Если заболевания нет, то тест выявляет отсутствие заболевания в среднем в 94% случаев. Известно, что в среднем тест оказывается положительным у 10% пациентов, направленных на тестирование. При обследовании некоторого пациента врач направил его на ПЦР-тест, который оказался положительным. Какова вероятность того, что пациент действительно имеет это заболевание?

198*. Докажите, что в формуле полной вероятности безусловная вероятность $P(A)$ всегда лежит в промежутке между минимальной и максимальной из условных вероятностей $P(A|H_i)$.

ГЛАВА 5

ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

Комбинаторика изучает различные виды ***комбинаций***, способы их перечисления и подсчёта. Само слово «комбинация» происходит от латинского *combino* — «соединяю». При получении любой комбинации мы получаем её из отдельных элементов, соединяя их друг с другом. Чаще всего эти элементы выбираются из некоторого конечного множества.
Комбинаторные методы могут встречаться в задачах, которые возникают в арифметике, алгебре, геометрии и других разделах математики. Рассматривая опыты с равновозможными исходами, мы уже столкнулись с необходимостью подсчёта комбинаций при вычислении вероятностей.
В этой главе мы познакомимся с наиболее важными типами комбинаций и выведем формулы для их подсчёта.

§ 11. Перестановки и размещения

1 Перебор комбинаций

Если искомых комбинаций немного, то их можно последовательно перебрать (т. е. выписать, перечислить) одну за другой.

Пример 1. Перечислим все двузначные числа, которые можно составить из цифр 0, 1, 2:

10, 11, 12,   20, 21, 22.

Очевидно, что других комбинаций нет: искомые двузначные числа могут начинаться с 1 (таких чисел три — 10, 11, 12) или с 2 (их тоже три — 20, 21, 22).

Пример 2. Выпишем все трёхбуквенные слова, которые можно составить, используя только буквы А и Б:

ААА, ААБ, АБА, АББ,   БАА, БАБ, ББА, БББ.

Сначала мы выписали все слова, которые начинаются на букву А, а потом те, которые начинаются на Б. Тех и других слов оказалось по четыре.

Чтобы не запутаться при перечислении комбинаций, важно установить некоторое правило, по которому они перечисляются. Самое универсальное правило — **перебирать все комбинации по порядку**. Для чисел перебор по порядку означает «по возрастанию», а для слов — «по алфавиту». Если бы все 8 слов из примера 2 содержались в русском языке, то именно в таком порядке они были бы перечислены в орфографическом словаре.

Во многих задачах приходится находить комбинации, в которых любой элемент может использоваться не более одного раза. Их перебор может оказаться более сложной задачей — ведь на каждом шаге нужно учитывать, какие элементы уже использовались.

Пример 3. В школьное расписание нужно поставить четыре урока: алгебру, геометрию, литературу и русский язык. Выпишем все способы, которыми можно это сделать.

Обозначим каждый из предметов соответствующей буквой: А, Г, Л, Р. Каждому способу будет соответствовать слово, составленное из этих четырёх букв, следующих в определённом порядке. Выпишем все эти слова по алфавиту:

АГЛР, АГРЛ, АЛГР, АЛРГ, АРГЛ, АРЛГ,
ГАЛР, ГАРЛ, ГЛАР, ГЛРА, ГРАЛ, ГРЛА,
ЛАГР, ЛАРГ, ЛГАР, ЛГРА, ЛРАГ, ЛРГА,
РАГЛ, РАЛГ, РГАЛ, РГЛА, РЛАГ, РЛГА.

Получилось 24 комбинации, и их перечисление оказалось довольно долгим.

Таким образом, когда комбинаций немного, их количество можно найти прямым перебором. Но чем больше будет длина комбинаций или число используемых элементов, тем сложнее будет осуществить такой перебор.

? ВОПРОСЫ

1. Перечислите все трёхзначные числа, которые можно составить из цифр 0 и 1.
2. Перечислите все двухбуквенные слова, которые можно составить из букв М и Ж.
3. Если выписывать по алфавиту все пятибуквенные слова из букв А, Б и В, то какое слово будет следовать за словом АБВВВ?

2 Правило умножения

При большом числе комбинаций для их подсчёта используют специальные комбинаторные правила, главным из которых является ***правило умножения***. Сформулируем его сначала для двух элементов: **если первый элемент в комбинации можно выбрать a способами, после чего второй элемент — b способами, то общее число комбинаций из двух элементов будет равно $a \cdot b$.**

Вернёмся к примерам, в которых мы перечисляли различные комбинации, и попробуем найти их количество без перечисления, пользуясь сформулированным правилом.

Пример 1. Сколько двузначных чисел можно составить, если использовать только цифры 0, 1, 2?

Подсчитаем количество комбинаций по правилу умножения: первую цифру для такого числа можно выбрать двумя способами (это 1 или 2); после этого вторую цифру можно выбрать тремя способами (0, 1 или 2). Всего таких комбинаций будет $2 \cdot 3 = 6$.

Правило умножения распространяется и на общий случай, когда количество элементов в комбинации больше двух.

Правило умножения

Если комбинация должна состоять из k элементов и при этом первый элемент в комбинации можно выбрать a_1 способами, после чего второй элемент — a_2 способами, третий элемент — a_3 способами и т. д., то общее число таких комбинаций будет равно произведению k сомножителей:

$$a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_k.$$

Пример 2. Сколько трёхбуквенных слов можно составить, используя только буквы А и Б? Первую букву такого слова можно выбрать двумя способами, вторую — также двумя способами и третью — тоже двумя. Всего таких комбинаций будет $2 \cdot 2 \cdot 2 = 2^3 = 8$.

Пример 3. Сколько слов можно составить из букв А, Г, Л, Р, переставляя их между собой? Первую букву можно выбрать четырьмя способами, после этого вторую — тремя способами (первую выбрать уже нельзя), третью — двумя способами и, наконец, четвёртую — только одним. Всего таких комбинаций будет $4 \cdot 3 \cdot 2 \cdot 1 = 24$.

Как видите, во всех трёх примерах мы получили те же результаты, что и раньше, но для этого не понадобилось перечислять комбинации. Преимущества такого подсчёта становятся особенно очевидными при большом числе комбинаций.

Пример 4. Сколько существует шестизначных чисел, в которых все цифры разные?

Первую цифру шестизначного числа можно выбрать 9 способами, так как нельзя выбирать 0. После этого вторую цифру — тоже 9 способами, поскольку нельзя использовать ту цифру, что была выбрана первой. Следующую цифру — 8 способами, затем — 7 способами и т. д. Всего таких чисел будет

$$9 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 136\,080.$$

Пример 5. В компьютере каждый символ (буква, цифра, специальный знак) кодируется последовательностью из восьми цифр 0 и 1, например:

01000110 — код буквы *F*;
00110010 — код цифры «2» и т. д.

Сколько различных символов можно закодировать таким образом? Другими словами, сколько существует двоичных кодов длины 8?

Первую цифру кода можно выбрать двумя способами, после чего вторую цифру — тоже двумя способами и т. д. Всего получаем $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \times \times 2 \cdot 2 \cdot 2 = 2^8 = 256$ различных двоичных кодов. Именно столько различных символов содержит таблица ASCII, которая использовалась для компьютерного кодирования символов ещё с 60-х годов прошлого века.

С появлением современных средств коммуникации и необходимостью обмениваться текстами на разных языках недостатки такой таблицы стали очевидны: 256 двоичных кодов не хватало для кодирования всех используемых в мире национальных алфавитов. Поэтому в 1991 г. вступил в действие новый стандарт кодирования символов, получивший название Unicode. Длина двоичного кода была увеличена всего вдвое — с 8 до 16 цифр. Но количество различных комбинаций увеличилось при этом в 256 раз и стало равняться $2^{16} = 65\,536$. Этот результат легко получить по тому же правилу умножения.

В следующей задаче догадаться о том, как использовать в ней правило умножения, не так просто.

Пример 6. Сколько различных натуральных делителей имеет число $N = 3^5 \cdot 5^2 \cdot 7^3$?

Конечно, можно вычислить, чему равно число N (это будет 2 083 725), а потом перебирать все его возможные делители от 1 до N и проверять, каждое из этих чисел на делимость. Можно значительно сократить этот перебор, ограничившись диапазоном от 1 до $\sqrt{N}$, и считать делители парами (если a — делитель числа N и $a \leqslant \sqrt{N}$, то $b = \dfrac{N}{a}$ — тоже делитель и при этом $b \geqslant \sqrt{N}$).

Но ещё лучше не вычислять данное произведение, а использовать то, что оно является **разложением числа N на простые множители**. Это значит, что любой делитель d числа N записывается в виде $d = 3^x \cdot 5^y \cdot 7^z$, где x, y, z — целые числа, для которых

$$0 \leqslant x \leqslant 5, \quad 0 \leqslant y \leqslant 2, \quad 0 \leqslant z \leqslant 3.$$

Таким образом, для выбора показателя степени x у нас 6 вариантов, для выбора y — 3 варианта, для выбора z — 4 варианта. Общее количество вариантов для выбора тройки x, y, z вычисляется по правилу умножения: $6 \cdot 3 \cdot 4 = 72$. Столько же будет и делителей. Оказалось, что с помощью правила умножения задачу о подсчёте количества делителей можно решить в уме!

Правда, если вам дано не готовое разложение на простые множители, а само число N, то нужно сначала это разложение найти. Например, чтобы применить приведённый способ к числу $N = 2024$, нужно получить, что $2024 = 2^3 \cdot 11 \cdot 23$, а затем найти количество делителей по правилу умножения:

$$4 \cdot 2 \cdot 2 = 16.$$

Рассмотрим ещё один курьёзный пример использования правила умножения, связанный на этот раз с литературой. В 1961 г. французский поэт Раймон Кено выпустил книгу «Сто тысяч миллиардов стихотворений». Это был сборник, содержащий всего 10 сонетов. Откуда же взялось столь необычное название? Дело в том, что каждый из этих сонетов был разрезан на строчки.

Произвольно выбрав по одной строчке из каждого сонета, можно было каждый раз получать новое стихотворение. Форма сонета, как известно, подразумевает 14 строк. Для каждой из этих строк есть 10 вариантов выбора. По правилу умножения получаем $10^{14} = 100\ 000\ 000\ 000\ 000$ различных стихотворений. Кено назвал своё произведение «машиной для производства стихов».

? ВОПРОСЫ

1. Сформулируйте комбинаторное правило умножения.
2. Сколько двоичных кодов длины 5?
3. Объясните, почему для способа подсчёта делителей из примера 7 важно иметь разложение числа N именно на **простые** множители.

3 Перестановки и факториал

Некоторые типы комбинаций играют в математике настолько важную роль и встречаются так часто, что для них введены специальные названия и выведены формулы для подсчёта. Начнём с комбинаций, которые называются ***перестановками***.

Перестановкой из N различных элементов называют комбинацию, в которой все эти N элементов расположены в определённом порядке.

Таким образом, перестановки из одного и того же набора элементов различаются между собой не составом (он у них одинаковый), а **порядком** следования элементов в комбинации.

Элементами, которые участвуют в перестановке, могут быть числа, буквы, шары и вообще любые объекты. Выпишем для примера все перестановки из чисел 1, 2, 3:

123, 132, 213, 231, 312, 321.

Рассмотрим все перестановки из трёх букв a, b, c:

abc, acb, bac, bca, cab, cba.

Можно взять в качестве элементов перестановок три разноцветных шара (рис. 85) или три смайлика (рис. 86).

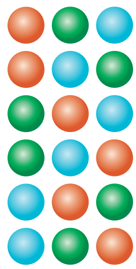


Рис. 85

Рис. 86

Конечно, во всех этих случаях количество перестановок будет одинаковым, ведь оно зависит только от того, сколько элементов участвует в перестановке. Мы видим, что из трёх элементов можно составить 6 разных перестановок. Понятно, что из двух элементов перестановок будет 2.

Выведем общую формулу для количества перестановок из N элементов. Будем использовать для этого правило умножения. Первый элемент перестановки можно выбрать N способами; после этого второй элемент — $N - 1$ способами (поскольку один элемент уже выбран); третий — $N - 2$ способами и т. д. до последнего элемента, который можно будет выбрать только одним способом (это единственный элемент, который ещё не был выбран). По правилу умножения общее количество комбинаций будет равно произведению $N \cdot (N - 1) \cdot (N - 2) \cdot ... \cdot 1$.

Мы получили, что количество всех перестановок из N элементов равно произведению всех натуральных чисел от 1 до N. Это произведение называется в математике ***факториалом*** числа N (от латинского *factorialis* — «умножающий») и обозначается $N!$:

$$\boxed{N! = 1 \cdot 2 \cdot 3 \cdot ... \cdot N}$$

Отметим также, что значение 0! полагается равным 1.

Выпишем несколько значений факториала для возрастающих значений N (табл. 50).

Таблица 50

N	***N*!**	***N***	***N*!**
0	1	8	40320
1	1	9	362880
2	2	10	3628800
3	6	11	39916800
4	24	12	479001600
5	120	13	6227020800
6	720	...	...
7	5040	25	15511210043330985984000000

Вы видите, что с ростом N факториал очень быстро растёт, а вместе с ним растёт и количество перестановок, которые можно получить из N элементов. В комбинаторике факториал является одной из важнейших функций и ещё не раз встретится нам в формулах для подсчёта комбинаций.

В электронных таблицах есть специальная функция для вычисления $N!$ — это ФАКТР(N). Но учтите, что начиная с некоторого N (обычно с $N = 15$) эта функция даёт приближённое значение факториала, поскольку разрядность полученного результата превышает максимально допустимую в электронной таблице.

Количество перестановок из N элементов обозначается P_N (буква P берётся от французского слова *permutation* — «перестановка»). Если использовать это обозначение, то мы доказали первую комбинаторную формулу:

$$\boxed{P_N = N!}$$

Полученное значение P_N даёт нам **число способов, которым можно упорядочить N заданных элементов**. Приведём несколько примеров использования этой формулы.

Пример 1. Сколькими способами можно расставить 8 участников финального забега на 8 беговых дорожках?

Каждый такой способ — это перестановка из 8 элементов. Всего таких перестановок будет

$$P_8 = 8! = 40\,320.$$

Пример 2. Полное собрание сочинений братьев Стругацких включает в себя 33 тома. Сколькими способами это собрание можно расставить на полке?

Каждый способ — это перестановка из 33 элементов. Всего таких перестановок будет

$$P_{33} = 33! = 8\,683\,317\,618\,811\,886\,495\,518\,194\,401\,280\,000\,000,$$

т. е. чуть меньше, чем 10^{37}.

? ВОПРОСЫ

1. Что такое факториал? Как он обозначается?
2. Найдите $\frac{10!}{8!}$.
3. Чему равно число перестановок из N элементов? Как оно обозначается?
4. Почему для больших значений N в конце числа $N!$ так много нулей?

4 Размещения

В предыдущем пункте нам нужно было расставить на полке все 33 тома братьев Стругацких. А если на полку влезает только 10 из них — как тогда посчитать количество способов, которыми можно её заполнить?

Воспользуемся снова правилом умножения: первую книгу можно выбрать и поставить на полку 33 способами, вторую — 32 способами, третью — 31 способом и т. д. до последней десятой книги. К моменту её выбора останется 24 невыбранные книги. По правилу умножения общее число способов можно получить как произведение:

$$33 \cdot 32 \cdot 31 \cdot 30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 \cdot 25 \cdot 24 = 335\,885\,501\,952\,000.$$

Каждая из полученных в этой задаче комбинаций называется ***размещением***.

Размещением из N элементов по k называют комбинацию, в которой любые k из этих элементов расположены в определённом порядке.

Заметим, что все элементы размещения, как и элементы перестановки, должны быть различными.

Размещения отличаются друг от друга как составом комбинации, так и порядком расположения элементов. Да и сам термин «размещение» означает, что нужно не только выбрать какие-то k элементов из N имеющихся, но и разместить их в определённом порядке.

Элементами размещений могут быть любые объекты. Выпишем для примера все размещения из чисел 1, 2, 3, 4 по 3 элемента:

123, 124, 132, 134, 142, 143, 213, 214, 231, 234, 241, 243,
312, 314, 321, 324, 341, 342, 412, 413, 421, 423, 431, 432,

Выпишем все размещения из букв a, b, c по 2 элемента:

$$ab, ac, ba, bc, ca, cb.$$

В примере с книгами мы нашли количество размещений из 33 по 10. Используя тот же метод, можно вывести общую формулу для подсчёта размещений из N по k. Первый элемент такого размещения можно выбрать N способами; второй — $(N - 1)$ способом; третий — $(N - 2)$ способами и т. д. до последнего k-го элемента размещения. К моменту его выбора останется $(N - (k - 1))$ невыбранных элементов, поэтому этот последний элемент можно будет выбрать $(N - k + 1)$ способом. По правилу умножения общее количество размещений будет равно произведению

$$N \cdot (N - 1) \cdot (N - 2) \cdot \ldots \cdot (N - k + 1).$$

Заметим, что количество сомножителей в этом произведении равно k — по числу выбранных элементов.

Полученное произведение можно записать в свёрнутом виде, если использовать факториалы чисел. В самом деле, если поделить меньший факториал на больший, то часть множителей сократится и останется произведение нужного нам вида, например: $\frac{13!}{8!} = 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9$. Отсюда

$$N \cdot (N - 1) \cdot (N - 2) \cdot \ldots \cdot (N - k + 1) = \frac{N!}{(N - k)!}.$$

Количество размещений из N по k обозначается в комбинаторике A_N^k (буква A берётся от французского слова *arrangement* — «расположение», «размещение»). Читается это обозначение как «а из эн по ка».

Если использовать новое обозначение, то полученную формулу можно записать так:

$$\boxed{A_N^k = \frac{N!}{(N - k)!}}$$

Значение A_N^k даёт нам **число способов, которым можно выбрать k из N заданных элементов с учётом порядка, в котором мы их выбираем.** Рассмотрим несколько примеров использования этой формулы.

Пример 1. В теннисном турнире участвует 12 спортсменов. Сколькими способами могут распределиться три призовых места?

Каждое распределение призовых мест — это размещение из 12 элементов по 3. В самом деле, мы должны выбрать трёх призёров, а потом разместить их по трём местам. Всего таких размещений будет

$$A_{12}^3 = \frac{12!}{9!} = 12 \cdot 11 \cdot 10 = 1320.$$

Пример 2. В текущей четверти девятиклассники изучают 15 предметов. Сколькими способами можно составить расписание из 5 уроков, если все предметы должны быть разные?

Каждое допустимое в этой задаче расписание — это размещение из 15 по 5. Число таких размещений равно

$$A_{15}^5 = \frac{15!}{10!} = 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 = 360\ 360.$$

Пример 3. В классе, где учится 28 школьников, нужно выбрать старосту и двух его заместителей. Сколькими способами это можно сделать?

Может показаться, что в этой задаче речь снова идёт о размещениях: нужно выбрать 3 человек из 28 и «разместить» их по трём должностям:

$$A_{28}^3 = \frac{28!}{25!} = 28 \cdot 27 \cdot 26 = 19\ 656.$$

Но две должности заместителя, судя по условию задачи, равноценны, поэтому при обмене местами двух заместителей выбранный вариант не меняется. Значит, каждый вариант был посчитан дважды, поэтому нужно поделить полученное число на 2:

$$\frac{A_{28}^3}{2} = \frac{19\ 656}{2} = 9828.$$

Последний пример показывает, что знать комбинаторные формулы важно, но они не могут охватить всё разнообразие возможных комбинаций.

Отметим два интересных свойства размещений.

1. Размещения из N элементов по N — это перестановки из N элементов, поэтому

$$A_N^N = P_N = N!$$

2. Размещений из N элементов по $(N - 1)$ столько же, сколько размещений из N по N:

$$A_N^{N-1} = A_N^N = P_N = N!$$

Второе свойство легко получить из формулы для подсчёта размещений:

$$A_N^{N-1} = \frac{N!}{(N-(N-1))!} = \frac{N!}{1!} = N!$$

? ВОПРОСЫ

1. Что такое размещение? Чем оно отличается от перестановки?
2. Чему равно число размещений из N по k? Как оно обозначается? Как читается это обозначение?
3. Какие свойства числа размещений A_N^k вы знаете?

УПРАЖНЕНИЯ

199. Если выписать в порядке возрастания все трёхзначные числа, в записи которых используются только 0, 2, 4, 6, то какое число будет следующим за 426? предшествовать ему?

200. Если выписать по возрастанию все двоичные коды длины 8, то какой код будет следовать за кодом 10101011? предшествовать коду 10001000?

201. Если выписать по алфавиту все шестибуквенные слова, которые можно составить из букв А, О, У, Ы, Э, то какое слово будет следовать за словом ОУУЫЫЭ? предшествовать слову УУУААА?

202. Сколькими способами могут встать в очередь к кассе 6 человек? Как называются все такие комбинации?

203. В чемпионате города по футболу играет 12 команд. Сколькими способами могут распределиться три призовых места? Как называются все такие комбинации?

204. Вычислите:

а) $5!$; б) $\frac{10!}{8!}$; в) A_5^3; г) A_{10}^5; д) A_{10}^9.

205. Сколько трёхзначных чисел можно записать, используя только цифры 0, 2, 4, 6?

206. Из нечётных цифр составляют все возможные числа, содержащие не более четырёх цифр. Сколько существует таких чисел?

207. Монету подбрасывают 11 раз подряд. Сколько разных исходов у этого случайного опыта?

208. Одновременно подбрасывают 5 кубиков. Сколько разных исходов у этого случайного опыта?

209. Из класса, в котором учится 15 девочек и 10 мальчиков, случайно выбирают двух дежурных. Сколько разных исходов у этого случайного опыта?

210. В меню школьной столовой 3 разных супа, 4 вторых блюда и 2 вида сока. Сколькими способами можно выбрать обед из 3 блюд?

211. На деловую встречу пришло 7 человек. Каждый с каждым обменялся рукопожатием. Сколько всего рукопожатий было совершено?

212. В конференции участвовало 40 человек. Каждый с каждым обменялся визитной карточкой. Сколько на такой обмен ушло карточек?

213. После хоккейного матча каждый игрок одной команды пожал руку каждому игроку другой. Сколько всего игроков присутствовало на площадке, если было совершено 323 рукопожатия?

214. У рояля 88 клавиш. Сколькими способами можно последовательно извлечь на нём 6 нот? одновременно извлечь какие-то 3 ноты?

215. Сколько разных слов можно получить, переставляя буквы слова АБРАКАДАБРА? Сколько из них начинается на букву К? В скольких из них обе буквы Б стоят рядом?

216. Сколькими способами можно составить расписание для 10 класса, если в этот день нужно поставить указанные 5 уроков и выполнить следующие требования:

а) алгебра, русский язык, литература, география, вероятность и статистика;

б) две алгебры, русский язык, литература, география;

в) две алгебры, которые должны стоять рядом, русский язык, литература, география;

г) алгебра, русский язык, литература, география, физкультура; при этом русский язык и литература должны стоять рядом, а физкультура — быть последним уроком?

217*. В автомобиле пять мест. Сколькими способами пять человек могут занять в ней места для путешествия, если водить машину могут только трое из них?

218*. Номера паспортов в России состоят из 6 цифр. Сколько из них:

а) содержат хотя бы две одинаковые цифры;

б) содержат, по крайней мере, два нуля;

в) являются палиндромами?

219*. Одновременно бросают N кубиков. С какой вероятностью максимум из выпавших чисел будет равен 3? С какой вероятностью минимум из выпавших чисел будет равен 3?

220*. Сколько различных натуральных делителей имеет число:

а) 13; б) 63; в) 1000; г) 1024; д) 663 552; е) 10 125 000?

§ 12. Сочетания и их свойства

1 Сочетания

Если правило умножения использовать для подсчёта комбинаций, в которых порядок следования элементов не важен, то полученный результат будет ошибочным: ведь в таких комбинациях не нужно учитывать какой элемент был выбран первым, какой — вторым и т. д. Для правильного ответа в таких случаях нужно использовать ***правило деления***: если при подсчёте по правилу умножения каждая из комбинаций была посчитана m раз, то полученный результат нужно поделить на m.

Это напоминает известную шутку про ленивого пастуха, который привык пасти стадо лёжа на земле. Когда его спросили, как ему удаётся, не вставая с земли, пересчитывать баранов в стаде, он не задумываясь ответил: «я пересчитываю их ноги, а потом делю на 4».

Пример 1. Из класса, в котором учится 25 человек, нужно выбрать двоих для участия в олимпиаде по краеведению. Сколькими способами это можно сделать?

Первого ученика можно выбрать 25 способами, после чего второго — 24 способами. Всего таких способов по правилу умножения будет $25 \cdot 24 = 600$. Однако при выборе пары учеников для участия в олимпиаде их порядок в паре не имеет значения. Получается, что каждую пару мы посчитали дважды, поэтому правильным ответом будет $\frac{25 \cdot 24}{2} = 300$.

Пример 2. Из 11 футболистов, участвовавших в матче, нужно выбрать троих для проведения антидопинговой пробы. Сколькими способами это можно сделать?

Первого игрока можно выбрать 11 способами, после чего второго — 10 и третьего — 9 способами. Всего таких способов будет $11 \cdot 10 \cdot 9 = 990$. Однако при выборе тройки футболистов порядок их выбора не имеет значения. Посмотрим, сколько же раз мы посчитали каждую тройку.

Для этого нужно найти, сколькими способами можно упорядочить тройку футболистов, не меняя её состава. На первое место можно поставить любого из трёх выбранных футболистов, на второе — любого из двух, и для третьего места остаётся только один вариант. Значит, мы посчитали каждую тройку футболистов $3 \cdot 2 \cdot 1 = 6$ раз.

Поэтому правильным ответом в этой задаче будет $\frac{11 \cdot 10 \cdot 9}{6} = 165$.

В двух приведённых примерах мы столкнулись с комбинациями, в которых порядок элементов не учитывается. Такие комбинации называются ***сочетаниями***.

Сочетанием из N элементов по k называют комбинацию, составленную из любых k этих элементов без учёта их порядка.

Все элементы сочетания, разумеется, должны быть различными. Сочетания отличаются друг от друга только составом комбинации — порядок элементов не имеет значения.

С точки зрения теории множеств, сочетание есть не что иное, как **любое k-элементное подмножество N-элементного множества**. Например, сочетания из 8 по 3 — это все трёхэлементные подмножества множества из 8 элементов.

В русском языке «сочетать» означает «соединять, объединять». Формируя сочетание из k элементов, мы объединяем их в одно подмножество, а в нём, как известно, порядок элементов не учитывается.

Элементами сочетаний, как и элементами множеств, могут быть любые объекты. Выпишем все сочетания из чисел 1, 2, 3, 4 по 3 элемента:

123, 124, 134, 234.

Обратите внимание, их оказалось немного — всего 4. В то время как размещений из этих же четырёх чисел по 3 элемента получилось в предыдущем пункте целых 24. Объясняется это тем, что мы не выписывали комбинации, отличающиеся только порядком. Если в размещениях комбинация, которая состоит из цифр 1, 2, 3, была выписана 6 раз:

123, 132, 213, 231, 312, 321,

то теперь только один: 123. И так произошло с каждой.

Отметим ещё, что при перечислении сочетаний мы выписывали только те комбинации, где цифры идут по возрастанию, — это позволило избежать повторов.

Сочетаний из трёх букв a, b, c по 2 элемента будет всего 3:

$$ab,\ ac,\ bc.$$

При выборе 3 футболистов из 11 мы нашли количество сочетаний из 11 по 3. Для этого мы сначала вычислили число размещений из 11 по 3, а потом поделили его на 3!. Тем же методом можно вывести общую формулу для подсчёта числа сочетаний. Покажем, как это сделать.

При подсчёте размещений мы сначала выбираем k элементов из N (т. е. формируем сочетание из N по k), а потом размещаем их в каком-то порядке. Разместить k элементов в определённом порядке можно $k!$ способами (это число перестановок из k элементов). Значит, при подсчёте размещений каждое сочетание повторяется $k!$ раз. Отсюда сочетаний из N по k будет в $k!$ раз меньше, чем размещений из N по k, и их количество будет равняться

$$\frac{A_N^k}{k!} = \frac{N!}{k!(N-k)!}.$$

Количество сочетаний из N по k обозначается в комбинаторике C_N^k (буква C берётся от французского слова *combination* — «комбинация»). Читается это обозначение как «цэ из эн по ка».

С использованием нового обозначения полученная комбинаторная формула выглядит так:

$$\boxed{C_N^k = \frac{N!}{k!(N-k)!}.}$$

Значение C_N^k даёт нам **число способов, которым можно выбрать k из N заданных элементов без учёта порядка, в котором мы их выбираем**.

При $k = 2$ сочетания превращаются в неупорядоченные пары элементов. Количество таких пар вычисляется по формуле

$$C_N^2 = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}.$$

В электронных таблицах есть специальная функция для вычисления числа сочетаний C_N^k — это ЧИСЛКОМБ(N; k). Например,

ЧИСЛКОМБ(10;2)=45.

Рассмотрим примеры, в которых используются формулы для подсчёта числа сочетаний.

Пример 3. В текущем учебном году десятиклассникам предлагаются на выбор 7 дополнительных предметов, из которых они должны выбрать для посещения 4. Сколькими способами это можно сделать?

В этой ситуации нам нужно выбрать 4 предмета из 7, причём порядок выбора значения не имеет — важен только состав этой четвёрки. Значит, речь идёт о сочетаниях из 7 по 4. Находим их количество по формуле

$$C_7^4 = \frac{7!}{4!3!} = \frac{5 \cdot 6 \cdot 7}{3!} = 35.$$

Обратите внимание, что при вычислении дроби мы не стали сразу находить значения трёх разных факториалов — большинство множителей удалось сократить. Так будет и в других ситуациях, когда нужно вычислить C_N^k.

Пример 4. Из класса, в котором учится 13 мальчиков и 17 девочек, нужно набрать команду для участия в школьной олимпиаде по спортивному многоборью. В команде должно быть 2 мальчика и 2 девочки. Сколькими способами это можно сделать?

Посчитаем сначала, сколькими способами можно выбрать 2 мальчиков из 13:

$$C_{13}^2 = \frac{13 \cdot 12}{2} = 78.$$

Теперь сколькими способами можно выбрать 2 девочек из 17:

$$C_{17}^2 = \frac{17 \cdot 16}{2} = 136.$$

Чтобы получить общее число способов, которым можно собрать команду, нужно воспользоваться правилом умножения:

$$C_{13}^2 \cdot C_{17}^2 = 78 \cdot 136 = 10\ 608.$$

Пример 5. Для поездки на экскурсию 30 участников олимпиады нужно распределить по трём микроавтобусам, каждый из которых вмещает 10 человек. Сколькими способами это можно сделать?

Сначала выберем 10 человек для первого микроавтобуса — это $C_{30}^{10} = \frac{30!}{10!20!} = 30\ 045\ 015$ способов. Осталось 20 человек, из которых снова нужно выбрать 10, — это $C_{20}^{10} = \frac{20!}{10!10!} = 184\ 756$ способов. Оставшиеся 10 человек едут на третьем микроавтобусе — здесь уже ничего выбирать не нужно, или можно сказать, что здесь всего один вариант: $C_{10}^{10} = \frac{10!}{10!0!} = 1$.

Чтобы получить ответ, используем правило умножения:

$$C_{30}^{10} \cdot C_{20}^{10} \cdot C_{10}^{10} = 30\ 045\ 015 \cdot 184\ 756 \cdot 1 = 5\ 550\ 996\ 791\ 340.$$

? ВОПРОСЫ

1. В каких случаях в комбинаторике используется правило деления?
2. Что такое сочетание? Чем оно отличается от размещения?
3. Чему равно число сочетаний из N по k? Как оно обозначается? Как читается это обозначение?
4. Найдите C_{10}^2, C_{100}^1, C_{100}^{100}.

2 Свойства чисел сочетаний

Числа сочетаний C_N^k используются в самых разных областях математики. Остановимся на их свойствах более подробно. Начнём с «крайних» случаев, когда сочетание содержит очень мало или, наоборот, очень много элементов.

Свойство 1. $C_N^0 = C_N^N = 1.$

Это свойство следует из общей формулы: $C_N^0 = C_N^N = \dfrac{N!}{0!N!} = 1.$

Но можно получить его и комбинаторно: выбрать 0 элементов из N можно только одним способом — ничего не взять. Точно так же выбрать N элементов из N можно только одним способом — выбрать все элементы.

Свойство 2. $C_N^1 = C_N^{N-1} = N.$

Здесь тоже есть два доказательства. Сначала подставим $k = 1$ и $k = N - 1$ в общую формулу для числа сочетаний: $C_N^1 = C_N^{N-1} = \dfrac{N!}{1!(N-1)!} = N.$

Теперь рассмотрим комбинаторное доказательство: выбрать 1 элемент из N можно N способами.

Чтобы выбрать $(N - 1)$ элементов из N, нужно указать один элемент, который не выбран, что можно сделать тоже N способами.

Последнее рассуждение можно применить и для доказательства свойства симметричности.

Свойство 3. $C_N^k = C_N^{N-k}$ для любого $0 \leqslant k \leqslant N$.

Выбрать $(N - k)$ элементов из N можно таким же количеством способов, каким указать k элементов, которые мы не выберем, т. е. $C_N^{N-k} = C_N^k$.

Конечно, это свойство можно доказать и непосредственно из общей формулы: ведь она симметрична относительно значений k и $(N - k)$.

Получить следующее равенство проще всего именно с помощью комбинаторных рассуждений.

Свойство 4. $C_N^0 + C_N^1 + C_N^2 + ... + C_N^{N-1} + C_N^N = 2^N.$

Сначала докажем, что **общее количество подмножеств в любом N-элементном множестве равно 2^N**. Любой из N элементов мы можем либо включить в подмножество, либо нет. Таким образом, при формировании любого подмножества у нас есть два возможных варианта для первого элемента, два варианта для второго и т. д. до последнего N-го элемента. Всего таких вариантов будет

$$\underbrace{2 \cdot 2 \cdot ... \cdot 2}_{N \text{ раз}} = 2^N.$$

Теперь посчитаем отдельно 0-элементые подмножества, 1-элементые подмножества, 2-элементые подмножества и т. д. Поскольку k-элементные подмножества есть не что иное, как сочетания из N по k, то их количество равно числу C_N^k. Поэтому сумма всех чисел C_N^k для k от 0 до N также даёт общее количество всех подмножеств, т. е. 2^N. Равенство доказано.

