СТАТИСТИКА

Первая книга

Alick Hartley

STATISTICS

Book One

AJANTA

IB **IMPART BOOKS**
invite you to visit their web site at
http://www.books.mid-wales.net/index.html

Алик Жартли

СТАТИСТИКА

Первая книга

Учебно-методическое пособие

Москва
«Финансы и статистика»
2004

УДК 311(075)
ББК 60.6я7
Х23

Перевод с английского
Н.В. Мамаевой и **М.Ю. Меженного**

Под редакцией
доктора экономических наук, профессора
О.Э. Башиной

Хартли Алик

Х23 Статистика. Первая книга: Учеб.-метод. пособие: Пер. с англ.; Под ред. О.Э. Башиной. – М.: Финансы и статистика, 2004. – 312 с.: ил. – Пер. изд.: Alick Hartley. Statistics. Book One. IMPART BOOKS, Gwelfryn, Great Britain, 1996, Revised Edition AJANTA BOOKS INTERNATIONAL, Delhi, India, 1998.

ISBN 5-279-02736-7

Приведена самая необходимая информация по методам и приемам статистики. Материал представлен с привлечением реальных примеров. Для наглядности и повышения степени усвоения широко использованы графики, таблицы и тесты. Часть цифрового материала при подготовке русского издания обновлена.

Для студентов колледжей, высших и средних учебных заведений гуманитарного профиля, предпринимателей, не имеющих базового экономического образования, а также для слушателей курсов повышения квалификации и студентов, обучающихся по специальностям «Финансы и кредит», «Бухгалтерский учет, анализ и аудит», «Коммерция».

Х $\frac{0702000000 - 119}{010(01) - 2004}$ 314 – 2004

УДК 311 (075)
ББК 60.6я7

ISBN 81-202-0308-9
ISBN 5-279-02736-7

К ЧИТАТЕЛЮ

Вашему вниманию предлагается книга английского ученого, статистика Алика Хартли «Статистика. Книга первая». В ней содержится самая основная и необходимая информация, которую должен знать человек, желающий изучить и пользоваться приемами и методами статистики.

Большой опыт работы Алика Хартли в качестве главного экзаменатора по статистике в ряде международных и региональных экзаменационных комиссий в Великобритании, безусловно, отразился на стиле и содержании книги.

Материал изложен необычайно просто, легок для запоминания и усвоения. Для наглядного изложения материала используются наиболее распространенные методы: графический, табличный и метод тестового запоминания. В пособии содержится много примеров, связанных с учебным процессом в школе и вузе, с бытовыми проблемами. Нетрадиционным для теоретической российской методологии является использование примеров с игральными картами, бросанием монеты и игральных костей.

Пособие предназначено для студентов колледжей, техникумов, высших учебных заведений гуманитарного профиля, предпринимателей, не имеющих базового экономического образования, слушателей курсов повышения квалификации. Им также могут воспользоваться и студенты высших учебных заведений экономического профиля для усвоения статистической методологии по основным дидактическим единицам, представленным по специальностям «Финансы и кредит», «Бухгалтерский учет, анализ и аудит», «Экономика и управление на предприятии» в Государственном образовательном стандарте второго поколения.

Доктор экономических наук,
профессор *О.Э. Башина*


A
B
C
D
E
?

ВВЕДЕНИЕ

Содержание этой книги ограничено самой основной и необходимой информацией по статистике, которую должны знать обучающиеся. В книге также представлены соответствующие примеры и упражнения для того, чтобы студенты могли проверить свои знания по каждой теме.

В заключение каждого раздела даются два параллельных упражнения с цифровыми ответами.

Некоторые темы можно использовать в качестве «визуального тестирования» на занятиях. По этим темам даются тесты для проверки в классе. Перед каждым тестом обучающимся раздаются карточки с большими буквами «A», «B», «C», «D», «E» и знаком вопроса «?». Для каждого вопроса предлагаются четыре или пять возможных ответов, а каждый ответ соответствует одной букве A, B, C, D или E.

Учащийся должен обдумать ответ на вопрос, поставленный преподавателем. Затем по команде преподавателя каждый должен поднять букву в соответствии с тем решением, которое считает правильным. В случае если он считает, что правильный ответ не представлен ни одной из букв A, B, C, D, E, или если учащемуся требуется подсказка или помощь преподавателя, он должен поднять вопросительный знак.

Буквы A, B, C, D, E и символ «?» для удобства могут быть нарисованы на большом картонном шестиугольнике так, как показано на рисунке на с. 6.

Комментарии автора относительно использования монет, игральных костей и карт в теории вероятностных методов

Автор хорошо понимает, что многие преподаватели и студенты по религиозным соображениям выступают против азартных игр на деньги, с использованием монет, игральных костей или карт. Такого же мнения придерживаются автор и многие экзаменаторы.

Тем не менее автор надеется, что преподаватели и студенты правильно воспримут использование им этих предметов в процессе изучения теории вероятности. Эти предметы позволяют проще и легче понять методы получения полностью случайных результатов, «абсолютно случайных» результатов.

Использование этих предметов для облегчения изучения и понимания теории вероятностей не имеет целью потворствовать или поощрять их применение в азартных играх.

Подробности, перечисленные ниже, даются специально для того, чтобы помочь тем, кто по различным причинам, упомянутым выше, не знаком с монетами, игральными костями и картами.

(a) Бросание монеты
Здесь есть два равновероятных результата – орел или решка.

(b) Обычная шестигранная игральная кость («кубик» – наиболее распространенное название)
Здесь при бросании возможны шесть равнозначно вероятных результатов: 1, 2, 3, 4, 5 и 6.
Иногда кубик может быть размечен иным образом. В этом случае в примере или вопросе будет дана информация.

(c) Стандартная колода из 52 карт
Полная колода состоит из 52 карт. Каждая карта имеет «масть» и «номинал».
Есть четыре масти – две красные и две черные:
черные: Трефы и Пики
красные: Бубны и Червы.
В каждой масти 13 номиналов:
изображения на картах (картинки):
Король, Дама и Валет;
нечетные очки: 1, или Туз, 3, 5, 7 и 9;
четные очки: 2, 4, 6, 8 и 10.

ОГЛАВЛЕНИЕ

Глава 1

ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ

1.1. Цель графического представления

Цель, или назначение, графического представления данных состоит в том, чтобы получить наиболее понятное ви́дение всей картины с первого взгляда, без необходимости рассмотрения и сопоставления большого числа данных.

Существует множество методов графического представления:
(a) условные обозначения, или пиктограммы,
(b) круговые диаграммы,
(c) столбиковые диаграммы (простые, двойные и составные),
(d) линейные диаграммы,
(e) частотные полигоны,
(f) гистограммы,
(g) диаграммы кумулятивных частот.

Метод, который мы используем, зависит от типа переменной.

1.2. Типы переменных

Переменные, или, можно сказать, любые меры, которые изменяются, классифицируются двумя различными способами:

(a) количественные или качественные,

(b) одномоментные или непрерывные.

Разницу между количеством и качеством выражает отличие между описанием на уроке родного языка, например «очень высокая девочка», и описанием на уроке математики, например «девочка ростом 195 см».

Качественные переменные описывают качество, в то время как количественные – нечто, что можно выразить в цифрах.

Одномоментными переменными называются переменные, которые существуют только в отдельных классах или величинах, например размер рубашки или обуви. Обувь можно купить только конечного количества различных размеров (таких как 4, 5, 5½, 6 и т.д.). Непрерывные переменные существуют в неограниченном количестве различных величин. Например, длина ступни ребенка проходит через неопределенное количество различных величин, поскольку длина ступни увеличивается от 10 до 25 см.

Таким образом, существуют четыре различные категории типов переменных. Примеры этих четырех категорий представлены в таблице.

Примеры четырех типов переменных

	Одномоментные	Непрерывные
Качественные	Сборка автобуса Имя водителя автобуса	Цвет автобуса
Количественные	Количество колес автобуса Количество сидений в автобусе Количество дверей в автобусе	Длина автобуса Скорость автобуса «Возраст» автобуса

Аудиторный тест № 1

Решите, какой из возможных ответов правильный. Выберите букву A, B, C, D, используя следующие коды:

A означает качество и одномоментность,
B означает качество и непрерывность,
C означает количество и одномоментность,
D означает количество и непрерывность.

По просьбе преподавателя поднимите вверх карточку с буквой так, чтобы преподаватель мог видеть ваш ответ. Если вы не знаете ответа, поднимите символ «?».

Номер вопроса	Переменная
1.	Количество учеников в классе
2.	Точный возраст ученика
3.	Имя ученика
4.	Длина ступни ученика
5.	Размер обуви ученика
6.	Цвет обуви ученика
7.	Число полных лет ученика
8.	Рост ученика
9.	Возраст ученика до ближайшего дня рождения
10.	Масса (вес) ученика
11.	Название газеты
12.	Количество страниц в газете
13.	Высота страницы в газете
14.	Число строк на странице газеты
15.	Стоимость газеты
16.	Масса (вес) газеты
17.	Толщина страницы газеты
18.	Точное время
19.	Название школы
20.	Количество учеников в школе

1.3. Условные обозначения, или пиктограммы

В условном обозначении, или пиктограмме, небольшой символ представляет как один, так и несколько предметов.

Пример 1. В классе могут быть 12 мальчиков и 18 девочек. Разделение класса в этом случае можно проиллюстрировать следующим образом:

Мальчики ☺☺☺☺☺☺☺☺☺☺☺☺
Девочки ☺☺☺☺☺☺☺☺☺☺☺☺☺☺☺☺☺☺

Ключ: символ ☺ представляет одного учащегося

Заметим, что в данном случае один символ используется для представления как мальчиков, так и девочек.

Пример 2. Число обучающихся в старших классах элитной школы [Здесь и далее «Первый год обучения» соответствует 7-му классу, «второй год» – 8-му и т.д. – *Прим. науч. ред. перев.*].

Первый год обучения	▲▲▲▲▲▲▲▲▲▲▲▲▲
Второй год обучения	▲▲▲▲▲▲▲▲▲▲▲
Третий год обучения	▲▲▲▲▲▲▲▲▲▲▲▲
Четвертый год обучения	▲▲▲▲▲▲▲▲▲▲
Пятый год обучения	▲▲▲▲▲▲▲▲

Ключ: символ ▲ представляет 10 обучающихся

Упражнение 1

1. Ниже приведены сведения о распределении работников крупной нефтяной компании в 1985 г. по странам и регионам:

Глазго (Шотландия)	1548
Абердин (Шотландия)	1120
Лондон (Англия)	36
США	39
Индонезия	11
Нидерланды	5

Источник: «Отчет Бритойл о занятом персонале 1985».

Проиллюстрируйте данные с помощью пиктограммы.

2. Ниже приведены данные о ежедневном производстве нефти в 1985 г. (баррель/день) на восьми основных месторождениях:

Тистл	78 300
Беатрис	52 800
Данлин	52 300
Childиан	234 200
Мурчисон	101 000
Статфиорд	595 800
Сент-Брэ	92 700
Хаттон	58 100

Источник: «Годовой отчет Бритойл 1985».

Проиллюстрируйте данные с помощью пиктограммы.

3. Приводится перечень курсов и их посещаемость студентами в определенном колледже:

Искусство	36
Искусство и наука	42
Наука	58
Техника	24
Коммерция	30

Источник: Условные данные.

Проиллюстрируйте данные с помощью пиктограммы.

1.4. Круговые диаграммы

Круговая диаграмма состоит из круговой линии, разделенной внутри на сектора так, чтобы площадь каждого сектора была пропорциональна числовому значению, которое она представляет.

Пример 1. Расписание для ученика включает 40 уроков, которые распределяются следующим образом: 6 – Математика, 5 – Английский язык, 5 – Французский язык, 4 – География, 4 – История, 5 – Физика, 4 – Химия, 21 – Статистика, 2 – Физкультура и 3 – Самостоятельное обучение.

Проиллюстрируйте указанные данные с помощью круговой диаграммы.

Решение. Первый шаг состоит в подсчете и записи углов секторов. Вся окружность составляет 360°, таким образом, 40 уроков соответствуют 360°, что означает, что один урок представлен 9°, два урока – 18° и т. д. Мы записываем эту информацию в таблице следующим образом:

Предмет	Число уроков	Угол сектора, град.
Математика	6	54
Английский язык	5	45
Французский язык	5	45
География	4	36
История	4	36
Физика	5	45
Химия	4	36
Статистика	2	18
Физкультура	2	18
Самостоятельное обучение	3	27
Всего	40	360

Источник: Расписание Жана Филиппа.

Примечание. Всегда рекомендуется проверить, что сумма всех углов равняется 360°.

Теперь можно начертить круговую диаграмму с использованием круга удобного радиуса.

Уроки Жана Филиппа за неделю.

Источник: Расписание Жана Филиппа.

Примечания.

(i) Каждый сектор помечен названием того предмета, которому он соответствует;

(ii) На самой круговой диаграмме не пишутся ни число уроков, ни угол сектора;

(iii) По причинам, которые будут изложены в разделе 1.17, сектора не должны изображаться разными цветами;

(iv) Круговая диаграмма должна иметь заголовок, с тем чтобы читатель мог понять, что она иллюстрирует;

(v) По возможности после диаграммы следует указать источник информации.

В примере 1 расчет углов сектора был довольно простым, поскольку один урок был представлен одним точным значением градусов. Но так бывает не всегда.

Пример 2. В школе «Сьюпербиа» было 424 ученика, 96 из них учились первый год, 101 – второй, 87 – третий, 93 – четвертый и 47 – пятый год (см. примечание к Примеру 1. – *Ред. перев.*).

Проиллюстрируйте указанные данные с помощью круговой диаграммы.

Решение. 424 ученика представлены 360 градусами, один ученик представлен 360 : 424 градусами, 96 учеников представлены $360 \times 96 : 424$ градусами и т. д.

Снова записываем всю информацию в таблицу.

Просматривая расчеты, мы видим, что когда мы складываем значения углов секторов, представленные с точностью до двух знаков после запятой, общая сумма составляет 360°, как и следует ожидать. Но когда мы складываем углы секторов, округленные до ближайшего целого градуса, сумма оказывается равной 361°. Это иногда бывает, если, чисто случайно, значения всех углов округляются в бульшую сторону. Нам необходимо использовать округление и в меньшую сторону. В настоящем примере значение 81,51 является значением угла, которое менее всех заслуживает округления в бульшую сторону, так что мы округляем его в меньшую сторону – до 81. Сумма углов, округленных до ближайшего целого, теперь составит 360°. Поэтому рекомендуется вначале вычислять достаточно точные значения углов, а уже потом округлять эти значения до ближайшего целого.

Теперь мы можем построить круговую диаграмму, используя значения углов с точностью до градуса.

Год обучения в школе	Число учеников	Угол сектора (с точностью до градуса)
Первый	96	360 × 96 : 424 = 81,51º = 82º
Второй	101	360 × 101 : 424 = 85,75º = 86º
Третий	87	360 × 87 : 424 = 73,87º = 74º
Четвертый	93	360 × 93 : 424 = 78,96º = 79º
Пятый	47	360 × 47 : 424 = 39,91º = 40º
Всего	424	360,00º = 361º

Пример 3. Приведенная круговая диаграмма иллюстрирует распределение 36 уроков Мари Селест, ученицы другой школы.

Измерим и запишем углы секторов на круговой диаграмме (с точностью до градуса). Затем рассчитаем, сколько уроков у Мари по каждому предмету.

Решение. Значения углов (с точностью до градуса) составят:

Математика	60°
Английский язык	50°
Французский язык	50°
Искусство	40°
Экономика	40°

Продолжение

История	40°
Физкультура	20°
Самостоятельное обучение	60°
Всего	360°

Примечание. Всегда проверяйте сумму углов, которая должна составлять 360°.

Теперь мы можем подсчитать число уроков Мари по каждому предмету.

360° представляет 36 уроков,
таким образом, 10° представляет 1 урок,
итак, 60° (математика) представляет 6 уроков.

Подобным образом осуществляем расчет по другим предметам:

Предмет	Число уроков
Математика	6
Английский язык	5
Французский язык	5
Искусство	4
Экономика	4
История	4
Физкультура	2
Самостоятельное обучение	6
Всего	36

И опять нам необходимо удостовериться, что сумма уроков составляет 36, что является общим числом уроков в расписании Мари.

Упражнение 2А *(ответы приведены в конце книги)*

1. В таблице показано, как девочка тратит свои карманные деньги (в произвольных, условных единицах).

Переезды	2,60
Одежда	2,40
Конфеты	1,20
Развлечения	2,00

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы радиусом 5 см.

2. В приведенной ниже таблице показано (в млн МВт), что происходит с 4 000 000 миллионами мегаватт солнечной энергии, которая в день доходит до Земли.

Поглощенная, а затем переизлученная	2 000 000
Потребленная при круговороте воды	790 000
Сразу отраженная	1 200 000
Пошедшая на образование ветров и волн	10 000

Проиллюстрируйте указанную выше информацию с помощью круговой диаграммы, отмечая секторы «Поглощенная/Излученная», «Круговорот воды», «Отраженная», «Ветер/Волны». (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

3. На условную дату население Земли приблизительно составляло (в млн чел.):

Африка	424
Восточная Азия	1 037
Европа	478
Латинская Америка	324
Северная Америка	242
Южная Азия и Океания	1 340
Россия	143

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

4. Мировое производство энергии (уголь, нефть и т.д.) в условный год было следующим (в %):

Уголь	34
Гидроэлектро- и атомная энергия	2
Природный газ	19
Нефть	45

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

5. В определенном году производство шерсти в мире было следующим (в %):

Аргентина	6
Австралия	27
Новая Зеландия	12
ЮАР	4
Россия	8
Остальной мир	43

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

Упражнение 2В

1. Ниже приводятся данные об акциях различных компаний США, которыми владеет некий американский инвестор (в долл.):

«Боинг»	2 000
«Крайслер»	5 000
«Кока-Кола»	3 000
«Дженерал Моторз»	4 000
«ИБМ»	3 000
«Пан Ам»	1 000

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы радиусом 5 см. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

2. Минимальная дневная потребность человека в еде приблизительно составляет 7150 единиц, распределяемых следующим образом:

Углеводы	550
Жиры	300
Минералы	50
Кислород	1 650
Протеин	150
Вода	4 450

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы радиусом 5 см. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

3. Ежедневный прирост численности населения мира примерно составляет (в тыс. чел.):

Африка	30
Восточная Азия	45
Европа	15
Латинская Америка	25
Северная Америка	5
Южная Азия и Океания	95

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы радиусом 5 см. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

4. Мировая потребность в продовольствии складывается из следующих продуктов (в %):

Мясо и рыба	11
Масла и жиры	9
Рис	21
Фрукты, овощи и орехи	10
Пшеница	20
Другие источники	29

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

5. Ниже приводится распределение использования серебра в мире (в %):

Ювелирные изделия	20
Электротехника	20
Промышленность	25
Фотография	30
Прочее	5

Проиллюстрируйте информацию, представленную в таблице, с помощью круговой диаграммы. (Прежде чем строить круговую диаграмму, составьте таблицу углов секторов.)

1.5. Две круговые диаграммы сравнения

Для отображения двух сопоставимых явлений нам может потребоваться использование двух круговых диаграмм.

Пример. Ниже приведены типы курсов, посещаемые студентами-старшекурсниками двух схожих колледжей или школ.

Тип курса	Школа A	Школа B
Только искусство	10	13
Искусство/Наука	9	7
Только наука	6	16
Всего (все курсы)	25	36

Проиллюстрируйте данные о посещаемости с помощью двух круговых диаграмм сравнения.

Решение. Когда мы рисуем две круговые диаграммы сравнения, площадь каждой диаграммы оказывается пропорциональна числу или количеству, представленному этой диаграммой.

Число, представляемое первой диаграммой (n), ~ Площади первой диаграммы = πr^2.

Число, представляемое второй диаграммой (N), ~ Площади второй диаграммы = πR^2.

Таким образом, $\dfrac{\pi r^2}{\pi R^2} = \dfrac{n}{N}$.

Следовательно, $\dfrac{r^2}{R^2} = \dfrac{n}{N}$.

Отсюда $\dfrac{r}{R} = \sqrt{\dfrac{n}{N}}$.

В данном примере: $\dfrac{\text{Радиус для } A}{\text{Радиус для } B} = \sqrt{\dfrac{25}{36}} = \dfrac{5}{6}$.

Предположим, что мы выбрали радиус круговой диаграммы по школе A равным 3 см, тогда радиус для круговой диаграммы на школе B должен быть равен 3 см $\times$ 6/5 = 3,6 см.

Рассчитаем углы секторов для каждой круговой диаграммы, как было изложено выше в разделе 1.4.

Тип курса	Школа A	Школа B
Только искусство	144°	130°
Искусство/Наука	130°	70°
Только наука	86°	160°

Теперь мы можем изобразить две круговые диаграммы сравнения.

Школа A

Школа B

Упражнение 3А *(ответы даны в конце книги)*

1. Общее количество растительного материала на суше ежедневно увеличивается приблизительно на 270 млн т.

Сумма распределяется следующим образом (в млн т):

Культивируемые земли	25
Леса	175
Луга и пастбища	40
Лесистая местность	10
Прочие территории	20

Общее количество растительного материала в океане ежедневно увеличивается на 150 млн т.

Эта сумма распределяется следующим образом (в млн т):

Континентальный шельф	25
Открытые моря	115
Рифы и устья рек	10

Проиллюстрируйте приведенные выше данные с помощью двух круговых диаграмм сравнения.

2. Круговая диаграмма иллюстрирует мировое производство чая в условном году.

(i) Измерьте и запишите углы секторов с точностью до градуса. Общее производство чайной продукции в данном году составило 1 080 000 т.
(ii) Рассчитайте производство чая в данном году по каждому из четырех регионов.
(iii) Измерьте и запишите радиус круговой диаграммы с точностью до 0,1 см.

В предыдущем году производство чая составило всего 800 000 т.

(iv) Рассчитайте радиус круговой диаграммы сравнения, которая представляла бы мировое производство чая в предыдущем году.
(v) Используя данные, приведенные в таблице, постройте круговую диаграмму сравнения для предыдущего года (в т).

Цейлон	200 000
Индия	330 000
Япония	90 000
Остальной мир	180 000

(vi) Сравнив круговые диаграммы, прокомментируйте изменение в производстве чая в данном году по сравнению с предыдущим.

Упражнение 3В

1. Годовой доход финансового сектора крупного города в 1984 г. составил в целом 720 млн ф.ст. Вклад разных видов деятельности отображен на диаграмме:

(i) Измерьте и запишите с точностью до 0,1 см радиус круговой диаграммы.
(ii) Измерьте и запишите углы пяти секторов с точностью до ближайшего целого числа. Теперь рассчитайте вклад каждого вида деятельности.

В 1985 г. доход от тех же пяти видов деятельности выглядел следующим образом (в млн ф.ст.):

Банки	180
Брокерские операции	160
Страхование	380
Торговля	140
Прочие операции	60

(iii) Удостоверьтесь, что общий доход в 1985 г. составил 920 млн ф.ст.
(iv) Рассчитайте радиус круговой диаграммы, сравниваемой с приведенной выше круговой диаграммой.
(v) Рассчитайте с точностью до градуса и запишите в таблице углы секторов круговой диаграммы для 1985 г.
(vi) Затем нарисуйте круговую диаграмму сравнения для того, чтобы проиллюстрировать доход в 1985 г.

2. Круговая диаграмма иллюстрирует число учеников в школе из 900 учащихся, которые ходят в школу пешком, ездят на велосипеде или на автобусе.

(i) Измерьте и запишите углы (с точностью до градуса) по каждому сектору.

(ii) Затем рассчитайте число учеников школы, которые:

(a) ходят в школу пешком,

(b) ездят на велосипеде,

(c) ездят на автобусе.

(iii) Измерьте и запишите с точностью до 0,1 см радиус круговой диаграммы.

В другой школе 1200 учеников.

(iv) Рассчитайте с точностью до 0,1 см радиус круговой диаграммы сравнения для представления учеников в школе с бульшим числом учащихся. В школе с бульшим числом учащихся 65% учеников ходят в школу пешком, 15% ездят на велосипеде, а остальные ездят на автобусе.

(v) Изобразите круговую диаграмму сравнения для представления учеников в школе с бульшим числом учащихся.

(vi) Сравните две круговые диаграммы и затем прокомментируйте различие выбора транспорта в двух школах.

1.6. Столбиковые диаграммы

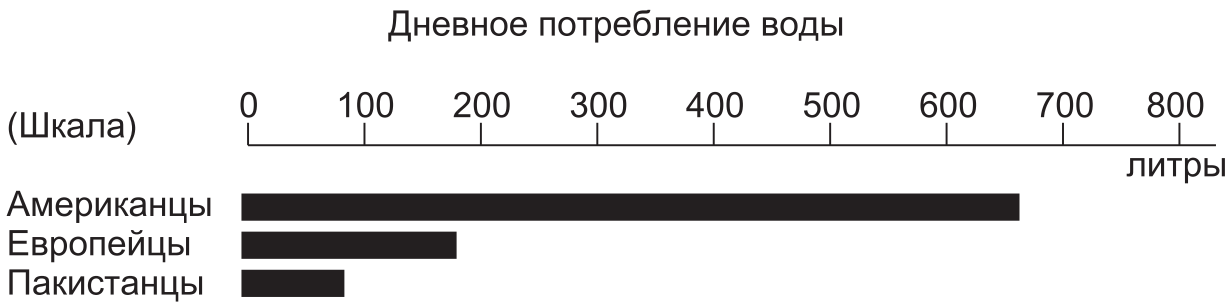
В столбиковых диаграммах столбцы, или графы, **равной** ширины используются для представления количеств или чисел.

Пример 1. Объем воды, используемой каждый день в целях санитарии и домашнего потребления представителями трех регионов, следующий (в л):

Американцы (США)	650
Европейцы	175
Пакистанцы	90

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

Решение.



Примечание.

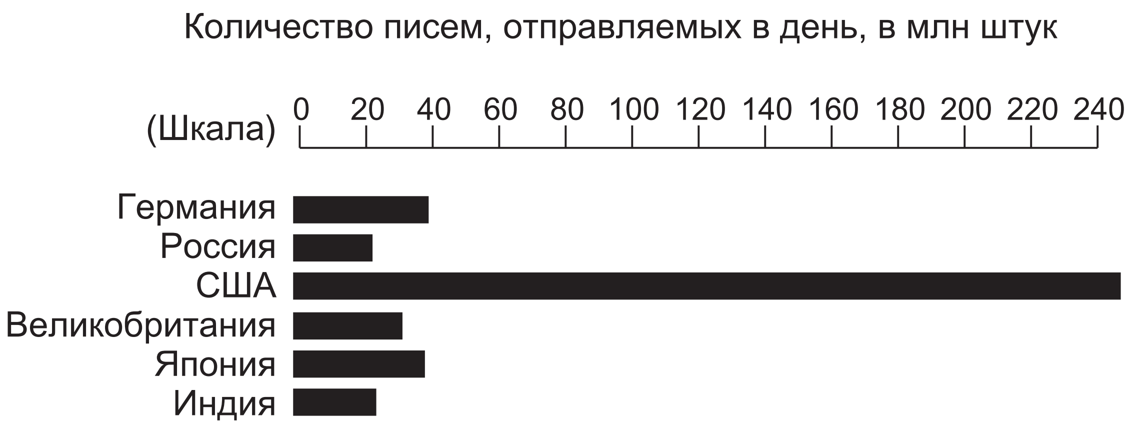
(i) Масштаб дается вдоль колонок, или столбцов.
(ii) Все столбцы одного цвета или одинаково затенены. (См. далее раздел 1.17.)

Пример 2. Число писем, отправляемых каждый день в шести странах мира, составляло в условном году (в млн штук):

Индия	20
Япония	38
Великобритания	24
США	243
Россия	19
Германия	38

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

Решение.

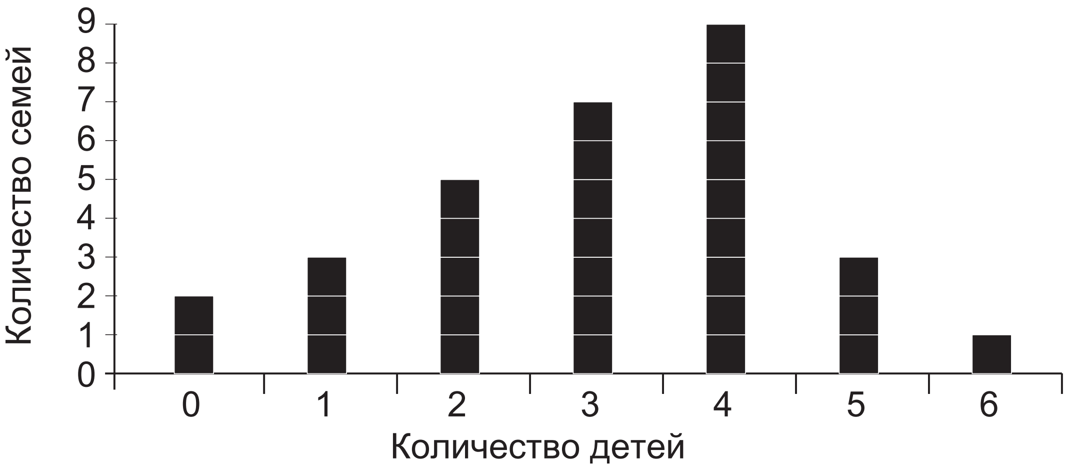


Пример 3. В таблице представлено количество детей в 30 различных семьях:

Количество детей	Количество семей
0	2
1	3
2	5
3	7
4	9
5	3
6	1

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

Решение.



Упражнение 4А *(ответы приведены в конце книги)*

1. Среднее количество денег, которые может потратить каждый человек в шести странах, составляет (в долл.):

Франция	17
Япония	13
Северная Америка	21
Швеция	24
Великобритания	11
Германия	20

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

2. Среднее количество энергии, потребляемой в день каждым человеком в четырех регионах мира, составляет (в условных единицах):

Центральная Америка	7
Европа	24
Океания	26
Северная Америка	62

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

Упражнение 4В

1. Среднее число шин, производимых в условном году для транспортных средств каждый день в пяти основных странах-производителях, следующее (в тыс. штук):

Франция	125
Япония	235
США	500
Россия	105
Германия	100

Проиллюстрируйте информацию с помощью столбиковой диаграммы.

2. Ниже приводятся основные цели, на которые пошло дневное производство стали в США (в тыс. т):

Приборы и оборудование	20
Консервные банки, канистры и контейнеры	20
Автомобили и прочие транспортные средства	100
Строительная промышленность	90
Машиностроение	70
Нефте- и газодобывающие отрасли	20

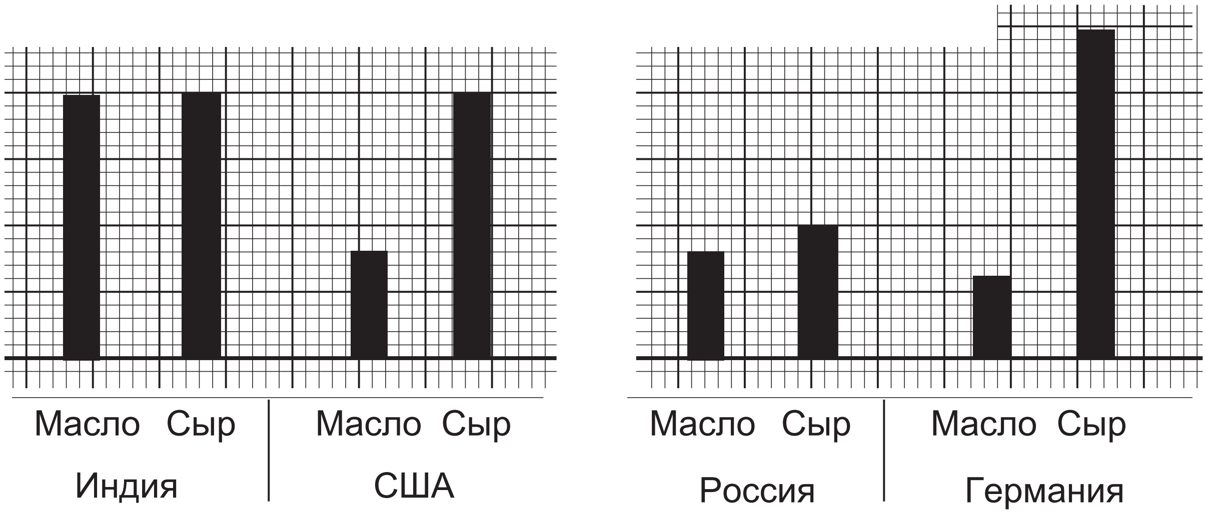
Проиллюстрируйте информацию с помощью столбиковой диаграммы.

1.7. Двойные столбиковые диаграммы

Двойные столбиковые диаграммы используются в случае, когда имеются две параллельные переменные и нам необходимо изучить соотношение связи между этими переменными, а также изменения по каждой из переменных.

Пример. Двойная столбиковая диаграмма иллюстрирует производство масла и сыра в четырех странах.

При рассмотрении двойных столбиковых диаграмм мы можем выявить не только страны с наибольшим производством, но также сравнить производство масла и сыра в каждой из четырех стран.



Упражнение 5

1. Ниже приводится посещаемость курсов учащимися в трех школах:

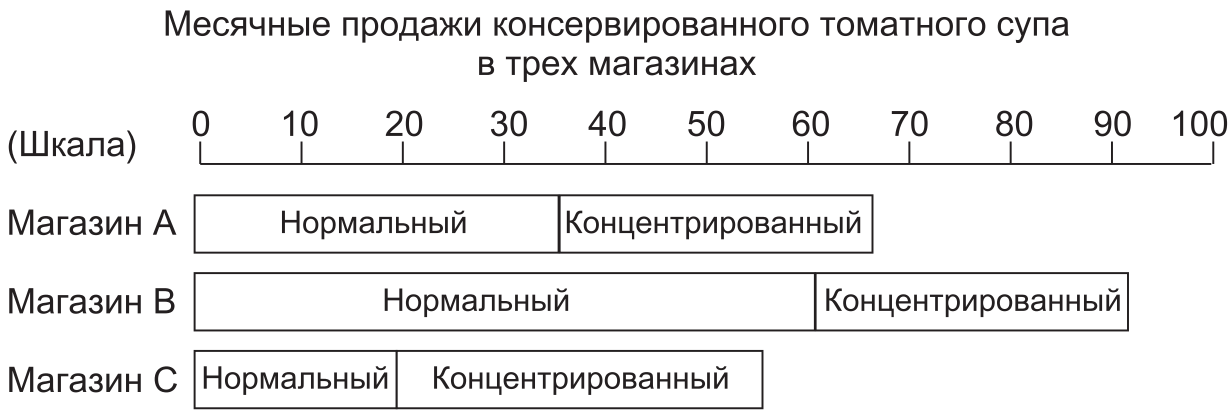
Школа A	Искусство 25	Наука 3
Школа B	Искусство 20	Наука 20
Школа C	Искусство 10	Наука 35

Проиллюстрируйте приведенные данные с помощью двойной столбиковой диаграммы.

1.8. Комбинированные столбиковые диаграммы

Комбинированные столбиковые диаграммы используются, когда изучаемое явление может быть разбито на некоторые группы. Например, консервированный томатный суп может быть нормальной консистенции (т.е. готовым к употреблению) или концентрированным (который сначала должен быть разбавлен).

Комбинированная столбиковая диаграмма для изображения продаж супа в трех магазинах или лавках будет выглядеть следующим образом.



Упражнение 6

1. Доход от премий (страховых взносов) и от процентов ведущей компании по страхованию жизни в 1981 – 1985 гг. (в млн ф.ст.) приведен ниже:

Годы	Проценты	Премии
1981	47,5	46,3
1982	49,5	71,0
1983	53,7	89,9
1984	62,4	113,9
1985	75,5	138,5

Источник: Отчет и счета за 1985 г. «Лондонской Ассоциации Жизни».

Проиллюстрируйте данные с помощью комбинированной столбиковой диаграммы.

2. Число работников на заводе за два последовательных года приводится ниже:

Персонал	1985	1986
Управленческий	18	25
Контролирующий	90	175
Производственный	360	625

Источник: условные данные.

Проиллюстрируйте данные с помощью комбинированной столбиковой диаграммы.

Прокомментируйте изменения, произошедшие за год.

1.9. Использование круговых и столбиковых диаграмм

Круговые диаграммы обычно используются для качественных переменных.

Столбиковые диаграммы чаще всего используются для качественных переменных, но могут быть использованы и для дискретных количественных переменных.

Круговые диаграммы предпочтительнее, когда количественное соотношение различных групп более важно, чем фактические значения в этих группах. Если информация представляется в процентах, а не в фактических значениях, круговая диаграмма будет наиболее подходящей. Например, мировое производство энергии, рассмотренное в вопросе 4 упражнения 2А (в %):

Уголь	34
Гидроэлектро- и атомная энергия	2
Природный газ	19
Нефть	45

Столбиковым диаграммам отдается предпочтение перед круговыми диаграммами, когда фактические значения более важны, чем количественные или процентные соотношения. Рассмотрим представленный в примере 1 параграфа 1.6 объем воды, потребляемой ежедневно в разных регионах (в л):

Американцы (США)	650
Европейцы	175
Пакистанцы	90

Поскольку условия жизни населения в США, Европе и Пакистане сильно отличаются, круговая диаграмма будет в данном случае лишена смысла.

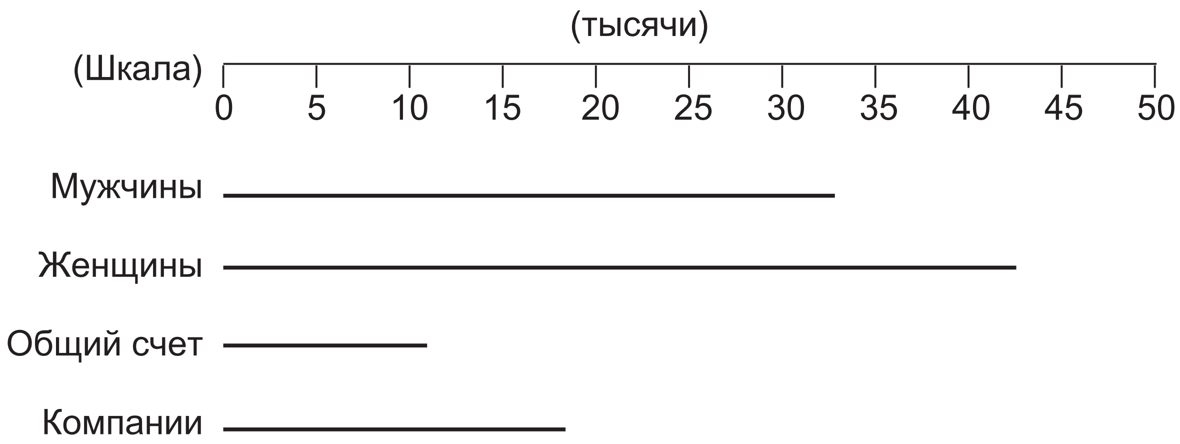
Можно обобщить вышесказанное следующим образом:

	Преимущества	Недостатки
Круговая диаграмма	Легко увидеть соотношения количеств	Сложно увидеть количественные значения
Столбиковая диаграмма	Легко увидеть количественные значения	Сложно увидеть соотношения количеств

1.10. Линейные диаграммы

Линейные диаграммы попросту являются альтернативой столбиковым диаграммам. «Столбики» сокращались в толщине до тех пор, пока не превратились в карандашную линию.

Пример. Анализ акционеров Национального Вестминстерского Банка (НВБ) на 31 декабря 1985 г.

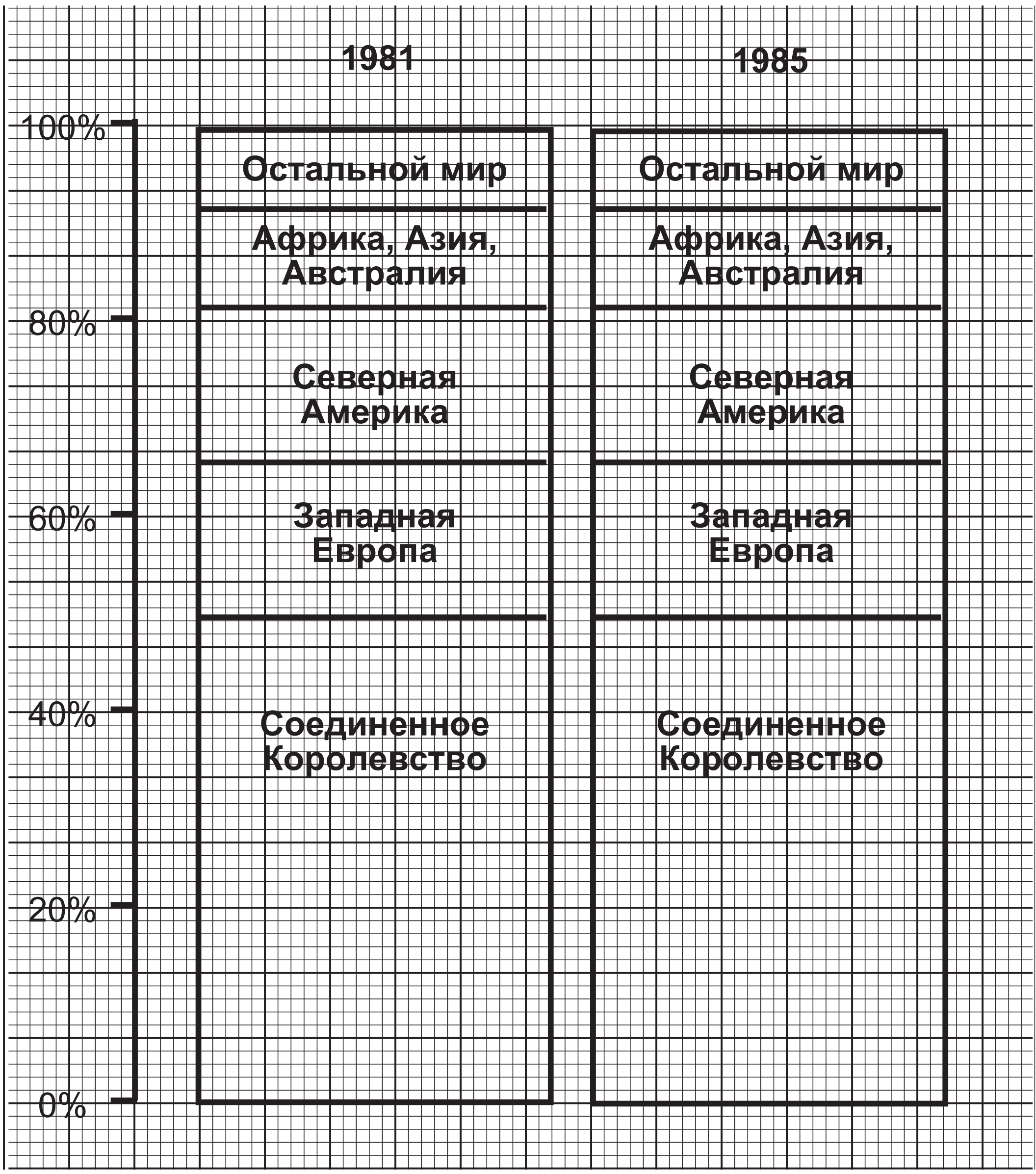


Источник: Годовой отчет Национального Вестминстерского Банка, 1985.

1.11. Групповые процентные столбиковые диаграммы

Перед построением диаграммы данные необходимо выразить в процентах

Пример. Распределение общего дохода НВБ по регионам.

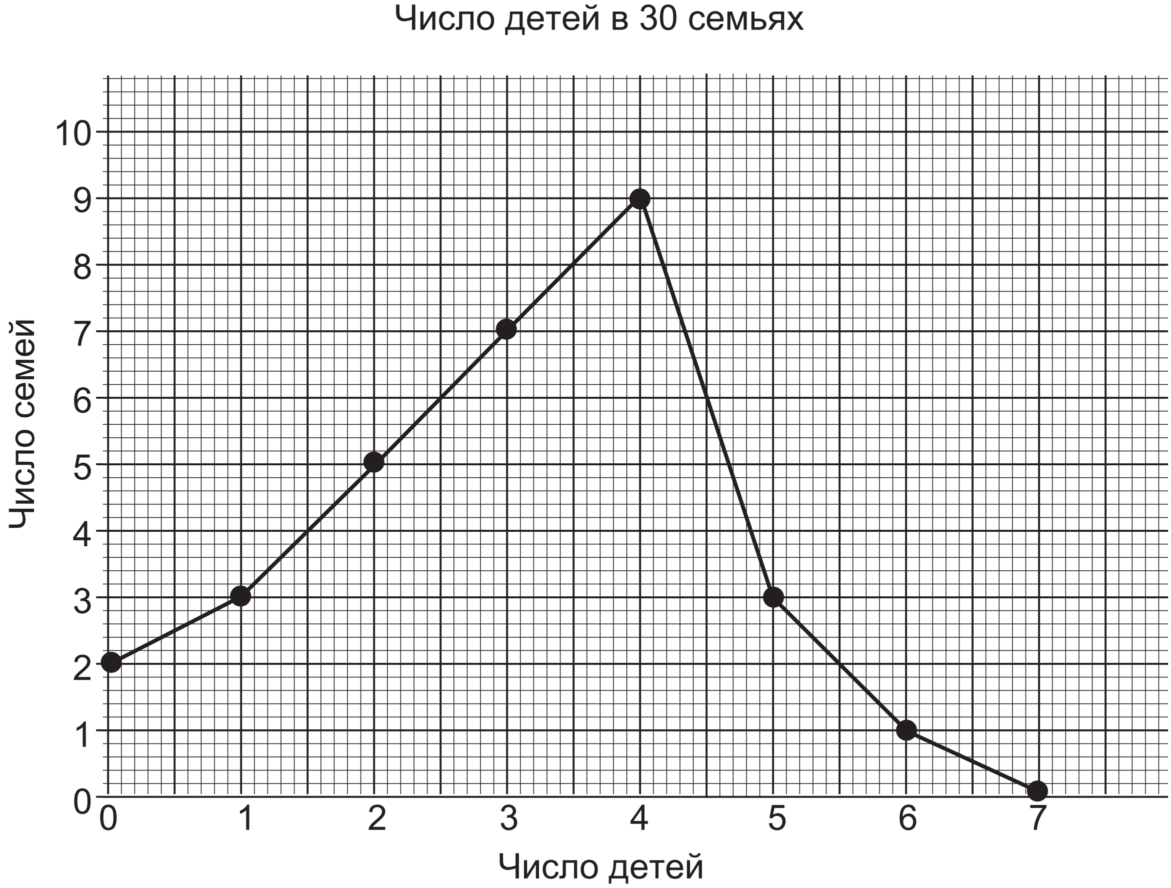


Источник: Годовой отчет Национального Вестминстерского Банка, 1985.

1.12. Частотные полигоны

Частотные полигоны используются для изображения распределения дискретных количественных переменных.

Пример. Число детей в 30 семьях, которое было проиллюстрировано столбиковой диаграммой в примере 3 параграфа 1.6, может также быть проиллюстрировано с помощью частотного полигона.



Примечание. Точки должны быть соединены отрезками прямых линий.

Упражнение 7

1. Ниже приводится число голов, забитых 30 футбольными командами в один и тот же день – субботу:

2	3	1	1	1	0	0	3	1	2
1	1	1	1	0	0	1	2	1	1
0	5	2	3	2	0	1	1	0	2

Постройте таблицу числа голов и проиллюстрируйте результат с помощью частотного полигона.

2. Ниже приводится количество лиц, проживающих в 15 домах в небольшом жилом районе:

Количество проживающих лиц	0	1	2	3	4	5	6 или более
Количество домов	0	2	3	5	4	1	0

Проиллюстрируйте данные с помощью частотного полигона.

1.13. Гистограммы

1.13.1. Введение в гистограммы

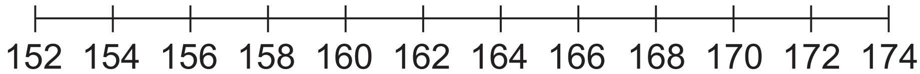
Интервальное количественное распределение иллюстрируется с помощью гистограммы. Площадь каждого столбца представляет частоту или число в этой группе. Высота столбца не всегда пропорциональна частоте или числу в этой группе.

Пример 1. Приведен рост девочек определенной возрастной группы.

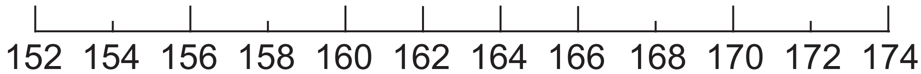
Группировка Рост X девочки (см)	*Частота* Количество девочек
$152 \le X < 156$	4
$156 \le X < 160$	8
$160 \le X < 162$	8
$162 \le X < 164$	14
$164 \le X < 166$	10
$166 \le X < 170$	12
$170 \le X \le 174$	4

Проиллюстрируйте приведенные данные с помощью гистограммы.

Решение. Начнем с изображения подходящей линии основания гистограммы, охватывающей интервал от 152 до 174 см.



Затем на линии основания отметим границы интервалов 152, 156, 160, 162, 164, 166, 170, 174 см.



Затем рассчитаем высоту каждого столбца с учетом того факта, что площадь столбца представляет число единиц, или частотность.

Основание или ширина первого столбца равна 4 см (156 см – – 152 см).

Площадь первого столбца представляет 4 девочки.

Таким образом, высота первого столбца равна 4 девочки : : 4 см = 1 девочка/см.

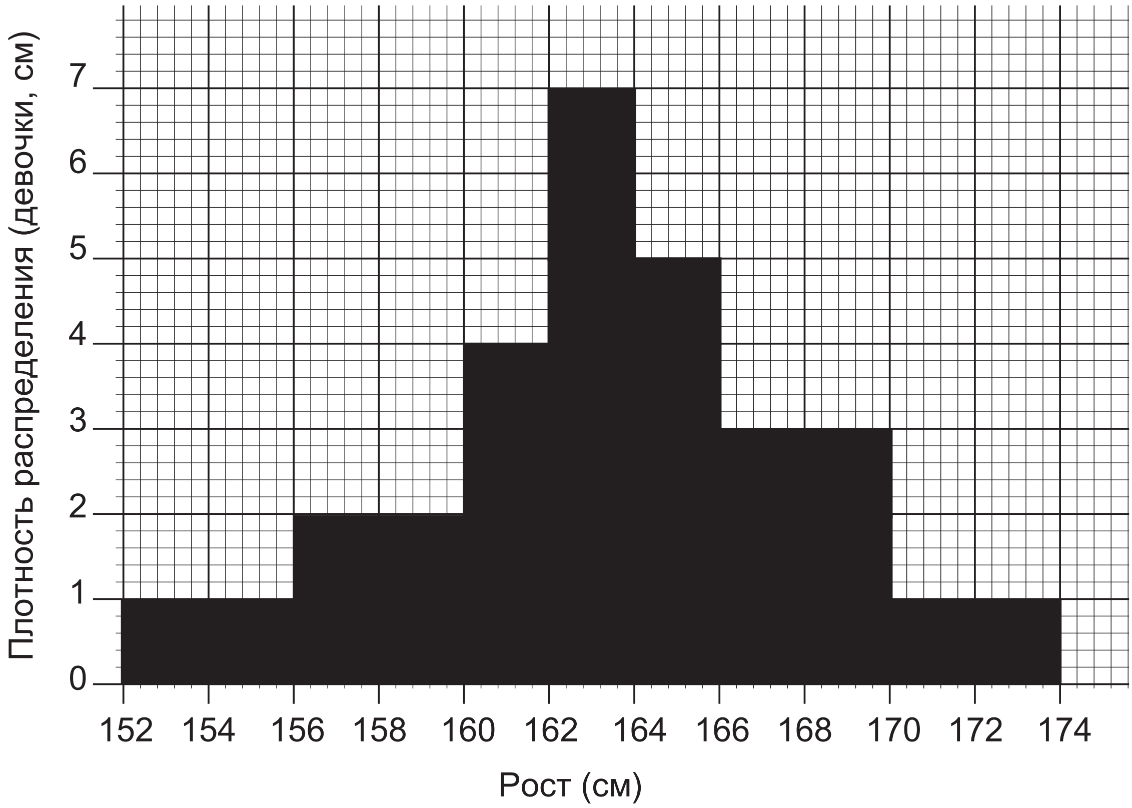
Переменная по оси Y называется **плотностью распределения.**

Для этого графика единицей измерения по оси Y является «девочка/см». Когда мы находим площадь первого столбца, мы умножаем основание на высоту (4 см × 1 девочка/см), что дает нам 4 девочки (точно так же, как 2 ч × 15 км/ч дает 30 км)

Таким же способом рассчитываем высоту всех столбцов.

Группировка Рост X (см)	Ширина столбца (см)	Частота, или количество девочек *f*	*Плотность распределения* Высота столбца Частотность/ширина (девочка/см)
152≤X<156	4	4	4/4=1
156≤X<160	8	8	8/4=2
160≤X<162	8	8	8/2=4
162≤X<164	14	14	14/2=7
164≤X<166	10	10	10/2=5
166≤X≤170	12	12	12/4=3
170≤X<174	4	4	4/4=1
Всего		60	

Теперь можно нарисовать гистограмму.



Пример 2. Были взвешены 260 животных; полученные результаты приведены ниже:

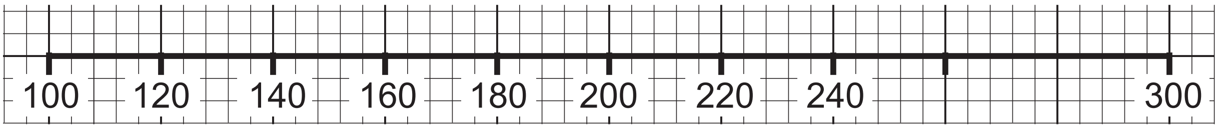
Масса животного (г)	100-	140-	160-	180-	200-	220-	240-300
Число животных	16	46	76	54	40	10	18

Проиллюстрируйте информацию с помощью гистограммы.

Примечание. В этой таблице 100- означает, что интервал начинается от 100 г и оканчивается на нижней границе следующего интервала, т.е. 140 г.

Таким образом, «100-» здесь означает то же, что и $100 \le X < 140$.