Вообще в комбинаторике часто используется такой приём доказательства тождеств: **мы считаем число каких-то комбинаций сначала одним способом, а потом другим — и получаем равенство найденных чисел**.

Следующее свойство чисел C_N^k даёт нам ***рекуррентный*** способ их вычисления, т. е. получение новых чисел с бóльшими значениями N через предыдущие числа с меньшими значениями N.

■ **Свойство 5.** $C_N^k = C_{N-1}^{k-1} + C_{N-1}^k$.

Это тождество можно доказать через общую формулу для чисел C_N^k, но мы оставим этот способ для упражнения 234*, а здесь приведём комбинаторное доказательство.

В левой части равенства стоит число k-элементных подмножеств некоторого N-элементного множества A. Попробуем посчитать все эти подмножества так. Выберем какой-то элемент $x \in A$ и поделим все подмножества множества A на два типа: те, что содержат x, и те, что не содержат x.

Чтобы получить любое k-элементное подмножество первого типа, нужно выбрать и добавить к элементу x ещё $(k - 1)$ элемент из $(N - 1)$ оставшихся элементов. Это можно сделать C_{N-1}^{k-1} способами.

Чтобы получить любое k-элементное подмножество второго типа, нужно выбрать k элементов из $(N - 1)$ элемента, поскольку элемент x выбирать нельзя. Это можно сделать C_{N-1}^k способами.

Вместе тех и других подмножеств будет $C_{N-1}^{k-1} + C_{N-1}^k$. Тождество доказано.

? ВОПРОСЫ

1. Чему равны числа C_{100}^0, C_{99}^{99}?
2. Чему равны числа C_{100}^1, C_{123}^{122}?
3. Чему равна сумма чисел $C_{10}^1 + C_{10}^2 + C_{10}^3 + \ldots + C_{10}^{10}$?

3 Треугольник Паскаля

Свойство 5, доказанное в предыдущем пункте, позволяет находить числа C_N^k вообще без использования факториалов. Для этого нужно последовательно применять полученную рекуррентную формулу при $N = 1, 2, 3$ и т. д. Если, например, мы уже нашли все значения C_4^0, C_4^1, C_4^2, C_4^3, C_4^4, то можем найти любое значение C_5^k по формуле $C_5^k = C_4^{k-1} + C_4^k$.

Покажем, как это делается, оформив вычисления в таблицу 51. Сначала заполним все «крайние» значения чисел C_N^k для $k = 0$ и $k = N$.

Таблица 51

N \ k	0	1	2	3	4	5	6
0	$C_0^0 = 1$						
1	$C_1^0 = 1$	$C_1^1 = 1$					
2	$C_2^0 = 1$		$C_2^2 = 1$				
3	$C_3^0 = 1$			$C_3^3 = 1$			
4	$C_4^0 = 1$				$C_4^4 = 1$		
5	$C_5^0 = 1$					$C_5^5 = 1$	
6	$C_6^0 = 1$						$C_6^6 = 1$
...	...	...	...	...	...	...	...

Теперь каждое число C_N^k в пустой клетке получается как сумма чисел C_{N-1}^{k-1} и C_{N-1}^k, стоящих в предыдущей строке (табл. 52).

Таблица 52

N \ k	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1
...	...	...	...	...	...	...	...

Полученная таблица, составленная из чисел C_N^k, носит название ***треугольника Паскаля*** по имени замечательного французского учёного Блеза Паскаля (1623—1662). Первое упоминание этой конструкции из чисел встречается гораздо раньше в трудах древнеиндийских математиков, но Паскаль посвятил ей «Трактат об арифметическом треугольнике», в котором обнаружил ряд других замечательных закономерностей, помимо рекуррентного соотношения между строками треугольника.

Чаще треугольник рисуют в симметричной форме в виде бесконечной таблицы, в вершине и по бокам которой стоят 1, а каждое другое число равно сумме двух чисел, расположенных над ним (рис. 87).

Симметричность полученного треугольника — это уже знакомое нам свойство 3 для чисел C_N^k: $C_N^k = C_N^{N-k}$.

1

1 1

1 2 1

1 3 3 1

1 3 6 3 1

1 5 10 10 5 1

1 6 15 20 15 6 1

1 7 21 35 35 21 7 1

1 8 28 56 70 56 28 8 1

Рис. 87. Треугольник Паскаля

Если сложить числа в любой строке, то по свойству 4 получится соответствующая степень двойки, например:

$$1 + 7 + 21 + 35 + 35 + 21 + 7 + 1 = 128 = 2^7.$$

У треугольника Паскаля есть и другие замечательные свойства, которые вы узнаете при выполнении упражнений.

? ВОПРОСЫ

1. Запишите 5 первых строк треугольника Паскаля.
2. Чему равна сумма чисел в каждой строке треугольника Паскаля?
3. Сколько чисел в 13 строке треугольника Паскаля?

4 Бином Ньютона

«Подумаешь, бином Ньютона!» — эта фраза одного из героев романа «Мастер и Маргарита» сделала понятие бинома синонимом чего-то очень сложного и непонятного. В рассказе «Последнее дело Холмса» о профессоре Мориарти и вовсе сказано, что «он написал трактат о биноме Ньютона, завоевавший ему европейскую известность».

Если в цитате из романа Булгакова сквозит ирония, то у Конан Дойла здесь явное преувеличение: вряд ли в XIX в. можно было завоевать европейскую известность таким трактатом. К этому времени формула бинома Ньютона давно входила в программы школ и колледжей.

В алгебре ***биномом Ньютона*** называется выражение вида $(a + b)^N$, а ***формулой бинома Ньютона*** — тождество, позволяющее раскрыть скобки в этом выражении и представить его как сумму одночленов. Вам хорошо известны частные случаи этой формулы для $N = 2$ и $N = 3$. Это формулы квадрата суммы:

$$(a + b)^2 = a^2 + 2ab + b^2$$

и куба суммы:

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Обратите внимание, что если выписать коэффициенты перед слагаемыми в правых частях этих тождеств, то получатся вторая и третья строки треугольника Паскаля:

1 2 1

1 3 3 1.

Проверим, сохранится ли эта закономерность для четвёртой степени:

$$(a + b)^4 = (a^2 + 2ab + b^2)(a^2 + 2ab + b^2) =$$
$$= a^4 + 2a^3b + a^2b^2 + 2a^3b + 4a^2b^2 + 2ab^3 + b^2a^2 + 2ab^3 + b^4 =$$
$$= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$$

Всё подтвердилось: 1 4 6 4 1. Эту закономерность обнаружили многие средневековые математики в Китае, Индии и на Ближнем Востоке. В 1665 г. Исаак Ньютон (1643—1727) обобщил эту формулу для дробных и отрицательных показателей степени. Поэтому с тех пор она и носит его имя.

Итак, сформулируем полученный результат в виде следующего тождества.

Для любого натурального значения *n* справедливо равенство, которое называется формулой бинома Ньютона:

$$(a + b)^N = C_N^0 a^0 b^N + C_N^1 a^1 b^{N-1} + C_N^2 a^2 b^{-2} + \ldots + C_N^{N-1} a^{N-1} b^1 + C_N^N a^N b^0.$$

Таким образом, коэффициентами для разложения бинома Ньютона степени N являются числа C_N^k, которые можно найти из треугольника Паскаля.

Из формулы бинома Ньютона можно вывести ещё несколько интересных свойств чисел C_N^k. Отметим среди них следующее.

Свойство 6. $C_N^0 - C_N^1 + C_N^2 - \ldots + (-1)^N C_N^N = 0.$

То есть знакопеременная сумма чисел C_N^k в любой строке треугольника Паскаля равна 0. Для нечётных значений N это следует из симметрии треугольника. Например, при $N = 7$ получаем:

$$1 - 7 + 21 - 35 + 35 - 21 + 7 - 1 = 0.$$

Для чётных N симметрия здесь не помогает, но свойство продолжает выполняться:

$$1 - 8 + 28 - 56 + 70 - 56 + 28 - 8 + 1 = 0.$$

Для доказательства тождества используем формулу бинома Ньютона. Положим в ней $a = -1$, $b = 1$. Тогда

$$0 = (-1 + 1)^N = C_N^0(-1)^0 + C_N^1(-1)^1 + C_N^2(-1)^2 + \ldots + C_N^{N-1}(-1)^{N-1} + C_N^N(-1)^N.$$

Свойство доказано.

ВОПРОСЫ

1. Что такое бином Ньютона?
2. Запишите формулу бинома Ньютона для $N = 3$; для $N = 4$.
3. Чему равна знакочередующаяся сумма чисел в любой строке треугольника Паскаля?

УПРАЖНЕНИЯ

221. Из двенадцати фильмов, номинированных на лучшую режиссёрскую работу, жюри кинофестиваля должно отобрать трёх финалистов. Сколькими способами оно может это сделать? Как называются все такие комбинации?

222. Выпишите все сочетания из цифр 1, 2, 3, 4 по три цифры. Сколько их получилось?

223. Вычислите:
а) C_{10}^2; б) C_5^3; в) C_{11}^9; г) C_{10}^5; д) C_{100}^{99}.

224. Из класса, в котором учится 15 девочек и 10 мальчиков, нужно выбрать делегацию на слёт из 3 девочек и 3 мальчиков. Сколькими способами это можно сделать?

225. Сколько диагоналей:
а) у четырёхугольника;
б) у пятиугольника;
в) у десятиугольника;
г) у N-угольника?

226. Сколько существует двоичных кодов (т. е. последовательностей из нулей и единиц) длины 10, в которых:
а) нет нулей;
б) ровно одна единица;
в) ровно 3 единицы;
г) хотя бы одна единица;
д) нулей и единиц поровну;
е) нулей больше, чем единиц?

227*. Пусть $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Посчитайте количество всех подмножеств множества A, которые:

а) состоят из трёх элементов;

б) состоят из пяти элементов и содержат 1;

в) состоят из шести элементов и не содержат 2;

д) состоят из трёх чётных и двух нечётных чисел;

е) содержат меньше пяти элементов.

228*. Отряд из 20 туристов нужно разделить на 3 поисковых группы численностью 3, 5 и 12 человек. Сколькими способами это можно сделать?

229*. Из отделения в 20 солдат каждую ночь выделяется наряд, состоящий из 3 человек. Сколько ночей подряд командир может выделять наряд, не совпадающий ни с одним из предыдущих? Сколько раз при этом сходит в наряд каждый солдат?

230*. Сколькими способами можно рассадить за 10 партами 10 мальчиков и 10 девочек так, чтобы за каждой партой сидели мальчик и девочка?

231*. Сколькими способами колоду из 36 карт можно разделить пополам так, чтобы:

а) красных и чёрных карт в каждой половине оказалось поровну;

б) карт каждой масти в каждой половине оказалось поровну?

232*. Сколько существует двоичных кодов длины n, в которых число единиц:

а) чётно; б) нечётно?

233*. В правлении банка 7 человек. Каково должно быть минимальное число замков от сейфа и как следует распределить ключи между членами правления, чтобы любое большинство правления могло открыть сейф, а любое меньшинство не могло?

Замечание: каждый член правления может получить ключи от нескольких замков.

234*. Докажите тождества:

а) $C_n^0 + C_{n+1}^1 + \ldots + C_{n+k}^k = C_{n+k+1}^k$;

б) $C_n^n + C_{n+1}^n + \ldots + C_{n+k}^n = C_{n+k+1}^{n+1}$.

Найдите, где расположены все эти числа в треугольнике Паскаля.

235. Раскройте скобки в выражении:

а) $(x + 1)^7$; б) $(x + 2)^6$; в) $(3x + 2y)^4$; г) $(x - 4y)^5$.

ГЛАВА 6

ИСПЫТАНИЯ БЕРНУЛЛИ. СЛУЧАЙНЫЙ ВЫБОР

В этой главе мы будем изучать ***испытания Бернулли***. Это одна из наиболее важных вероятностных моделей, которая используется для решения многих практических задач. Впервые её рассмотрел в своих работах швейцарский математик Якоб Бернулли (1655—1705), один из представителей многочисленного математического семейства Бернулли.

Мы уже не раз сталкивались с испытаниями Бернулли, сами того не подозревая. Подбрасывание монеты или кубика, стрельба в мишень, проверка деталей на соответствие стандарту и многие другие серии случайных опытов — всё это примеры таких испытаний.

Разберёмся, что объединяет все эти примеры и почему для их изучения можно рассматривать одну и ту же вероятностную модель.

§ 13. Независимые испытания

1 Успех и неудача

В теории вероятностей *испытаниями Бернулли* называют серию одинаковых независимых случайных опытов, каждый из которых может завершиться одним из двух исходов — *успехом* или *неудачей*.

Названия «успех» и «неудача» носят условный характер и не несут никакой эмоциональной окраски. Что, например, считать успехом при подбрасывании монеты — «орла» или «решку» — вы можете решить сами. Если речь идёт о проверке детали на соответствие стандарту качества, может оказаться удобным считать «успехом» обнаружение брака.

Вероятности успеха и неудачи обозначают буквами p и q. Поскольку в каждом испытании возможны только эти два исхода, то $p + q = 1$. Важным требованием к испытаниям Бернулли является то, что вероятности p и q **остаются постоянными** и не изменяются в процессе проведения испытаний.

Второе важное требование — **независимость** испытаний. Это означает, что результаты предыдущих испытаний никак не влияют на вероятности появления успеха и неудачи в последующих испытаниях. Понятно, что с монетой дело обстоит именно так. Но если речь идёт о стрельбе в мишень на каких-то важных соревнованиях, то в независимость может вмешаться психологический фактор: не исключено, что, совершив промах, спортсмен занервничает, и вероятность успеха в следующем выстреле может уменьшиться. Может быть и наоборот — он соберётся, и тогда вероятность успеха увеличится. Здесь, как всегда при построении математической модели, мы сами должны решить: существенна эта зависимость или ею можно пренебречь.

В начале этой главы мы привели в качестве примера испытаний Бернулли подбрасывание кубика. У вас, наверное, уже появился вопрос: как же так, ведь у этого опыта не 2, а 6 возможных исходов. Это так, но если нас будет интересовать конкретное случайное событие, например A = «на кубике выпадет шесть очков», мы вполне можем считать его успехом, а противоположное событие $\overline{A}$ — неудачей. И тогда серия испытаний с кубиком тоже превратится в испытания Бернулли.

Таким образом, если в случайном опыте нас интересует какое-то случайное событие A, мы можем назвать наступление этого события успехом, а наступление противоположного события $\overline{A}$ — неудачей. После этого серию таких опытов (при условии их независимости) можно будет считать испытаниями Бернулли и применять к ним те результаты, которые мы получим в этой главе.

Рассмотрим примеры испытаний Бернулли.

Пример 1. Несколько раз подбрасывают монету.

Это пример симметричных испытаний Бернулли, в которых $p = q = \frac{1}{2}$.

Пример 2. Несколько раз подбрасывают канцелярскую кнопку.

Будем считать успехом выпадение кнопки остриём вверх, а неудачей — остриём вниз (рис. 88). В этом опыте $p \neq q$, а определить их приближённые значения можно только экспериментально, используя в качестве оценок частоту успехов и неудач.

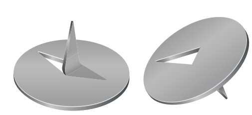


Рис. 88. Опыт с подбрасыванием кнопки

Пример 3. Несколько раз подбрасывают игральный кубик.

Чтобы превратить этот опыт в испытание Бернулли, будем считать успехом выпадение шестёрки. Тогда $p = \frac{1}{6}$, $q = \frac{5}{6}$.

В этих трёх классических примерах все требования к испытаниям Бернулли очевидно выполнены. В более реальных ситуациях, как мы уже говорили, могут изменяться вероятности p и q или нарушаться независимость. Если эти отклонения небольшие, то ими можно пренебречь.

Пример 4. Плохо подготовленный ученик отвечает на вопросы теста. К каждому вопросу предлагаются 4 варианта ответа, из которых только один вариант верный.

Будем считать успехом, если ученик выбрал верный ответ, и неудачей — неверный. С некоторой натяжкой это можно считать испытаниями Бернулли, в которых $p = \frac{1}{4}$, $q = \frac{3}{4}$.

В этом примере чем хуже подготовлен ученик, тем ближе описанная ситуация к испытаниям Бернулли. Реальный ученик какие-то темы наверняка знает лучше, поэтому вероятность p для них может повышаться. Да и вопросы в тесте бывают разной сложности. Но для первого приближения схема Бернулли здесь вполне пригодна. И мы ещё воспользуемся ею в дальнейшем при обсуждении ситуации с тестированием.

Вот в следующем примере зависимость испытаний настолько очевидна, что пренебрегать ею уже нельзя.

Пример 5. В корзинке лежат 5 пирожков с капустой и 5 с повидлом. Бабушка любит пирожки с капустой и считает выбор такого пирожка большим успехом. Она наугад достаёт из корзинки и съедает 4 пирожка. Будут ли её действия испытаниями Бернулли?

В первом испытании вероятность успеха равна 0,5. Но уже во втором испытании она будет зависеть от того, чем закончилось первое испытание. Если первым был вынут пирожок с капустой, то условная вероятность успеха во втором испытании снизится до $\frac{4}{9} \approx 0{,}44$, а если с повидлом, то повысится до $\frac{5}{9} \approx 0{,}56$. Таким образом, испытания будут **зависимыми**: каждый новый успех будет уменьшать вероятность успеха в следующем испытании. Поэтому считать случайный выбор ***без возвращения*** (ведь выбранные пирожки съедаются) испытаниями Бернулли нельзя. Мы ещё вернёмся к таким случайным экспериментам после изучения испытаний Бернулли.

? ВОПРОСЫ

1. Перечислите требования, которым должны удовлетворять испытания Бернулли.
2. Приведите примеры испытаний Бернулли.
3. Почему случайный выбор без возвращения нельзя считать испытаниями Бернулли?

2 Формула Бернулли

Первая важная задача, которую мы решим для испытаний Бернулли, заключается в следующем: **пусть проводится N испытаний Бернулли, с какой вероятностью в такой серии произойдёт ровно k успехов?**

Договоримся обозначать эту вероятность $P_N(k)$. Понятно, что при этом $0 \leqslant k \leqslant N$. Прежде чем вывести общую формулу, рассмотрим несколько конкретных примеров этой задачи.

Пример 1. Монету бросают 8 раз. С какой вероятностью при этом выпадет ровно 4 орла?

Здесь $p = \frac{1}{2}$ и нужно найти $P_8(4)$.

Пример 2. Кубик бросают 10 раз. С какой вероятностью в этих опытах выпадет ровно 2 шестёрки?

Успех — это выпадение шестёрки, поэтому $p = \frac{1}{6}$ и нужно найти $P_{10}(2)$.

Пример 3. Ученик отвечает на 20 вопросов теста, выбирая каждый раз наугад один из четырёх вариантов ответа. С какой вероятностью он правильно ответит на 12 вопросов?

Поскольку ученик выбирает наугад один из четырёх вариантов ответа, то $p = \frac{1}{4}$ и нужно найти $P_{20}(12)$.

Теперь перейдём к выводу общей формулы. Рассмотрим всю серию из N испытаний Бернулли как один случайный опыт. Выясним, какие у него возможные исходы и сколько их.

Каждый исход такого «длинного» опыта — это последовательность из N успехов и неудач, следующих друг за другом в произвольном порядке. Если обозначить успех цифрой 1, а неудачу — цифрой 0, то получим последовательность из N произвольно чередующихся 0 и 1. Если перейти на язык информатики, то это будут ***двоичные коды*** длины N.

Вот так, например, будут выглядеть все возможные исходы для серии из трёх испытаний:

000, 001, 010, 011, 100, 101, 110, 111.

Как видим, таких исходов при $N = 3$ получилось 8. В общем случае их количество легко посчитать по правилу умножения: для выбора первой цифры 2 варианта, для второй — тоже 2 и т. д. до последней N-й цифры. Значит, общее количество исходов будет

$$\underbrace{2 \cdot 2 \cdot \ldots \cdot 2}_{N \text{ раз}} = 2^N.$$

Итак, количество исходов мы нашли, но будут ли все они равновозможны? В общем случае нет. Однако вероятность каждого исхода можно легко вычислить, пользуясь ***формулой произведения вероятностей*** для независимых событий — ведь испытания Бернулли независимые.

Покажем, как это сделать для случая $N = 3$ (табл. 53).

Таблица 53

Исходы	Вероятности
000	$qqq = q^3$
001	$qqp = pq^2$
010	$qpq = pq^2$
011	$qpp = p^2q$
100	$pqq = pq^2$
101	$pqp = p^2q$
110	$ppq = p^2q$
111	$ppp = p^3$

Мы видим, что вероятность каждой серии зависит только от того, **сколько в ней успехов и неудач, а не от того, как они между собой чередуются**. В общем случае, если серия содержит k успехов и $(N - k)$ неудач, то её вероятность будет равна $p^k q^{N-k}$.

Чтобы получить вероятность $P_N(k)$, остаётся сделать последний шаг: посчитать, сколько всего серий, которые содержат ровно k успехов. Другими словами, нам нужно найти количество двоичных кодов длины N, которые содержат ровно k единиц.

Для формирования такого кода нужно выбрать k из N мест, на которых будут стоять единицы, а на оставшиеся места поставить нули. Выбрать k из N мест можно C_N^k способами. Значит, всего серий, в которых ровно k единиц, будет C_N^k.

Поскольку вероятность каждой такой серии одна и та же и равна $p^k q^{N-k}$, то получаем следующее выражение для вероятности получить k успехов в N испытаниях:

$$\boxed{P_N(k) = C_N^k \cdot p^k q^{N-k}}$$

Полученная формула носит название ***формулы Бернулли*** и была доказана швейцарским математиком Якобом Бернулли более 300 лет тому назад.

Используем эту формулу, чтобы ответить на вопросы из примеров 1—3, рассмотренных в начале этого пункта.

Пример 1. Монету бросают 8 раз. С какой вероятностью выпадет ровно 4 орла?

$$P_8(4) = C_8^4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^4 = \frac{8!}{4!4!}\cdot\frac{1}{256} \approx 0{,}273.$$

Пример 2. Кубик бросают 10 раз. С какой вероятностью выпадет ровно 2 шестёрки?

$$P_{10}(2) = C_{10}^2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^8 = \frac{10!}{2!8!}\cdot\frac{5^8}{6^{10}} \approx 0{,}291.$$

Пример 3. Ученик отвечает на 20 вопросов теста, выбирая каждый раз наугад один из четырёх вариантов ответа. С какой вероятностью он правильно ответит ровно на 12 вопросов?

$$P_{20}(12) = C_{20}^{12}\left(\frac{1}{4}\right)^{12}\left(\frac{3}{4}\right)^8 = \frac{20!}{12!8!}\cdot\frac{3^8}{4^{20}} \approx 0{,}000752.$$

Как видите, решение довольно сложных задач превращается благодаря формуле Бернулли в обычное арифметическое упражнение. Но будьте внимательны: в этой формуле много параметров, в которых легко запутаться и получить неверный ответ.

Для вычисления вероятностей $P_N(k)$ удобно использовать электронную таблицу, в которой есть функция БИНОМРАСП(). С её помощью можно найти как одну вероятность $P_N(k)$, так и сумму таких вероятностей от 0 до заданного k. Какой именно из этих двух вариантов выбрать, указывает последний, четвёртый параметр этой функции:

$P_N(k)$ = БИНОМРАСП(k; N; p; 0);

$P_N(0) + P_N(1) + \ldots + P_N(k)$ = БИНОМРАСП(k; N; p; 1).

? ВОПРОСЫ

1. Сколько всего двоичных кодов длины N?
2. Сколько двоичных кодов длины N, в которых ровно k единиц?
3. Если в серии из 10 испытаний Бернулли 4 успеха и 6 неудач, чему равна её вероятность?
4. Запишите формулу Бернулли.

3 Наиболее вероятное число успехов*

Вы, наверное, обратили внимание, что в рассмотренных только что примерах получились довольно разные по величине ответы: в примерах 1 и 2 вероятности оказались достаточно большими (0,273 и 0,291), а в примере 3 — всего 0,000752.

Конечно, такую маленькую вероятность можно списать на то, что число успехов в третьем примере меняется в большем диапазоне, чем в первых двух: оно может принимать значения от 0 до 20, между которыми нужно «распределить» всю вероятность, суммарно равную 1.

Но в примере с монетой и кубиком этот диапазон меньше примерно в два раза, а вероятность при этом больше в сотни раз.

Дело в том, что в задаче с монетой мы вычисляли вероятность $P_8(4)$, т. е. вероятность, что в 8 испытаниях выпадет 4 орла. Поскольку монета симметрична, то это, по-видимому, и есть **самое вероятное число успехов**.

Во всяком случае, оно должно быть больше, чем $P_0(8)$ и $P_8(8)$: вряд ли можно ожидать, что все 8 раз выпадет решка или все 8 раз выпадет орёл.

Это и в самом деле так: даже без формулы Бернулли легко получить, что $P_8(0) = P_8(8) = \frac{1}{2^8} = \frac{1}{256} \approx 0{,}00391.$

При подбрасывании кубика вероятность выпадения шестёрки равна $\frac{1}{6}$. Значит, в N таких испытаниях шестёрка должна выпадать примерно в $\frac{N}{6}$ случаях. Поскольку в примере 2 из предыдущего пункта кубик бросали 10 раз, то $\frac{10}{6} \approx 1{,}7$ — т. е. 2 успеха из 10 в этой серии весьма вероятны. Поэтому и вероятность получилась немаленькой — 0,291.

В примере с угадыванием ответов, если использовать те же рассуждения, следует ожидать около $\frac{20}{4} = 5$ правильных ответов. А мы вычисляли вероятность для 12 правильных ответов, поэтому и получили такое маленькое число. Кстати, вероятность для 5 успехов будет в этом примере гораздо больше:

$$P_{20}(5) = C_{20}^{5}\frac{20!}{5!5!} \cdot \left(\frac{1}{4}\right)^{5}\left(\frac{3}{4}\right)^{15} \approx 0{,}202.$$

На этих примерах мы видим, что наиболее вероятное число успехов в серии из N испытаний Бернулли лежит где-то около числа $N \cdot p$, где p — вероятность успеха. По мере удаления числа успехов от $N \cdot p$ оно становится всё менее вероятным.

С помощью компьютера можно вычислить значения $P_N(k)$ при $k = 0$, 1, 2, ... , N и построить по ним график (этому будет посвящена специальная лабораторная работа). На рисунке 89 такой график построен для серии испытаний, в которых 10 раз бросают кубик. Хорошо видно, что максимального значения вероятность достигает в точке $k = 1$, как мы примерно и ожидали.

Можно доказать следующую теорему, которая даёт точное значение наиболее вероятного числа успехов.

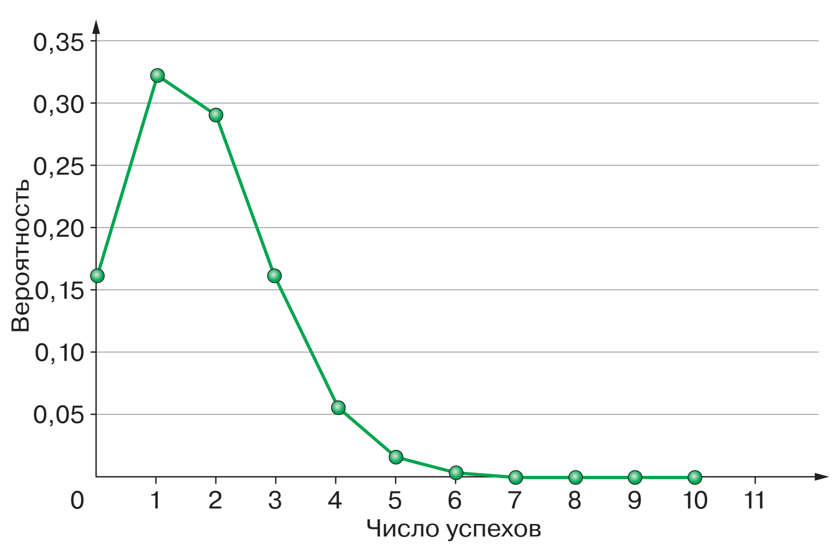


Рис. 89. Вероятность получить k успехов в 10 испытаниях с кубиком

■ **Теорема.** **(О наиболее вероятном числе успехов.)** В серии из N испытаний Бернулли с вероятностью успеха p наиболее вероятное число успехов x удовлетворяет неравенству

$$Np - q \leqslant x \leqslant Np - q + 1.$$

Заметим, что поскольку левая и правая границы этого неравенства отличаются на 1, то на этот отрезок попадают либо два целых числа (когда эти границы целые), либо одно целое число (когда границы дробные).

Таким образом, если число $Np - q$ не целое, то существует только одно наиболее вероятное число успехов, а если целое, то таких чисел два.

Вы докажете теорему в упражнении 257*. Сейчас воспользуемся ею, чтобы найти наиболее вероятное число успехов в наших трёх примерах.

Пример 1. Монету бросают 8 раз.

$Np - q = 8 \cdot \frac{1}{2} - \frac{1}{2} = 3{,}5 \Rightarrow$ наиболее вероятное число орлов равно 4.

Пример 2. Кубик бросают 10 раз.

$Np - q = 10 \cdot \frac{1}{6} - \frac{5}{6} = \frac{5}{6} \Rightarrow$ наиболее вероятное число шестёрок равно 1.

Пример 3. Ученик отвечает на 20 вопросов теста, выбирая каждый раз наугад один из четырёх вариантов ответа.

$Np - q = 20 \cdot \frac{1}{4} - \frac{3}{4} = 4{,}25 \Rightarrow$ наиболее вероятное число угаданных ответов равно 5.

? ВОПРОСЫ

1. Монету бросают 5 раз. Какое число орлов наиболее вероятно?
2. Монету бросают 6 раз. Какое число орлов наиболее вероятно?
3. Кубик бросают 6 раз. Какое число двоек наиболее вероятно?

УПРАЖНЕНИЯ

236. Стрелок в тире пять раз стреляет по мишеням. Вероятность попадания в мишень при одном выстреле равна 0,6. С какой вероятностью он:
а) все пять раз попадёт в мишень;
б) первые четыре раза попадёт, а последний — промахнётся;
в) первые три раза промахнётся, а последние два — попадёт?

237. У вас в кармане 7 орехов. Вероятность, что орех будет пустым, равна 0,4. Какова вероятность того, что у вас по крайней мере 3 полных ореха?

238. На склад поступают детали, из которых около 5% бракованных. Найдите вероятность того, что среди 10 наугад взятых деталей:
а) не будет бракованных;
б) будет 2 бракованных;
в) будет не больше 2 бракованных.

239. Владелец интернет-сайта вывесил на главной странице рекламный баннер. Ежедневно сайт посещает около 1000 человек. Вероятность, что посетитель кликнет по баннеру, равна 0,005. С какой вероятностью за день:
а) не будет сделано ни одного клика;
б) будет сделано больше одного клика?

240. В подъезде горят 5 лампочек. Вероятность, что любая лампочка не перегорит в течение ближайшего месяца, равна 0,2. Какова вероятность того, что в течение месяца:
а) сгорят все лампочки;
б) сгорит ровно одна лампочка;
в) останутся гореть по крайней мере 3 лампочки?

241. Два игральных кубика бросают 5 раз. Найдите вероятности событий:
а) сумма 7 выпадет по крайней мере дважды;
б) сумма 12 выпадет по крайней мере дважды;
в) сумма 6 выпадет 3 раза.

242. С какой вероятностью орлов выпадет больше, чем решек, если монету бросают:
а) 4 раза;
б) 5 раз;
в) 20 раз?

243. В гараже имеется 12 автомобилей. Вероятность выхода на линию каждого из них равна 0,9. Найдите вероятность нормальной работы гаража в ближайший день, если для этого необходимо иметь на линии не менее 10 автомобилей.

244*. В продукции завода по производству свёрл брак составляет около 4%. Каждое сверло проходит проверку, при которой брак, если он есть, выявляется с вероятностью 0,8. Свёрла пакуются в коробки по 20 штук. Какой процент коробок, поступающих в продажу, содержит бракованные свёрла?

245. Тест состоит из 10 вопросов. К каждому вопросу предлагается 4 варианта ответа, из которых ровно один верный. С какой вероятностью ученик, не знающий предмета, правильно ответит хотя бы на один вопрос?

246*. В предыдущей задаче к каждому вопросу предлагается 4 варианта ответа, из которых **любое количество** может быть верным (в том числе ни одного). Ответ считается верным, если указаны все правильные варианты и нет ни одного неправильного. С какой вероятностью ученик, не знающий предмета, правильно ответит хотя бы на один вопрос?

247*. Вероятность, что саженец дерева приживётся, равна 0,8. Сколько саженцев нужно посадить, чтобы наивероятнейшее число прижившихся деревьев было равно 20?

248*. Проводится серия из $2n$ испытаний Бернулли с вероятностью успеха p. Найдите границы, в которых должно находиться значение p, чтобы наиболее вероятное число успехов во всей серии было равно k.

Следующие три задачи ведут свою историю с XV–XVII в. Все они касаются игр, которые были основным источником первых вопросов и проблем, связанных с вероятностью.

249*. **Задача шевалье де Мере.** Что вероятнее: в 4 бросаниях одной игральной кости получить хотя бы один раз 1 или в 24 бросаниях двух костей хотя бы один раз получить две 1? Как вы думаете, почему эту задачу часто называют парадоксом де Мере?

250*. **Задача Сэмуэля Пипса.** Будут ли одинаковыми шансы на успех следующих трёх человек, если первому нужно получить хотя бы одну шестёрку при 6 бросаниях кости, второму — не менее 2 шестёрок при 12 бросаниях, третьему — не менее 3 шестёрок при 18 бросаниях?

251*. **Задача Луки Пачоли о разделе ставки.** Двое играют в игру, состоящую из отдельных партий без ничьих. Вероятность выигрыша для обоих игроков одинаковая — 0,5. Победителем считается тот, кто первым выиграет 6 партий. В каком отношении следует разделить приз, если игра была внезапно прервана при счёте:
а) 5 : 3; б) 4 : 3; в) 4 : 2?

Указание. Эта задача была впервые сформулирована более 500 лет тому назад. За её долгую историю было решено, что делить приз следует пропорционально вероятностям, с которыми игроки довели бы прерванную игру до победы. Их и нужно найти в этой задаче.

252*. Что вероятнее — выиграть у равного по силе противника:
а) три партии из шести или четыре партии из восьми;
б) более трёх партий из шести или более четырёх партий из восьми;
в) более n партий из $2n$ или более $n + 1$ партий из $2n + 2$?