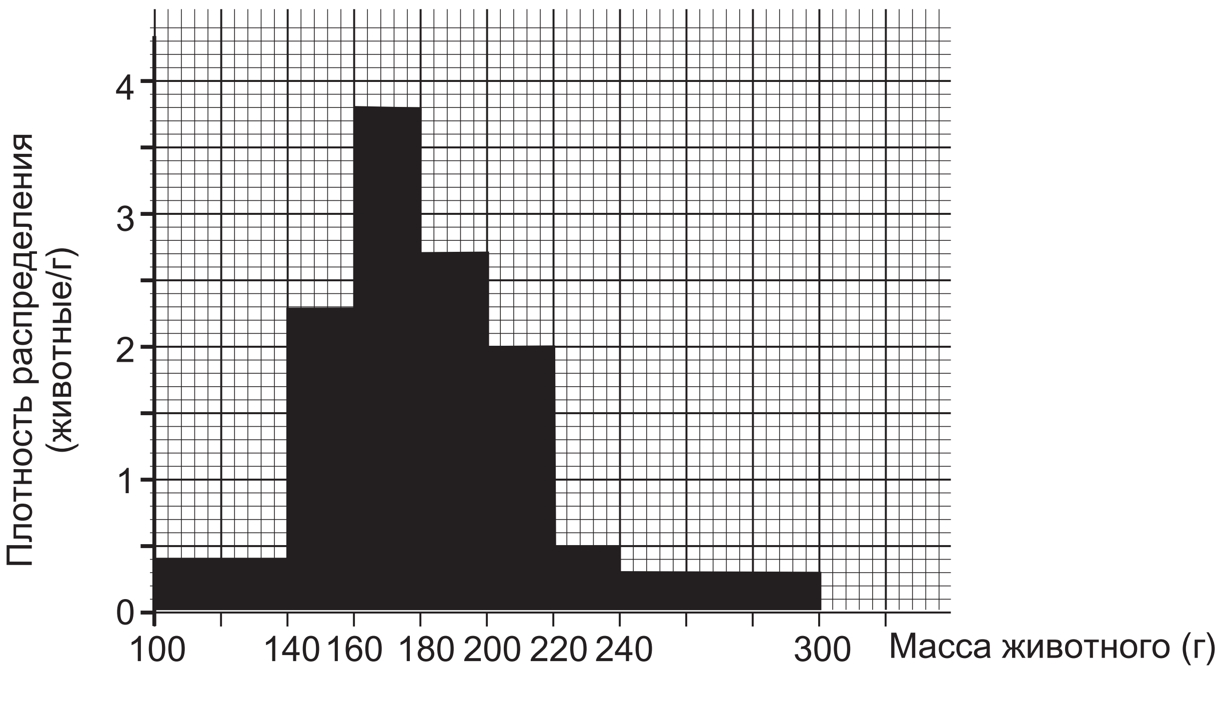
Решение. Начнем с того, что нарисуем линию основания от 100 до 300 г, а затем отметим границы интервалов в 100, 140, 160, 180, 200, 220 и 240 г.



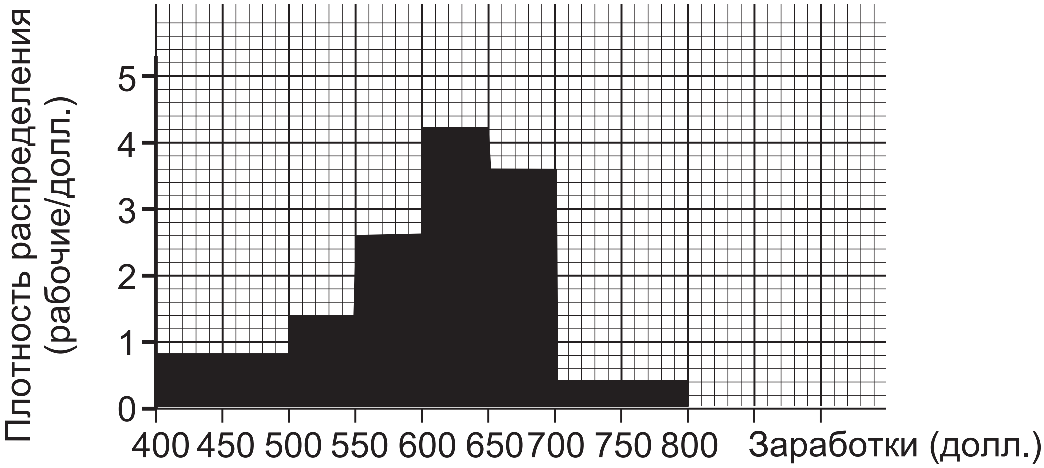
Затем рассчитаем высоту каждого столбца, как это делалось ранее.

Группировка Масса животного (г)	Ширина столбца (см)	Частота, или количество животных *f*	*Плотность распределения* Высота столбца Частота/Ширина
100-	40	16	16/40=0,4
140-	20	46	46/20=2,3
160-	20	76	76/20=3,8
180-	20	54	54/20=2,7
200-	20	40	40/20=2,0
220-	20	10	10/20=0,5
240-300	60	18	18/60=0,3
Всего (Проверка)		260	

Теперь можно нарисовать гистограмму.



Пример 3. Приведенная ниже гистограмма иллюстрирует распределение рабочих завода по заработкам.



Определите количество работников в каждом интервале.

Решение.

Заработки (долл.)	Ширина столбца (долл.)	*Высота столбца* Количество рабочих/долл.	Количество работников (Частотность)
400≤X<500	100	0,8	80
500≤X<550	50	1,4	70
550≤X<600	50	2,4	120
600≤X<650	50	4,2	210
650≤X<700	50	3,6	180
700≤X<800	100	0,4	40

Общее количество работников составляет 700.

1.13.2. Гистограмма по «округленным» данным

Иногда данные временно́го ряда представляются в «округленном», или «скорректированном», виде – до целого числа, до одной десятой или каким-либо другим образом. Когда мы рисуем гистограмму по округленным данным, эти округленные данные вначале следует заменить действительными границами интервала.

Например, 5 г (с точностью до целого грамма) означает интервал от 4,500 до 5,500 г. Меньшее значение, которое может быть округлено до 5, это 4,500, в то время как большее значение, например 5,500 г, может также быть округлено до 5 г (хотя традиционно 5,5 округляется до 6, в статистике 5,5 можно округлять в любую сторону).

Пример 1. Рост 20 девочек был измерен с точностью до сантиметра, и результаты представлены в таблице:

Рост (с точностью до см)	152	153	154	155	156
Количество девочек	3	5	7	4	1

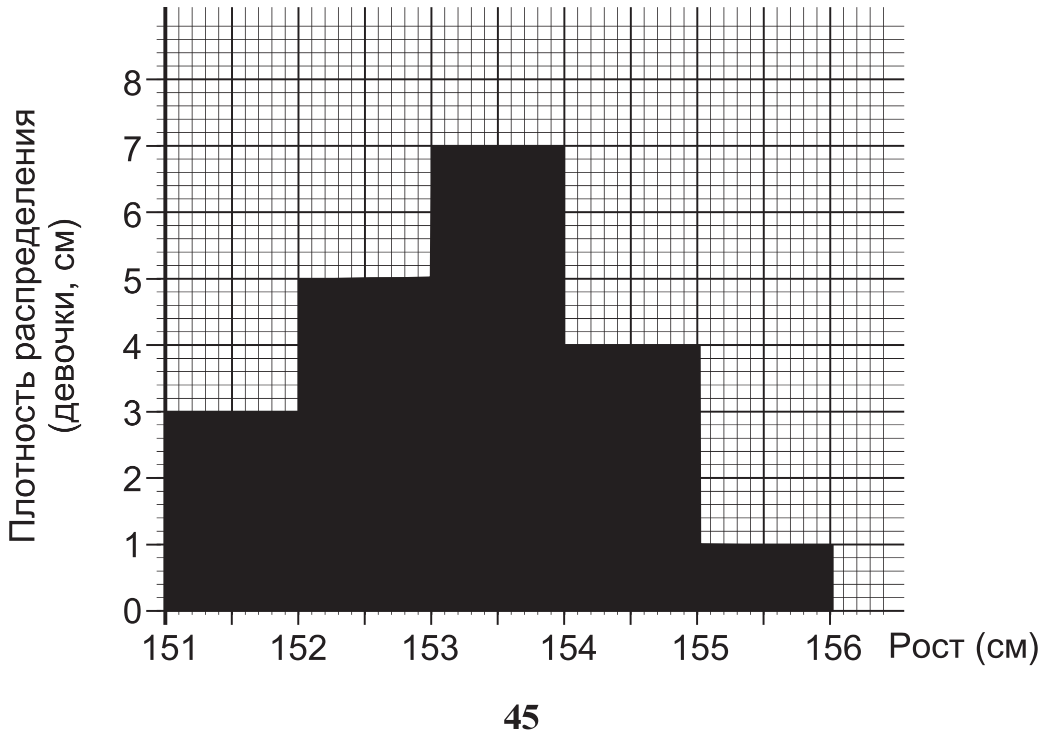
Проиллюстрируем распределение с помощью гистограммы.

Решение. 152 см с точностью до ближайшего целого сантиметра означает 151,5≤X<152,5, точно так же и по другим интервалам.

Таким образом, распределение становится следующим:

Группировка Рост X (см)	Ширина столбца (см)	*Частота* Количество девочек *f*	*Плотность распределения* Высота столбца Частота/Ширина
151,5≤X<152,5	1,0 см	3	3
152,5≤X<153,5	1,0 см	5	5
153,5≤X<154,5	1,0 см	7	7
154,5≤X<155,5	1,0 см	4	4
155,5≤X≤156,5	1,0 см	1	1

Получаем следующую гистограмму.



Аудиторный тест № 2

До проведения теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте.

Решите, какой из вариантов ответа, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
	Возраст (с точностью до года) 5–14, 15–24, 25–34					
1.	Нижняя граница 1-го интервала	4,5	4,9	4,95	5,0	5,05
2.	Верхняя граница 2-го интервала	23,5	23,9	24,0	24,5	24,99999
	Число полных лет 7–10, 11–14, 15–18					
3.	Нижняя граница 1-го интервала	6,5	6,9	7,0	7,5	7,99999
4.	Верхняя граница 2-го интервала	13	13,5	13,9	14	14,99999
	Площадь в гектарах $10 \le X < 20$, $20 \le X < 30$, $30 \le X < 40$					
5.	Нижняя граница 1-го интервала	9,5	9,9	10,0	10,05	10,5
6.	Верхняя граница 2-го интервала	30	25	29	29,5	29,99999

Упражнение 8A

1. Ниже приведены результаты измерений 80 брусков, изготовленных на неком станке.

Длина бруска, мм	130-	140-	145-	150-	155-	160-180
Число брусков	10	10	20	16	12	12

Изобразите приведенные данные при помощи гистограммы.

2. Ниже приведены данные о площадях сельскохозяйственных угодий в регионе:

Площадь (га)	0-	25-	50-	60-	70-	80-100
Число сельскохозяйственных угодий	25	50	30	60	45	10

Изобразите приведенные данные при помощи гистограммы.

Упражнение 8B

1. Имеются следующие данные о распределении населения города по пяти возрастным группам:

Возрастная группа	0-	10-	20-	40-	60-100
Число жителей	600	900	2000	3500	3000

Изобразите эти данные при помощи гистограммы.

2. В таблице приведены результаты измерения длины 90 деталей, сошедших с конвейера (результаты измерений приведены с точностью до мм).

Длина (с точностью до мм)	70-71	72-73	74-75	76-80	81-90
Число деталей	18	28	14	20	10

Изобразите эти данные при помощи гистограммы.

1.14. Диаграмма кумулятивных частот

1.14.1. Введение в диаграммы кумулятивных частот

До сих пор мы иллюстрировали «плотность распределения», т.е. численность, или частоту, в каждой категории или интервале. Иногда бывает полезно проиллюстрировать распределение кумулятивных частот.

Пример 1. Постройте распределение кумулятивных частот данного ряда распределения.

Заработки (долл.)	Частота (Количество работников)
$400 \le X < 500$	80
$500 \le X < 550$	70
$550 \le X < 600$	120
$600 \le X < 650$	210
$650 \le X < 700$	180
$700 \le X \le 800$	40
Всего	700

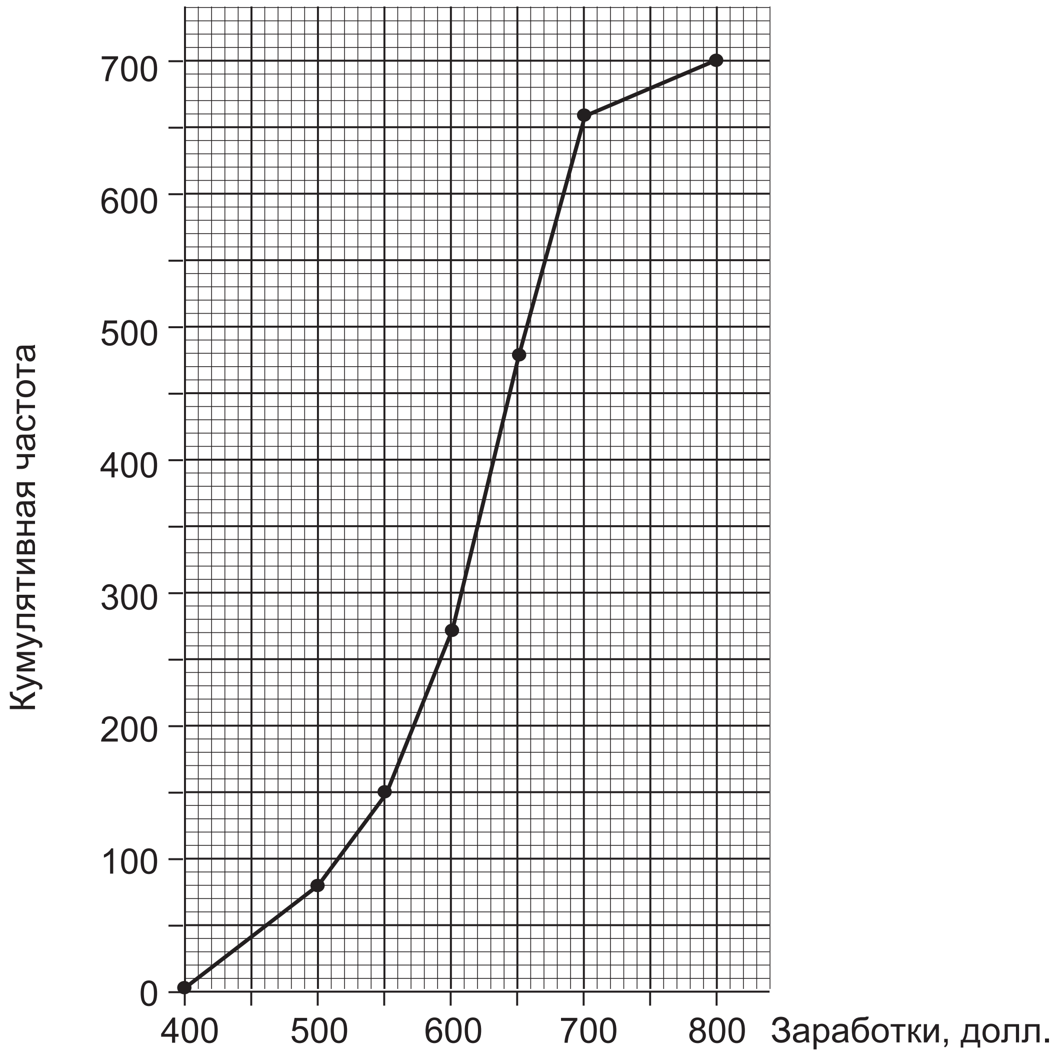
Решение. Сначала мы составляем **таблицу кумулятивных частот**, суммируя частоты в разных интервалах. Это всегда следует делать в новой таблице. Попытка произвести сложение в имеющейся таблице может привести к серьезным ошибкам.

Заработки (долл.) «до», или «меньше чем»	Кумулятивная частота
400	0
500	0 + 80 = 80
550	80 + 70 = 150
600	150 + 120 = 270
650	270 + 210 = 480
700	480 + 180 = 660
800	660 + 40 = 700

Примечания.

1. В первой строке таблицы записывается значение нижней границы первого интервала (в нашем примере 400), а кумулятивная частота равна нулю (0).
2. В последней строке таблицы записывается значение верхней границы последнего интервала (здесь 800), а кумулятивная частота равна общей, т.е. сумме частот по группам (в данном случае 700).

Затем наносим полученные значения на график.



Если мы соединим соседние точки отрезками прямых, то получим **полигон кумулятивных частот.**

Если мы соединим все точки плавной кривой, то получим **кривую кумулятивных частот**, или **огиву.**

Пример 2. Постройте кривую кумулятивных частот для иллюстрации следующего распределения частот.

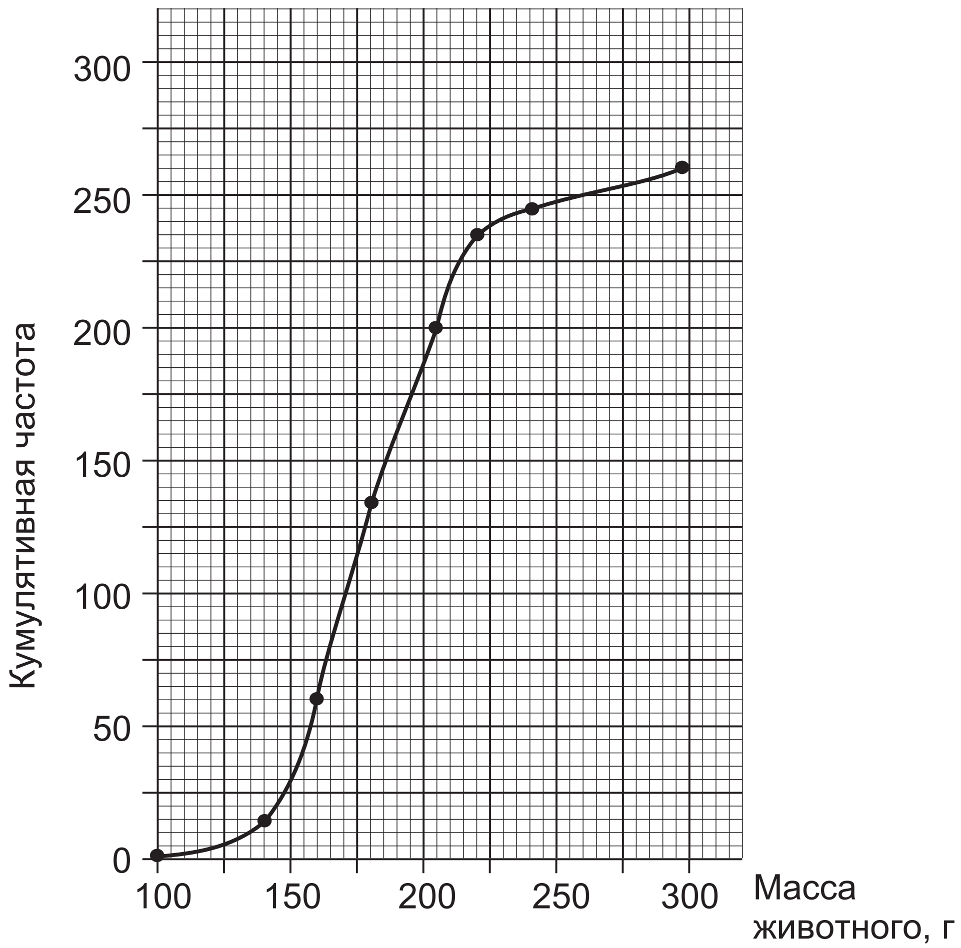
Масса животного (г)	Частота (количество животных)
$100 \leq X < 140$	16
$140 \leq X < 160$	46
$160 \leq X < 180$	76
$180 \leq X < 200$	54
$200 \leq X < 220$	40
$220 \leq X < 240$	10
$240 \leq X \leq 300$	18
Суммарная частота	260

Решение.

Построим таблицу кумулятивных частот.

Масса животного (г) «до» или «меньше чем»	Кумулятивная частота	
100	0	
140	0 + 16	16
160	16 + 46	62
180	62 + 76	138
200	138 + 54	192
220	192 + 40	232
240	232 + 10	242
300	242 + 18	260

Проверяем, что значение последней кумулятивной частоты (260) совпадает с суммарной частотой. Теперь строим огиву, или кривую кумулятивных частот.



Кумулятивная частота
300
250
200
150
100
50
0
100
150
200
250
300
Масса животного, г

1.14.2. График кумулятивных частот по округленным данным

Иногда данные интервального ряда распределения представлены в «округленном», или «скорректированном», виде – до целого числа, до одного десятичного знака и т.п. При построении графика кумулятивных частот по округленным данным эти округленные данные вначале должны быть заменены *действительными границами интервала*.

Например, 5 г (с точностью до грамма) означает интервал от 4,500 г до 5,500 г. Минимальное значение массы, которое можно округлить до 5 г, – это 4,500 г, тогда как максимальное – 5,500 г. (Хотя общепринято округлять 5,5 до 6, в статистике 5,5 можно округлять в любую сторону.)

Пример 1. Ниже приведены результаты измерения роста 20 девочек (с точностью до сантиметра):

Рост (см)	X	152	153	154	155	156
Количество девочек	*f*	3	5	7	4	1

Постройте полигон кумулятивных частот для иллюстрации распределения.

Решение.

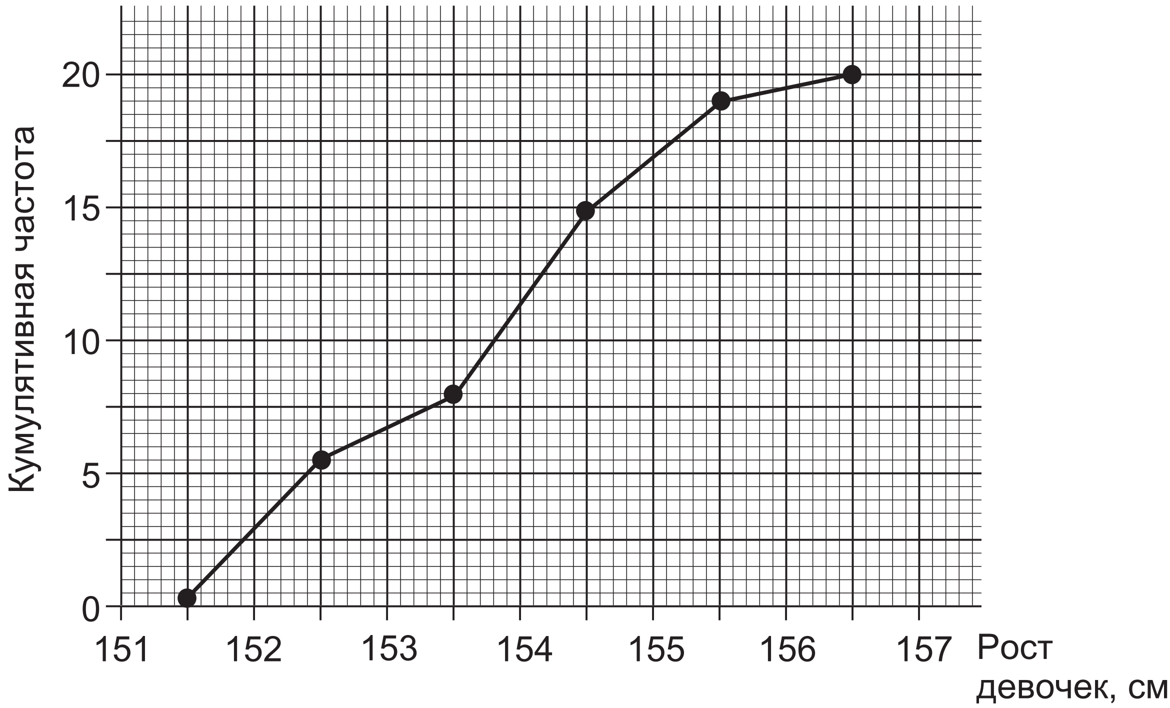
В нижнем интервале 152 см с точностью до сантиметра означает интервал $151{,}5 \le X < 152{,}5$, следовательно, кумулятивная частота 0 на графике соответствует значению 151,5.

В верхнем интервале 156 см с точностью до сантиметра означает интервал $155{,}5 \le X < 156{,}5$, следовательно, кумулятивная частота 20 на графике соответствует значению 156,5.

Аналогично находятся границы остальных интервалов.

Строим таблицу кумулятивных частот:

Рост девочек (см) «до», или «меньше чем»	Кумулятивная частота
151,5	0
152,5	3
153,5	8
154,5	15
155,5	19
156,5	20



Теперь строим полигон кумулятивных частот.

1.14.3. График кумулятивных частот сгруппированных дискретных данных

Пример. Построим полигон кумулятивных частот приведенного ниже сгруппированного ряда распределения.

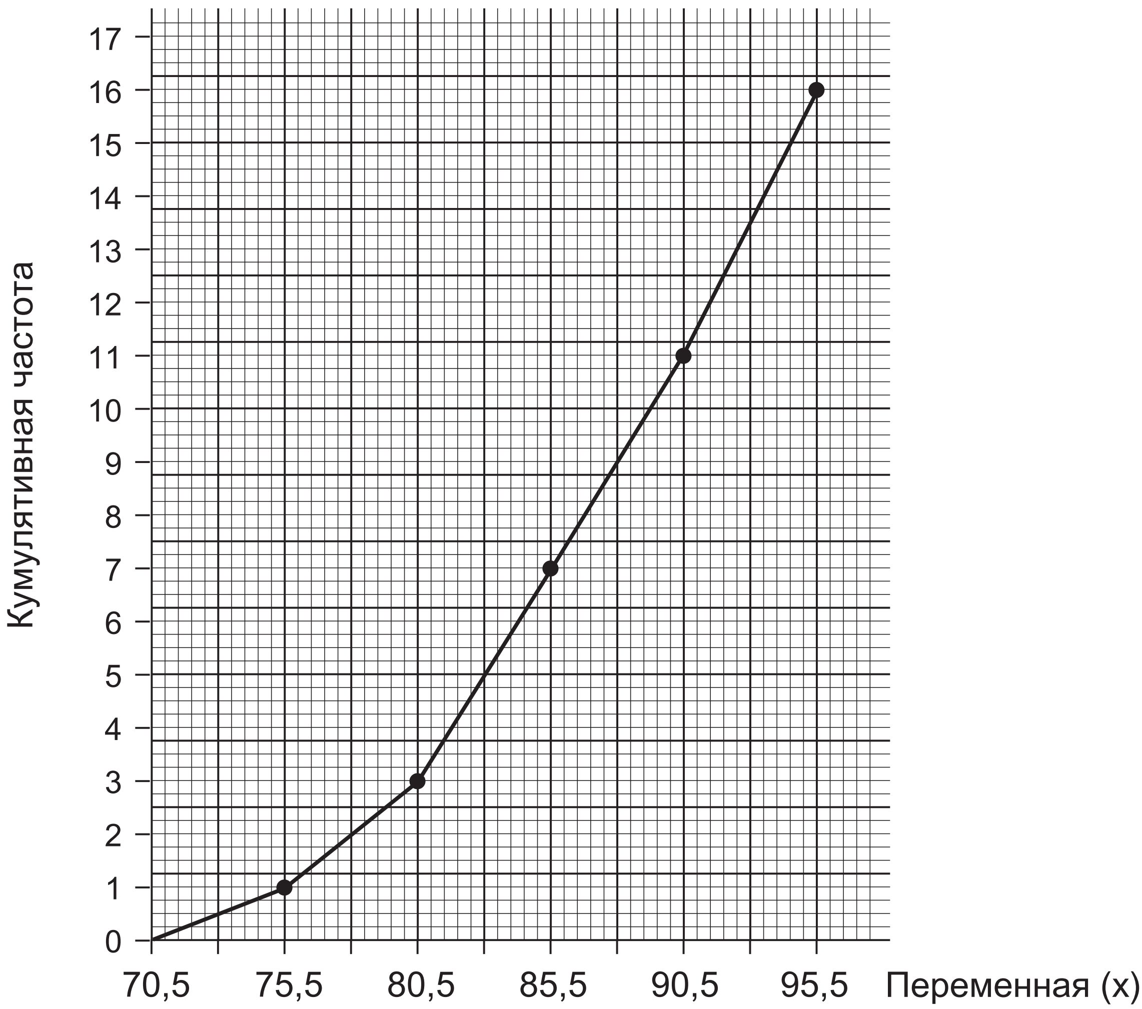
Значения переменной	Частота
71-75	1
76-80	2
81-85	4
86-90	4
91-95	5
Суммарная частота	16

Решение. Интервал 71-75 трактуется как интервал значений от 70,5 до 75,5 (аналогично для других интервалов). Далее построим таблицу кумулятивных частот.

Проверка: последняя кумулятивная частота равна суммарной частоте.

Переменная «до», или «меньше чем»	Кумулятивная частота
70,5	0
75,5	1
80,5	3
85,5	7
90,5	11
95,5	16

Изобразим полигон кумулятивных частот:



1.14.4. График кумулятивных частот дискретных данных

График кумулятивных частот дискретных данных имеет вид «ступенек».

Пример. Построить график кумулятивных частот следующего ряда распределения.

Переменная	X	0	1	2	3	4
Частота	*f*	2	7	8	2	1

Решение.

Суммарная частота равна 20.

Для X = 0 кумулятивная частота изменяется от 0 до 2.

Для X = 1 кумулятивная частота изменяется от 2 до 9.

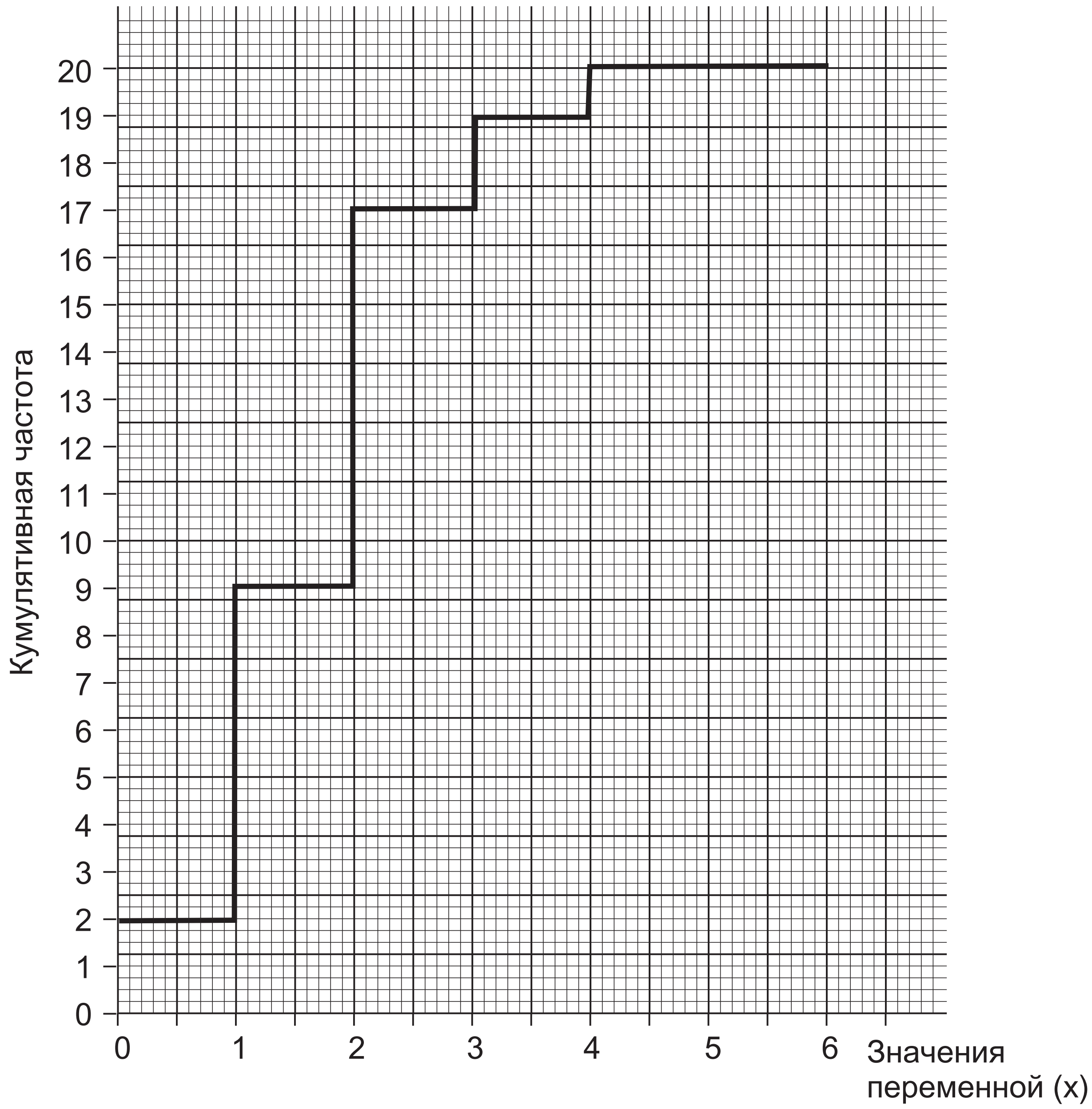
Для X = 2 кумулятивная частота изменяется от 9 до 17.

Для X = 3 кумулятивная частота изменяется от 17 до 19.

Для X = 4 кумулятивная частота изменяется от 19 до 20.

Проверка: последняя кумулятивная частота равна суммарной частоте.

Ниже приведен график кумулятивных частот.


Кумулятивная частота
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
0
1
2
3
4
5
6
Значения переменной (x)

Упражнение 9А

1. В один день посеяли 80 семян одного сорта овощной культуры. Спустя некоторое время была измерена высота всходов. Получены следующие результаты:

Высота растения (см)	0-	5-	6-	7-	8-	9-	10-12
Количество растений	0	6	8	13	19	26	8

Постройте таблицу кумулятивных частот данного ряда распределения.

Затем постройте полигон кумулятивных частот.

2. На экзамене по статистике за две письменные работы 200 экзаменуемых получили следующие баллы:

Количество баллов	Количество экзаменуемых	
	Работа №1	Работа №2
11–20	12	0
21–30	20	0
31–40	28	0
41–50	36	0
51–55	20	0
56–60	16	40
61–70	22	48
71–80	20	64
81–90	26	36
91–100	0	12

Постройте два полигона кумулятивных частот для иллюстрации информации.

Прокомментируйте различие между двумя письменными работами в отношении:

(a) трудностей, с которыми сталкивались экзаменуемые при наборе баллов;

(b) эффективности письменных работ для «отбора» экзаменуемых.

Упражнение 9В

1. Изучение характеристик 40 судов, стоящих на якоре в гавани приморского города, дало следующие результаты:

Максимальная скорость (км/ч)	0-	10-	12-	14-	15-	16-	18-	20-25
Количество судов	0	4	10	7	4	6	4	5

Постройте таблицу кумулятивных частот ряда распределения. Затем постройте кривую кумулятивных частот (огиву).

2. Распределение рабочих сборочного завода по доходам в текущем году представлено в таблице (в ф.ст.)

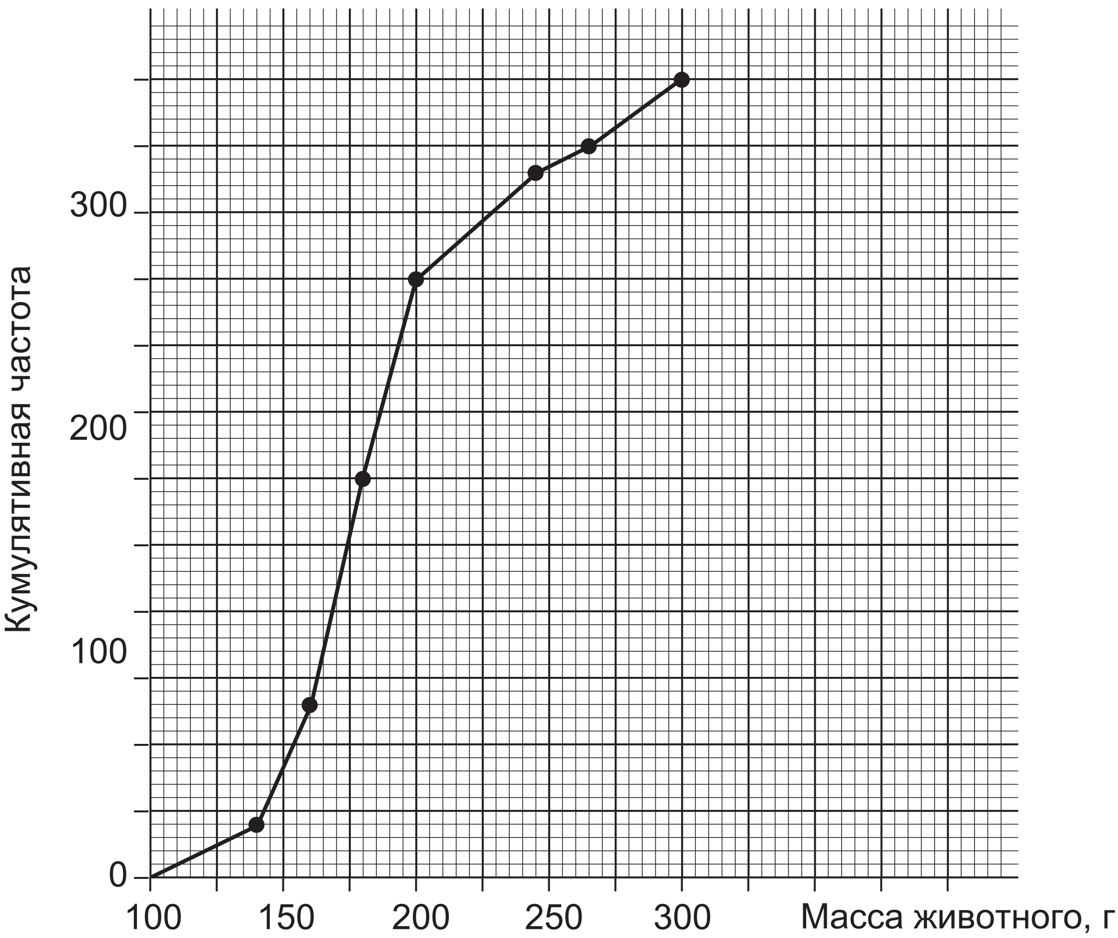
Доход	3500-	4000-	4200-	4400-	4600-	5000-
Количество рабочих	12	20	24	36	8	0

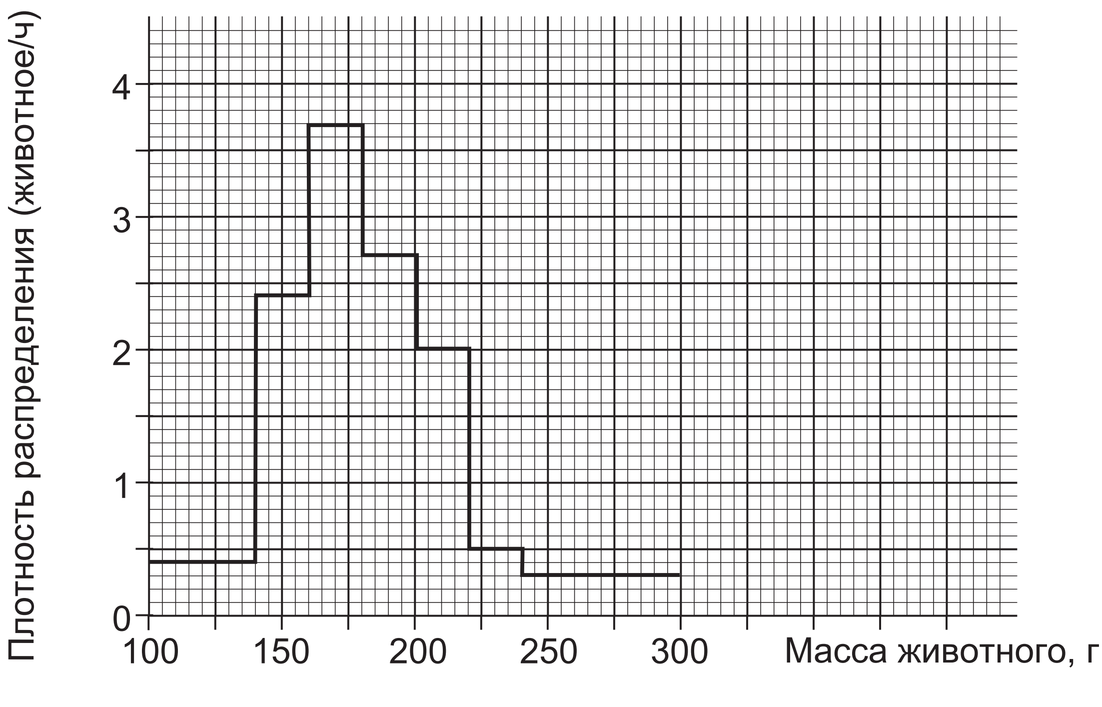
Проиллюстрируйте информацию с помощью кривой кумулятивных частот (огивы).

1.15. Соотношение между полигоном кумулятивных частот и гистограммой

Далее изображены полигон кумулятивных частот и гистограмма к примеру 2 из параграфа 1.14.1. Шкала по горизонтальной оси X на обоих графиках согласована таким образом, что одинаковые значения расположены строго друг под другом.

Мы видим, что самая крутая часть (с максимальным градиентом) полигона кумулятивных частот соответствует самому высокому столбцу гистограммы. Таким образом, высота столбца пропорциональна градиенту соответствующей части полигона.

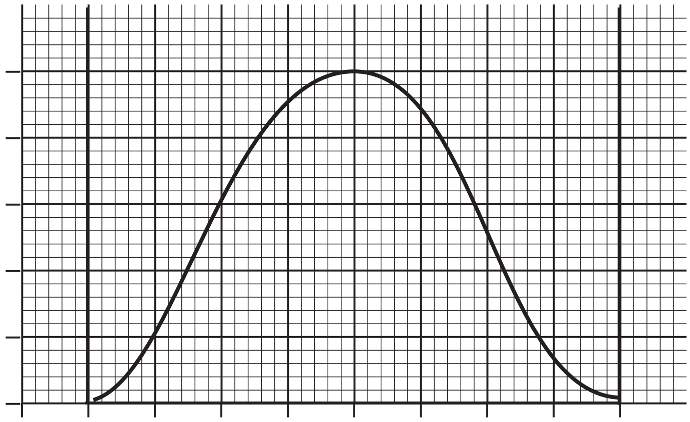

Кумулятивная частота
0
100
200
300
100
150
200
250
300
Масса животного, г


Плотность распределения (животное/ч)
0
1
2
3
4
100
150
200
250
300
Масса животного, г

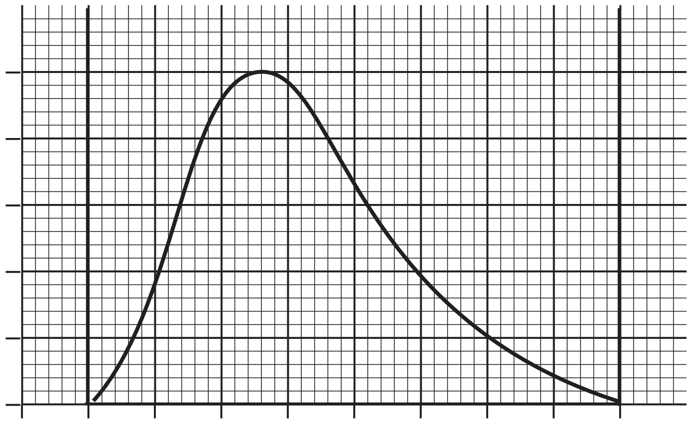
1.16. Виды распределений

Вид гистограммы дает нам представление о виде распределения, с которым мы имеем дело. Ниже приведены шесть наиболее распространенных видов.

(i) нормальное распределение

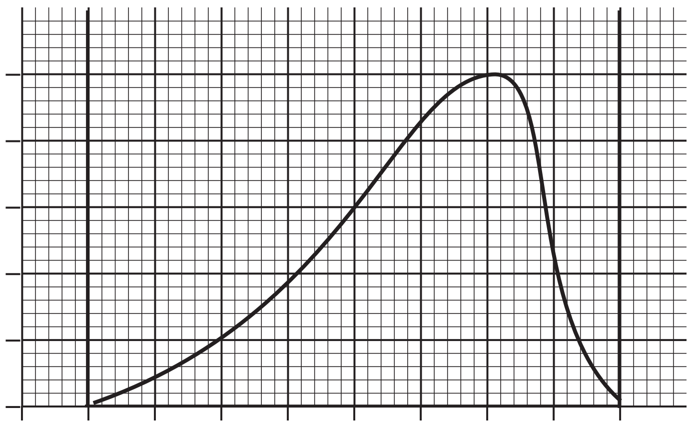


(ii) распределение с правосторонней асимметрией



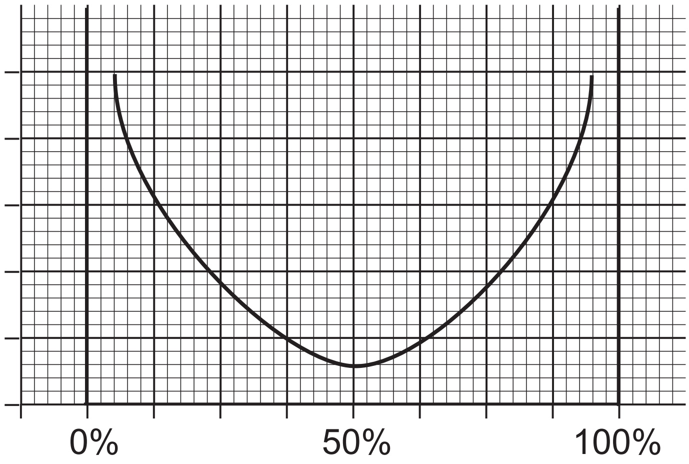
Пример. Возраст вступления в брак

(iii) распределение с левосторонней асимметрией



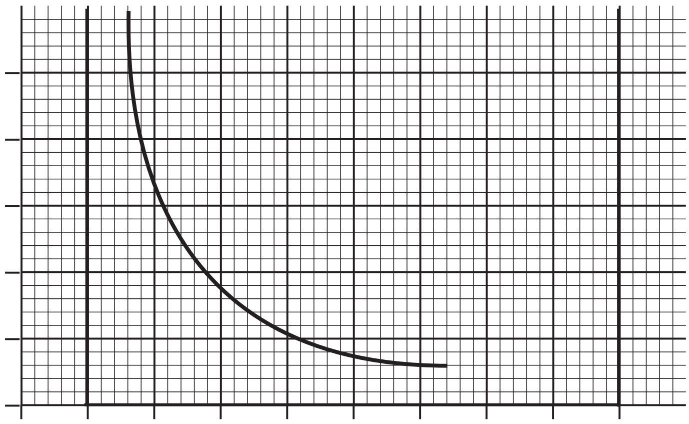
Пример. Возраст умирающих

(iv) U-образное распределение

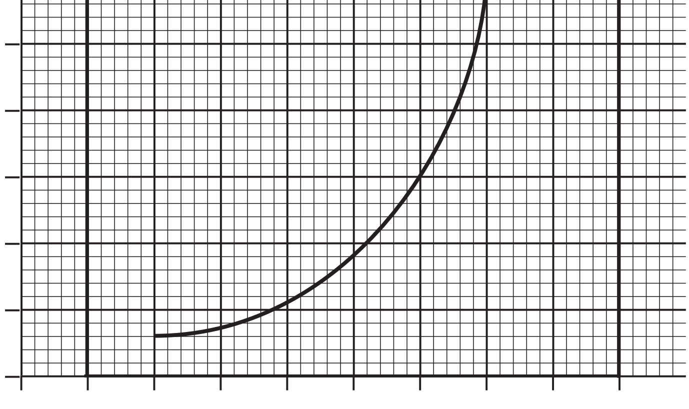


Пример. Облачность над Кью-Гарденс в Лондоне

(v) гиперболическое распределение



(vi) экспоненциальное распределение



1.17. Обманчивое представление данных

Иногда графическое представление данных может ввести в заблуждение. Это может быть обдуманным шагом для того, чтобы представить односторонний взгляд на предмет обсуждения.

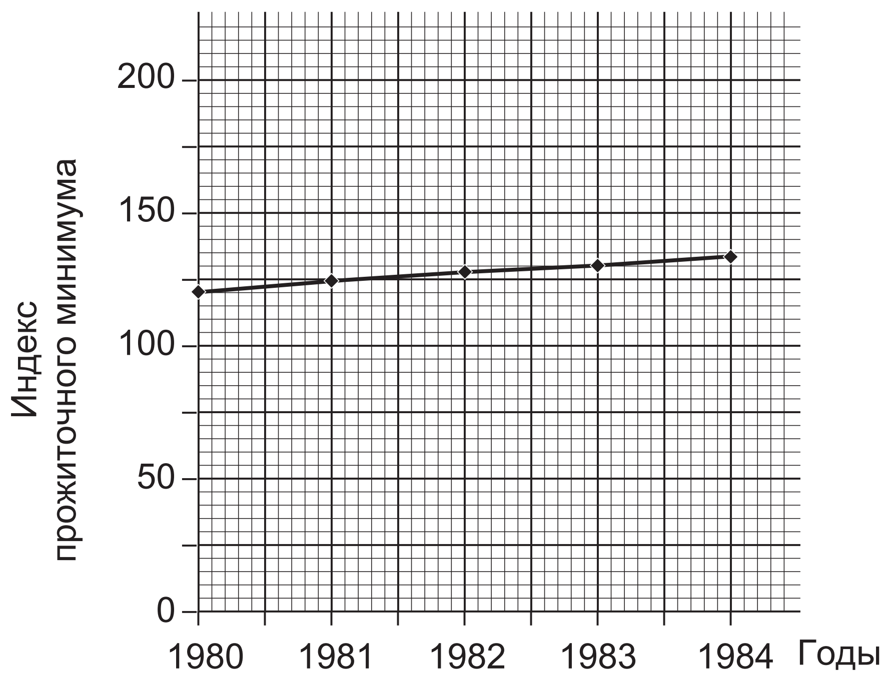
Наиболее распространенными способами обманчивого представления данных являются:

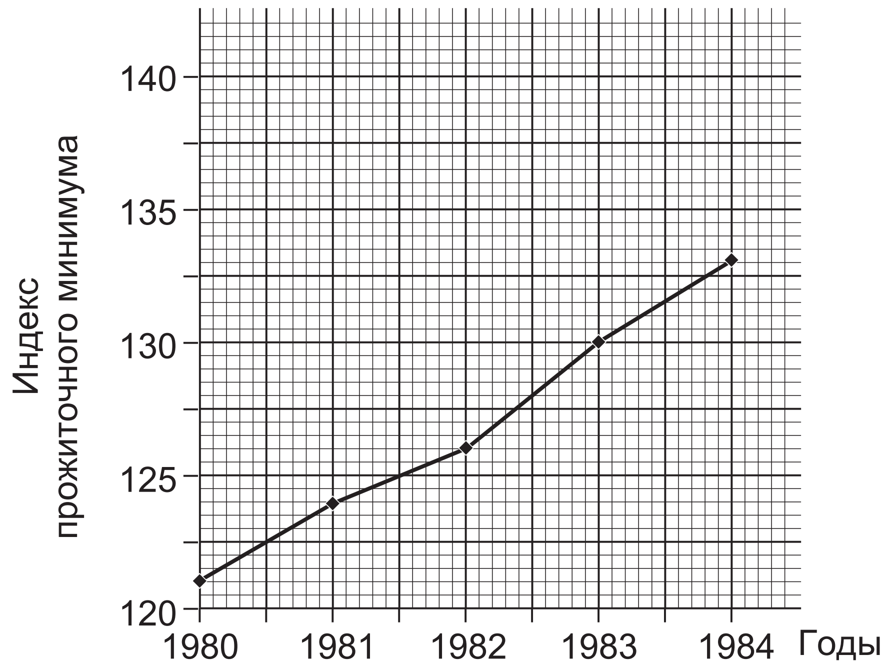
– использование разрыва масштабной шкалы;
– использование цветовой гаммы при построении круговых или столбиковых диаграмм;
– использование площадей квадратов или кругов для представления количественных данных.

(i) Использование разрыва масштабной шкалы

Используя «разрыв масштабной шкалы», т.е. шкалу, которая начинается не с нуля, можно как завысить, так и занизить величину вариации. Например, приведенные ниже данные могут быть графически представлены двумя способами, производящими совершенно разное впечатление.

Год	1980	1981	1982	1983	1984
Индекс прожиточного минимума	121	124	126	130	133

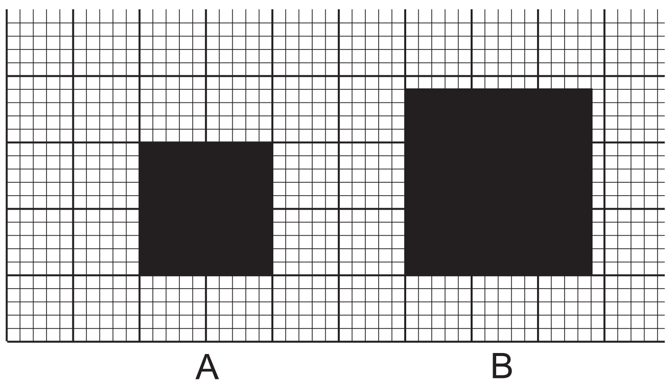




(ii) Использование цветовой гаммы

Яркие цвета, такие как красный, оранжевый и желтый, имеют свойство визуально увеличивать объекты. Если в секторной диаграмме один сектор закрашен ярко-красным, а остальные – тусклыми, неяркими цветами, то ярко-красный сектор будет казаться больше, чем он есть на самом деле. Таким образом, доля переменной, представленной ярко-красным сектором, будет казаться больше, чем в действительности. Следовательно, такая секторная диаграмма введет в заблуждение.

(iii) Использование площадей квадратов или кругов для представления количественных данных



В представленной выше диаграмме площадь квадрата B точно в два раза больше площади квадрата A. Однако для большинства людей квадрат B не выглядит в два раза больше квадрата A. Это потому, что его сторона лишь в 1,414 раза больше стороны квадрата A. В результате так называемого «оптического обмана» создается впечатление, что площади различаются как нечто среднее между отношением сторон (1,414) и площадей (2).

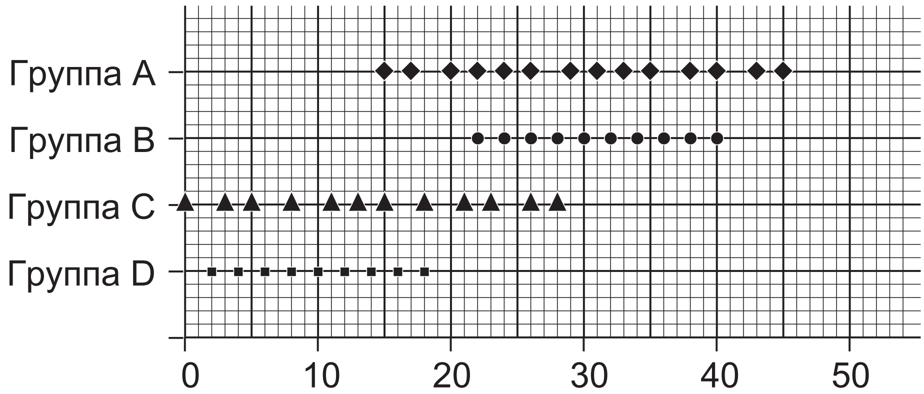
Так, многие подумали бы, что значение, представленное площадью квадрата B, примерно в 1,7 раза больше значения, выраженного квадратом A. Следовательно, диаграмма обманчива.

Сказанное выше справедливо и для случая двух кругов, даже если это обычный метод построения двух сопоставимых круговых диаграмм, который мы рассматривали в разделе 1.5.

Глава 2

СПОСОБЫ ВЫЧИСЛЕНИЯ СРЕДНИХ

Приведенная ниже диаграмма иллюстрирует оценки, полученные студентами четырех параллельных групп за один и тот же тест.



Сравнивая эти оценки, мы видим, что все четыре группы различаются.

Группа A получила в целом высокие оценки, но они сильно варьируются.

Группа B получила в целом высокие оценки, и они располагаются тесно друг к другу.