253*. Петя бросил монету 100 раз, а Вася — 101 раз. С какой вероятностью орлов у Васи выпало больше, чем у Пети?

254*. Монету бросают $2n$ раз. С какой вероятностью орлов в любой момент будет не меньше, чем решек?

255*. На борту воздушного судна, вместимость которого составляет 150 мест, пассажирам предлагается 2 вида обедов: с курицей и с рыбой. Каждый пассажир может с равной вероятностью выбрать курицу или рыбу, поэтому на борт загрузили 75 обедов с курицей и рыбой.
а) С какой вероятностью все пассажиры будут довольны?
б) С какой вероятностью недовольных пассажиров будет больше одного?
в) Какое будет наиболее вероятное число недовольных пассажиров?

256*. **Задача Стефана Банаха.** Некий математик всегда носит с собой две коробки со спичками. Каждый раз, когда нужно зажечь огонь, он достаёт наугад одну из двух коробок и поджигает спичку. Через какое-то время он обнаруживает, что одна из коробок пуста. С какой вероятностью в другой коробке в это время лежит k спичек? В начальный момент в каждой коробке было по n спичек.

С. Банах (1892—1945) — польский математик, один из создателей ***функционального анализа*** — нового раздела математики, связанного с теорией вероятностей и математической статистикой.

257*. Докажите теорему о наиболее вероятном числе успехов в N испытаниях Бернулли.

§ 14. Испытания до первого успеха

1 Когда же наступит успех?

Можно вспомнить много ситуаций, в которых вы с нетерпением ждали, когда же вам повезёт: ваш любимый форвард попадёт, наконец, в ворота (а лучше — забьёт гол); вас вызовут (или не вызовут?) к доске; купленный в очередной раз лотерейный билет выиграет и т. д.

Теперь вы знаете, что серию таких испытаний, каждое из которых может закончиться успехом или неудачей, называют испытаниями Бернулли и для их анализа можно использовать полученные ранее результаты, например формулу Бернулли.

Но в описанных примерах ситуация несколько иная, чем была до этого. Здесь нет фиксированного числа испытаний N, как в формуле Бернулли. И мы не вычисляем, с какой вероятностью в этих испытаниях произойдёт ровно k успехов. Мы ждём теперь только первого успеха и хотим вычислить, как долго может продлиться это ожидание.

Если сформулировать задачу более точно, то она будет выглядеть так. Проводится серия испытаний Бернулли с вероятностью успеха p и вероятностью неудачи $q = 1 - p$. Испытания проводятся до появления первого успеха. Нужно вычислить вероятность, с которой первый успех появится на первом шаге, на втором шаге и т. д.

С одной стороны, ситуация даже проще: параметр N отсутствует, поэтому искомая вероятность зависит только от номера шага k, на котором происходит первый успех, и от вероятности успеха p. С другой — первый успех может наступить на любом шаге, поэтому этих вероятностей бесконечно много. Обозначим их по аналогии с вероятностями в формуле Бернулли $P(1)$, $P(2)$, ... и найдём общую формулу для $P(k)$.

Но сначала рассмотрим несколько примеров.

Пример 1. Монету подбрасывают до тех пор, пока не выпадет орёл. С какой вероятностью это произойдёт на первом шаге? на втором? на k-м?

Пример 2. Кубик подбрасывают до тех пор, пока не выпадет шестёрка. С какой вероятностью это произойдёт на первом шаге? на втором? на k-м?

Пример 3. Ученик отвечает на вопросы теста, выбирая каждый раз наугад один из четырёх вариантов ответа. Тестирование продолжается до тех пор, пока ученик не даст первый правильный ответ. С какой вероятностью это произойдёт на первом вопросе? на втором? на k-м?

Вернёмся к общему случаю. Вероятность, что успех наступит уже в первом испытании, равна p, поэтому $P(1) = p$.

Чтобы первый успех наступил во втором испытании, нужно, чтобы сначала выпала неудача, а потом успех, поэтому $P(2) = qp$.

Первый успех наступит в третьем испытании, если первые два испытания были неудачными, а третье — успешным, поэтому $P(3) = q \cdot q \cdot p = q^2p$ и т. д.

Если вспомнить способ, которым мы кодировали серии испытаний Бернулли в предыдущем разделе, то возможными исходами нашего опыта будут теперь такие последовательности:

$$1,\ 01,\ 001,\ 0001,\ 00001,\ \ldots$$

(напомним, что цифра 1 соответствует успеху, а цифра 0 — неудаче). Заметим, что в этом опыте исходов **бесконечно много** — ведь мы не можем гарантировать, что успех обязательно выпадет до какого-то определённого шага.

Вероятностями этих исходов будут:

$$p,\ qp,\ q^2p,\ q^3p,\ q^4p,\ \ldots .$$

Общая формула для вероятности получения первого успеха на k-м шаге, где $k \geqslant 1$, будет выглядеть так:

$$\boxed{P(k) = q^{k-1}p}$$

Найдём с её помощью вероятности в примерах 1—3.

Пример 1. Для монеты $p = q = \frac{1}{2}$, поэтому $P(k) = \left(\frac{1}{2}\right)^k$. Найдём несколько первых значений $P(k)$ (табл. 54).

Таблица 54

k	1	2	3	4	5	...
P(k)	0,5	0,25	0,125	0,0625	0,0313	...

Пример 2. Для кубика $p = \frac{1}{6}$, $q = \frac{5}{6}$, поэтому $P(k) = \frac{5^{k-1}}{6^k}$ (табл. 55).

Таблица 55

k	1	2	3	4	5	...
$P(k)$	0,167	0,139	0,116	0,0965	0,0804	...

Пример 3. В задаче с тестированием $p = \frac{1}{4}$, $q = \frac{3}{4}$, поэтому $P(k) = \frac{3^{k-1}}{4^k}$ (табл. 56).

Таблица 56

k	1	2	3	4	5	...
$P(k)$	0,25	0,186	0,141	0,105	0,0791	...

Вы заметили, что во всех примерах найденные вероятности с ростом k убывают и довольно быстро. В этом нет ничего удивительного: если внимательно посмотреть на формулу, то можно увидеть, что числа $P(k)$ образуют бесконечно убывающую геометрическую прогрессию со знаменателем $q < 1$ и первым членом, равным p.

Вспомним формулу для суммы такой прогрессии, известную вам из курса алгебры:

$$p + qp + q^2p + q^3p + q^4p + ... = \frac{p}{1-q} = \frac{p}{p} = 1.$$

Получили интересный факт: сумма всех найденных вероятностей равна 1. Это значит, что первый успех **обязательно наступит на каком-то шаге** и его не придётся ждать бесконечно долго.

Согласитесь — весьма утешительный результат.

? ВОПРОСЫ

1. С какой вероятностью при подбрасывании монеты первый орёл выпадет во втором испытании? в третьем испытании? вообще никогда не выпадет?
2. С какой вероятностью при подбрасывании кубика первая четвёрка выпадет в первом же испытании? во втором испытании? вообще никогда не выпадет?
3. Две монеты подбрасывают до тех пор, пока они выпадут на одну сторону. С какой вероятностью это произойдёт во втором испытании? в третьем испытании? вообще никогда не произойдёт?

2 Сколько испытаний провести?

Испытания Бернулли до первого успеха — первая модель случайного опыта, в котором мы столкнулись **с бесконечным числом неравновозможных исходов.** Это приводит к тому, что при вычислении вероятностей многих случайных событий приходится вычислять суммы, содержащие бесконечное число слагаемых.

Одну из таких сумм, равную 1, мы только что нашли — она оказалась суммой бесконечно убывающей геометрической прогрессии. Рассмотрим ещё несколько примеров, которые приводят к подобным вычислениям.

Пример 1. По каналу связи, в котором есть помехи, передаются сообщения. Вероятность правильной передачи сообщения составляет 0,7. Если сообщение передано с ошибками, оно повторяется. С какой вероятностью для правильной передачи понадобится более 3 попыток?

Каждая передача сообщения — это испытание Бернулли. Если считать успехом передачу без ошибок, то $p = 0{,}7$; $q = 0{,}3$. Чтобы найти вероятность интересующего нас события, нужно сложить вероятности $P(4)$, $P(5)$, ... и т. д. до бесконечности. Запишем эту сумму, используя выведенную ранее формулу: $q^3p + q^4p + q^5p + \ldots$. Мы видим, что это сумма бесконечно убывающей геометрической прогрессии со знаменателем q и первым членом q^3p, поэтому она равна

$$\frac{q^3p}{1-q} = \frac{q^3p}{p} = q^3 = 0{,}3^3 = 0{,}027.$$

Полученный простой ответ подсказывает и более простое решение. В самом деле, чтобы понадобилось более 3 попыток, необходимо и достаточно, чтобы первые три попытки были неудачными, а вероятность этого по правилу умножения для независимых событий равна q^3.

Есть и ещё одно решение — через вероятность противоположного события $\overline{A}$ = «понадобится не более трёх попыток»:

$$P(\overline{A}) = p + qp + q^2p = p\frac{1-q^3}{1-q} = 1 - q^3.$$

Отсюда $P(A) = 1 - P(\overline{A}) = 1 - (1 - q^3) = q^3$.

Пример 2. Антон и Борис по очереди бросают монету. У кого первого выпадет орёл — тот и выиграл. Антон бросает первым. Справедлива ли такая игра?

Чтобы ответить на этот вопрос, нужно вычислить вероятности выигрыша для каждого из игроков.

Игру можно представить как испытания Бернулли до первого успеха — первого орла. При этом, если орёл выпадает в первом, третьем, пятом и вообще в любом нечётном испытании, то выигрывает Антон, а если в любом чётном — то Борис. Найдём вероятность того, что выиграет Антон, сложив соответствующие вероятности:

$$p + q^2p + q^4p + \ldots = \frac{p}{1-q^2} = \frac{p}{(1-q)(1+q)} = \frac{p}{p(1+q)} = \frac{1}{1+q} = \frac{2}{3}.$$

Мы воспользовались здесь тем, что это сумма бесконечно убывающей геометрической прогрессии со знаменателем q^2 и первым членом p, а также тем, что $p = q = \frac{1}{2}$. Аналогично получаем вероятность выигрыша для Бориса:

$$p + q^3p + q^5p + \ldots = \frac{qp}{1-q^2} = \frac{qp}{(1-q)(1+q)} = \frac{qp}{p(1+q)} = \frac{q}{1+q} = \frac{1}{3}.$$

Видим, что игра несправедливая: шансы Антона в два раза выше. Заодно мы убедились, что в этой игре не бывает ничьих, так как $\frac{2}{3} + \frac{1}{3} = 1$ (точнее, что вероятность ничьей равна 0).

Ещё один пример связан с так называемой ***обратной задачей***, когда нужно не по числу испытаний найти вероятность, а наоборот — по заданной вероятности найти, сколько потребуется испытаний.

Пример 3. Вероятность, что тест обнаружит ошибку, которая содержится в компьютерной программе, составляет около 0,2. Сколько тестов нужно подготовить, чтобы выявить ошибку (разумеется, если она там есть) с вероятностью, большей 0,99?

Испытания Бернулли состоят в последовательном запуске программы на различных тестах. Успехом считается обнаружение ошибки, поэтому $p = 0{,}2$; $q = 0{,}8$.

Получается, что нам нужно найти такое наименьшее значение k, при котором

$$p + qp + q^2p + q^3p + \dots + q^{k-1}p > 0{,}99.$$

Сумму в левой части неравенства можно найти либо как сумму конечной геометрической прогрессии, либо вычитанием из 1 суммы бесконечно убывающей геометрической прогрессии (мы уже делали это в примере 1):

$$1 - q^k > 0{,}99 \quad \text{или} \quad q^k < 0{,}01.$$

Поскольку q^k с ростом k приближается к нулю, то нужное значение k можно найти простым перебором, последовательно возводя 0,8 во 2, 3 и т. д. степени:

$$0{,}8^2 = 0{,}64$$
$$\dots$$
$$0{,}8^{20} > 0{,}011$$
$$0{,}8^{21} < 0{,}0093.$$

Получилось, что нужно проверить программу на 21-м тесте.

Для более быстрого способа решения таких неравенств, без использования перебора, можно использовать ***логарифмирование***.

? ВОПРОСЫ

1. Антон и Борис из примера 2 решили по очереди бросать не монету, а кубик. Выигрывает тот, у кого первого выпадет 6 очков. С какой вероятностью Антон выиграет?
2. Сколько раз нужно подбросить монету, чтобы вероятность появления орла была больше 0,9?
3. Сколько кубиков нужно подбросить, чтобы вероятность появления хотя бы одной единицы была больше 0,9?

3 Испытания Бернулли в электронной таблице

Для вычисления вероятностей случайных событий, связанных с испытаниями Бернулли, в электронных таблицах есть три функции:

ФАКТР(N) — вычисляет факториал $N!$;

ЧИСЛКОМБ(N; k) — вычисляет число сочетаний C_N^k;

БИНОМРАСП(k; N; p; 0 или 1) — вычисляет вероятность получить k успехов в N испытаниях Бернулли.

При этом последний, четвёртый параметр функции БИНОМРАСП() имеет следующий смысл:

- если он равен 0, то вычисляется вероятность получить ровно k успехов;
- если он равен 1, то вычисляется вероятность получить от 0 до k успехов.

Таким образом, в принятых выше обозначениях

БИНОМРАСП(k; N; p; 0) $= P_N(k)$;

БИНОМРАСП(k; N; p; 1) $= P_N(0) + P_N(1) + \dots + P_N(k)$.

Покажем, как вычислить все вероятности $P_N(k)$ для испытаний с кубиком.

Пример 1. Кубик бросают 20 раз. Вычислим вероятность получить k шестёрок для $k = 0, 1, \dots, 20$.

	A	B
1	Число успехов	Вероятность
2	0	0,026084053
3	1	0,104336213
4	2	0,198238805
5	3	0,237886566
6	4	0,202203581
7	5	0,129410292
8	6	0,064705146
9	7	0,025882058
10	8	0,008411669
11	9	0,002243112
12	10	0,000493485
13	11	8,97245E-05
14	12	1,34587E-05
15	13	1,65645E-06
16	14	1,65645E-07
17	15	1,32516E-08
18	16	8,28226E-10
19	17	3,89753E-11
20	18	1,29918E-12
21	19	2,73511E-14
22	20	2,73511E-16

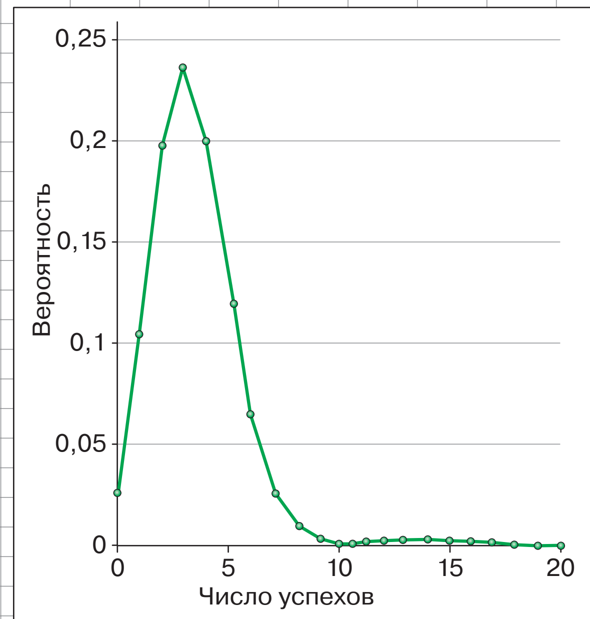


Рис. 90. Вероятности получить k успехов для $k = 0, 1, \ldots, 20$

В столбце A запишем все значения k от 0 до 20. В столбце B с помощью функции БИНОМРАСП() вычислим соответствующие вероятности. На рисунке 90 показан получившийся результат. Здесь же построена диаграмма, которая показывает, как ведут себя найденные вероятности при изменении k.

Хорошо видно, что после $k = 7$ вероятности быстро убывают и становятся практически нулевыми. Своего наибольшего значения вероятность достигает при $k = 3$. Это соответствует теореме о наиболее вероятном числе успехов в N испытаниях Бернулли.

С помощью рассмотренной ранее функции СЛЧИС() можно **смоделировать** серию испытаний Бернулли с заданной вероятностью успеха p.

Пример 2. Смоделируем серию из 20 испытаний Бернулли с заданной вероятностью успеха p. При $p = \frac{1}{6}$ это будет соответствовать испытаниям с кубиком, которые мы рассмотрели в примере 1.

Запишем в ячейку D2 вероятность успеха $p = \frac{1}{6}$. В столбце A ввведём номера испытаний от 1 до 20. Чтобы получить результат одного испытания Бернулли, разыграем с помощью функции СЛЧИС() случайное число от 0 до 1: если оно окажется меньше p, то будем считать, что испытание завершилось успехом, если нет — то неудачей. Для этого введём в ячейку B2 формулу:

=ЕСЛИ(СЛЧИС()<D2;1;0)

и скопируем её вниз на 20 ячеек. Обратите внимание, что адрес D2 записан как абсолютный, чтобы при копировании формулы ссылка на ячейку, где записана вероятность успеха p, не изменялась.

В ячейке D3 с помощью функции СЧЁТЕСЛИ() найдём число успехов в серии. На рисунке 91 показаны результаты моделирования: шестёрка выпала 3 раза — в испытаниях 5, 7 и 12.

	A	B	C	D
1	№ испытания	Результат	p =	0,166667
2	1	0	k =	3
3	2	0		
4	3	0		
5	4	0		
6	5	1		
7	6	0		
8	7	1		
9	8	0		
10	9	0		
11	10	0		
12	11	0		
13	12	1		
14	13	0		
15	14	0		
16	15	0		
17	16	0		
18	17	0		
19	18	0		
20	19	0		
21	20	0		

Рис. 91. Моделирование испытаний Бернулли

Чтобы провести новую серию испытаний, достаточно нажать кнопку F9 — лист электронной таблицы обновится, и все случайные числа изменятся. Так можно проводить неограниченное число повторений и следить за тем, сколько успехов (т. е. сколько шестёрок) при этом выпадает.

? ВОПРОСЫ

1. Какие функции позволяют вычислить вероятность заданного числа успехов?
2. Какие аргументы имеет функция БИНОМРАСП()?
3. Как смоделировать в электронной таблице серию испытаний Бернулли?

ЛАБОРАТОРНАЯ РАБОТА № 3

Испытания Бернулли: подсчёт вероятностей и моделирование

Задание 1. За сколько ответов ставить зачёт?

Тест на зачёте состоит из 20 вопросов. К каждому из них предлагаются 4 варианта ответа, ровно один из которых верный. За сколько правильных ответов можно ставить зачёт, чтобы не знающий предмета студент смог пройти тест с вероятностью не больше 0,5? 0,1? 0,01?
Чтобы ответить на этот вопрос, вычислите вероятности соответствующих событий в электронной таблице, используя для этого функцию БИНОМРАСП().

Этапы выполнения задания:

1. Заполните столбец A возрастающими значениями k от 0 до 20.
2. Найдите в каждой ячейке столбца B вероятность ответить ровно на k вопросов теста. Используйте для этого формулу =БИНОМРАСП(20; k; 1/4; 0).
3. Вычислите в каждой ячейке столбца C суммарную вероятность **ответить на k и больше вопросов** теста.
4. С помощью полученных в столбце C вероятностей ответьте на поставленные вопросы (на рисунке 92 найденные ответы выделены цветом).

	A	B	C
1	0	0,003171212	1
2	1	0,021141413	0,996828788
3	2	0,066947808	0,975687788
4	3	0,133895615	0,908739568
5	4	0,189685455	0,774843952
6	5	0,202331152	0,585158497
7	6	0,168609293	0,382827346
8	7	0,112406195	0,214218052
9	8	0,060886689	0,101811857
10	9	0,027060751	0,040925168
11	10	0,009922275	0,013864417
12	11	0,00300675	0,003942142
13	12	0,000751688	0,000935392
14	13	0,000154192	0,000183704
15	14	2,56987E-05	2,9511E-05
16	15	3,4265E-06	3,81303E-06
17	16	3,56927E-07	3,86532E-07
18	17	2,79942E-08	2,9605E-08
19	18	1,55524E-09	1,61072E-09
20	19	5,45697E-11	5,54792E-11
21	20	9,09495E-13	9,09495E-13

Рис. 92. Пример выполнения задания 1

	A	B	C
1	0	0,25	5
2	0	6	3
3	0		1
4	0		2
5	0		7
6	0		4
7	1		5
8	0		7
9	0		5
10	0		5
11	1		6
12	0		5
13	1		
14	1		
15	0		
16	0		
17	0		
18	1		
19	1		
20	0		

Рис. 93. Фрагмент задания 2

Задание 2. Моделирование испытаний Бернулли

Смоделируйте испытания Бернулли из задания 1 в электронной таблице и проверьте полученные в этом задании ответы экспериментально.

Этапы выполнения задания:

1. Введите в ячейку B1 значение $p = 0{,}25$.
2. Смоделируйте в столбце A результаты 20 испытаний Бернулли с вероятностью успеха $p = 0{,}25$.
3. Найдите в ячейке B2 число успехов (т. е. количество верных ответов).
4. Нажмите несколько раз клавишу F9 и понаблюдайте, как изменяется число успехов от серии к серии.
5. Последовательно заполните ячейки столбца C значениями из ячейки B2. После заполнения каждой очередной ячейки C **результаты случайных испытаний будут автоматически обновляться** и можно будет заполнять очередную ячейку новым значением. На рисунке 93 таким образом заполнены 12 ячеек C1:C12.
6. Найдите частоты каждого числа верных ответов.
7. Сравните полученные результаты с ответами, полученными в задании 1.

УПРАЖНЕНИЯ

258. Монету бросают до тех пор, пока не появится орёл. С какой вероятностью придётся сделать:
а) больше 3 бросков; б) меньше 5 бросков; в) от 3 до 5 бросков?

259. Кубик бросают до тех пор, пока не выпадет 5 очков. С какой вероятностью придётся сделать:
а) больше 3 бросков; б) меньше 5 бросков; в) от 3 до 5 бросков?

260. Телефон передаёт СМС-сообщение. В случае неудачи телефон делает следующую попытку. Вероятность безошибочной передачи сообщения в каждой попытке равна 0,6. С какой вероятностью для передачи сообщения потребуется не больше двух попыток?

261. Максим заходит на сайт музея, чтобы купить билеты на престижную выставку. Из-за большого числа обращений вероятность успешного оформления билета составляет 0,4. После четырёх неудачных попыток сайт блокируется. С какой вероятностью Максиму удастся купить билет?

262. Баскетболист на тренировке бросает мяч со штрафного до тех пор, пока не попадёт в корзину. Вероятность попадания одного броска равна 0,3. С какой вероятностью ему придётся сделать:
а) меньше 3 бросков; б) больше 3 бросков; в) чётное число бросков?

263. Маша коллекционирует принцесс из шоколадных яиц. Всего в коллекции 10 разных принцесс. У Маши уже есть две разные принцессы из коллекции. Какова вероятность того, что для получения следующей принцессы Маше придётся купить ещё:
а) 1 яйцо; б) больше 1 яйца; в) больше 2 яиц?

264. Игра состоит в набрасывании колец на колышек. Игрок получает 6 колец и бросает их до первого попадания. Вероятность точного броска равна 0,2. С какой вероятностью хотя бы одно кольцо останется неиспользованным?

265. Сколько раз нужно бросить монету, чтобы вероятность появления хотя бы одного орла стала больше 0,9? больше 0,99? больше 0,999?

266. Сколько раз нужно бросить кубик, чтобы вероятность появления хотя бы одной единицы была больше 0,9? больше 0,99? больше 0,999?

267. Вероятность рождения мальчика составляет 0,54. Сколько детей должно быть в семье, чтобы вероятность рождения хотя бы одного мальчика была не меньше:
а) 0,9; б) 0,99; в) 0,999?

268. Стрелок в тире стреляет по мишени до тех пор, пока не поразит её. Известно, что он попадает в цель с вероятностью 0,3 при каждом отдельном выстреле. Какое наименьшее количество патронов нужно дать стрелку, чтобы он поразил цель с вероятностью не менее 0,9?

269*. Вы решили найти незнакомого человека, у которого день рождения совпадает с вашим. Сколько незнакомцев вам нужно опросить, чтобы вероятность встретить такого человека была больше 0,5?

270*. Профессор Злобин не хочет ставить положительную оценку студенту Разгуляеву и задаёт ему вопросы до тех пор, пока на очередном вопросе студент не ошибётся. Вероятность правильного ответа на любой вопрос составляет 0,8. Сколько вопросов нужно иметь в запасе профессору, чтобы их хватило для выставления двойки с вероятностью не менее 0,95?

271*. Процент всхожести семян редкого сорта цветов составляет 0,2. Сколько семян нужно посадить, чтобы с вероятностью 0,99 взошёл хотя бы 1 цветок? хотя бы 2 цветка?

272*. Мастер должен собрать в квартире заказчика книжный шкаф. При сборке шкафа используется 20 стяжек, для каждой из которых нужен специальный болт. Эти болты имеют обыкновение ломаться при сборке с вероятностью 0,2. Сколько таких болтов нужно взять с собой мастеру, чтобы собрать шкаф с вероятностью:

а) не менее 0,9; б) не менее 0,95; в) не менее 0,99?

273*. Сколько раз нужно бросить кубик, чтобы вероятность получения хотя бы двух одинаковых чисел была больше 0,5?

274*. Петя и Вася бросают монету, пока не появятся подряд ОО или ОР (буква О соответствует орлу, буква Р — решке). Если раньше появится ОО — выигрывает Вася, если ОР — Петя.
а) У кого больше шансов на выигрыш?
б) Петя изменил свою тактику и поставил на РО. Изменится ли ответ?

275*. Артём бросает монету 2 раза, а Богдан — 3 раза. У кого выпало больше орлов, выигрывает. Если орлов выпало поровну, то броски повторяются до победы одного из игроков. С какой вероятностью каждый из них выиграет?

276*. Вам с товарищем необходимо бросить жребий, кто из вас двоих будет мыть посуду. У вас имеется только канцелярская кнопка, для которой вероятности двух исходов неодинаковые и вам неизвестны. Придумайте, как с её помощью бросить справедливый жребий.

277*. Частица совершает случайное блуждание на числовой прямой. В начальный момент времени частица находится в точке 0. Каждую секунду она может перепрыгнуть на 1 вправо с вероятностью $p = 0{,}8$ или на 1 влево с вероятностью $q = 1 - p = 0{,}2$. Какова вероятность того, что рано или поздно частица попадёт в точку –1?

§ 15. Случайный выбор*

1 Случайный выбор без возвращения

Рассмотрим ещё одну модель испытаний, очень популярную в теории вероятностей и важную с практической точки зрения. Вспомним для начала пример, который мы уже рассматривали раньше при знакомстве с испытаниями Бернулли.

Пример 1. В корзинке лежат 5 пирожков с капустой и 5 с повидлом. Бабушка любит пирожки с капустой и считает выбор такого пирожка большим успехом. Она выбирает из корзинки и съедает 4 пирожка. С какой вероятностью все они окажутся с капустой?

Мы уже выясняли, что считать такую серию из 4 случайных опытов испытаниями Бернулли нельзя, так как результаты опытов зависимы друг от друга. Тем не менее даже для зависимых событий мы можем использовать формулу умножения вероятностей. Обозначим A_1, A_2, A_3, A_4 события, которые состоят в том, что соответственно в первом, втором, третьем и четвёртом испытаниях будет выбран пирожок с капустой. Нам нужно найти вероятность пересечения этих четырёх событий. По формуле умножения получаем:

$$P(A_1A_2A_3A_4) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)P(A_4 \mid A_1A_2A_3) =$$
$$= \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{3}{8} \cdot \frac{2}{7} = \frac{1}{42} \approx 0{,}0238.$$

Можно решить ту же задачу по-другому. Исходами опыта, в котором из корзинки с 10 пирожками друг за другом вынимают 4 пирожка, будут размещения из 10 по 4. Общее количество таких исходов равно

$$A_{10}^4 = \frac{10!}{6!} = 10 \cdot 9 \cdot 8 \cdot 7.$$

Благоприятными из этих исходов будут те, в которых все 4 пирожка выбраны из тех 5, которые с капустой. Это размещения из 5 по 4:

$$A_5^4 = \frac{5!}{1!} = 5 \cdot 4 \cdot 3 \cdot 2.$$

Как видим, отношение числа благоприятных исходов к числу всех равновозможных исходов даёт тот же результат.

Рассмотренная в этом примере схема случайного выбора называется выбором без возвращения. Уточним, как такой выбор происходит.

Пусть имеется ящик (коробка, «урна»), содержащий n неразличимых на ощупь предметов. ***Случайным выбором без возвращения*** называется последовательность из k испытаний ($k \leqslant n$), в каждом из которых наугад вынимается один из находящихся в ящике предметов. Перед следующим испытанием вынутый предмет в ящик не возвращается. **Элементарными исходами такого выбора являются размещения из n по k.**

Понятно, почему в этом определении $k \leqslant n$: после n таких опытов ящик опустеет. Разумеется, «ящик» и «предмет» понимаются здесь в обобщённом смысле. Мы можем, например, случайно выбирать людей из какой-то группы или вообще разыгрывать такой выбор с помощью компьютера.

Случайный выбор без возвращения нельзя считать испытаниями Бернулли, поскольку условия каждого следующего опыта будут зависеть от результатов предыдущих опытов. Однако вопросы, которые здесь возникают, очень похожи.

Попробуем, например, ответить в рассмотренном выше примере 1 на такой вопрос. С какой вероятностью ровно половина вынутых пирожков будет с капустой? На языке испытаний Бернулли это была бы вероятность получить два успеха в четырёх испытаниях. Но здесь формула Бернулли, как мы уже выяснили, нам не поможет.

Очень часто для решения этой задачи предлагается следующее **неверное решение**:

$$P = \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{5}{8} \cdot \frac{4}{7} = \frac{5}{63} \approx 0{,}0794.$$

На первый взгляд, оно может показаться правильным: мы перемножаем вероятности двух успехов и двух неудач. Но ведь эти два успеха и две неудачи, как и в испытаниях Бернулли, могут чередоваться в разном порядке! Другими словами, мы нашли только вероятность $P(A_1 A_2 \overline{A}_3 \overline{A}_4)$, а есть ведь ещё $P(A_1 \overline{A}_2 A_3 \overline{A}_4)$, $P(A_1 \overline{A}_2 \overline{A}_3 A_4)$ и т. д.

Несложно догадаться, что все эти вероятности будут одинаковыми. Нужно только посчитать, сколькими способами можно расположить два «успеха» и две «неудачи» между собой. А такой подсчёт мы уже делали, когда доказывали формулу Бернулли, — это будет число сочетаний из 4 по 2:

$$C_4^2 = \frac{3 \cdot 4}{2} = 6.$$

Для получения правильного ответа нужно домножить полученную ранее вероятность на это число:

$$P = C_4^2 \cdot \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{5}{8} \cdot \frac{4}{7} = \frac{10}{21} \approx 0{,}476.$$

Теперь и результат получился более правдоподобный — ведь расклад 2 + 2 среди 4 выбранных пирожков наиболее вероятный, а значит, его вероятность не могла быть такой маленькой, как была получена в начале.

Вообще, решая задачи на вычисление вероятностей, не забывайте сверять полученные ответы со здравым смыслом: чтобы вероятности маловероятных событий не получались близкими к 1, а вероятности почти достоверных — к 0.

И ещё один замечательный пример, который называют часто «парадоксом дней рождения». Придумал его австрийский математик Р. Мизес (1883—1953), обосновавший в своих работах частотный подход к определению вероятности.

Пример 2. **(Парадокс дней рождения.)** В классе учится 25 человек. С какой вероятностью хотя бы двое из них празднуют день рождения в один день?

Мы будем считать, что каждый из учеников может с равной вероятностью родиться в любой из 365 дней года. Правда, бывают ещё и високосные годы, в которых на один день больше, но его учёт внесёт лишние технические трудности и почти не скажется на ответе.

Общее число равновозможных исходов можно определить по правилу умножения:

$$\underbrace{365 \cdot 365 \cdot \ldots \cdot 365}_{25\text{ раз}} = 365^{25}.$$

Как всегда, когда в задаче встречаются слова «хотя бы...», лучше перейти к противоположному событию и посчитать неблагоприятные исходы. Для такого исхода нам нужно выбрать для каждого из 25 человек свой день рождения из 365 дней года, т. е. произвести случайный выбор 25 дней из 365 без возвращения. Любой такой выбор будет размещением из 365 по 25, поэтому неблагоприятных исходов будет

$$A_{365}^{25} = \frac{365!}{340!} = 365 \cdot 364 \cdot \ldots \cdot 341.$$

Чтобы найти отношение этих чисел, лучше использовать компьютер:

$$\frac{A_{365}^{25}}{365^{25}} = \frac{365}{365} \cdot \frac{364}{365} \cdot \ldots \cdot \frac{341}{365} \approx 0{,}4313.$$

Отсюда вероятность, что в классе из 25 человек найдутся двое, родившиеся в один день, будет

$$1 - 0{,}4313 = 0{,}5867 > 0{,}5.$$

Собственно, в этом и состоит парадокс: людей немного, а вероятность совпадающих дней рождения больше половины! Разумеется, парадокс этот только кажущийся. Вероятность совпадения дней рождения у пары конкретных людей действительно небольшая — всего $\frac{1}{365} \approx 0{,}0027$. Но пар в этом классе достаточно много: $\frac{25 \cdot 24}{2} = 300$. Про этот факт и забывает «обычная интуиция», которая подсказывает, что вероятность должна быть небольшой.

? ВОПРОСЫ

1. Объясните, что такое случайный выбор без возвращения.
2. Что можно считать элементарными исходами случайного выбора без возвращения?
3. Из ящика, в котором лежат 2 красных носка и 4 чёрных, вытаскивают наугад 2 носка. С какой вероятностью они оба будут красными? оба будут чёрными?

2 Одновременный случайный выбор

Модель случайного выбора без возвращения часто заменяют на другую, в которой предметы вынимаются из ящика не друг за другом, а одновременно. Уточним, как это происходит.

Пусть имеется ящик (коробка, «урна»), содержащий n неразличимых на ощупь предметов. ***Одновременным случайным выбором*** называется одно испытание, в котором из ящика наугад вынимаются k предметов ($k \leqslant n$). **Элементарными исходами такого выбора являются сочетания из n по k.**

Таким образом, в одновременном случайном выборе вместо того, чтобы k раз вытягивать из ящика по одному предмету, вынимают все k предметов разом. Случайный выбор без возвращения всегда можно заменить одновременным, если, конечно, нас не интересуют вероятности событий, связанных с порядком, в которых вынимались предметы.