Группа C в целом получила низкие оценки, и они сильно варьируются.

Группа D в целом получила низкие оценки, но они располагаются тесно друг к другу.

Различия между группами можно анализировать по двум факторам:

(i) **среднему значению** оценок;

(ii) **колеблемости,** или **вариации,** оценок.

Результаты такого анализа могут быть представлены в виде следующей таблицы.

Среднее значение	Колеблемость, или вариация	
	Малая	Большая
Малое (низкое)	D	C
Большое (высокое)	B	A

Большинство статистических работ включает в себя расчет двух основных показателей: средних величин и колеблемости, или вариации.

Вначале мы рассмотрим способы вычисления среднего значения.

2.1. Мода

Мода представляет собой самое распространенное значение; значение, повторяющееся чаще остальных, т.е. с наибольшей частотой.

2.1.1. Мода дискретного ряда распределения

Пример 1. В группе из 11 студентов получены следующие оценки за тест (максимальное число – 10 баллов).

5, 4, 3, 7, 9, 5, 6, 2, 0, 5, 6.

Определите моду данного ряда распределения.

Решение. Имеются 3 оценки в 5 баллов, 2 оценки в 6 баллов, а остальные встречаются только один раз. Следовательно, мода равна 5.

Пример 2. Имеются следующие данные о количестве братьев и сестер у 20 учеников в классе:

2, 0, 1, 3, 1, 2, 2, 2, 2, 1, 0, 1, 2, 1, 3, 1, 4, 1, 2, 2.

Определите моду ряда распределения.

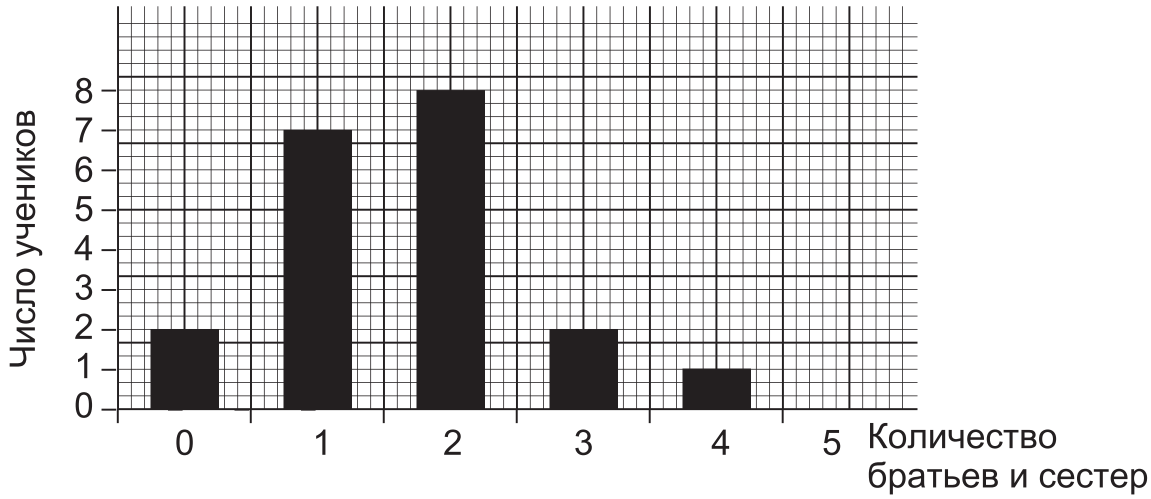
Решение. В этом примере сложно сразу определить, какое значение встречается чаще всего. Поэтому мы запишем распределение частот по исходным данным. (Обратите внимание, что символ *////* означает 5 единиц.)

Количество братьев/сестер, X	Единица счета	Частота, f
0	//	2
1	~~////~~ //	7
2	~~////~~ ///	8
3	//	2
4	/	1
Суммарная частота		20 (проверка)

Теперь намного легче увидеть, что значением с наибольшей частотой является 2.

Следовательно, мода данного ряда распределения – 2.

Данное распределение можно проиллюстрировать с помощью столбиковой гистограммы.



Мы видим, что самый высокий столбец и является модальным значением.

Пример 3. Имеются данные о размере обуви 11 девочек в классе.

5, 3, 4, 3, 5, 4, 3, 4, 2, 6, 2.

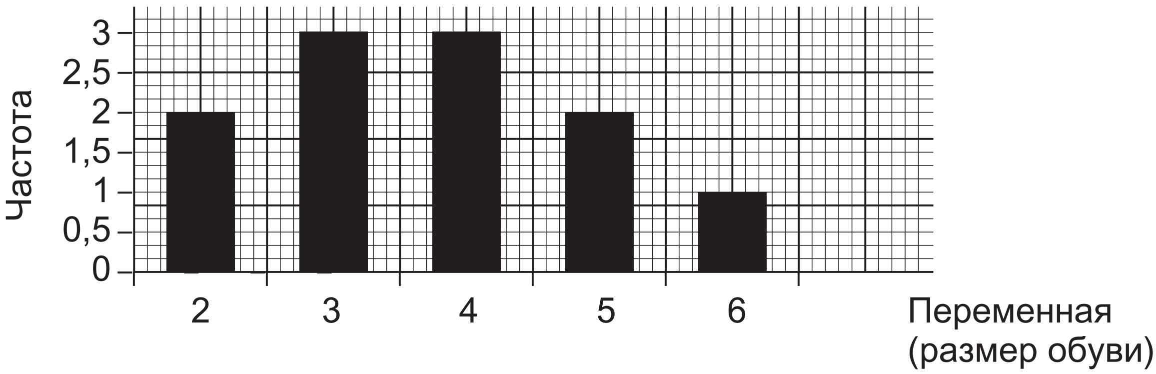
Определить модальный размер обуви (моду размеров).

Решение. Запишем в таблицу полученное распределение частот.

Размер обуви, X	Единица счета	Частота, f
2	//	2
3	///	3
4	///	3
5	//	2
6	/	1
Суммарная частота		11 (проверка)

Поскольку наибольшую частоту имеют два соседних размера обуви 3 и 4, то модальное значение будет равно $3^1/_2$.

Снова проиллюстрируем распределение с помощью столбиковой гистограммы.



Аудиторный тест № 3

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По команде преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите моду данных десяти чисел.

(Будьте внимательны и не включайте в данный ряд чисел номер вопроса.)

Вопрос		**Варианты ответов**				
		A	B	C	D	E
1.	1, 2, 3, 1, 2, 4, 5, 2, 0, 6	1	2	3	4	5
2.	8, 7, 4, 0, 6, 4, 5, 1, 2, 3	2	3	4	5	6
3.	2, 2, 3, 6, 1, 5, 6, 7, 6, 4	2	3	4	5	6
4.	4, 3, 5, 4, 5, 4, 5, 5, 4, 5	2	3	4	5	6
5.	6, 8, 7, 8, 7, 8, 6, 8, 6, 9	5	6	7	8	9
6.	0, 1, 3, 0, 2, 3, 1, 2, 1, 1	0	1	2	3	6
7.	6, 4, 7, 6, 7, 2, 7, 5, 6, 7	2	4	5	6	7
8.	8, 9, 7, 7, 6, 9, 5, 8, 6, 9	5	6	7	8	9
9.	0, 3, 2, 2, 2, 3, 3, 1, 0, 3	0	1	2	3	4
10.	4, 5, 3, 0, 5, 0, 4, 3, 0, 2	0	2	3	4	5
11.	2, 3, 1, 0, 4, 3, 0, 1, 3, 2	0	1	2	3	4
12.	9, 8, 8, 8, 9, 7, 6, 9, 6, 9	5	6	7	8	9
13.	8, 9, 10, 7, 6, 8, 10, 9, 8, 6	6	7	8	9	10
14.	2, 3, 1, 4, 5, 1, 3, 1, 2, 6	1	2	3	4	5
15.	6, 6, 7, 5, 5, 7, 8, 8, 9, 7	5	6	7	8	9
16.	4, 5, 3, 2, 5, 1, 4, 3, 6, 5	1	2	3	4	5
17.	2, 1, 0, 3, 2, 3, 4, 1, 0, 2	0	1	2	3	4
18.	4, 5, 5, 5, 4, 3, 3, 4, 2, 4	1	2	3	4	5
19.	5, 6, 7, 8, 9, 4, 6, 7, 5, 6	5	6	7	8	9
20.	9, 7, 6, 6, 7, 6, 7, 8, 7, 9	5	6	7	8	9

Аудиторный тест № 4

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По команде преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найти моду данного распределения частот.

Вопрос			Варианты ответов				
			A	B	C	D	E
1.	x f	2 3 4 5 6 6 5 5 4 2	2	3	4	5	6
2.	x f	1 2 3 4 5 3 5 4 2 1	1	2	3	4	5
3.	x f	1 2 3 4 5 2 4 5 3 1	1	2	3	4	5
4.	x f	3 4 5 6 7 8 9 2 3 5 7 4 2 1	5	6	7	8	9
5.	x f	4 5 6 7 8 9 10 1 3 4 4 4 2 1	6	6½	7	7½	8
6.	x f	5 6 7 8 9 10 11 3 5 8 8 6 4 2	6½	7	7½	8	8½
7.	x f	7 8 9 10 11 12 13 1 2 5 8 9 6 3	8	9	10	11	12
8.	x f	4 5 6 7 8 9 10 5 9 8 6 4 2 1	4	5	6	7	8

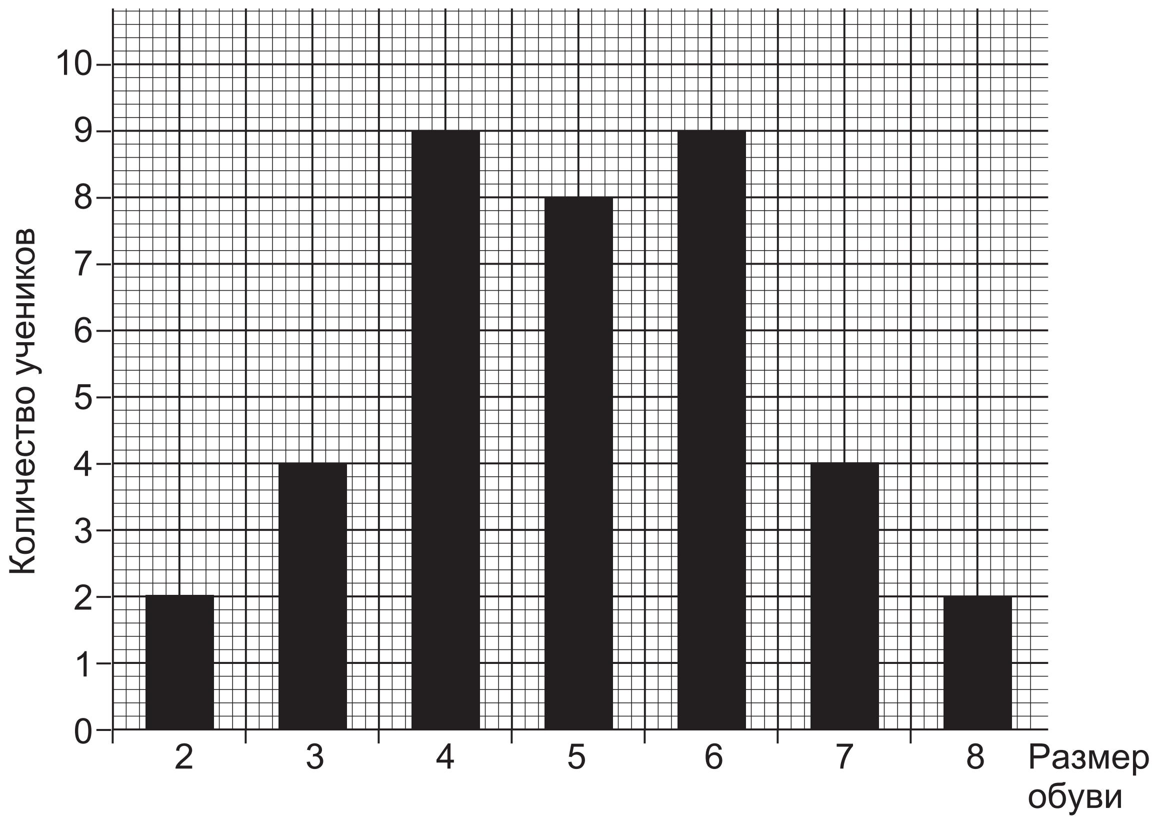
2.1.2. Бимодальные распределения

Некоторые распределения являются бимодальными. Это означает, что они имеют две моды. Как правило, это наблюдается в том случае, когда распределение является суммой, или комбинацией, двух различных распределений.

Пример. Имеются данные о размере обуви 38 учеников в классе:

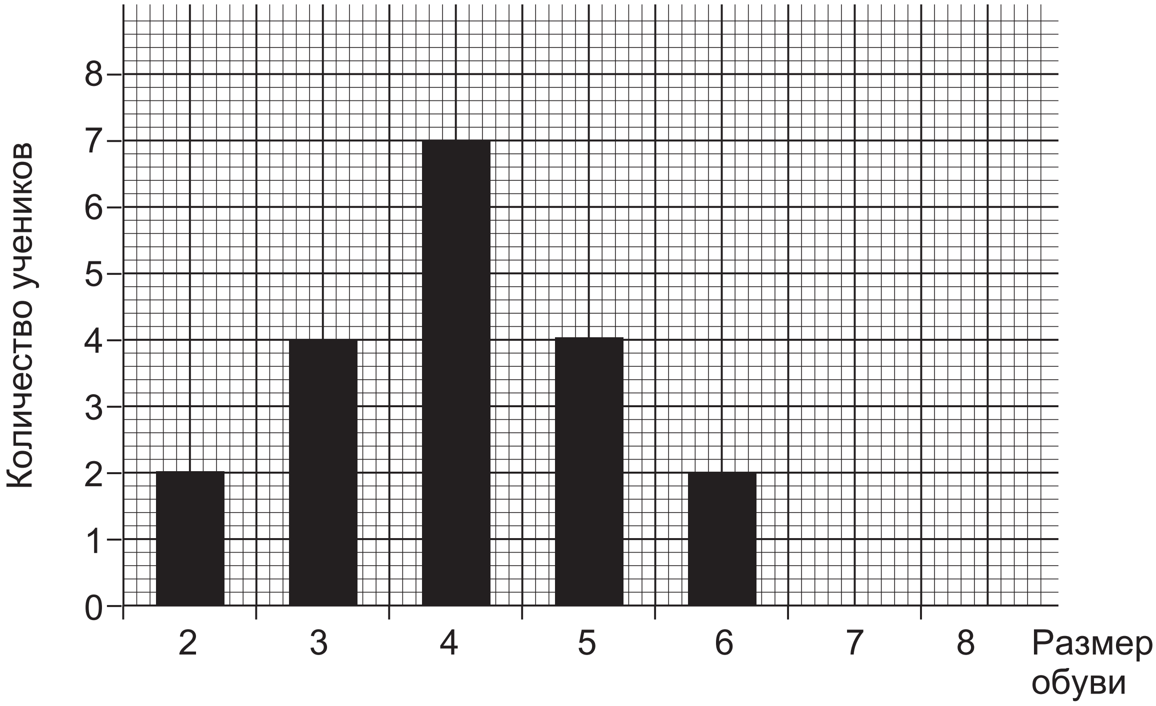
Размер обуви, *x*	2	3	4	5	6	7	8
Частота, *f*	2	4	9	8	9	4	2

Приведенная ниже столбиковая диаграмма иллюстрирует это распределение.

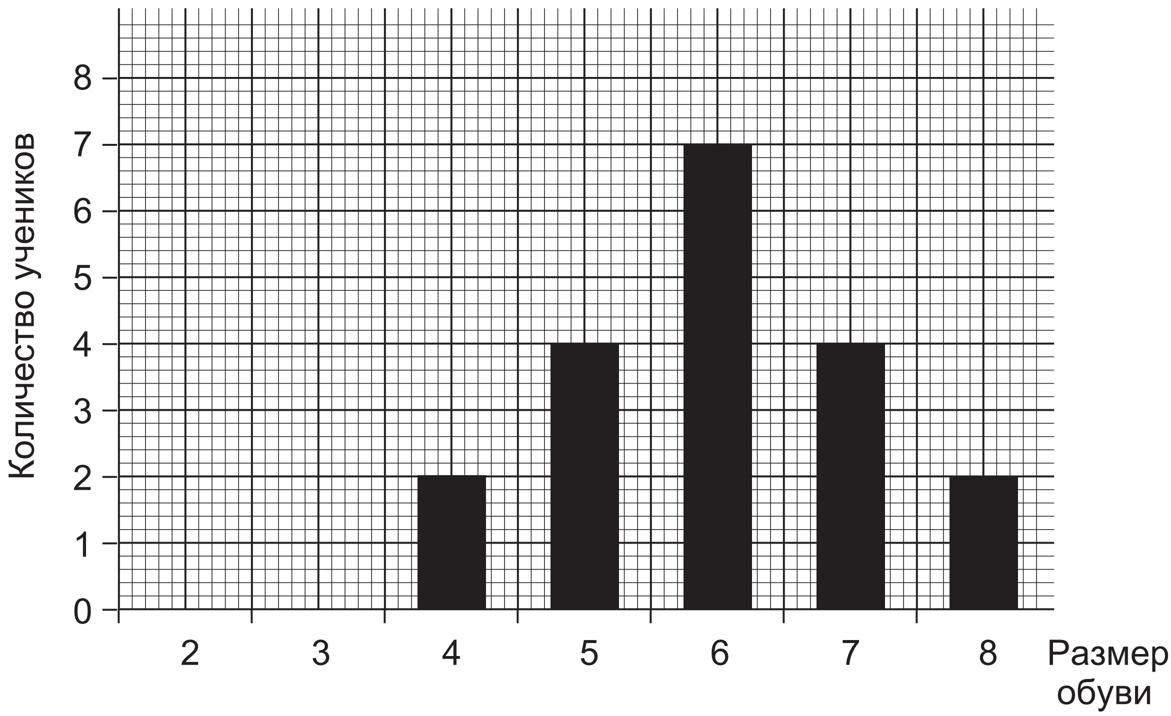


При дальнейшем исследовании обнаруживается, что распределение является суммой двух различных распределений: для мальчиков и для девочек.

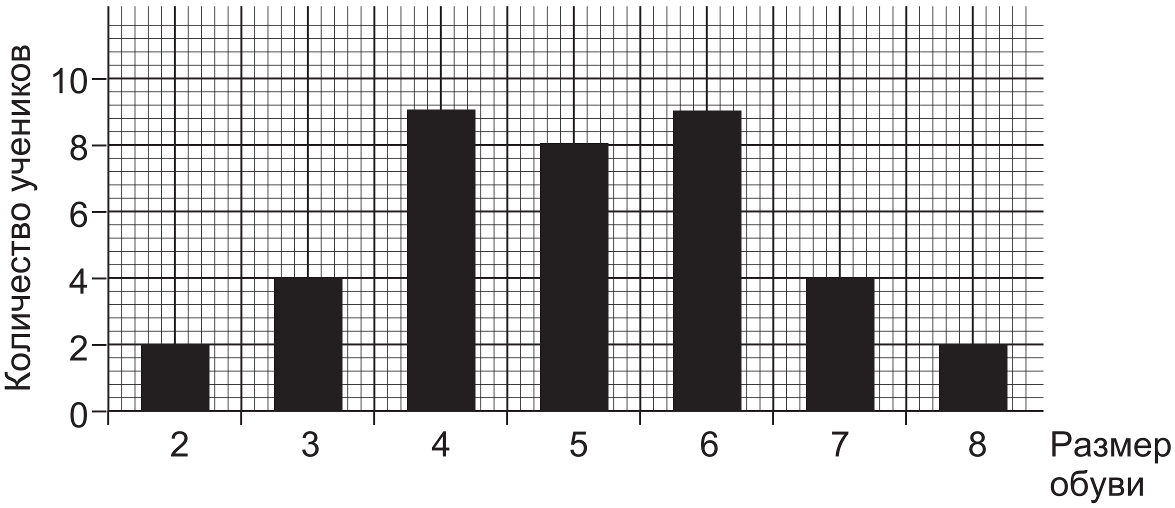
Столбиковая диаграмма распределения для девочек.



Столбиковая диаграмма распределения для мальчиков.



Столбиковая диаграмма распределения для мальчиков и девочек.



Упражнение 10А

1. Определите модальное значение для чисел:

2, 3, 0, 6, 5, 6, 1, 7, 6.

2. Постройте ряд распределения частот для следующих 18 чисел:

7, 12, 9, 9, 11, 10, 10, 7, 8, 8, 10, 11, 9, 10, 9, 11, 9, 8.

Затем определите моду полученного ряда.

3. Постройте ряд распределения по числу букв в слове в следующем предложении, содержащем семь слов:

«Модой называется значение признака с наибольшей частотой».

Затем найдите модальное число букв в слове (моду числа букв в слове).

Упражнение 10В

1. Определите моду ряда чисел:

7, 8, 0, 3, 9, 7, 8, 2, 10, 8.

2. Постройте распределение частот для следующих 30 чисел:

6, 12, 7, 12, 9, 7, 11, 10, 8, 7, 8, 11, 8, 11, 10,
10, 11, 9, 9, 10, 10, 10, 8, 9, 12, 12, 11, 11, 10, 9.

Затем определите моду распределения.

3. Постройте ряд распределения по числу букв в слове в следующем предложении, содержащем 12 слов:

«Цель графического представления данных – сделать общую модель легко воспринимаемой с первого взгляда».

Затем найдите модальное число букв в слове (моду числа букв в слове).

2.2. Мода интервального ряда распределения

Моду интервального ряда распределения можно определить двумя способами:

(i) построением гистограммы;

(ii) при помощи расчета.

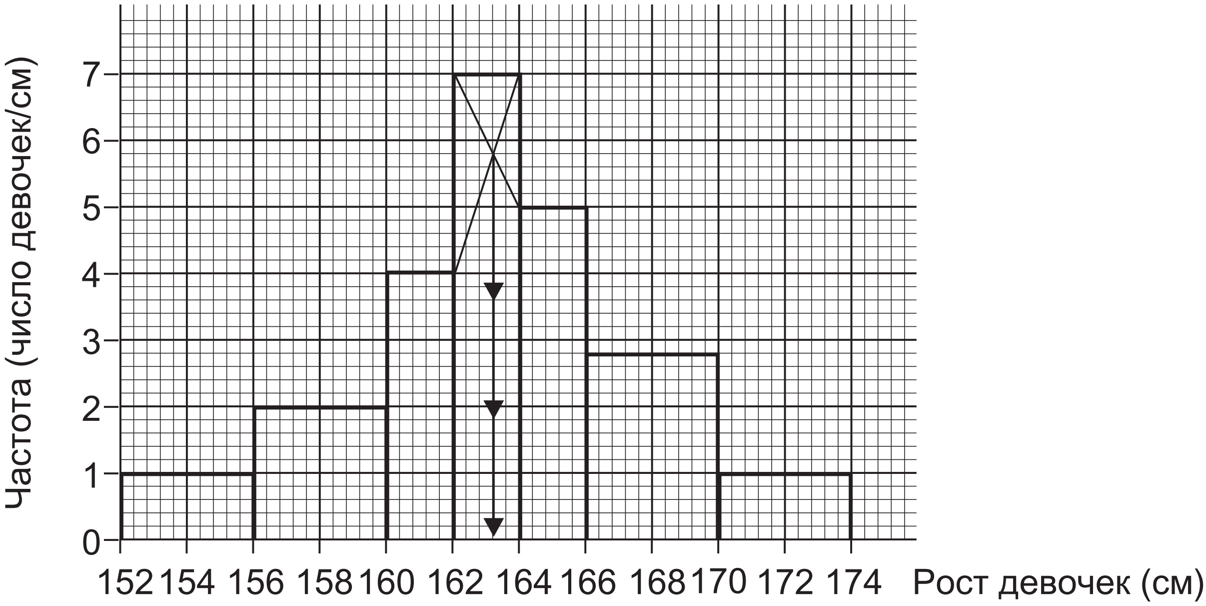
Пример 1. Определить моду данного ряда распределения.

Рост девочек условной возрастной группы.

Группировка Рост девочек (см)	*Частота* Количество девочек
$152 \leq X < 156$	4
$156 \leq X < 160$	8
$160 \leq X < 162$	8
$162 \leq X < 164$	14
$164 \leq X < 166$	10
$166 \leq X < 170$	12
$170 \leq X \leq 174$	4

Решение.

(i) Построением гистограммы

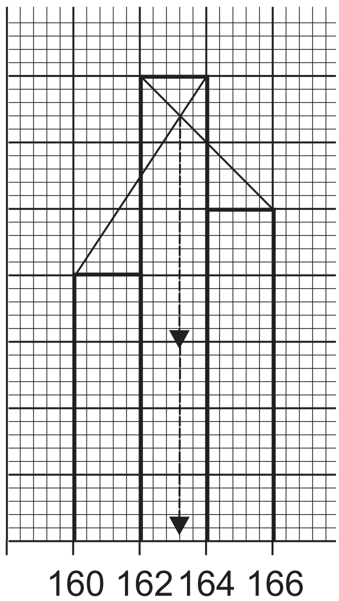


Модальным интервалом называется интервал, которому соответствует столбец гистограммы с максимальной высотой. В нашем примере это интервал $162 \leq X < 164$.

Для определения моды проведем диагонали через вершины самого высокого столбца (модального интервала), как показано на гистограмме.

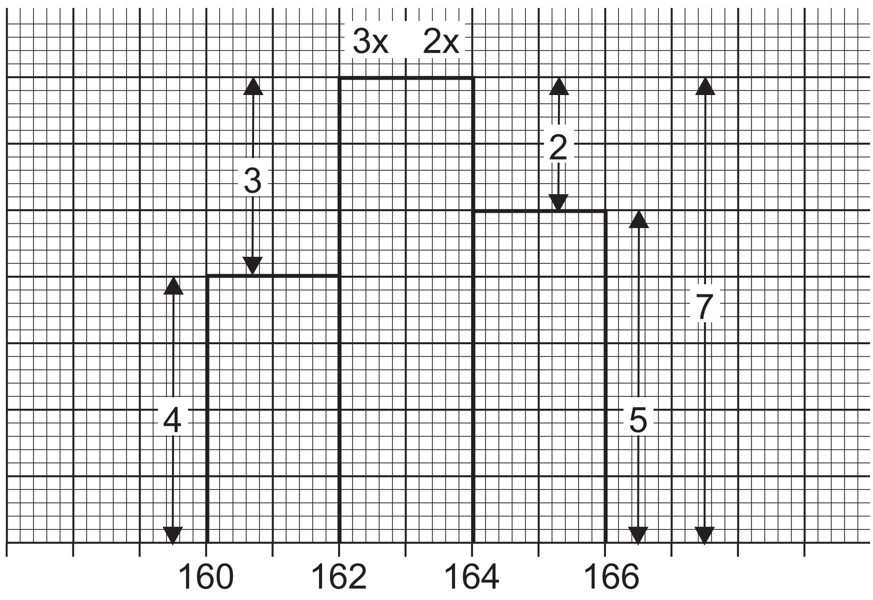
Мода определяется точкой пересечения этих двух линий. В нашем примере мода равна 163,2 см.

Если соседние с модальным интервалом столбцы имеют одинаковую ширину **(но только в этом случае)**, можно определить моду другим построением, как показано на гистограмме ниже.



Второе построение может служить хорошей проверкой.

(ii) При помощи расчета



Мы делим расстояние от 162 до 164 (границы модального интервала) в отношении разностей высот модального и соседних с ним столбцов.

В нашем примере расстояние от 162 до 164 равно 2.

Разности высот равны соответственно 7 – 4 = 3 и 7 – 5 = 2.

Разделим расстояние от 162 до 164 в отношении 3 : 2, так что части, или доли, составляют 3/5 и 2/5.

Мода равна $162 + 3/5 \times 2 = 162 + 1{,}2 = 163{,}2$ или $164 - 2/5 \times 2 = 164 - 0{,}8 = 163{,}2$.

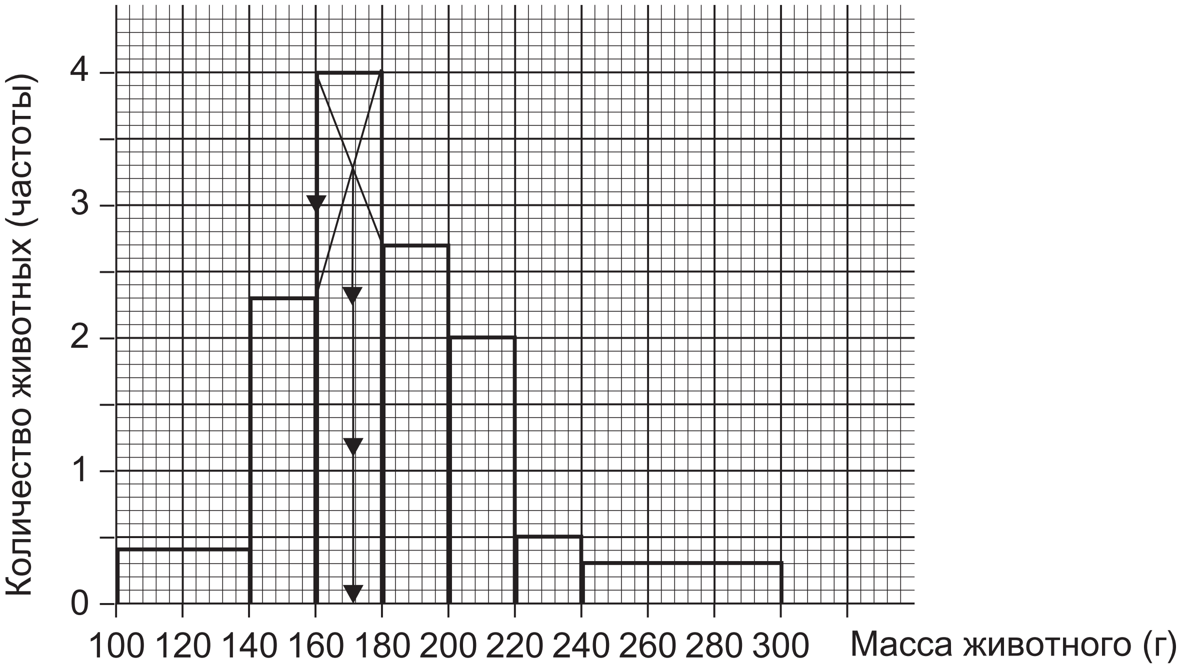
Определение моды двумя способами является эффективной проверкой.

Пример 2. Определите моду данного ряда распределения.

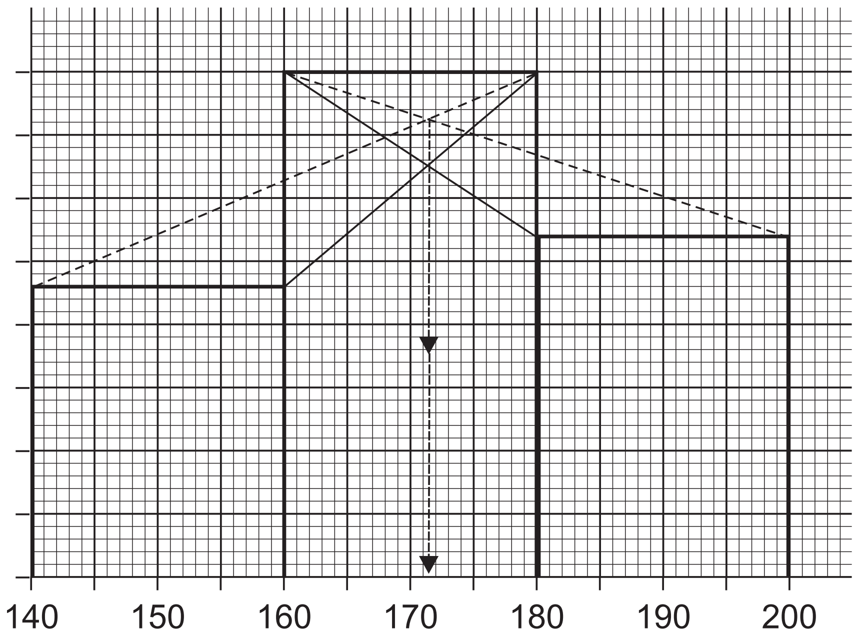
Масса животного (г)	100-	140-	160-	180-	200-	220-	240-300
Количество животных	16	46	76	54	40	10	18

Решение.

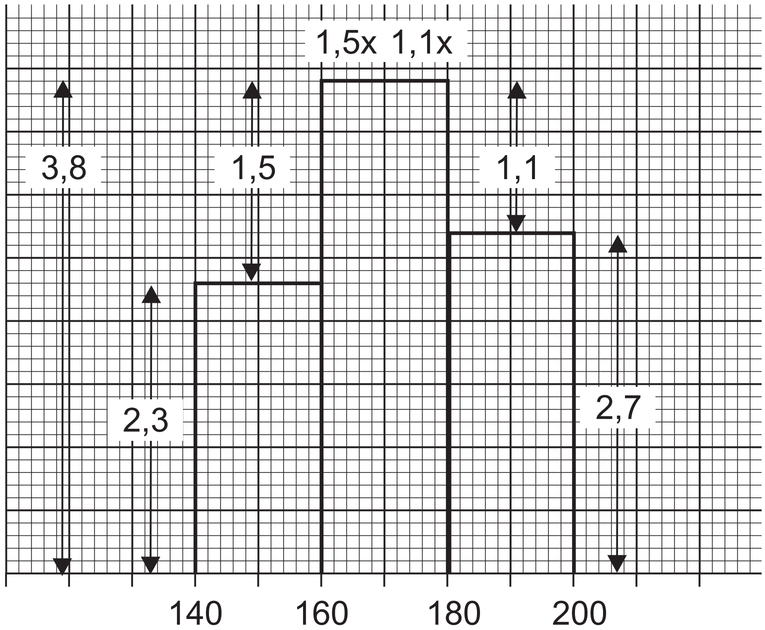
(i) Построением гистограммы



Мода = 171,5 г



(ii) При помощи расчета



Разделим расстояние от 160 до 180 (границы модального интервала) в отношении разностей высот модального и соседних с ним столбцов.

В нашем примере расстояние от 160 до 180 равно 20.

Разности высот равны соответственно 3,8 – 2,3 = 1,5 и 3,8 – 2,7 = = 1,1.

Мы делим расстояние от 160 до 180 в отношении 1,5 : 1,1 (или 15:11), так что части, или доли, составляют 15/26 и 11/26.

Мода равна 160 + 15/26 × 20 = 171,5 или 180 – 11/26 × 20 = = 171,5.

Определение моды двумя способами является эффективной проверкой.

Упражнение 11 А

1. Ниже приведено распределение рабочих некоего завода по заработкам.

Заработки (ф.ст.)	3500-	3600-	3700-	3800-	3900-	4000-4100
Количество рабочих	4	8	18	30	13	7

a) Проиллюстрируйте распределение с помощью гистограммы. Запишите границы модального интервала и определите графически моду данного ряда распределения.

b) Проверьте ответ, рассчитав модальное значение.

2. В таблице даны результаты опроса 40 сотрудников городской конторы о времени, затрачиваемом ими утром на дорогу до работы.

Время (с точностью до минуты)	0-9	10-19	20-24	25-29	30-34	35-39	40-50
Количество сотрудников	2	6	8	9	6	5	4

a) Запишите границы модального интервала.

b) Проиллюстрируйте распределение с помощью гистограммы.

c) По гистограмме определите модальное значение времени в пути.

d) Рассчитайте модальное значение времени в пути.

Упражнение 11 В

1. Был проведен опрос 400 школьников о времени, которое у них ежедневно уходит на просмотр телепередач. Полученные результаты приведены ниже в таблице.

Время просмотра (мин)	0-	20-	30-	40-	50-	60-100
Количество детей	40	40	60	90	50	120

a) Проиллюстрируйте распределение с помощью гистограммы.

b) Запишите границы модального интервала и определите моду ряда распределения.

c) Определите модальное значение времени просмотра телепередач.

d) Рассчитайте модальное значение и сравните результаты.

2. Ниже приведены результаты 50 участников легкоатлетического соревнования по метанию, измеренные с точностью до метра.

Расстояние (м)	9-14	15-17	18-20	21-23	24-26	27-33
Количество спортсменов	5	9	15	10	7	4

a) Запишите границы всех интервалов распределения.

b) Проиллюстрируйте распределение с помощью гистограммы.

c) Запишите границы модального интервала и определите моду распределения.

d) Определите модальное значение по гистограмме.

e) Рассчитайте модальное значение.

2.3. Медиана

Медианой распределения является такое значение, что одна половина распределения или единиц совокупности находится ниже, а другая половина распределения или единиц совокупности – выше этого значения.

Медиана дискретного ряда распределения

2.3.1. Случай НЕЧЕТНОГО числа членов ряда, или значений

Если n – нечетное число, то после расположения всех значений в порядке возрастания медианой будет $\frac{1}{2}(n+1)$-е значение.

Пример 1. Найдите медиану чисел

2, 0, 2, 7, 5, 8, 10.

Решение. Вначале расположим числа в порядке возрастания:

0, 2, 2, 5, 7, 8, 10.

Имеется 7 значений, т.е. $n = 7$.

Медианой является $\frac{1}{2}(n+1)$-е значение.

В нашем примере это $\frac{1}{2}(7+1)$-е, т.е. 4-е значение, которое равно 5. Следовательно, медиана равна 5.

Пример 2. Найдите медиану чисел

7, 2, 6, 3, 9, 8, 5, 4, 7.

Решение. Переставим числа в порядке возрастания:

2, 3, 4, 5, 6, 7, 7, 8, 9.

У нас имеется 9 значений, т.е. $n = 9$.

Медианой является $\frac{1}{2}(n+1)$-е значение.

В нашем примере это $\frac{1}{2}(9+1)$-е, т.е. 5-е значение, которое равно 6. Следовательно, медиана равна 6.

Медианой ряда распределения с **НЕЧЕТНЫМ** числом значений после расположения всех чисел в порядке возрастания является центральное значение.

2.3.2. Случай ЧЕТНОГО числа членов ряда, или значений

Если n – четное число, то после расположения всех значений в порядке возрастания медиана будет равна простой средней $n/2$-го и $(n/2+1)$-го значений.

Пример 1. Найдите медиану чисел

2, 5, 10, 0, 8, 2, 7, 10.

Решение. Расположим числа в порядке возрастания:

0, 2, 2, 5, 7, 8, 10, 10.

Ряд содержит 8 значений, т.е. $n = 8$.

Медиана равна простой средней $n/2$-го и $(n/2+1)$-го значений. В нашем примере медианой является простая средняя 4-го и 5-го значений.

4-е значение равно 5;

5-е значение равно 7.

Таким образом, медиана ряда распределения равна $\frac{1}{2}(5+7) = 6$.

Пример 2. Найдите медиану чисел

8, 6, 5, 3, 9, 2, 8, 6, 5, 4, 2, 1.

Решение. Сперва расположим числа в порядке возрастания:

1, 2, 2, 3, 4, 5, 5, 6, 6, 8, 8, 9.

Ряд содержит 12 значений, т.е. $n = 12$.

Медиана равна простой средней $n/2$-го и $(n/2+1)$-го значений. В нашем примере медиана равна простой средней 6-го и 7-го значений.

6-е значение равно 5.

7-е значение равно 5.

Медиана ряда распределения равна 5.

В случае **ЧЕТНОГО** числа после расположения всех значений в порядке возрастания имеется не одно, а два центральных значения.

Тогда медианой будет простая средняя двух центральных значений (их следует сложить и разделить на 2) после расположения всех значений в порядке возрастания.

Аудиторный тест № 5

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По команде преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Определите медиану данных чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	11, 7, 13, 9, 12	8	9	10	11	12
2.	4, 3, 2, 2, 4	1	2	3	4	5
3.	8, 4, 8, 5, 8, 5, 7	4	5	6	7	8
4.	3, 9, 8, 4, 6, 7, 4	4	6	7	8	9
5.	5, 4, 5, 3, 1, 2, 5	1	2	3	4	5
6.	6, 5, 5, 7, 8, 5, 5	5	6	7	8	9
7.	10, 1, 9, 7, 3, 9, 6, 10	6	7	8	9	10
8.	12, 8, 4, 9, 10	8	9	10	11	12
9.	10, 7, 7, 3, 5, 8, 9, 7	7	8	9	10	11
10.	13, 9, 7, 6, 9, 12	6	7	8	9	10
11.	13, 9, 7, 6, 7, 12	6	7	8	9	10
12.	1, 4, 5, 1, 6, 7, 1	3	4	5	6	7
13.	5, 4, 7, 8, 9, 7, 2, 3, 6	5	6	7	8	9
14.	7, 5, 9, 8, 5, 5	5	6	7	8	9
15.	12, 8, 6, 2, 4, 10	4	5	6	7	8
16.	4, 0, 4, 1, 4, 2, 3	0	1	2	3	4
17.	11, 8, 9, 10, 11, 11, 11, 12	7	8	9	10	11
18.	13, 10, 8, 12, 12, 8, 6, 10	6	8	10	12	13
19.	12, 5, 7, 7, 7, 10, 9, 9, 8	5	6	7	8	9
20.	10, 2, 9, 3, 8, 4, 7, 6	6	7	8	9	10

2.3.3. Медиана распределения частот дискретных значений

Пример 1. Найдите медиану ряда распределения, приведенного в таблице.

Значение, x	2	3	4	5	6	7
Частота, f	5	7	9	13	6	3

Решение. Суммарная частота равна 43, т.е. $n = 43$.

Имеется нечетное число значений, поэтому медиана равна $\frac{1}{2}(n+1)$-му, т.е. 22-му значению.

Из таблицы распределения частот легко можно увидеть, что

первые 5 значений, т.е. с 1-го по 5-е, все равны 2;
следующие 7 значений с 6-го по 12-е все равны 3;
следующие 9 значений с 13-го по 21-е все равны 4;
следующие 13 значений с 22-го по 34-е все равны 5.

Таким образом, 22-е значение равно 5, следовательно, медиана распределения равна 5.

Пример 2. Найдите медиану приведенного ниже ряда распределения.

Значение, x	5	6	7	8	9	10
Частота, f	3	10	5	3	2	1

Решение. Суммарная частота равна 24, т.е. $n = 24$.

Мы имеем четное число значений, поэтому медиана равна простой средней $n/2$-го и $(n/2+1)$-го значений.

В нашем примере медианой будет простая средняя 12-го и 13-го значений.

Из таблицы видно, что
первые 3 значения, т.е. с 1-го по 3-е все равны 5;
следующие 10 значений с 4-го по 13-е все равны 6;
следующие 6 значений с 14-го по 19-е все равны 7.

Таким образом, 12-е значение равно 6, 13-е значение также 6. Следовательно, медиана распределения равна 6.

Пример 3. Найдите медиану приведенного ниже ряда распределения.

Значение, x	9	10	11	12	13	14
Частота, f	4	5	9	7	6	5

Решение. Суммарная частота равна 36, т.е. $n = 36$.

Мы имеем четное число значений, поэтому медиана равна простой средней $n/2$-го и $(n/2+1)$-го значений.

В нашем примере медианой будет простая средняя 18-го и 19-го значений.

Из таблицы видно, что

первые 4 значения, т.е. с 1-го по 4-е все равны 9;
следующие 5 значений с 5-го по 9-е все равны 10;
следующие 9 значений с 10-го по 18-е все равны 11;
следующие 7 значений с 19-го по 25-е все равны 12.

Таким образом, 18-е значение равно 11, а 19-е значение – 12.

Медианой является простая средняя 11 и 12, т.е. ½(11+12), или 11,5.

Медиана распределения равна 11,5.

Аудиторный тест № 6

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Определите медиану данного распределения частот.

Вопрос			**Варианты ответов**				
			A	B	C	D	E
1.	x	2 3 4 5 6	2	3	4	5	6
	f	6 5 5 4 2					
2.	x	1 2 3 4 5	1	2	3	4	5
	f	3 5 4 2 1					
3.	x	1 2 3 4 5	1	2	3	4	5
	f	2 4 5 3 1					
4.	x	3 4 5 6 7 8 9	5	6	7	8	9
	f	2 3 5 7 4 2 1					
5.	x	4 5 6 7 8	5	5½	6	6½	7
	f	4 6 10 12 8					
6.	x	4 5 6 7 8	5	5½	6	6½	7
	f	9 12 8 7 5					
7.	x	0 1 2 3 4	0	1	2	3	4
	f	3 5 7 6 2					
8.	x	4 5 6 7 8	4	5	6	7	8
	f	4 6 8 12 7					

2.4. Медиана интервального распределения частот

Медиану интервального распределения частот можно определить:

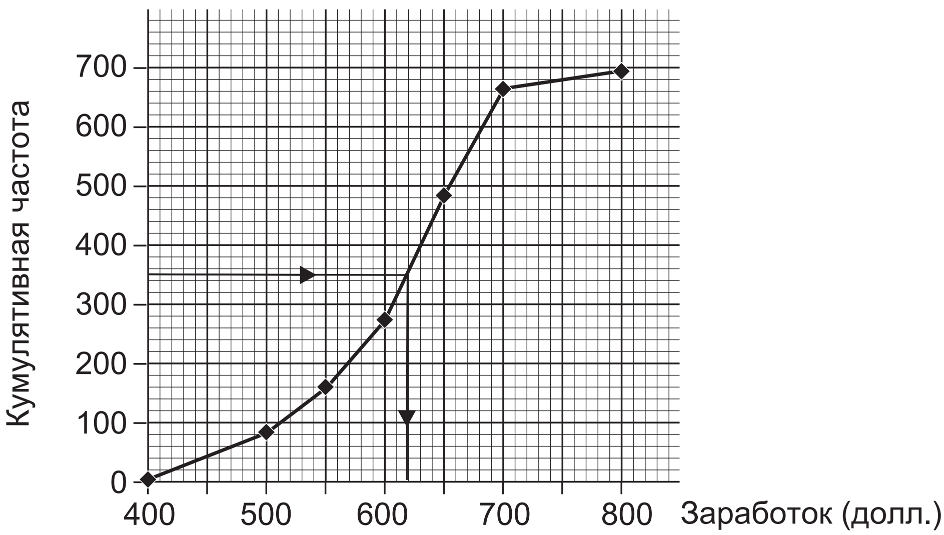
– по диаграмме кумулятивных частот;
– расчетами по таблице кумулятивных частот.

2.4.1. Определение медианы по диаграмме кумулятивных частот

Пример 1. Найдите медиану следующего интервального распределения частот.

Группировка Заработок (долл.)	*Частота* Количество работников
$400 \le X < 500$	80
$500 \le X < 550$	70
$550 \le X < 600$	120
$600 \le X < 650$	210
$650 \le X < 700$	180
$700 \le X \le 800$	40
Всего	700

Решение. Сначала составим таблицу кумулятивных частот, а затем построим полигон кумулятивных частот, который представлен ниже.



Чтобы определить медиану, проведем горизонтальную линию до пересечения с полигоном на высоте, строго равной половине суммарной частоты.

В нашем примере суммарная частота равна 700.

Проведем горизонтальную прямую, соответствующую кумулятивной частоте 350. Обратите внимание, **ровно посередине**: не 350½ и не 351, а именно 350. Из точки пересечения с полигоном проведем вертикальную линию вниз до пересечения с горизонтальной осью X.

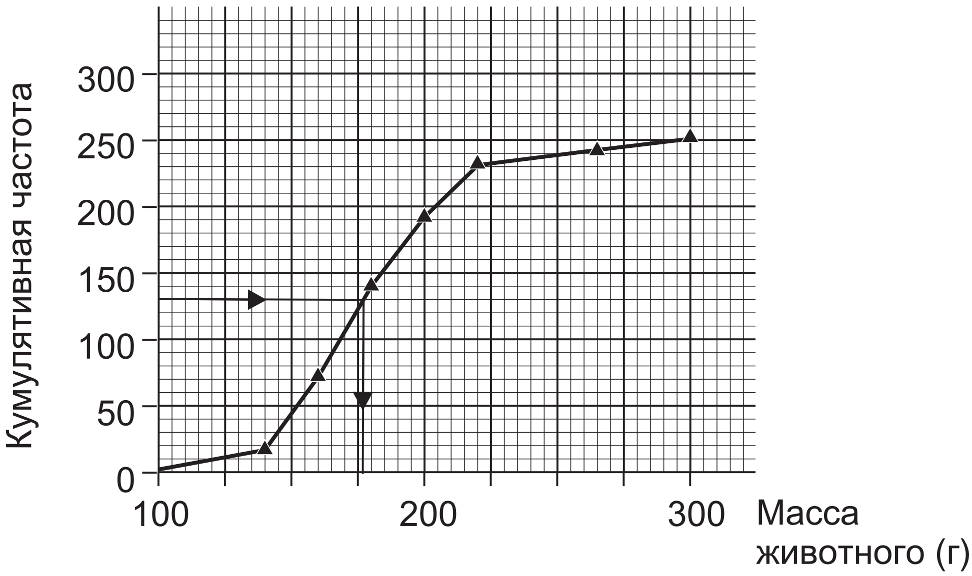
Это и есть значение медианы.

В нашем примере медиана равна 619 долл.

Пример 2. Найдите медиану представленного ниже интервального распределения частот.

Группировка Масса животного (г)	*Частота* Количество животных
$100 \le X < 140$	16
$140 \le X < 160$	46
$160 \le X < 180$	76
$180 \le X < 200$	54
$200 \le X < 220$	40
$220 \le X < 240$	10
$240 \le X \le 300$	18
Суммарная частота	260

Решение. Сначала составим таблицу кумулятивных частот, а затем построим полигон кумулятивных частот, который изображен ниже.



Для определения медианы проведем горизонтальную линию до пересечения с полигоном на высоте, строго равной половине суммарной частоты.

В нашем примере суммарная частота равна 260.

Проведем горизонтальную прямую, соответствующую кумулятивной частоте 130. Обратите внимание, что это **ровно половина**: 130, а не 130½ и не 131. Затем из точки пересечения с полигоном проведем вертикальную линию вниз до пересечения с горизонтальной осью X.

Это и есть значение медианы.

В нашем примере медиана равна 177,9.

2.4.2. Расчет медианы по таблице кумулятивных частот

Пример 1. Определите медиану представленного ниже интервального распределения частот.

Группировка Заработки (долл.)	*Частота* Количество работников
$400 \le X < 500$	80
$500 \le X < 550$	70
$550 \le X < 600$	120
$600 \le X < 650$	210
$650 \le X < 700$	180
$700 \le X \le 800$	40
Суммарная частота	700

Решение.

Сначала составим таблицу кумулятивных частот.

Заработки (долл.) «до», или «меньше чем»	Кумулятивная частота
400	0
500	80
550	150
600	270
650	480
700	660
800	700

Суммарная частота равна 700, половина будет составлять 350.

Из таблицы кумулятивных частот мы видим, что кумулятивная частота 350 находится между 600 долл. (кумулятивная частота 270) и 650 долл. (кумулятивная частота 480).

600 долл. 270
>80
? 350
>130
650 долл. 480

Для того чтобы найти значение заработка, соответствующее кумулятивной частоте 350, разделим расстояние от 600 до 650 долл. в том же отношении, в каком 350 делит расстояние между кумулятивными частотами 270 и 480, т.е. 80 : 130, которое можно упростить как 8 : 13.

Доли, или части, равны соответственно 8/21 и 13/21.

Таким образом, медиана равна $600 + 8/21 \times 50 = 619$ долл.

или $650 - 13/21 \times 50 = 619$ долл.

Расчет двумя способами является эффективной проверкой.

Пример 2. Определите медиану представленного ниже интервального распределения частот.

Группировка Масса животного (г)	*Частота* Количество животных
$100 \le X < 140$	16
$140 \le X < 160$	46
$160 \le X < 180$	76
$180 \le X < 200$	54
$200 \le X < 220$	40
$220 \le X < 240$	10
$240 \le X \le 300$	18
Суммарная частота	260

Решение.

Сначала мы составляем таблицу кумулятивных частот.

Масса животного (г) «до», или «меньше чем»	Кумулятивная частота
100	0
140	16
160	62
180	138
200	192
220	232
240	242
300	260

Суммарная частота равна 260, половина будет составлять 130.

Анализируя таблицу кумулятивных частот, мы видим, что кумулятивная частота 130 находится между 160 г (кумулятивная частота 62) и 180 г (кумулятивная частота 138).

160 г	62	
?	130	>68
180 г	138	>8

Для того чтобы найти значение массы, соответствующее кумулятивной частоте 130, разделим расстояние от 160 до 180 г в том же отношении, в котором 130 делит расстояние между кумулятивными частотами 62 и 138, т.е. 68 : 8, которое можно упростить 17 : 2.

Доли, или части, равны соответственно 17/19 и 2/19.

Таким образом, медиана равна $160 + 17/19 \times 20 = 177{,}9$ г

или $180 - 2/19 \times 20 = 177{,}9$ г.

Расчет двумя способами является эффективной проверкой.

Заметьте, что

(i) значение медианы, полученное с помощью полигона кумулятивных частот, должно **в точности** совпадать с расчетным значением медианы;

(ii) значение медианы, полученное с помощью кривой кумулятивных частот (огивы), должно быть **приблизительно** равным расчетному значению медианы.

2.4.3. Определение медианы интервального распределения частот по гистограмме

Медиана делит площадь гистограммы распределения пополам. Следовательно, вертикальная линия, делящая площадь гистограммы на две равные части, указывает нам медиану.

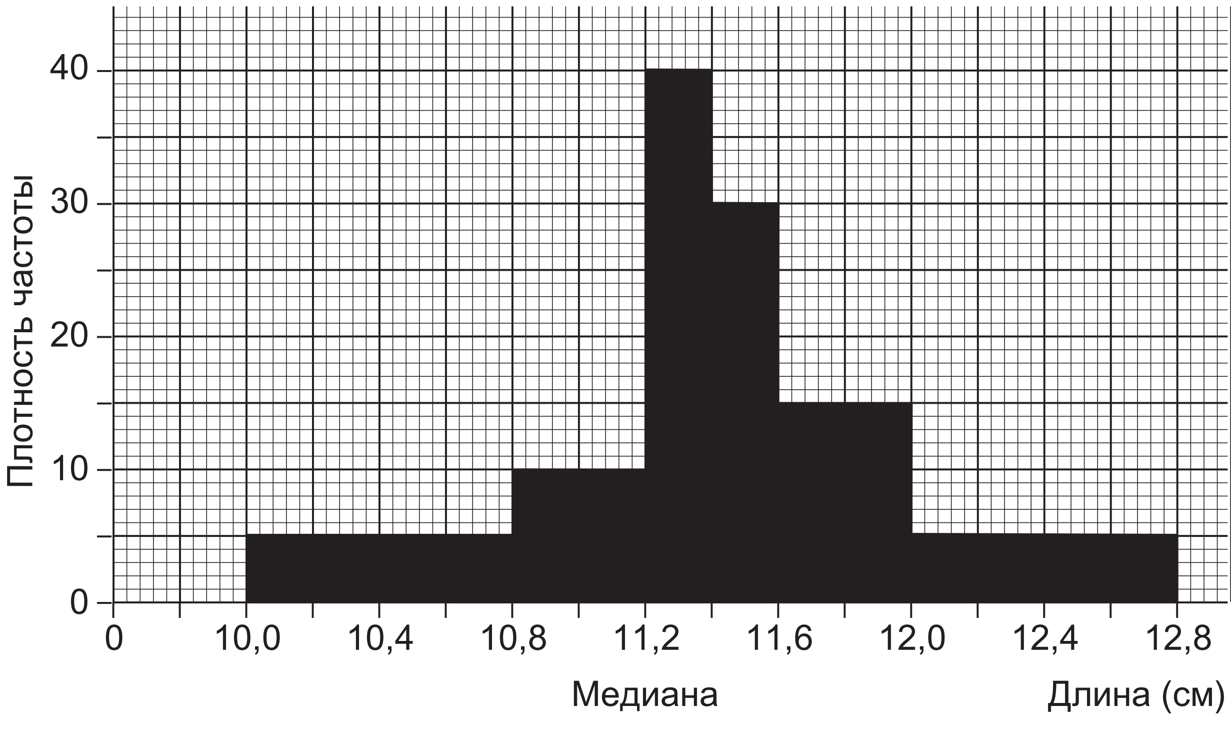
Пример 1. Постройте гистограмму для иллюстрации следующего интервального распределения частот. Затем проведите прямую линию, указывающую медиану.