Чтобы объяснить, как происходит переход от одной модели к другой, вернёмся снова к примеру с пирожками.

Пример 1. В корзинке лежат 5 пирожков с капустой и 5 с повидлом. Бабушка выбирает из корзинки и съедает 4 пирожка. Рассмотрим следующие случайные события:

A = «все вынутые пирожки окажутся с капустой»;
B = «два вынутых пирожка окажутся с капустой»;
C = «вторым будет съеден пирожок с капустой».

Из этих трёх событий только событие C использует порядок, в котором вынимаются пирожки, поэтому для вычисления его вероятности понадобится выбор с возвращением.

Найдём вероятности A и B, используя одновременный случайный выбор. Элементарными исходами такого опыта являются сочетания из 10 по 4, поэтому их количество будет равно

$$C_{10}^{4} = \frac{10!}{4!6!} = 210.$$

Благоприятными исходами для события A будут сочетания из 5 пирожков с капустой по 4: $C_5^4 = 5$. Чтобы получить исход, благоприятный для события B, нужно выбрать 2 пирожка из 5 с капустой (это $C_5^2 = 10$ способов) и 2 пирожка из 5 с повидлом (тоже $C_5^2 = 10$ способов). Всего благоприятных исходов для B будет $C_5^2 \cdot C_5^2 = 10 \cdot 10 = 100$.

Отсюда находим вероятности событий A и B:

$$P(A) = \frac{C_5^4}{C_{10}^4} = \frac{5}{210} = \frac{1}{42} \approx 0{,}0238;$$

$$P(B) = \frac{C_5^2 \cdot C_5^2}{C_{10}^4} = \frac{100}{210} = \frac{10}{21} \approx 0{,}476.$$

Как видите, вероятности совпали с найденными ранее. Это подтверждает эквивалентность двух моделей.

Вообще, когда говорят «из n предметов наугад выбирают k», то, как правило, имеют в виду одновременный случайный выбор. Приведём ещё несколько примеров, в которых используется эта модель выбора.

Пример 2. В 10 «А» классе разыгрывают по жребию 6 билетов в театр. С какой вероятностью в театр пойдёт поровну мальчиков и девочек, если в классе учатся 10 мальчиков и 15 девочек?

«Разыгрывают по жребию» в данном случае означает то же самое, что наугад выбирают шестерых учеников, которые пойдут в театр. Всего в классе учится 25 ребят, поэтому общее количество равновозможных исходов такого опыта равно $C_{25}^6 = \dfrac{25!}{6!19!} = 177\ 100$.

Поскольку билетов 6, то «поровну» означает, что будут выбраны 3 мальчика и 3 девочки. Троих мальчиков выбирают из десяти, поэтому количество таких способов будет равно $C_{10}^3 = \dfrac{10!}{3!7!} = 120$. Троих девочек выбирают из пятнадцати, поэтому количество способов будет равно $C_{15}^3 = \dfrac{15!}{3!12!} = 455$.

Общее количество благоприятных исходов можно найти по правилу умножения:

$$C_{10}^3 \cdot C_{15}^3 = \frac{10!}{3!7!} \cdot \frac{15!}{3!12!} = 54\ 600.$$

Для вычисления вероятности остаётся найти отношение числа благоприятных исходов к числу всех исходов опыта:

$$\frac{C_{10}^3 \cdot C_{15}^3}{C_{25}^6} = \frac{10!}{3!7!} \cdot \frac{15!}{3!12!} \cdot \frac{6!19!}{25!} = \frac{10}{25} \cdot \frac{9}{24} \cdot \frac{8}{23} \cdot \frac{15}{22} \cdot \frac{14}{21} \cdot \frac{13}{20} \cdot 20 \approx 0{,}308.$$

Мы специально сократили факториалы именно так: полученное произведение 6 дробей подсказывает и второе решение, в котором используется последовательный выбор без возвращения. Посчитаем вероятность того, что при последовательном выборе 6 человек из 20 будет получен такой результат: МММДДД, где буква М означает, что на очередном шаге выбран мальчик, а буква Д — девочка:

$$\frac{10}{25} \cdot \frac{9}{24} \cdot \frac{8}{23} \cdot \frac{15}{22} \cdot \frac{14}{21} \cdot \frac{13}{20}.$$

Но это ещё не ответ на вопрос задачи. Ведь нас устраивает любая последовательность из букв М и Д, в которой ровно три буквы М и три буквы Д. Мы уже не раз считали количество подобных последовательностей — их будет $C_6^3 = 20$. Все они будут иметь ту же вероятность, что и последовательность МММДДД. Остаётся умножить её на 20, и будет получен тот же самый ответ, что и в первом способе.

Пример 3. На диске «Последний герой» группы «Кино» записаны 10 композиций. Борис ставит в музыкальном плеере режим shuffle, в котором все треки проигрываются в случайном порядке. С какой вероятностью в число первых пяти треков попадёт песня «В наших глазах»? попадут песни «Группа крови» и «Мы ждём перемен»?

Здесь опыт состоит в том, что плеер выбирает случайный порядок проигрывания 10 песен на диске. Но мы можем заменить его на случайный выбор 5 песен из 10. Тогда общее число равновозможных исходов будет равно $C_{10}^5 = 252$.

Чтобы исход был благоприятным для первого события, нужно выбрать песню «В наших глазах», а потом ещё любые 4 песни из 9 оставшихся. Это можно сделать $C_9^4 = 126$ способами. Отсюда вероятность первого события будет равна

$$\frac{C_9^4}{C_{10}^5} = \frac{126}{252} = \frac{1}{2}.$$

По тому же принципу для второго события вероятность равна

$$\frac{C_8^3}{C_{10}^5} = \frac{56}{252} = \frac{2}{9}.$$

Полученные ответы подсказывают ещё одно, более простое решение этой задачи. Представим, что у плеера есть 10 пустых строчек, на которых он должен разместить в случайном порядке 10 песен. Он берёт песню «В наших глазах» и выбирает для неё случайную строчку. Вероятность, что будет выбрана одна из первых пяти строчек, будет равна, конечно, $\frac{5}{10} = \frac{1}{2}$.

Для второго события нужно выбрать две строчки, и обе они должны оказаться среди первых пяти. Эту вероятность можно найти по правилу умножения: $\frac{5}{10} \cdot \frac{4}{9} = \frac{2}{9}$.

? ВОПРОСЫ

1. Объясните, что такое одновременный случайный выбор.
2. Что можно считать элементарными исходами одновременного случайного выбора?
3. Сколькими способами из 6 фильмов можно выбрать для просмотра 2?
4. Чему равна вероятность события C из примера 1?

3 Случайный выбор с возвращениями

Ещё одна схема случайного выбора принципиально отличается от двух рассмотренных и возвращает нас снова к испытаниям Бернулли.

Пример 1. В реке живут 25 щук, из которых 3 щуки — долгожительницы (старше 30 лет). Ихтиолог изучает возраст рыб. Для этого он вылавливает щуку, определяет её возраст и отпускает обратно в пруд.

Договоримся считать испытание успешным, если ихтиологу попалась щука-долгожительница. Тогда в первом испытании вероятность успеха $p = \frac{3}{25}$. И такой же она останется во всех последующих, поскольку состав водоёма не меняется. При таком случайном выборе испытания будут независимыми, поэтому вполне могут считаться испытаниями Бернулли.

Посчитать вероятности различных событий, связанных с этим опытом, можно по формуле Бернулли. Например, вероятность того, что все три выловленные щуки будут долгожительницами, равна

$$P_3(3) = \left(\frac{3}{25}\right)^3 \approx 0{,}00173,$$

а что таких щук будет две —

$$P_3(2) = C_3^2 \cdot \left(\frac{3}{25}\right)^2 \cdot \left(\frac{22}{25}\right)^1 \approx 0{,}038.$$

Использованная здесь схема случайного выбора называется ***случайным выбором с возвращением***. Опишем её в общем виде.

Пусть имеется ящик (коробка, «урна»), содержащий n неразличимых на ощупь предметов. ***Случайным выбором с возвращением*** называется последовательность из k испытаний, в каждом из которых наугад вынимается один из находящихся в ящике предметов. Перед следующим испытанием вынутый предмет возвращается обратно в ящик. **Такую последовательность случайных опытов можно считать испытаниями Бернулли.**

Для выбора с возвращением ограничение $k \leqslant n$ снимается, так как испытаний можно провести сколько угодно.

В социологии выборка, полученная в результате случайного выбора без возвращения, называется ***бесповторной***, а в результате выбора с возвращением — ***повторной***. Если объём генеральной совокупности, из которой производится случайный выбор, очень большой, то различие между бесповторной и повторной выборкой стирается. Это показывает следующий пример.

Пример 2. По данным последней переписи населения, около 35% работающих людей в России имеют высшее образование. Для проведения социологического опроса производят **бесповторную** выборку из 100 работающих человек. Нас интересует, сколько человек и с какой вероятностью среди них будут с высшим образованием.

Понятно, что в этой ситуации успех означает, что выбранный человек имеет высшее образование. Но можно ли назвать такой случайный выбор испытаниями Бернулли?

Если опрос проводится для большой совокупности людей, например всего работающего населения России, какого-то региона или хотя бы города, то случайный выбор из неё одного или даже сотни людей не меняет состава такой совокупности. Вероятность успеха после каждого испытания практически не меняется и остаётся равной 0,35. Значит, такую выборку можно считать повторной и применять к ней те результаты и формулы, которые мы получили для испытаний Бернулли.

? ВОПРОСЫ

1. Объясните, что такое случайный выбор с возвращением.
2. Почему случайный выбор с возвращением можно считать испытаниями Бернулли?
3. Что такое повторная и бесповторная выборка?
4. При каких условиях бесповторная выборка может считаться повторной?

4 Комбинаторика перестановок

На практике часто встречаются ситуации, в которых исходами случайного опыта являются перестановки каких-то объектов. Например, несколько человек случайным образом выстраиваются в очередь в кассу; несколько писем случайно раскладываются по конвертам; музыкальный плеер выбирает случайный порядок проигрывания песен и т. д.

При подсчёте исходов в таких опытах нужно помнить, что общее количество перестановок из n элементов равно $n!$ — т. е. факториалу числа n. Но одной этой формулы для решения подобных задач бывает обычно недостаточно. Приходится использовать с ней другие комбинаторные правила и формулы.

Рассмотрим, как это делается, на примерах.

Пример 1. Маленький Коля, не умеющий читать, играет с карточками, на которых написаны буквы. Он берёт:

а) 4 карточки с буквами А, Г, Р, Ф и случайно выкладывает их в ряд. С какой вероятностью получится слово «ГРАФ»?

б) 4 карточки с буквами А, А, П, Р и случайно выкладывает их в ряд. С какой вероятностью получится слово «ПАРА»?

в) 4 карточки с буквами А, А, М, М, и случайно выкладывает их в ряд. С какой вероятностью получится слово «МАМА»?

г) 10 карточек с буквами А, А, А, Е, И, К, М, М, Т, Т и случайно выкладывает их в ряд. С какой вероятностью получится слово «МАТЕМАТИКА»?

Во всех четырёх опытах равновозможными исходами являются случайные перестановки из n карточек. В первых трёх опытах их будет $4! = 24$, а в четвёртом — $10! = 3\,628\,800$.

В слове «ГРАФ» все буквы разные, поэтому ему соответствует только одна перестановка. Значит, вероятность получить это слово будет равна $\frac{1}{24}$.

В слове «ПАРА» есть две одинаковые буквы «А», от обмена которых ничего не меняется, поэтому ему соответствуют две перестановки. Значит, вероятность получить это слово будет равна $\frac{2}{24} = \frac{1}{12}$.

Слово «МАМА» содержит две пары одинаковых букв. В каждой паре карточки можно поменять местами, а слово при этом не изменится. Поэтому слову «МАМА» соответствуют $2 \cdot 2 = 4$ перестановки. Вероятность получить это слово будет равна $\frac{4}{24} = \frac{1}{6}$.

Наконец, слово «МАТЕМАТИКА» содержит три буквы А и по две буквы М и Т. Внутри каждой группы из одинаковых букв карточки можно менять местами, не изменяя слова. Всего таких обменов можно сделать $3! \cdot 2! \cdot 2! = 24$. Поэтому слову «МАТЕМАТИКА» соответствуют 24 перестановки. Вероятность получить это слово будет равна $\frac{24}{3\,628\,800} = \frac{1}{151\,200}$.

Заметим, что при решении этой задачи нам пришлось иметь дело с перестановками, в которых участвуют одинаковые элементы — в данном случае буквы. В комбинаторике такие перестановки называются ***перестановками с повторениями***. Чтобы посчитать количество таких перестановок, нужно сначала вычислить $n!$, а потом применить правило деления, чтобы не считать одинаковые перестановки по несколько раз.

Так, например, если нам нужно посчитать, сколько разных слов можно получить, переставляя буквы слова «МАТЕМАТИКА», то ответом будет

$$\frac{10!}{3!2!2!} = 151\,200.$$

Пример 2. Три господина пришли в ресторан в одинаковых шляпах и сдали их в гардероб. Уходя домой, они разобрали шляпы наугад. С какой вероятностью все ушли из ресторана в чужих шляпах?

Исходами этого опыта также будут перестановки. Чтобы это увидеть, пронумеруем шляпы, присвоив шляпе первого господина номер 1, второго — номер 2 и третьего — номер 3. Тогда каждый исход опыта можно представить как перестановку из чисел 1, 2, 3.

Например, перестановка 123 будет означать, что все три господина надели свои шляпы, а перестановка 213 — что первый господин обменялся шляпой со вторым, а третий надел свою шляпу.

Итак, этот опыт имеет $3! = 6$ равновозможных исходов. Чтобы найти количество благоприятных исходов (т. е. таких, когда все трое надели чужие шляпы), придётся выписать все 6 перестановок и каждую из них проанализировать (табл. 57).

Таблица 57

Исход	Благоприятный или нет
123	нет
132	нет
213	нет
231	да
312	да
321	нет

Благоприятными для нашего события оказались все перестановки, в которых каждое число стоит не на своём месте — это 231 и 312. В комбинаторике такие перестановки называются ***беспорядками***. Оказалось, что число беспорядков из трёх элементов равно 2. Следовательно, вероятность того, что все три господина уйдут из ресторана в чужих шляпах, равна $\frac{2}{6} = \frac{1}{3}$.

Можно решить аналогичную задачу и для произвольного количества человек n, но для этого нужно вывести общую формулу для числа беспорядков из n элементов. Это число называется в комбинаторике ***субфакториалом*** и обозначается $!n$. Формула для вычисления субфакториала гораздо сложнее, чем для факториала и выглядит так:

$$!n = n! - \frac{n!}{1!} + \frac{n!}{2!} - \frac{n!}{3!} + \ldots + (-1)^n\frac{n!}{n!}.$$

Вывести её можно, используя известную вам ***формулу включений и исключений***. Мы оставим этот вывод в качестве упражнения 304 (довольно сложного).

Заметим ещё, что приведённая задача о шляпах часто приписывается уже знакомому вам Леонарду Эйлеру. Связано это с тем, что с увеличением количества человек n отношение $\frac{!n}{n!}$ приближается к числу $\frac{1}{e}$, где $e = 2{,}71828...$ — одна из самых известных в математике констант — константа Эйлера. Вы ещё не раз встретитесь с ней в самых разных областях математики.

Следующий пример показывает, что иногда полезно использовать равновозможность не только исходов, но и случайных событий. Это может значительно упростить решение задачи.

Пример 3. На музыкальном конкурсе выступают артисты — по одному от каждой из n заявленных стран. Порядок выступления определяется жребием. Какова вероятность того, что артист из России будет выступать после артистов из Белоруссии и Казахстана?

Итак, перед началом конкурса проводится жеребьёвка, которая определяет порядок выступления артистов. Следовательно, равновозможными исходами опыта можно считать все возможные перестановки из n элементов. Всего таких исходов $n!$.

Чтобы найти благоприятные исходы, присвоим всем артистам номера: артисту из России — 1, из Белоруссии — 2, из Казахстана — 3. Тогда благоприятными для нашего события исходами будут такие перестановки из n чисел, в которых число 1 стоит после чисел 2 и 3. Конечно, можно найти количество таких перестановок, выразить его через n и потом поделить полученный результат на $n!$ — но мы попробуем рассуждать по-другому.

Рассмотрим случайное событие, которое состоит в том, что очерёдность наших трёх артистов в результате жеребьёвки будет такая: ...1...2...3... (многоточия обозначают остальных возможных участников жеребьёвки). Обозначим это событие A_{123}. Заметим, что это именно случайное событие — ему соответствует много перестановок, в которых 1 стоит раньше 2, а 2 — раньше 3.

Всего подобных событий будет 6 — столько же, сколько перестановок из трёх чисел:

$$A_{123}, A_{132}, A_{213}, A_{231}, A_{312}, A_{321}.$$

Очевидно, что эти события исчерпывают **все возможные варианты, не пересекаются и имеют одинаковую вероятность**. Отсюда вероятность каждого из них равна $\frac{1}{6}$.

Событие A, о котором спрашивается в нашей задаче, равно объединению двух из этих событий:

$$A = A_{231} \cup A_{321}.$$

Отсюда

$$P(A) = P(A_{231}) + P(A_{321}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

? ВОПРОСЫ

1. Чему равно количество перестановок из n элементов?
2. Что такое перестановки с повторениями?
3. Какие перестановки называются беспорядками?

УПРАЖНЕНИЯ*

278. У маленькой Маши 2 одинаковые пары варежек. Собираясь на прогулку, она наугад выбирает 2 варежки.

а) С какой вероятностью они окажутся парными (т. е. на разные руки)?

б) Маша потеряла одну варежку, и их осталось 3. Собираясь на прогулку, она снова наугад выбирает 2 варежки. С какой вероятностью они окажутся парными на этот раз?

279. В студенческой группе 15 девушек и 10 юношей.

а) По жребию выбирают одного. Найдите вероятность того, что будет выбран юноша.

б) По жребию выбирают двоих. Найдите вероятность того, что будут выбраны юноша и девушка.

280. В ящике находится 20 деталей, среди которых 4 бракованных. Из него вынимают 4 детали. С какой вероятностью:

а) все они будут исправные;

б) все они будут бракованые;

в) хотя бы одна из них будет бракованная;

г) будет поровну исправных и бракованных?

281. В шкафу находятся 4 пары ботинок. Из них случайно выбирают 2 ботинка. С какой вероятностью они окажутся парными, если:

а) все 4 пары ботинок одинаковые;

б) это пары 41, 42, 43 и 44 размеров?

282. В ящике лежат 8 синих, 6 красных и 4 зелёных носка. Из него случайным образом выбирают 2 носка. С какой вероятностью они будут одного цвета?

283. В ящике находятся a красных и b синих шаров. Из него вынимают наугад два шара. Могут ли вынутые шары иметь одинаковый цвет с вероятностью 0,5? Если да, то при каких значениях a, b?

284. На книжной полке стоят 5 книг по математике и 4 книги по физике. С полки снимают наугад 3 книги. С какой вероятностью среди них есть книги и по математике, и по физике?

285. На полке в библиотеке стоял четырёхтомник сочинений Толстого и двухтомник Гоголя. Какие-то три из этих шести книг случайно переставили на другую полку. Найдите вероятность того, что двухтомник Гоголя стоит на одной полке.

286. На замке 10 кнопок с цифрами от 0 до 9. Замок открывается нажатием одновременно на какие-то три кнопки. С какой вероятностью человек, не знающий кода этого замка, откроет его с первого раза?

287. В предыдущей задаче человек начинает беспорядочно нажимать коды из трёх цифр. Каждую секунду он нажимает наугад новый код, возможно повторяя предыдущие. С какой вероятностью он откроет замок в течение минуты? двух минут?

288. Накануне контрольной работы по математике преподаватель опубликовал 10 задач и сообщил, что даст каждому ученику 5 задач, случайно выбранных из этих 10. Андрей знает решение семи из этих задач. С какой вероятностью он решит на контрольной:
а) все 5 задач; б) 4 задачи?

289. Вы ждёте свой багаж у ленты транспортёра. Всего с этого рейса загружено на ленту в случайном порядке 50 чемоданов. С какой вероятностью ваши два чемодана:
а) окажутся первыми;
б) один чемодан будет первым, а другой — последним;
в) будут расположены на ленте рядом друг с другом;
г) окажутся в первой десятке?

290. С помощью датчика случайных чисел получают одно за другим k целых чисел из диапазона от 1 до n. С какой вероятностью:
а) первое число будет 1, второе — 2, третье — 3;
б) первые три числа будут 1, 2, 3 в любом порядке;
в) все k раз появится одно и то же число;
г) все k чисел будут разными?

291. Шесть незнакомых друг с другом пассажиров зашли в лифт на первом этаже одиннадцатиэтажного дома. Каждый из них может с равной вероятностью выйти на любом из десяти этажей. С какой вероятностью:
а) на первых трёх этажах не выйдет ни один из пассажиров;
б) все пассажиры выйдут на первых шести этажах;
г) все пассажиры выйдут на одном этаже;
д) все пассажиры выйдут на разных этажах;
е) на каждом этаже, где будет останавливаться лифт, выйдут по два пассажира?

292. В соревнованиях по бобслею участвует 12 экипажей из 8 стран. Норвегию представляют два экипажа, Россию — один экипаж. Очерёдность выступления определяется жребием. С какой вероятностью российский экипаж будет выступать после двух норвежских?

293. Маленький ребёнок, не умеющий читать, играет с карточками разрезной азбуки. Перед ним лежит слово «КОТ». Он перемешал карточки и снова выложил их в ряд. С какой вероятностью получится слово «ТОК»? Ответьте на тот же вопрос для слов:
а) «РОСИНКА» — «СОРИНКА»;
б) «КУКЛА» — «КУЛАК»;
в) «КАБАН» — «БАНКА»;
г) «СТАРОРЕЖИМНОСТЬ» — «НЕРАСТОРЖИМОСТЬ».

294. Учитель дал четверым десятиклассникам контрольную, а потом собрал работы и случайным образом раздал их обратно, чтобы ученики проверили её друг у друга. С какой вероятностью кому-нибудь из них попадётся его собственная работа?

295. Рядом с вахтёром висит доска с 10 ключами от школьных кабинетов. Каждый ключ на своём крючке. Доска упала, и ключи рассыпались. Вахтёр собрал их и развесил на крючках в случайном порядке. С какой вероятностью:
а) все ключи висят на своих крючках;
б) ровно один ключ висит на своём крючке;
в) хотя бы один ключ висит на своём крючке;
г) все ключи висят не на своих крючках?

296. В сёлах России существовало когда-то такое гадание. Девушка зажимала в руке 6 травинок так, чтобы их концы торчали сверху и снизу. Подруга разбивала 6 верхних концов на пары и связывала каждую пару травинок между собой. Потом то же самое делала с нижними концами. Рука разжималась. Если все 6 травинок оказывались связанными в одно кольцо, девушке предсказывали выйти замуж. С какой вероятностью все 6 травинок свяжутся в одно кольцо?

297. **Продолжение парадокса дней рождения.** Какое наименьшее количество человек должно быть в классе, чтобы вероятность совпадения хотя бы двух дней рождения была:
а) больше 0,5; б) больше 0,9; в) больше 0,99; г) равна 1?

298. С какой вероятностью в компании из пяти человек найдутся два человека, родившиеся в одном месяце? три таких человека?

299. Артём должен сдать тест по вероятности и статистике, который проводится на компьютере. В тесте 10 вопросов, из которых компьютер случайно выбирает для каждого ученика 3 вопроса. Артём выучил ровно половину вопросов. В каком случае его шансы ответить на все три вопроса выше: когда компьютер выбирает их с возвращением (т. е. вопросы могут случайно повториться) или без возвращения (все три вопроса обязательно разные)?

300. Исследуйте полученные в предыдущей задаче вероятности, если Артём выучил k вопросов из 10.

301. В группе дошкольников 10 мальчиков и 10 девочек. Перед выходом на прогулку их случайным образом расставили в пары. С какой вероятностью в каждой паре оказались мальчик и девочка?

302. На конференцию приехало 20 учёных. Перед конференцией 10 из них узнали о новом научном открытии. В гостинице их случайным образом поселили в двухместных номерах. Каждый, кто знал об открытии, рассказал о нём своему соседу по номеру. С какой вероятностью на следующий день о новом открытии знали:
а) все 20 учёных; б) ровно 13 учёных; в) ровно 12 учёных?

303. У кассы кинотеатра стоят в очереди $2n$ человек. Из них n человек имеют только 100-рублёвые купюры, а остальные — только 50-рублёвые. Билет стоит 50 р. Каждый покупает по одному билету. В начальный момент в кассе нет денег. С какой вероятностью каждый покупатель не будет ждать сдачу?

304. Докажите приведённую в этом разделе формулу для числа беспорядков.

Указание. Найдите число перестановок, где хотя бы одно из чисел стоит на своём месте; используйте для этого формулу включений-исключений.

ГЛАВА 7

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И РАСПРЕДЕЛЕНИЯ

До сих пор предметом нашего изучения были ***случайные события*** и их вероятности. В этой главе мы перейдём к изучению ***случайных величин***, играющих в теории вероятностей и её приложениях не менее важную роль, чем случайные события.

Случайное событие может произойти, а может не произойти — это зависит от того исхода, которым закончится случайный эксперимент. Точно так же от этого исхода зависит значение случайной величины. Мы научимся определять вероятности этих значений и вычислять различные характеристики случайной величины. Такая информация позволяет прогнозировать поведение случайной величины в будущем и решать практические задачи.

§ 16. Понятие случайной величины

1 Что такое случайная величина?

Событие называется случайным, поскольку может произойти, а может не произойти в результате случайного эксперимента. Величина называется случайной, потому что её значение зависит от исхода случайного эксперимента. Таким образом, зная исход, которым завершился эксперимент, мы можем точно вычислить значение случайной величины. То есть каждому исходу соответствует вполне определённое значение случайной величины.

В опыте с подбрасыванием двух кубиков можно рассмотреть случайную величину, равную сумме выпавших на кубике чисел. Если опыт завершится исходом (3; 4), она будет равна 7, а при исходе (2; 1) её значением будет 3.

Случайной величиной называется числовая величина, значение которой однозначно определяется исходом случайного эксперимента.

В алгебре числовые величины обозначаются обычно строчными (маленькими) латинскими буквами, например: x, y, t, s. Чтобы не путать случайные величины с неслучайными, мы будем обозначать их заглавными латинскими буквами: X, Y, T, S и т. д.

Рассмотрим некоторые примеры случайных величин, связанных с уже знакомыми вам случайными экспериментами.

Пример 1. Подбрасывают два игральных кубика.

Исходом такого опыта является пара чисел, выпавших на кубиках. С этой парой можно связать много разных случайных величин, например:

X — число, выпавшее на первом кубике;
Y — число, выпавшее на втором кубике;

S — сумма чисел, выпавших на кубиках;
R — произведение чисел, выпавших на кубиках;
W — наибольшее из чисел, выпавших на кубиках;
V — наименьшее из чисел, выпавших на кубиках

(когда на кубиках выпадает одно и то же число, то оно считается и наибольшим, и наименьшим).

Как видите, с одним и тем же опытом можно связать много разных случайных величин — так же как, впрочем, и случайных событий.

Если мы укажем конкретный исход, которым завершился случайный опыт, то можно будет найти значения всех перечисленных величин. Например, для исхода (3; 5) значения этих величин будут следующими:

$$X = 3;\ Y = 5;\ S = 8;\ R = 15;\ W = 5;\ V = 3.$$

Отметим ещё, что одна случайная величина может выражаться через другую или несколько других величин. В нашем примере можно увидеть такие соотношения между величинами:

$$S = X + Y = W + V,\ R = X \cdot Y = W \cdot V.$$

Пример 2. Монету подбрасывают 10 раз подряд.

Исходом такого опыта является последовательность из 10 орлов и решек, произвольно чередующихся между собой. Здесь интересно будет рассмотреть такие случайные величины:

X — число выпавших орлов;
Y — число выпавших решек;
L — длина самой длинной серии из идущих подряд орлов.

Для исхода ОРОООРОРОР эти случайные величины примут следующие значения:

$$X = 6;\ Y = 4;\ L = 3.$$

Если говорить о каких-то «неслучайных» соотношениях между этими величинами, то можно заметить, что $X + Y = 10$, $L \leqslant X$. Конечно, все эти величины лежат в промежутке от 0 до 10.

Ещё один важный пример случайной величины имеет специальное название.

Случайная величина I называется *индикатором* случайного события A, если она принимает только два значения 0 и 1 и при этом

$$I = \begin{cases} \textbf{1, если событие } A \textbf{ произошло;} \\ \textbf{0, если событие } A \textbf{ не произошло.} \end{cases}$$

По значению индикатора можно узнать, произошло или нет соответствующее событие — отсюда и его название. Индикатор называют также ***бинарной случайной величиной***, поскольку она может принимать всего два значения — 0 и 1.

В опыте с двумя кубиками можно рассмотреть, например, такой индикатор:

$$I = \begin{cases} \text{1, если на кубиках выпали одинаковые числа;} \\ \text{0, если на кубиках выпали разные числа.} \end{cases}$$

Тогда для исхода (5; 5) значением случайной величины I будет 1, а для исхода (6; 2) — 0.

? ВОПРОСЫ

1. Какая величина называется случайной?
2. Приведите примеры случайных величин.
3. Какая случайная величина называется индикатором случайного события?

2 Случайные величины вокруг нас

Со случайными величинами мы постоянно сталкиваемся в окружающей нас жизни. Рассмотрим некоторые из них.

Пример 1. Случайная величина U — напряжение в бытовой сети.

В результате случайного опыта — замера напряжения в сети в разные моменты времени с помощью вольтметра — можно получить ряд случайных значений этой величины в вольтах, например:

218, 218, 220, 216, 221, 225, 222, 219, 223, 227, 218, 223.

Если вы проведёте свои измерения, то у вас получатся другие значения — эта величина действительно случайная.

В отличие от величин из примеров с кубиком или монетой здесь нельзя заранее указать точный набор значений, которые может принимать данная величина. Но можно ограничить их каким-то диапазоном, за который эти значения заведомо не выйдут — например, от 0 до 380 В.

Пример 2. Случайная величина V — вес новорождённого малыша.

При рождении детей вес у всех разный. Он колеблется обычно в диапазоне от 2,5 до 4,5 кг, но может выходить и за эти пределы. Его тоже можно рассматривать как случайную величину.

Пример 3. Случайная величина Z — стоимость литра бензина на автозаправке.

Если вы когда-нибудь отправлялись с семьёй в автомобильное путешествие, то могли заметить, что на разных автозаправках цены на бензин разные. Их общий уровень зависит, конечно, от общего положения в экономике, но при этом на каждой конкретной заправке отклонения от этой цены носят непредсказуемый характер. Поэтому цену на бензин тоже можно рассматривать как случайную величину.

Пример 4. Случайная величина T — продолжительность телефонного разговора.

Ваш сотовый телефон фиксирует длительность всех телефонных разговоров с точностью до секунды. Если вы откроете страницу с этой информацией, вы увидите самые разные значения, которые тоже носят случайный характер.

Как видите, в отличие от примеров с кубиком и монетой случайные величины, которые встречаются в жизни, чаще всего измеряются в каких-то единицах: вольтах, килограммах, рублях, секундах и т. д.

В статистике мы изучали различные методы, которые позволяют нам ***по наблюдениям случайной величины*** делать какие-то выводы о её поведении в будущем. Например, можно составить таблицу частот и выяснить, какие значения в этих наблюдениях встречались чаще, а какие реже. Примерно такую частоту этих значений следует ожидать и в дальнейшем.

Много полезной информации о случайной величине содержат числовые показатели, которые можно найти по наблюдениям: среднее арифметическое, медиана, дисперсия и другие.

В теории вероятностей мы научились вычислять вероятность многих случайных событий без каких-либо статистических данных, основываясь, например, на основе предположений о равновозможности исходов, независимости событий и т. д. Такой же подход можно использовать и для случайных величин.

? ВОПРОСЫ

1. Приведите примеры случайных величин из реальной жизни.
2. Вы подошли к автобусной остановке. Приведите примеры случайных величин, которые вы сможете наблюдать.
3. В вашем классе завтра будет проводиться контрольная работа по вероятности и статистике. Приведите примеры случайных величин, связанных с этим случайным опытом.

3 Дискретные и непрерывные величины

В предыдущих примерах мы уже не раз обсуждали, какой набор значений может принимать та или иная случайная величина. Для одних величин этих значений немного и их легко перечислить, для других они занимают весь числовой промежуток.

В зависимости от того, как устроено это множество возможных значений, случайные величины делятся на несколько типов. Самые важные из них — ***дискретные*** и ***непрерывные*** случайные величины.

Слово «дискретный» происходит от латинского *discretus* — «разделённый», «прерывистый».

Случайная величина называется дискретной, если множество её возможных значений состоит из отдельно отстоящих друг от друга чисел, не сливающихся в интервалы или отрезки.

Чаще всего этим множеством будет какое-то конечное подмножество целых чисел, например $\{1, 2, 3, 4, 5, 6\}$.

Слово «непрерывный» — это антоним к слову «дискретный». Толковый словарь русского языка объясняет его как «длящийся без перерыва, без разделения на части, сплошной линией».

Случайная величина называется непрерывной, если множеством её возможных значений является непрерывный промежуток (или несколько промежутков) числовой прямой.

Иногда этим множеством будет вся числовая прямая.

Посмотрим, какие из рассмотренных ранее величин являются дискретными, а какие — непрерывными.

Пример 1. Подбрасывают два игральных кубика.

Напомним, что мы рассмотрели в этом опыте шесть разных случайных величин:

X — число, выпавшее на первом кубике;
Y — число, выпавшее на втором кубике;
S — сумма чисел;
R — произведение чисел;
W — наибольшее из чисел;
V — наименьшее из чисел.

Найдём для каждой из них множество возможных значений. Очевидно, что величины X, Y могут принимать значения 1, 2, 3, 4, 5, 6. Сумма чисел S может равняться любому целому числу от 2 до 12. Произведение R всегда является целым числом и лежит в диапазоне от 1 до 36. При этом в отличие от суммы произведение может принимать не все значения из этого диапазона: например, произведение 7 получить нельзя. Случайные величины W и V, как и числа на кубиках, могут равняться любому целому числу от 1 до 6.