Длина (см)	Частота
10,0-	4
10,8-	4
11,2-	8
11,4-	6
11,6-	6
12,0-12,8	4

Решение. Вначале рассчитаем высоту столбцов гистограммы.

Группировка Длина (см)	Ширина столбца (см)	*Частота*	*Плотность частоты* Высота столбца (Частота/Ширина)
10,0 ≤ X < 10,8	0,8	4	4/0,8 = 5
10,8 ≤ X < 11,2	0,4	4	4/0,4 = 10
11,2 ≤ X < 11,4	0,2	8	8/0,2 = 40
11,4 ≤ X < 11,6	0,2	6	6/0,2 = 30
11,6 ≤ X < 12,0	0,4	6	6/0,4 = 15
12,0 ≤ X < 12,8	0,8	4	4/0,8 = 5

Теперь можем построить гистограмму.



Мы получим суммарную частоту равной 32 (ее представляют 32 квадрата со стороной ½ см).

Отсчитывая 16 таких квадратов с левого края гистограммы, найдем, что медиана равна 11,4 см.

(Для проверки отсчитаем и 16 квадратов справа от медианы.)

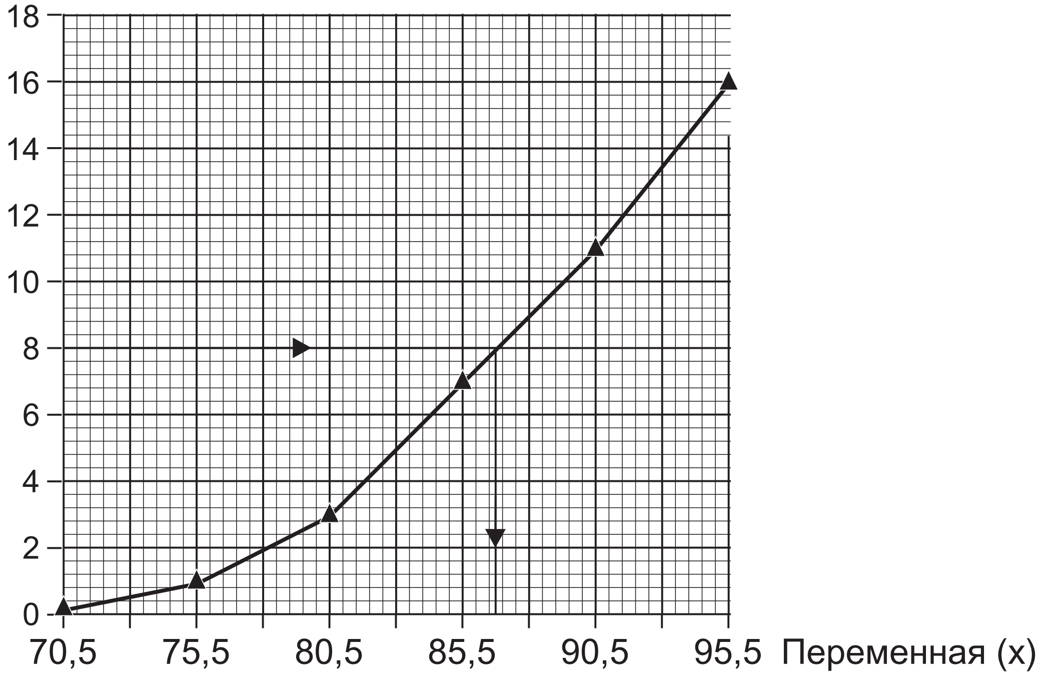
2.4.4. Определение медианы по сгруппированным частотам дискретного распределения

Пример. Построив полигон кумулятивных частот, определите медиану по сгруппированным частотам следующего распределения.

Переменная	Частота
71-75	1
76-80	2
81-85	4
86-90	4
91-95	5
Суммарная частота	16

Решение. Составляем таблицу кумулятивных частот и строим полигон.

«Меньше чем»	Кумулятивная частота
70,5	0
75,5	1
80,5	3
85,5	7
90,5	11
95,5	16



Для определения медианы проведем горизонтальную линию на высоте, равной половине кумулятивной частоты (в нашем примере 8). Затем из точки пересечения этой прямой с полигоном кумулятивных частот опустим перпендикуляр на ось X. Точка пересечения с осью X и даст нам значение медианы распределения (в данном случае 86,75).

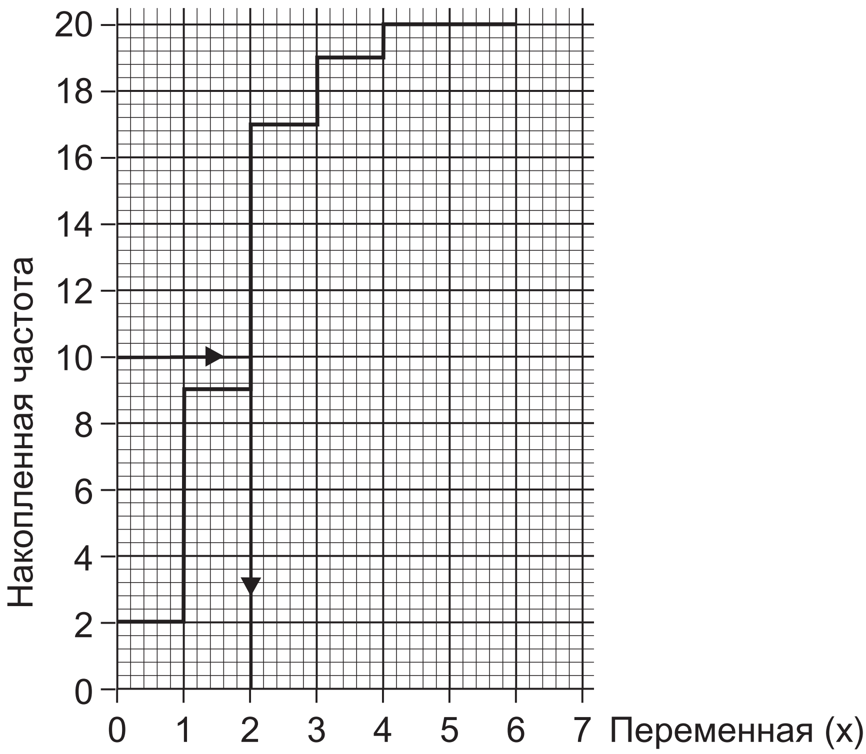
2.4.5. Определение медианы дискретного частотного ряда с помощью графика

Медиану дискретного частотного ряда можно рассчитать непосредственно по ряду распределения. Но можно найти ее и по графику распределения.

Пример. Имеется дискретный ряд распределения.

Переменная, *x*	0	1	2	3	4
Частота, *f*	2	7	8	2	1

Ниже изображен график кумулятивных частот.



Линии построения показывают, как определить медиану по графику. В нашем примере медиана равна 2.

Записав ряд распределения в развернутом виде, получим:

0, 0, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 4.

Медиана равна простой средней 10-го и 11-го значений (каждое из которых равно 2), так что медиана равна 2, что совпадает с результатом, полученным графическим способом.

Упражнение 12A

1. Был проведен анализ возрастного состава населения города. Полученные результаты приведены в таблице:

Возраст (лет)	0-	15-	30-	45-	65-	90-
Число жителей	480	450	420	600	150	0

a) Составьте таблицу кумулятивных частот.

b) Проиллюстрируйте распределение, построив полигон кумулятивных частот.

c) По графику определите медианный возраст жителей города.

d) Рассчитайте медиану по таблице кумулятивных частот.

2. Протяженность 100 отрезков автомагистрали была измерена с точностью до ближайшего километра. В результате получилась следующая таблица.

Длина (км)	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-
Количество отрезков	6	10	13	16	26	18	11	0

a) Запишите точную верхнюю границу интервала 0-4.

b) Составьте таблицу кумулятивных частот.

c) Проиллюстрируйте распределение, построив полигон кумулятивных частот.

d) По графику определите медианную длину.

e) Рассчитайте медиану по таблице кумулятивных частот.

Упражнение 12B

1. В качестве эксперимента каждого ученика в классе попросили нарисовать прямую линию длиной 25 см без использования линейки. Результаты последующего измерения 32 прямых приведены в таблице.

Длина (см)	20-	22-	24-	25-	26-	28-30
Число учеников	3	5	7	9	6	2

a) Составьте таблицу кумулятивных частот.

b) Проиллюстрируйте распределение, построив полигон кумулятивных частот.

c) По графику определите медианную длину.

d) Рассчитайте медиану по таблице кумулятивных частот.

2. В таблице записано количество книг, взятых в общественной библиотеке 50 жителями в течение одного месяца.

Количество книг	0-4	5-9	10-14	15-19	20-24	25-29	30-
Число жителей	3	8	13	15	7	4	0

a) Запишите верхнюю границу интервала 0-4.

b) Составьте таблицу кумулятивных частот.

c) Проиллюстрируйте распределение, построив полигон кумулятивных частот.

d) По графику определите медиану числа книг, взятых в библиотеке в течение одного месяца.

e) Рассчитайте медиану по таблице кумулятивных частот.

2.5. Средняя арифметическая

Средняя арифметическая ряда распределения вычисляется как сумма всех значений, деленная на количество значений.

2.5.1. Средняя арифметическая для небольшого числа малых значений

Средняя арифметическая рассчитывается по формуле $\bar{x} = \frac{\sum x}{n}$,

где Σx – сумма всех значений x;
n – количество значений x.

Пример 1. Найдите среднюю арифметическую следующих пяти чисел:

2, 0, 3, 8, 6.

Решение. Сумма всех значений равна: $\Sigma x = 2 + 0 + 3 + 8 + 6 = 19$.
Количество значений $n = 5$.
Теперь рассчитаем среднюю арифметическую:

$$\bar{x} = \frac{\sum x}{n} = \frac{19}{5} = 3{,}8.$$

Пример 2. Найдите среднюю арифметическую девяти значений:

8, 0, 3, 7, 2, 6, 5, 9, 4.

Решение. Сумма всех значений равна: $\Sigma x = 44$.
Количество значений $n = 9$.

Средняя арифметическая: $\bar{x} = \frac{\sum x}{n} = \frac{44}{9} = 4{,}9$ (с точностью до одного десятичного знака).

Аудиторный тест № 7

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вычислите среднюю арифметическую данных чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	8, 6, 9, 2, 5	5	6	7	8	9
2.	2, 9, 8, 7, 4, 3, 2	5	6	7	8	9
3.	3, 7, 5, 8, 4, 15	5	6	7	8	9
4.	8, 9, 7, 11, 12, 6, 10, 9	5	6	7	8	9
5.	0, 2, 0, 3, 8, 9, 6	1	2	3	4	5
6.	0, 3, 0, 4, 3, 8, 0, 2, 7	1	2	3	4	5
7.	5, 8, 0, 9, 10, 11, 13	7	8	9	10	11
8.	13, 7, 8, 7, 15	7	8	9	10	11
9.	7, 5, 4, 9, 8, 6	6	6½	7	7½	8
10.	5, 9, 8, 6, 0, 10, 8, 3, 14	6	6½	7	7½	8
11.	0, 4, 8, 2, 3, 5, 4, 18	3½	4	4½	5	5½
12.	4, 0, 5, 10, 2, 6, 3, 2, 4, 9	4½	5	5½	6	6½
13.	7, 8, 2, 0, 10, 10, 9, 6, 20	5	6	7	8	9
14.	5, 7, 7, 8, 2, 3, 10	5	6	7	8	9
15.	6, 8, 5, 9, 3, 7, 6, 8	5½	6½	7½	8½	9½
16.	8, 10, 7, 11, 6, 12, 9	5	6	7	8	9
17.	7, 3, 0, 10, 11, 20	5½	6½	7½	8½	9½
18.	10, 0, 12, 8, 10, 9, 13, 10	7½	8	8½	9	9½
19.	4, 3, 0, 2, 5, 7, 4, 3	1½	2	2½	3	3½
20.	7, 6, 0, 9, 6, 12, 12, 9, 8, 6	5½	6½	7½	8½	9½

2.5.2. Средняя арифметическая для небольшого числа больших значений

Если имеется калькулятор, то среднюю арифметическую можно вычислить описанным выше способом (см. параграф 2.5.1).

В отсутствие калькулятора для определения средней арифметической можно применить следующий способ.

Пример 1. Найди средний возраст в месяцах (среднюю арифметическую) восьми учеников, возраст которых приведен ниже (в месяцах):

192, 195, 197, 193, 195, 196, 197, 199.

Решение. Сначала сделаем предположение относительно среднего возраста. Пусть он будет равен 196.

Мы говорим, что **условный средний возраст** восьми учеников равен 196.

Теперь найдем отклонения фактических значений от нашего условного среднего, т.е. от 196.

Значения больше 196 имеют положительное отклонение (+).

Значения меньше 196 имеют отрицательное отклонение (–).

Соответственно отклонения будут равны: –4, –1, +1, –3, –1, 0, +1, +3.

Произведем упрощение, сокращая одинаковые по модулю отклонения с противоположными знаками:

–1 и +1 (дважды);

–3 и +3.

У нас остаются: –4 и 0, что в сумме составляет (–4).

Затем рассчитаем среднее отклонение от нашего условного среднего:

$$-4 : 8 = -0{,}5.$$

Прибавим полученную величину к предполагаемому среднему значению:

$$196 + (-0{,}5) = 195{,}5.$$

Таким образом, мы получили, что в действительности средняя арифметическая (средний возраст учеников) равна 195,5.

Рассмотрим тот же пример еще раз, выбрав другое условное среднее значение. Теперь в качестве условного среднего возьмем 194.

Рассчитаем отклонения фактических значений от 194:

$$-2, +1, +3, -1, +1, +2, +3, +5.$$

Произведем сокращения:

$$-2 \text{ и } +2;$$
$$+1 \text{ и } -1.$$

У нас остаются +3, +1, +3 и +5, что в сумме составляет +12.

Среднее отклонение от нашего условного среднего: 12:8 = 1,5.

Прибавим полученную величину к условному среднему значению:

$$194 + (+1{,}5) = 195{,}5.$$

Таким образом, мы получили, что в действительности средняя арифметическая равна 195,5.

Как видно, оба рассмотренных варианта решения дают один и тот же ответ.

2.6. Средняя арифметическая распределения частот

2.6.1. Средняя арифметическая для большого числа малых значений

Пример 1. Найдите среднюю арифметическую чисел:

0, 1, 2, 2, 3, 1, 2, 1, 2, 1, 2, 1, 2, 1, 3, 0, 2, 2, 4, 1.

Решение. Сначала составим таблицу распределения частот.

Значение x	Подсчет	Частота f
0	//	2
1	~~////~~ //	7
2	~~////~~ ///	8
3	//	2
4	/	1
Суммарная частота		20

Вместо суммирования 20 чисел 0 + 1 + 2 + 2 и т.д. воспользуемся распределением частот.

Сумма 20 чисел выражается формулой $\sum fx$, т.е. суммой произведений каждого значения на его частоту.

Значение x	Частота f	Частота × Значение fx
0	2	0
1	7	7
2	8	16
3	2	6
4	1	4
Сумма	$\sum f = 20$	$\sum fx = 33$

Средняя арифметическая рассчитывается как сумма всех значений, деленная на количество значений.

Таким образом, средняя арифметическая распределения частот определяется формулой

$$\bar{x} = \frac{\sum fx}{\sum f}.$$

В нашем примере средняя арифметическая равна $\frac{33}{20} = 1,65.$

Если нет возможности использовать калькулятор, многие находят, что проще пользоваться этим методом.

Некоторые считают, что с помощью калькулятора проще сложить 20 исходных чисел. Однако важно знать этот метод, если данные представлены в виде распределения частот. Вычислить $\sum fx$ на калькуляторе очень просто.

Пример 2. Найдите среднюю арифметическую следующих чисел:

0, 1, 3, 4, 2, 6, 7, 2, 0, 5, 5, 1, 3, 3, 3, 2, 0, 1, 4, 1, 6, 7, 6, 1, 3.

Решение.

Значение x	Подсчет	Частота f	Частота × × Значение fx
0	///	3	0
1	~~////~~	5	5
2	///	3	6
3	~~////~~	5	15
4	//	2	8
5	//	2	10
6	///	3	18
7	//	2	14
Суммарная частота		$\Sigma f = 20$	$\Sigma f x = 76$

Средняя арифметическая распределения частот $= \dfrac{\sum fx}{\sum f}$.

В нашем примере средняя арифметическая $= \dfrac{76}{25} = 3{,}04$.

Аудиторный тест № 8

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По команде преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите среднюю арифметическую приведенных распределений частот.

Вопрос			**Варианты ответов**				
			A	B	C	D	E
1.	x	0 1 2 3 4 5	1	2	3	4	5
	f	2 7 9 8 9 10					
2.	x	0 1 2 3 4 5	1	2	3	4	5
	f	7 5 4 3 2 4					
3.	x	2 3 4 5	1	2	3	4	5
	f	3 7 6 13					
4.	x	2 3 4 5 6 7	1	2	3	4	5
	f	2 1 4 5 8 2					
5.	x	0 1 2 3 4	1	2	3	4	5
	f	10 4 3 2 1					
6.	x	0 2 4 6 8 10	5	6	7	8	9
	f	2 2 5 3 1 7					
7.	x	0 2 6 9 12	5	6	7	8	9
	f	1 4 6 4 10					
8.	x	2 4 6 8 10 12	5	6	7	8	9
	f	1 3 11 5 5 1					

2.6.2. Средняя арифметическая для большого числа больших значений

Пример 1. Найдите среднюю арифметическую чисел

218, 217, 216, 217, 215, 218, 217, 216, 215, 214, 219, 216, 218, 217.

Решение. Вначале составляем таблицу распределения частот.

Значение x	Подсчет	Частота f
214	/	1
215	//	2
216	///	3
217	////	4
218	///	3
219	/	1
Суммарная частота		$\Sigma f = 14$

Теперь сделаем предположение, чему может быть равна средняя арифметическая. Положим ее равной 217.

Значение x	Отклонение от предполагаемой средней D	Частота, f	Частота × × Отклонение fD
214	–3	1	–3
215	–2	2	–4
216	–1	3	–3
217	0	4	0
218	+1	3	+3
219	+2	1	+2
Сумма		$\Sigma f = 14$	$\Sigma f x = -5$

Среднее отклонение от условного среднего значения определяется формулой $\dfrac{\sum fD}{\sum f}$.

В нашем примере среднее отклонение от условного среднего значения равно:

$$\frac{-5}{14} = -0,357.$$

Реальная средняя арифметическая = Условное среднее значение + + Среднее отклонение от среднего значения.

Таким образом, реальная средняя арифметическая равна:

$$217 + (-0,357) = 216,643.$$

Еще раз напомним, что при использовании калькулятора может быть проще найти сумму четырнадцати исходных чисел.

2.6.3. Средняя арифметическая для сгруппированных частот дискретного распределения

Иногда распределение имеет настолько большую вариацию, что работать с исходными данными нецелесообразно. Вместо этого данные объединяют в группы. Например, экзаменационные оценки, изменяющиеся от 0 до 100, могут быть объединены в группы 0–4, 5–9, 10–14, 15–19 и т.д. Здесь знак «–» означает интервал «до», а не «минус». Средняя арифметическая, вычисленная после такой группировки данных, получается с некоторой погрешностью, но для большинства статистических задач точность оказывается достаточной.

Пример. Определите среднюю арифметическую по сгруппированному частотному распределению.

Значение	Частота
71-75	1
76-80	2
81-85	4
86-90	4
91-95	5
Суммарная частота	16

Решение. Метод 1

Прежде всего определим середину каждого интервала. Затем рассчитаем среднюю арифметическую, используя полученные значения середин интервалов.

Значение	Частота f	Середина интервала x	Частота × × Отклонение, fx
71-75	1	73	73
76-80	2	78	156
81-85	4	83	332
86-90	4	88	352
91-95	5	93	465
Сумма	$\Sigma f = 16$		$\Sigma fx = 1378$

Средняя арифметическая = Σfx / Σf = 1378/16 = 86,125.

Для условной средней используем одно из значений середины интервала. Мы выбираем среднее значение, которое имеет или наибольшую частоту, или достаточно большую. В нашем примере среднее значение 93 имеет наибольшую частоту, но оно не центральное. Поэтому вместо него мы возьмем среднее значение 83, которое является центральным и имеет большую частоту.

Решение. Метод 2

Сперва предположим, каким может быть ответ, и используем это предположение для выбора условного среднего значения. В нашем примере мы возьмем значение 83. Затем рассчитаем все отклонения от условного среднего.

Значение	Частота	Середина интервала	Отклонение от предполагаемого среднего
71-75	1	73	–10
76-80	2	78	–5
81-85	4	83	0
86-90	4	88	+5
91-95	5	93	+10

Мы видим, что все числа в четвертом столбце делятся на 5. Поэтому разделим их на 5 и получаем коэффициент $X = \frac{(x-83)}{5}$.

Далее рассчитаем произведение fx для каждого значения x, а затем найдем сумму Σfx.

Таким образом, мы вычислили среднюю арифметическую X.

Значение	Частота	x	x–83	X = (x–83)/5	fX
71-75	1	73	–10	–2	–2
76-80	2	78	–5	–1	–2
81-85	4	83	0	0	0
86-90	4	88	+5	+1	+4
91-95	5	93	+10	+2	+10
Сумма	$\Sigma f = 16$				$\Sigma fx = 10$

Средняя арифметическая X равна $\frac{\sum fX}{\sum f} = \frac{10}{16} = 0,625.$

Если $X = \frac{(x-83)}{5}$, то 5X = x – 83 или x = 5X + 83.

Таким образом, средняя арифметическая x равна $5 \times 0{,}625 + 83 =$ $= 86{,}125$.

При наличии калькулятора первый метод может быть быстрее и проще.

В противном случае быстрее и проще решать вторым способом.

2.6.4. Средняя арифметическая интервального распределения частот

Пример 1. Используем данные из примера 1 раздела 1.12.

Рассчитайте среднюю арифметическую для интервального распределения частот.

Рост девочек (см)	Число девочек
$152 \leq X < 156$	4
$156 \leq X < 160$	8
$160 \leq X < 162$	8
$162 \leq X < 164$	14
$164 \leq X < 166$	10
$166 \leq X < 170$	12
$170 \leq X \leq 174$	4

Решение. Метод аналогичен методу определения средней арифметической для сгруппированных частот дискретного распределения, описанному в параграфе 2.6.3.

Метод 1. С использованием калькулятора

Рост девочек (см)	Частота f	Среднее значение x	fx
$152 \leq X < 156$	4	154	616
$156 \leq X < 160$	8	158	1264
$160 \leq X < 162$	8	161	1288
$162 \leq X < 164$	14	163	2282
$164 \leq X < 166$	10	165	1650
$166 \leq X < 170$	12	168	2016
$170 \leq X \leq 174$	4	172	688
Сумма	$\Sigma f = 60$		$\Sigma fx = 9804$

Средняя арифметическая x, $\bar{x} = \dfrac{\sum f \cdot x}{\sum f} = \dfrac{9804}{60} = 163{,}4.$

Метод 2. Без использования калькулятора

Рост девочек (см)	Частота f	Среднее значение x	$X = x - 163$	fX
$152 \le X < 156$	4	154	–9	–36
$156 \le X < 160$	8	158	–5	–40
$160 \le X < 162$	8	161	–2	–16
$162 \le X < 164$	14	163	0	0
$164 \le X < 166$	10	165	+2	+20
$166 \le X < 170$	12	168	+5	+60
$170 \le X \le 174$	4	172	+9	+36
Сумма	$\Sigma f = 60$			$\Sigma fX = 24$

В данном примере числа в четвертом столбце не имеют подходящего общего кратного (делителя). Поэтому мы не производим деления.

Средняя арифметическая X, $\overline{X} = \dfrac{\sum fX}{\sum f} = \dfrac{24}{60} = 0,4.$

$X = x - 163$, следовательно, $\overline{x} = \overline{X} + 163 = 0,4 + 163 = 163,4.$

Средняя арифметическая распределения равна 163,4, что подтверждает результат, полученный первым методом.

Пример 2. Рассчитайте среднюю арифметическую по данным примера 2 из параграфа 1.13.1

Рассчитайте среднюю арифметическую для интервального распределения частот.

Масса животного (г)	Частота
$100 \le X < 140$	16
$140 \le X < 160$	46
$160 \le X < 180$	76
$180 \le X < 200$	54
$200 \le X < 220$	40
$220 \le X < 240$	10
$240 \le X \le 300$	18

Решение.

Метод 1. С использованием калькулятора

Масса животного (г)	Частота f	Среднее значение x	fx
$100 \le X < 140$	16	120	1 920
$140 \le X < 160$	46	150	6 900
$160 \le X < 180$	76	170	12 920
$180 \le X < 200$	54	190	10 260
$200 \le X < 220$	40	210	8 400
$220 \le X < 240$	10	230	2 300
$240 \le X \le 300$	18	270	4 860
Сумма	$\Sigma f = 260$		$\Sigma fx = 47\ 560$

Средняя арифметическая $x, \bar{x} = \dfrac{\sum fx}{\sum f} = \dfrac{47\ 560}{260} = 182,92.$

Метод 2. Без использования калькулятора

Масса животного (г)	Частота f	Среднее значение x	$x - 190$	$X = (x - 190)/20$	fX
$100 \le X < 140$	16	120	–70	–3,5	–56
$140 \le X < 160$	46	150	–40	–2	–92
$160 \le X < 180$	76	170	–20	–1	–76
$180 \le X < 200$	54	190	0	0	0
$200 \le X < 220$	40	210	+20	+1	+40
$220 \le X < 240$	10	230	+40	+2	+20
$240 \le X \le 300$	18	270	+80	+4	+72
Сумма	$\Sigma f = 260$				$\Sigma fX = -92$

Средняя арифметическая $X, \overline{X} = \dfrac{\sum fX}{\sum f} = \dfrac{-92}{260} = -0,3539.$

$X = (x - 190)/20$, следовательно
$20X = x - 190$ или $x = 20X + 190$.
Средняя арифметическая равна $\bar{x} = 20(-0,3539) + 190 = 182,92$.
Это подтверждает результат, полученный первым методом.

Аудиторный тест № 9

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите среднее значение интервального ряда, взятого по сгруппированным данным дискретного распределения.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	7, 8, 9, 10	8	8,5	9	9,5	10
2.	12, 13, 14	13	13,25	13,5	13,75	14
3.	0,4 0,5 0,6 0,7 0,8	0,4	0,5	0,6	0,7	0,8
4.	6-10	7	7,5	8	8,5	9
5.	20-24	21	22	23	24	44
6.	40-49	44	44,5	45	45,5	89
7.	51, 52, 53, 54, 55	51,5	52	52,5	53	53,5
8.	3½, 4, 4½, 5	3,75	4	4,25	4,5	4,75
9.	60-69	63,75	64	64,25	64,5	64,75
10.	80-99	89	89,25	89,5	89,75	90

Найдите среднее значение данного интервала, взятого из интервального распределения частот.

	Вопрос	Варианты ответов				
		A	B	C	D	E
11.	$140 < X \leq 160$	145	150	155	160	165
12.	$17 < X \leq 19$	17	17,5	18	18,5	19
13.	$0,3 < X \leq 0,4$	0,31	0,32	0,33	0,34	0,35
14.	$200 < X \leq 250$	220	225	230	235	240
15.	$2,0 < X \leq 2,2$	2,05	2,1	2,15	2,2	4,2
16.	$24 < X \leq 26$	23	24	25	26	50
17.	$300 < X < 400$	345	350	355	360	365
18.	$3,0 < X \leq 3,5$	3,05	3,1	3,15	3,2	3,25
19.	$1200 < X \leq 1400$	1240	1250	1260	1280	1300
20.	$0,8 < X \leq 0,9$	0,84	0,85	0,86	0,87	1,7

Упражнение 13A

1. В таблице указано время, затрачиваемое на дорогу сотрудниками фирмы.

Время (мин)	Количество сотрудников
0-	6
5-	11
10-	17
15-	13
20-	3
25-	0

(i) Запишите средние значения каждого интервала.
(ii) Вычислите среднее значение времени в пути (среднюю арифметическую).

2. В таблице указано время просмотра школьниками телепередач за неделю (с точностью до часа).

Время (ч)	Число школьников
0-3	16
4-6	8
7-9	4
10-12	2

(i) Запишите нижнюю границу, верхнюю границу и среднее значение интервала 0-3.
(ii) Запишите нижнюю границу, верхнюю границу и среднее значение интервала 4-6.
(iii) Вычислите среднюю арифметическую ряда распределения.

Упражнение 13B

1. Для испытания нового корма было отобрано некоторое количество животных. Перед экспериментом каждое животное было точно взвешено, и эти веса приведены в таблице.

Масса (кг)	Количество животных
85-	14
95-	23
105-	37
115-	17
125-135	6

(i) Запишите средние значения каждого интервала.
(ii) Вычислите среднюю арифметическую ряда распределения.

2. В таблице указано время, затраченное школьниками на выполнение домашней работы (с точностью до часа).

Время (ч)	Количество школьников
4	2
5	3
6	5
7	6
8	8
9	5
10	2

Вычислите среднюю арифметическую данного ряда распределения.

2.7. Средняя геометрическая

2.7.1. Средняя геометрическая двух чисел

Рассмотрим два числа x и y.

СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ

Средняя арифметическая = $= \frac{1}{2}(x + y)$

Пример 1. 4 и 64.

Средняя арифметическая = $= \frac{1}{2}(4 + 64) = 34$

Пример 2. 25 и 625.

Средняя арифметическая = $= \frac{1}{2}(25 + 625)=325$

Пример 3. 3 и 5.

Средняя арифметическая = $= \frac{1}{2}(3 + 5) = 4$

СРЕДНЯЯ ГЕОМЕТРИЧЕСКАЯ

Средняя геометрическая равна квадратному корню из произведения x и y

Средняя геометрическая $= \sqrt{x \cdot y}$

Пример 1. 4 и 64.

Средняя геометрическая = $= \sqrt{4 \cdot 64} = \sqrt{256} = 16$

Пример 2. 25 и 625.

Средняя геометрическая = $= \sqrt{25 \cdot 625} = \sqrt{15625} = 125$

Пример 3. 3 и 5.

Средняя геометрическая = $= \sqrt{3 \cdot 5} = \sqrt{15} = 3{,}873$

(расчет на калькуляторе)

Пример 4. Данные о населении Соединенного Королевства (с точностью до тысячи человек):

В 1931 г. численность населения была 46 038 000 чел.

В 1941 г. перепись не проводилась, данных нет.

В 1951 г. численность населения – 50 225 000 чел.

Лучшей оценкой численности населения в 1941 г. является средняя геометрическая данных за 1931 и 1951 гг.

Средняя геометрическая =

$$= \sqrt{46\,038\,000 \times 50\,225\,000} = 48\,086\,000.$$

Аудиторный тест № 10

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите среднюю геометрическую двух чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	1 и 4	1	1,5	2	2,5	3
2.	1 и 9	1	2	3	4	5
3.	1 и 16	1	2	4	8	8,5
4.	1 и 100	2	5	10	20	50,5
5.	1 и 25	2	5	10	13	15
6.	4 и 9	5	6	6,5	13	36
7.	2 и 8	3	4	5	10	16
8.	1 и 49	4	7	25	49	50
9.	4 и 25	7	10	14,5	29	100
10.	2 и 32	8	16	32	34	64
11.	3 и 27	9	15	27	30	81
12.	4 и 16	8	10	20	32	64
13.	3 и 12	4	6	15	18	36
14.	5 и 20	10	12,5	25	50	100
15.	2 и 18	6	7	8	9	10
16.	2 и 50	6	7	8	9	10
17.	1 и 36	2	4	6	8	10
18.	1 и 81	1	3	9	27	81
19.	4 и 36	4	6	8	10	12
20.	6 и 24	12	15	30	72	144

2.7.2. Связь между средней геометрической и средней арифметической

Средняя арифметическая = ½(x + y) log(Средней геометрической) = = ½(log$x$ + log$y$).

Видно, что логарифм средней геометрической равен средней арифметической логарифмов двух чисел.

Арифметический ряд

Если мы рассмотрим ряд чисел 4, 34 и 64, то увидим, что **прибавляя** 30 к 4, мы получим 34, а прибавляя 30 к 34, получим 64.

$$4 \xrightarrow{+30} 34 \xrightarrow{+30} 64$$

Такой ряд называется арифметическим рядом. В арифметическом ряду каждый член ряда является средней арифметической двух соседних. Здесь 34 – средняя арифметическая 4 и 64, как мы видели в примере 1 выше.

Пример 2 дает нам еще один арифметический ряд: 25, 325, 625.

В данном случае переход от одного члена ряда к другому осуществляется прибавлением 300.

Геометрический ряд

Если мы рассмотрим ряд чисел 4, 16 и 64, то увидим, что **умножая** 4 на 4, получим 16, а умножая 16 на 4, получим 64.

$$4 \xrightarrow{\text{Ч}\,4} 16 \xrightarrow{\text{Ч}\,4} 64$$

Такой ряд называется геометрическим рядом. В геометрическом ряду каждый член ряда является средней геометрической двух соседних. Здесь 16 – средняя геометрическая 4 и 64, как мы видели в примере 1 выше.

Пример 2 дает нам еще один геометрический ряд: 25, 125, 625.

В данном случае переход от одного члена ряда к другому осуществляется умножением на 5.

2.7.3. Средняя геометрическая более чем двух чисел

Средняя геометрическая n чисел рассчитывается как корень степени n из произведения этих чисел.

Пример 1. Вычислите среднюю геометрическую чисел 2, 8, 32.

Решение. Найдем произведение этих чисел: $2 \times 8 \times 32 = 512$.

Всего у нас имеется 3 числа, следовательно, средняя геометрическая равна корню кубическому из 512, что составляет 8.

Пример 2. Вычислите среднюю геометрическую чисел 2, 5, 7, 11 и 13.

Решение. Найдем произведение этих чисел: $2 \times 5 \times 7 \times 11 \times 13 =$ $= 10\,010$.

Всего у нас имеется 5 чисел, следовательно, средняя геометрическая равна корню 5-й степени из 10 010, что составляет 6,311, или 6,31 с точностью до двух десятичных знаков.

2.7.4. Применение средней геометрической

(i) Когда переменная возрастает в геометрической прогрессии.

Пример. Население.

Население города в 1930 г. составляло 500 000 человек. Население того же города в 1950 г. составляло 700 000 человек. Оцените население города в 1940 г., когда проведение переписи было невозможно.

Лучшей оценкой является средняя геометрическая 500 000 и 700 000.

Вычислим произведение: $5 \times 10^5 \times 7 \times 10^5 = 35 \times 10^{10}$.

Тогда средняя геометрическая равна $5{,}91608 \times 10^5 = 591\,608$. Результат можно округлить: 591 600.

(ii) Уменьшить влияние аномальных наблюдений (значений).

Пример. Рассмотрим данные об оценках, полученных 10 девочками:

5, 4, 4, 4, 6, 5, 4, 3, 12, 4.

Средняя арифметическая равна $51/10 = 5{,}1$, что недостоверно представляет распределение, поскольку 8 из 10 девочек сдали ниже, чем на 5,1.

Произведение этих 10 чисел равно 5 529 600. Средняя геометрическая рассчитывается как корень 10-й степени из 5 529 600, т.е. 4,724. Это значение намного достовернее.

Аудиторный тест № 11

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите среднюю геометрическую следующих чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	2, 4, 1	1	2	2,33	3	4
2.	2, 4, 8	1	2	3	4	4,67
3.	9, 1, 3	1	2	3	4	4,33
4.	25, 1, 5	5	6	7	10	10,33
5.	2, 8, 32	6	8	10	12	14
6.	2, 2, 4, 16	2	3	4	5	6
7.	36, 1, 6	4	6	8	14	14,33
8.	1, 1, 8	1,5	2	2,5	3	3,5
9.	1, 8, 8	4	4,5	5	5,67	6
10.	1, 49, 7	7	14	18	19	21
11.	16, 2, 24	4	5	6	6,67	7
12.	1, 81, 9	5	6	7	8	9
13.	3, 9, 9, 27	5	6	7	8	9
14.	2, 5, 100	2	5	10	20	36
15.	1, 2, 4, 2	2	3	4	5	6
16.	1, 27, 1	3	4	5	6	9
17.	16, 1, 4	3	4	5	6	7
18.	1, 125, 1, 5	1	3	5	7	9
19.	6, 12, 24	6	8	10	12	14
20.	2, 2, 36	4	6	8	10	12

Упражнение 14А

Найдите среднюю геометрическую для каждого ряда значений. (***Разрешается пользоваться калькулятором.***)

Приближенные ответы округляйте до двух десятичных знаков.

1. 49 и 81
2. 16 и 121
3. 2,5 и 4,9
4. 47 и 53
5. 89 и 81
6. 48 и 52
7. 38, 41 и 44
8. 53, 57 и 61
9. 80, 85 и 90
10. 16, 64 и 256
11. 24, 36 и 54
12. 35, 35 и 40
13. 35, 40 и 40
14. 20, 30 и 45
15. 20, 20 и 30

Упражнение 14В

Найдите среднюю геометрическую для каждого ряда значений. (***Разрешается пользоваться калькулятором.***)

Приближенные ответы округляйте до двух десятичных знаков.

1. 36 и 49
2. 25 и 81
3. 8,1 и 0,9
4. 31 и 37
5. 67 и 71
6. 65 и 75
7. 41, 44 и 47
8. 53, 58 и 63
9. 60, 70 и 80
10. 27, 81 и 243
11. 15, 75 и 375
12. 40, 40 и 50
13. 40, 50 и 50
14. 20, 40 и 80
15. 35, 40 и 40

2.8. Обзор видов средних величин

(a) Мода

(i) Легко определить на взгляд, по гистограмме или расчетами.
(ii) Зависимость от экстремальных значений отсутствует.
(iii) Не используется при более углубленном статистическом анализе.

(b) Медиана

(i) Легко определить на взгляд, по диаграмме кумулятивных частот или расчетами.
(ii) Зависимость от экстремальных значений отсутствует.
(iii) Не используется при более углубленном статистическом анализе.

(c) Средняя арифметическая

(i) Вычисляется сложнее.
(ii) Зависимость от экстремальных значений сильная.
(iii) Часто используется при углубленном статистическом анализе.

(d) Средняя геометрическая

(i) Вычисляется значительно сложнее.
(ii) Зависимость от экстремальных значений слабая.
(iii) Не очень часто используется при углубленном статистическом анализе.

2.9. Расположение моды, медианы и средней арифметической на графике

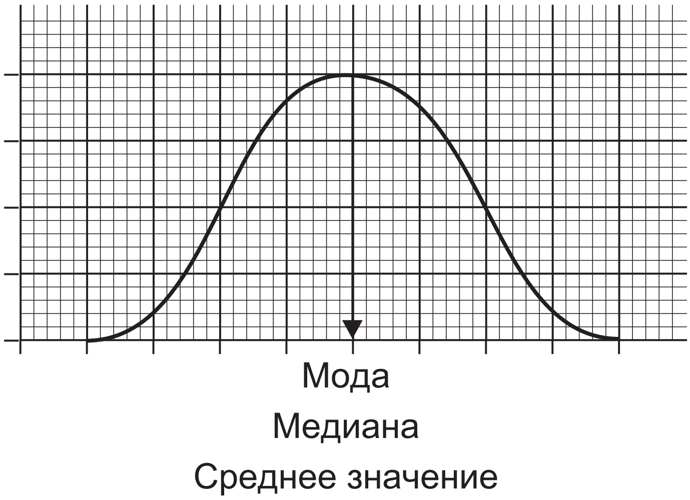
Мода ряда распределения – это значение, соответствующее самой высокой точке кривой.

Медиана ряда распределения делит пополам площадь под графиком распределения.

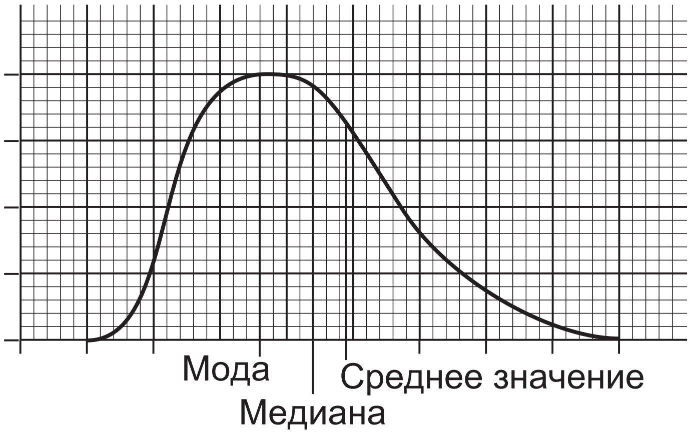
Средняя арифметическая ряда распределения лежит на линии равновесия площадей, если график нарисовать на картоне, а затем вырезать и подвесить.

Относительное расположение моды, медианы и средней арифметической зависит от формы распределения: нормальное (колоколообразное), с правосторонней или левосторонней асимметрией.

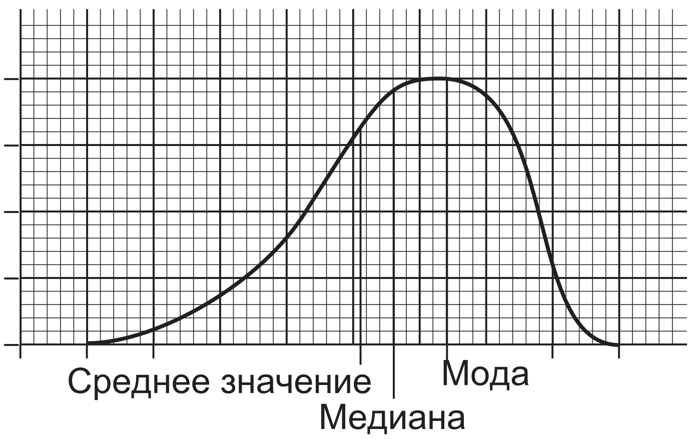
Для нормального распределения **мода**, **медиана** и **среднее значение** совпадают.



Для распределения с правосторонней асимметрией средние величины располагаются в порядке возрастания следующим образом: **мода**, **медиана**, **среднее значение.**



Для распределения с левосторонней асимметрией средние величины располагаются в порядке возрастания следующим образом: **среднее значение**, **медиана**, **мода.**



Глава 3

МЕРЫ КОЛЕБЛЕМОСТИ ИЛИ ВАРИАЦИИ

Мерами колеблемости, или вариациями, являются:

(i) размах;
(ii) межквартильный размах;
(iii) среднее отклонение;
(iv) стандартное отклонение и дисперсия.

3.1. Размах

Размах – это расстояние от минимального значения до максимального или разница между максимальным и минимальным значениями.

Пример 1. Найдите размах следующего ряда значений:

5, 2, 7, 9, 15, 8, 6.

Решение. Минимальное значение равно 2.
Максимальное значение равно 15.
15 – 2 = 13. Размах равен 13.

Обратите внимание, что в статистике размах – это просто число или величина. В рассмотренном выше примере размах равен просто 13, а не интервалу $2 \le x \le 15$.

Пример 2. Найдите размах следующего ряда значений:

8, 7, 3, 5, 2, 1, 29, 17, 16, 9.

Решение. Минимальное значение равно 1.
Максимальное значение равно 29.
29 – 1 = 28. Размах равен 28.

Аудиторный тест № 12

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Найдите размах данных рядов чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	2, 8, 6, 9, 12, 23, 14, 17	7	8	15	21	23
2.	4, 8, 3, 7, 13, 2, 5, 16	2	12	13	14	16
3.	8, 5, 3, 4, 9, 12, 7, 13, 2	6	8	10	11	12
4.	7, 8, 3, 10, 9, 5, 6	3	4	5	6	7
5.	14, 3, 2, 17, 5, 9,8	3	6	8	14	15
6.	4, 7, 1, 5, 12, 4, 9	5	7	9	11	13
7.	3, 0, 7, 0, 8, 8, 2	0	2	3	7	8
8.	6, 1, 15, 17, 20, 9, 7	1	15	19	20	21
9.	4, 16, 8, 7, 9, 5, 23	4	5	8	19	22
10.	19, 4, 6, 12, 8, 15, 2	2	4	8	17	19
11.	17, 13, 6, 2, 14, 7, 9	14	15	16	17	18
12.	14, 24, 36, 18, 12, 10	4	8	12	16	26
13.	13, 20, 48, 62, 17, 12	1	36	49	50	60
14.	23, 18, 19, 34, 51, 62	33	43	44	45	46
15.	71, 80, 92, 37, 64, 47	24	25	35	45	55
16.	48, 19, 23, 57, 18, 36	12	17	23	39	48
17.	43, 12, 20, 26, 37, 50	31	38	48	50	52
18.	21, 32, 43, 54, 65, 76	55	57	59	62	65
19.	80, 22, 71, 47, 63, 19	41	51	52	61	71
20.	21, 81, 18, 72, 64, 45	24	48	54	60	63

3.2. Межквартильный размах

3.2.1. Понятие межквартильного размаха

Нижний квартиль – это такое значение, что одна четверть наблюдений или распределения находится ниже него, а остальные три четверти наблюдений или распределения находятся выше него.

Нижний квартиль – это $\frac{1}{4}(n+1)$-е наблюдение.

Верхний квартиль – это такое значение, что три четверти наблюдений или распределения находятся ниже него, а одна четверть наблюдений или распределения находится выше него.

Верхний квартиль – это $\frac{3}{4}(n+1)$-е наблюдение.

Межквартильный размах – это расстояние от нижнего квартиля до верхнего квартиля.

Пример 1. Определите межквартильный размах по 7 значениям:

15, 12, 17, 23, 35, 5, 14.

Решение. Прежде всего надо расположить числа в порядке возрастания.

5, 12, 14, 15, 17, 23, 35.

Нижний квартиль – это $\frac{1}{4}(7+1)$-е значение, т.е. второе, которое равно 12.

Верхний квартиль – это $\frac{3}{4}(7+1)$-е значение, т.е. шестое, которое равно 23.

Межквартильный размах равен 23 – 12, т.е. 11.

Помните, что как размах, так и межквартильный размах – это просто числа.

Пример 2. Определите межквартильный размах по 9 значениям:

1, 5, 18, 4, 7, 13, 20, 25, 22.

Решение. Прежде всего надо расположить числа в порядке возрастания.

1, 4, 5, 7, 13, 18, 20, 22, 25.

Нижний квартиль – это ¼(9 + 1)-е значение, т.е. (2½)-е значение. Мы используем простую среднюю из 2-го и 3-го значений, которые равны 4 и 5. Таким образом, нижний квартиль равен 4,5.

Верхний квартиль – это ¾(9 + 1)-е значение, т.е. (7½)-е значение. Мы используем простую среднюю из 7-го и 8-го значений, которые равны 20 и 22. Таким образом, верхний квартиль равен 21.

Межквартильный размах равен 21 – 4,5, т.е. 16,5.

Пример 3. Определите межквартильный размах по 8 значениям:

12, 5, 4, 23, 17, 32, 27, 15.

Решение. Прежде всего надо расположить числа в порядке возрастания.

4, 5, 12, 15, 17, 23, 27, 32.

Нижний квартиль – это ¼(8+1)-е значение, т.е. (2¼)-е значение. Надо определить значение, которое находится на ¼ расстояния от 2-го значения до 3-го значения.

2-е значение равно 5.

3-е значение равно 12.

Расстояние от 5 до 12 равно 7. Таким образом, нижний квартиль равен 5 + ¼ × 7, т.е. 6,75.

Верхний квартиль – это ¾(8+1)-е значение, т.е. (6¾)-е значение. Надо определить значение, которое находится на ¾ расстояния от 6-го значения до 7-го значения.

6-е значение равно 23.

7-е значение равно 27.

Расстояние от 23 до 27 равно 4. Таким образом, верхний квартиль равен 23 + ¾ × 4, т.е. 26.

Межквартильный размах равен 26 – 6,75, т.е. 19,25.

Аудиторный тест № 13

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Определите межквартильный размах в каждой из приведенных серий чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	2, 4, 4, 5, 6, 7, 8	2	3	4	5	6
2.	3, 4, 4, 6, 7, 8, 9, 9, 10, 12, 14	4	5	6	7	8
3.	4, 6, 7, 7, 8, 9, 9, 10, 11, 12, 13, 14, 15, 16, 18	7	8	9	10	11
4.	3, 3, 6, 7, 7	3½	4	4½	5	5½
5.	4, 5, 5, 6, 8, 9, 10, 10, 12	3½	4	4½	5	5½
6.	3, 4, 5, 8, 8	3½	4	4½	5	5½
7.	2, 3, 3, 4, 5, 5, 6, 7, 8	3½	4	4½	5	5½
8.	3, 5, 5, 6, 7, 9, 9, 10	3½	3¾	4	4¼	4½
9.	2, 3, 5, 6, 6, 7, 8, 9, 9, 11	3½	3¾	4	4¼	4½
10.	4, 5, 5, 6, 6, 7, 9	1	2	3	4	5
11.	5, 5, 6, 8, 10	1	2	3	4	5
12.	2, 3, 4, 6, 7, 8, 9, 9, 12	4½	5	5½	6	6½
13.	3, 4, 4, 5, 6, 8, 9, 11	4½	4¾	5	5¼	5½
14.	2, 3, 3, 5, 5, 6, 7, 8, 8, 9	4½	4¾	5	5¼	5½
15.	2, 3, 5, 6, 7, 9, 10, 12, 14, 16, 17, 18, 19, 20, 20	8	9	10	11	12
16.	3, 5, 5, 6, 8, 8, 10, 11, 13, 14	4½	5	5½	6	6½
17.	4, 6, 7, 7, 8, 8, 9, 10, 11, 12, 12, 14, 15	4½	5	5½	6	6½
18.	2, 3, 5, 6, 7, 8, 8, 10	4½	5	5½	6	6½
19.	2, 4, 5, 5, 6, 7, 7, 8, 8, 9, 11, 12, 13	4½	5	5½	6	6½
20.	3, 4, 5, 7, 8, 9, 9, 10, 10, 11, 11, 12, 12	4½	5	5½	6	6½

Упражнение 15 А

Определите межквартильный размах в каждой из приведенных серий чисел.

1. 9, 1, 3, 8, 3, 9, 1
2. 3, 4, 5, 6, 7, 9, 12, 13, 11, 8, 6, 6, 5, 3, 2
3. 5, 8, 8, 6, 5
4. 5, 5, 3, 2, 1
5. 3, 12, 4, 9, 8, 5, 8, 6
6. 11, 3, 5, 10, 10, 9, 6, 7, 8, 8
7. 2, 9, 8, 6, 4, 5, 4, 5
8. 3, 4, 5, 6, 14, 15, 16, 17, 7, 13, 8, 12, 11, 10, 9
9. 0, 19, 1, 17, 2, 4, 5, 15, 13, 11, 9, 8, 7
10. 2, 19, 18, 16, 3, 5, 5, 6, 14, 7, 12, 9, 11

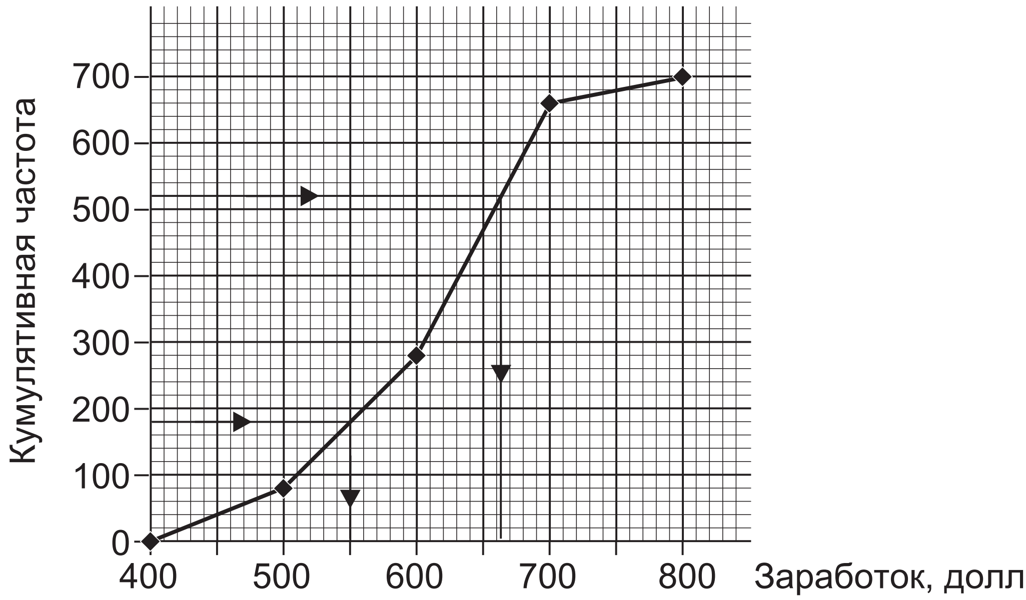
Упражнение 15 В

Определите межквартильный размах в каждой из приведенных серий чисел.

1. 0, 1, 3, 4, 6, 7, 8, 10, 12, 14, 16
2. 12, 3, 11, 4, 6, 8, 9
3. 9, 1, 8, 3, 3, 8, 5, 6, 7
4. 2, 9, 7, 3, 2
5. 11, 5, 6, 9, 9, 6, 8, 7
6. 2, 13, 11, 3, 11, 5, 9, 7, 6
7. 5, 13, 6, 12, 6, 12, 11, 7, 9, 8
8. 3, 13, 4, 12, 4, 10, 9, 8, 7, 5
9. 5, 14, 6, 13, 7, 13, 11, 9
10. 4, 21, 5, 20, 19, 6, 19, 8, 9, 11, 13, 15, 17

3.2.2. Межквартильный размах в интервальном ряду распределения

В параграфе 1.14 мы показали, как представить интервальный ряд распределения при помощи полигона кумулятивных частот. Мы можем использовать полигон кумулятивных частот для определения нижнего квартиля, верхнего квартиля и межквартильного размаха непрерывного (интервального) распределения.