Таким образом, все шесть рассмотренных в этом примере величин относятся к дискретному типу.

Пример 2. Случайная величина V — вес новорождённого малыша.

Вес — величина непрерывная. Он может принимать любое значение из некоторого промежутка. Определить этот промежуток для данной величины V можно только исходя из опыта. Теоретически можно всегда взять его «с запасом», например от 0 до 5 кг, считая при этом, что вероятность очень малых и очень больших значений мала или вовсе равна нулю.

Пример 3. Случайная величина Z — стоимость литра бензина на автозаправке.

Здесь возникает некоторая сложность. С одной стороны, минимальный шаг изменения цены равен одной копейке — дробного числа копеек в цене на бензин не бывает (хотя в курсах валют или акций, заметим, такая точность используется).

С другой стороны, цены с шагом в одну копейку расположены так плотно, что вполне можно считать этот ряд значений непрерывным. С такой «двойственностью» мы уже сталкивались в статистике, когда решали, нужно ли группировать значения числового ряда или лучше обойтись без группировки. Окончательный выбор типа величины часто зависит от задачи, которую предстоит решить.

Как видите, чётко провести грань между дискретной и непрерывной моделью получается не всегда — особенно для величин из реальной жизни.

В «чистом» виде непрерывные величины встречаются в случайных опытах, где речь идёт о геометрической вероятности.

Пример 4. Внутри круга радиуса 10 случайным образом выбирают точку M. Случайная величина R — расстояние от точки M до центра круга O.

Очевидно, что эта величина непрерывная. Её возможные значения заполняют весь промежуток [0; 10] (рис. 94). При этом отметим, что каждое конкретное значение из этого промежутка величина R принимает с вероятностью 0.

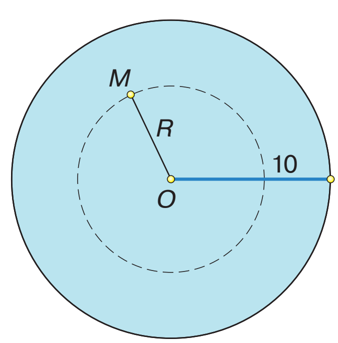


Рис. 94. Случайная величина R – расстояние до центра круга

В самом деле: если взять произвольное число $x \in [0; 10]$, то равенство $R = x$ будет выполняться для точек, лежащих на окружности радиуса x. Вероятность попадания случайной точки на эту окружность равна 0 (вспомните *геометрическое определение вероятности*).

Мы уже сказали в начале этого пункта, что «дискретный» и «непрерывный» в обычном языке — это антонимы. Значит ли это, что рассмотренные два типа исчерпывают все возможные величины? Оказывается, это не так.

Как в теории вероятностей, так и в реальной жизни встречаются величины смешанного типа, в которых присутствует и дискретная, и непрерывная составляющая.

Пример 5. Случайная величина T — продолжительность телефонного разговора.

Время, как известно, непрерывно, поэтому непрерывной должна быть и величина T, которая теоретически может принимать любое значение из промежутка $[0; +\infty)$. Но в этом промежутке есть одно «дискретное» значение, у которого ненулевая вероятность, — это 0.

В самом деле, из своего опыта вы знаете, что часть звонков сбрасывается без установки соединения. При этом звонок фиксируется, но его продолжительность оказывается равной 0.

Аналогичная картина наблюдается и с некоторыми другими случайными величинами, например сроком службы электрической лампочки, временем ожидания зелёного света на светофоре (объясните почему).

Для изучения непрерывных величин требуются некоторые понятия из математического анализа, которые будут изучаться уже в 11 классе. Поэтому в этой главе речь будет идти только о **дискретных случайных величинах**.

? ВОПРОСЫ

1. Чем дискретные величины отличаются от непрерывных?
2. Приведите пример дискретной случайной величины.
3. Приведите пример непрерывной случайной величины.

УПРАЖНЕНИЯ

305. Вспомните, сколько в вашем классе мальчиков и сколько девочек. Какие значения могут принимать следующие случайные величины:
X — количество отсутствующих на уроке мальчиков;
Y — количество отсутствующих на уроке девочек;
Z — количество отсутствующих на уроке учеников.
Какие соотношения между ними выполняются?

306. Вы подошли к автобусной остановке, на которой останавливаются автобусы маршрутов № 13, 34, 119 и 230. Вам нужен автобус № 119. Какие значения могут принимать следующие случайные величины:
а) время ожидания № 119;
б) количество пассажиров в подъехавшем № 119;
в) количество автобусов, подошедших к остановке раньше, чем № 119;
г) температура в салоне № 119.
Какие из этих величин, на ваш взгляд, являются дискретными, а какие — непрерывными?

307. Бросают три игральных кубика. Какие значения могут принимать случайные величины:
S — сумма трёх чисел на кубиках;
K — количество различных чисел на трёх кубиках;
W — наибольшее из трёх чисел на кубиках;
V — наименьшее из трёх чисел на кубиках;
M — среднее из трёх чисел на кубиках (медиана).
Какие соотношения между ними выполняются?

308. Вы выбираете наугад слово из орфографического словаря русского языка. Случайная величина L равна количеству букв в этом слове. Какие значения она может принимать?

309. Вы раскрываете эту книгу наугад. Какие значения может принимать случайная величина X, если она равна:
а) номеру страницы слева n;
б) номеру страницы справа m;
в) остатку от деления n на 10;
г) остатку от деления m на 10;
д) сумме цифр числа n?

310. Проводятся N испытаний Бернулли. Случайная величина X равна числу успехов; случайная величина Y — числу неудач. Какие значения они могут принимать? Как одна из них выражается через другую?

311. Проводятся испытания Бернулли до первого успеха. Случайная величина I равна 1, если первый успех наступит на чётном испытании, 0 — если на нечётном. Какое из этих двух значений более вероятно?

312. В волейболе матч продолжается, пока одна из команд не одержит победы в трёх партиях. Случайная величина X равняется количеству партий в волейбольном матче. Какие значения она может принимать? Ничьих в волейбольных партиях не бывает.

313. Два шахматиста играют финальный матч на звание чемпиона мира до победы в шести партиях. Случайная величина X равняется количеству партий в финальном матче. Какие значения она может принимать? Ничья в шахматной партии вполне возможна.

314*. Бросают четыре кубика. Случайная величина K равна количеству различных чисел, которые при этом выпали. Какие значения она может принимать? Найдите вероятность каждого из этих значений.

315*. На борту воздушного судна, вместимость которого составляет 150 мест, пассажирам предлагается 2 вида обедов: с курицей и с рыбой. Каждый пассажир может с равной вероятностью выбрать курицу или рыбу, поэтому на борт загрузили 75 обедов с курицей и рыбой. Случайная величина X равна количеству недовольных пассажиров. Какие значения она может принимать? Какое из этих значений самое вероятное?

§ 17. Распределение вероятностей

1 Закон распределения вероятностей

Итак, важной характеристикой случайной величины является множество её возможных значений. Для дискретных величин, напомним, оно состоит из множества отдельно отстоящих друг от друга, чаще всего целых, чисел.

Однако даже указание всех возможных значений ещё не даёт полного представления о поведении случайной величины. Вернёмся к примеру с двумя кубиками и рассмотрим уже знакомые нам случайные величины.

Пример 1. Подбрасывают два игральных кубика. Пусть:

X — число, выпавшее на первом кубике;
Y — число, выпавшее на втором кубике;
S — сумма чисел, выпавших на кубиках;
W — наибольшее из чисел, выпавших на кубиках;
V — наименьшее из чисел, выпавших на кубиках.

Четыре из этих пяти перечисленных величин имеют одно и то же множество возможных значений: 1, 2, 3, 4, 5, 6. Но если величины X и Y ведут себя действительно очень похоже (ведь и та, и другая — это число очков, выпавших на игральном кубике), то поведение величин W и V будет совсем другим.

Величины X и Y принимают все свои значения с одинаковой вероятностью $\frac{1}{6}$. Для величины W значение 6 явно вероятнее, чем 1: ведь наибольшее из чисел будет равно шести для одиннадцати исходов этого опыта (перечислите их), а единице — только для одного. Для величины V всё будет с точностью до наоборот.

Таким образом, важно знать не только возможные значения случайной величины, но и вероятности этих значений. Эта информация называется ***распределением вероятностей*** случайной величины или ***законом распределения, распределением*** случайной величины.

Распределением вероятностей **случайной величины называется закон, который описывает все возможные значения случайной величины, а также вероятности, с которыми она их принимает.**

Название «распределение вероятностей» объясняется тем, что этот закон показывает, как вся вероятность, равная 1, распределена между отдельными значениями случайной величины.

Для дискретных величин с небольшим количеством значений самым удобным способом представления такого закона служит таблица, состоящая из двух строк: в первой строке указываются по возрастанию все возможные значения, а во второй — их вероятности.

Если случайная величина X принимает значения $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$, то её закон распределения будет выглядеть так (табл. 58):

Таблица 58

Значения *X*	x_1	x_2	x_3	…	x_n
Вероятность	p_1	p_2	p_3	…	p_n

Вернёмся к примеру 1 и найдём законы распределения рассмотренных в нём случайных величин.

Проще всего будут выглядеть распределения величин X и Y: мы уже заметили, что здесь все значения имеют одинаковую вероятность (табл. 59):

Таблица 59

Значения *X* (*Y*)	1	2	3	4	5	6
Вероятность	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Чтобы найти закон распределения случайной величины S и любой другой величины, связанной с этим опытом, удобно сначала заполнить таблицу, в которой будут перечислены все 36 возможных исходов этого опыта и для каждого исхода указано значение интересующей нас величины. Для суммы S такая таблица будет выглядеть так (табл. 60):

Таблица 60

X \ Y	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Из таблицы 60 хорошо видно, сколько раз встречалось каждое значение суммы: например, сумма 5 встречалась 4 раза, а сумма 7 — 6 раз. Остаётся построить по этим данным закон распределения для случайной величины S (табл. 61):

Таблица 61

Значения *S*	2	3	4	5	6	7	8	9	10	11	12
Вероятность	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Мы специально не стали сокращать дроби в строке вероятностей, так как дроби с одинаковым знаменателем легче сравнивать между собой.

Из таблицы 61 хорошо видны некоторые свойства этого закона: например, его симметрия, а также то, что наиболее вероятное значение суммы

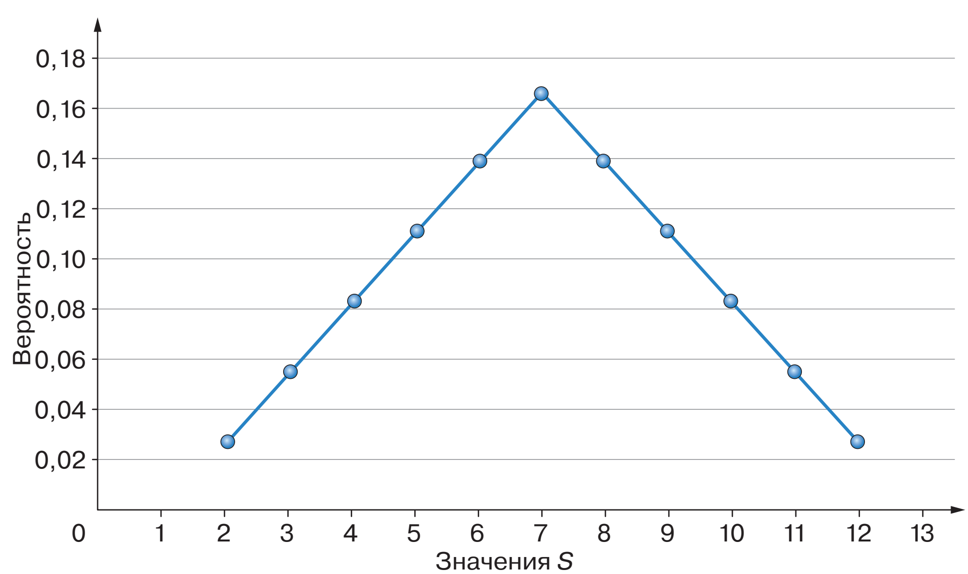


Рис. 95. Полигон распределения для суммы S

равно 7. Ещё лучше эти и некоторые другие свойства можно увидеть, если представить полученный закон распределения графически.

Для этого используют обычно диаграмму, на которой по оси абсцисс откладывают все возможные значения случайной величины, а по оси ординат — их вероятности. На диаграмму наносят точки с координатами $(x_1; p_1)$, $(x_2; p_2)$, ... , $(x_n; p_n)$, а затем соединяют их ломаной линией. Полученную диаграмму называют ***полигоном распределения*** или ***диаграммой распределения вероятностей***.

Диаграмма распределения для суммы S построена на рисунке 95.

Для случайной величины W, равной наибольшему из двух чисел на кубиках, можно воспользоваться тем же приёмом, что и для суммы. Сначала заполним вспомогательную таблицу, в которую занесём значения величины W для каждого исхода опыта (табл. 62).

Таблица 62

X \ Y	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

Теперь построим по этим данным закон распределения для случайной величины W (табл. 63).

Таблица 63

Значения *W*	1	2	3	4	5	6
Вероятность	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Полигон распределения для случайной величины W изображён на рисунке 96.

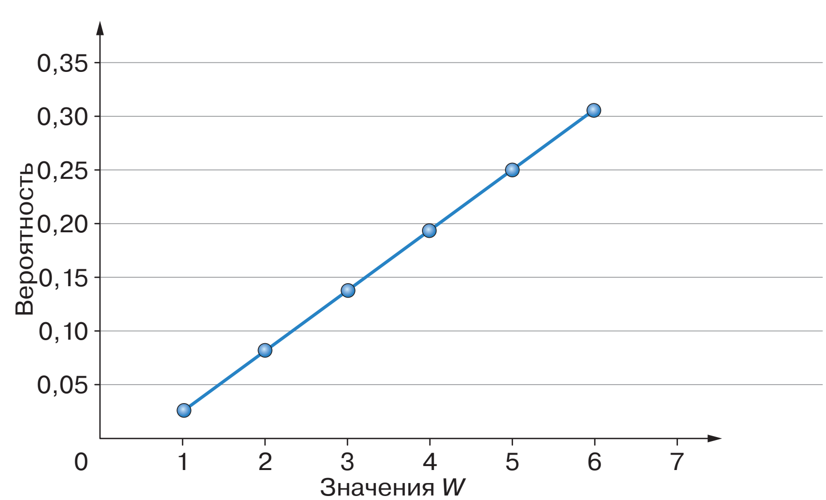


Рис. 96. Полигон распределения для наибольшего из чисел W

На рисунке 96 хорошо видно, что распределение вероятностей асимметрично: более вероятными являются бо́льшие значения W.

Пример 2. Найдём закон распределения бинарной случайной величины I, которая является индикатором некоторого случайного события A. Напомним, что она равна 1, если событие A происходит, и равна 0, если не происходит. Если $P(A) = p$, то закон распределения I выглядит так (табл. 64):

Таблица 64

Значения *I*	0	1
Вероятность	$1 - p$	p

Как видите, законы распределения дискретных случайных величин могут быть самые разные. Но в любом из них всегда выполняется важное свойство: **сумма всех вероятностей в любом законе распределения всегда равна 1**:

$$p_1 + p_2 + \ldots + p_n = 1.$$

Если оно не выполнено, значит, либо в таблицу попали не все возможные значения случайной величины, либо произошла ошибка при вычислении какой-то из вероятностей.

Можете убедиться, что во всех найденных нами законах это свойство выполнено.

В каких-то задачах проверка этого свойства позволит вам найти ошибку в найденном распределении, а в каких-то — избавит от вычисления последней (иногда самой трудной) вероятности. Её можно будет найти, вычитая сумму всех остальных вероятностей из 1.

? ВОПРОСЫ

1. Что такое закон распределения вероятностей? В какой форме он задаётся для дискретных случайных величин?
2. Какое основное свойство для дискретного закона распределения вы знаете?
3. Могут ли все значения случайной величины иметь вероятности 0,3? 0,2? 0,25?

2 Биномиальное распределение

Вернёмся к примеру, в котором речь шла об испытаниях с монетой и количестве выпавших при этом орлов.

Пример 1. Монету подбрасывают 10 раз подряд. Случайная величина X равна числу испытаний, в которых на монете выпал орёл.

Найдём закон распределения этой величины. Очевидно, что возможные значения X — это целые числа от 0 до 10. Чтобы найти вероятности этих значений, рассмотрим наши десять опытов как серию из 10 испытаний Бернулли. Если считать выпадение орла успехом, то X — это число успехов в 10 испытаниях Бернулли. Получается, что вероятность события $\{X = k\}$ для любого значения k от 0 до 10 можно посчитать по формуле Бернулли:

$$P(X = k) = C_{10}^{k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{10-k} = \frac{C_{10}^{k}}{2^{10}}.$$

Остаётся подставить в эту формулу значения $k = 0, 1, \ldots, 10$ и записать полученные значения в таблицу 65.

Таблица 65

Значения *X*	0	1	2	3	4	5	6	7	8	9	10
Вероятность	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$	$\frac{210}{1024}$	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$

Полученное распределение будет симметричным, а максимальная вероятность будет достигаться для значения $X = 5$. Все эти свойства хорошо видны на графике (рис. 97).

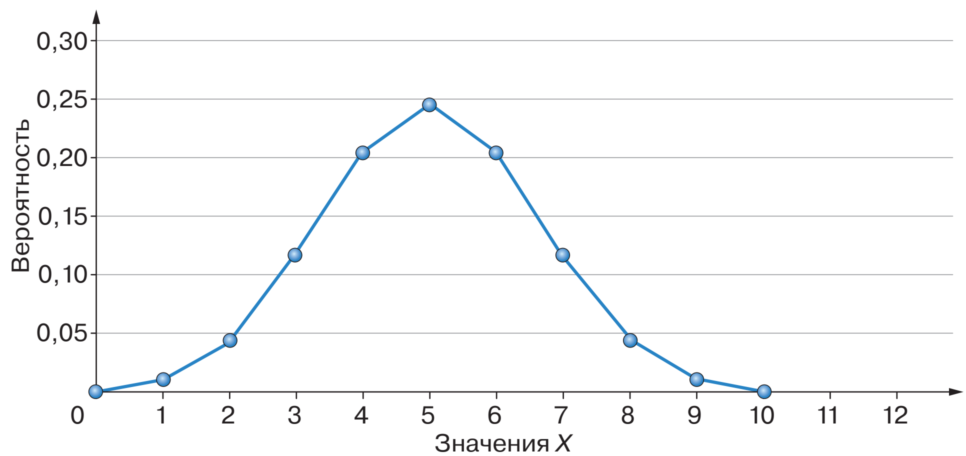


Рис. 97. Распределение числа орлов при десятикратном бросании монеты

Теперь рассмотрим пример, в котором бросают не монету, а кубик.

Пример 2. Кубик подбрасывают 8 раз подряд. Случайная величина Y равна числу испытаний, в которых на кубике выпала шестёрка.

Найдём закон распределения величины Y. Её возможными значениями будут целые числа от 0 до 8. Распределение вероятностей между этими значениями, конечно, будет уже не симметричным. Вряд ли можно ожидать, что наиболее вероятным количеством шестёрок будет 4.

Тем не менее нам снова поможет формула Бернулли. Если считать выпадение шестёрки успехом, то Y — это число успехов в 8 испытаниях Бернулли. Вероятность успеха в каждом испытании равна $\frac{1}{6}$. Применяя формулу Бернулли, получаем, что вероятность события $\{Y = k\}$ будет равна

$$P(Y = k) = C_8^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{8-k}$$

Вычислим эти вероятности для $k = 0, 1, \ldots, 8$ с помощью электронной таблицы. В главе 5 «Испытания Бернулли. Случайный выбор» мы уже пользовались для подобных расчётов функцией БИНОМРАСП(). Применим её и для нашего примера (табл. 66).

Таблица 66

Значения *Y*	0	1	2	3	4	5	6	7	8
Вероятность	0,233	0,372	0,261	0,104	0,026	0,004	0,0004	$2{,}4 \cdot 10^{-05}$	$5{,}95 \cdot 10^{-07}$

Вероятности двух последних значений получились такими маленькими, что пришлось записывать их в так называемом ***научном формате***, с умножением на десять в отрицательной степени.

Наибольшая вероятность достигается здесь для значения $Y = 1$. Это соответствует теореме о наиболее вероятном числе успехов, которую мы приводили в той же главе об испытаниях Бернулли. Диаграмма распределения построена на рисунке 98.

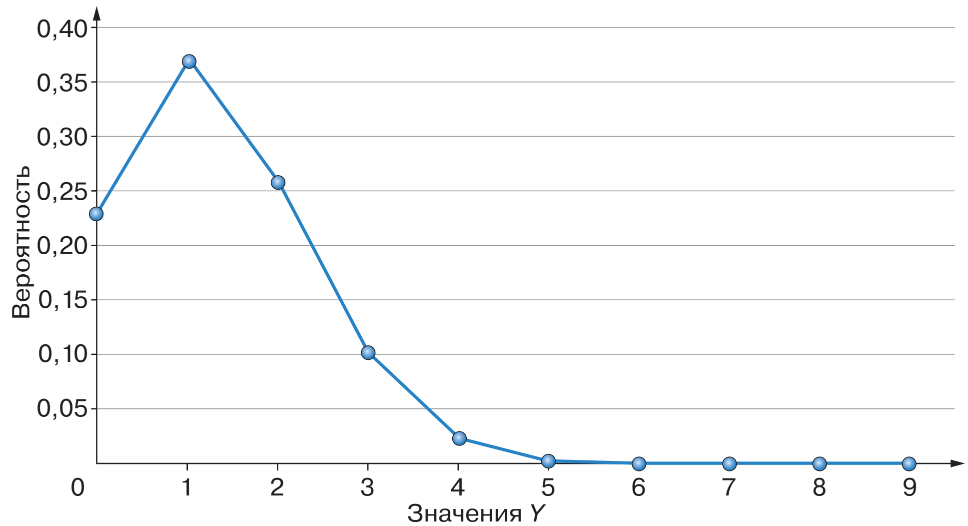


Рис. 98. Распределение числа шестёрок при восьмикратном бросании кубика

Распределение вероятностей, полученное в примерах 1 и 2, называется ***биномиальным***. Такое распределение имеет любая случайная величина, равная числу успехов в серии испытаний Бернулли. Название «биномиальное» объясняется тем, что вероятности $C_N^k p^k q^{N-k}$ есть не что иное, как слагаемые в разложении бинома Ньютона:

$$1 = (p + q)^N = C_N^0 p^0 q^N + C_N^1 p^1 q^{N-1} + C_N^2 p^2 q^{N-2} + \ldots + C_N^{N-1} p^{N-1} q^1 + C_N^N p^N q^0.$$

Если вы забыли эту замечательную формулу, советуем вернуться к пункту «Бином Ньютона», где она была доказана.

Теперь понятно, почему функция электронной таблицы для вычисления вероятностей по формуле Бернулли называется БИНОМРАСП().

На самом деле биномиальный закон — это целое семейство распределений с двумя параметрами N и p, где N — количество испытаний, p — вероятность успеха в одном испытании. Биномиальный закон распределения имеет специальное обозначение: $\mathbf{Bin}(N, p)$.

Случайная величина X из первого примера имела распределение $\mathbf{Bin}(10;\ 0{,}5)$, а случайная величина Y из второго примера — $\mathbf{Bin}\left(8;\ \frac{1}{6}\right)$.

Биномиальное распределение — одно из важнейших в теории вероятностей и очень часто встречается при решении самых разных задач.

? ВОПРОСЫ

1. Что такое биномиальный закон распределения? Почему он так называется?
2. От каких параметров зависит биномиальный закон распределения?
3. Сколько разных значений имеет случайная величина с распределением $\mathbf{Bin}(15;\ 0{,}3)$?
4. Случайная величина X имеет распределение $\mathbf{Bin}\left(5;\ \frac{1}{3}\right)$. С какой вероятностью она равна 1?

3 Геометрическое распределение

С испытаниями Бернулли связано ещё одно распределение вероятностей, которое называется геометрическим.

Рассмотрим серию испытаний Бернулли до первого успеха. Пусть X — это число опытов, которое потребовалось при этом провести. Тогда X — **дискретная случайная величина с бесконечным множеством возможных значений**.

В самом деле, возможными значениями этой величины будут все натуральные числа от 1 до бесконечности, поскольку мы не можем установить верхнюю границу для числа испытаний, при достижении которой обязательно наступит первый успех.

Чтобы вычислить вероятности всех этих значений, вернёмся на минутку к параграфу 12 «Испытания до первого успеха». Там мы вывели формулу для вероятности получения первого успеха на k-м шаге:

$$P(k) = q^{k-1}p.$$

Она и даёт нам искомое распределение вероятностей (табл. 67).

Таблица 67

Значения X	1	2	3	4	5	...
Вероятность	p	qp	q^2p	q^3p	q^4p	...

Таблица 67 продолжается вправо до бесконечности. Тем не менее основное свойство закона распределения всё равно сохраняется: сумма всех вероятностей равна 1. Мы уже доказали это, когда изучали испытания Бернулли:

$$p + qp + q^2p + q^3p + q^4p + ... = \frac{p}{1-q} = \frac{p}{p} = 1.$$

Напомним, что это сумма бесконечно убывающей геометрической прогрессии со знаменателем q. Отсюда и название этого распределения — ***геометрическое***.

Геометрическое распределение, как и биномиальное, представляет собой целое семейство распределений только не с двумя, а с одним параметром p, равным вероятности успеха в одном испытании Бернулли. Геометрическое распределение с параметром p обозначается $\mathbf{Geom}(p)$.

Рассмотрим несколько примеров геометрических распределений.

Пример 1. Монету подбрасывают до появления первого орла. Случайная величина X, равная числу проведённых опытов, имеет распределение $\mathbf{Geom}\left(\frac{1}{2}\right)$.

Выпишем несколько первых значений с их вероятностями (табл. 68):

Таблица 68

Значения X	1	2	3	4	5	...
Вероятность	0,5	0,25	0,125	0,0625	0,03125	...

Пример 2. Кубик подбрасывают до появления первой шестёрки. Случайная величина Y, равная числу проведённых опытов, имеет распределение $\mathbf{Geom}\left(\frac{1}{6}\right)$.

Вот начало таблицы для этого распределения (табл. 69):

Таблица 69

Значения *Y*	1	2	3	4	5	...
Вероятность	0,167	0,0833	0,0417	0,0208	0,0104	...

Пример 3. Студент сдаёт зачёт по теории вероятностей до тех пор, пока не решит какую-нибудь задачу. Вероятность решения любой задачи этим студентом равна 0,2. Случайная величина Z равна числу задач, которые он получит, пока, наконец, сдаст зачёт. Она имеет распределение **Geom**(0,2).

Вот начало таблицы для этого распределения (табл. 70):

Таблица 70

Значения *Y*	1	2	3	4	5	...
Вероятность	0,2	0,1	0,05	0,025	0,0125	...

На рисунке 99 показаны полигоны всех трёх распределений из примеров 1, 2 и 3. Видно, что все они имеют максимальное значение вероятности в нуле и быстро убывают. Несложно понять, что этим свойством обладают все геометрические распределения, поскольку вероятности в них образуют бесконечно убывающую геометрическую прогрессию со знаменателем $q < 1$.

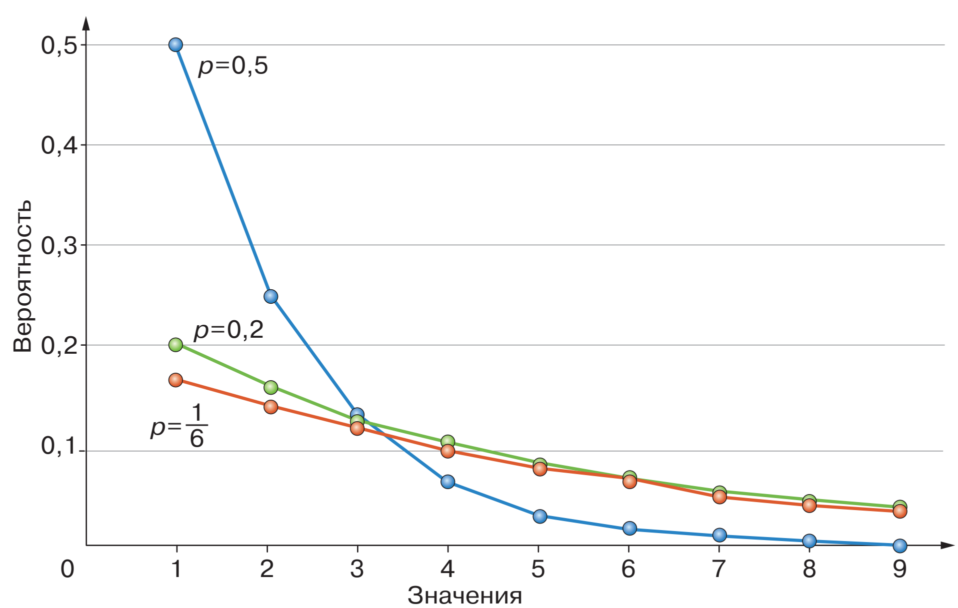


Рис. 99. Геометрические законы распределения для разных значений p

Конечно, биномиальное и геометрическое распределения не исчерпывают всего разнообразия законов распределения случайных величин. Некоторые из них вы встретите в упражнениях, с какими-то познакомитесь позже, изучая курс вероятности и статистики в 11 классе.

ВОПРОСЫ

1. Что такое геометрический закон распределения? Почему он так называется?
2. От каких параметров зависит геометрический закон распределения?
3. Сколько разных значений имеет случайная величина с распределением $\mathbf{Geom}(0{,}3)$?
4. Случайная величина X имеет распределение $\mathbf{Geom}\left(\frac{1}{3}\right)$. С какой вероятностью она равна 1? равна 2?

УПРАЖНЕНИЯ

316. Распределение вероятностей дискретной случайной величины X задано таблицами 71 и 72:

а)

Таблица 71

Значения X	1	5	10	11	12
Вероятность	0,2	0,12	0,1	p	0,4

б)

Таблица 72

Значения X	–2	–1	0	1	2
Вероятность	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	p	$\frac{1}{16}$

Найдите неизвестное значение вероятности p и постройте полигон распределения вероятностей.

317. Распределение вероятностей дискретной случайной величины X задано таблицей 73:

Таблица 73

Значения X	0	3	5	10	50
Вероятность	0,1	0,3	0,4	0,15	0,05

Найдите вероятности случайных событий:

а) $X < 5$; б) $X \geqslant 3$; в) $4 < X < 8$; г) $X < 0$; д) $X > 0$.

318. Закон распределения дискретной случайной величины X задан таблицей 74:

Таблица 74

Значения X	–10	–1	0	1	10
Вероятность	0,3	0,15	0,1	0,15	0,3

Найдите законы распределения случайных величин $Y = |X|$ и $Z = X^2$.

319. В партии 10% нестандартных деталей. Из неё наугад отбирают 4 детали. Случайная величина X равна числу нестандартных деталей среди отобранных. Найдите закон распределения случайной величины X. Какое число нестандартных деталей в этой выборке наиболее вероятно?

320. Профессор Злобин не хочет ставить положительную оценку студенту Разгуляеву и задаёт ему вопросы до тех пор, пока на очередном вопросе студент не ошибётся. Вероятность правильного ответа на любой вопрос составляет 0,8. Случайная величина X равна числу вопросов, на которые правильно ответит студент перед тем, как получит двойку. Найдите закон распределения случайной величины X.

321*. Биатлонист должен поразить 3 мишени пятью выстрелами. На каждый выстрел он тратит 8 с и попадает в цель с вероятностью 0,6. Случайная величина T — общее время, которое он проведёт на огневом рубеже. Найдите распределение вероятностей T.

322*. В ящике лежат 5 белых и 5 чёрных шаров. Из него наугад вытаскивают 4 шара. Случайная величина X равна числу белых шаров в полученной выборке. Найдите закон распределения X.

323*. Три джентльмена пришли в ресторан в одинаковых шляпах и сдали их в гардероб. Уходя домой, они разобрали шляпы наугад. Случайная величина X равна количеству джентльменов, ушедших домой в своих шляпах. Найдите её распределение вероятностей.

324*. Решите предыдущую задачу для четырёх джентльменов.

325*. Пусть N — выбранное наугад двузначное случайное число, а случайная величина X равна сумме его цифр. Найдите распределение вероятностей случайной величины X.

326*. Решите предыдущую задачу, если X — это наибольшая из двух цифр.

327*. В волейболе матч продолжается, пока одна из команд не одержит победы в трёх партиях. Случайная величина X равняется количеству партий в волейбольном матче между командами «Чайка» и «Буревестник», шансы которых на победу одинаковые. Ничьих в волейбольной партии не бывает. Найдите закон распределения случайной величины X.

328*. Шахматисты А и Б играют финальный матч за звание чемпиона до победы в двух партиях. Случайная величина X равняется количеству партий в финальном матче. Найдите закон её распределения, если: любая партия может закончиться вничью с вероятностью 0,5, а А и Б имеют равные шансы на выигрыш.

329*. Бросают три кубика. Случайная величина K равна количеству различных чисел, которые при этом выпали. Найдите закон распределения K.

330*. Подбрасывают одновременно n кубиков. Случайная величина W равна наибольшему из выпавших чисел. Найдите закон распределения W. Нарисуйте диаграммы распределения вероятностей для $n = 2, 3, 5$ и 10.