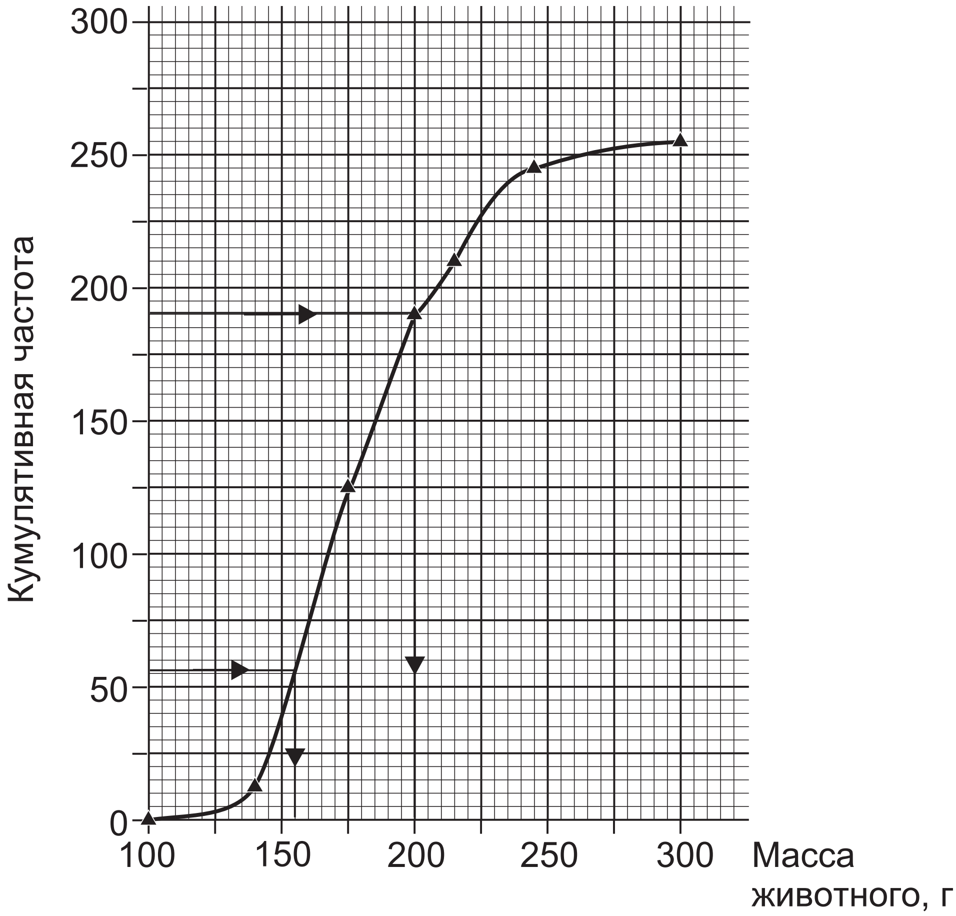
Ниже воспроизведен график, полученный в примере 1 из раздела 1.14.

Для определения нижнего квартиля найдем значение заработка (в долл.), которое соответствует кумулятивной частоте, в точности равной одной четверти суммарных частот. В этом примере одна четверть суммарной частоты равна $700 : 4 = 175$, а соответствующий заработок равен 560 долл.

Для определения верхнего квартиля найдем значение заработка, которое точно соответствует кумулятивной частоте, равной трем четвертям суммарной частоты. В этом примере три четверти суммарной частоты будут равны $700 \times 3 : 4 = 525$, а соответствующий заработок равен 663 долл.

Межквартильный размах равен 663–560 = 103 долл.

Ниже приведен график из примера 2 параграфа 1.14.1.



Для определения нижнего квартиля найдем значение массы, которое точно соответствует кумулятивной частоте, равной одной четверти суммарных частот. В этом примере одна четверть суммарной частоты равна 260 : 4 = 64, а соответствующая масса равна 161 г.

Для определения верхнего квартиля найдем значение массы, которое точно соответствует кумулятивной частоте, равной трем четвертям суммарной частоты. В этом примере три четверти суммарной частоты будут равны $260 \times 3 : 4 = 195$, а соответствующая масса равна 202 г.

Межквартильный размах равен 202 – 161 = 41 г.

Можно также вычислять межквартильный размах интервального распределения при помощи расчетов, аналогичных расчету медианы.

Пример 1.

Заработки (долл.) «до», или «меньше чем»	Кумулятивная частота
400	0
500	80
550	150
600	270
650	480
700	660
800	700

Для определения нижнего квартиля найдем значение заработка, которое **точно** соответствует одной четверти суммарной частоты. В данном примере одна четверть суммарной частоты равна:

$$\frac{700}{4} = 175.$$

Из таблицы видно, что кумулятивная частота, равная 175, должна находиться между 550 долл. (кумулятивная частота равна 150) и 600 долл. (кумулятивная частота равна 270).

500 долл. 150
? .... 175 >25
600 долл. .... 270 >95

Для определения значения заработка, которое соответствует кумулятивной частоте, равной 175, делим расстояние между 550 и 600 долл. в таком же соотношении, в котором 175 делит расстояние между кумулятивными частотами 150 и 270, т.е. в отношении 25 : 95.

Таким образом, нижний квартиль равен $550 + 50 \times \frac{25}{120} =$ = 560 долл. (округляем до целого значения).

Для определения верхнего квартиля найдем значение заработка, которое **точно** соответствует трем четвертям суммарной частоты. В данном примере три четверти суммарной частоты равны:

$$700 \times \frac{3}{4} = 525.$$

Из таблицы видно, что кумулятивная частота, равная 525, должна находиться между 650 долл. (кумулятивная частота равна 480) и 700 долл. (кумулятивная частота равна 660).

650 долл. 480
? 525
700 долл. 660
>45
>135

Для определения значения заработка, которое соответствует кумулятивной частоте, равной 525, разделим расстояние между 650 и 700 долл. в таком же соотношении, в котором 525 делит расстояние между кумулятивными частотами 480 и 660, т.е. в отношении 45 : 135.

Таким образом, верхний квартиль равен $650 + 50 \times \frac{45}{180} =$ $= 662{,}5$ долл.

Межквартильный размах равен 662,5 – 560 = 102,5 долл.

Результат совпадает со значением, полученным с помощью графика.

Пример 2.

Масса животного (г) «до», или «меньше чем»	Кумулятивная частота
100	0
140	16
160	62
180	138
200	192
220	232
240	242
300	260

Для определения нижнего квартиля найдем значение массы, которое **точно** соответствует одной четверти суммарной частоты. В данном примере одна четверть суммарной частоты равна 260 : 4 = 65.

Из таблицы видно, что кумулятивная частота, равная 65, находится между 160 г (кумулятивная частота равна 62) и 180 г (кумулятивная частота равна 138).

160 г 62
? .... 65 >3
180 г .... 138 >73

Для определения значения массы, которое соответствует кумулятивной частоте, равной 65, разделим расстояние от 160 до 180 г в таком же соотношении, в котором 65 делит расстояние между кумулятивными частотами 62 и 138, т.е. в отношении 3 : 73.

Таким образом, нижний квартиль равен $160 + 20 \times \frac{3}{76} = 161$ г (округляем до целого значения).

Для определения верхнего квартиля найдем значение массы, которое **точно** соответствует трем четвертям суммарной частоты. В данном примере три четверти суммарной частоты равны:

$$260 \times \frac{3}{4} = 195.$$

Из таблицы видно, что кумулятивная частота, равная 195, должна находиться между 200 г (кумулятивная частота равна 192) и 220 г (кумулятивная частота равна 232).

200 г 192
? .... 195 >3
220 г .... 232 >37

Для определения значения массы, которое соответствует кумулятивной частоте, равной 195, разделим расстояние от 200 до 220 г в таком же соотношении, в котором 195 делит расстояние между кумулятивными частотами 192 и 232, т.е. в отношении 3 : 37.

Таким образом, верхний квартиль равен $200 + 20 \times \frac{3}{40} =$
$= 201{,}5$ г.

Межквартильный размах равен 201,5 г – 161 г = 40,5 г.

Результат совпадает со значением, полученным с помощью графика.

Обратите внимание, что:

(i) значения квартилей, полученные при помощи полигона кумулятивных частот, должны **в точности** совпадать с расчетными значениями;

(ii) значения квартилей, полученные при помощи кривой кумулятивных частот (огивы), должны быть **приблизительно** равны расчетным значениям.

Упражнение 16А

1. В таблице приведены результаты исследования **точного** возраста жителей города.

Возраст (годы)	0-	15-	30-	45-	65-	90-
Число жителей	150	210	280	330	70	0

(a) Для изображения распределения постройте полигон кумулятивных частот.

(b) На основе построенного графика определите нижний и верхний квартили, а также межквартильный размах.

(c) Проверьте полученные результаты, произведя расчеты на основе таблицы кумулятивных частот.

2. В таблице приведены результаты измерения протяженности 160 отрезков дороги в некотором регионе (результаты представлены с точностью до целых чисел).

Длина отрезков (км)	0-5	6-10	11-15	16-20	21-25	26-30	31-35	36-
Число отрезков	11	15	23	29	38	23	21	0

(a) Для изображения распределения постройте полигон кумулятивных частот.

(b) На основе построенного графика определите нижний и верхний квартили, а также межквартильный размах.

(c) Проверьте полученные результаты, произведя расчеты на основе таблицы кумулятивных частот.

Упражнение 16B

1. При проведении эксперимента ученикам дали задание нарисовать окружность радиусом 7,0 см без использования линейки. Радиусы 36 нарисованных окружностей были измерены, а результаты представлены в таблице.

Радиус (см)	5,0-	6,0-	7,0-	8,0-	9,0-	10,0-	11,0-
Число учеников	3	5	8	10	7	3	0

(a) Для изображения распределения постройте полигон кумулятивных частот.
(b) На основе построенного графика определите нижний и верхний квартили, а также межквартильный размах.
(c) Проверьте полученные результаты, произведя расчеты на основе таблицы кумулятивных частот.

2. Ученикам другого класса дали задание нарисовать квадрат со стороной 8,0 см без использования линейки. Длины оснований квадратов округлили до целых чисел. Результаты приведены ниже.

Длина основания (см)	6	7	8	9	10
Число учеников	3	6	10	5	4

(a) Для изображения распределения постройте полигон кумулятивных частот.
(b) На основе построенного графика определите нижний и верхний квартили, а также межквартильный размах.
(c) Проверьте полученные результаты, произведя расчеты на основе таблицы кумулятивных частот.

3.3. Перцентили

Квартили, включая медиану, – это значения переменной, которые делят совокупность на четверти.
(a) Одна четверть совокупности находится до нижнего квартиля.
(b) Одна четверть совокупности находится между нижним квартилем и медианой.
(c) Одна четверть совокупности находится между медианой и верхним квартилем.

(d) Одна четверть совокупности находится после верхнего квартиля.

По аналогии, перцентили делят совокупность в процентном отношении.

(i) 10-й (десятый) перцентиль – это такое значение переменной, что 10% совокупности находится до него, а 90% – после него.

(ii) 20-й (двадцатый) перцентиль – это такое значение переменной, что 20% совокупности находится до него, а 80% – после него.

(iii) 70-й (семидесятый) перцентиль – это такое значение переменной, что 70% совокупности находится до него, а 30% – после него.

Отсюда следует, что между 30-м (тридцатым) и 70-м (семидесятым) перцентилями находится 40% совокупности.

Перцентили можно вычислять таким же способом, как медиану и квартили:

(a) определив их на основе графика кумулятивных частот;

(b) рассчитав их на основе таблицы кумулятивных частот.

Пример.

Определите 10-й, 20-й и 40-й перцентили распределения, приведенного в примере 1 раздела 1.14.

Решение.

В совокупности 700 единиц.

10% от 700 составляет 70, т.е. 10-й перцентиль соответствует кумулятивной частоте, равной 70.

20% от 700 составляет 140, т.е. 20-й перцентиль соответствует кумулятивной частоте, равной 140.

40% от 700 составляет 280, т.е. 40-й перцентиль соответствует кумулятивной частоте, равной 280.

(a) На основе полигона кумулятивных частот из параграфа 3.22 определяем:

10-й перцентиль (кумулятивная частота 70) приблизительно равен 490 долл.;

20-й перцентиль (кумулятивная частота 140) приблизительно равен 540 долл.;

40-й перцентиль (кумулятивная частота 280) приблизительно равен 604 долл.

(ii) Таблица кумулятивных частот из параграфа 1.14:

Заработки (долл.) «до», или «меньше чем»	Кумулятивная частота
400	0
500	80
550	150
600	270
650	480
700	660
800	700

10-й перцентиль (кумулятивная частота 70) равен:

$$400+\frac{70}{80}\times(500-400) = 487{,}5 \text{ долл.};$$

20-й перцентиль (кумулятивная частота 140) равен:

$$500+\frac{60}{70}\times(550-500) = 542{,}9 \text{ долл.};$$

40-й перцентиль (кумулятивная частота 280) равен:

$$600+\frac{10}{210}\times(650-600) = 602{,}4 \text{ долл.}$$

Аудиторный тест № 14

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Все вопросы относятся к распределениям кумулятивных частот.

В вопросах с 1 по 10 приведены суммарные частоты. Необходимо определить кумулятивные частоты, которые соответствуют приведенным мерам.

	Вопрос		Варианты ответов				
	Суммарные частоты	Меры	A	B	C	D	E
1.	120	Медиана	50	50½	60	60½	70
2.	140	Нижний квартиль	25	25½	35	75	105
3.	180	Верхний квартиль	45	75	75½	135	135½
4.	180	10-й перцентиль	10	18	25	45	162
5.	120	20-й перцентиль	10	12	20	24	30
6.	320	Медиана	25	50	80	160	240
7.	240	Верхний квартиль	60	75	120	180	216
8.	180	70-й перцентиль	30	54	70	126	135
9.	360	Нижний квартиль	25	60	90	180	270
10.	550	60-й перцентиль	33	60	275	330	385

В вопросах с 11 по 20 приведены соответствующие значения переменных и кумулятивных частот. Необходимо найти значение переменной, которое соответствует выделенной справа кумулятивной частоте.

	Вопрос						Варианты ответов				
	Переменная	Кумулятивная частота	Переменная	Кумулятивная частота	Переменная	Кумулятивная частота	A	B	C	D	E
11.	60	45	80	55		50	65	70	72	74	75
12.	78	20	86	30		25	79	80	81	82	83
13.	86	65	94	85		75	90	91	92	93	94
14.	60	40	90	70		50	65	68	70	72	75
15.	36	40	60	70		60	44	48	52	56	60
16.	42	36	70	64		43	46	47	48	49	50
17.	54	56	70	72		68	66	68	75	78	83
18.	33	35	68	85		45	40	47	54	61	68
19.	33	35	68	85		65	40	47	54	61	68
20.	33	35	68	85		75	40	47	54	61	68

Упражнение 17А

1. Найдите 10-й, 20-й и 70-й перцентили при помощи полигона кумулятивных частот, построенного при решении задачи 1 в упражнении 16А. Проверьте ответ при помощи вычислений на основе таблицы кумулятивных частот.

2. Найдите 30-й, 60-й и 80-й перцентили при помощи полигона кумулятивных частот, построенного при решении задачи 2 в упражнении 16А. Проверьте ответ при помощи вычислений на основе таблицы кумулятивных частот.

Упражнение 17В

1. Найдите 40-й, 70-й и 90-й перцентили при помощи полигона кумулятивных частот, построенного при решении задачи 1 в упражнении 16В. Проверьте ответ при помощи вычислений на основе таблицы кумулятивных частот.

2. Найдите 10-й, 30-й и 80-й перцентили при помощи полигона кумулятивных частот, построенного при решении задачи 2 в упражнении 16В. Проверьте ответ при помощи вычислений на основе таблицы кумулятивных частот.

3.4. Среднее отклонение

3.4.1. Расчет среднего отклонения по индивидуальным значениям

Среднее отклонение – это средняя арифметическая модулей отклонений от средней арифметической (модуль отклонения – это величина отклонения без учета знака «+» или «–»).

Среднее отклонение $= \frac{\sum|d|}{n}$, где $|d|$ означает модуль d.

Пример 1. Определите среднее отклонение чисел 3, 6, 9, 10, 12.

Решение. Средняя арифметическая чисел равна $= \frac{\sum x}{n} = \frac{40}{5} = 8.$

Отклонения чисел от средней арифметической составляют

$$-5, -2, +1, +2, +4.$$

На этом этапе важно проверить, что $\Sigma d = 0$.

В нашем примере $\Sigma d = +7 - 7 = 0$.

Модули отклонений чисел от их средней арифметической равны

$$5, 2, 1, 2, 4.$$

Средняя арифметическая модулей равна $\frac{\sum|d|}{n} = \frac{14}{5} = 2{,}8.$

Следовательно, среднее отклонение равно 2,8.

Пример 2. Определите среднее отклонение чисел 5, 4, 7, 6, 2, 0, 3, 0, 8, 9.

Решение. Средняя арифметическая чисел равна $\frac{\sum x}{n} = \frac{46}{10} = 4{,}6.$

Отклонения чисел от средней арифметической составляют

$$+0{,}4;\ -0{,}6;\ +2{,}4;\ +1{,}4;\ -2{,}6;\ -4{,}6;-1{,}6;\ +3{,}4;\ +4{,}4.$$

На этом этапе важно проверить, что $\Sigma d = 0$.

В нашем примере $\Sigma d = +12 - 14 = -2$.

Это говорит о том, что что-то неверно!

Пересчитав среднюю арифметическую, находим, что она равна $\frac{44}{10} = 4{,}4.$

Пересчитываем отклонения чисел от средней арифметической:

$$+0{,}6;\ -0{,}4;\ +2{,}6;\ +1{,}6;\ -2{,}4;\ -4{,}4;\ -1{,}4;\ +3{,}6;\ +4{,}6.$$

Проверяем заново Σd. На этот раз $\Sigma d = +13 - 13 = 0$.

Вычисляем модули отклонений чисел от их средней арифметической:

$$+0{,}6;\ -0{,}4;\ +2{,}6;\ +1{,}6;\ -2{,}4;\ -4{,}4;\ -1{,}4;\ +3{,}6;\ +4{,}6.$$

Средняя арифметическая модулей равна $\frac{\sum|d|}{n} = \frac{26}{10} = 2{,}6.$

Следовательно, среднее отклонение равно 2,6.

Это означает, что среднее расстояние между значениями и их средней составляет 2,6.

Аудиторный тест № 15

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Определите среднее отклонение в каждой из приведенных последовательностей чисел.

	Вопрос	**Варианты ответов**				
		A	B	C	D	E
1.	3, 4, 5, 6, 7	1	1,2	1,4	1,6	1,8
2.	12, 13, 15, 17, 18	1,2	1,4	1, 6	1, 8	2
3.	22, 24, 25, 26, 28	1,2	1,4	1, 6	1, 8	2
4.	32, 32, 35, 38, 38	2	2,2	2,4	2,6	2,8
5.	12, 17, 23, 36, 42	10,2	10,4	10,6	10,8	11
6.	5, 6, 8, 9	1,1	1,2	1,3	1,4	1,5
7.	7, 9, 15, 17	3,4	3,6	3,8	4	1,5
8.	11, 12, 13, 14	0,75	1	1,25	1,5	1,75
9.	15, 15, 20, 26	4	4,25	4,5	4,75	5
10	21, 23, 29, 31	4	4,25	4,5	4,75	5
11.	3, 6 ,9	1	1,33	1,67	2	5
12.	15, 18, 30	5,33	5,67	6	6,33	6,67
13.	35, 37, 48	5,33	5,67	6	6,33	6,67
14.	30, 30, 60	12	12,3	12,7	13	13,3
15.	20, 40, 60	12	12,3	12,7	13	13,3
16.	3, 3, 8, 8, 8	2	2,2	2,4	2,6	2,8
17.	13, 19	2	3	4	5	6
18.	45, 50, 60, 65	7,25	7,5	7,75	8	8,25
19.	57, 59, 64	2	2,33	2,67	3	3,33
20.	12, 13, 14, 15, 21	2,2	2,4	2,6	2,8	3

Упражнение 18A

Определите (i) среднюю арифметическую;
(ii) среднее отклонение
в каждой из приведенных последовательностей чисел:

1. 9, 1, 3, 8, 3, 9, 2
2. 3, 4, 5, 6 ,7, 9, 12, 13, 11, 10
3. 5, 8, 8, 6, 8
4. 5, 5, 3, 2, 0
5. 3, 12, 4, 9, 4, 8, 5, 8, 1
6. 11, 3, 5, 10, 10, 9, 6, 7, 8, 11
7. 2, 9, 8, 6, 4, 5, 4, 10
8. 3, 4, 5, 6, 14, 15, 16, 17
9. 0, 19, 1, 17, 2, 4, 5, 15, 0
10. 2, 19, 18, 16, 3, 5, 7

Упражнение 18B

Определите (i) среднюю арифметическую;
(ii) среднее отклонение
в каждой из приведенных последовательностей чисел:

1. 2, 9, 7, 3, 2, 1
2. 0, 1, 3, 4, 6, 7, 8, 10, 15
3. 12, 3, 11, 4, 6, 8, 9, 3
4. 5, 5, 3, 2, 0, 5, 1
5. 11, 5, 6, 9, 9, 6, 8, 7, 2, 7
6. 2, 13, 11, 3, 1
7. 7, 13, 6, 12, 6, 4
8. 3, 13, 4, 12, 4, 10, 9, 8, 0
9. 0, 19, 1, 17, 2, 4, 5, 15, 0
10. 4, 20, 5, 20, 19, 6, 19, 8, 9, 0

3.4.2. Расчет среднего отклонения по сгруппированным значениям

Среднее отклонение по сгруппированным значениям вычисляется по формуле $\frac{\sum f|d|}{\sum f}$.

Пример 1. Определите среднее отклонение в приведенном ниже частотном распределении.

x	f
10	2
9	5
8	10
7	7
6	1

Решение. В первую очередь надо определить среднюю величину.

Найдем суммы: $\Sigma f = 25$ и $\Sigma fx = 200$.

x	f	fx
10	2	20
9	5	45
8	10	80
7	7	49
6	1	6

Средняя арифметическая $= \frac{\sum fx}{\sum f} = 200 : 25 = 8.$

Затем определим отклонения и умножим каждое отклонение на соответствующую частоту.

Найдем суммы: $\Sigma f = 25$ и $\Sigma fd = +9 - 9 = 0$, что говорит о правильности расчетов.

x	f	d	fd	$f\|d\|$
10	2	+2	+4	4
9	5	+1	+5	5
8	10	0	0	0
7	7	–1	–7	7
6	1	–2	–2	2

Затем найдем $\Sigma f|d| = 18$.

Среднее отклонение $= \dfrac{\sum f|d|}{\sum f} = \dfrac{18}{25} = 0{,}7.$

Следовательно, среднее отклонение равно 0,7.

Пример 2. Определите среднее отклонение в приведенном ниже частотном распределении.

x	f
18	1
19	4
20	7
21	9
22	6
23	3

Решение. В первую очередь надо определить среднюю арифметическую.

x	f	fx
18	1	18
19	4	76
20	7	140
21	9	189
22	6	132
23	3	69
Сумма	$\Sigma f = 30$	$\Sigma fx = 624$

Средняя арифметическая равна $\frac{624}{30} = 20,8.$

Затем определим отклонения и умножим каждое отклонение на соответствующую частоту.

x	f	d	fd	$f\|d\|$
18	1	–2,8	–2,8	2,8
19	4	–1,8	–7,2	7,2
20	7	–0,8	–5,6	5,6
21	9	+0,2	+1,8	1,8
22	6	+1,2	+7,2	7,2
23	3	+2,2	+6,6	6,6

Найдем суммы: $\Sigma f = 30$ и $\Sigma fd = +15,6 - 15,6 = 0$, что говорит о правильности расчетов.

Затем найдем $\Sigma f|d| = 31,2$.

Среднее отклонение = $\frac{\sum f|d|}{\sum f} = \frac{31,2}{30} = 1,04.$

Упражнение 19А

Определите среднюю арифметическую и среднее отклонение (от средней арифметической) по каждому из приведенных частотных распределений.

1.	Переменная	x	3	4	5	6	7	
	Частота	f	2	6	9	4	3	
2.	Переменная	x	4	5	6	7	8	9
	Частота	f	1	3	8	14	9	4
3.	Переменная	x	1	2	3	4	5	6
	Частота	f	3	6	9	5	2	1
4.	Переменная	x	4	5	6	7	8	
	Частота	f	3	6	7	1	1	

Упражнение 19В

Определите среднюю арифметическую и среднее отклонение (от средней арифметической) по каждому из приведенных частотных распределений.

1.	Переменная	x	5	6	7	8	9	
	Частота	f	3	3	8	5	2	
2.	Переменная	x	4	5	6	7	8	9
	Частота	f	3	7	9	8	1	1
3.	Переменная	x	1	2	3	4	5	6
	Частота	f	1	2	9	11	8	4
4.	Переменная	x	4	5	6	7	8	
	Частота	f	1	3	10	9	5	

3.5. Дисперсия и стандартное отклонение

3.5.1. Понятие дисперсии и стандартного отклонения

Дисперсия распределения – это средняя арифметическая квадратов отклонений от средней арифметической.

Дисперсия $= \dfrac{\sum d^2}{n}$.

Стандартное отклонение – это квадратный корень из дисперсии.

Стандартное отклонение – это «среднее квадратическое отклонение от средней».

Стандартное отклонение равно $\sqrt{\dfrac{\sum d^2}{n}}$.

Пример 1. Определить стандартное отклонение чисел 1, 2, 3, 4, 5.

Решение. В первую очередь определяем среднюю арифметическую.

Средняя арифметическая $x = \dfrac{\sum x}{n} = \dfrac{15}{5} = 3$ (что очевидно).

Отклонения от средней арифметической составляют: –2, –1, 0, +1, +2.

Проверяем, чтобы сумма отклонений была равна 0: –3 + 3 = 0.

Проверка подтверждает правильность расчетов.

Затем находим квадраты отклонений: 4, 1, 0, 1, 4.

(Обратите внимание, что все квадраты являются положительными.)

Сумма квадратов равна $\Sigma d^2 = 10$.

Дисперсия $= \frac{\sum d^2}{n} = \frac{10}{5} = 2.$

Стандартное отклонение – квадратный корень из 2, что составляет 1,41... .

Пример 2. Определить стандартное отклонение чисел

$$1, 6, 5, 7, 8, 2, 4, 7, 9, 1.$$

Решение. В первую очередь определяем среднюю арифметическую.

Средняя арифметическая $\frac{\sum x}{n} = \frac{50}{10} = 5.$

Отклонения от средней арифметической составляют

$$-4, +1, 0, +2, +3, -3, -1, +2, +4, -4.$$

Проверяем, чтобы сумма отклонений была равна нулю: –12 + 12 = 0.

Проверка подтверждает правильность расчетов.

Затем находим квадраты отклонений: 16, 1, 0, 4, 9, 9, 1, 4, 16, 16.

(Обратите внимание, что все квадраты являются положительными.)

Сумма квадратов $\Sigma d^2 = 76$.

Дисперсия $= \frac{\sum d^2}{n} = \frac{76}{10} = 7{,}6.$

Стандартное отклонение – это квадратный корень из 7,6, что составляет 2,76.

3.5.2. Единицы измерения стандартного отклонения и дисперсии

Стандартное отклонение имеет те же единицы измерения, что и исходные данные, например, см, г, кг и т.д.

Дисперсия измеряется в единицах, которые соответствуют квадрату единиц измерения исходных данных.

Примеры.	Исходные данные	Стандартное отклонение	Дисперсия
	см	см	$см^2$
	г	г	$г^2$
	кг	кг	$кг^2$
	км	км	$км^2$

3.5.3. Расчет дисперсии и стандартного отклонения по индивидуальным значениям

Метод 1. Использование средней арифметической

Дисперсия $= \dfrac{\sum d^2}{n}$.

Стандартное отклонение равно квадратному корню из $\dfrac{\sum d^2}{n}$.

Примеры 1 и 2 из параграфа 3.5.1 выполнены на основе данного метода.

Метод 2. Использование отклонений от условной средней

Дисперсия $= \dfrac{\sum D^2}{n} - \left(\dfrac{\sum D}{n}\right)^2$.

Стандартное отклонение равно $\sqrt{\dfrac{\sum D^2}{n} - \left(\dfrac{\sum D}{n}\right)^2}$.

Пример 1. Определить среднее отклонение чисел 1, 2, 3, 4, 5.

Решение. Для того чтобы проиллюстрировать этот метод, примем число 2 в качестве условной средней.

Отклонения от условной средней составят: –1, 0, +1, +2, +3.

(В этом методе проверка не производится.)

Квадраты отклонений от условной средней составят: 1, 0, 1, 4, 9.

Таким образом, $n = 5$, $\Sigma D = +5$, $\Sigma D^2 = 15$.

На основе приведенной выше формулы дисперсия равна

$$\frac{15}{5}-\left(\frac{5}{5}\right)^2=3-1=2.$$

Таким образом, стандартное отклонение равно квадратному корню из 2, т.е. 1,41.

Эти результаты совпадают с полученными в примере 1 из параграфа 3.5.1.

Пример 2. Определить среднее отклонение чисел 1, 6, 5, 7, 8, 2, 4, 7, 9, 1.

Решение. Принимаем число 7 в качестве условной средней.

Отклонения от условной средней составят: –6, –1, –2, 0, +1, –5, –3, 0, +2, –6.

(В этом методе проверка не производится.)

Квадраты отклонений от условной средней составят: 36, 1, 4, 0, 1, 25, 9, 0, 4, 36.

Таким образом, $n = 10$, $\Sigma D = -20$, $\Sigma D^2 = 116$.

На основе приведенной выше формулы дисперсия равна

$$\frac{116}{10}-\left(\frac{20}{10}\right)^2=11,6-4=7,6.$$

Таким образом, стандартное отклонение равно квадратному корню из 7,6, т.е. 2,76.

Эти результаты совпадают с полученными в примере 2 из параграфа 3.5.1.

Метод 3. Использование исходных данных

Дисперсия $= \dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2$.

Стандартное отклонение равно $\sqrt{\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2}$.

Пример 3. Определите среднее отклонение чисел 1, 2, 3, 4, 5.

Решение. Квадраты исходных данных равны 1, 4, 9, 16, 25.
Таким образом, $n = 5$, $\Sigma x^2 = 55$, $\Sigma x = 15$.

Используя приведенную выше формулу, получаем, что дисперсия $= \dfrac{55}{10} - \left(\dfrac{15}{5}\right)^2 = 11 - 9 = 2.$

Таким образом, стандартное отклонение равно квадратному корню из 2, т.е. 1,41.

Эти результаты совпадают с полученными в примере 1 из параграфа 3.5.1 и примере 1 из параграфа 3.5.3.

Пример 4. Определить среднее отклонение чисел 1, 6, 5, 7, 8, 2, 4, 7, 9, 1.

Решение. Квадраты исходных данных равны 1, 36, 25, 49, 64, 4, 16, 49, 81, 1.
Таким образом, $n = 10$, $\Sigma x^2 = 326$, $\Sigma x = 50$.

Используя приведенную выше формулу, получаем, что дисперсия $= \dfrac{326}{10} - \left(\dfrac{50}{10}\right)^2 = 32{,}6 - 25 = 7{,}6.$

Таким образом, стандартное отклонение равно квадратному корню из 7,6, т.е. 2,76.

Эти результаты совпадают с полученными в примере 2 из параграфа 3.5.1 и примере 2 из параграфа 3.5.3.

Метод 4. Использование инженерного калькулятора

Многие современные калькуляторы позволяют вычислить стандартное отклонение исходных данных. В калькуляторах заложены разные принципы расчета. Если у вас возникнут трудности при использовании калькулятора для расчета стандартного отклонения, обратитесь за помощью к преподавателю. Для определения дисперсии при помощи калькулятора надо вначале вычислить стандартное отклонение и затем полученный результат возвести в квадрат.

Упражнение 20A

В этом упражнении все ответы давайте с точностью до двух знаков после запятой.

Для каждой из приведенных последовательностей чисел определите:

(i) среднюю арифметическую;

(ii) дисперсию и стандартное отклонение, используя отклонения от средней арифметической (метод 1).

1. 9, 5, 1, 4, 6, 2, 8, 3, 7

2. 4, 3, 6, 2, 9, 8, 5, 7, 0, 16

Для каждой из приведенных последовательностей чисел определите дисперсию и стандартное отклонение, используя отклонения от условной средней (метод 2).

3. 7, 6, 4, 2, 8, 1, 3, 5 условная средняя = 4

4. 9, 8, 5, 0, 6, 7, 2, 4, 4, 10 условная средняя = 5

Для каждой из приведенных последовательностей чисел определите дисперсию и стандартное отклонение по исходным данным (метод 3). Если есть возможность, используйте калькулятор для проверки ответов.

5. 2, 6, 0, 7, 9, 12, 8, 5

6. 1, 0, 3, 7, 4, 6, 8, 9, 2

7. 4, 3, 2, 7, 8, 9, 5, 3, 0

Упражнение 20В

В этом упражнении все ответы давайте с точностью до двух знаков после запятой.

Для каждой из приведенных последовательностей чисел определите:

(i) среднюю арифметическую;

(ii) дисперсию и стандартное отклонение, используя отклонения от средней арифметической (метод 1).

1. 8, 6, 2, 3, 5, 1, 7, 4, 9

2. 16, 1, 8, 6, 9, 5, 4, 7, 3, 11

Для каждой из приведенных последовательностей чисел определите дисперсию и стандартное отклонение, используя отклонения от условной средней (метод 2).

3. 6, 5, 3, 1, 4, 1, 0, 8 условная средняя = 4

4. 10, 9, 6, 1, 7, 8, 3, 4, 5, 12 условная средняя = 6

Для каждой из приведенных последовательностей чисел определите дисперсию и стандартное отклонение по исходным данным (метод 3). Если есть возможность, используйте калькулятор для проверки ответов.

5. 4, 0, 7, 5, 3, 6, 2, 9, 1

6. 8, 7, 9, 8, 7, 6, 9, 9, 10

7. 7, 9, 4, 2, 0, 1, 6, 5, 12

3.5.4. Расчет дисперсии и стандартного отклонения для дискретного ряда распределения

Метод 1. Использование отклонений от средней арифметической

Дисперсия = $\dfrac{\sum fd^2}{\sum f}$. Среднее отклонение равно $\sqrt{\dfrac{\sum fd^2}{\sum f}}$.

Пример 1. Определить дисперсию и стандартное отклонение частотного распределения:

x	3	11	21
f	2	3	1

Решение. В первую очередь надо определить среднюю арифметическую.

x	f	fx
3	2	6
11	3	33
21	1	21
Сумма	$\Sigma f = 6$	$\Sigma fx = 60$

Средняя арифметическая $= \frac{60}{6} = 10.$

x	f	d	fd
3	2	–7	–14
11	3	+1	+3
21	1	+11	+11
	$\Sigma f = 6$		$\Sigma fd = +14 – 14 = 0$

Затем определяем отклонения и умножаем каждое отклонение на соответствующую частоту.

На этом этапе проверяется, что $\Sigma fd = 0$, что говорит о правильности расчетов.

Затем вычисляются d^2 и fd^2.

x	f	d	d^2	fd^2
3	2	–7	49	98
11	3	+1	1	3
21	1	+11	121	121
	$\Sigma f = 6$			$\Sigma fd^2 = 222$

Дисперсия $= \frac{\sum fd^2}{\sum f} = \frac{222}{6} = 37.$

Стандартное отклонение = квадратный корень из дисперсии =
= квадратный корень из 37 = 6,083.

Метод 2. Использование отклонений от условной средней

Дисперсия $= \frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2.$

Стандартное отклонение равно $\sqrt{\frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2}.$

Пример 2. Определить дисперсию и стандартное отклонение дискретного ряда распределения:

x	3	11	21
f	2	3	1

Решение. Принимаем число 11 в качестве условной средней. Затем вычисляем отклонения от условной средней и – D^2.

x	*F*	*D*	*fD*	D^2	fD^2
3	2	–8	–16	64	128
11	3	0	0	0	0
21	1	+10	+10	100	100
	$\Sigma f = 6$		$\Sigma fD = -6$		$\Sigma fD^2 = 228$

Дисперсия $= \frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2 = \frac{228}{6} - \left(\frac{-6}{6}\right)^2 = 38 - 1 = 37.$

Стандартное отклонение равно квадратному корню из дисперсии, т.е. квадратному корню из 37, что составляет 6,083.

Этот ответ совпадает с результатом, полученным методом 1.

Метод 3. Использование исходных данных

Дисперсия $= \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2.$

Стандартное отклонение равно $\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}.$

Пример 3. Определить дисперсию и стандартное отклонение дискретного ряда распределения:

x	3	11	21
f	2	3	1

Решение.

x	f	fx	x^2	fx^2
3	2	6	9	18
11	3	33	121	363
21	1	21	441	441
	$\Sigma f = 6$	$\Sigma fx = 60$		$\Sigma fx^2 = 822$

$$\text{Дисперсия } = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2 = \frac{822}{6} - \left(\frac{60}{6}\right)^2 = 137 - 100 = 37.$$

Стандартное отклонение равно квадратному корню из дисперсии, т.е. квадратному корню из 37, что составляет 6,083.

Этот ответ совпадает с результатом, полученным методами 1 и 2.

Обзор формул для вычисления дисперсии и стандартного отклонения

	Индивидуальные значения	Частотное распределение
Использование отклонений от средней арифметической	$\sqrt{\frac{\sum d^2}{n} - \cancel{\left(\frac{\sum d}{n}\right)^2}}$ второе слагаемое равно 0, так как $\Sigma d = 0$	$\sqrt{\frac{\sum fd^2}{\sum f} - \cancel{\left(\frac{\sum fd}{\sum f}\right)^2}}$ второе слагаемое равно 0, так как $\Sigma d = 0$
Использование отклонений от условной средней	$\sqrt{\frac{\sum D^2}{n} - \left(\frac{\sum D}{n}\right)^2}$	$\sqrt{\frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2}$
Использование исходных данных	$\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$	$\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$

Метод 4. Использование инженерного калькулятора

Многие современные калькуляторы позволяют вычислить стандартное отклонение для дискретного ряда распределения. В калькуляторах заложены разные принципы расчета. Если у вас возникнут трудности при использовании калькулятора для расчета стандартного отклонения, обратитесь за помощью к преподавателю. Для определения дисперсии при помощи калькулятора надо вначале вычислить стандартное отклонение и затем полученный результат возвести в квадрат.

Упражнение 21А

В этом упражнении все ответы давайте с точностью до двух знаков после запятой.

Для каждого из приведенных дискретных частотных распределений определите:

(i) среднюю арифметическую;

(ii) дисперсию и стандартное отклонение, используя отклонения от средней арифметической (метод 1).

1.	Переменная	x	4	5	6	7	8	
	Частота	f	2	6	9	4	3	
2.	Переменная	x	1	2	3	4	5	6
	Частота	f	1	3	8	14	9	4
3.	Переменная	x	0,5	0,6	0,7	0,8	0,9	1,0
	Частота	f	3	6	9	5	2	1
4.	Переменная	x	6	7	8	9	10	
	Частота	f	3	6	7	1	1	

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение, используя отклонения от условной средней (метод 2).

5.	Переменная	x	7	8	9	10	11		условная
	Частота	f	2	4	7	6	3		средняя = 9
6.	Переменная	x	7	8	9	10	11	12	условная
	Частота	f	1	3	4	8	5	2	средняя = 10
7.	Переменная	x	0	0,1	0,2	0,3	0,4	0,5	условная
	Частота	f	3	5	7	4	2	1	средняя = 0,2
8.	Переменная	x	3	4	5	6	7		условная
	Частота	f	3	4	7	8	5		средняя = 5

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение по исходным данным (метод 3).

9.	Переменная	x	5	6	7	8	9	
	Частота	f	2	3	4	7	5	
10.	Переменная	x	1	2	3	4	5	6
	Частота	f	2	3	4	7	3	1
11.	Переменная	x	1,1	1,2	1,3	1,4	1,5	
	Частота	f	3	4	7	6	2	
12.	Переменная	x	2	3	4	5	6	
	Частота	f	2	4	7	5	3	

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение, используя инженерный калькулятор (метод 4).

13.	Переменная	x	18	19	20	21	22	
	Частота	f	5	8	13	9	4	
14.	Переменная	x	21	22	23	24	25	26
	Частота	f	2	7	13	17	8	3
15.	Переменная	x	2,3	2,4	2,5	2,6	2,7	
	Частота	f	8	14	17	16	5	
16.	Переменная	x	32	33	34	35	36	
	Частота	f	3	5	7	4	2	

Упражнение 21B

В этом упражнении все ответы давайте с точностью до двух знаков после запятой.

Для каждого из приведенных дискретных частотных распределений определите:

(i) среднюю арифметическую;

(ii) дисперсию и стандартное отклонение, используя отклонения от средней арифметической (метод 1).

1.	Переменная	x	7	8	9	10	11	
	Частота	f	3	3	8	5	2	
2.	Переменная	x	7	8	9	10	11	12
	Частота	f	3	7	9	8	1	1
3.	Переменная	x	0	0,1	0,2	0,3	0,4	0,5
	Частота	f	1	2	9	11	8	4
4.	Переменная	x	3	4	5	6	7	
	Частота	f	1	3	10	9	5	

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение, используя отклонения от условной средней (метод 2).

5.	Переменная	x	4	5	6	7	8		условная средняя = 9
	Частота	f	3	4	9	6	2		
6.	Переменная	x	1	2	3	4	5	6	условная средняя = 10
	Частота	f	4	9	14	8	3	1	
7.	Переменная	x	0,5	0,6	0,7	0,8	0,9	1,0	условная средняя = 0,2
	Частота	f	1	3	6	9	5	2	
8.	Переменная	x	6	7	8	9	10		условная средняя = 5
	Частота	f	1	3	8	12	5		

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение по исходным данным (метод 3).

9.	Переменная	x	1	2	3	4	5	
	Частота	f	2	3	8	4	1	
10.	Переменная	x	3	4	5	6	7	8
	Частота	f	1	1	8	9	7	3
11.	Переменная	x	0,3	0,4	0,5	0,6	0,7	0,8
	Частота	f	1	4	7	6	3	2
12.	Переменная	x	3	4	5	6	7	
	Частота	f	5	9	10	3	1	

Для каждого из приведенных дискретных частотных распределений определите дисперсию и стандартное отклонение, используя инженерный калькулятор (метод 4).

13.	Переменная	x	11	12	13	14	15	
	Частота	f	2	5	9	8	3	
14.	Переменная	x	23	24	25	26	27	28
	Частота	f	3	8	10	15	9	4
15.	Переменная	x	0,3	0,4	0,5	0,6	0,7	0,8
	Частота	f	2	5	7	9	6	3
16.	Переменная	x	43	44	45	46	47	
	Частота	f	5	9	17	12	6	

3.5.5. Расчет дисперсии и стандартного отклонения для интервального ряда распределения

Для определения дисперсии интервального ряда распределения можно использовать любой из четырех методов, рассмотренных в разделе 3.5.4 для дискретного ряда распределения, предварительно определив середину каждого интервала.

Пример 1. Определить дисперсию и стандартное отклонение для интервального ряда распределения.

Переменная Рост девочек (см) x	*Частота* Число девочек f
154-	1
160-	2
162-	7
164-	1
166-	3
168-	2
170–176	2

Решение.

В первую очередь определяем середину каждого интервала.

154-	160-	162-	164-	166-	168-	170-176
157	161	163	165	167	169	173

Метод 1. Использование средней арифметической

x	f	fx	d	fd	d^2	fd^2
157	1	157	–8	–8	64	64
161	2	322	–4	–8	16	32
163	7	1141	–2	–14	4	28
165	1	165	0	0	0	0
167	3	501	+2	+6	4	12
169	2	338	+4	+8	16	32
173	2	346	+8	+16	64	128
	$\Sigma f = 18$	$\Sigma fx = 2970$		$\Sigma fd = 0$		$\Sigma fd^2 = 296$

Средняя арифметическая $= \frac{\sum fx}{\sum f} = \frac{2970}{18} = 165.$

$\Sigma fd = 0$, что говорит о правильности расчетов.

Дисперсия $= \frac{\sum fd^2}{\sum f} = \frac{296}{18} = 16,444.$

Стандартное отклонение равно квадратному корню из дисперсии.

Стандартное отклонение равно квадратному корню из 16,444, т.е. 4,055.

Метод 2. Использование отклонений от условной средней (выбрано число 163)

x	f	D	fD	d^2	fD^2
157	1	–6	–6	36	36
161	2	–2	–4	4	8
163	7	0	0	0	0
165	1	+2	+2	4	4
167	3	+4	+12	16	48
169	2	+6	+12	36	72
173	2	+10	+20	100	200
	$\Sigma f = 18$		$\Sigma fD = 36$		$\Sigma fD^2 = 368$

Дисперсия равна $\frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2 = \frac{368}{18} - \left(\frac{36}{18}\right)^2 = 20{,}4444 - 4 =$

$= 16{,}4444.$

Стандартное отклонение равно квадратному корню из дисперсии.

Стандартное отклонение равно квадратному корню из 16,444, т.е. 4,055.

Результат совпадает со значением, полученным методом 1.

Метод 3. Использование исходных данных

x	f	fx	x^2	fx^2
157	1	157	24 649	24 649
161	2	322	25 921	51 842
163	7	1141	26 569	185 983
165	1	165	27 225	27 225
167	3	501	27 089	83 667
169	2	338	28 561	57 122
173	2	346	29 929	59 858
	$\Sigma f = 18$	$\Sigma fx = 2970$		$\Sigma fx^2 = 490\ 346$

$$\text{Дисперсия} = \frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2 = \frac{490\ 346}{18} - \left(\frac{2970}{18}\right)^2 =$$

$$= 27\ 241{,}444 - 27\ 225 = 16{,}444.$$

Стандартное отклонение равно квадратному корню из дисперсии.

Стандартное отклонение равно квадратному корню из 16,444, т.е. 4,055.

Результат совпадает со значением, полученным методами 1 и 2.

Метод 4. Использование инженерного калькулятора

Стандартное отклонение = 4,055.

Дисперсия равна квадрату стандартного отклонения, т.е. 16,444.

Упражнение 22А

Ответ дайте с точностью до двух знаков после запятой.

Определите дисперсию и стандартное отклонение в каждом из приведенных интервальных рядов распределения.

1.	Масса животного (г)	180-	200-	220-	240-	260-	280-	300-	320-
	Количество животных	5	8	12	15	22	13	6	0
2.	Высота растения (см)	5,0-	5,2-	5,4-	5,6-	5,8-	6,0-	6,2-	6,4-
	Количество растений	3	7	12	18	15	11	6	0
3.	Срок службы лампы (ч)	700-	750-	800-	850-	900-	950-	1000-	
	Количество ламп	13	18	46	72	63	23	0	
4.	Время в пути (ч)	1,0-	1,2-	1,4-	1,6-	1,8-	2,0-		
	Количество сотрудников	3	7	13	18	9	0		

Упражнение 22B

Ответ дайте с точностью до двух знаков после запятой.

Определите дисперсию и стандартное отклонение в каждом из приведенных интервальных распределений.

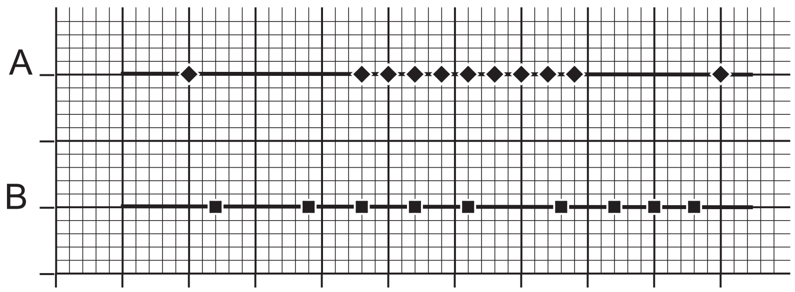
1.	Длина цилиндра (см)	4,0-	4,1-	4,2-	4,3-	4,4-	4,5-	4,6-
	Количество цилиндров	2	7	13	21	15	7	0
2.	Длина детали (см)	3,1-	3,2-	3,3-	3,4-	3,5-	3,6-	3,7-
	Количество деталей	3	9	20	10	7	4	0
3.	Рост ребенка (см)	150-	155-	160-	165-	170-	175-	180-
	Количество детей	0	4	9	15	7	3	0
4.	Диаметр шарикоподшипника (мм)	5,9-	6,0-	6,1-	6,2-	6,3-	6,4-	
	Количество шарикоподшипников	13	46	81	39	20	0	

3.6. Обзор показателей вариации, или колеблемости

(а) Размах

Размах легко вычислить, но он сильно зависит от экстремальных значений и, следовательно, не надежен.

Пример.

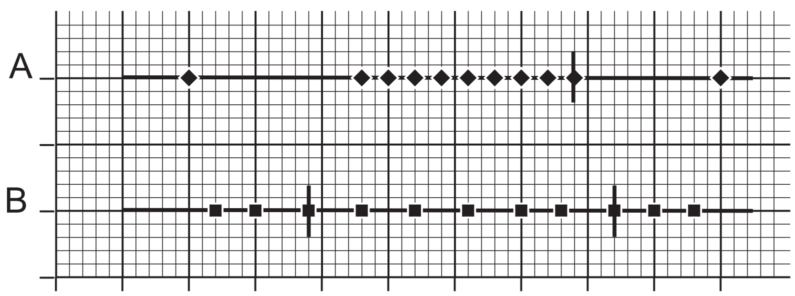


Размах в группе A больше размаха в группе B, но группа A в целом более плотная, чем группа B.

(b) Межквартильный размах

Межквартильный размах легко вычислить, он, конечно, является более надежной оценкой, чем размах, но редко используется при углубленном статистическом анализе.

Пример.



(c) Среднее отклонение

Среднее отклонение достаточно легко вычислить, но оно используется крайне редко.

(d) Стандартное отклонение и дисперсия

Вычисление стандартного отклонения и дисперсии утомительно, но они являются наиболее используемыми мерами вариации (применение инженерного калькулятора сильно сокращает время их расчета).

Глава 4

ЛИНЕЙНОЕ ПРЕОБРАЗОВАНИЕ ДАННЫХ

4.1. Необходимость нормирования экзаменационных оценок

В таблице приведены исходные оценки, полученные тремя экзаменуемыми по трем предметам:

	Анна	Бетти	Клара
Математика	95	5	50
Французский язык	43	67	55
Английский язык	54	60	66

Суммарный балл каждой экзаменуемой составил соответственно: 192 балла у Анны, 132 – у Бетти, 171 – у Клары.

Тогда расположение девушек в порядке величины суммарного балла будет следующим: Анна, Клара, Бетти.

Однако при более детальном изучении оценок получаем следующие значения среднего балла и среднего отклонения оценок по трем дисциплинам:

	Средний балл	Среднее отклонение
Математика	50	30
Французский язык	55	8
Английский язык	60	4

Поскольку размах, или колеблемость, оценок по математике значительно больше, чем колеблемость оценок по английскому языку, то при ранжировании девушек в порядке величины суммарного балла по трем дисциплинам преувеличивается значение оценок по математике по сравнению с оценками по английскому и французскому языкам.

Для уравнивания значимости оценок по всем дисциплинам нам необходимо добиться, чтобы оценки по всем трем дисциплинам имели одинаковый разброс. Сопоставим оценкам по каждой дисциплине «нормированную среднюю», равную 45, и «нормированное среднее отклонение», равное 10. (Можно было бы выбрать любые другие удобные числа, скажем, между 40 и 50 для средней и между 10 и 20 для величины среднего отклонения.)

Существует несколько способов, но в данном примере используется простейший для понимания способ.

<u>По математике</u>

Оценка Анны составила 95 баллов, что на 45 баллов больше среднего значения, равного 50.

Среднее отклонение равно 30, следовательно, 45 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Анны на 1½ величины нормированного среднего отклонения превышает нормированную среднюю и равна $45 + 1½ \times 10 = 60$.

Бетти получила 5 баллов, что на 45 баллов меньше среднего значения 50.

Среднее отклонение равно 30, следовательно, 45 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Бетти на 1½ величины нормированного среднего отклонения ниже нормированной средней и равна $45 – 1½ \times 10 = 30$.

Оценка Клары составила 50 баллов, что в точности равно среднему значению.

Следовательно, нормированная оценка Клары в точности равна нормированной средней, т.е. 45.

<u>По французскому языку</u>

Оценка Анны составила 43 балла, что на 12 баллов меньше среднего значения, равного 55.

Среднее отклонение равно 8, следовательно, 12 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Анны на 1½ величины нормированного среднего отклонения ниже нормированной средней и равна $45 - 1½ \times 10 = 30$.

Бетти получила 67 баллов, что на 12 баллов больше среднего значения.

Среднее отклонение равно 8, следовательно, 12 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Бетти на 1½ величины нормированного среднего отклонения превышает нормированную среднюю и равна $45 + 1½ \times 10 = 60$.

Оценка Клары составила 55 баллов, что в точности равно среднему значению.

Следовательно, нормированная оценка Клары в точности равна нормированной средней, т.е. 45.

<u>По английскому языку</u>

Оценка Анны составила 54 балла, что на 6 баллов меньше среднего значения, равного 60.

Среднее отклонение равно 4, следовательно, 6 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Анны на 1½ величины нормированного среднего отклонения ниже нормированной средней и равна $45 - 1½ \times 10 = 30$.

Оценка Бетти составила 60 баллов, что в точности равно среднему значению.

Следовательно, нормированная оценка Клары в точности равна нормированной средней, т.е. 45.

Клара получила 66 баллов, что на 6 баллов больше среднего значения.

Среднее отклонение равно 4, следовательно, 6 баллов составляет 1½ величины среднего отклонения. Таким образом, «нормированная» оценка Клары на 1½ величины нормированного среднего отклонения превышает нормированную среднюю и равна $45 + 1½ \times 10 = 60$.

Суммарный нормированный балл трех экзаменуемых по этим дисциплинам составил: 120 баллов у Анны, 135 – у Бетти и 145 – у Клары.

Теперь расположение девушек в порядке величины суммарного нормированного балла будет следующим: Клара, Бетти, Анна.

Рассмотренный пример показывает, что если для определения суммарного балла складываются баллы по различным дисциплинам, необходимо убедиться, что размах, или колеблемость, оценок по этим дисциплинам одинаковая. И здесь уже не имеет значения, что средние баллы по разным дисциплинам не равны.

4.2. Другие способы линейного преобразования, или нормирования, оценок

(i) Воспользуемся формулой

$$\frac{\text{Исходный балл} - \text{Средний балл}}{\text{Среднее отклонение}} =$$

$$= \frac{\text{Нормированный балл} - \text{Нормированный средний бал л}}{\text{Нормированное среднее отклонение}}$$

или

$$\frac{\text{Исходный балл} - \text{Средний балл}}{\text{Стандартное отклонение}} =$$

$$= \frac{\text{Нормированный балл} - \text{Нормированный средний бал л}}{\text{Нормированное стандартное отклонение}}.$$

Пример 1. Среднее значение и стандартное отклонение оценок равны соответственно 54 и 6. Оценки нормированы со средним значением 50 и стандартным отклонением 10.

Вычислите: a) нормированную оценку S, соответствующую 69 баллам;

b) оценку R, соответствующую нормированной оценке 40.

Решение.

a)
$$\frac{\text{Исходный балл} - \text{Средний балл}}{\text{Стандартное отклонение}} =$$

$$= \frac{\text{Нормированный балл} - \text{Нормированный средний ба лл}}{\text{Нормированное стандартное отклонение}},$$

следовательно, $\frac{(69-54)}{6} = \frac{(S-50)}{10}$.