ЛАБОРАТОРНАЯ РАБОТА № 4

Закон распределения вероятностей

Задание 1. Биномиальное распределение

Постройте с помощью электронной таблицы биномиальное распределение вероятностей **Bin**(N, p) с $N = 20$ и разными вероятностями успеха $p = 0{,}1; 0{,}2; 0{,}5; 0{,}8; 0{,}9$ (рис. 100).
Постройте на одной диаграмме полученные распределения вероятности.
Изучите, как ведёт себя биномиальное распределение при изменении вероятности успеха p.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	0	0,121577	0,011529	9,54E-05	1,05E-14	1E-20		0,1	0,2	0,5	0,8	0,9			
2	1	0,27017	0,057646	1,91E-05	8,39E-13	1,8E-18									
3	2	0,28518	0,136909	0,000181	3,19E-11	1,54E-16									
4	3	0,19012	0,205364	0,001087	7,65E-10	8,31E-15									
5	4	0,089779	0,218199	0,004621	1,3E-08	3,18E-13									
6	5	0,031921	0,17456	0,01486	1,66E-07	9,15E-12									
7	6	0,008867	0,1091	0,036964	1,66E-06	2,06E-10									
8	7	0,00197	0,05455	0,073929	1,33E-05	3,71E-09									
9	8	0,000356	0,022161	0,120134	8,66E-05	5,42E-08									
10	9	5,27E-06	0,007387	0,160179	0,000462	6,51E-07									
11	10	6,44E-06	0,002031	0,176197	0,002031	6,44E-06									
12	11	6,51E-07	0,000462	0,160179	0,007387	5,27E-05									
13	12	5,42E-08	8,66E-05	0,120134	0,022161	0,000356									
14	13	3,71E-09	1,33E-05	0,073929	0,05455	0,00197									
15	14	2,06E-10	1,66E-06	0,036964	0,1091	0,008867									
16	15	9,15E-12	1,66E-07	0,014786	0,17456	0,031921									
17	16	3,18E-13	1,3E-08	0,04621	0,218199	0,089779									
18	17	8,31E-15	7,65E-10	0,001087	0,205364	0,19012									
19	18	1,54E-16	3,19E-11	0,000181	0,136909	0,28518									
20	19	1,8E-18	8,39E-13	1,91E-05	0,057646	0,27017									
21	20	1E-20	1,05E-14	9,94E-07	0,011529	0,121577									

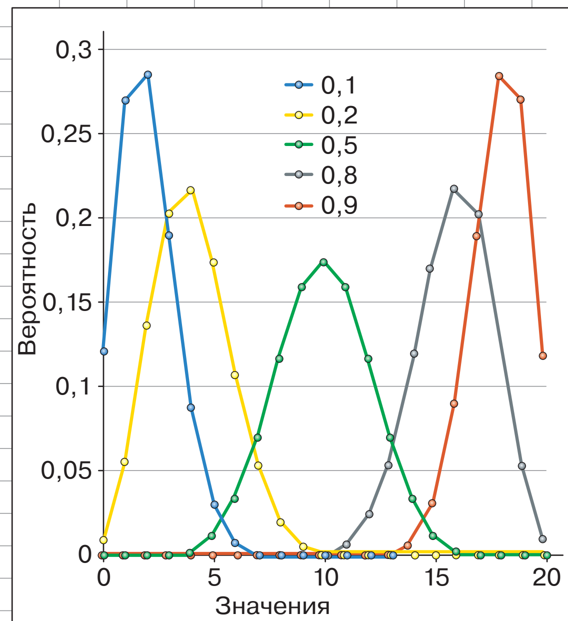


Рис. 100. Пример выполнения задания 1

Задание 2. Геометрическое распределение

Постройте с помощью электронной таблицы геометрическое распределение вероятностей **Geom**(p) с разными вероятностями успеха p = 0,1; 0,2; 0,3; 0,4; 0,5 (рис. 101). Постройте на одной диаграмме полученные распределения вероятности. Изучите, как ведёт себя геометрическое распределение при изменении вероятности успеха p.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	1	0,1	0,2	0,3	0,4	0,5		0,1	0,2	0,3	0,4	0,5			
2	2	0,09	0,16	0,21	0,24	0,25									
3	3	0,081	0,128	0,147	0,144	0,125									
4	4	0,0729	0,1024	0,1029	0,0864	0,0625									
5	5	0,06561	0,08192	0,07203	0,05184	0,03125									
6	6	0,059019	0,065536	0,050421	0,031104	0,015625									
7	7	0,053144	0,052429	0,035295	0,018662	0,007813									
8	8	0,04783	0,041943	0,024706	0,011197	0,003906									
9	9	0,043047	0,033554	0,017294	0,006718	0,001953									
10	10	0,038742	0,026844	0,012106	0,004031	0,000977									
11	11	0,034868	0,021475	0,008474	0,002419	0,000488									
12	12	0,031381	0,01718	0,005932	0,001451	0,000244									
13	13	0,028243	0,013744	0,004152	0,000871	0,000122									
14	14	0,025419	0,010995	0,002907	0,000522	6,1E-05									
15	15	0,022877	0,008796	0,002035	0,000313	3,05E-05									
16	16	0,020589	0,007037	0,001424	0,000188	1,53E-05									
17	17	0,01853	0,005629	0,000997	0,000113	7,63E-06									
18	18	0,16677	0,004504	0,000698	6,77E-05	3,81E-06									
19	19	0,015009	0,003603	0,000489	4,06E-05	1,91E-06									
20	20	0,013509	0,002882	0,000342	2,44E-05	9,54E-07									
21															

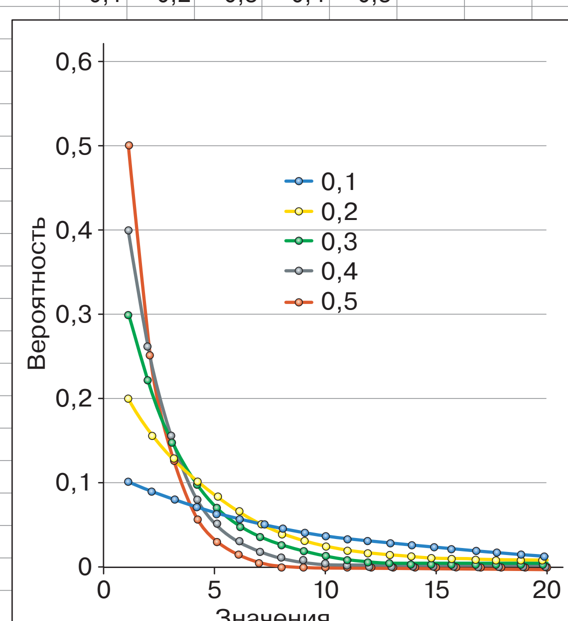


Рис. 101. Пример выполнения задания 2

Задание 3. Лотерея «Спортлото»

В проводившейся в СССР с 70-х годов прошлого века лотерее «Спортлото 6 из 49» нужно было угадать, какие 6 из 49 номеров выиграют в очередном тираже. Для этого участник зачёркивал в карточке этой лотереи любые 6 номеров из 49. Если было угадано 3 и больше номеров, то участника ожидал денежный приз.
Найдите с помощью электронной таблицы закон распределения случайной величины X, равной числу угаданных номеров. Вычислите по этим данным вероятность получения денежного приза.

	A	B	C	D	E
1	0	6096454	0,435965	n=	13983816
2	1	5775588	0,4130195	p=	0,018637
3	2	1851150	0,132378		
4	3	246820	0,0176504		
5	4	13545	0,0009686		
6	5	258	1,845E-05		
7	6	1	7,151E-08		

Рис. 102. Фрагмент задания 3

Этапы выполнения задания:
1. Заполните столбец A возможными значениями случайной величины X.
2. Вычислите в ячейке E1 количество всех возможных исходов лотереи «Спортлото 6 из 49».
3. Найдите в столбце B число благоприятных исходов для каждого из значений X от 0 до 6.
4. Найдите в столбце C вероятности этих значений.
5. Вычислите в ячейке E2 вероятность события A = «будет получен денежный приз».
6. Постройте диаграмму полученного распределения вероятностей.

§ 18. Математическое ожидание

Закон распределения содержит наиболее полную информацию о случайной величине.

Однако во многих ситуациях эту информацию требуется сократить. Так уже было в статистике, когда мы использовали для описания полученных статистических данных среднее арифметическое, медиану, дисперсию и другие числовые характеристики.

Теперь мы познакомимся с аналогичными характеристиками для случайных величин.

1 Что такое математическое ожидание?

Вы уже не раз писали Всероссийские проверочные работы и знаете, что их результаты бывают достаточно непредсказуемы. Рассмотрим случайную величину X, равную оценке учеников на ВПР по математике. Предположим, что на основе проведённых статистических наблюдений для этой величины было получено следующее распределение вероятностей (табл. 75):

Таблица 75

Значения *X*	2	3	4	5
Вероятность	0,1	0,4	0,3	0,2

Какой средний балл следует ожидать на очередной ВПР?

Для ответа на этот вопрос вспомним, что частота любого случайного события при большом числе опытов приближается к его вероятности. Значит, если проверочную работу будут писать, скажем, 1000 учеников, то примерно 100 из них получат отметку 2, 400 учеников — отметку 3, 300 учеников — отметку 4 и 200 учеников — отметку 5. Поэтому средний балл, посчитанный по этим ученикам, составит:

$$\frac{2 \cdot 100 + 3 \cdot 400 + 4 \cdot 300 + 5 \cdot 200}{1000} = 2 \cdot 0{,}1 + 3 \cdot 0{,}4 + 4 \cdot 0{,}3 + 5 \cdot 0{,}2 = 3{,}6.$$

Получается, что для прогнозирования среднего балла нам достаточно знать только закон распределения случайной величины: нужно умножить каждое значение случайной величины на его вероятность и затем сложить эти произведения.

Полученное выражение называется ***математическим ожиданием*** случайной величины и является одной из важнейших его характеристик.

Математическим ожиданием дискретной случайной величины X с законом распределения, указанным в таблице 76, называется число, которое обозначается $E(X)$ и вычисляется по формуле

$$E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \ldots + x_np_n.$$

Таблица 76

Значения X	x_1	x_2	x_3	…	x_n
Вероятность	p_1	p_2	p_3	…	p_n

Термин «математическое ожидание» объясняется тем, что **в среднем именно такое значение X мы ожидаем получить**, если будем проводить испытания с этой величиной. Обозначение $E(X)$ происходит от первой буквы английского слова *expectation* — «ожидание». Математическое ожидание называют также *средним значением* случайной величины X.

В электронной таблице математическое ожидание удобно вычислять с помощью функции

СУММПРОИЗВ(Диапазон 1; Диапазон 2),

где в качестве первого диапазона нужно указать диапазон значений x_1, x_2, … , x_n, а в качестве второго — диапазон вероятностей p_1, p_2, … , p_n.

Рассмотрим несколько примеров на вычисление математического ожидания.

Пример 1. Подбрасывают игральный кубик. Случайная величина X — число очков, выпавших на кубике.

Найдём математическое ожидание $E(X)$. Поскольку все значения случайной величины X имеют вероятность $\frac{1}{6}$, то

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3{,}5.$$

Таким образом, можно сказать, что **на кубике в среднем выпадает 3,5 очка**. Возможно, поначалу это звучит странно, но имеет очень простое объяснение: если долго бросать кубик, а потом вычислить среднее арифметическое всех выпавших результатов, то получится число, близкое к 3,5.

Пример 2. Подбрасывают два кубика. Случайная величина S — сумма очков, выпавших на кубиках.

Найдём математическое ожидание $E(S)$. Вспомним её закон распределения (табл. 77):

Таблица 77

Значения S	2	3	4	5	6	7	8	9	10	11	12
Вероятность	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Отсюда

$$E(S) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \ldots + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = \frac{252}{36} = 7.$$

Мы получили ответ 7, который можно было угадать и без вычислений — ведь распределение вероятностей симметрично относительно значения 7. К этому свойству мы ещё вернёмся в следующем пункте и обоснуем использование симметрии для вычисления математического ожидания.

Пример 3. Рассмотрим бинарную случайную величину I, которая является индикатором случайного события A. Если $P(A) = p$, то закон распределения I выглядит так (табл. 78):

Таблица 78

Значения I	0	1
Вероятность	$1 - p$	p

Её математическое ожидание будет равно вероятности события A:

$$E(I) = 0 \cdot (1 - p) + 1 \cdot p = p.$$

Пример 4. Представьте, что вы покупаете лотерейный билет стоимостью 50 р. При этом условия лотереи таковы, что вы можете выиграть призы от тысячи до миллиона рублей. Как рассчитать свои шансы и стоит ли участвовать в такой лотерее?

Чтобы ответить на этот вопрос, нужно знать закон распределения для случайной величины X, равной выигрышу на один билет. Иногда эту информацию можно получить из условий проведения лотереи (мы ещё рассмотрим такие лотереи позже).

Предположим, что закон распределения X нам известен и выглядит так (табл. 79):

Таблица 79

Значения X	0	1000	10 000	1 000 000
Вероятность	0,98899	0,01	0,001	0,00001

Значения X даны в рублях. Вроде бы условия лотереи неплохие: выигрыш в 1 000 000 р. получает в среднем один человек на 100 000. Если лотерея массовая, то в большом городе один, а то и несколько его жителей в каждом тираже могут выиграть по миллиону. Но не будем торопиться с выводами.

Посчитаем математическое ожидание выигрыша $E(X)$:

$$E(X) = 0 \cdot 0{,}98889 + 1000 \cdot 0{,}01 + 10\,000 \cdot 0{,}001 + 1\,000\,000 \cdot 0{,}00001 = 30.$$

Получилось, что средний выигрыш на один билет составляет 30 р. при цене билета 50 р. Это значит, что организаторы лотереи получают в среднем по 20 р. с каждого билета. А вы в среднем эти 20 р. проигрываете.

Наверняка у вас готово возражение: но ведь кто-то же выигрывает по 1 000 р., по 10 000 р. и даже по 1 000 000 р.! Да, конечно — поэтому можно сыграть в лотерею один-два раза. Убытки будут небольшие, а шанс выиграть миллион (правда, очень мизерный) всё-таки будет. Но если вы начнёте играть регулярно, то начнёт действовать ***закон больших чисел***, о котором мы поговорим подробнее уже в 11 классе, и ваш средний выигрыш с учётом цены билета начнёт приближаться к (–20) р. Умножьте его на количество купленных билетов и получите ожидаемый убыток...

Учтите, что так устроены все лотереи — средний выигрыш всегда меньше стоимости билета. Иначе лотерея была бы убыточной для её организаторов. Разница в 20 р., которая у нас получилась, ещё не самая плохая. В любом случае, покупая билет лотереи, сначала разберитесь в её правилах и, если это возможно, попробуйте оценить математическое ожидание своего выигрыша.

Мы уже сталкивались с дискретными величинами, которые имеют бесконечное количество возможных значений, например с геометрическим распределением вероятностей.

В этом случае выражение для математического ожидания содержит бесконечное число слагаемых и может как существовать, так и не существовать. Например, математическое ожидание случайной величины, имеющей геометрическое распределение с параметром $p = 0{,}5$ существует и равно 2 (мы докажем это чуть позже). Для закона распределения (табл. 80) математического ожидания не существует, поскольку соответствующая сумма будет бесконечно большой:

$$E(X) = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + \ldots = 1 + 1 + 1 + \ldots .$$

Таблица 80

Значения X	2	4	8	16	32	...
Вероятность	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	...

Так что величины, которые не имеют математического ожидания, тоже встречаются и даже используются при решении некоторых задач, но это скорее исключение, чем правило.

? ВОПРОСЫ

1. Дайте определение математического ожидания дискретной случайной величины.
2. Как обозначается математическое ожидание случайной величины X?
3. Случайная величина X принимает значения –1 и 1 с вероятностями 0,25 и 0,75. Чему равно её математическое ожидание?
4. Приведите пример закона распределения, для которого математическое ожидание равно 1.

2 Физический смысл математического ожидания*

При вычислении математического ожидания суммы чисел, которая выпадает на двух кубиках, мы получили ответ 7. Это то значение, относительно которого распределение вероятности симметрично. Почему так получилось и будет ли это верным для других симметричных распределений?

Чтобы ответить на этот вопрос, давайте вспомним физику. Представим, что на числовой прямой в точках $x_1, x_2, \ldots, x_n$ расположили грузики с массами $m_1, m_2, \ldots, m_n$ (рис. 103). Где поставить «опору», чтобы вся прямая находилась в равновесии?

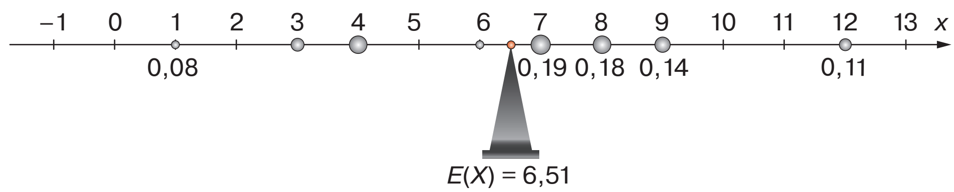


Рис. 103. Центр масс

Из курса физики вы знаете, что такая точка называется ***центром масс*** этой системы. Конечно, её положение можно найти экспериментально, и вы наверняка проделывали подобные опыты. Но оказывается, её координату на числовой оси можно вычислить по следующей формуле:

$$x = \frac{x_1 \cdot m_1 + x_2 \cdot m_2 + \ldots + x_n \cdot m_n}{m_1 + m_2 + \ldots + m_n}.$$

Числитель этой дроби очень похож на формулу для математического ожидания, мешает только знаменатель. Но представим, что суммарная масса всех грузиков равна 1. Тогда знаменатель превращается в 1, а вся формула для координаты центра масс превращается в формулу математического ожидания.

Таким образом, **математическое ожидание можно рассматривать как центр масс системы материальных точек с координатами x_1, x_2, ..., x_n и массами $p_1, p_2, \ldots, p_n$.**

Если вся система материальных точек (с учётом их координат и масс) симметрична относительно некоторой точки x_0, то эта точка и будет её центром масс, а значит, число x_0 будет математическим ожиданием соответствующего распределения вероятностей.

Теперь понятно, почему математическое ожидание случайной величины S оказалось равно 7.

Конечно, это свойство можно доказать и чисто математически. В самом деле, пусть для закона распределения случайной величины X нашлась такая точка x_0, что все значения $x_1, \ldots, x_n$ можно поделить на две симметричные относительно точки x_0 половины. Найдём разность математического ожидания и x_0:

$$E(X) - x_0 = x_1 \cdot p_1 + \ldots + x_n \cdot p_n - x_0 =$$
$$= x_1 \cdot p_1 + \ldots + x_n \cdot p_n - x_0(p_1 + \ldots + p_n) =$$
$$= (x_1 - x_0)p_1 + \ldots + (x_n - x_0)p_n.$$

Но если распределение симметрично относительно точки x_0, то

$$x_1 - x_0 = -(x_n - x_0),\ p_1 = p_n,$$
$$x_2 - x_0 = -(x_{n-1} - x_0),\ p_2 = p_{n-1} \text{ и т. д.}$$

Все слагаемые сократятся, а их сумма получится равной 0, т. е. $E(X) = x_0$.

Не забудьте про это свойство при вычислении математического ожидания симметричных распределений.

Пример. Нужно найти математическое ожидание случайной величины X, имеющей следующее распределение вероятностей (табл. 81):

Таблица 81

Значения X	1	2	4	6	8	9
Вероятность	0,012	0,152	0,336	0,336	0,152	0,012

Из таблицы 81 и рисунка 104 хорошо видно, что распределение симметрично относительно точки 5, поэтому $E(X) = 5$. Обратите внимание, что сама эта точка не является одним из значений случайной величины — это требование вовсе не обязательно.

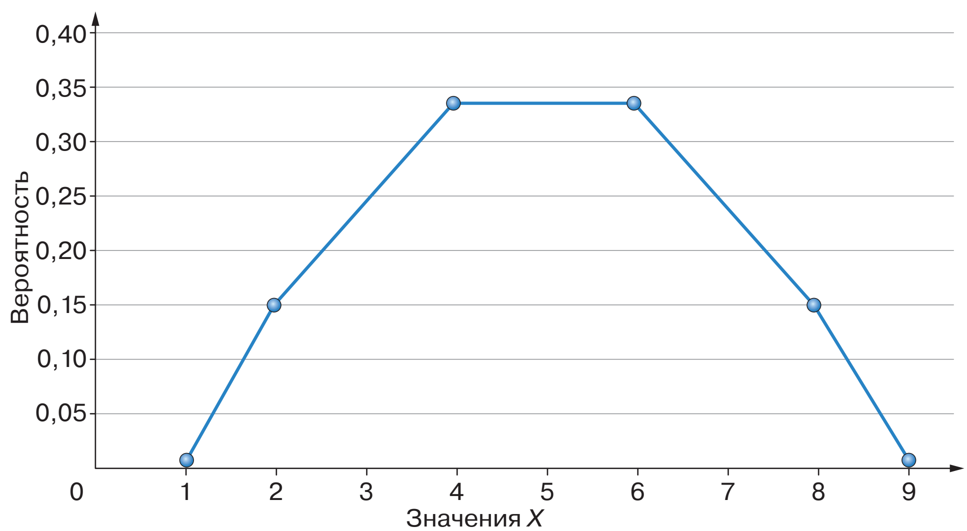


Рис. 104. Симметричное распределение вероятностей

? ВОПРОСЫ

1. Объясните физический смысл математического ожидания.
2. Как связаны среднее арифметическое и математическое ожидание?
3. Какой закон распределения называется симметричным? Чему равно его математическое ожидание?

3 Свойства математического ожидания

Одно свойство математического ожидания, связанное с симметрией закона распределения, мы уже установили. Есть и другие, не менее важные и интересные.

■ **Свойство 1.** Математическое ожидание константы равно этой константе:

$$E(a) = a.$$

Заметим, что константу тоже можно рассматривать как некую «вырожденную» случайную величину, которая принимает всего одно значение, и вероятность этого значения 1.

Свойство 2. Для любой константы a и случайной величины X:

$$E(aX) = a \cdot E(X).$$

Про это свойство говорят ещё, что постоянный множитель можно выносить за знак математического ожидания. Доказать его совсем не сложно. Рассмотрим произвольный закон распределения случайной величины X (табл. 82):

Таблица 82

Значения X	x_1	x_2	x_3	...	x_n
Вероятность	p_1	p_2	p_3	...	p_n

Тогда закон распределения величины aX будет следующим (табл. 83):

Таблица 83

Значения aX	ax_1	ax_2	ax_3	...	ax_n
Вероятность	p_1	p_2	p_3	...	p_n

Обратите внимание, что все значения умножились на a, а вероятности остались прежними. Отсюда

$$E(aX) = ax_1p_1 + ax_2p_2 + \ldots + ax_np_n = a(x_1p_1 + x_2p_2 + \ldots + x_np_n) = aE(X).$$

Так же просто доказывается и следующее свойство.

Свойство 3. Для любых констант a, b и случайной величины X:

$$E(aX + b) = a \cdot E(X) + b.$$

Свойство 4 доказать сложнее. Его доказательство мы отложим до 11 класса. Но использовать это свойство в решении задач будем очень часто. Обратите на него самое пристальное внимание.

Свойство 4. Для любых случайных величин X, Y:

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

т. е. **математическое ожидание суммы случайных величин равно сумме их математических ожиданий**.

Наконец, последнее, тоже очень важное свойство выполнено уже не для всех случайных величин, а только для ***независимых***.

Напомним, что случайные события A и B называются независимыми, если наступление одного из них не влияет на вероятность другого. Более точно независимость случайных событий выражается равенствами

$$P(A \mid B) = P(A)$$

или

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

Последнее равенство обычно и принимают за определение независимости.

Независимость случайных величин определяется через независимость связанных с ними случайных событий.

Дискретные случайные величины X и Y называются независимыми, если независимыми являются любые два события вида $\{X = x_i\}$ и $\{Y = y_j\}$, где
x_i — любое из возможных значений случайной величины X;
y_j — любое из возможных значений случайной величины Y.

Обычно независимость случайных величин, как и независимость событий, следует из условий проведения опыта. Например, если бросают два кубика и X — число очков на первом кубике, а Y — число очков на втором, то эти величины будут независимыми.

Вернёмся к формулировке последнего свойства математического ожидания.

■ **Свойство 5.** Для независимых случайных величин X, Y:

$$E(X \cdot Y) = E(X) \cdot E(Y),$$

т. е. **математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий**.

Это свойство мы также оставим без доказательства.

Посмотрим, как полученные свойства можно использовать при решении задач. Начнём с примера, в котором рассматриваются два кубика.

Пример 1. Подбрасывают два игральных кубика. Нужно найти математические ожидания следующих случайных величин:

X — число, выпавшее на первом кубике;
Y — число, выпавшее на втором кубике;
S — сумма чисел, выпавших на кубиках;
R — произведение чисел, выпавших на кубиках;
W — наибольшее из чисел, выпавших на кубиках;
V — наименьшее из чисел, выпавших на кубиках.

Для величин X, Y мы уже решали эту задачу, используя определение математического ожидания:

$$E(X) = \frac{1+2+3+4+5+6}{6} = 3{,}5.$$

Теперь можно получить тот же результат из симметрии: распределения X и Y симметричны относительно точки 3,5. Для S математическое ожидание получится сразу из свойства 4:

$$E(S) = E(X + Y) = E(X) + E(Y) = 3{,}5 + 3{,}5 = 7.$$

Таким образом, у нас есть целых три способа вычисления $E(S)$: по определению, через симметрию распределения и с использованием свойства математического ожидания. Поскольку величины X и Y независимые, то для вычисления $E(R)$ можно воспользоваться свойством 5:

$$E(R) = E(X \cdot Y) = E(X) \cdot E(Y) = 3{,}5 \cdot 3{,}5 = 12{,}25.$$

Заметим, что мы получили этот результат, вообще не рассматривая закон распределения случайной величины R, который, кстати говоря, будет довольно сложным.

К сожалению, для математического ожидания случайной величины W никакого подходящего свойства нет, но мы уже находили закон распределения W, поэтому посчитаем $E(W)$ по определению:

$$E(W) = \frac{1}{36} \cdot 1 + \frac{3}{36} \cdot 2 + \frac{5}{36} \cdot 3 + \frac{7}{36} \cdot 4 + \frac{9}{36} \cdot 5 + \frac{11}{36} \cdot 6 = \frac{161}{36} \approx 4{,}47.$$

Зато для случайной величины V теперь снова можем воспользоваться свойством 4. В самом деле, поскольку

$$S = X + Y = W + V,$$

то

$$E(V) = E(S) - E(W) = 7 - \frac{161}{36} = \frac{91}{36} \approx 4{,}47.$$

Пример 2. Два стрелка делают по одному выстрелу в одну и ту же мишень. Вероятность попадания первого равна p_1, вероятность попадания второго — p_2. Случайная величина X равна числу пуль, попавших в мишень. Нужно найти её математическое ожидание.

Конечно, можно сначала найти закон распределения случайной величины X, а потом вычислить математическое ожидание $E(X)$ по определению, но мы поступим по-другому.

Рассмотрим две вспомогательные случайные величины X_1 и X_2, каждая из которых будет индикатором попадания в мишень соответствующего стрелка:

$$X_1 = \begin{cases} 1, & \text{если первый стрелок попал в мишень;} \\ 0, & \text{если первый стрелок не попал в мишень.} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{если второй стрелок попал в мишень;} \\ 0, & \text{если второй стрелок не попал в мишень.} \end{cases}$$

Мы уже знаем, что математическое ожидание индикатора равно вероятности соответствующего случайного события, т. е.

$$E(X_1) = p_1, \quad E(X_2) = p_2.$$

Случайная величина X равна сумме индикаторов: $X = X_1 + X_2$. В самом деле, ведь в этой сумме, состоящей из двух индикаторов, ровно столько единиц, сколько пуль оказалось в мишени.

Отсюда

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = p_1 + p_2.$$

Приём, который мы здесь использовали, называется ***метод индикаторов***. Мы ещё вернёмся к нему в следующем пункте.

? ВОПРОСЫ

1. Чему равно математическое ожидание константы?
2. Если $E(X) = 6$, то чему равно математическое ожидание случайной величины $3X$?
3. Чему равно математическое ожидание суммы случайных величин?
4. Для каких случайных величин $E(X \cdot Y) = E(X) \cdot E(Y)$?

4 Математическое ожидание биномиального распределения

В предыдущем параграфе мы рассмотрели два важных распределения, связанных с испытаниями Бернулли: биномиальное и геометрическое. Найдём, чему равны их математические ожидания, и узнаем, как они выражаются через параметры этих распределений.

Напомним, что биномиальное распределение вероятностей обозначается **Bin**(N, p) и зависит от двух параметров N и p: где N — число проведённых испытаний, а p — вероятность успеха в одном испытании Бернулли.

Случайная величина X, имеющая биномиальное распределение, может принимать целые значения от 0 до N. При этом вероятность того, что $X = k$, вычисляется по формуле Бернулли

$$P(X = k) = C_N^k p^k q^{N-k}.$$

Значит, чтобы найти математическое ожидание X, нужно посчитать такую сумму:

$$E(X) = 0 \cdot C_N^0 p^0 q^N + 1 \cdot C_N^1 p^1 q^{N-1} + 2 \cdot C_N^2 p^2 q^{N-2} + \ldots + N \cdot C_N^N p^N q^0.$$

Конечно, для конкретных значений N, p можно вычислить все эти слагаемые и получить ответ, но как найти эту сумму в общем случае? Как выразить её через N и p?

Чтобы решить эту задачу, мы снова, как и в примере 2 из предыдущего пункта, воспользуемся ***методом индикаторов***. Изложим сначала общую идею этого метода.

Напомним, что ***индикатором*** случайного события A называется такая случайная величина I, которая может принимать всего два значения — 0 и 1, и при этом

$$I = \begin{cases} 1, \text{ если событие } A \text{ произошло;} \\ 0, \text{ если событие } A \text{ не произошло.} \end{cases}$$

Если обозначить вероятность $P(A)$ через p, то, как мы уже замечали раньше, $E(I) = p$.

Метод индикаторов применяется для вычисления математического ожидания произвольной случайной величины X следующим образом. Мы пытаемся представить случайную величину X в виде суммы индикаторов каких-то случайных событий. Как найти такие события — зависит каждый раз от конкретной задачи. Если это удаётся, то мы представляем X в виде такой суммы:

$$X = X_1 + X_2 + \ldots + X_N,$$

находим математическое ожидание каждого из индикаторов $X_1, X_2, \ldots, X_N$ (для этого достаточно найти вероятности соответствующих событий), а потом вычисляем $E(X)$ как сумму математических ожиданий индикаторов.

Конечно, такой метод работает далеко не всегда, но если удаётся его применить, то решение задачи сильно упрощается.

Применим метод индикаторов к вычислению математического ожидания биномиального распределения. Поскольку случайная величина X, имеющая этот закон распределения, равна числу успехов в серии из N испытаний Бернулли, то естественно рассмотреть следующие N индикаторов:

$$X_i = \begin{cases} 1, \text{ если в } i\text{-м испытании был успех;} \\ 0, \text{ если в } i\text{-м испытании была неудача,} \end{cases}$$

где $i = 1, 2, \ldots, N$.

Если, например, $N = 4$ и первые два испытания Бернулли закончились успехом, а следующие два испытания — неудачей, то получим:

$$X_1 = 1,\ X_2 = 1,\ X_3 = 0,\ X_4 = 0.$$

Очевидно, что общее число успехов в серии будет равно сумме этих индикаторов:

$$X = X_1 + X_2 + \ldots + X_N,$$

ведь в этой сумме столько же единиц, сколько было успехов в данной серии.

Математическое ожидание любого из N индикаторов равно вероятности успеха p:

$$E(X_1) = E(X_2) = ... = E(X_N) = p.$$

Теперь можем найти математическое ожидание общего числа успехов в N испытаниях:

$$E(X) = E(X_1) + E(X_2) + ... + E(X_N) = \underbrace{p + p + ... + p}_{N \text{ раз}} = Np.$$

Таким образом, мы получили важную формулу:

$$\boxed{E(X) = Np.}$$

Этот результат очень близок к наиболее вероятному числу успехов (вспомните теорему из главы 5 «Испытания Бернулли. Случайный выбор»). Но если наиболее вероятное число успехов было обязательно целым числом, то Np вполне может быть дробным.

Пример 1. Монету бросают 250 раз. Чему равно математическое ожидание числа выпавших орлов?

Это испытания Бернулли с $N = 250$ и $p = \frac{1}{2}$. Если X — это число выпавших орлов, то

$$E(X) = Np = 250 \cdot \frac{1}{2} = 125.$$

Пример 2. Кубик бросают 300 раз. Чему равно математическое ожидание числа выпавших единиц?

Это испытания Бернулли с $N = 300$ и $p = \frac{1}{6}$. Если X — это число выпавших единиц, то

$$E(X) = Np = 300 \cdot \frac{1}{6} = 50.$$

Пример 3. Баскетболист попадает в корзину с вероятностью 0,6. За игру он сделал 47 бросков. Чему равно математическое ожидание числа попаданий?

Это испытания Бернулли с $N = 47$ и $p = 0{,}6$. Если X — это число точных бросков, то

$$E(X) = Np = 47 \cdot 0{,}6 = 28{,}2.$$

? ВОПРОСЫ

1. Чему равно математическое ожидание биномиального распределения?
2. Монету подбрасывают 3 раза. Чему равно математическое ожидание числа орлов?
3. Кубик подбрасывают 3 раза. Чему равно математическое ожидание числа шестёрок?
4. При стрельбе из положения стоя биатлонист совершает около 20% промахов. На огневом рубеже 5 мишеней. Чему равно математическое ожидание числа сбитых мишеней?

5 Математическое ожидание геометрического распределения*

Рассмотрим теперь геометрическое распределение. Напомним, что такое распределение имеет случайная величина X, равная числу испытаний Бернулли до первого успеха. Она может принимать любые значения от 1 до бесконечности. При этом вероятность того, что $X = k$, вычисляется по формуле

$$P(X = k) = q^{k-1}p.$$

Здесь решение задачи осложняется тем, что сумма в выражении математического ожидания содержит бесконечное число слагаемых:

$$E(X) = 1 \cdot p + 2 \cdot qp + 3 \cdot q^2p + 4 \cdot q^3p + \ldots .$$

Мы уже знаем, что такие суммы тоже могут быть конечными, например сумма бесконечно убывающей геометрической прогрессии. Но здесь выражение сложнее: перед каждым членом прогрессии стоят множители 1, 2, 3, 4, ...

Попробуем выписать слагаемые этой суммы в виде следующей бесконечной треугольной таблицы (табл. 84):

Таблица 84

p	qp	q^2p	q^3p	q^4p	...
	qp	q^2p	q^3p	q^4p	...
		q^2p	q^3p	q^4p	...
			q^3p	q^4p	...
				q^4p	...
					...

Вы видите, что каждый член геометрической прогрессии входит в эту таблицу столько раз, каков множитель, стоящий перед ним в нужной нам сумме. Значит, если мы сможем просуммировать все числа в этой таблице, то это и будет искомым математическим ожиданием.

Заметим, что сумма чисел в первой строке будет равна сумме бесконечно убывающей геометрической прогрессии со знаменателем q и первым членом p, т. е. $\frac{p}{1-q}$.

Сумма чисел во второй строке будет равна сумме бесконечно убывающей геометрической прогрессии со знаменателем q и первым членом qp, т. е. $\frac{qp}{1-q}$.