Перепишем формулу следующим образом: $15 \times 10 = 6(S - 50)$, откуда получаем $S = 75$.

b) $$\frac{\text{Исходный балл} - \text{Средний балл}}{\text{Стандартное отклонение}} =$$

$$= \frac{\text{Нормированный балл} - \text{Нормированный средний бал л}}{\text{Нормированное стандартное отклонение}},$$

следовательно, $\dfrac{(R-54)}{6} = \dfrac{(40-50)}{10}$.

Из уравнения находим $R - 54 = -6$.

Таким образом, $R = 48$.

(ii) Графический способ

По горизонтальной оси X будем откладывать исходные оценки, а по вертикальной оси Y – нормированные.

Строим прямую по следующим трем точкам:

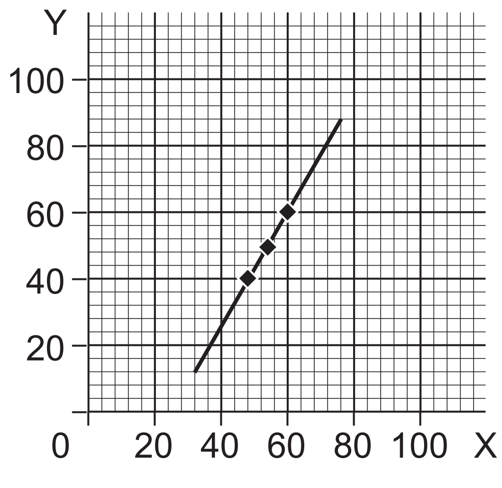
(средний балл по исходным оценкам; нормированный средний балл);

(средний балл + 1 отклонение по исходным данным; нормированный средний балл + 1 нормированное отклонение);

(средний балл – 1 отклонение по исходным данным; нормированный средний балл – 1 нормированное отклонение).

В примере 1 координаты этих точек равны соответственно (54; 50), (60; 60), (48; 40).

На графике ниже изображена прямая, построенная по этим точкам.



По этому графику мы легко определим, что:

a) нормированная оценка, соответствующая 69 баллам, равна 75;

b) оценка, соответствующая нормированной оценке 40, равна 48 баллам.

4.3. Уравнение линейного преобразования

Линейное преобразование может быть представлено уравнением вида $y = Ax + B$, где A и B – константы (фиксированные числа). При нормировании оценок можно обозначить нормированную оценку буквой s, а исходную оценку – буквой r. Тогда получим следующее уравнение:

$$s = Ar + B.$$

Вернемся к примеру 1 из параграфа 4.2.

Нормированная оценка 75 соответствует исходной оценке 69 баллов.

Подставляем эти данные в уравнение:

$$75 = A \times 69 + B. \qquad \text{(i)}$$

Нормированная оценка 40 соответствует исходной оценке 48 баллов.

Получаем второе уравнение:

$$40 = A \times 48 + B. \qquad \text{(ii)}$$

Решая систему этих двух уравнений, найдем коэффициенты A и B.

Вычитаем уравнение (ii) из уравнения (i):

$35 = 21A$, следовательно $A = \dfrac{35}{21} = \dfrac{5}{3}$.

Подставляя в уравнение (i) вместо A его значение $\dfrac{5}{3}$, получаем

$75 = \dfrac{5}{3} \times 69 + B$. Отсюда следует, что $B = –40$.

Уравнение линейного преобразования запишется следующим образом:

$$s = \frac{5}{3} \cdot r - 40 \cdot$$

Проверка.

(i) Если $r = 69$, то $s = \frac{5}{3} \times 69 - 40 = 75$, что совпадает с полученным ранее результатом.

(ii) Если $r = 48$, то $s = \frac{5}{3} \times 48 - 40 = 40$, что также совпадает с полученным ранее результатом.

Пример 1.

В линейном преобразовании $x = 40$ соответствует $y = 65$,
$x = 48$ соответствует $y = 77$.

(i) Выразите y через x.
(ii) Найдите значение y, если $x = 50$.
(iii) Найдите значение x, если $y = 59$.
(iv) Найдите $\overline{y}$ (среднюю арифметическую значений y), если известна средняя арифметическая значений x, равная $\overline{x} = 38$.

Решение.

(i) Запишем уравнение линейного преобразования $y = Ax + B$. Тогда

$$65 = 40A + B; \qquad \text{(i)}$$

$$77 = 48A + B. \qquad \text{(ii)}$$

Вычитая (i) из (ii), получаем:

$$12 = 8A, \text{ откуда } A = 1{,}5.$$

Подставляя вместо А его значение 1,5 в (i), находим значение B:

$$65 = 40 \times 1{,}5 + B, \text{ откуда } B = 5.$$

Следовательно, $y = 1{,}5x + 5$.

Проверка.

Подставляя $x = 40$, получаем $y = 1{,}5 \times 40 + 5 = 65$, что отвечает начальному условию.

Подставляя $x = 48$, получаем $y = 1{,}5 \times 48 + 5 = 77$, что также отвечает начальному условию.

(ii) Если $x = 50$, то $y = 1{,}5 \times 50 + 5 = 80$.
(iii) Пусть $y = 59$. Тогда получаем уравнение $59 = 1{,}5x + 5$. Откуда $1{,}5x = 54$ и $x = 36$.
(iv) $\overline{y} = 1{,}5\,\overline{x} + 5$,
значит, $\overline{y} = 1{,}5 \times 38 + 5 = 57 + 5 = 62$.

Аудиторный тест № 16

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

В этом задании отклонение может быть или средним отклонением, или стандартным. Знак «–» означает, что значение меньше среднего, знак «+» указывает, что значение больше среднего.

Вопрос				**Варианты ответов**				
				A	B	C	D	E
Вычислите величину отклонения оценки от среднего значения.								
	Среднее значение	Отклонение	Оценка					
1.	50	3	59	–3	–2	+1	+2	+3
2.	47	2	43	–3	–2	+1	+2	+3
3.	45	1	47	3	–2	+1	+2	+3
4.	51	4	39	3	–2	+1	+2	+3
5.	48	6	54	3	–2	+1	+2	+3
6.	52	5	67	3	–2	+1	+2	+3
7.	53	2,5	58	3	–2	+1	+2	+3
8.	49	1,5	46	3	–2	+1	+2	+3
9.	50	4	46	3	–2	+1	+2	+3
10.	46	3,5	53	3	–2	+1	+2	+3

Найдите оценку, имеющую заданную величину отклонения от среднего значения.

	Среднее значение	Отклонение	Величина отклонения					
11.	50	5	+3	35	47	53	55	65
12.	47	6	–2	35	41	45	53	59
13.	45	4	+2	37	47	49	53	55
14.	51	1	–3	48	49	50	52	54
15.	48	2	+1,5	49	49,5	50	50,5	51
16.	52	3	–3	43	46	49	55	61
17.	53	1,5	+4	49	51,5	54,5	57	59
18.	49	2,5	–2	44	46,5	47	51,5	56

Упражнение 23А

1. Среднее значение и стандартное отклонение всех экзаменационных оценок равны соответственно 40 и 10. Оценки нормированы со средним значением 50 и стандартным отклонением 12.

Вычислите: (i) нормированную оценку учащегося B, получившего 46 баллов;
(ii) оценку, полученную учащимся C, чья нормированная оценка составила 38.

2. По результатам экзамена среднее значение и стандартное отклонение всех оценок равны соответственно 45 и 14. Оценки нормированы со средним значением 50 и стандартным отклонением 16.

Вычислите: (i) оценку учащегося D, чья нормированная оценка составила 42 балла;
(ii) нормированную оценку учащегося E, получившего 66 баллов.

3. По результатам экзамена среднее значение и стандартное отклонение всех оценок равны соответственно 47 и 12. Оценки были нормированы со средним значением 50 и стандартным отклонением 16. Оценка экзаменуемого S была на 6 баллов выше оценки экзаменуемого T. На сколько баллов оценка экзаменуемого S больше оценки экзаменуемого T после нормирования оценок?

Упражнение 23B

1. По результатам экзамена среднее значение и стандартное отклонение всех оценок равны соответственно 48 и 12. Оценки нормированы со средним значением 45 и стандартным отклонением 15.

Вычислите: (i) оценку учащегося G, чья нормированная оценка составила 40 баллов;
(ii) нормированную оценку учащегося H, получившего 56 баллов.

2. Среднее значение и стандартное отклонение всех экзаменационных оценок равны соответственно 45 и 6. Оценки нормированы со средним значением 48 и стандартным отклонением 10.

Вычислите: (i) нормированную оценку учащегося J, получившего 33 баллов;
(ii) оценку, полученную учащимся K, чья нормированная оценка составила 73 балла.

3. Экзаменационные оценки двух экзаменуемых равны соответственно 51 и 63 балла, тогда как их нормированные оценки составили 47 и 65 баллов. Стандартное отклонение исходных оценок равно 8. Вычислите стандартное отклонение нормированных оценок.

Глава 5

ПРОСТЕЙШИЙ АНАЛИЗ ВРЕМЕННЫ́Х РЯДОВ

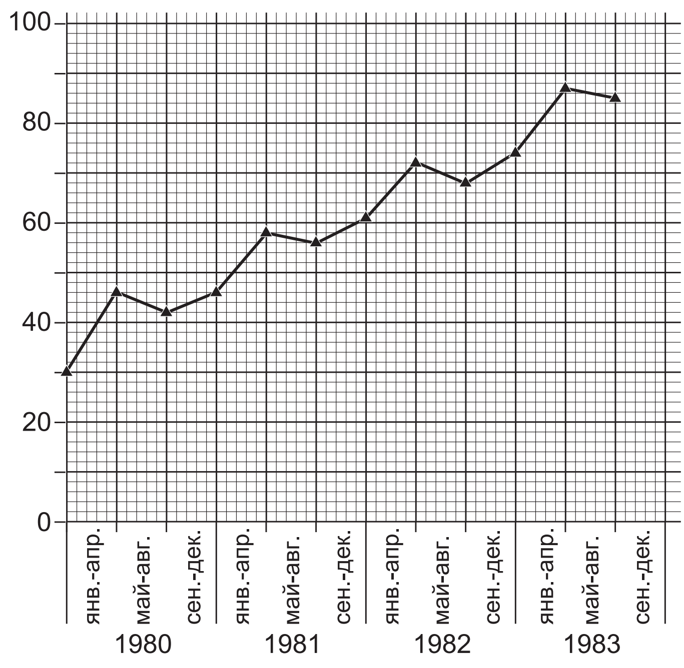
5.1. Понятие временны́х рядов

Временно́й ряд – это серия показателей, характеризующих изменение переменной во времени. Например, еженедельный объем продаж магазина или ежемесячный объем производства на заводе.

Пример 1. Объем продаж в магазине:

Год	январь – апрель	май – август	сентябрь – декабрь
1980	30	45	42
1981	45	58	56
1982	61	72	70
1983	74	87	86

Ниже изображен график, построенный по данным таблицы.



При анализе временны́х рядов изучают следующие его компоненты:

(i) тренд, или постоянная компонента

Тренд – это долговременное изменение, такое, как с 1980 по 1983 г.

(ii) сезонные колебания

Сезонные колебания состоят из подъемов и спадов, зависящих от времени года.

Пример:

	Подъем	Спад
Продажа мороженого	Лето (жаркая погода)	Зима (холодная погода)
Продажа зонтов	Зима (погода с осадками)	Лето (сухая погода)

(iii) остаточные, или случайные, колебания

Непредсказуемые колебания, которые не объясняются ни одной из перечисленных выше причин.

Примеры: смерть почитаемого монарха или президента; серьезная авария на производстве или землетрясение.

(iv) периодические (циклические) колебания

Колебания с периодом, большим или меньшим одного года.

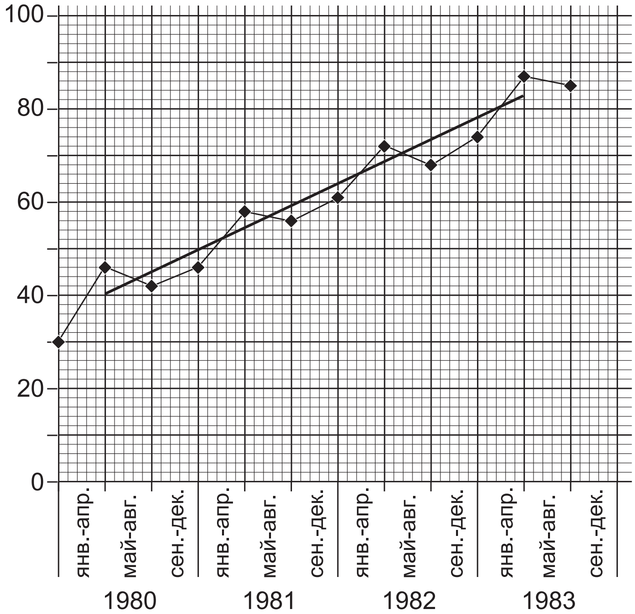
Примеры:
(a) долговременные колебания с периодом 5–6 лет между выборами;
(b) ежедневный объем продаж в магазине может меняться с периодом 5, 6 или 7 дней в зависимости от числа дней в неделю, когда магазин открыт.

5.2. Анализ временны́х рядов

Прежде всего мы выделяем тренд. Для этого применяется метод **скользящей** средней. В нашем примере имеется по **три значения** за каждый год, поэтому мы используем трехчленную скользящую среднюю. Это означает, что мы находим простую среднюю первых трех значений, затем простую среднюю следующих трех значений и т.д., как показано в таблице.

Годы	Периоды	Значение	Трех-членная скользящая сумма	Тренд (Трехчленная скользящая средняя) (с точностью до целого числа)
1980	январь – апрель	30		
	май – август	45	117	39
	сентябрь – декабрь	42	132	44
1981	январь – апрель	45	145	48
	май – август	58	159	53
	сентябрь – декабрь	56	175	58
1982	январь – апрель	61	189	63
	май – август	72	203	68
	сентябрь – декабрь	70	216	72
1983	январь – апрель	74	231	77
	май – август	87	247	82
	сентябрь – декабрь	86		

Ниже представлен график тренда, или скользящей средней, построенный по данным таблицы. Обратите внимание, что каждое значение трехчленной скользящей средней на графике соответствует центральному значению интервала сглаживания.



5.3. Выделение сезонных колебаний

После выделения методом скользящей средней тренда мы можем устранить или вычесть его из исходных данных, а затем перейти к выделению сезонных колебаний.

Годы	Периоды	Значение V	Тренд (Трехчленная скользящая средняя) T	$V - T$
1980	январь – апрель	30		
	май – август	45	39	+6
	сентябрь – декабрь	42	44	–2
1981	январь – апрель	45	48	–3
	май – август	58	53	+5
	сентябрь – декабрь	56	58	–2
1982	январь – апрель	61	63	–2
	май – август	72	68	+4
	сентябрь – декабрь	70	72	–2
1983	январь – апрель	74	77	–3
	май – август	87	82	+5
	сентябрь – декабрь	86		

Для определения (среднего значения) сезонных колебаний мы находим простую среднюю (арифметическую среднюю) величины (V–T) для одноименных сезонов. (Результат округляем до целого числа.)

Среднее значение сезонных колебаний за период январь–апрель равно сумме (– 3 – 2 – 3), деленной на 3, т.е. «–3».

Среднее значение сезонных колебаний за период май–август равно сумме (+6 +5 +4 +5), деленной на 4, т.е. «+5».

Среднее значение сезонных колебаний за период сентябрь–декабрь равно сумме (– 2 – 2 – 2), деленной на 3, т.е. «–2».

Сумма всех сезонных колебаний должна быть близка к нулю. В нашем примере сумма –3 +5 –2 в точности равна 0.

5.4. Выделение случайных, или остаточных, колебаний

Если мы вычтем средние сезонные колебания из значений в столбце *V–T*, то получим величину случайных, или остаточных, колебаний.

Годы	Периоды	Значение, *V*	Тренд, *T*	*V–T*	Сезонные колебания, *SV*	Остаточные колебания, *V – T – SV*
1980	январь – апрель	30				
	май – август	45	39	+6	+5	+1
	сентябрь – декабрь	42	44	–2	–2	0
1981	январь – апрель	45	48	–3	–3	0
	май – август	58	53	+5	+5	0
	сентябрь – декабрь	56	58	–2	–2	0
1982	январь – апрель	61	63	–2	–3	+1
	май – август	72	68	+4	+5	–1
	сентябрь – декабрь	70	72	–2	–2	0
1983	январь – апрель	74	77	–3	–3	0
	май – август	87	82	+5	+5	0
	сентябрь – декабрь	86				

Сумма случайных, или остаточных, колебаний должна быть близкой к нулю. В нашем примере сумма +1 +1 –1 равна +1.

Таким образом, значение большинства переменных складывается из трех компонентов:

Значение	=	Тренд (скользящая средняя)	+	Средние сезонные колебания	+	Случайные, или остаточные, колебания

Использование временны́х рядов для прогнозирования

Если тренд устойчивый и ярко выражен, то временной ряд можно использовать для осуществления прогнозов.

Прогнозное значение	=	Тренд (скользящая средняя)	+	Средние сезонные колебания

В нашем примере для получения оценки прогнозного значения на период январь–апрель 1984 г. продолжим линию тренда до соответствующего временного периода: январь–апрель 1984 г. Полученное значение равно 92.

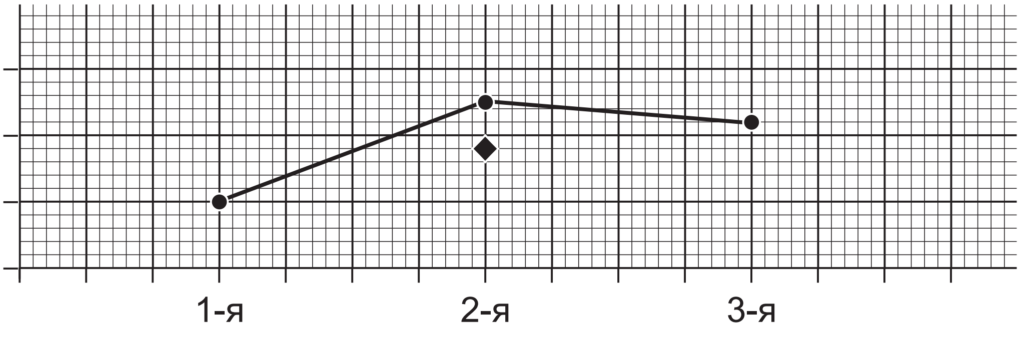
Для оценки прогнозного значения на январь–апрель 1984 г. добавим к тренду величину сезонных колебаний за январь–апрель, равную «–3».

Таким образом, оценка прогнозного значения на январь–апрель 1984 г. равна: 92 – 3 = 89.

5.5. Нечетное и четное число звеньев скользящей средней

5.5.1. Скользящая средняя с нечетным числом звеньев

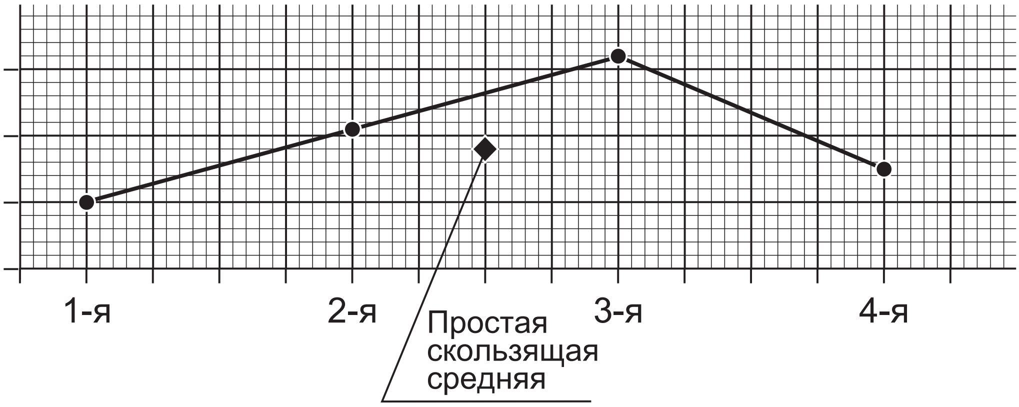
В рассмотренном выше примере имелось по 3 значения за год, поэтому мы применяли трехчленную скользящую среднюю. Мы суммировали последовательные группы из трех значений. В данном случае речь идет о скользящей средней с нечетным числом звеньев, поскольку 3 – нечетное число (не делится на 2). В этом случае мы суммируем последовательные группы из нечетного числа значений, находим скользящие средние и наносим их на график в моменты времени, соответствующие середине интервала сглаживания. Этот метод получил название простой скользящей средней.



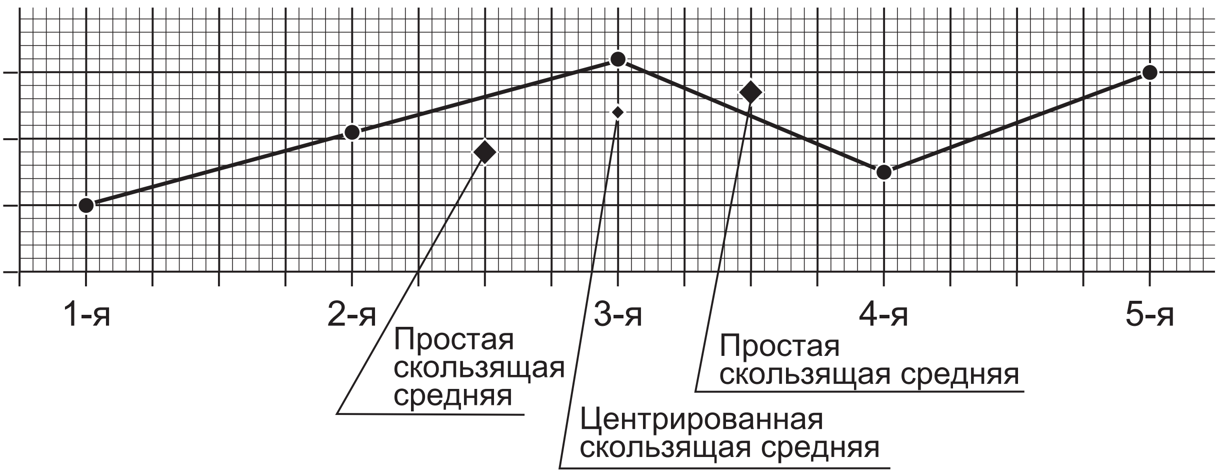
5.5.2. Скользящая средняя с четным числом звеньев

В случае со скользящей средней с четным числом звеньев имеется не одно, а два центральных значения. Далее можно действовать двумя способами.

(a) Рассчитываем скользящие средние, как в случае с нечетным числом звеньев, а затем наносим значения на график в момент времени посередине между двумя центральными значениями, как показано на графике.



(b) Рассчитываем центрированную скользящую среднюю. Для расчета центрированной скользящей средней суммируем значения двух соседних простых скользящих средних и делим на 2. Затем наносим на график значения центрированных скользящих средних в моменты времени посередине между двумя значениями простых скользящих средних, как показано ниже.



5.6. Вычисление центрированной скользящей средней (с четным числом звеньев)

(i) Расчет простой средней двух соседних скользящих средних

Пример. Найдите центрированную скользящую среднюю следующего временного ряда: 10, 11, 13, 16, 30, 31, 33, 36.

Решение.

Значение	Четырехчленная скользящая сумма	Четырехчленная скользящая средняя	Сумма двух соседних скользящих средних	Центрированная скользящая средняя
10				
11				
	50	12,5		
13			30	15
	70	17,5		
16			40	20
	90	22,5		
30			50	25
	110	27,5		
31			60	30
	130	32,5		
33				
36				

Обратите внимание, что

(i) Первые значения четырехчленной скользящей суммы и четырехчленной скользящей средней записаны посередине между 2-м и 3-м уровнями временного ряда, т.е. на строке 2½. Аналогично, вторые значения четырехчленной скользящей суммы и четырехчленной скользящей средней записываются между 3-м и 4-м уровнями на строке 3½.

(ii) Сумма двух соседних значений скользящей средней, расположенных на строках 2½ и 3½, записывается на 3-й строке.

(iii) Подобным образом первое центрированное значение четырехчленной скользящей средней соответствует моменту времени 3, второе центрированное значение соответствует на графике моменту времени 4 и т.д.

(ii) Сложение двух соседних скользящих сумм

Значение	Четырехчленная скользящая сумма	Сумма двух соседних скользящих средних	Центрированная скользящая средняя, деленная на 8
10			
11			
	50		
13		120	15
	70		
16		160	20
	90		
30		200	25
	110		
31		240	30
	130		
33			
36			

Обратите внимание, что

(i) Первое значение четырехчленной скользящей суммы записано посередине между 2-м и 3-м уровнями временного ряда, т.е. на строке 2½. Аналогично, второе значение четырехчленной скользящей средней записывается на строке 3½.

(ii) Сумма двух соседних значений скользящей средней, расположенных на строках 2½ и 3½, записывается на 3-й строке. Аналогично центрированное значение четырехчленной скользящей средней также записывается на 3-й строке.

(iii) Подобным образом первое центрированное значение четырехчленной скользящей средней соответствует моменту времени 3, второе центрированное значение соответствует на графике моменту времени 4 и т.д.

5.7. Выбор числа звеньев скользящей средней

Для некоторых временных рядов число звеньев скользящей средней определяется легко.

Например:

(1) квартальный отчет или потребление электричества или газа;

4 значения за каждый год, следовательно, применяем четырехчленную скользящую среднюю.

(2) ежемесячные объемы продаж или производства;

12 значений за каждый год, следовательно, применяем двенадцатичленную скользящую среднюю.

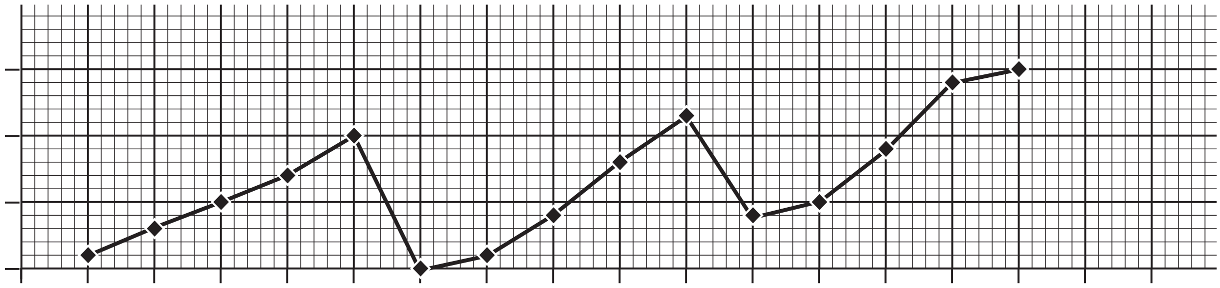
(3) семестровые расходы в школе (3 семестра в год);

3 значения за каждый год, следовательно, применяем трехчленную скользящую среднюю.

(4) ежедневный объем продаж в магазине, работающем 5 дней в неделю; 5 значений за каждую неделю, следовательно, применяем пятичленную скользящую среднюю.

Однако мы можем располагать только рядом значений без указания периода. В таком случае лучше всего начать с построения графика или хотя бы грубого наброска временного ряда. Возможно, это поможет определить период повторяющихся колебаний.

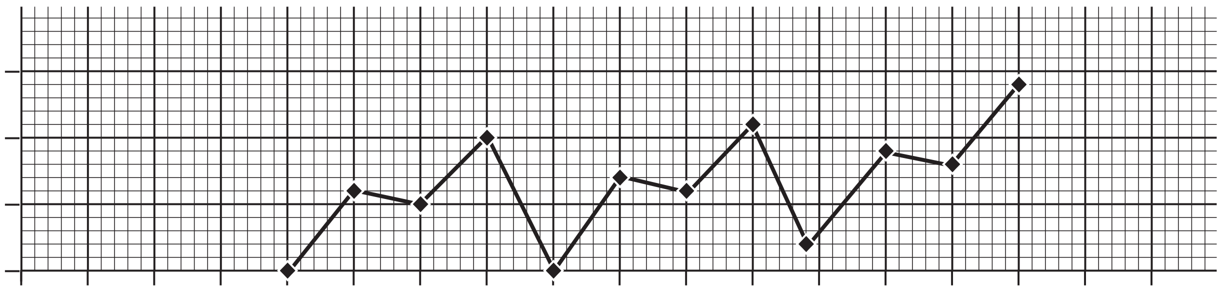
Пример 1.



По графику видно, что подъемы имеют место в каждом 5-м временном периоде, поэтому надо применять пятичленную скользящую среднюю.

Пример 2.

Подъемы имеют место в каждом 4-м временном периоде, поэтому надо применять четырехчленную скользящую среднюю.



Аудиторный тест № 17

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

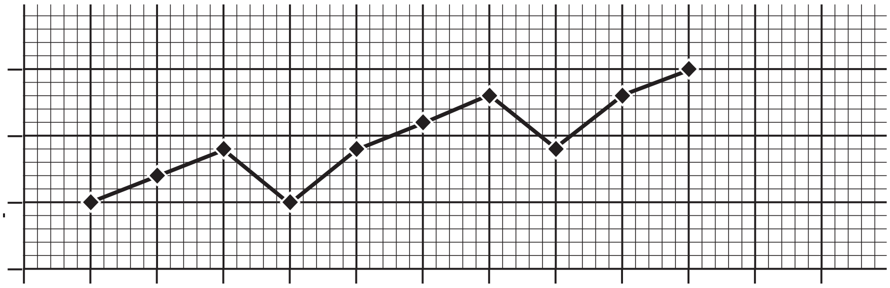
Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Выберите подходящее число звеньев скользящей средней для каждого временного ряда, приведенного ниже.

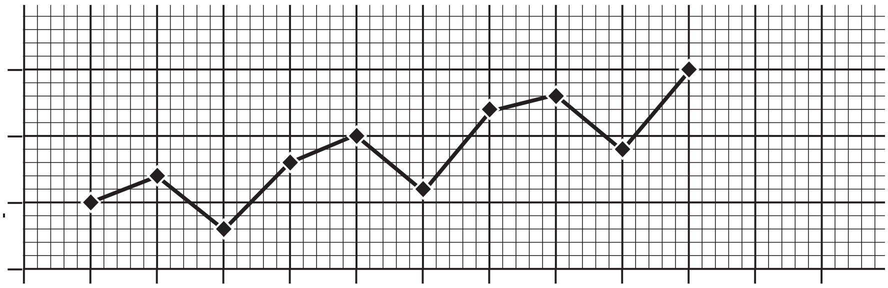
Ответ A означает трехчленную скользящую среднюю, B – четырехчленную, C – пятичленную, D – шестичленную скользящую среднюю и E – семичленную скользящую среднюю.

Вопрос.

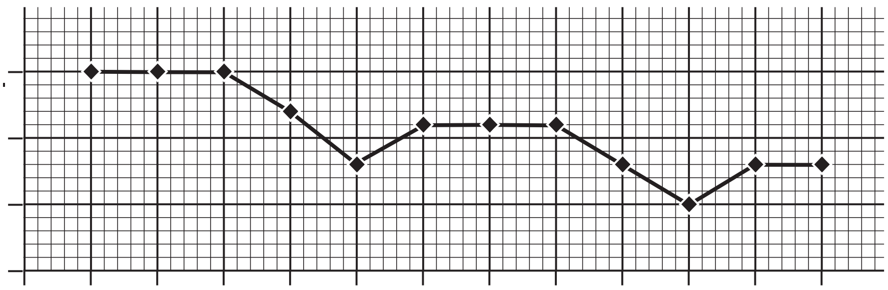
1.



2.



3.



4. 7, 10, 12, 7, 7, 8, 13, 16, 18, 13, 13, 14
5. 19, 14, 18, 16, 11, 15, 13, 8, 12, 10, 5, 9
6. 7, 10, 7, 9, 11, 15, 11, 13, 15, 19, 15, 17
7. 18, 17, 16, 15, 16, 11, 12, 25, 24, 23, 22, 21
8. 21, 21, 22, 6, 17, 17, 18, 2, 13, 13, 14
9. 11, 11, 5, 14, 14, 8, 17, 17, 11, 20, 20
10. 8, 13, 17, 10, 12, 18, 23, 27, 20, 22, 28

Упражнение 24А

1. В таблице приведены данные за 1984, 1985 и 1986 гг. по объему воды, поступавшей в резервуар.

Годы	Объем воды (1000 м^3)		
	январь – апрель	май – август	сентябрь – декабрь
1984	5,6	5,6	6,5
1985	6,8	6,8	7,7
1986	8,0	8,0	8,9

(i) Нанесите данные на график в следующем масштабе: по оси X 2 см $\Rightarrow$ 4 месяца; по оси Y 2 см $\Rightarrow$ 1000 м^3.
(ii) Выделите тренд, рассчитав значения трехчленной скользящей средней.
(iii) Постройте график тренда (скользящей средней).
(iv) Вычислите средние сезонные колебания для периодов январь – апрель, май – август, сентябрь – декабрь.
(v) Оцените объем воды, поступившей в резервуар в январе – апреле 1987 г.

2. В таблице приведены значения среднего еженедельного объема продаж в магазине за период с 1984 по 1986 г. Все суммы указаны в долларах.

Годы	январь – март	апрель – июнь	июль – сентябрь	октябрь – декабрь
1984	735	705	720	720
1985	655	625	640	640
1986	575	545	560	560

(i) Нанесите данные на график в следующем масштабе: по оси X 1см $\Rightarrow$ 3 месяца; по оси Y 1см $\Rightarrow$ 50 долл.
(ii) Рассчитайте соответствующие центрированные значения скользящей средней для выделения тренда.
(iii) Постройте график тренда.
(iv) Вычислите средние сезонные колебания для каждого квартала.
(v) Оцените еженедельный объем продаж в магазине в январе – марте 1987 г.

Упражнение 24В

1. В таблице приведены данные о количестве сельскохозяйственных машин, произведенных на заводе в период с 1984 по 1986 г.

Годы	январь – апрель	май – август	сентябрь – декабрь
1984	145	189	206
1985	245	279	296
1986	325	369	386

(i) Нанесите данные на график в следующем масштабе: по оси X 1 см $\Rightarrow$ 4 месяца; по оси Y 1 см $\Rightarrow$ 20 машин.

(ii) Рассчитайте значения трехчленной скользящей средней для выделения тренда. Значения скользящей средней округляйте до целого числа.

(iii) На имеющемся графике исходных данных нанесите значения скользящей средней, а затем на нем же постройте график тренда.

(iv) Вычислите средние сезонные колебания для периодов январь – апрель, май – август, сентябрь – декабрь, округляя полученные значения до целого числа.

(v) Оцените количество машин, произведенных в каждом периоде 1987 г.

2. В таблице приведены данные о количестве жилых домов, построенных за период с 1984 по 1986 г.

Годы	январь – март	апрель – июнь	июль – сентябрь	октябрь – декабрь
1984	157	163	168	172
1985	197	203	208	212
1986	237	243	248	252

(i) Нанесите данные на график в следующем масштабе: по оси X 1 см $\Rightarrow$ 3 месяца; по оси Y 1 см $\Rightarrow$ 20 жилых домов.

(ii) Рассчитайте соответствующие центрированные значения скользящей средней для выделения тренда.

(iii) Постройте график тренда.

(iv) Вычислите средние сезонные колебания для каждого квартала.

(v) Оцените количество домов, построенных за первые три месяца 1987 г.

Глава 6

СРЕДНИЕ ВЗВЕШЕННЫЕ

6.1. Понятие средней взвешенной

Предположим, что на одном экзамене средний балл (средняя арифметическая) успеваемости в классе 5А равен 36, а средний балл успеваемости в классе 5В равен 42. Было бы неверно утверждать, что средний балл успеваемости в двух классах равен (36 + 42) / 2, т.е. 39. Такой расчет не учитывает тот факт, что в классах может быть разное число учеников. Для нахождения среднего балла для двух классов мы должны использовать среднюю взвешенную, при расчете которой средний балл каждого класса «взвешивается» с учетом числа учеников в классе.

Для определения среднего балла мы действуем следующим образом:

Класс	Число учеников	Средний балл	Общий балл
5A	25	36	25×36 = 900
5B	35	42	35×42 = 1470
Оба класса	60		900 + 1470 = 2370

Средний балл успеваемости в двух классах вместе рассчитывается как общий балл, деленный на суммарное число учеников.

Средний балл успеваемости в двух классах равен 2370 : 60 = 39,5.

Обратите внимание, что правильно рассчитанный средний балл несколько ближе к среднему баллу успеваемости в классе с бо́льшим числом учеников, чем неверно рассчитанное значение среднего балла.

6.2. Другие повседневные примеры средней взвешенной

Пример 1. Средний балл успеваемости в трех классах по результатам одного экзамена.

Класс	Число учеников	Средний балл	Общий балл
3A	24	45	1080
3B	36	42	1512
3C	32	50	1600
Три класса	92		4192

Средний балл успеваемости в трех классах равен 4192 : 92 = 45,6.

Пример 2. Средняя скорость во время двухэтапного путешествия.

Способ передвижения	Время	Средняя скорость	Пройденный путь
Поезд	0,5 ч	50 км/ч	0,5×50 = 25 км
Ходьба	0,2 ч	5 км/ч	0,2× 5 = 1 км
Вместе	0,7 ч		25 + 1 = 26 км

Средняя скорость во время всего путешествия равна 26 : 0,7 = = 31,7 км/ч.

Пример 3. Плотность сплава из двух металлов.

Металл	Объем	Плотность	Масса
Медь	2 см^3	8,92 г/см^3	17,84 г
Цинк	1 см^3	7,14 г/см^3	7,14 г
Сплав	3 см^3		24,98 г

Плотность сплава равна 24,98 : 3 = 8,33 г/см^3.

Пример 4. Цена смеси двух сортов чая.

Чай	Количество	Цена	Стоимость
«Небесный» «Божественный»	3 кг 7 кг	25 ф.ст. за 1 кг 15 ф.ст. за 1 кг	75 ф.ст. 105 ф.ст.
«Смешанный»	10 кг		180 ф.ст.

Цена смеси двух сортов чая равна 180 : 10, что составляет 18 ф.ст. за 1 кг.

Пример 5. На экзамене средний балл успеваемости 30 учеников класса 4А составил 40. Средний балл успеваемости 50 учеников классов 4А и 4В вместе равен 44. Вычислите средний балл успеваемости 20 учеников класса 4В.

Класс	Число учеников	Средний балл	Общий балл
4А и 4В	50	44	2200
Только 4А	30	40	1200
Вычитаем:			
Только 4В	20		1000

Следовательно, средний балл успеваемости учеников 4В класса равен 1000 : 20 = 50.

Упражнение 25А

1. Путешествие мальчика состояло из часовой поездки на поезде со скоростью 40 км/ч и 30 мин ходьбы со скоростью 6 км/ч. Найдите среднюю скорость за все путешествие.

2. Средний балл успеваемости 30 учеников 2А класса равен 56. Средний балл успеваемости 40 учеников 2В класса равен 52. Вычислите средний балл успеваемости всех 70 учеников. Ответ округлите до целого числа.

3. Для изготовления сплава было взято 3 см^3 меди и 2 см^3 цинка. Известно, что плотность меди и цинка равна 8,9 г/см^3 и 7,1 г/см^3 соответственно. Вычислите плотность сплава.

4. Для приготовления смеси чая взяли 4 кг чая «Республика» по цене 20 ф.ст. за 1 кг и 6 кг чая «Содружество». Цена смеси равна 18,80 ф.ст. за 1 кг. Найдите цену за 1 кг чая «Содружество».

Упражнение 25B

1. Девочка ехала 30 мин на автобусе со скоростью 35 км/ч, а затем шла пешком 1 ч со скоростью 6 км/ч. Найдите среднюю скорость девочки за все путешествие.

2. Средний балл успеваемости 20 учеников 4A класса равен 48. Средний балл успеваемости 30 учеников 4B класса равен 62. Вычислите средний балл успеваемости всех 50 учеников. Ответ округлите до целого числа.

3. Для изготовления сплава было взято 4 см^3 меди и 3 см^3 цинка. Известно, что плотность меди и цинка равна 8,92 г/см^3 и 7,14 г/см^3 соответственно. Вычислите плотность сплава.

4. Для приготовления смеси чая взяли 3 кг чая «Роскошный» по цене 20 ф.ст. за 1 кг и 2 кг чая «Любимый». Цена смеси равна 22 ф.ст. за 1 кг. Найдите цену за 1 кг чая «Любимый».

6.3. Средняя геометрическая взвешенная

В параграфе 2.7.3 мы находили среднюю геометрическую более чем двух чисел. Например, средняя геометрическая чисел 2, 3, 5, 7 и 10 равна корню 5-й степени из произведения $2 \times 3 \times 5 \times 7 \times 10 = 2100$, что составляет 4,62.

Иногда бывает, что одно или более чисел повторяются, например, 4, 4, 4, 5, 5, 6. В этом случае мы можем найти среднюю геометрическую двумя способами:

(i) Игнорируя повторение значений и вычисляя корень 6-й степени из произведения $4 \times 4 \times 4 \times 5 \times 5 \times 6$. Корень 6-й степени из 9600 равен 4,61.

(ii) Вычисляя корень 6-й степени из произведения $4^3 \times 5^2 \times 6$, что опять дает корень 6-й степени из 9600. Снова получаем 4,61.

Вычисление средней геометрической вторым способом и является определением средней геометрической взвешенной, поскольку числа 4 и 5 взяты с весом 3 и 2 соответственно.

Пример 1. Двое рабочих получили по 20 долл., а пять других – по 15 долл. Найдите среднюю геометрическую взвешенную заработка всех 7 рабочих.

Решение.

Средняя геометрическая взвешенная равна корню 7-й степени из произведения $20^2 \times 15^5$, т.е. корню 7-й степени из 303 750 000, что составляет 16,285 долл.

Проверка. Произведение $20 \times 20 \times 15 \times 15 \times 15 \times 15 \times 15$ тоже дает 303 750 000.

6.4. Индекс цен

Индекс цен товара в текущем году по сравнению с другим годом (обычно называемым **базисным**) определяется формулой

$$\text{Индекс цен} = \frac{\text{Цена товара в текущем году}}{\text{Цена товара в базисном году}} \times 100.$$

Пример 1. В 1978 г. ракетка для игры в бадминтон стоила 8 ф.ст. Цена такой же ракетки в 1984 г. составила 15 ф.ст.

Вычислите индекс цен (i) в 1984 г. по сравнению с 1978 г.;
(ii) в 1978 г. по сравнению с 1984 г.

Решение.

(i) Индекс цен в 1984 г. по сравнению с 1978 г. равен:

$$\frac{15 \text{ ф.ст.}}{8 \text{ ф.ст.}} \times 100 = 187{,}5.$$

(ii) Индекс цен в 1978 г. по сравнению с 1984 г. равен:

$$\frac{8 \text{ ф.ст.}}{15 \text{ ф.ст.}} \times 100 = 53{,}3.$$

Примечание. Хотя индекс цен выражается в процентах, символ «%» не указывается.

Пример 2. В 1983 г. цена обуви составляла 25 ф.ст. за пару. В 1984 г. цена такой же пары была уже 28 ф.ст.

Вычислите индекс цен (i) в 1984 г. по сравнению с 1983 г.;
(ii) в 1983 г. по сравнению с 1984 г.

Решение.

(i) Индекс цен в 1984 г. по сравнению с 1983 г. равен:

$$\frac{28 \text{ ф.ст.}}{25 \text{ ф.ст.}} \times 100 = 112.$$

(ii) Индекс цен в 1983 г. по сравнению с 1984 г. равен:

$$\frac{25 \text{ ф.ст.}}{28 \text{ ф.ст.}} \times 100 = 89{,}3.$$

Аудиторный тест № 18

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вопрос **Варианты ответа**

Вычислите индекс цен для второй цены по сравнению с первой ($ означает доллары США, £ – фунты стерлингов).

	Первая цена	Вторая цена	A	B	C	D	E
1.	$100	$125	0,8	80	1,25	12,5	125
2.	$200	$250	0,8	80	1,25	12,5	125
3.	£500	£700	1,4	7,0	140	200	700
4.	$250	$300	1,2	1,25	120	125	130
5.	$900	$450	0,5	2,0	50	200	450
6.	$48	$60	0,6	1,2	60	112	125
7.	£2,50	£2,75	0,25	1,1	110	257	550
8.	$36	$27	0,75	1,33	27	75	133
9.	$1,20	$1,50	0,25	0,8	1,25	80	125
10.	$3,00	$2,40	0,25	0,8	1,25	80	125

Вычислите индекс цен для первой цены по сравнению со второй.

	Первая цена	Вторая цена	A	B	C	D	E
1.	$5	$4	0,8	1,25	40	80	125
2.	$10	$16	0,6	1,6	60	62,5	160
3.	$8	$12	50	60	63	65	67
4.	£32	£24	75	80	120	125	133
5.	$48	$40	75	80	120	125	160
6.	$50	$60	77	80	83	87	90
7.	£50	£80	60	62,5	65	67,5	70
8.	$8	$4	40	50	80	200	400

6.5. Метод цепных индексов

В методе цепных индексов цена товара в последующем году выражается в ценах предыдущего года.

Например, цена книги составляла $4 в 1984 г., $5 в 1985 г. и $6 в 1986 г.

По методу цепных индексов:

индекс цен в 1985 г. (по сравнению с 1984 г.) равен $5 / $4 × × 100 = 125;

индекс цен в 1986 г. (по сравнению с 1985 г.) равен $6 / $5 × × 100 = 120.

Простой индекс цен в 1986 г. по сравнению с 1984 г. может быть вычислен двумя способами: умножением 100 × 1,25 × 1,20 или $6 / $4 × 100 = 150.

Пример 1.

По методу цепных индексов индекс цен для некоторого товара на 1 июля 1985 г., 1986 г. и 1987 г. составил соответственно 100, 80 и 75. Вычислите простой индекс цен на 1 июля 1987 г., используя в качестве базисного периода 1 июля 1985 г.

Решение.

Индекс цен для 1987 г. равен 100 × 0,80 × 0,75 = 60.

Пример 2.

Если принять 1985 г. за базисный год, индексы цен для товара в 1986 г. и 1987 г. окажутся равными соответственно 115 и 92. Представьте эту информацию, используя метод цепных индексов.

Решение.

Индекс цен в 1986 г. (по сравнению с 1985 г.) равен 115.

Индекс цен в 1987 г. (по сравнению с 1986 г.) равен:

$$100 \times \frac{92}{115} = 80.$$

Пример 3.

По методу цепных индексов относительная цена книги в 1985 г., 1986 г. и 1987 г. была равна соответственно 100, 95 и 80. В 1985 г. действительная цена книги составляла \$8. Вычислите цену книги в 1986 г. и 1987 г.

Решение.

Цена в 1986 г. равна $\$8 \times \frac{95}{100} = \$7,60.$

Цена в 1987 г. равна $\$7,60 \times \frac{80}{100} = \$6,08.$

Упражнение 26А

1. Цена учебника по статистике составляла \$8 в 1985 г. и \$10 в 1986 г.

Вычислите:

(i) индекс цен для учебника в 1986 г., взяв 1985 г. за базисный год;

(ii) индекс цен для учебника в 1985 г., взяв 1986 г. за базисный год.

2. Цена предмета одежды составляла £5 в 1985 г., £6 в 1986 г. и £5,40 в 1987 г. Вычислите:

(i) простой индекс цен для 1986 г. и 1987 г., взяв 1985 г. за базисный год;
(ii) методом цепных индексов индекс цен для 1986 г. и 1987 г.

3. Цепные индексы цен пары обуви в 1985 г., 1986 г. и 1987 г. были равны соответственно 100, 80 и 110. Известно, что в 1985 г. такая пара обуви стоила $12. Вычислите:

(i) цену пары обуви в 1986 г. и 1987 г.;
(ii) индекс цен для 1987 г., взяв 1985 г. за базисный год.

Упражнение 26В

1. Цена учебника по математике составляла $7,50 в 1986 г. и $9,00 в 1987 г. Вычислите:

(i) индекс цен учебника в 1987 г., взяв 1986 г. за базисный год;
(ii) индекс цен учебника в 1986 г., взяв 1987 г. за базисный год.

2. Цена кухонной посуды составляла £3 в 1985 г., £3,60 в 1986 г. и £5,40 в 1987 г. Вычислите:

(i) простой индекс цен для 1986 г. и 1987 г., взяв 1985 г. за базисный год;
(ii) методом цепных индексов индекс цен для 1986 г. и 1987 г.

3. Цепные индексы цен калькулятора в 1985 г., 1986 г. и 1987 г. были равны соответственно 100, 65 и 50. Известно, что в 1985 г. такой калькулятор стоил $12. Вычислите:

(i) цену калькулятора в 1986 г. и 1987 г.;
(ii) индекс цен для 1985 г., взяв 1987 г. за базисный год.

Результат округлите до целого числа.

6.6. Индекс прожиточного минимума

Индекс прожиточного минимума является взвешенной средней различных индексов цен рассматриваемых предметов. Каждый индекс цен взвешивается по доле доходов, которая тратится на этот предмет.

Пример 1. Прожиточный минимум молодежи в 1987 г. по сравнению с 1986 г.

Наименование расходов	Доля (%)	Индекс цен в 1987 г. по сравнению с 1986 г.	Доля × Индекс цен
Проезд	25	140	25×140 = 3500
Одежда	45	120	45×120 = 5400
Книги	30	115	30×115 = 3450
Суммарные расходы	100		12 350

Индекс прожиточного минимума равен 12 350 : 100 = 123,5.

Пример 2. Стоимость набора продовольственных товаров.

Наименование расходов	Доля	Индекс цен в 1987 г. по сравнению с 1986 г.	Доля × Индекс цен
Мясо	5	112	5×112 = 560
Молочные продукты	10	105	10×105 = 1050
Овощи	25	108	25×108 = 2700
Суммарные расходы	40		4310

Индекс прожиточного минимума равен 4310 : 40 = 107,75.

Если нам известно действительные количество и цены каждого предмета, мы можем получить значение индекса, используя действительные цены.

Пример 3.

Ежегодно строитель использует 80 т мелкого песка, 150 т крупнозернистого песка и 250 т гравия. В 1985 г. цена этих строительных материалов составляла £6,00, £6,50 и £8,00 за тонну соответственно. В 1986 г. цена поднялась до £7,00, £7,20 и £9,00 за тонну соответственно.

Вычислите индекс затрат на строительство в 1986 г. по сравнению с 1985 г.

Решение.

Материал	1985 г.			1986 г.		
	Количество (т)	Цена за 1 т	Стоимость	Количество (т)	Цена за 1 т	Стоимость
Мелкий песок	80	6,00	48,00	80	7,00	56,00
Крупнозернистый песок	150	6,50	975,00	150	7,20	1080,00
Гравий	250	8,00	2000,00	250	9,00	2250,00
Все материалы			3023,00			3386,00

Таким образом, индекс $=100\times\frac{3386}{3023}=112,0.$

Альтернативный индекс может быть вычислен по индексам цен.

Материал	Цены в 1985 г. (за 1 т)	Цены в 1986 г. (за 1 т)	Индекс цен 1987 г. по сравнению с 1986 г.	Доля	Доля × × Индекс цен
Мелкий песок	6,00	7,00	100×7,00/6,00 = = 116,7	80	9 336
Крупнозернистый песок	6,50	7,20	100×7,20/6,50 = = 110,8	150	16 620
Гравий	8,00	9,00	100×9,00/8,00 = = 112,5	250	28 125
Все материалы				480	54 081

Таким образом, индекс $=100\times\frac{54\,081}{480}=112,7.$

Второй индекс известен как индекс цен Ласпейреса.

6.7. Процентные изменения

Обратите внимание на следующие процентные изменения:

(i) рост на 5% означает, что индекс цен второго года по сравнению с первым равен 105;
(ii) увеличение цены в два раза означает, что индекс цен второго года по сравнению с первым равен 200;
(iii) уменьшение цены в два раза означает, что индекс цен второго года по сравнению с первым равен 50;
(iv) уменьшение цены на 8% означает, что индекс цен второго года по сравнению с первым равен 92.

Упражнение 27А

1. Вычислите сводный индекс по следующим данным:

Наименование	Доля	Индекс
Продукты питания	50	107
Одежда	30	112
Дорога	20	90

2. За период с 1986 по 1987 г. стоимость продуктов питания снизилась на 5%, стоимость проезда снизилась на 10%, стоимость одежды снизилась на 15%. Взяв долю продуктов питания равной 20, проезда 13 и одежды 7, рассчитайте с точностью до одного знака после запятой:

(i) сводный индекс в 1987 г. по сравнению с 1986 г.;
(ii) сводный индекс в 1986 г. по сравнению с 1987 г.

3.

Наименование	A	B	C
Индекс	110	106	X
Доля	4	3	2

Известно, что сводный индекс равен 108. Найдите значение X.

Упражнение 27В

1. Вычислите сводный индекс по следующим данным:

Наименование	Доля	Индекс
Одежда	35	109
Дорога	45	105
Продукты питания	20	96

2. За период с 1983 по 1984 г. стоимость проезда увеличилась на 8%, стоимость продуктов питания возросла на 4%, а стоимость одежды снизилась на 6%. Взяв долю проезда равной 18, продуктов питания 13 и одежды 9, рассчитайте с точностью до одного знака после запятой:

(i) сводный индекс в 1983 г. по сравнению с 1984 г.;

(ii) сводный индекс в 1984 г. по сравнению с 1983 г.

3.

Наименование	A	B	C
Индекс	120	80	X
Доля	5	3	2

Известно, что сводный индекс равен 104. Найдите значение X.

6.8. Коэффициенты смертности, рождаемости и брачности

6.8.1. Общий коэффициент смертности

Общий коэффициент смертности в городе или на территории рассчитывается как число смертей, зарегистрированных в данном городе или на территории, на 1000 человек населения. Знак ‰ означает, что показатель рассчитывается на 1000 человек (проми́лле).

Пример 1. Население города составляет 23 500 человек. В 1984 г. число смертей в городе было равно 223. Рассчитайте общий коэффициент смертности.

Решение. Общий коэффициент смертности $=\frac{223}{23\,500}\times 1000 = 9{,}49$ на 1000 жителей.

Пример 2. Население острова равно 8200 человек. В 1982 г. общий коэффициент смертности на острове составлял 8,78‰ . Определите число смертей на острове в 1982 г.

Решение. Число смертей = 8200 × 8,78 : 1000 = 71,996.

Так как количество смертей может быть только целым числом, то округляем результат до 72.

Пример 3. В 1981 г. количество смертей на некоторой территории было равно 29, а общий коэффициент смертности равен 7,84‰. Определите численность населения этой территории с точностью до сотен.

Решение. Численность населения $= \frac{29}{7,84} \times 1000 = 3700$ человек.

6.8.2. Сравнение коэффициентов смертности двух территорий

Показатель смертности используется для сравнения риска или вероятности умереть, проживая на той или иной территории.

К сожалению, общий коэффициент смертности зависит от двух факторов:

(i) безопасности или экологичности окружающей среды города или территории;

(ii) возраста населения, проживающего в городе или на территории.

Вероятность умереть изменяется при переходе от одной возрастной группы населения к другой. Например, пусть в некоторой стране территория A является промышленной. Как результат наличия производства, атмосфера загрязнена, экология не очень хорошая, но проживающее там население относительно молодое. Территория B является уютным жилым районом с благоприятной экологией, но проживают там в основном пожилые люди, пенсионеры.