Аналогично сумма чисел в третьей строке будет равна $\frac{q^2p}{1-q}$, в четвёртой — $\frac{q^3p}{1-q}$ и т. д.

Получаем, что

$$E(X) = \frac{p}{1-q} + \frac{qp}{1-q} + \frac{q^2p}{1-q} + \frac{q^3p}{1-q} + \ldots = 1 + q + q^2 + q^3 + \ldots = \frac{1}{1-q} = \frac{1}{p}.$$

Таким образом, перестановка и перегруппировка слагаемых в бесконечной сумме позволила нам свести задачу к вычислению суммы геометрической прогрессии. Конечно, у вас может возникнуть вопрос: насколько законно выполнять такие операции с суммами, содержащими бесконечное число слагаемых? Оказывается, это действительно можно делать не всегда. Но если все слагаемые в сумме положительные, то от перестановки и перегруппировки слагаемых сумма не меняется.

Итак, для геометрического распределения

$$\boxed{E(X) = \frac{1}{p}.}$$

Есть другой способ вычисления математического ожидания для геометрического распределения. Он, как и в предыдущем пункте, основан на методе индикаторов. Мы оставим этот способ в виде самостоятельного упражнения, а сейчас рассмотрим несколько примеров.

Пример 1. Сколько раз в среднем нужно бросить монету, чтобы на ней выпал орёл?

Речь здесь идёт о математическом ожидании случайной величины X, имеющей геометрическое распределение с параметром $p = \frac{1}{2}$. Отсюда $E(X) = \frac{1}{p} = 2$.

Пример 2. Сколько раз нужно в среднем бросить кубик, чтобы на нём выпала шестёрка?

Здесь нужно вычислить математическое ожидание случайной величины Y, имеющей геометрическое распределение с параметром $p = \frac{1}{6}$. Поэтому $E(Y) = \frac{1}{p} = 6$.

Как видите, ответы в этих двух примерах вполне ожидаемые. Вообще, **если успех — это один из n равновозможных исходов одного испытания Бернулли, то ждать его придётся в среднем n испытаний**.

Пример 3. Студент сдаёт зачёт по теории вероятностей до тех пор, пока не решит какую-нибудь задачу. Вероятность решения любой задачи этим студентом равна 0,2. Сколько задач в среднем он получит?

В этой задаче геометрическое распределение имеет параметр $p = 0{,}2$. Поэтому математическое ожидание равно $\frac{1}{p} = \frac{1}{0{,}2} = 5$. Таким образом, студенту следует ожидать в среднем 5 задач до получения зачёта.

Если знать, сколько времени даётся на решение одной задачи, можно вычислить, сколько времени в среднем будет длиться такой зачёт.

? ВОПРОСЫ

1. Чему равно математическое ожидание геометрического распределения?
2. Кубик подбрасывают до появления чётного числа. Чему равно математическое ожидание числа испытаний?
3. Иван Иванович узнал, что экзамен по вождению автомобиля в его автошколе сдаёт около 40% экзаменуемых. С какого раза в среднем ему следует ожидать сдачу экзамена?

УПРАЖНЕНИЯ

В некоторых задачах этого раздела вы можете использовать результаты, полученные в упражнениях § 17 «Распределение вероятностей».

331. Распределение вероятностей дискретной случайной величины X задано таблицами 85—87:

а)

Таблица 85

Значения X	1	2	3	4	5
Вероятность	0,2	0,1	0,1	0,2	0,4

б)

Таблица 86

Значения X	-2	-1	0	1	2
Вероятность	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{16}$

в)

Таблица 87

Значения X	3	7	11	15	19
Вероятность	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{16}$

Найдите математическое ожидание $E(X)$.

332. Закон распределения дискретной случайной величины X задан таблицей 88:

Таблица 88

Значения X	-10	-1	0	1	10
Вероятность	0,3	0,15	0,1	0,15	0,3

Найдите математическое ожидание случайных величин $Y = |X|$ и $Z = X^2$.

333. Страховой полис на один год стоит 8000 р. Страховые случаи делятся на 3 категории: с малым (10 000 р.), средним (50 000 р.) и большим (500 000 р.) ущербом. При этом каждый из этих случаев возникает в течение года с вероятностями 0,1, 0,05 и 0,002 соответственно. Найдите ожидаемый доход страховой компании от продажи одного полиса.

334. В партии 10% нестандартных деталей. Из неё наугад отбирают 4 детали. Случайная величина X равна числу нестандартных деталей среди отобранных. Найдите математическое ожидание случайной величины X.

335. Профессор Злобин не хочет ставить положительную оценку студенту Разгуляеву и задаёт ему вопросы до тех пор, пока на очередном вопросе студент не ошибётся. Вероятность правильного ответа на любой вопрос составляет 0,8. Случайная величина X равна числу вопросов, на которые правильно ответит студент, перед тем, как получит двойку. Найдите математическое ожидание случайной величины X.

336*. Биатлонист должен поразить 3 мишени пятью выстрелами. На каждый выстрел он тратит 8 с и попадает в цель с вероятностью 0,6. Случайная величина T — общее время, которое он проведёт на огневом рубеже. Найдите математическое ожидание T.

337*. В ящике лежат 5 белых и 5 чёрных шаров. Из него наугад вытаскивают 4 шара. Случайная величина X равна числу белых шаров в полученной выборке. Найдите математическое ожидание X.

338*. Три джентльмена пришли в ресторан в одинаковых шляпах и сдали их в гардероб. Уходя домой, они разобрали шляпы наугад. Случайная величина X равна количеству джентльменов, ушедших домой в своих шляпах. Найдите её математическое ожидание.

339*. Решите предыдущую задачу для четырёх джентльменов.

340*. Пусть N — выбранное наугад двузначное случайное число, а случайная величина X равна сумме его цифр. Найдите математическое ожидание случайной величины X.

341*. Решите предыдущую задачу, если X — это наибольшая из двух цифр двузначного числа N.

342*. В волейболе матч продолжается, пока одна из команд не одержит победу в трёх партиях. Случайная величина X равняется количеству партий в волейбольном матче между командами «Чайка» и «Буревестник». Ничьих в волейбольных партиях не бывает. Найдите математическое ожидание случайной величины X, если:
а) в каждой партии команды имеют равные шансы на победу;
б) в каждой партии «Чайка» выигрывает у «Буревестника» с вероятностью 0,8.

343*. Бросают 3 кубика. Случайная величина K равна количеству различных чисел, которые при этом выпали. Найдите математическое ожидание K.

344*. Подбрасывают одновременно n кубиков. Случайная величина M равна наибольшему из выпавших чисел. Найдите математическое ожидание M.

345*. На новогодний вечер собралось 12 детей, каждый из которых пришёл со своим подарком. Все подарки были повешены на ёлку, а в конце вечера случайно разыграны. Найдите математическое ожидание числа детей, получивших свои подарки.

346*. 10 юношей и 10 девушек случайным образом садятся за круглый стол. Будем считать юношу счастливым, если рядом с ним сидит хотя бы одна девушка. Найдите математическое ожидание числа счастливых юношей.

347*. На конференцию приехало $2n$ учёных. Перед конференцией n из них узнали о новом научном открытии. В гостинице их случайным образом поселили в двухместных номерах. Каждый, кто знал об открытии, рассказал о нём своему соседу по номеру. Найдите математическое ожидание числа учёных, знавших на следующий день о новом открытии.

348*. Требуется протестировать большое количество человек N на наличие в крови некоторого вируса. Каждый тест обходится очень дорого, поэтому предлагается такая схема тестирования. N человек разбиваются на группы размера k (можно считать, что N делится на k), и сначала проводится анализ крови, собранной во всей группе. Если тест отрицательный, то вся группа считается здоровой. Если тест положительный, то тестируется отдельно каждый человек из этой группы. Вероятность обнаружения вируса в крови любого человека равна p. Чему равно ожидаемое число тестов при такой схеме тестирования? Будет ли оно меньше N?

§ 19. Дисперсия и стандартное отклонение

1 Что такое дисперсия?

В статистике среднее арифметическое $\overline{x}$ далеко не всегда даёт объективную информацию о поведении числового набора. Мы уже не раз обсуждали среднюю зарплату, «среднюю температуру по больнице» и многие другие курьёзные ситуации, когда среднее значение нельзя рассматривать без учёта других числовых характеристик ряда.

Примерно такая же ситуация наблюдается и в теории вероятностей. На рисунке 105 представлены законы распределения с одним и тем же математическим ожиданием, равным 5, но остальные их свойства абсолютно разные.

Вы помните, что для числового набора, помимо средних характеристик, есть ещё и характеристики разброса. Простейшая из них — размах, т. е. разность наибольшего и наименьшего значений.

В теории вероятностей подобная характеристика практически не используется: во-первых, наибольшее и наименьшее значения часто имеют очень маленькую вероятность, а во-вторых, у некоторых распределений

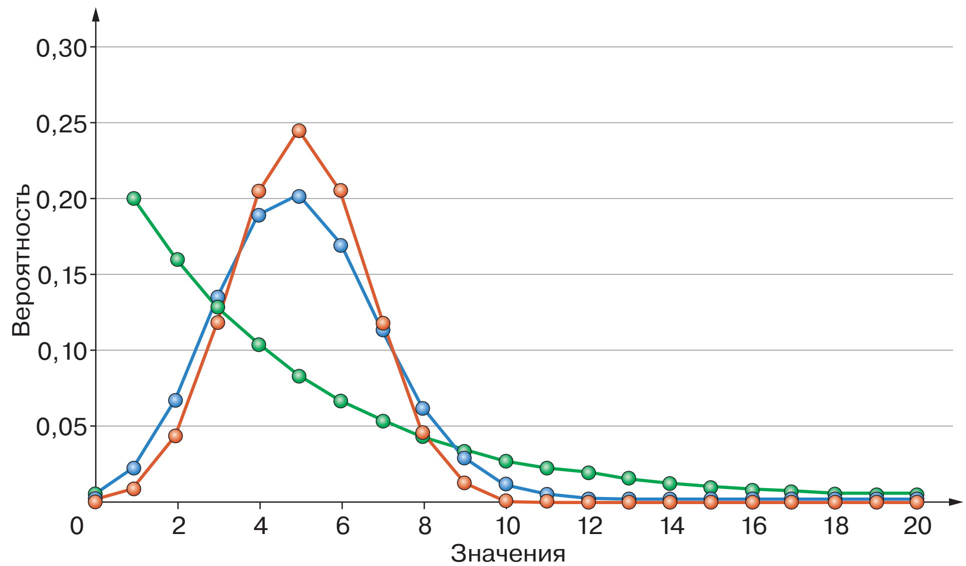


Рис. 105. У всех этих распределений математическое ожидание равно 5

таких значений вовсе нет. Например, в известном вам геометрическом распределении нет наибольшего значения.

Изучая статистику в основной школе, вы познакомились ещё с одной величиной, которая характеризует разброс данных, — ***дисперсией*** числового набора. Напомним, что так называется в статистике **среднее арифметическое квадратов отклонений от среднего арифметического** $\overline{x}$:

$$S^2 = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n}.$$

В дальнейшем будем называть её ***выборочной дисперсией***, чтобы не путать с понятием дисперсии, которое мы введём для случайной величины.

Поскольку аналогом среднего арифметического для случайной величины является математическое ожидание, то естественным будет следующее определение дисперсии случайной величины.

Дисперсией случайной величины X называется математическое ожидание квадрата отклонения X от её математического ожидания $E(X)$:

$$D(X) = E(X - E(X))^2.$$

Слово «дисперсия» и обозначение $D(X)$ происходят от латинского *dispersio* — «рассеяние».

Выражение для вычисления дисперсии дискретной случайной величины можно записать через её закон распределения. Пусть распределения вероятностей случайной величины X задаётся таблицей 89.

Таблица 89

Значения X	x_1	x_2	…	x_n
Вероятность	p_1	p_2	…	p_n

Обозначим через m математическое ожидание $E(X)$. Тогда закон распределения случайной величины $(X - m)^2$ будет таким (табл. 90):

Таблица 90

Значения $(X - m)^2$	$(x_1 - m)^2$	$(x_2 - m)^2$	...	$(x_n - m)^2$
Вероятность	p_1	p_2	...	p_n

Математическое ожидание этой величины и будет дисперсией случайной величины X:

$$D(x) = (x_1 - m)^2 p_1 + (x_2 - m)^2 p_2 + ... + (x_n - m)^2 p_n.$$

Как видите, формула для вычисления дисперсии случайной величины $D(X)$ отличается от формулы для выборочной дисперсии S^2 тем, что здесь **каждый квадрат отклонения входит в сумму со своим весом, равным соответствующей вероятности**.

Попробуем найти дисперсию какой-нибудь величины по этой формуле.

Пример. Пусть X — число очков, которое выпадает при подбрасывании игрального кубика. Найдём дисперсию этой случайной величины.

Мы уже знаем, что $E(X) = 3{,}5$. Поэтому

$$D(X) = (1 - 3{,}5)^2 \cdot \frac{1}{6} + (2 - 3{,}5)^2 \cdot \frac{1}{6} + (3 - 3{,}5)^2 \cdot \frac{1}{6} + (4 - 3{,}5)^2 \cdot \frac{1}{6} +$$

$$+ (5 - 3{,}5)^2 \cdot \frac{1}{6} + (6 - 3{,}5)^2 \cdot \frac{1}{6} = 6{,}25 \cdot \frac{1}{6} + 2{,}25 \cdot \frac{1}{6} + 0{,}25 \cdot \frac{1}{6} +$$

$$+ 0{,}25 \cdot \frac{1}{6} + 2{,}25 \cdot \frac{1}{6} + 6{,}25 \cdot \frac{1}{6} = 17{,}5 \cdot \frac{1}{6} = \frac{35}{12} \approx 2{,}917.$$

Как видите, вычисления оказались довольно громоздкими даже для такого простого примера. Но уже в следующем пункте мы научимся вычислять дисперсию, используя другие формулы и способы.

? ВОПРОСЫ

1. Дайте определение дисперсии случайной величины.
2. Как обозначается дисперсия случайной величины X?
3. Как вычислить дисперсию дискретной случайной величины по её закону распределения?
4. Случайная величина X принимает значения –1 и 1 с одинаковыми вероятностями 0,5. Чему равна её дисперсия?

2 Свойства дисперсии

Как и математическое ожидание, дисперсия имеет ряд свойств, которые могут значительно упростить её вычисление.

■ **Свойство 1.** Дисперсия случайной величины может быть вычислена по формуле

$$D(X) = E(X)^2 - E^2(X).$$

Эту формулу можно прочитать так: **дисперсия случайной величины равна разности математического ожидания квадрата этой величины и квадрата её математического ожидания**.

Для доказательства этой формулы будем использовать уже известные нам свойства математического ожидания. Введём обозначение $E(X) = m$. Тогда $D(X) = E(X - m)^2 = E(X - 2Xm + m^2)$.

Пока мы просто возвели выражение $(X - m)$ в квадрат. Теперь воспользуемся свойствами математического ожидания:

$$E(X^2 - 2Xm + m^2) = E(X^2) - 2mE(X) + m^2 = E(X^2) - 2m^2 + m^2 = E(X^2) - m^2.$$

Формула доказана.

В чём её преимущество перед формулой, которая была дана в определении дисперсии? Дело в том, что $E(X^2)$ вычислять обычно проще, чем $E(X - E(X))^2$. Убедимся в этом, вернувшись к примеру с кубиком из предыдущего пункта.

Пример 1. Найдём дисперсию случайной величины X, равной числу очков, которое выпадает при подбрасывании игрального кубика. Будем использовать для этого только что полученную формулу.

Мы знаем, что $E(X) = 3{,}5$. Найдём $E(X^2)$:

$$E(X^2) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6}.$$

На этот раз дело обошлось почти без дробей, поскольку все значения X были целыми числами. Остаётся подставить найденные значения в формулу:

$$D(X) = E(X^2) - E^2(X) = \frac{91}{6} - \left(\frac{1}{2}\right)^2 = \frac{35}{12}.$$

Вычисления получились менее трудоёмкими.

■ **Свойство 2.** Дисперсия любой случайной величины X неотрицательна: $D(X) \geqslant 0$.

При этом она равняется нулю только для константы.

Доказательство следует сразу из определения дисперсии, поскольку все слагаемые в этом определении неотрицательные. Это свойство помогает обнаружить ошибки в вычислениях, если вдруг вы получили отрицательное значение дисперсии по формуле из свойства 1.

■ **Свойство 3.** Для любой константы a и случайной величины X: $D(aX) = a^2 \cdot D(X)$.

Таким образом, если постоянный множитель выносят за знак дисперсии, то он возводится в квадрат.

Чтобы доказать это свойство, рассмотрим произвольный закон распределения случайной величины X (табл. 91):

Таблица 91

Значения X	x_1	x_2	…	x_n
Вероятность	p_1	p_2	…	p_n

Пусть $E(X) = m$. Тогда $E(aX) = aE(X) = a \cdot m$, и закон распределения величины $(aX - am)^2$ будет выглядеть так (табл. 92):

Таблица 92

Значения $(aX - am)^2$	$(ax_1 - am)^2$	$(ax_2 - am)^2$	…	$(ax_n - am)^2$
Вероятность	p_1	p_2	…	p_n

Вычислим её математическое ожидание, которое и будет равно $D(aX)$:

$$E(aX - am)^2 = a^2(x_1 - m)^2 p_1 + \ldots + a^2(x_n - m)^2 p_n = a^2 D(X).$$

Таким же образом можно доказать и следующее свойство.

Свойство 4. Для любых констант a, b и случайной величины X:

$$D(aX + b) = a^2 \cdot D(X).$$

А вот дисперсия суммы будет равна сумме дисперсий только для независимых величин!

Свойство 5. Для независимых случайных величин X, Y:

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

т. е. **дисперсия суммы независимых случайных величин равна сумме их дисперсий**.

При доказательстве этого свойства будем использовать формулу для вычисления дисперсии, полученную в свойстве 1:

$$D(X + Y) = E((X + Y)^2) - (E(X + Y))^2 = E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2.$$

Введём обозначения: $E(X) = m_X$, $E(Y) = m_Y$. Поскольку случайные величины X и Y независимы, то по одному из свойств математического ожидания $E(XY) = m_X m_Y$. Отсюда получаем:

$$E(X^2 + 2XY + Y^2) - (m_X + m_Y)^2 = E(X^2) + 2m_X m_Y + E(Y^2) -$$
$$- (m_X^2 + 2m_X m_Y + m_Y^2) = E(X^2) - m_X^2 + E(Y^2) - m_Y^2 = D(X) + D(Y).$$

Доказанное свойство является одним из важнейших свойств дисперсии. Именно на нём основаны многие вероятностные законы, в том числе уже упоминавшийся не раз ***закон больших чисел***.

Решим ещё несколько задач с использованием свойств дисперсии.

Пример 2. Подбрасывают два игральных кубика. Чему равна дисперсия случайной величины S, равной сумме выпавших очков?

Если X — число, выпавшее на первом кубике, а Y — число, выпавшее на втором, то $S = X + Y$. Дисперсии X и Y мы нашли выше в примере 1:

$$D(X) = D(Y) = \frac{35}{12}.$$

Из независимости X, Y получаем:

$$D(S) = D(X + Y) = D(X) + D(Y) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6} \approx 5{,}8.$$

Пример 3. Найдём дисперсию бинарной случайной величины I, которая является индикатором некоторого случайного события A. Если $P(A) = p$, то законы распределения I и I^2 будут одинаковыми (табл. 93 и 94):

Таблица 93

Значения *I*	0	1
Вероятность	$1 - p$	p

Таблица 94

Значения I^2	0	1
Вероятность	$1 - p$	p

Так произошло потому, что $0^2 = 0$ и $1^2 = 1$. Отсюда

$$E(I^2) = E(I) = p$$
$$D(I) = E(I^2) - E^2(I) = p - p^2 = p(1 - p).$$

Пример 4. Два стрелка делают по одному выстрелу в одну и ту же мишень. Вероятность попадания первого равна p_1, вероятность попадания второго — p_2. Случайная величина X равна числу пуль, попавших в мишень. Чему равна её дисперсия?

Напомним, что при вычислении $E(X)$ мы использовали метод индикаторов. Для этого мы представили случайную величину X как суммы двух случайных величин $X = X_1 + X_2$, где X_1 и X_2 — индикаторы попадания в мишень для первого и второго стрелка. Формулу для дисперсии индикаторов мы получили в предыдущем примере:

$$D(X_1) = p_1(1 - p_1), \quad D(X_2) = p_2(1 - p_2).$$

Поскольку случайные величины X_1 и X_2 независимы, то

$$D(X) = D(X_1 + X_2) = D(X_1) + D(X_2) = p_1(1 - p_1) + p_2(1 - p_2).$$

? ВОПРОСЫ

1. По какой формуле можно вычислять дисперсию?
2. Может ли дисперсия быть отрицательной? Чему равна дисперсия константы?
3. Если $D(X) = 6$, то чему равна дисперсия случайной величины $3X$?
4. Для каких случайных величин $D(X + Y) = D(X) + D(Y)$?

3 Стандартное отклонение

Если числовой набор имеет какие-то единицы измерения, то его дисперсия измеряется в «квадратных» единицах, поскольку она равняется среднему арифметическому **квадратов** отклонений чисел от их среднего арифметического.

Чтобы вернуться к исходным единицам, мы ввели в статистике такую меру разброса, как ***стандартное***, или ***среднеквадратичное***, ***отклонение***, которое равнялось корню из дисперсии:

$$S = \sqrt{S^2}.$$

Такая же картина наблюдается и в теории вероятностей. Если случайная величина X измеряется, например, в сантиметрах, то её математическое ожидание $E(X)$ измеряется в тех же единицах, а дисперсия — в квадратных сантиметрах, поскольку $D(X) = E(X - E(X))^2$.

Поэтому для случайной величины вводят аналогичную характеристику, равную корню из дисперсии.

Стандартным, или среднеквадратичным, отклонением случайной величины называется корень из её дисперсии. Для обозначения стандартного отклонения используется греческая буква σ — «сигма»:

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

Термины «стандартное» и «среднеквадратичное» говорят о том, что в среднем значения случайной величины X отклоняются от $E(X)$ примерно на величину σ. Есть значения, которые отклоняются от $E(X)$ меньше, чем на σ, есть которые больше.

Более точно связь между отклонением X от математического ожидания и величиной σ выражается следующим неравенством.

Неравенство «трёх сигм». Для любой случайной величины X вероятность того, что она отклонится от своего математического ожидания более, чем на 3σ, не превосходит $\frac{1}{9}$:

$$P(|X - E(X)| \geqslant 3\sigma) \leqslant \frac{1}{9}.$$

Перейдя к противоположному событию, можно записать это неравенство по-другому:

$$P(|X - E(X)| < 3\sigma) \geqslant \frac{8}{9}.$$

Значит, есть не меньше 8 шансов из 9, что значение случайной величины X окажется в интервале

$$(E(X) - 3\sigma;\ E(X) + 3\sigma).$$

На самом деле для многих распределений эта вероятность ещё больше. Таким образом, **интервал «трёх сигм» включает в себя почти все вероятные значения случайной величины** X. Иногда в него попадают вообще все возможные значения X.

Мы докажем приведённое выше неравенство в 11 классе, а пока вычислим стандартное отклонение для рассмотренных ранее величин.

Пример 1. Случайная величина X равна числу очков, выпавших при подбрасывании игрального кубика. Найдём её стандартное отклонение.

Поскольку $D(X) = \frac{35}{12} \approx 2{,}9$, то $\sigma = \sqrt{D(X)} = \sqrt{\frac{35}{12}} \approx 1{,}7$.

Так как $E(X) = 3{,}5$ и $3\sigma = 5{,}1$, то интервалом «трёх сигм» будет $(3{,}5 - 5{,}1;\ 3{,}5 + 5{,}1)$. В нём лежат все значения случайной величины X.

Пример 2. Пусть теперь S — сумма очков, выпавших при подбрасывании двух кубиков. Найдём её стандартное отклонение.

Мы уже вычислили $D(S) = \frac{35}{6} \approx 5{,}8$, поэтому $\sigma = \sqrt{D(S)} = \sqrt{\frac{35}{6}} \approx 2{,}4$.

Так как $E(S) = 7$ и $3\sigma = 7{,}2$, то интервалом «трёх сигм» будет $(7 - 7{,}2;\ 7 + 7{,}2)$. В нём также лежат все значения случайной величины S.

Пример 3. Вернёмся к задаче, в которой два стрелка делают по одному выстрелу в мишень. Если вероятности попадания для них равны p_1 и p_2, то случайная величина X, равная общему числу попаданий, имеет следующие характеристики:

$$E(X) = p_1 + p_2,\quad D(X) = p_1(1 - p_1) + p_2(1 - p_2).$$

Отсюда $\sigma = \sqrt{p_1(1 - p_1) + p_2(1 - p_2)}$. Если взять, например, $p_1 = 0{,}99$ и $p_2 = 0{,}01$, то получим:

$$E(X) = 1;\quad \sigma \approx 0{,}14;\quad 3\sigma \approx 0{,}42$$

и интервалом «трёх сигм» будет интервал $(0{,}58;\ 1{,}42)$. Он будет включать только одно из трёх возможных значений случайной величины X (напомним, она может принимать значения 0, 1 и 2). Но вероятность этого значения будет

$$P(X = 1) = 0{,}99 \cdot 0{,}99 + 0{,}01 \cdot 0{,}01 = 0{,}98 > \frac{8}{9}.$$

Так что неравенство «трёх сигм» здесь тоже выполняется.

? ВОПРОСЫ

1. Что такое стандартное отклонение?
2. Если случайная величина X измеряется в метрах, то в каких единицах измеряется её дисперсия? стандартное отклонение?
3. Запишите неравенство «трёх сигм».

4 Дисперсия биномиального распределения

Дисперсию биномиального распределения можно найти всё тем же ***методом индикаторов***. Напомним, что при вычислении математического ожидания мы представили случайную величину X, равную числу успехов в N испытаниях Бернулли, как сумму индикаторов:

$$X = X_1 + X_2 + \ldots + X_N.$$

Каждая из величин $X_1, X_2, \ldots, X_N$ равна 1, если соответствующее ей испытание завершилось успехом, и равна 0 — если неудачей. Отсюда

$$D(X_1) = D(X_2) = \ldots = D(X_N) = p(1 - p) = pq.$$

Из независимости испытаний Бернулли следует независимость случайных величин $X_1, X_2, \ldots, X_N$, поэтому

$$D(X) = D(X_1) + D(X_2) + \ldots + D(X_N) = \underbrace{pq + pq + \ldots + pq}_{N \text{ раз}} = Npq.$$

Таким образом, мы получили формулу для дисперсии биномиального распределения:

$$\boxed{D(X) = Npq.}$$

Рассмотрим её использование на примерах.

Пример 1. Монету бросают 25 раз. Чему равна дисперсия числа выпавших орлов?

Это испытания Бернулли с $N = 25$ и $p = q = \frac{1}{2}$. Если X — число выпавших орлов, то

$$E(X) = Np = 25 \cdot \frac{1}{2} = 12{,}5;\ D(X) = Npq = 25 \cdot \frac{1}{2} \cdot \frac{1}{2} = 6{,}25.$$

Отсюда $\sigma = 2{,}5$ и $3\sigma = 7{,}5$. Интервалом «трёх сигм» будет интервал (5; 20). В него попадают значения от 6 до 19, что составляет около 54% всех возможных значений величины X.

Если монету подбрасывают 250 раз, то

$$E(X) = Np = 250 \cdot \frac{1}{2} = 125;\ D(X) = Npq = 250 \cdot \frac{1}{2} \cdot \frac{1}{2} = 62{,}5.$$

Теперь $\sigma \approx 7{,}9$ и $3\sigma \approx 23{,}7$. Интервалом «трёх сигм» будет интервал (101,3; 148,7), содержащий значения от 102 до 148, что составляет около 19% всех возможных значений величины X. Мы видим, что с увеличением N доля вероятных значений величины X уменьшилась.

Увеличим число испытаний до 2500:

$$E(X) = Np = 2500 \cdot \frac{1}{2} = 1250;\ D(X) = Npq = 2500 \cdot \frac{1}{2} \cdot \frac{1}{2} = 625.$$

Поскольку $\sigma = 25$ и $3\sigma = 75$, то интервалом «трёх сигм» будет интервал (1175; 1325). В него попадает всего 151 значение случайной величины X, что составляет только 6% от всех возможных её значений.

Таким образом, с ростом числа опытов доля вероятных значений случайной величины X уменьшается. Увеличивая N, эту долю можно сделать сколь угодно малой, поскольку диапазон возможных значений случайной величины растёт пропорционально N, а длина интервала «трёх сигм» растёт как $\sqrt{N}$. Мы ещё не раз используем это свойство в дальнейшем.

Пример 2. Кубик бросают 2500 раз. Чему равна дисперсия числа выпавших единиц?

Это испытания Бернулли с $N = 2500$ и $p = \frac{1}{6}$. Если X — это число выпавших единиц, то

$$E(X) = 2500 \cdot \frac{1}{6} = Np \approx 416{,}7; \; D(X) = Npq = 2500 \cdot \frac{1}{6} \cdot \frac{5}{6} \approx 347{,}2.$$

Как видим, дисперсия биномиального распределения при $p \neq q$ меньше, чем для симметричного случая $p = q$. Это легко объяснить: выражение $pq = p(1 - p)$ является относительно переменной p квадратным трёхчленом и достигает максимума в точке $p = 0{,}5$ (рис. 106). При $p = q$ мы имеем максимальную неопределённость, а при $p = 0$ и $p = 1$ дисперсия обращается в 0.

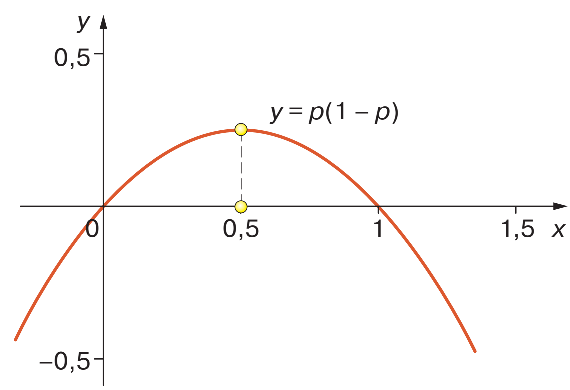


Рис. 106. График функции $y = p(1 - p)$

В этом примере $\sigma \approx 18{,}6$ и $3\sigma \approx 56$. Интервал «трёх сигм» включает значения от 361 до 472, что составляет чуть больше 4% от всех возможных её значений.

? ВОПРОСЫ

1. Чему равна дисперсия биномиального распределения?
2 Монету подбрасывают 5 раз. Чему равна дисперсия числа орлов? стандартное отклонение?
3. Кубик подбрасывают 5 раз. Чему равна дисперсия числа шестёрок? стандартное отклонение?

5 Дисперсия геометрического распределения*

Для вычисления дисперсии геометрического распределения воспользуемся тем же приёмом, который мы использовали при вычислении математического ожидания. Только теперь его придётся применить дважды.

Чтобы вычислить дисперсию $D(X)$ для случайной величины X, имеющей геометрическое распределение, нужно найти $E(X^2)$. Выражение для этой величины, исходя из геометрического закона распределения, будет выглядеть так:

$$E(X^2) = 1^2 \cdot p + 2^2 \cdot qp + 3^2 \cdot q^2p + 4^2 \cdot q^3p + \ldots .$$

Запишем эти слагаемые в таблицу (табл. 95):

Таблица 95

p	qp	q^2p	…
	qp	q^2p	…
	qp	q^2p	…
	qp	q^2p	…
		q^2p	…
		q^2p	…
		q^2p	…
		q^2p	…
		q^2p	…
			…

Она устроена немного сложнее, чем для математического ожидания: в ней любой k-й столбец содержит на $k^2 - (k-1)^2 = 2k - 1$ больше чисел, чем предыдущий. Поэтому

$$E(X^2) = \frac{p}{1-q} + \frac{3qp}{1-q} + \frac{5q^2p}{1-q} + \frac{7q^3p}{1-q} + \ldots = 1 + 3q + 5q^2 + 7q^3 + \ldots .$$

Для этой суммы составим ещё одну таблицу (табл. 96):

Таблица 96

1	q	q^2	…
	q	q^2	…
	q	q^2	…
		q^2	…
		q^2	…
			…

Отсюда

$$E(X^2) = \frac{1}{1-q} + \frac{2q}{1-q} + \frac{2q^2}{1-q} + \frac{2q^3}{1-q} + \frac{2q^4}{1-q} + \ldots =$$

$$= \frac{1}{p} + \frac{2}{p}(q + q^2 + q^3 + \ldots) = \frac{1}{p} + \frac{2q}{p(1-q)} = \frac{p+2q}{p^2} = \frac{1+q}{p^2}.$$

Теперь можно найти дисперсию:

$$D(X) = E(X^2) - E^2(X) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Получили формулу для дисперсии геометрического распределения:

$$\boxed{D(X) = \frac{q}{p^2}.}$$

Рассмотрим несколько примеров применения этой формулы.

Пример 1. Найдём дисперсию числа испытаний с монетой до появления первого орла.

Поскольку здесь $p = q = \frac{1}{2}$, то дисперсия будет равна $\frac{q}{p^2} = \frac{4}{2} = 2$.

Отсюда $\sigma = \sqrt{2} \approx 1{,}41$. Поскольку математическое ожидание числа испытаний до первого успеха равно $\frac{1}{p} = 2$, то интервал «трёх сигм» получится от 1 до 6 (мы сдвинули его левую границу, получившуюся отрицательной, в 1). То есть, хотя в среднем понадобится всего два испытания, ожидание «успеха» может вполне продлиться до 6 испытаний.

Посмотрим, как обстоит дело с выпадением шестёрки на кубике.

Пример 2. Найдём дисперсию числа испытаний с кубиком до появления первой шестёрки.

Здесь $p = \frac{1}{6}$, $q = \frac{5}{6}$, поэтому дисперсия будет равна $\frac{q}{p^2} = \frac{5}{6} \cdot 36 = 30$.

Отсюда $\sigma = \sqrt{30} \approx 5{,}48$. Математическое ожидание равно $\frac{1}{p} = 6$, а интервал «трёх сигм» получится от 1 до 22. Как видим, с кубиком дело обстоит ещё хуже: хотя в среднем успех придёт на 6 испытании, его вполне можно ждать и до 22!

Таким образом, особенностью геометрического распределения является большая дисперсия, которая тем больше, чем меньше вероятность успеха.