Общий коэффициент смертности обеих территорий может быть приблизительно равным. Когда мы сравниваем две территории, нас интересует только фактор (i). Следовательно, нам надо устранить влияние фактора (ii), сравнивая **стандартизованные коэффициенты смертности** вместо общих коэффициентов смертности.

Можно сказать, что приближенно:

$$\text{Общий коэффициент смертности} = \\ = (\text{Возраст населения}) \times (\text{Экологичность территории});$$

$$\text{Стандартизованный коэффициент смертности} = \\ = (\text{Стандартное население}) \times (\text{Экологичность территории}).$$

Если мы возьмем одинаковое стандартное население для обеих рассматриваемых территорий, то сравнение двух стандартизованных коэффициентов смертности позволит нам сравнить экологичность этих территорий.

Город или территория с пожилым населением, вероятнее всего, будет иметь высокий общий коэффициент смертности. Чем благополучнее экологическая обстановка и чем выше шанс выживания, тем ниже стандартизованный коэффициент смертности.

6.8.3. Стандартизованный коэффициент смертности

Стандартизованный коэффициент смертности в городе или на территории рассчитывается как число зарегистрированных смертей на 1000 человек, если бы население города или территории было «стандартным».

Для вычисления стандартизованного коэффициента смертности в городе или на территории нам необходимо знать:

(i) анализ или распределение населения города или территории по возрастным группам;
(ii) число смертей в каждой возрастной группе или общий коэффициент смертности для каждой возрастной группы (этот показатель удобно называть повозрастным коэффициентом смертности);
(iii) анализ или распределение по возрастным группам стандартного населения.

Для определения стандартизованного показателя смертности мы рассчитываем взвешенные средние для каждой группы, используя повозрастной коэффициент смертности для каждой возрастной группы, взвешенный в соответствии с каждой возрастной группой стандартного населения.

Пример 1. Вычислим:
(i) общий коэффициент смертности,
(ii) стандартизованный коэффициент смертности для территории А, используя данные таблицы.

Возрастная группа	Территория А		Стандартное население
	Население	Число смертей	
0-	2000	10	25%
20-	1000	8	30%
40-	3000	12	20%
60-	1000	15	25%

Решение.

(i) Общий коэффициент смертности =

$$=\frac{\text{Суммарное число смертей}}{\text{Суммарное число жителей}}\times 1000=\frac{45}{7000}=6,43 \text{ на тысячу жителей.}$$

(ii) Сперва рассчитаем повозрастной коэффициент смертности для каждой возрастной группы.

Возрастная группа	Население	Число смертей	Повозрастной коэффициент смертности
0-	2000	10	10 / 2 = 5
20-	1000	8	8 / 1 = 8
40-	3000	12	12 / 3 = 4
60-	1000	15	15 / 1 = 15

Затем умножим каждый повозрастной коэффициент смертности на долю соответствующей группы стандартного населения.

Возрастная группа	Повозрастной коэффициент смертности	Стандартное население	Повозрастной коэффициент смертности × × Стандартное население
0-	5	25%	125
20-	8	30%	240
40-	4	20%	80
60-	15	25%	375
Все группы		100%	820

Таким образом, стандартизованный коэффициент смертности = 820 : 100 = 8,2 на тысячу жителей.

Пример 2. Вычислим

(i) общий коэффициент смертности,

(ii) стандартизованный коэффициент смертности для территории B, используя данные таблицы.

Возрастная группа	Территория B		Стандартное население
	Население	Число смертей	
0-	7000	21	1000
20-	8000	56	2000
40-	5000	45	3000
60-	3000	36	1500

Решение.

(i) Общий коэффициент смертности равен

$$\frac{\text{Суммарное число смертей}}{\text{Суммарное число жителей}} \times 1000 = \frac{158}{23\ 000} = 6,87 \text{ на тысячу жителей.}$$

(ii) Сперва рассчитаем повозрастной коэффициент смертности для каждой возрастной группы.

Возрастная группа	Население	Число смертей	Повозрастной коэффициент смертности
0-	7000	21	21 / 7 = 3
20-	8000	56	56 / 8 = 7
40-	5000	45	45 / 5 = 9
60-	3000	36	36 / 3 = 12

Затем умножим каждый повозрастной коэффициент смертности на долю соответствующей группы стандартного населения.

Возрастная группа	Повозрастной коэффициент смертности	Стандартное население	Повозрастной коэффициент смертности × Стандартное население
0-	3	1000	3 000
20-	7	2000	14 000
40-	9	3000	27 000
60-	12	1500	18 000
Все группы		7500	62 000

Таким образом, стандартизованный коэффициент смертности = 62 000 : 7500 = 8,27 на тысячу жителей.

Аудиторный тест № 19

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вопрос			**Варианты ответа**				
			A	B	C	D	E

Вычислите общий или повозрастной коэффициент смертности по имеющимся данным.

	Население	Число смертей	A	B	C	D	E
1.	1000	5	0,5	5	20	50	200
2.	2000	16	0,8	1,25	8	12,5	32
3.	5000	20	0,4	1,0	2,5	4	25
4.	6000	30	0,5	2	5	18	20
5.	4500	9	0,2	2	4,05	5	20

По имеющимся данным рассчитайте число смертей.

	Население	Коэффициент смертности	A	B	C	D	E
6.	3000	12‰	25	36	40	250	360
7.	7000	7‰	5	10	49	100	490
8.	4000	8‰	2	5	20	32	50
9.	8000	4‰	2	5	20	32	50
10.	9000	3‰	3	27	30	33	270

Вопрос			**Варианты ответа**				
			A	B	C	D	E

Рассчитайте численность населения.

	Число смертей	Коэффициент смертности					
11.	20	5‰	250	400	1000	2500	4000
12.	15	3‰	2000	3000	4000	4500	5000
13.	18	3‰	540	600	1700	5400	6000
14.	60	7,5‰	1250	4500	6000	7500	8000
15.	14	7‰	200	500	2000	5000	9800

Найдите значение Х.

	Население	Число смертей	Коэффициент смертности					
16.	5000	35	Х	7‰	14‰	18‰	35‰	70‰
17.	6000	Х	4‰	15	24	36	48	67
18.	Х	32	8‰	2000	2500	2560	4000	5000
19.	7000	35	Х	2‰	2,45‰	5‰	20‰	50‰
20.	8000	Х	6‰	13	48	75	133	480

Упражнение 28А

1. В таблице представлены данные по двум соседним территориям D и E.

Возрастная группа	Территория D		Территория E		Стандартное население, %
	Численность населения	Число смертей	Процент населения	Коэффициент смертности на 1000 жителей	
0-	1500	12	35	7	35
20-	2500	15	20	5	40
40-	1000	10	45	9	25

Вычислите:
(i) общий коэффициент смертности для территории D;
(ii) коэффициент смертности каждой возрастной группы для территории D;
(iii) стандартизованный коэффициент смертности для территории D;
(iv) общий коэффициент смертности для территории E;
(v) коэффициент смертности каждой возрастной группы для территории E;
(vi) стандартизованный коэффициент смертности для территории E.
Сделайте вывод, где условия для жизни лучше?

2. В таблице представлены данные по двум соседним регионам R и S.

Возрастная группа	Регион R		Регион S		Стандартное население, %
	Численность населения	Коэффициент смертности на 1000 жителей	Численность населения	Число смертей	
0-	2000	10	3000	30	40
20-	5000	9	5000	45	35
40-	3000	13	2000	28	25

Вычислите:
(i) число смертей в каждой возрастной группе для региона R;
(ii) общий коэффициент смертности для региона R;
(iii) стандартизованный коэффициент смертности для региона R;
(iv) общий коэффициент смертности для региона S;
(v) коэффициент смертности каждой возрастной группы для региона S;
(vi) стандартизованный коэффициент смертности для региона S.

Прокомментируйте результаты.

Упражнение 28В

1. В таблице представлены данные по двум соседним районам M и N.

Возрастная группа	Район M		Район N		Стандартное население, %
	Процент населения	Коэффициент смертности на 1000 жителей	Численность населения	Число смертей	
0-	40	8	2000	14	30
20-	35	7	1000	6	45
40-	25	10	3000	30	25

Вычислите:
(i) общий коэффициент смертности для района M;
(ii) стандартизованный коэффициент смертности для района M;
(iii) общий коэффициент смертности для района N;
(iv) коэффициент смертности каждой возрастной группы для района N;
(v) стандартизованный коэффициент смертности для района N.

Сделайте вывод, в каком районе условия для жизни лучше.

2. В таблице представлены данные по двум соседним городам J и K.

Возрастная группа	Город J		Город K		Стандартное население, %
	Численность населения	Число смертей	Численность населения	Коэффициент смертности на 1000 жителей	
0-	3000	27	2000	9	40
20-	5000	40	6500	8	25
40-	2000	22	1500	12	35

Вычислите:

(i) коэффициент смертности каждой возрастной группы для города J;

(ii) общий коэффициент смертности для города J;

(iii) стандартизованный коэффициент смертности для города J;

(iv) число смертей в каждой возрастной группе для города K;

(v) общий коэффициент смертности для города K;

(vi) стандартизованный коэффициент смертности для города K.

Прокомментируйте результаты.

Глава 7

ДИАГРАММЫ РАССЕИВАНИЯ

7.1. Понятие диаграммы рассеивания

Можно изучать взаимосвязь двух переменных, нанеся на график соответствующие значения каждой переменной. Значения независимой переменной, которые контролируются или выбираются исследователем, откладываются по оси X, или в направлении X. Значения другой переменной, называемой зависимой, откладываются по оси Y, или в направлении Y.

Пример 1.

Постройте диаграмму рассеивания оценок 10 учеников по математике и физике (рекомендуемый масштаб 1см $\Rightarrow$ 10 баллов). Прокомментируйте результат.

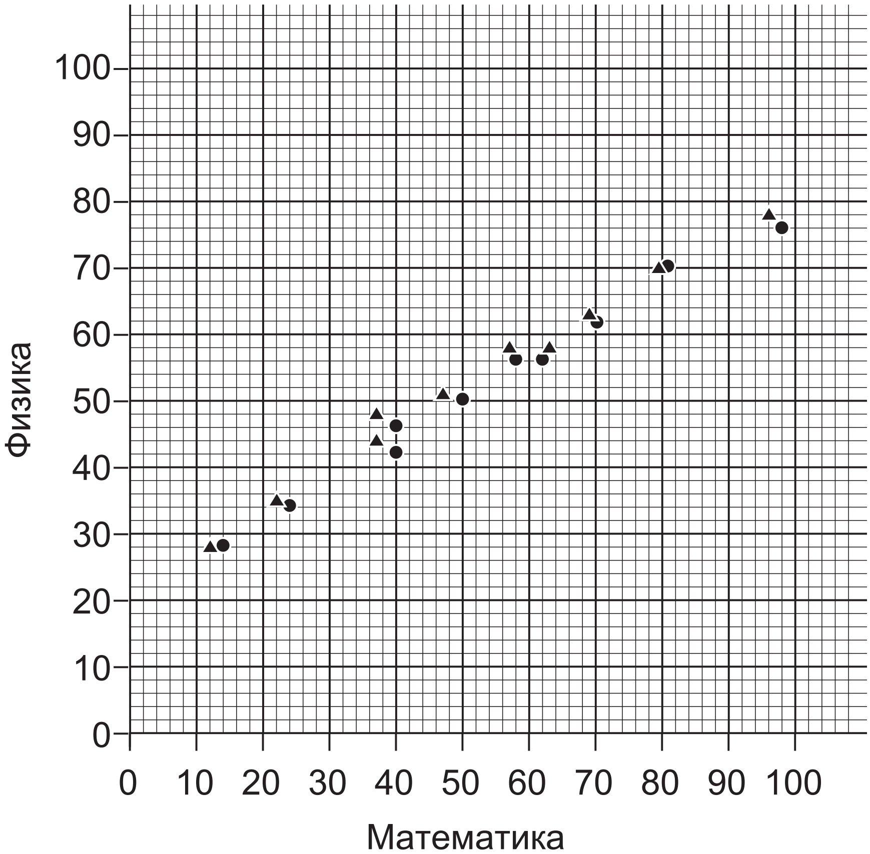
Ученик	A	B	C	D	E	F	G	H	I	J
Математика	50	58	14	40	70	96	80	24	40	62
Физика	50	56	28	42	62	76	70	34	46	56

Решение.

Нанеся на график оценки по физике (по оси Y) и оценки по математике (по оси X), получим диаграмму рассеивания (см. с. 214).

Замечание. Диаграммы рассеивания отличаются от обычных математических графиков тем, что одному значению x может соответствовать более одного значения y. В нашем примере это точки (40, 42) и (40, 46).

Аналогично одному значению y может соответствовать несколько значений x. В нашем примере это точки (58, 56) и (62, 56).



Комментарий. По диаграмме рассеивания видно, что малые значения y соответствуют малым значениям x, тогда как большие значения y соответствуют большим значениям x.

Это говорит о наличии **положительной корреляции** между x и y. Далее, поскольку значения или точки на диаграмме лежат на прямой линии, мы говорим, что имеется **положительная линейная корреляция** между x и y.

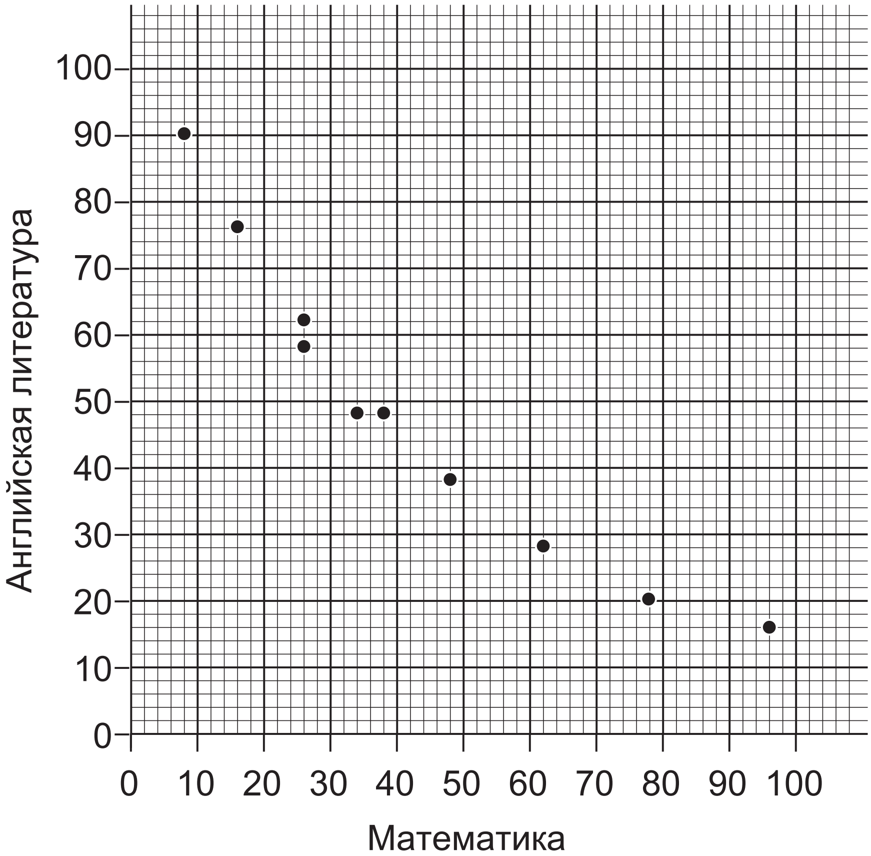
Пример 2.

Постройте диаграмму рассеивания оценок 10 учеников по математике и английской литературе (рекомендуемый масштаб 1 см $\Rightarrow$ 10 баллов). Прокомментируйте результат.

Ученик	A	B	C	D	E	F	G	H	I	J
Математика	26	34	8	48	96	26	62	16	78	38
Английская литература	58	48	90	38	14	62	28	76	20	48

Решение.

Нанеся на график оценки по английской литературе (по оси Y) и оценки по математике (по оси X), получим диаграмму рассеивания.



Комментарий. По диаграмме рассеивания видно, что большие значения y соответствуют низким значениям x, тогда как низкие значения y соответствуют большим значениям x.

Это говорит о наличии **отрицательной корреляции** между x и y. Далее, поскольку значения или точки на графике не лежат на прямой линии, в данном случае не имеется **отрицательной линейной корреляции** между x и y.

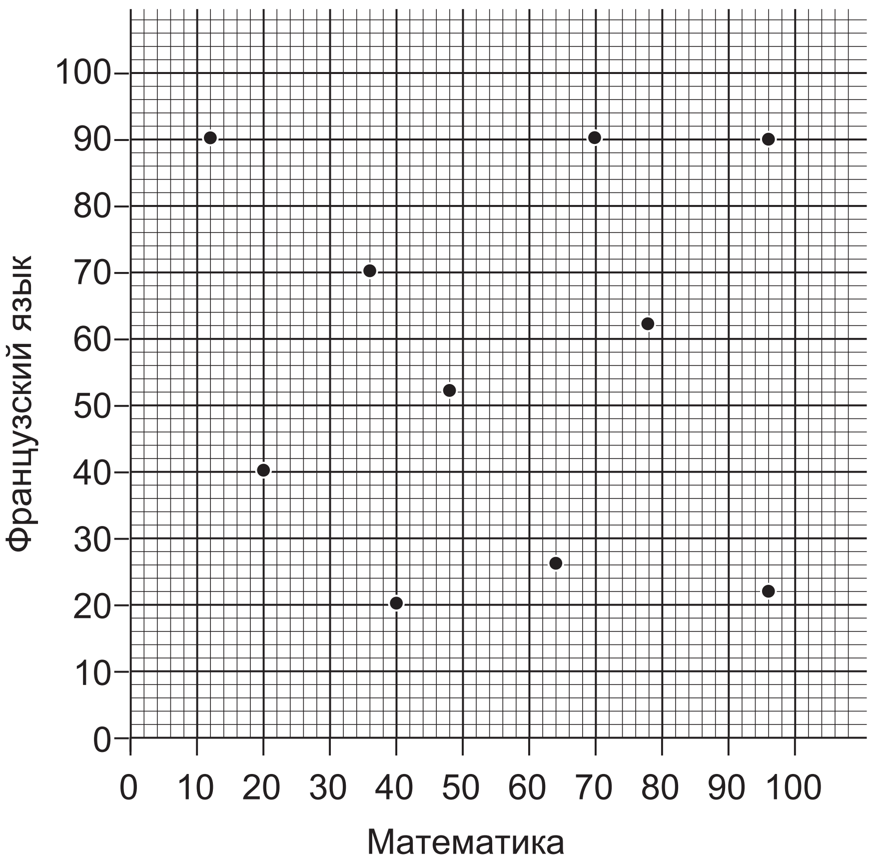
Пример 3.

Постройте диаграмму рассеивания оценок 10 учеников по математике и французскому языку (рекомендуемый масштаб 1 см ⇒ 10 баллов). Прокомментируйте результат.

Ученик	A	B	C	D	E	F	G	H	I	J
Математика	20	64	40	76	12	96	48	70	36	96
Французский язык	40	26	20	62	90	22	52	90	70	90

Решение.

Нанеся на график оценки по французскому языку (по оси Y) и оценки по математике (по оси X), получим диаграмму рассеивания.



Комментарий. Точки, или значения, на диаграмме не создают какой-либо особенной структуры. Следовательно, мы приходим к выводу, что **нет корреляции** между x и y.

7.2. Линейная корреляция

Если значения, или точки, лежат на прямой линии, пусть даже это и не строгая прямая, диаграмма рассеивания указывает на наличие **линейной корреляции**. Можно найти математическое соотношение между двумя переменными, построив **линию регрессии,** или линию наилучшего соответствия.

Для построения линии регрессии, или линии наилучшего соответствия, надо:

(i) Рассчитать среднюю арифметическую значений x и среднюю арифметическую значений y.

(ii) Нанести на график точку с координатами (среднее значение x, среднее значение y), или $(\overline{x}, \overline{y})$.

(iii) Рассчитать среднее значение x и среднее значение y для точек, расположенных левее точки с координатами (среднее значение x, среднее значение y). В случае положительной корреляции эти точки называются **нижним массивом**.

(iv) Нанести эту точку на график.

(v) Рассчитать координаты точки (среднее значение x, среднее значение y) для **верхнего массива** и нанести эту точку на график.

(vi) Провести прямую, проходящую через точку с координатами $(\overline{x},\overline{y})$ и, по возможности, через точки с координатами (среднее значение x, среднее значение y) для верхнего и нижнего массивов.

(vii) Уравнение прямой $y = mx + c$. Из уравнения видно, что тангенс угла наклона прямой равен m. Таким образом, определение тангенса угла наклона позволит нам найти значение коэффициента m из уравнения.

(viii) Если сможем определить значение y при $x = 0$ (при пересечении с осью Y, то получим значение коэффициента c (свободный член) из уравнения.

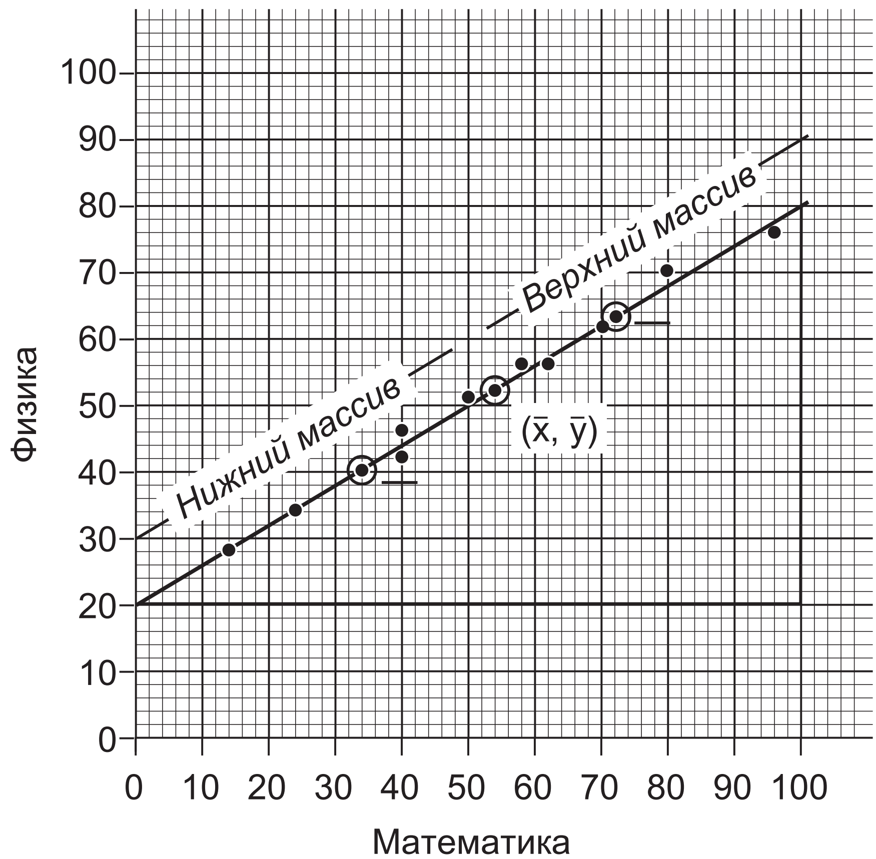
(ix) Если мы не можем определить значение y при $x = 0$ (у прямой нет пересечения с положительной осью Y), то вычислим значение m, подставляя в уравнение $y = mx + c$ средние значения x и y.

В примере 1 из параграфа 7.1 средняя арифметическая x равна 53,4, а средняя арифметическая y равна 52. Рекомендуется точки с координатами (среднее значение x, среднее значение y) обводить кружочками, а исходные значения наносить на график крестиками.

Нижний массив, т.е. точки левее $(\bar{x}, \bar{y})$ – это точки с координатами (50, 50), (14, 28), (40, 42), (24, 34) и (40, 46). Среднее значение x для этих 5 точек равно 33,6. Среднее значение y для этих 5 точек равно 40.

Верхний массив, т.е. точки правее, – это точки с координатами (58, 56), (70, 62), (96, 76), (80, 70) и (62, 56). Среднее значение x для этих 5 точек равно 73,2. Среднее значение y для этих 5 точек равно 64.

На диаграмме рассеивания отмечены все эти точки и проведена линия наилучшего соответствия.



Тангенс угла наклона прямой определяется из большого треугольника, изображенного на диаграмме.

Изменение величины y составляет 80 – 20 = 60.

Изменение величины x составляет 100 – 0 = 100.

Тангенс угла наклона равен 60 : 100 = 0,60. Таким образом, $m = 0{,}60$.

Можно легко определить значение y при $x = 0$. Свободный член равен 20, т.е. $c = 20$.

Теперь уравнение линии регрессии запишется следующим образом:

$$y = 0{,}60x + 20.$$

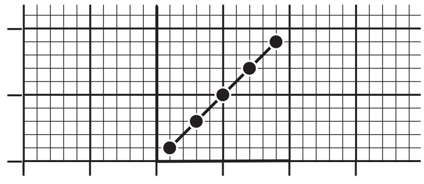
Замечание. Для определения тангенса угла наклона рекомендуется использовать наибольший из возможных треугольников. Это позволит для вычисления тангенса угла наклона выбрать длину основания треугольника (изменение переменной x) равной удобному целому числу, такому как 5 или 10.

7.3. Меры корреляции

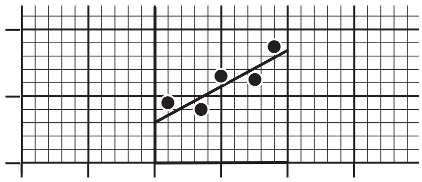
Как мы уже видели, если малые значения y соответствуют малым значениям x, а большие значения y соответствуют большим значениям x, то имеется положительная корреляция.

Мы измеряем корреляцию по шкале от –1 до +1.

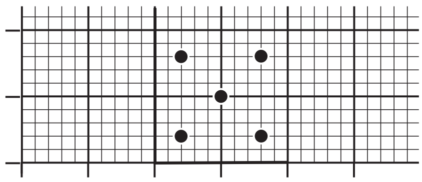
(i) Если точки на диаграмме рассеивания лежат **строго** на прямой линии с положительным тангенсом угла наклона, корреляция равна +1. Это называется полной положительной корреляцией.



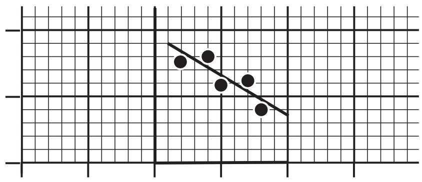
(ii) Если точки расположены вдоль прямой линии с положительным тангенсом угла наклона, но не лежат на этой прямой, то корреляция будет около +0,8.



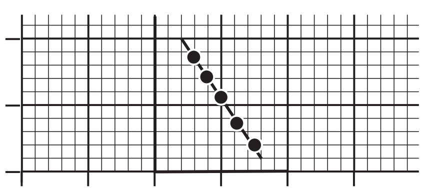
(iii) Если точки не образуют какой-либо четкой структуры, корреляция равна нулю.



(iv) Если точки расположены вдоль прямой линии с отрицательным тангенсом угла наклона, но не лежат на этой прямой, то корреляция будет около –0,8.



(v) Если точки на диаграмме рассеивания лежат **строго** на прямой линии с отрицательным тангенсом угла наклона, корреляция равна –1. Это называется полной отрицательной корреляцией.



Аудиторный тест № 20

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

В данном тесте:

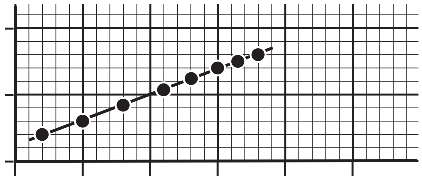
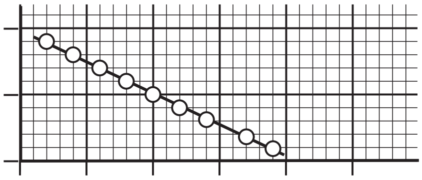
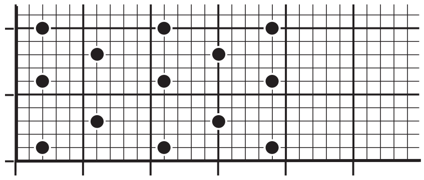
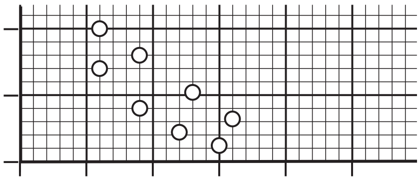
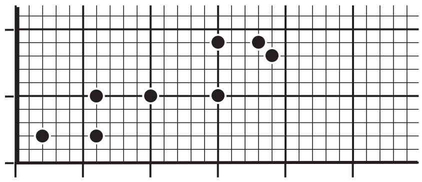
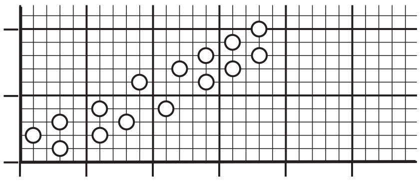
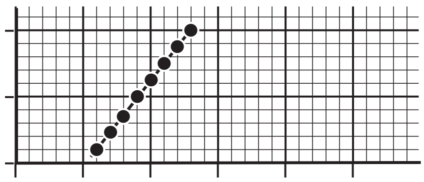
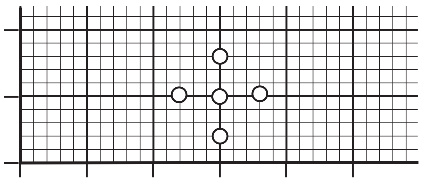
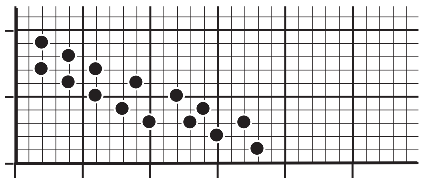
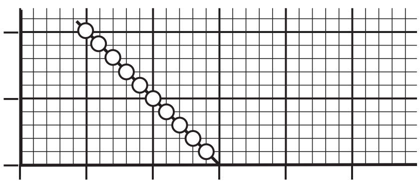
ответ A означает, что корреляция равна +1, т.е. полная положительная корреляция;

ответ B означает, что корреляция равна +0,5, т.е. умеренная положительная корреляция;

ответ C означает, что корреляция равна 0, т.е. переменные не зависят друг от друга;

ответ D означает, что корреляция равна –0,5, т.е. умеренная отрицательная корреляция;

ответ E означает, что корреляция равна –1, т.е. полная отрицательная корреляция.

№ вопроса	Диаграмма	№ вопроса	Диаграмма
1.		**2.**	
3.		**4.**	
5.		**6.**	
7.		**8.**	
9.		**10.**	

Упражнение 29А

1. Результаты анализа взаимосвязи между двумя переменными x и y приведены в таблице.

x	14,0	23,0	21,0	29,0	17,0	26,0
y	23,5	38,0	34,5	47,0	29,0	41,5

(i) Постройте диаграмму рассеивания по имеющимся данным, используя масштаб 2 см $\Rightarrow$ 4 ед. по оси X и 2 см $\Rightarrow$ 5 ед. по оси Y.
(ii) Вычислите средние арифметические x и y и полусредние (средние значения для нижнего и верхнего массивов).
(iii) Затем проведите на глаз линию наилучшего соответствия.
(iv) Найдите уравнение прямой в виде $y = mx + c$, где m и c определите с точностью до двух значащих цифр.
(v) По уравнению найдите значение y при $x = 18{,}0$.
(vi) Полученный ответ проверьте по графику.

2. Геологом собраны образцы минералов, которые предположительно являются или баритом, или кварцем. Масса и объем образцов приведены в таблице.

Образец	A	B	C	D	E	F	G	H	I	J	K	L
Объем ($см^3$)	1,8	2,3	4,9	5,8	3,0	3,5	2,7	3,6	3,9	4,7	5,5	2,1
Масса (г)	4,8	10,2	13,0	10,2	8,0	15,6	12,0	9,6	17,4	21,0	14,6	9,4

(i) Постройте диаграмму рассеивания массы от объема, используя масштаб 2 см $\Rightarrow$ 1 $см^3$ по оси X и 2 см $\Rightarrow$ 2 г по оси Y.
(ii) Масса 1 $см^3$ барита равна 4,5 г. Проведя соответствующую прямую линию на графике, перечислите образцы барита.
(iii) Перечислите образцы кварца.
(iv) Определите образец, который не является ни баритом, ни кварцем.
(v) Вычислите средний объем и среднюю массу образцов кварца.
(vi) Затем проведите на глаз линию наилучшего соответствия через точки, представляющие образцы кварца.

(vii) По линии наилучшего соответствия найдите с точностью до 1 знака после запятой:
a) массу образца кварца объемом 4 см^3;
b) объем кварцевого образца массой 16 г;
c) плотность кварца (г/см^3).

Упражнение 29B

1. В таблице приведены данные по количеству собранного ячменя в Соединенном Королевстве за девятилетний период. Площадь полей указана в миллионах гектаров, а урожай – в миллионах тонн.

Площадь, x	1,6	1,9	2,0	2,2	2,5	2,4	2,4	2,4	2,2
Урожайность, y	5,8	6,7	7,5	8,2	8,7	9,2	8,3	8,8	7,5

(i) Постройте диаграмму рассеивания, характеризующую зависимость урожая от площади, используя масштаб 5 см $\Rightarrow$ 1 млн га по оси X и 2 см $\Rightarrow$ 1 млн т по оси Y.
(ii) Вычислите средние арифметические x и y и полусредние (средние значения для нижнего и верхнего массивов).
(iii) Объясните, почему в данном случае линия наилучшего соответствия должна проходить через начало координат.
(iv) Затем проведите на глаз линию наилучшего соответствия.
(v) Запишите уравнение прямой в виде $y = mx + c$, где m и c определите с точностью до одного знака после запятой.
(vi) Объясните не на статистическом языке смысл m в данном уравнении.

2. В ходе агрономического эксперимента посевная площадь была поделена на девять равных участков. Затем было подсчитано число растений двух видов с кодовыми наименованиями X и Y, произрастающих на каждом участке. Результаты приведены в таблице.

Образец	A	B	C	D	E	F	G	H	I
Растение X	5	8	13	2	15	12	3	7	10
Растение Y	7	11	18	3	21	17	4	10	14

(i) Постройте диаграмму рассеивания числа растений вида Y от числа растений вида X, используя масштаб 1 см = = 1 растение.
(ii) Вычислите средние арифметические x и y и полусредние (средние значения для нижнего и верхнего массивов).
(iii) Затем проведите на глаз линию наилучшего соответствия.
(iv) Запишите уравнение прямой в виде $y = mx + c$, где m и c определите с точностью до одного знака после запятой.
(v) Объясните, почему все данные приведены с точностью до целого числа.
(vi) Дайте обоснованный ответ, способствуют ли условия роста растений вида X росту растений вида Y.
(vii) Какой вид растений целесообразнее выращивать в таких условиях?

Глава 8

ВЕРОЯТНОСТЬ

8.1. Виды событий

8.1.1. Независимые события

Два события G и H являются независимыми, если исход, или результат, одного события (G) не влияет на исход, или результат, другого события (H).

Пример 1.

Если мы бросаем две игральные кости, одну красную и одну синюю, то число, выпавшее на верхней грани красной кости, не влияет на число, выпавшее на верхней грани синей кости. Следовательно, эти два события являются независимыми.

Пример 2.

Если мы вытягиваем по одной карте из двух колод, то карта, взятая из первой колоды, не повлияет на карту, взятую из другой. Таким образом, эти два события оказываются независимыми.

8.1.2. Зависимые события

Два события J и K являются зависимыми, если исход, или результат, одного из них (J) влияет на исход, или результат, другого события (K).

Пример. Если мы вытягиваем две карты из одной и той же колоды, первая карта влияет на вторую. Если первым вытянули туза треф, то второй картой не может снова оказаться туз треф. Если первой мы вытянули карту красной масти, то вторая карта наиболее вероятно окажется черной масти и менее вероятно красной. Следовательно, два события зависимы.

8.1.3. Взаимоисключающие события

Два исхода M и N называются взаимоисключающими, если один исход, или результат, M делает исход, или результат, N невозможным.

Пример. Если карта, взятая из обычной колоды в 52 карты, оказалась масти «червы», то она не может быть черной. Таким образом, для одной карты «червы» и «черная» – взаимоисключающие события.

Аудиторный тест № 21

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

В этом тесте рассмотрим вытягивания карты из обычной колоды в 52 карты.

Ответ A означает «События X и Y – взаимоисключающие события».

Ответ B означает «Событие X влечет за собой появление события Y».

Ответ C означает «События X и Y являются зависимыми событиями».

Ответ D означает «События X и Y – независимые события».

Номер вопроса	Событие X	Событие Y
1.	туз	семерка
2.	бубны	красная масть
3.	красная масть	червы
4.	трефы	туз
5.	черная масть	картинка
6.	червы	черная масть
7.	нечетное число	пятерка
8.	червы	красная масть
9.	пики	черная масть
10.	картинка	дама
11.	картинка	девятка
12.	четное число	бубны
13.	шестерка	черная масть
14.	четное число	пятерка
15.	король	картинка
16.	четверка	картинка
17.	шестерка	тройка
18.	черная масть	трефы
19.	красная масть	пятерка
20.	туз	четное число

8.2. Теоретическая вероятность

Вероятность свершения того или иного исхода, или результата, может быть вычислена *на чисто теоретической основе*.

Пример 1. Рассмотрим шестигранную игральную кость, на грани которой нанесены цифры 1, 2, 3, 4, 5, 6. Предположим, что кость имеет одинаковую плотность, т.е. вероятность выпадения каждой грани при броске кости одинакова. Поскольку имеется 6 граней, или 6 равновероятных результатов, то вероятность выпадения какой-либо определенной цифры равна $^1/_6$.

$$P(1) = {}^1/_6,\ P(2) = {}^1/_6,\ P(3) = {}^1/_6 \text{ и т.д.}$$

Пример 2. Рассмотрим обыкновенную монету. Одна сторона называется «орел» (поскольку на ней изображен орел), другая – «решка». Снова предполагаем, что монета «честная» – имеет одинаковую плотность, т.е. вероятность выпадения каждой из сторон одинакова. Так как имеются две стороны, или два равновероятных события, вероятность выпадения «орла» равна ½ и вероятность выпадения «решки» равна ½.

$$P(\text{орел}) = ½ \text{ и } P(\text{решка}) = ½.$$

Пример 3. Рассмотрим стандартную колоду из 52 игральных карт. Если взять из колоды 1 карту **наугад** (т.е. случайным образом, без какой-либо системы или предпочтений), вероятность вытянуть любую карту будет одинаковой. Поскольку в колоде 52 карты, вероятность достать какую-то определенную карту будет равна $^1/_{52}$.

$$P(\text{туз пик}) = {}^1/_{52}.$$

Если все исходы эксперимента или события равновероятны, то мы можем сказать, что:

$$P(\text{исхода определенного типа}) = \frac{\text{Число исходов этого типа}}{\text{Общее число исходов}}.$$

Пример 4. Предположим, что мы вытянули 1 карту из колоды в 52 карты. Всего имеется 52 карты, т.е. 52 равновероятных события.

В колоде 4 короля, что составляет 4 из этих равновероятных событий.

Таким образом, вероятность достать короля составляет $^4/_{52} = {}^1/_{13}$.

Общая вероятность

В любой ситуации или эксперименте общая вероятность (сумма всех частных вероятностей) должна быть всегда равна 1.

Это дает очень важную и полезную проверку.

В примере 1:

$$P(1) = {}^{1}/_{6},\ P(2) = {}^{1}/_{6},\ P(3) = {}^{1}/_{6},\ P(4) = {}^{1}/_{6},\ P(5) = {}^{1}/_{6},\ P(6) = {}^{1}/_{6}.$$

Общая вероятность, или вероятность того, что произойдет какое-либо событие, равна:

$${}^{1}/_{6} + {}^{1}/_{6} + {}^{1}/_{6} + {}^{1}/_{6} + {}^{1}/_{6} + {}^{1}/_{6} = 1.$$

В примере 2 P(орел) = ½, P(решка) = ½.

Суммарная, или общая, вероятность равна ½ + ½ = 1.

Аудиторный тест № 22

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вопросы 1–6 относятся к эксперименту с бросанием игральной кости.

Вопрос	**Варианты ответов**				
	A	B	C	D	E
1. P(3)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
2. P(четное число)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
3. P(нечетное число)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
4. P(больше или равно 3)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
5. P(меньше или равно 3)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
6. P(меньше 3)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$

В вопросах 7–18 идет речь о результатах вытягивания игральной карты из колоды в 52 карты.

Нечетные числа – 1(туз), 3, 5, 7, 9

Четные числа – 2, 4, 6, 8, 10

Вопрос	**Варианты ответов**				
	A	B	C	D	E
7. P(король)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
8. P(карта красной масти)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
9. P(нечетное число)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
10. P(картинка)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
11. P(пики)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
12. P(карта черной масти)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
13. P(четное число)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
14. P(бубны)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
15. P(бубновый король)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^4/_{13}$	$^1/_{52}$
16. P(меньше или равно 5)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
17. P(меньше 5)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
18. P(картинка красной масти)	$^1/_{13}$	$^2/_{13}$	$^1/_{26}$	$^3/_{26}$	$^5/_{26}$

8.3. Теорема сложения для взаимоисключающих событий

Если два события или исхода X и Y являются взаимоисключающими, то

$$P(X \text{ или } Y) = P(X) + P(Y).$$

Пример 1. Рассмотрим пример с вытягиванием карты из колоды. Вытягивание карты «червы» и «черной масти» являются взаимоисключающими событиями, поскольку карта «червы» не может быть черной, а карта «черной масти» не может быть «червы».

Таким образом,

$$P(\text{червы или черная масть}) = P(\text{червы}) + P(\text{черная масть}) = \\ = {}^1/_4 + {}^1/_2 = {}^3/_4.$$

Пример 2. Рассмотрим пример с бросанием обычной игральной кости.

Выпадение на верхней грани кости числа «3» и «четного числа» являются взаимоисключающими событиями, так как число «3» не является четным, а «четное число» не может быть равно 3.

Следовательно,

$$P(\text{3 или четное число}) = P(3) + P(\text{четное число}) = \\ = {}^1/_6 + {}^1/_2 = {}^4/_6 = {}^2/_3.$$

Замечание. Запомните, что теорема сложения применима только к **взаимоисключающим событиям**.

Пример 1. Рассмотрим пример с вытягиванием карты из колоды. Вероятность вытянуть карту масти червы или короля **не равна** сумме вероятностей достать карту масти червы или достать короля, поскольку эти события не являются взаимоисключающими (существует карта король червей).

$$P(\text{масть червы или король}) \neq P(\text{масть червы}) + P(\text{король}).$$

Пример 2. Рассмотрим эксперимент с бросанием обычной игральной кости. Вероятность выпадения «3» или «нечетного числа» **не равна** сумме вероятностей выпадения «3» или выпадения «нечетного числа», так как эти события не являются взаимоисключающими (3 является нечетным числом).

$$P(\text{3 или нечетное число}) \neq P(3) + P(\text{нечетное число}).$$

Аудиторный тест № 23

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вычислите вероятность перечисленных ниже событий. Помните, что два указанных события могут не быть взаимоисключающими.

Вопросы 1–3 относятся к эксперименту с бросанием игральной кости.

Вопрос	**Варианты ответов**				
	A	B	C	D	E
1. P(1 или 2)	$^1/_6$	$^1/_3$	$^1/_2$	$^2/_3$	$^5/_6$
2. P(1 или нечетное число)	$^1/_6$	$^1/_3$	$^1/_2$	$^2/_3$	$^5/_6$
3. P(1 или четное число)	$^1/_6$	$^1/_3$	$^1/_2$	$^2/_3$	$^5/_6$

В вопросах 4–12 идет речь о результатах вытягивания игральной карты из колоды в 52 карты.

Нечетные числа – 1(туз), 3, 5, 7, 9

Четные числа – 2, 4, 6, 8, 10

4. P(король или дама)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
5. P(валет или нечетное число)	$^4/_{13}$	$^5/_{13}$	$^6/_{13}$	$^7/_{13}$	$^8/_{13}$
6. P(червы или черной масти)	$^1/_4$	$^1/_2$	$^3/_4$	$^1/_{13}$	$^2/_{13}$
7. P(червы или красной масти)	$^1/_4$	$^1/_2$	$^3/_4$	$^1/_{13}$	$^2/_{13}$
8. P(картинка или четное число)	$^4/_{13}$	$^5/_{13}$	$^6/_{13}$	$^7/_{13}$	$^8/_{13}$
9. P(четное или нечетное число)	$^7/_{13}$	$^8/_{13}$	$^9/_{13}$	$^{10}/_{13}$	$^{13}/_{13}$
10. P(картинка или туз)	$^1/_{13}$	$^2/_{13}$	$^3/_{13}$	$^4/_{13}$	$^5/_{13}$
11. P(картинка или красной масти)	$^8/_{13}$	$^9/_{13}$	$^{10}/_{13}$	$^{11}/_{13}$	$^{12}/_{13}$
12. P(четное число или король)	$^4/_{13}$	$^5/_{13}$	$^6/_{13}$	$^7/_{13}$	$^8/_{13}$

8.4. Теорема умножения для независимых событий

Если два события F и G независимые, то **вероятность свершения обоих событий F и G** равна **вероятности свершения события F,** умноженной на **вероятность свершения события G.**

Это можно записать следующим образом: P(F и G) = P(F) × P(G).

Пример 1. Рассмотрим пример с вытягиванием карты из колоды. Вероятность вытянуть пятерку красной масти равна вероятности вытягивания карты красной масти, умноженной на вероятность вытягивания карты достоинством 5.

P(пятерка красной масти) = P(красная масть) × P(пятерка).

В колоде из 52 игральных карт имеется 26 карт красной масти, т.е. P(красная масть) = $^{26}/_{52} = {}^{1}/_{2}$.

В колоде имеется 4 карты достоинством 5. Тогда P(5) = $^{4}/_{52} = {}^{1}/_{13}$.

Таким образом, вероятность того, что вытянутая карта окажется пятеркой красной масти, равна:

P(пятерка красной масти) = $^{1}/_{2} \times {}^{1}/_{13} = {}^{1}/_{26}$.

Полученный результат легко проверить.

В колоде из 52 карт имеется 2 пятерки красной масти (5 червей и 5 бубен), следовательно, P(пятерка красной масти) = $^{2}/_{52} = {}^{1}/_{26}$.

Пример 2. Снова рассмотрим эксперимент с вытягиванием игральной карты из колоды. Вероятность вытянуть картинку черной масти равна вероятности вытянуть карту черной масти, умноженной на вероятность достать картинку.

P(картинка черной масти) = P(черная масть) × P(картинка).

В колоде из 52 игральных карт имеется 26 карт черной масти, т.е. P(черная масть) = $^{26}/_{52} = {}^{1}/_{2}$.

В колоде имеется 12 картинок. Тогда P(картинка) = $^{12}/_{52} = {}^{3}/_{13}$.

Таким образом, вероятность того, что вытянутая карта окажется картинкой черной масти, равна:

P(картинка черной масти) = $^{1}/_{2} \times {}^{3}/_{13} = {}^{3}/_{26}$.

Снова проверяем результат.

В колоде из 52 карт имеется 6 картинок черной масти (король треф, дама треф, валет треф, король пик, дама пик, валет пик), следовательно, P(картинка черной масти) = $^{6}/_{52} = {}^{3}/_{26}$.

Замечание. Помните, что теорема умножения применима только к **независимым событиям**.

Пример. Обратимся к эксперименту с вытягиванием карты из колоды.

P(красная масть червы) не равна P(красная масть) × P(червы),

поскольку вытягивание карты «красной масти» и вытягивание карты «червы» не являются независимыми событиями. Карта красной масти с большей вероятностью окажется «червы», чем любая карта, взятая случайным образом.

Карта «червы» должна быть картой «красной масти».

Аудиторный тест № 24

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По команде преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вычислите вероятность перечисленных ниже событий. Помните, что два указанных события могут не быть взаимоисключающими.

Вопросы 1–5 относятся к эксперименту с одновременным бросанием двух игральных костей.

Вопрос	**Варианты ответов**				
	A	B	C	D	E
1. P(2 шестерки)	$^1/_6$	$^1/_3$	$^1/_2$	$^1/_4$	$^1/_{36}$
2. P(2 нечетных числа)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
3. P(2 четных числа)	$^1/_6$	$^1/_5$	$^1/_4$	$^1/_3$	$^1/_2$
4. P(2 одинаковых числа)	$^1/_6$	$^1/_3$	$^1/_2$	$^1/_4$	$^1/_{36}$
5. P(1 четное и 1 нечетное число)	$^1/_6$	$^1/_3$	$^1/_2$	$^1/_4$	$^1/_{36}$

В вопросах 6–10 идет речь о результатах вытягивания игральной карты из колоды в 52 карты.

Нечетные числа – 1 (туз), 3, 5, 7, 9

Четные числа – 2, 4, 6, 8, 10

6. P(король красной масти)	$^1/_2$	$^1/_4$	$^1/_{13}$	$^1/_{26}$	$^1/_{52}$
7. P(дама червы)	0	$^1/_4$	$^1/_{13}$	$^1/_{26}$	$^1/_{52}$
8. P(нечетное число пик)	$^1/_{26}$	$^5/_{26}$	$^1/_{52}$	$^3/_{52}$	$^5/_{52}$
9. P(черная карта пик)	0	$^1/_2$	$^1/_4$	$^1/_{13}$	$^1/_{169}$
10. P(красная карта треф)	0	$^1/_2$	$^1/_4$	$^1/_{13}$	$^1/_{169}$

Вопросы 11 и 12 относятся к эксперименту с подбрасыванием двух «честных» монет.

11. P(2 орел)	0	$^1/_4$	$^1/_3$	$^1/_2$	1
12. P(1 орел и 1 решка)	0	$^1/_4$	$^1/_3$	$^1/_2$	1

8.5. Комбинация независимых событий

Для определения вероятности свершения двух независимых событий бывает удобно составить таблицу.

Пример 1. Запишем в таблицу все 36 равновероятных исходов в случае, когда одновременно бросают две стандартные шестигранные кости (красную и синюю), а затем суммируют цифры, выпавшие на верхних гранях. Затем найдем вероятность каждого возможного исхода.

Решение.

		Первая кость (красная)					
	Сумма	1	2	3	4	5	6
Вторая кость (синяя)	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

В таблице представлены 36 чисел, 36 равновероятных исходов. Эти 36 чисел включают одну 2, две 3, три 4 и т.д.

Ниже приведены вероятности 11 различных исходов (от 2 до 12):

$P(2) = {}^1/_{36}$, $P(3) = {}^2/_{36}$, $P(4) = {}^3/_{36}$, $P(5) = {}^4/_{36}$, $P(6) = {}^5/_{36}$, $P(7) = {}^6/_{36}$,

$P(8) = {}^5/_{36}$, $P(9) = {}^4/_{36}$, $P(10) = {}^3/_{36}$, $P(11) = {}^2/_{36}$, $P(12) = {}^1/_{36}$.

Проверка. Вычисляем сумму вероятностей. Она равна ${}^{36}/_{36} = 1$.

Для осуществления проверки полезно вначале представить все вероятности в виде дробей с одинаковым знаменателем (в нашем примере это 36), а потом сокращать.

Пример 2. Запишем в таблицу все 36 равновероятных исходов в случае, когда одновременно бросают две стандартные шестигранные кости (красную и синюю), а затем перемножают цифры, выпавшие на верхних гранях. Затем найдем вероятность каждого возможного исхода.

Решение.

	Произведение	Первая кость (красная)					
		1	2	3	4	5	6
Вторая кость (синяя)	1	1	2	3	4	5	6
	2	2	4	6	8	10	12
	3	3	6	9	12	15	18
	4	4	8	12	16	20	24
	5	5	10	15	20	25	30
	6	6	12	18	24	30	36

В таблице представлены 36 чисел, 36 равновероятных исходов. Эти 36 чисел включают одну 1, две 2, две 3, три 4 и т.д.

Ниже приведены вероятности 18 различных исходов (от 1 до 36):

$P(1) = {}^1/_{36}$, $P(2) = {}^2/_{36}$, $P(3) = {}^2/_{36}$, $P(4) = {}^3/_{36}$, $P(5) = {}^2/_{36}$,
$P(6) = {}^4/_{36}$, $P(8) = {}^2/_{36}$, $P(9) = {}^1/_{36}$, $P(10) = {}^2/_{36}$, $P(12) = {}^1/_{36}$,
$P(15) = {}^2/_{36}$, $P(16) = {}^1/_{36}$, $P(18) = {}^2/_{36}$, $P(20) = {}^2/_{36}$,
$P(25) = {}^1/_{36}$, $P(30) = {}^2/_{36}$, $P(36) = {}^1/_{36}$.

Примечание. Вероятности всех других исходов равны 0.

8.6. Комбинация зависимых событий

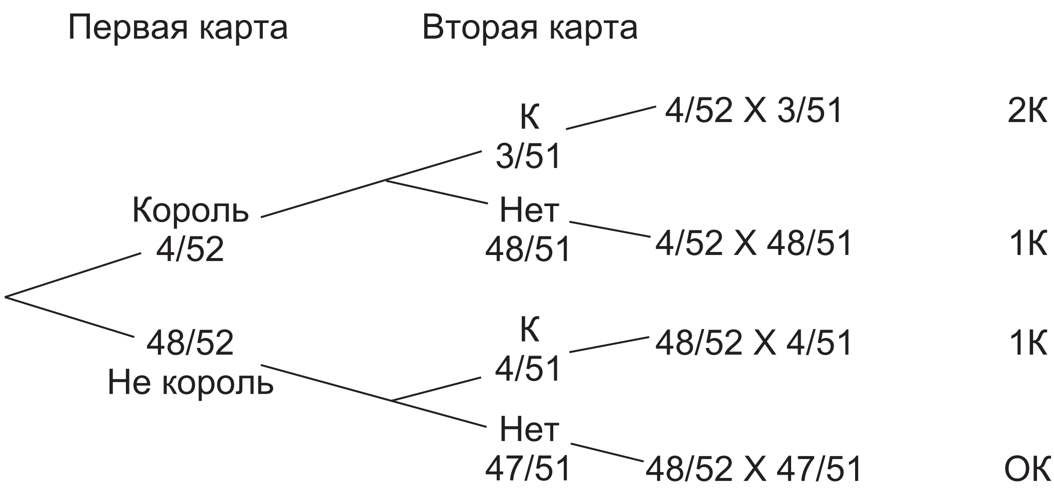
Для нахождения вероятности свершения двух событий, не являющихся независимыми, бывает полезно построить граф.

Пример 1. Найдите вероятность вытянуть (i) двух королей, (ii) только одного короля, (iii) ни одного короля в случае, когда одновременно достают 2 карты из одной колоды в 52 игральных карты.

(Вопрос (ii) «только одного короля» подразумевает вытягивание только одного короля и не включает случай вытягивания двух королей.)

Решение. Проиллюстрируем условие задачи с помощью графа. Вероятность свершения каждого из событий указана в конце каждой ветви.

Граф



Если к одному и тому же исходу (результату) ведут 2 или более ветви, то конечные вероятности этих ветвей складываются (суммируются).