Поэтому, если вы собираетесь стрелять в тире по мишени до первого попадания и ваша меткость p невелика, запасите не $\frac{1}{p}$ патронов, а побольше. Какой именно взять запас, легко найти из неравенства «трёх сигм»: для геометрического распределения $3\sigma = \frac{3\sqrt{q}}{p}$, поэтому правой границей интервала «трёх сигм» будет $\frac{1}{p} + \frac{3\sqrt{q}}{p}$.

Если, например, вы попадаете в мишень с вероятностью $p = 0{,}2$, то лучше взять не 5 патронов, а

$$\frac{1 + 3\sqrt{q}}{p} = \frac{1 + 3\sqrt{0{,}8}}{0{,}2} \approx 18.$$

? ВОПРОСЫ

1. Чему равна дисперсия геометрического распределения?
2. Кубик подбрасывают до появления чётного числа. Чему равна дисперсия числа испытаний?
3. Иван Иванович узнал, что экзамен по вождению автомобиля в его автошколе сдаёт около 40% экзаменуемых. С какого в среднем раза ему следует ожидать сдачу экзамена? На что рассчитывать в худшем случае?

УПРАЖНЕНИЯ

В некоторых задачах этого раздела вы можете использовать результаты, полученные в упражнениях § 17 «Распределение вероятностей» и § 18 «Математическое ожидание».

349. Распределение вероятностей дискретной случайной величины X задано таблицами 97—99:

а)

Таблица 97

Значения X	1	2	3	4	5
Вероятность	0,2	0,1	0,1	0,2	0,4

б)

Таблица 98

Значения X	–2	–1	0	1	2
Вероятность	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{16}$

в)

Таблица 99

Значения X	3	7	11	15	19
Вероятность	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{2}$	$\frac{3}{16}$	$\frac{1}{16}$

Найдите дисперсию и стандартное отклонение случайной величины X.

350. Закон распределения дискретной случайной величины X задан таблицей 100:

Таблица 100

Значения X	–10	–1	0	1	10
Вероятность	0,3	0,15	0,1	0,15	0,3

Найдите дисперсию и стандартное отклонение случайных величин $Y = |X|$ и $Z = X^2$.

351. В партии 10% нестандартных деталей. Из неё наугад отбирают 4 детали. Случайная величина X равна числу нестандартных деталей среди отобранных. Найдите дисперсию и стандартное отклонение случайной величины X.

352. Страховой полис на один год стоит 8000 р. Страховые случаи делятся на 3 категории: с малым (10 000 р.), средним (50 000 р.) и большим (500 000 р.) ущербом. При этом каждый из этих случаев возникает в течение года с вероятностями 0,1, 0,05 и 0,002 соответственно. Найдите дисперсию и стандартное отклонение страховой выплаты по полису в течение года. Сделайте то же самое для дохода, который компания получает с одного полиса.

353. Профессор Злобин не хочет ставить положительную оценку студенту Разгуляеву и задаёт ему вопросы до тех пор, пока на очередном вопросе студент не ошибётся. Вероятность правильного ответа на любой вопрос составляет 0,8. Случайная величина X равна числу вопросов, на которые правильно ответит студент перед тем, как получит двойку. Найдите дисперсию и стандартное отклонение случайной величины X. Как вы думаете, какое количество вопросов следует запасти профессору?

354*. В ящике лежат 5 белых и 5 чёрных шаров. Из него наугад вытаскивают 4 шара. Случайная величина X равна числу белых шаров в полученной выборке. Найдите дисперсию и стандартное отклонение X.

355*. Монету бросают до тех пор, пока орёл и решка не выпадут хотя бы по разу. Найдите дисперсию и стандартное отклонение числа испытаний. Вычислите верхнюю границу числа испытаний по правилу «трёх сигм».

356*. Пусть N — выбранное наугад двузначное случайное число, а случайная величина X равна сумме его цифр. Найдите дисперсию и стандартное отклонение случайной величины X.

357*. Решите предыдущую задачу, если X — это наибольшая из двух цифр.

358*. При каком значении параметра p дисперсия бинарной случайной величины будет наибольшей? наименьшей?

В следующих задачах для вычисления математического ожидания представьте случайную величину как сумму других, более простых случайных величин.

359*. Сколько в среднем раз нужно бросить кубик, чтобы «увидеть» все его стороны (т. е. чтобы каждая из шести граней выпала хотя бы по разу)?

360*. Машина младшая сестра Ирина собирает фигурки динозавров, скрытые в шоколадных яйцах. Полная коллекция фигурок составляет 25 штук. Сколько в среднем шоколадных яиц нужно купить, чтобы собрать всю коллекцию?

361*. В предыдущей задаче Ирине подарили сразу k шоколадных яиц. Сколько разных фигурок в них можно ожидать (найдите математическое ожидание этой величины)?

362*. На первом этаже 11-этажного дома в лифт вошли 12 человек. Считая, что любой из них с равной вероятностью может выйти на любом из 10 этажей, найдите среднее число остановок лифта.

363*. Случайным образом независимо друг от друга разыгрываются n целых чисел из диапазона от 1 до n. Чему равно среднее количество полученных при этом различных чисел?

364*. В руке зажимают $2n$ травинок так, чтобы их концы торчали сверху и снизу. Верхние $2n$ концов случайным образом разбивают на пары и связывают каждую пару травинок между собой. Потом то же самое делают с нижними концами. Рука разжимается. Чему равно математическое ожидание числа получившихся колец?

365*. Шахматисты А и Б играют финальный матч за звание чемпиона до победы в двух партиях. Случайная величина X равняется количеству партий в финальном матче. Найдите её математическое ожидание, если любая партия может закончиться вничью с вероятностью 0,5, а А и Б имеют равные шансы на выигрыш.

Математическое ожидание и дисперсия*

Задание 1. Выигрыш в «Спортлото»

В предыдущей лабораторной работе мы рассматривали лотерею «Спортлото 6 из 49», проводившуюся в СССР несколько десятилетий тому назад. Напомним, что в этой лотерее нужно было угадать, какие 6 из 49 номеров выиграют в очередном тираже. Мы уже нашли закон распределения вероятностей случайной величины X, равной числу угаданных номеров. Напомним, что приз давался за 3 и больше угаданных номера. Денежный размер приза колебался, но обычно он составлял:

- 3 р. — за 3 угаданных номера;
- 50 р. — за 4 номера;
- 2500 р. — за 5 номеров;
- 10 000 р. — за 6 номеров.

Билет лотереи стоил 30 коп. Найдите математическое ожидание, дисперсию и стандартное отклонение выигрыша на один билет.

Этапы выполнения задания:

1. Заполните столбец A числами $k = 0, 1, \ldots, 6$.
2. Найдите в столбце B число благоприятных исходов для каждого из событий A_0, A_1, ... , A_6, где A_k = «будет угадано k номеров».
3. Вычислите в ячейке B8 количество всех возможных исходов лотереи «Спортлото 6 из 49», как сумму всех чисел в столбце B.
4. Найдите в столбце C вероятности событий A_0, A_1, ... , A_6. Для контроля вычислите в ячейке C8 их сумму.
5. Введите в столбец D размер денежного приза за каждое число угаданных номеров от 0 до 6.
6. Вычислите в ячейке D8 математическое ожидание выигрыша. Используйте для этого функцию СУММПРОИЗВ().
7. Найдите в столбце E квадраты значений из столбца D.
8. Вычислите в ячейке E8 математическое ожидание квадрата выигрыша.
9. Найдите в ячейке E9 дисперсию, а в ячейке E10 — стандартное отклонение выигрыша (рис. 107).

	A	B	C	D	E
1	0	6096454	0,435965	0	0
2	1	5775588	0,4130195	0	0
3	2	1851150	0,132378	0	0
4	3	246820	0,0176504	3	9
5	4	13545	0,0009686	50	2500
6	5	258	1,845E-05	2500	6250000
7	6	1	7,151E-05	10000	1E+08
8		13983816	1	0,148222	125,0434
9					125,0214
10					11,1813

Рис. 107. Пример выполнения задания 1

Задание 2. Математическое ожидание и дисперсия биномиального распределения

Постройте с помощью электронной таблицы биномиальное распределение вероятностей $\mathbf{Bin}(N, p)$ с $N = 100$ и произвольной вероятностью успеха p.

Найдите его математическое ожидание, дисперсию и стандартное отклонение. Вычислите границы интервала «трёх сигм».
Отразите полученные результаты на диаграмме вероятностей. Отметьте на ней математическое ожидание и интервал «трёх сигм».
Изучите поведение полученного графика при изменении вероятности успеха p.

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Число успехов**	**Вероятность**										
2	0	1,09951E-08	P	0,6		5,427329	12	18,57267				
3	1	3,29853E-07	E(X)	12	0,25	0	0	0				
4	2	4,70041E-06	D(X)	4,8								
5	3	4,23037E-05	σ	2,19089								
6	4	0,000269686										
7	5	0,001294494										
8	6	0,004854351										
9	7	0,014563052										
10	8	0,03549744										
11	9	0,070994879										
12	10	0,117141551										
13	11	0,159738478										
14	12	0,179705788										
15	13	0,165882266										
16	14	0,124411699										
17	15	0,07464702										
18	16	0,03499079										
19	17	0,012349691										
20	18	0,003087423										
21	19	0,000487488										
22	20	3,65616E-05										

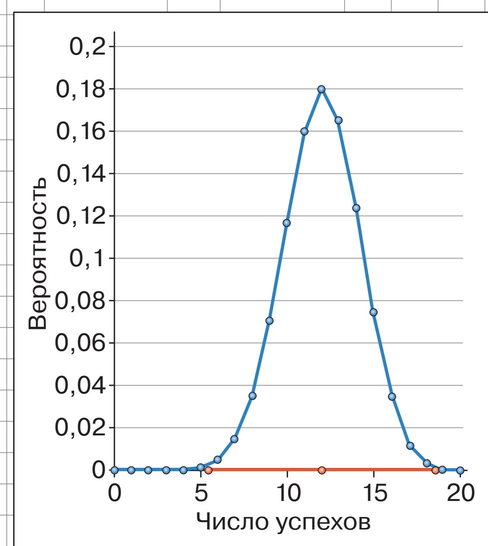


Рис. 108. Пример выполнения задания 2
($N = 20$, $p = 0{,}6$)

Задание 3. Математическое ожидание и дисперсия геометрического распределения

Проведите аналогичное исследование с геометрическим распределением.

Ответы

Глава 1. Представление данных и описательная статистика

8*. ABCEF, длина пути 9. **9*.** а) Команда из 4 человек: Давид, Ева, Кирилл, Элина; б) на 4 команды. **10*.** а) Семашко; б) 2; в) 4; г) Чиковани. **15.** С самого большого сектора по часовой стрелке: Тихий, Южный, Атлантический, Индийский, Северный Ледовитый. **25.** а) 13; б) нет моды. **26.** а) 38; б) 5,5. **27.** а) 30; б) 24; в) 13; г) 1,6. **28.** 177 и 7; 178 и 5. **29.** а) 3; б) –3. **30.** а) 26; б) 3. **31.** а) 10; б) 2,5; в) –3. **32.** Мода = 4, медиана = 4, среднее = 3,5. **33.** а) 1,9; б) 2; в) 2. **35.** а) Нет; б) нет; в) да; г) нет. **36.** а) Среднее = 125 000; мода = 100 000; медиана = 100 000. **39.** Среднее ≈ 2,83, мода = 3, медиана = 3; изменится только среднее, которое станет ≈ 3,17. **40.** 1) Среднее = 20, моды нет, медиана = 20; а) среднее = 30, моды нет, медиана = 30; б) среднее = 200, моды нет, медиана = 200. 2) а) Ко всем средним характеристикам прибавится это число; б) все средние характеристики умножатся на это число. **44.** а) 4; б) 11,3. **45.** а) 5,125; б) 5,5; в) 10. **46.** а) Среднее; б) медиана; в) мода; г) размах.

Глава 2. Элементы теории графов

67. а) Да; б) да; в) да; г) нет; д) нет; е) да. **68.** Нет. **69.** 12. **72.** а) и в). **73.** б) и в). **74.** 4; 6. **75.** Нет. **76.** Нет. **82.** а) 1; б) 6; в) 3. **83.** а) 3; б) 7; в) 2. **84.** а) — существует; б), в), г) — не существует. **85.** 19. **88.** Да. **91.** 2 и 4; 2 и 49; 2 и 2023. **92.** *n*. **94.** 5. **95.** 82. **96.** 30 000. **97.** 12. **99.** 6. **101.** а), б) — непланарные; в) — планарный. **102.** 9; 12.

Глава 3. Случайные события и их вероятности

103. $A = \{1;\ 2\}$, $B = \{1;\ 2;\ 3;\ 4;\ 5;\ 6\}$, $C = \{2;\ 3;\ 5\}$, $D = \{6\}$. **104.** $A = \{$ОР, ОО$\}$, $B = \{$ОР, РО$\}$, $C = \{$ОО$\}$, $D = \{$ОР, РО, ОО$\}$. **105.** 4; 9; 1; 18. **106.** 12. **107.** 3. **108.** Всего исходов — 8; A — 4; B — 0; C — 4; D — 5. **109.** 6. **110.** Во втором, пятом и шестом. **111.** $\frac{1}{5} = 0{,}2$. **112.** $\frac{19}{25} = 0{,}76$. **113.** $\frac{195}{200} = 0{,}975$. **114.** $\frac{85}{100} = 0{,}85$. **115.** 10. **116.** 0,9; 0,7. **117.** а) $\frac{3}{12} = \frac{1}{4} = 0{,}25$; б) $\frac{15}{40} = \frac{3}{8} = 0{,}375$; в) $\frac{475}{500} = \frac{19}{20} = 0{,}95$. **118.** $P(A) = \frac{4}{36} = \frac{1}{9}$; $P(B) = \frac{1}{4}$; $P(C) = \frac{1}{36}$; $P(D) = \frac{1}{9}$; $P(E) = \frac{8}{9}$; $P(F) = \frac{1}{2}$. **119.** а) $\frac{8}{30} = \frac{4}{15}; \frac{1}{15}; \frac{2}{3}$; б) 23; 29; 11. **120.** а) $\frac{45}{90} = \frac{1}{2} = 0{,}5$; б) $\frac{18}{90} = \frac{1}{5} = 0{,}2$; в) $\frac{21}{90} = \frac{7}{30} = 0{,}23(3)$; г) $\frac{6}{90} = \frac{1}{15} = 0{,}06(6)$. **121.** а) $\frac{2}{20} = \frac{1}{10}$; б) $\frac{3}{20}$; в) $\frac{10}{20} = \frac{1}{2}$. **122.** $P(A) = \frac{6}{36} = \frac{1}{6}$; $P(B) = \frac{30}{36} = \frac{5}{6}$; $P(C) = \frac{18}{36} = \frac{1}{2}$; $P(D) = \frac{5}{36}$; $P(E) = \frac{11}{36}$. **123.** $P(A) = \frac{1}{2}$; $P(B) = 0$; $P(C) = \frac{1}{2}$; $P(D) = \frac{5}{8}$. **124.** $\frac{1}{12}$; $\frac{1}{4}$. **125.** а) $\frac{648}{900} = \frac{18}{25} = 0{,}72$; б) $\frac{648}{900} = \frac{18}{25} = 0{,}72$; в) $\frac{252}{900} = \frac{7}{25} = 0{,}28$; г) $\frac{225}{900} = \frac{1}{4} = 0{,}25$. **126.** а) $\frac{1}{3}$; б) $\frac{2}{6} = \frac{1}{3}$; в) $\frac{4}{10} = \frac{2}{5}$. **127.** $\frac{2}{4} = \frac{1}{2}$. **128.** $\frac{2}{3}$. **129.** $\frac{2}{4} = \frac{1}{2}$. **130.** а) $\frac{4}{10} = \frac{2}{5}$; б) $\frac{2}{10} = \frac{1}{5}$; в) $\frac{6}{10} = \frac{3}{5}$. **131.** а) $\frac{1}{6}$; б) $\frac{2}{6} = \frac{1}{3}$; в) $\frac{3}{6} = \frac{1}{2}$; г) 0. **132.** а) $\frac{1}{6}$; б) $\frac{4}{6} = \frac{2}{3}$; в) $\frac{2}{6} = \frac{1}{3}$; г) $\frac{2}{6} = \frac{1}{3}$. **133*.** $\frac{7}{15}$; $\frac{8}{15}$. **134*.** $\frac{15}{36}$; $\frac{21}{36}$; 0. **135*.** При любом выборе первого игрока, второй игрок может выбрать кубик, который будет лучше: шансы первого и второго относятся как 15 : 21, второго и третьего — 15 : 21; третьего и первого — 11 : 25.

Глава 4. Сложение и умножение вероятностей

139. $B_0 = \overline{A_1}\,\overline{A_2}\,\overline{A_3}$; $B_1 = A_1\overline{A_2}\,\overline{A_3} + \overline{A_1}A_2\overline{A_3} + \overline{A_1}\,\overline{A_2}A_3$; $B_2 = A_1A_2\overline{A_3} + A_1\overline{A_2}A_3 + \overline{A_1}A_2A_3$; $B_3 = A_1A_2A_3$. **140.** $A = A_1A_2$; $B = \overline{A_1}\,\overline{A_2}$; $C = A_1A_2 + \overline{A_1}\,\overline{A_2}$; $D = A_1\overline{A_2} + \overline{A_1}A_2$. **141*.** а) Знатоки выиграют ровно одну игру; б) знатоки выиграют ровно две игры; в) знатоки выиграют хотя бы одну игру; г) знатоки выиграют первую, но проиграют вторую игру; д) все игры выиграют телезрители; е) знатоки выиграют все игры. **142*.** $A = A_1A_2 + \overline{A_1}\,\overline{A_2}$; $B = A_1\overline{A_2} + \overline{A_1}A_2$; $C = A_1 + A_2$; $D = \overline{A_1}\,\overline{A_2}$. **143*.** а) $A = A_1 + A_2 + A_3$; б) $A = A_1A_2A_3 + A_4A_5A_6$; в) $A = A_1(A_2 + A_3)A_4$; г) $(A_1 + A_2)(A_3 + A_4)$; д) $A = (A_1A_2 + A_3 + A_4)A_5$; е) $A = A_1A_3 + A_2A_4 + {}$ $+A_1A_5A_4 + A_2A_5A_3$. **144.** $P(\overline{A}) = 0{,}7; P(\overline{B}) = 0{,}6; P(A \cap B) = 0; P(A \cup B) = 0{,}7$. **145.** а) Да; б) да; в) нет; г) да. **148.** $P(A \cup B) = 0{,}6$; $P(\overline{A \cup B}) = 0{,}4$; $P(\overline{A} \cup \overline{B}) = 0{,}7$; $P(\overline{A} \cap \overline{B}) = 0{,}4$; $P(A \cap \overline{B}) = 0{,}1$; $P(\overline{A} \cup B) = 0{,}2$. **149.** а) Нет; б) да. **150.** а) 0,66; б) 0,94; в) 0,34; г) 0,24; д) 0,04. **151.** $P(A \cap B)$: наибольшее — 0,3, наименьшее — 0; $P(A \cup B)$: наибольшее — 0,7, наименьшее — 0,4. **152.** 83%. **153.** 0,09. **154.** 0,3. **155.** 0,623. **156.** 0,5. **157.** 0,95. **158.** От 0,1 до 0,5. **159.** 0,3. **160.** 0,6. **161.** 0,15; 0,65. **162.** $p_1 = 0{,}18$; $p_2 = 0{,}28$; $p_3 = 0{,}42$; $p_4 = 0{,}12$. **163.** 0,7; 0,6; 0,12; 0,58. **164.** а) $\frac{5}{51}$; б) $\frac{22}{51}$; в) $\frac{9}{17}$; г) $\frac{8}{17}$. **165.** $\frac{10}{21}$. **166.** $\frac{7}{40}$. **167.** $\frac{63}{125} = 0{,}504$. **168.** $\frac{105}{512} \approx 0{,}205$. **169.** ≈ 0,665. **170.** а) $\frac{2}{5}$; б) $\frac{1}{3}$. **171.** ≈ 0,531; ≈ 0,472. **172.** 0,271. **173.** $\frac{1}{36}$; $\frac{20}{36}$; $\frac{15}{36}$. **174*.** $P(B) = \frac{7}{12}$; $P(B \mid A) = \frac{1}{6}$ $\Rightarrow$ A и B зависимы; $P(C) = \frac{1}{2}; P(C \mid A) = \frac{1}{2}$ $\Rightarrow$ A и B независимы; $P(D) = \frac{3}{4}; P(D \mid A) = \frac{1}{2}$ $\Rightarrow$ A и B зависимы. **175*.** 0,384; 0,144. **176*.** Петров — 0,2; Иванов — 0,1. **177*.** ≈ 0,819. **178*.** 0,33. **179*.** 0,6. **180*.** $\frac{1}{5}$. **181*.** 2. **182*.** 0,8. **183*.** 0,125. **184.** а) $\frac{1}{3}$; б) $\frac{1}{3}$; в) $\frac{1}{3}$. **185.** 0,2952. **186.** ≈ 0,407. **187.** 0,077. **188.** 0,7625. **189*.** ≈ 0,553; ≈ 0,445. **190*.** 0,8. **191*.** 0,6. **192*.** ≈ 0,452; ≈ 0,315; ≈ 0,863. **193*.** 0,3; 0,7. **194*.** ≈ 0,165. **195*.** ≈ 0,152; ≈ 0,848. **196*.** ≈ 0,736; ≈ 0,225; ≈ 0,038; ≈ 0,0005. **197*.** 0,43.

Глава 5. Элементы комбинаторики

199. 440; 424. **200.** 10101100; 10000111. **201.** ОУУЫЭА; УУОЭЭЭ. **202.** 720. **203.** 1320. **204.** а) 120; б) 90; в) 60; г) 30 240; д) 3 628 800. **205.** 48. **206.** 780. **207.** 2048. **208.** 7776. **209.** 300. **210.** 24. **211.** 21. **212.** 1560. **213.** 36. **214.** 390 190 489 920; 109 736. **215.** 83 160; 7560; 15 120. **216.** а) 120; б) 60; в) 24; г) 6. **217*.** 72. **218*.** а) 848 800; б) 114 265; в) 1000. **219*.** $\frac{3^n - 2^n}{6^n}$; $\frac{4^n - 3^n}{6^n}$. **220*.** а) 2; б) 4; в) 16; г) 11; д) 70; е) 140. **221.** 220. **223.** а) 45; б) 10; в) 55; г) 252; д) 100. **224.** 54 600. **225.** а) 2; б) 5; в) 35; г) $\frac{N(N-1)}{2} - N$. **226.** а) 1; б) 10; в) 120; г) 1023; д) 252; е) 386. **227*.** а) 120; б) 126; в) 84; г) 21; д) 100; е) 386. **228*.** 7 054 320. **229*.** 1140; 171. **230*.** 13 484 225 986 560 000. **231*.** а) 2 363 904 400; б) 0. **232*.** а) 2^{n-1}; б) 2^{n-1}. **233*.** 35. **235.** а) $x^7 + 7x^6 + 21x^5 + 35x^4 + 35x^3 + 21x^2 + 7x + 1$; б) $x^6 + 12x^5 + 60x^4 + 160x^3 + 240x^2 + 192x + 64$; в) $81x^4 + 216x^3y + 216x^2y^2 + 96xy^3 + 16y^4$; г) $x^5 - 20x^4y + 160x^3y^2 - 640x^2y^3 + 1280xy^4 - 1024y^5$.

Глава 6. Испытания Бернулли. Случайный выбор

236. а) 0,7776; б) 0,05184; в) 0,02304. **237.** ≈ 0,904. **238.** а) ≈ 0,599; б) ≈ 0,0746; в) ≈ 0,988. **239.** а) ≈ 0,0067; б) ≈ 0,993. **240.** а) ≈ 0,328; б) 0,0064; в) ≈ 0,058. **241.** а) ≈ 0,196; б) ≈ 0,0073; в) ≈ 0,0199. **242.** а) $\frac{5}{16} = 0{,}3125$; б) 0,5; в) ≈ 0,412. **243.** ≈ 0,889. **244*.** ≈ 14,8%. **245.** ≈ 0,944. **246*.** ≈ 0,476. **247*.** 25. **248*.** $\frac{k}{n+1} < p < \frac{k+1}{n+1}$. **249*.** Первая вероятность больше второй: 0,518 > 0,491. **250*.** Не будут. Эти шансы разные: 0,665; 0,619; 0,597. **251*.** а) 7:1; б) 11:5; в) 13:3. **252*.** а) Три партии из шести; б) более четырёх партий из восьми; в) более $n + 1$ партий из $2n + 2$. **253*.** 0,5. **254*.** а) ≈ 0,065; б) ≈ 0,807; в) 0. **256*.** $\frac{C_{2n-k}^{n-k}}{2^{2n-k}}$. **258.** а) $\frac{1}{8} = 0{,}125$; б) $\frac{15}{16} = 0{,}9375$; в) $\frac{7}{32} \approx 0{,}219$. **259.** а) ≈ 0,579; б) ≈ 0,518; в) ≈ 0,293. **260.** 0,84. **261.** 0,8704. **262.** а) 0,51; б) 0,343; в) ≈ 0,412. **263.** а) 0,8; б) 0,2; в) 0,04. **264.** ≈ 0,672. **265.** 4; 7; 10. **266.** 13; 26; 38. **267.** а) 3; б) 6; в) 9. **268.** 7. **269*.** 253. **270*.** 14. **271*.** 21; 31. **272*.** а) 28; б) 29; в) 32. **273*.** 4. **274*.** а) Шансы одинаковые; б) у Пети шансы на выигрыш $\frac{3}{4}$, у Васи — $\frac{1}{4}$. **275*.** В одной партии с вероятностью $\frac{3}{16}$ выигрывает Артём; с вероятностью $\frac{8}{16}$ — Борис, с вероятностью $\frac{5}{16}$ — ничья. Во всей игре с вероятностью $\frac{3}{11}$ выигрывает Артём, с вероятностью $\frac{8}{11}$ — Борис. **276*.** Можно, например, бросить кнопку два раза подряд. Если кнопка первый раз выпадет остриём вверх, а второй раз — вниз, то мыть посуду будет первый; если наоборот — то второй. А если оба раза кнопка выпадет остриём вверх или оба раза остриём вниз, то жребий нужно будет перебросить. **277*.** 0,25. **278.** а) $\frac{2}{3}$; б) $\frac{2}{3}$. **279.** а) $\frac{2}{5}$; б) $\frac{1}{2}$. **280.** а) ≈ 0,376; б) ≈ 0,00021; в) ≈ 0,624; г) ≈ 0,149. **281.** а) $\frac{4}{7}$; б) $\frac{1}{7}$. **282.** $\frac{49}{153} \approx 0{,}32$. **283.** Да могут, если $a = \frac{n(n-1)}{2}$, $b = \frac{n(n+1)}{2}$ или, наоборот, (n = 2, 3, 4, ...). **284.** $\frac{5}{6} = 0{,}83(3)$. **285.** 0,4. **286.** $\frac{1}{120}$. **287.** ≈ 0,395; ≈ 0,634. **288.** а) 0,083(3); б) 0,416(6). **289.** а) ≈ 0,0082; б) ≈ 0,00082; в) 0,04; г) ≈ 0,0367. **290.** а) $\frac{1}{n^3}$; б) $\frac{6}{n^3}$; в) $\frac{1}{n^{k-1}}$; г) $\frac{A_n^k}{n^k}$ при $k \leqslant n$, 0 при $k > n$. **291.** а) $\frac{7^6}{10^6} \approx 0{,}118$; б) $\frac{6^6}{10^6} \approx 0{,}0467$; в) $\frac{10}{10^6} = 0{,}00001$; г) $\frac{189}{1250} = 0{,}1512$; д) $\frac{27}{2500} = 0{,}0108$. **292.** $\frac{1}{3}$. **293.** $\frac{1}{6}$; а) $\frac{1}{5040} \approx 0{,}0002$; б) и в) $\frac{1}{60} = 0{,}016(6)$; г) ≈ $1{,}2 \cdot 10^{-11}$. **294.** $\frac{15}{24} = \frac{5}{8} = 0{,}625$. **295.** а) ≈ $2{,}8 \cdot 10^{-7}$; б) ≈ 0,368; в) ≈ 0,632; г) ≈ 0,368. **296.** $\frac{8}{15} = 0{,}53(3)$. **297.** а) 23; б) 41; в) 57; г) 367. **298.** $\frac{89}{144} \approx 0{,}618$; $\frac{9}{64} \approx 0{,}141$. **299.** Шансы выше с возвращением: 0,125 > 0,08(3). **300.** При $k = 0$ и $k = 10$ шансы одинаковые; при остальных k шансы для выбора с возвращением выше. **301.** $\frac{256}{46189} \approx 0{,}00554$. **302.** а) $\frac{256}{46189} \approx 0{,}00554$; б) 0; в) $\frac{3150}{46\,189} \approx 0{,}0682$. **303.** $\frac{1}{n+1}$.

Глава 7. Случайные величины и распределения

307. $S \in [3; 18]$, $K \in [1; 3]$, $M \in [1; 6]$, $N \in [1; 6]$, $L \in [1; 6]$. **310.** От 0 до N; $Y = N - X$. **311.** 0. **312.** От 3 до 5. **313.** От 6 до ∞. **314*.** От 1 до 4; 1 — $\frac{1}{216}$; 2 — $\frac{35}{216}$; 3 — $\frac{5}{9}$; 4 — $\frac{5}{18}$. **315*.** От 0 до 75. Самое вероятное значение — 0. **316.** а) $p = 0{,}18$; б) $p = \frac{1}{2}$. **317.** а) 0,4; б) 0,9; в) 0,4; г) 0; д) 0,9.

318.

а)

Значения $\lvert X \rvert$	0	1	10
Вероятность	0,1	0,3	0,6

б)

Значения X^2	0	1	100
Вероятность	0,1	0,3	0,6

319. Биномиальное распределение **Bin**(4; 0,1):

Значения X	0	1	2	3	4
Вероятность	0,6561	0,2916	0,0486	0,0036	0,0001

320. Геометрическое распределение **Geom**(0,2):

Значения X	1	2	3	4	5	...
Вероятность	0,2	0,16	0,128	0,1024	0,08192	...

321*.

Значения T	24	32	40
Вероятность	0,216	0,2592	0,5248

322*.

Значения X	0	1	2	3	4
Вероятность	0,0238	0,2381	0,4762	0,2381	0,0238

323*.

Значения X	0	1	3
Вероятность	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

324*.

Значения *X*	0	1	2	4
Вероятность	$\frac{3}{8}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{24}$

325*.

Значения *X*	1	2	3	...	8	9	10	11	...	16	17	18
Вероятность	$\frac{1}{90}$	$\frac{2}{90}$	$\frac{3}{90}$	...	$\frac{8}{90}$	$\frac{9}{90}$	$\frac{9}{90}$	$\frac{8}{90}$	...	$\frac{3}{90}$	$\frac{2}{90}$	$\frac{1}{90}$

326*.

Значения *X*	1	2	3	4	5	6	7	8	9
Вероятность	$\frac{2}{90}$	$\frac{4}{90}$	$\frac{6}{90}$	$\frac{8}{90}$	$\frac{10}{90}$	$\frac{12}{90}$	$\frac{14}{90}$	$\frac{16}{90}$	$\frac{18}{90}$

327*.

Значения *X*	3	4	5
Вероятность	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$

328*.

Значения *X*	2	3	4	5	6	...	k	...
Вероятность	$\frac{2}{16}$	$\frac{6}{32}$	$\frac{12}{64}$	$\frac{20}{128}$	$\frac{30}{256}$	...	$\frac{k(k-1)}{2^{k+2}}$	...

329*.

Значения *X*	1	2	3
Вероятность	$\frac{1}{36}$	$\frac{5}{12}$	$\frac{5}{9}$

330*.

Значения *W*	1	2	3	4	5	6
Вероятность	$\frac{1}{6^n}$	$\frac{2^n-1}{6^n}$	$\frac{3^n-2^n}{6^n}$	$\frac{4^n-3^n}{6^n}$	$\frac{5^n-4^n}{6^n}$	$\frac{6^n-5^n}{6^n}$

331. а) 3,5; б) $\frac{3}{8}$; в) 11. **332.** 6,3; 60,3. **333.** 3500 р. **334.** 0,4. **335.** 5. **336*.** $\approx$ 34,47. **337*.** 2. **338*.** 1. **339*.** 1. **340*.** 9,5. **341*.** $\frac{19}{3} = 6{,}3(3)$. **342*.** а) $\frac{33}{8} = 4{,}125$; б) 3,6336.

344*. $6-\left(\frac{1}{6}\right)^n-\left(\frac{2}{6}\right)^n-\left(\frac{3}{6}\right)^n-\left(\frac{4}{6}\right)^n-\left(\frac{5}{6}\right)^n$. **345*.** 1. **346*.** $\frac{150}{19}\approx 7{,}9$. **347*.** $\frac{n(3n-1)}{2n-1}$. **348*.** $N(1-(1-p)^k+\frac{N}{k}$. Оно будет меньше N при $k(1-p)^k > 1$. **349.** а) 2,45 и 1,57; б) 0,984 и 0,992; в) 14 и 3,742. **350.** 20,61 и 4,54; 2364,21 и 48,62. **351.** 0,36 и 0,6. **352.** Для выплаты дисперсия ≈ 615, стандартное отклонение ≈ 24,8. Для дохода дисперсия и стандартное отклонение будут такими же. **353.** 5 и $\sqrt{5}\approx 2{,}24$. С учётом интервала «трёх сигм» следует запасти $5 + 3 \cdot 2{,}24 \approx 12$ вопросов. **354*.** $\frac{2}{3}$ и $\frac{\sqrt{2}}{\sqrt{3}}\approx 0{,}816$. **355*.** $E(X) = 3$; $D(X) = 2$; $s(X) = \sqrt{2}\approx 1{,}41$. Верхняя граница по правилу «трёх сигм» — $3 + 3\sqrt{2} \approx 7{,}24$. **356*.** 14,9 и 3,86. **357*.** 4,9 и 2,21. **358*.** Наибольшей — при $p = 0{,}5$; наименьшей — при $p = 0$ и $p = 1$. **359*.** 14,7. **360*.** ≈ 95,4. **361*.** ≈ 16. **362*.** ≈ 7,2. **363*.** $n\left(1-\left(\frac{n-1}{n}\right)^n\right)$. **364*.** $1+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{2n-1}$.

Предметный указатель

Оглавление