Таким образом,

P(два короля) = ${}^{4}/_{52} \times {}^{3}/_{51}$ $= {}^{1}/_{221}$;

P(только один король) = ${}^{4}/_{52} \times {}^{48}/_{51} + {}^{48}/_{52} \times {}^{4}/_{51}$ $= {}^{32}/_{221}$;

P(ни одного короля) = ${}^{48}/_{52} \times {}^{47}/_{51}$ $= {}^{188}/_{221}$.

Эти три результата являются единственно возможными (поэтому говорят, что они **исчерпывающие**), следовательно, сумма их вероятностей должна быть равна 1.

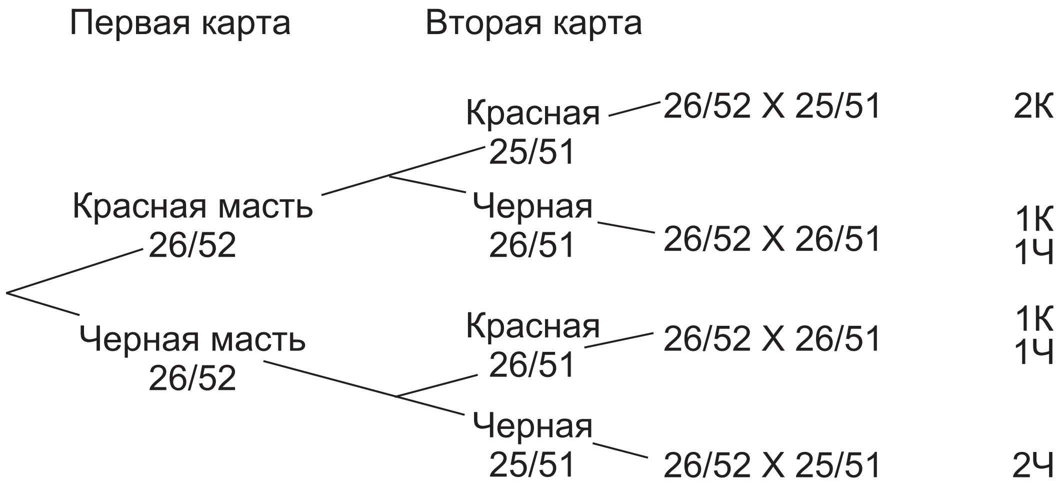
Как и прежде, это является очень важной и полезной проверкой.

Пример 2. Из колоды в 52 карты одновременно достают 2 карты. Найдите вероятность вытянуть:

(i) две карты красной масти;
(ii) 1 карту красной масти и 1 карту черной масти;
(iii) две карты черной масти.

Решение.

Граф



Таким образом,

P(две карты красной масти) $= {}^{26}/_{52} \times {}^{25}/_{51} = {}^{25}/_{102}$;

P(1 красная карта и 1 черная) $= {}^{26}/_{52} \times {}^{26}/_{51} + {}^{26}/_{52} \times {}^{26}/_{51} = {}^{52}/_{102}$;

P(две карты черной масти) $= {}^{26}/_{52} \times {}^{25}/_{51} = {}^{25}/_{102}$.

Эти три результата являются единственно возможными (поэтому говорят, что они **исчерпывающие**), следовательно, сумма их вероятностей должна быть равна 1.

Как и прежде, это является очень важной и полезной проверкой.

8.7. Математическое ожидание

Средняя арифметическая распределения вероятностей называется математическим ожиданием или ожидаемым значением.

Средняя арифметическая $= \dfrac{\sum fx}{\sum f}$, где f – частота.

Математическое ожидание равно $\dfrac{\sum px}{\sum p}$, где p – вероятность.

Поскольку суммарная вероятность Σp равна 1, то формула для математического ожидания упрощается до Σpx.

Пример 1.

Вычислите математическое ожидание, или ожидаемое число очков, при бросании **обычной** шестигранной игральной кости.

Решение.

Число, x	Вероятность, p	Число × Вероятность, px
1	$^1/_6$	$^1/_6$
2	$^1/_6$	$^2/_6$
3	$^1/_6$	$^3/_6$
4	$^1/_6$	$^4/_6$
5	$^1/_6$	$^5/_6$
6	$^1/_6$	$^6/_6$
Сумма	$\Sigma p = 1$	$\Sigma px = {}^{21}/_6$

Таким образом, математическое ожидание равно $\Sigma px = {}^{21}/_6 = 3{,}5$.

В данном примере распределение абсолютно симметрично. Легко увидеть без вычислений, что величина математического ожидания равна 3,5.

Пусть игроку, кидающему кубик, платят $1, если он выбросил цифру 1, $2, если он выбросил цифру 2, $3, если он выбросил цифру 3, и т.д. Тогда при большом числе испытаний он может ожидать выигрыш в среднем $3,50 за бросок. Игра будет честной, если вступительный взнос установить равным $3,50.

Пример 2.

Игра заключается в бросании двух костей. Вступительный взнос равен $1.

Если выпадают 2 шестерки, то игрок получает $5. При выпадении любого другого дубля игрок получает $2.

В других случаях игрок не получает ничего.

Рассчитайте ожидаемый выигрыш или проигрыш игрока.

Решение.

Результат	Выигрыш	Вероятность	ВыигрышЧ Ч Вероятность
Дубль 6	5	1/36	5/36
Другой дубль	2	5/36	10/36
Остальные результаты	0	30/36	0/36
Сумма		$\Sigma p = 36/36 = 1$	$\Sigma px = 15/36 = 5/12$

Из таблицы видно, что при большом числе испытаний игрок получит \$5/12 с одного доллара за игру.

Поскольку вступительный взнос равен \$1, то ожидаемый результат – проигрыш по \$7/12 с одного доллара.

Замечания.

(i) Строка «Остальные результаты» не вносит вклада в сумму «Выигрыш × Вероятность», но позволяет проверить, что сумма всех вероятностей Σp равна 1.

(ii) Проверку проще производить, если все вероятности выражены дробями с одинаковым знаменателем.

8.8. Честные игры

Игра называется «честной», если математическое ожидание игрока равно нулю, т.е. игрок при большом числе испытаний не ожидает ни выигрыша, ни проигрыша.

Пример 1.

При вытягивании игральной карты игрок получает следующие призы: \$20 за туз, \$10 за картинку, \$2 за карту достоинством 10. Если игрок достает любую другую карту, то не получает ничего.

Вычислите вступительный взнос, чтобы игра была честной.

Решение.

Результат	Выигрыш	Вероятность	Выигрыш Ч Ч Вероятность
Туз	20	1/13	20Ч1/13 = 20/13
Картинка	10	3/13	10Ч3/13 = 30/13
Десятка	2	1/13	2Ч1/13 = 2/13
Остальные результаты	0	8/13	0Ч8/13 = 0
Сумма		$\Sigma p = 13/13 = 1$	$\Sigma px = 52/13$

Таким образом, ожидаемый выигрыш, или математическое ожидание, равно 52/13 = 4.

Игра будет честной при вступительном взносе в $4.

Упражнение 30А

1. В сумке находятся 3 красных диска и 7 синих дисков. Из сумки одновременно вытаскивают 2 диска без замены.

Вычислите вероятность того, что:

(i) оба диска красные;

(ii) один диск красный, а второй – синий;

(iii) оба диска синие.

2. В сумке А находятся 2 красных и 3 синих диска, а в сумке В находятся 1 красный и 4 синих диска. Мальчик достает случайным образом диск из сумки А и кладет его в сумку В. Затем девочка вытаскивает диск из сумки В и кладет его в сумку А.

С помощью графа или другим способом вычислите вероятность того, что:

(i) в сумке А теперь находится 1 красный диск;

(ii) в сумке А теперь находятся 2 красных диска;

(iii) в сумке А теперь находятся 3 красных диска.

3. Одновременно бросают две шестигранные кости. Игра заключается в подсчете суммы выпавших цифр.

(i) Составьте таблицу 6×6 с результатами всех 36 исходов.

(ii) По таблице или иначе запишите плотность распределения 11 возможных исходов.

(iii) Запишите наиболее вероятный результат.

4. Человек случайным образом вытягивает карту из колоды. Вычислите вероятность того, что эта карта:

(i) туз;
(ii) картинка;
(iii) четное число;
(iv) черной масти.

Затем он возвращает карту в колоду и снова вытягивает карту. Рассчитайте вероятность того, что:

(i) обе карты – тузы;
(ii) обе карты одного цвета;
(iii) обе карты одинакового достоинства;
(iv) одна карта – черной масти, а вторая – красной.

Упражнение 30В

1. В сумке находятся 5 черных дисков и 7 белых дисков. Мальчик случайным образом вытаскивает из сумки диск. Диск кладется обратно в сумку, после чего девочка случайным образом вытаскивает из сумки один диск.

Вычислите вероятность того, что:

(i) оба диска черные;
(ii) один диск черный, а второй – белый;
(iii) оба диска белые.

2. В сумке X находятся 2 зеленых и 4 желтых диска, а в сумке Y находятся 3 зеленых и 3 желтых диска. Девочка достает случайным образом диск из сумки X и кладет его в сумку Y. Затем мальчик вытаскивает диск из сумки Y и кладет его в сумку X.

С помощью графа или другим способом вычислите вероятность того, что:

(i) в сумке X теперь находится 1 зеленый диск;
(ii) в сумке X теперь находятся 2 зеленых диска;
(iii) в сумке X теперь находятся 3 зеленых диска.

3. Одновременно бросают две шестигранных кости. Игра заключается в подсчете произведения выпавших цифр.

(i) Составьте таблицу 6×6 с результатами всех 36 исходов.
(ii) По таблице или иначе найдите вероятность того, что:
  (a) результат равен 2;
  (b) результат равен 6;
  (c) результат равен 12.
(iii) Запишите наиболее вероятный результат.

4. Человек случайным образом вытягивает карту из колоды в 52 карты. Вычислите вероятность того, что эта карта:

(i) король;

(ii) нечетное число (туз – нечетное число);

(iii) картинка;

(iv) красной масти.

Не возвращая карту в колоду, он вытягивает еще одну карту из колоды. Рассчитайте вероятность того, что:

(i) обе карты – тузы;

(ii) обе карты одного цвета;

(iii) обе карты – картинки;

(iv) обе карты – четные числа;

(v) одна карта – черной масти, а вторая – красной.

Глава 9

ОКРУГЛЕНИЕ И ОШИБКИ

9.1. Тенденциозное и нетенденциозное округление

Округление называется *тенденциозным*, если каждое значение округляется, или аппроксимируется, в одну и ту же сторону.

Пример 1. Для определения среднего возраста учеников в классе опросили всех учеников. Все ученики назвали «возраст последнего дня рождения», т.е. число полных лет. В данном случае мы имеем *тенденциозное* округление, поскольку каждый ученик *преуменьшает свой возраст*. Возраст одних учеников, может быть, превышает указанный ими на 1 месяц, других – почти на 12 месяцев, но все ученики занизили свой возраст. Средний возраст, рассчитанный по этим данным, будет на 6 месяцев, или на ½ года, меньше, чем фактический средний возраст.

Округление называется *нетенденциозным*, если около половины значений округляется в сторону уменьшения, а вторая половина округляется в сторону увеличения.

Пример 2. Для определения среднего возраста учеников в классе школьников спросили о «возрасте в ближайший день рождения». В данном случае мы имеем дело с нетенденциозным округлением, поскольку около половины учеников будут округлять свой возраст в сторону уменьшения, т.е. занижать свой возраст, тогда как остальные ученики будут округлять свой возраст в сторону увеличения, т.е. завышать свой возраст. Средний возраст, рассчитанный по этим данным, будет очень близок к фактическому среднему возрасту.

9.2. Наименьшее и наибольшее возможные значения округленных данных

Пример 1. Минимальное число, которое может быть корректно округлено в большую сторону до целого числа 17, равно 16,50. Таким образом, минимальным, или наименьшим возможным, значением 17 при корректном округлении до целого числа является 16,50.

Пример 2. Минимальное число, которое может быть корректно округлено в большую сторону до целой сотни до 800, равно 750. Таким образом, минимальным, или наименьшим возможным, значением 800 при корректном округлении до сотен является 750.

Пример 3. Максимальное число, которое может быть корректно округлено в сторону уменьшения до целого числа 13, равно 13,50. Таким образом, максимальным, или наиболее возможным, значением 13 при корректном округлении до целого числа является 13,50.

Аудиторный тест № 25

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

Вопрос	Варианты ответов				
	A	B	C	D	E

Из перечисленных ниже чисел выберите **наименьшее** возможное значение, корректно округленное с заданной степенью точности.

1. 56: до целого	55,4	55,5	55,6	55,9	56,5
2. 900: до сотен	840	849	850	890	899
3. 80: до десятков	75	79	79,5	79,95	79,99
4. 4,57: до трех значащих цифр	4,5	4,55	4,56	4,565	4,569
5. 19,2: до 1 знака после запятой	19,14	19,15	19,19	19,25	19,29
6. 4000: до тысяч	3490	3495	3500	3595	3599
7. 73: до целого	72	72,4	72,45	72,49	72,4
8. 130: до десятков	134	134,5	134,9	135	139

Из перечисленных ниже чисел выберите **наибольшее** возможное значение, корректно округленное с заданной степенью точности.

9. 89: до целого	88,4	88,9	89,4	89,49	89,5
10. 1300: до сотен	1345	1349	1350	1395	1399
11. 120: до десятков	124	124,5	124,9	125	129
12. 2,43: до трех значащих цифр	2,431	2,435	2,438	2,439	2,440
13. 23,6: до 1 знака после запятой	23,65	23,66	23,67	23,68	23,69
14. 7000: до тысяч	7450	7495	7499	7500	7501
15. 129: до целого	128,5	128,9	129,5	129,9	130
16. 170: до десятков	165	174	174,5	174,9	175

9.3. Наименьшее и наибольшее возможные значения комбинаций округленных данных

9.3.1. Наименьшее и наибольшее возможные значения суммы

Предположим, что ученик стоит на столе.

Известно, что рост ученика, стоящего на полу, равен 150 см с точностью до десятков.

Высота стола от пола равна 60 см с точностью до десятков.

Высота ученика над полом равна сумме двух этих высот.

Наименьшим возможным значением является сумма наименьшего возможного значения роста ученика, 145 см, и наименьшего возможного значения высоты стола, 55 см, что составляет 200 см. Наибольшим возможным значением является сумма наибольшего возможного значения роста ученика, 155 см, и наибольшего возможного значения высоты стола, 65 см, что составляет 220 см.

9.3.2. Наименьшее и наибольшее возможные значения разности

Теперь предположим, что ученик стоит рядом со столом.

Наименьшим возможным значением возвышения ученика над столом является наименьшее возможное значение разности двух данных высот.

Наименьшее возможное значение разности равно разности наименьшего возможного значения роста ученика, 145 см, и наибольшего возможного значения высоты стола, 65 см, что составляет 80 см.

Наибольшим возможным значением возвышения ученика над столом является наибольшее возможное значение разности двух данных высот.

Наибольшее возможное значение разности равно разности наибольшего возможного значения роста ученика, 155 см, и наименьшего возможного значения высоты стола, 55 см, что составляет 100 см.

9.3.3. Наименьшее и наибольшее возможные значения произведения

Пример. Размер ковра с точностью до целого метра составляет 5 м на 3 м. Вычислите наименьшее и наибольшее возможные значения площади ковра.

Решение. Длина ковра 5 м, измеренная с точностью до целого метра, означает, что в действительности она может составлять от 4,5 до 5,5 м. Ширина ковра 3 м, измеренная с точностью до целого метра, означает, что в действительности она может составлять от 2,5 до 3,5 м.

Наименьшее возможное значение площади, которая определяется как произведение длины ковра на его ширину, будет равно произведению наименьшего возможного значения длины и наименьшего возможного значения ширины, т.е. $4{,}5 \times 2{,}5 = 11{,}25$ м^2.

Наибольшее возможное значение площади, вычисляемой как произведение длины ковра на ширину, будет равно произведению наибольшего возможного значения длины и наибольшего возможного значения ширины, т.е. $5{,}5 \times 3{,}5 = 19{,}25$ м^2.

9.3.4. Наименьшее и наибольшее возможные значения частного

Пример. Автомобиль проехал расстояние 400 км (с точностью до десятка километров) за 5 ч (с точностью до часа). Вычислите наименьшее и наибольшее возможные значения скорости автомобиля.

Решение. Расстояние 400 км, измеренное с точностью до десятка километров, означает, что в действительности оно может составлять от 395 до 405 км. Время 5 ч, измеренное с точностью до часа, означает, что в действительности оно может составлять от 4,5 до 5,5 ч.

Наименьшее возможное значение скорости, которая определяется как частное от деления расстояния на время, будет равно наименьшему возможному значению расстояния, деленному на наибольшее возможное значение времени, т.е. $395 : 5{,}5 = 71{,}8$ км/ч.

Наибольшее возможное значение скорости, которая определяется как частное от деления расстояния на время, будет равно наибольшему возможному значению расстояния, деленному на наименьшее возможное значение времени, т.е. $405 : 4{,}5 = 90{,}0$ км/ч.

9.3.5. Обзор наименьших и наибольших возможных значений

В таблице приняты следующие обозначения:
L() – наименьшее возможное значение;
G() – наибольшее возможное значение.

	Наименьшее возможное значение	Наибольшее возможное значение
A + B	L(A) + L(B)	G(A) + G(B)
A – B	L(A) – G(B)	G(A) – L(B)
A × B	L(A) × L(B)	G(A) × G(B)
A : B	L(A) : G(B)	G(A) : L(B)
A : (B – C)	L(A) : [G(B) – L(C)]	G(A) : [L(B) – G(C)]

9.4. Абсолютные и относительные ошибки

9.4.1. Абсолютные ошибки

Абсолютная ошибка измерения определяется как разность между правильным, или истинным, значением и записанным значением.

Абсолютная ошибка измерения = Записанное значение –
– Правильное, или Истинное, значение.

Пример. Длина 12,9 м записана как 13 м. Найдите абсолютную ошибку.

Решение. Абсолютная ошибка равна 13 – 12,9 = 0,1 м.

Обратите внимание, что абсолютная ошибка имеет ту же единицу измерения, что и измеряемое значение.

9.4.1. Относительные ошибки

Относительная ошибка измерения определяется как частное от деления величины абсолютной ошибки измерения на истинное значение.

Относительная ошибка измерения =
= Абсолютная ошибка : Истинное значение.

Пример. Длина 12,9 м записана как 13 м. Найдите относительную ошибку.

Решение. Относительная ошибка равна 0,1 м : 12,9 м = 0,078.

Обратите внимание, что относительная ошибка является безразмерной величиной (доля или часть).

9.5. Наибольшая возможная ошибка измерения

Если измеренное значение округлено до целого числа, то наибольшая возможная ошибка измерения составляет ±0,5.

Если значение измерено с точностью до целого десятка, то наибольшая возможная ошибка измерения составляет ±5.

9.5.1. Наибольшая возможная ошибка суммы двух измерений

Наибольшая возможная ошибка суммы двух измерений определяется как сумма наибольших возможных ошибок этих двух измерений.

Пример. Размер ковра с точностью до целого числа составляет 5 м на 3 м. Найдите наибольшую возможную ошибку суммы длины и ширины.

Решение. Длина ковра равна 5 м ± 0,5 м. Наибольшая возможная ошибка составляет ±0,5 м.

Ширина ковра равна 3 м ± 0,5 м. Наибольшая возможная ошибка составляет ±0,5 м.

Наибольшая возможная ошибка суммы длины и ширины составляет ±(0,5 + 0,5), или ±1 м.

9.5.2. Наибольшая возможная ошибка разности двух измерений

Наибольшая возможная ошибка разности двух измерений определяется как сумма наибольших возможных ошибок этих двух измерений.

Пример. Длины двух кусков материи для платья равны соответственно 160 м и 70 м. Первое значение указано с точностью до 20 м, а второе – до 10 м. Найдите наибольшую возможную ошибку разности двух длин.

Решение. Длина первого куска материи равна 160 м $\pm$ 10 м. Длина второго куска материи равна 70 м $\pm$ 5 м.

Наибольшая возможная ошибка разности двух длин составляет $\pm(10 + 5)$, или ±15 м.

9.5.3. Наибольшая возможная ошибка произведения двух измерений

Наибольшая возможная ошибка произведения двух измерений равна приблизительной **сумме относительных ошибок** этих измерений.

Пример. Размер ковра с точностью до целого числа составляет 20 м на 10 м. Найдите наибольшую возможную ошибку вычисления площади ковра, рассчитанной как произведение длины на ширину.

Решение.

Наибольшая возможная ошибка длины равна 0,5 м.

Относительная ошибка измерения длины равна 0,5 : 20 = 0,025.

Наибольшая возможная ошибка ширины равна 0,5 м.

Относительная ошибка измерения ширины равна 0,5 : 10 = 0,05.

Относительная ошибка вычисления площади (произведения) равна 0,025 + 0,05 = 0,075.

Тогда наибольшая возможная ошибка вычисления площади (произведения) равна $0{,}075 \times 200 = 15$.

Таким образом, площадь ковра приблизительно равна 200 $\pm$ $\pm$ 15 м^2.

9.5.4. Наибольшая возможная ошибка частного двух измерений

Наибольшая возможная ошибка частного двух измерений приблизительно равна **сумме относительных ошибок** этих двух измерений.

Пример. Автомобиль проехал расстояние 350 км за 8 ч. Расстояние указано с точностью до десяти километров, а время – с точностью до получаса. Найдите наибольшую возможную ошибку вычисления скорости автомобиля (частное).

Решение.

Наибольшая возможная ошибка расстояния равна 5 км.

Относительная ошибка измерения расстояния равна 5 : 350 = = 0,014.

Наибольшая возможная ошибка измерения времени равна 0,25 ч.

Относительная ошибка измерения времени равна 0,25 : 8 = = 0,031.

Относительная ошибка вычисления скорости (частного) равна 0,014 + 0,031 = 0,045.

Тогда наибольшая возможная ошибка вычисления скорости (частного) равна $0{,}045 \times 43{,}75 = 1{,}97$.

Таким образом, скорость автомобиля приблизительно равна 43,75±1,97 км/ч.

9.5.5. Обзор наибольших возможных ошибок

Для нахождения наибольших возможных ошибок **суммы** или **разности** следует **сложить** наибольшие возможные **абсолютные** ошибки.

Для нахождения наибольших возможных ошибок **произведения** или **частного** следует **сложить** наибольшие возможные **относительные** ошибки.

Аудиторный тест № 26

Перед проведением теста подготовьте большие буквы A, B, C, D, E и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

В тесте используются следующие обозначения:
L() – наименьшее возможное значение;
G() – наибольшее возможное значение.
Все данные приведены с точностью до целого числа.

	Вопрос	**Варианты ответа**				
		A	B	C	D	E
1.	G(12)	11,5	11,9	12,1	12,5	12,99
2.	L(15)	14,5	14,6	14,9	15,4	15,9
3.	G(14 × 16)	209,25	222,75	224	224,75	239,25
4.	L(17 × 19)	305,25	321,75	323	323,75	341,25
5.	G(14 + 16)	29	29,5	30	30,5	31
6.	L(28 – 23)	4	4,5	5	5,5	6
7.	G(21 – 19)	1	1,5	2	2,5	3
8.	L(13 + 15)	27	27,5	28	28,5	29
9.	G(25 : 5)	4,45	4,64	5	5,44	5,67
10.	L(36 : 6)	5,46	5,62	6	6,45	6,64
11.	G(7 : 2)	3,5	4	4,5	5	5,5
12.	L(8 : 2)	3	3,5	4	4,5	5
13.	G(9 + 6)	14,5	15	15,5	16	16,5
14.	L(7 + 5)	10,5	11	11,5	12	12,5
15.	G(3 × 4 : 5)	1,94	2,4	2,86	3,5	3,75
16.	L(2 × 3 : 5)	0,55	0,68	0,83	1,2	1,94
17.	G[(2 × 3) + 5]	11	11,5	13,5	14,25	14,75
18.	G[(2 × 3) : 5]	1,2	1,5	1,75	1,94	2,17
19.	G[(5 : 3) – 2]	–0,33	0,2	0,7	1,2	1,7
20.	G[(8 : 5) + 3]	4,39	4,6	4,89	5,1	5,39

Упражнение 31A

1. Известно, что из 11 000 студентов, сдававших экзамен, 52% сдали успешно. Полагая, что оба числа указаны с точностью до двух значащих цифр, найдите:

(i) минимальное число студентов, которые могли сдать экзамен;

(ii) максимальное число студентов, которые могли сдать экзамен.

2. Автомобилист проехал 180 км за 4 ч. Известно, что расстояние измерено с точностью до километра, а время – с точностью до часа.

Вычислите:

(i) наибольшую возможную среднюю скорость;

(ii) наименьшую возможную среднюю скорость.

3. Длина прямоугольника была измерена с точностью 3%, а ширина – с точностью 4%. Найдите точность, с которой можно вычислить площадь прямоугольника.

Упражнение 31B

1. Стороны прямоугольника равны 5 м и 8 м. Результаты измерений указаны с точностью до метра.

Вычислите:

(i) наибольшее возможное значение площади прямоугольника;

(ii) наименьшее возможное значение площади прямоугольника.

2. Расстояние, которое проехал автомобилист, было измерено с точностью до 4%, а время в пути – с точностью до 2%. Найдите точность, с которой можно вычислить среднюю скорость автомобилиста.

3. Величина γ была измерена с точностью до 2%. Определите точность вычисления следующих величин:

(i) 5γ;

(ii) γ^2;

(iii) γ^3.

Глава 10

ВЫБОРКА И ПРОГРАММА НАБЛЮДЕНИЯ

10.1. Цель формирования выборки

Целью формирования выборки является оценка информации о всей совокупности, когда нет возможности оценить, измерить и изучить всю совокупность.

Единицей совокупности могут быть люди, кролики, металлические детали, транзисторы, электрические лампочки и т.д., практически все что угодно.

10.2. Пилотные обследования

Пилотные обследования – это маломасштабные обследования, проводимые перед полномасштабными опросами или обследованиями. Предполагается, что пилотные обследования выявят недостатки или ошибки в выбранном методе опроса или в анкете.

10.3. Случайная выборка

Это метод отбора, согласно которому каждая единица совокупности имеет равный шанс быть выбранной.

Метод случайного отбора включает:

(i) вытягивание имен из шляпы случайным образом;

(ii) использование выборки случайных чисел.

Использование выборки случайных чисел

Выборки случайных чисел представляют собой таблицы, в которых любое из чисел 0, 1, 2, 3, 4, 5, 6, 7, 8 или 9 имеет равную вероятность быть выбранным.

Таблицы случайных чисел включаются в статистические издания.

Пример 1.

Ниже приведены регистрационные номера и инициалы 20 учеников класса.

Номер	Инициалы	Номер	Инициалы	Номер	Инициалы
1.	N. A.	**8.**	D. J.	**15.**	W. S.
2.	R. B.	**9.**	R. K.	**16.**	D. T.
3.	M. C.	**10.**	N. L.	**17.**	F. V.
4.	E. D.	**11.**	D. M.	**18.**	D. W.
5.	G. E.	**12.**	E. N.	**19.**	R. W.
6.	M. G.	**13.**	F. P.	**20.**	C. Y.
7.	K. H.	**14.**	T. R.		

Используя приведенную ниже таблицу двухзначных случайных чисел, выберите четырех учеников из класса.

44	20	43	37	88	21	30	21	10	32
77	45	59	35	25	26	38	01	87	75
57	40	89	99	99	68	60	32	25	49
71	51	64	16	08	86	60	74	16	93

Решение.

Мы перемещаемся по таблице сверху вниз и слева направо. Если двухзначное число относится к регистрационным, т.е. является числом от 01 до 20 включительно, то мы отбираем ученика с соответствующим регистрационным номером. Если число больше 20, то мы его пропускаем.

Первое двухзначное число, являющееся регистрационным, равно 20. Следовательно, мы выбираем ученика под номером 20, инициалы которого – C.Y.

Второе двухзначное число, являющееся регистрационным, равно 10. Следовательно, мы выбираем ученика под номером 10, инициалы которого – N.L.

Третье двухзначное число, являющееся регистрационным, равно 01. Следовательно, мы выбираем ученика под номером 01, инициалы которого – N.A.

Четвертое двухзначное число, являющееся регистрационным, равно 16. Следовательно, мы выбираем ученика под номером 16, инициалы которого – D.T.

Таким образом, четырьмя отобранными учениками являются ученики с инициалами C.Y., N.L., N.A. и D.T.

10.4. Стратифицированная случайная выборка

В этом случае выборка состоит из конечного числа страт, или слоев, выбранных случайным образом.

Пример 2.

В школе 79 учеников первого года обучения, 122 ученика второго года обучения, 98 – третьего года обучения, 101 – четвертого года обучения и 95 – пятого года обучения. Сформируйте стратифицированную выборку из 25 учеников, представляющих всех учеников школы.

Решение.

В школе всего 495 учеников. Нам требуется отобрать 25, т.е. приблизительно 1 из 20 учеников школы.

Наша выборка, следовательно, состоит из:

4 учеников первого года обучения;

6 учеников второго года обучения;

5 учеников третьего года обучения;

5 учеников четвертого года обучения;

5 учеников пятого года обучения.

Из всех учеников первого года обучения выбираем случайным образом 4 учеников. Аналогично выбираем необходимое количество учеников из других групп.

10.5. Долевая выборка

Формирование долевой выборки основывается на понятии «выборочной сетки» – количестве представителей в каждой секции или группе, пропорциональной соответствующему количеству во всей совокупности.

Пример 3.

Ниже представлен состав учеников школы.

	Мальчики	**Девочки**
Первый год обучения	36	43
Второй год обучения	59	63
Третий год обучения	49	52
Четвертый год обучения	61	38
Пятый год обучения	49	68

Методом долевой выборки требуется избрать «школьный совет» из 52 учеников, представляющий школу. Рассчитайте, сколько представителей надо избрать в совет от каждой группы учеников.

Решение.

Общее число учеников школы равно 518. В состав совета должны быть избраны 52 ученика, что приблизительно составляет по 1 ученику от 10.

Тогда необходимое число представителей от каждой группы учеников будет следующим:

	Мальчики	**Девочки**
Первый год обучения	4	4
Второй год обучения	6	6
Третий год обучения	5	5
Четвертый год обучения	5	4
Пятый год обучения	5	7

10.6. Систематическая выборка

Систематическая выборка – это выборка, сформированная согласно какой-то определенной системе, например, выбор каждой *n*-й единицы совокупности. **Первая** единица выбирается случайным образом.

Пример 4.

Ниже приведены регистрационные номера и инициалы 20 учеников класса.

Номер	Инициалы	Номер	Инициалы	Номер	Инициалы
1.	N. A.	**8.**	D. J.	**15.**	W. S.
2.	R. B.	**9.**	R. K.	**16.**	D. T.
3.	M. C.	**10.**	N. L.	**17.**	F. V.
4.	E. D.	**11.**	D. M.	**18.**	D. W.
5.	G. E.	**12.**	E. N.	**19.**	R. W.
6.	M. G.	**13.**	F. P.	**20.**	C. Y.
7.	K. H.	**14.**	T. R.		

Сформируйте систематическую выборку из 4 учеников.

Решение.

Поскольку нам надо отобрать 4 учеников из 20, то мы выбираем каждого 5-го.

Мы должны случайным образом выбрать число от 1 до 5 включительно. Для этого бросаем кубик.

Если при первом броске выпадает 6, то мы игнорируем этот результат и бросаем снова.

Пусть на этот раз у нас выпало 3. Значит первым мы выбираем ученика с регистрационным номером 3. Его инициалы – M.C.

Теперь мы выбираем каждого 5-го ученика, т.е.:

Номер ученика	$3+5 = 8$	D.J.
Номер ученика	$8+5 = 13$	F.P.
Номер ученика	$13+5 = 18$	D.W.

Таким образом, 4 выбранных нами ученика имеют инициалы: M.C., D.J., F.P. и D.W.

10.7. Тенденциозная выборка

Тенденциозная выборка – это метод выборки, который дает только одностороннее представление об изучаемом явлении.

Например, если провести предвыборный опрос голосующих, используя телефон, то получим предвзятые результаты, так как наличие телефона подразумевает хороший доход. Таким образом, результаты опроса будут тенденциозными и больше отражать мнение людей с более высоким доходом.

10.8. Правила по составлению программы проведения наблюдения

Начинать разработку программы наблюдения следует с постановки цели исследования и составления вопросов программы наблюдения. По возможности покажите, что заполнение анкеты программы наблюдения будет выгодно человеку, заполняющему ее.

Вопросы программы:

(i) должны быть сформулированы вежливо;
(ii) должны быть сформулированы недвусмысленно, т.е. четко;
(iii) не должны быть многочисленными;
(iv) не должны вызывать затруднений и неудобств при ответе, например, требовать длинных вычислений;
(v) не должны быть трудными для наименее интеллектуальных и образованных из опрашиваемых;
(vi) должны предполагать четкий ответ, например, ДА/НЕТ, число, название местности, область, дата и т.д.;
(vii) должны быть составлены таким образом, чтобы при помощи ответов на одни вопросы контролировать ответы на другие вопросы;
(viii) должны иметь отношение к опрашиваемому;
(ix) НЕ должны склонять к какому-то определенному ответу (это вызовет *предвзятость*);
(x) НЕ должны требовать специальных знаний (если опрос не проводится среди специалистов в данной области).

Вопросы следует располагать в логической последовательности.

По возможности включайте «проверочные» вопросы для подтверждения предыдущих вопросов.

Анкету завершайте фразой «Благодарим Вас за помощь».

Упражнение 32

1. (i) Объясните кратко понятие «*Тенденциозная выборка*».
(ii) Приведите пример выборочного метода, который может быть тенденциозным.

2. Объясните, как следует создать систематическую выборку из 20 учеников школы, в которой обучаются 400 учеников.

3. Что является целью выборки?

4. Объясните кратко, что такое *долевая выборка.*

5. Перечислите 6 правил разработки составления программы наблюдения.

6. Что такое *пилотное обследование*? В чем заключается его цель?

Глава 11

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ РАНГОВ СПИРМЕНА

11.1. Введение

В главе 7, параграф 7.3, мы рассматривали понятие корреляции с помощью диаграмм рассеивания. Однако мы можем определить корреляцию количественно, рассчитав коэффициент корреляции рангов Спирмена.

$$\text{Коэффициент корреляции рангов Спирмена} = 1 - \frac{6 \cdot \sum D^2}{n(n^2 - 1)},$$

где ΣD^2 – сумма квадратов разностей рангов;
n – общее число индивидов или предметов, имеющих оба признака.

Пример 1. В таблице указан возраст и экзаменационные оценки 8 учениц одного года обучения.

Ученица	A	B	C	D	E	F	G	H
Возраст (год – месяц)	16-2	16-11	16-6	16-1	16-3	16-4	16-0	16-5
Экзаменационная оценка	16	23	24	7	8	15	12	18

Рассчитайте коэффициент корреляции рангов Спирмена между возрастом и экзаменационной оценкой. Прокомментируйте результат.

Решение. Прежде всего нам надо произвести «ранжирование» по возрасту и экзаменационным оценкам, т.е. расположить их по величине. Обычно ранжирование производится от большего значения к меньшему.

Ученица	A	B	C	D	E	F	G	H	$\sum$
Возраст (год – месяц)	16-2	16-11	16-6	16-1	16-3	16-4	16-0	16-5	
Ранг по возрасту	6	1	2	7	5	4	8	3	36
Ранг по экзаменационным оценкам	4	2	1	8	7	5	6	3	36
Экзаменационная оценка	16	23	24	7	8	15	12	18	
Разность рангов	+2	–1	+1	–1	–2	–1	+2	0	+5–5=0
D^2	4	1	1	4	4	1	4	0	16

Коэффициент корреляции рангов Спирмена $= 1 - \dfrac{6 \cdot \sum D^2}{n(n^2 - 1)} =$

$$= 1 - \frac{6 \times 16}{8(64 - 1)} =$$

$$= 1 - \frac{4}{21} = \frac{17}{21} = 0,810.$$

Комментарий. Значение коэффициента 0,810 говорит о наличии положительной корреляции между возрастом и экзаменационной оценкой: чем старше ученица (в рамках одного года), тем лучше она сдает экзамен.

Два важных способа проверки

(i) Сумма каждого ряда рангов данных должна быть одинаковой. В нашем примере сумма каждого ряда рангов равна 36.

В действительности сумма ряда рангов данных определяется длиной ряда рангов в соответствии с приведенной ниже таблицей.

Число наблюдений	2	3	4	5	6	7	8	9	10	11	12
Сумма рангов	3	6	10	15	21	28	36	45	55	66	78

(ii) Сумма разностей рангов всегда должна быть равна 0. В нашем примере сумма разностей рангов равна: +5 – 5 = 0.

11.2. Связные ранги

(i) Иногда два значения признака имеют одинаковую количественную оценку, например:

25, 27, 27, 29, 22, 23.

В этом случае двум равным значениям присваивается одинаковый ранг. Первый ранг (или 1) следует присвоить значению 29.

Два значения 27 были бы ранжированы как 2 и 3, если бы они не были равны. Ввиду их равенства этим значениям присваивается одинаковый ранг 2½ (средняя арифметическая, или простая средняя, из рангов 2 и 3).

Следующий ранг будет 4.

Окончательно ранжированный ряд выглядит следующим образом:

Значение признака	25	27	27	29	22	23
Ранг	4	2½	2½	1	6	5

Сумма рангов: $\Sigma = 21$.

(ii) Бывает, что три значения признака имеют одинаковую количественную оценку, например:

25, 27, 27, 27, 29, 22, 23.

В этом случае трем равным значениям присваивается одинаковый ранг. Первый ранг (или 1) следует присвоить значению 29.

Три значения 27 были бы ранжированы как 2, 3 и 4, если бы они не были равны. Ввиду их равенства этим значениям присваивается одинаковый ранг 3 (средняя арифметическая, или простая средняя, из рангов 2, 3 и 4).

Следующий ранг будет 5.

Окончательно ранжированный ряд выглядит следующим образом:

Значение признака	25	27	27	27	29	22	23
Ранг	5	3	3	3	1	7	6

Сумма рангов: Σ = 28.

Пример 2. В таблице указан возраст и экзаменационные оценки 8 мальчиков одного года обучения

Ученик	A	B	C	D	E	F	G	H
Возраст (год – месяц)	16-2	16-11	16-4	16-2	16-3	16-3	16-1	16-9
Экзаменационная оценка	15	23	24	7	8	15	12	18

Рассчитайте коэффициент корреляции рангов Спирмена между возрастом и экзаменационной оценкой. Прокомментируйте результат.

Решение. Прежде всего проведем «ранжирование» по возрасту и экзаменационным оценкам, т.е. расположим их по величине. Обычно ранжирование проводится от большего значения к меньшему.

Ученик	A	B	C	D	E	F	G	H	Σ
Возраст (год – месяц)	16-2	16-11	16-4	16-2	16-3	16-3	16-1	16-9	
Ранг по возрасту	6½	1	3	6½	4½	4½	8	2	36
Ранг по экзаменационным оценкам	4½	2	1	8	7	4½	6	3	36
Экзаменационная оценка	15	23	24	7	8	15	12	18	
Разность рангов	+2	–1	+2	–1½	–2½	0	+2	–1	+6 – 6=0
D^2	4	1	4	2,25	6,25	0	4	1	22,5

$$\text{Коэффициент корреляции рангов Спирмена} = 1 - \frac{6 \cdot \sum D^2}{n(n^2 - 1)} =$$

$$= 1 - \frac{6 \times 22,5}{8(64 - 1)} =$$

$$= 1 - 0,2679 = 0,7321.$$

Комментарий. Значение коэффициента 0,7321 говорит о наличии положительной корреляции между возрастом и экзаменационной оценкой: чем старше ученик (в рамках одного года), тем лучше он сдает экзамен.

11.3. Отрицательная корреляция

Пример 3. Двое судей ранжировали 6 игроков в бадминтон следующим образом:

Игрок	A	B	C	D	E	F
Судья X	1	3	5	2	6	4
Судья Y	5	4	1	6	2	3

Рассчитайте коэффициент корреляции рангов Спирмена и прокомментируйте результат.

Решение. Исходные данные уже ранжированы, поэтому сразу переходим к вычислению разностей между рангами.

Разность рангов	–4	–1	+4	–4	+4	+1	У = +9 – 9 = 0
D^2	16	1	16	16	16	1	У = 66

$$\text{Коэффициент корреляции рангов Спирмена} = 1 - \frac{6 \cdot \sum D^2}{n(n^2 - 1)} =$$

$$= 1 - \frac{6 \times 66}{6(36 - 1)} =$$

$$= 1 - 1,8857 = -0,8857.$$

Комментарий. Значение коэффициента –0,8857 говорит о наличии отрицательной корреляции: налицо несогласованность мнений двух судей.

11.4. Полная положительная корреляция

Когда результаты двух ранжирований полностью совпадают, коэффициент ранговой корреляции равен +1.

Пример 4. В таблице приведены результаты ранжирования 6 теннисистов двумя судьями.

Игрок	A	B	C	D	E	F	
Судья M	1	5	6	3	2	4	У = 21
Судья N	1	5	6	3	2	4	У = 21

Разность рангов	0	0	0	0	0	0	У = 0
D^2	0	0	0	0	0	0	У = 0

Коэффициент корреляции рангов Спирмена $=1-\dfrac{6\times 0}{6(36-1)}=+1$ точно.

11.5. Полная отрицательная корреляция

Когда результаты двух ранжирований полностью различаются, коэффициент ранговой корреляции равен –1.

Пример 5. В таблице приведены результаты ранжирования 6 игроков в поло двумя судьями.

Игрок	A	B	C	D	E	F	
Судья P	1	2	3	4	5	6	У = 21
Судья Q	6	5	4	3	2	1	У = 21

Разность рангов	–5	–3	–1	+1	+3	+5	У –9+9= 0
D^2	25	9	1	1	9	25	У = 70

Коэффициент корреляции рангов Спирмена $=1-\dfrac{6\times 70}{6(36-1)}=$

= 1 –2 = –1 точно.

11.6. Корреляция рангов является лишь грубой оценкой

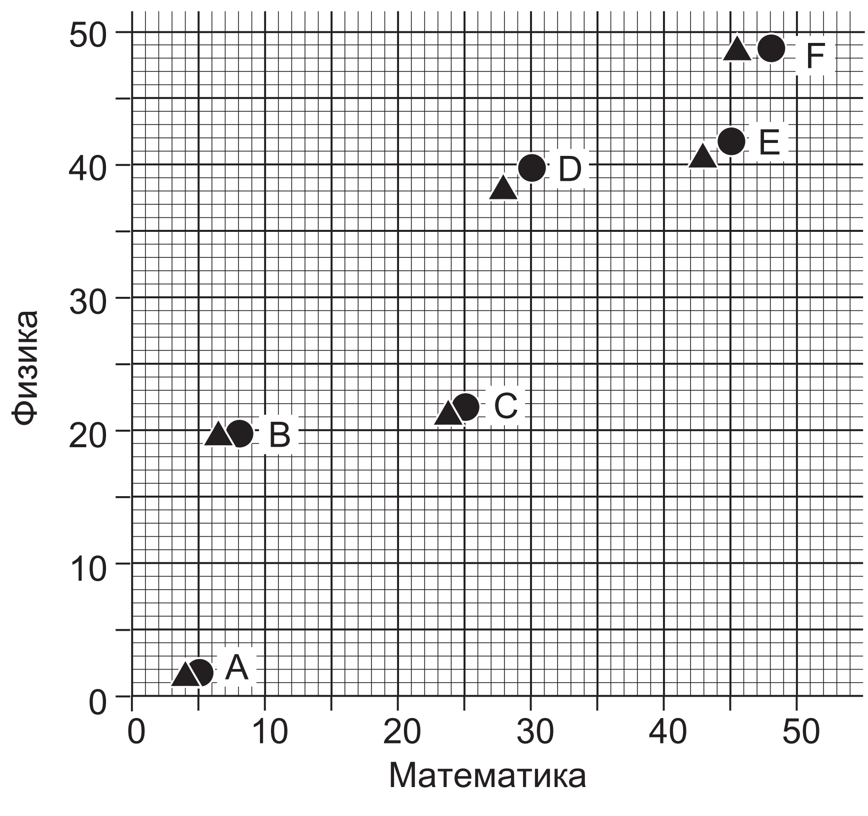
Может случиться, что ранговый коэффициент корреляции свидетельствует о полной корреляции, тогда как в действительности корреляция оказывается меньшей.

Пример 6.

Ученик	A	B	C	D	E	F
Оценка по математике	5	10	25	30	46	48
Оценка по физике	2	20	22	40	42	49
Ранг оценок по математике	6	5	4	3	2	1
Ранг оценок по физике	6	5	4	3	2	1

Результаты ранжирования оценок согласуются полностью, следовательно, ранговый коэффициент корреляции равен +1.

Однако по диаграмме рассеивания мы видим, что в действительности корреляция меньше +1.



Аудиторный тест № 27

Перед проведением теста подготовьте большие буквы A, B, C, D, E, и символ «?». Обозначения букв даются в каждом тесте. Решите, какой из вариантов ответов, представленных буквами A, B, C, D, E, является, по вашему мнению, правильным. По просьбе преподавателя поднимите ту букву, которая, по вашему мнению, соответствует правильному ответу.

Если вы не знаете ответ, то воспользуйтесь символом «?» (вместо угадывания или копирования чужого ответа). Это позволит преподавателю помочь вам.

	Вопрос	**Варианты ответа**				
		A	B	C	D	E

В каждом вопросе определите ранг первого значения (ранжирование осуществляется от наибольшего значения к наименьшему). Вопросы 1–10 НЕ включают связные ранги.

1.	84, 76, 95, 96, 80, 71, 56	1	2	3	4	5
2.	49, 23, 37, 20, 32, 31, 18	1	2	3	4	5
3.	65, 72, 68, 63, 69, 62, 60	1	2	3	4	5
4.	36, 31, 38, 29, 21, 27, 35	1	2	3	4	5
5.	41, 52, 49, 40, 47, 53, 39	1	2	3	4	5
6.	29, 83, 79, 65, 38, 49, 60	3	4	5	6	7
7.	16, 18, 15, 10, 12, 19, 17	3	4	5	6	7
8.	26, 38, 20, 36, 38, 35, 29	3	4	5	6	7
9.	75, 78, 72, 68, 65, 76, 61	3	4	5	6	7
10.	61, 68, 58, 66, 42, 64, 67	3	4	5	6	7

Вопрос		Варианты ответа				
		A	B	C	D	E
Вопросы 11–20 включают связные ранги.						
11.	25, 22, 18, 18, 12, 25, 21	1½	2	2½	3	3½
12.	47, 41, 25, 49, 39, 39, 47	1½	2	2½	3	3½
13.	56, 19, 54, 50, 59, 13, 54	1½	2	2½	3	3½
14.	70, 67, 76, 56, 71, 45, 70	1½	2	2½	3	3½
15.	76, 70, 57, 89, 76, 69, 76	1½	2	2½	3	3½
16.	85, 73, 91, 85, 68, 85, 80	2½	3	3½	4	4½
17.	82, 79, 74, 82, 60, 85, 90	2½	3	3½	4	4½
18.	75, 62, 86, 68, 87, 70, 75	2½	3	3½	4	4½
19.	43, 39, 47, 43, 48, 37, 45	2½	3	3½	4	4½
20.	64, 62, 91, 76, 60, 47, 87	2½	3	3½	4	4½

Упражнение 33А

1. В ходе сельскохозяйственного эксперимента вся посевная площадь была поделена на десять равных квадратных участков. Затем было подсчитано количество двух видов растений с кодовыми названиями X и Y, произрастающих на каждом квадрате. Результаты приведены в таблице.

Участок	A	B	C	D	E	F	G	H	I	J
Растение X	10	19	23	2	20	5	13	18	3	26
Растение Y	13	28	40	5	37	50	25	48	2	11

(i) Рассчитайте коэффициент корреляции рангов Спирмена между числом растений X и числом растений Y.
(ii) Прокомментируйте результат.

2. В таблице даны оценки 8 учащихся за два теста по математике.

Учащийся	A	B	C	D	E	F	G	H
Тест 1	18	16	13	13	19	12	10	20
Тест 2	10	14	16	17	9	18	20	15

(i) Рассчитайте коэффициент корреляции рангов Спирмена между оценками в тесте 1 и оценками в тесте 2.
(ii) Прокомментируйте результат.

Упражнение 33B

1. В таблице приведены данные о том, сколько учащихся выбирало учебные предметы для изучения в течение двух лет.

Предмет	Математика	Физика	Химия	Статистика	Английский язык	География
1-й год	53	27	23	18	36	29
2-й год	57	32	28	38	38	26

(i) Рассчитайте коэффициент корреляции рангов Спирмена между числом учащихся, изучавших каждый из этих предметов в 1-й год, и числом учащихся, изучавших тот же предмет во 2-й год.
(ii) Прокомментируйте результат.
(iii) По какому из предметов наблюдается большее различие?

2. В школе были сопоставлены оценки, полученные 8 учащимися на пробных экзаменах, с их экзаменационными оценками (высший балл 1, низший 8).

Учащийся	A	B	C	D	E	F	G	H
Пробная оценка	93	56	28	72	81	46	97	32
Экзаменационная оценка	2	5	8	4	3	5	1	7

(i) Рассчитайте коэффициент корреляции рангов Спирмена между пробными и действительными оценками.
(ii) Оцените эффективность пробного экзамена для предсказания результатов настоящего экзамена.

КРАТКИЕ ОБЗОРНЫЕ ВОПРОСЫ

Упражнения А

1. Назовите меру средней, которая проще (легче) всего определяется по диаграмме кумулятивных частот.

2. Средняя геометрическая 5 и числа x равна 15. Найдите число x.

3. Ниже приведен ряд температур, измеренных в °C:

$$7,\ 16,\ 13,\ 0,\ 2,\ -3,\ 4,\ 7,\ 4,\ 8,\ 7.$$

Найдите: (i) медиану, (ii) моду, (iii) размах, (iv) нижний квартиль, (v) верхний квартиль, (vi) межквартильный размах.

4. В таблице представлены оценки учащихся, полученные за небольшой тест.

Количество баллов	0	1	2	3	4	5	6	7	8	9
Число учащихся	2	5	7	10	14	13	9	7	6	2

Найдите: (i) медиану, (ii) нижний квартиль, (iii) верхний квартиль, (iv) межквартильный размах.

5. В таблице представлены оценки, полученные учащимися по математике, статистике и английскому языку.

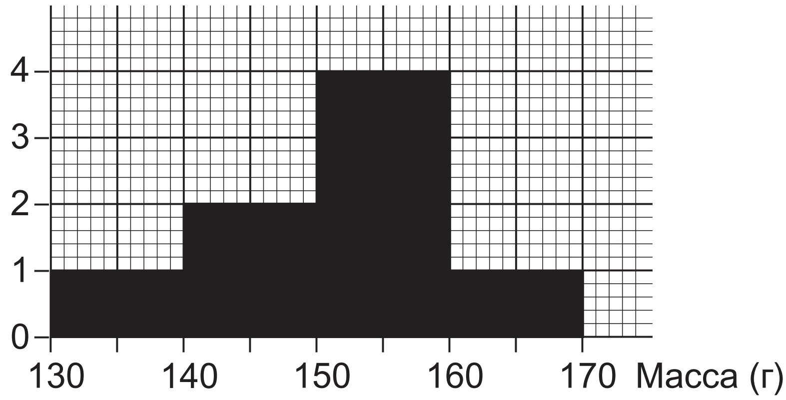
Учащийся	A	B	C	D	E
Математика	23	34	45	56	67
Статистика	32	43	54	65	76
Английский язык	67	56	45	34	23

Не производя вычислений, запишите

(i) коэффициент ранговой корреляции между оценками по математике и статистике;

(ii) коэффициент ранговой корреляции между оценками по математике и английскому языку.

6. Приведенная ниже гистограмма иллюстрирует распределение апельсинов по массе.



(i) Сколько апельсинов весят меньше 140 г?

(ii) Сколько апельсинов весят от 130 до 160 г?

(iii) Сколько апельсинов весят свыше 150 г?

7. Вычислите (i) среднюю арифметическую, (ii) среднее отклонение (от средней) следующего ряда чисел (значений):

$$32, 34, 37, 41, 45, 48, 50.$$

8. Вычислите (i) среднюю арифметическую, (ii) дисперсию, (iii) стандартное отклонение следующего ряда чисел (значений):

$$1, 7, 13, 19, 25, 31.$$

9. Назовите три компоненты, в отношении которых можно анализировать временны́е ряды.

10. На заводе в 1985 и 1986 гг. было произведено 4000 и 5000 деталей соответственно. Если сложившиеся тенденции роста сохранятся, то сколько деталей будет произведено в 1987 г.?

11. На спортивных соревнованиях двое судей, X и Y, выставили 7 атлетам за выступления следующие оценки:

Атлет	A	B	C	D	E	F	G
Судья X	25	18	15	12	14	19	21
Судья Y	23	18	11	16	13	18	22

(i) Ранжируйте атлетов по оценкам, поставленным судьей X.

(ii) Ранжируйте атлетов по оценкам, поставленным судьей Y.

(iii) Рассчитайте коэффициент корреляции рангов Спирмена.

(iv) Прокомментируйте степень согласованности мнений двух судей.

12. Куб с ребром 2 см представляет сумму в $2000. Какой должна быть длина ребра куба, представляющего сумму в $16 000?

13. Ниже даны индексы цен некоторого товара, рассчитанные по методу цепных индексов:

1985	1986	1987
100	108	80

Вычислите простой индекс цен в 1987 г., взяв за базу 1985 г.

14. Вычислите (i) среднюю арифметическую, (ii) среднюю геометрическую трех чисел 15, 60 и 240.

15. Во время рабочего дня компания фиксировала продолжительность телефонных разговоров сотрудников (в сек). Результаты приведены в таблице.

Продолжительность разговора (сек)	0-	25-	50-	75-	100-	125-
Количество звонков	16	22	40	30	12	0

Вычислите:

(i) общее количество звонков;

(ii) число звонков длительностью свыше 100 сек;

(iii) процент (долю) звонков длительностью свыше 1 мин.

16. Из стандартной колоды игральных карт случайным образом вытягивается одна карта. Найдите вероятность того, что эта карта:

(i) масти червы;
(ii) пятерка;
(iii) красной масти;
(iv) картинка;
(v) картинка черной масти.

17. Стандартное отклонение оценок, полученных на экзамене, равно 10. Оценки нормированы, средняя по ним составляет 45, а стандартное отклонение 12. Оценка в 60 баллов была пересчитана в нормированную оценку 75. Вычислите среднее значение исходных оценок.

18. Сосчитайте (i) количество слов и (ii) количество букв в следующем предложении[1]:

«Быстрая рыжая лисичка перепрыгнула через ленивую собаку».

Затем вычислите среднее число букв в слове.

19. Мода некоторого распределения равна 18, среднее значение – 10, а медиана – 14. Дайте статистическое описание формы распределения.

20. Найдите значение *второй* пятичленной скользящей средней данного временного ряда:

21, 28, 31, 29, 36, 51, 43, 46, 50.

21. На секторной диаграмме, иллюстрирующей сбор ячменя, овса и пшеницы на некоторой территории, углы секторов равны соответственно 110°, 135° и 115°. Определите сбор зерна каждого вида, если известно, что валовой сбор составил 7200 кг.

22. Двое судей оценивали четырех участников соревнований. Первый судья распределил места следующим образом:

---

[1] Фраза на английском языке («The quick brown fox jumped over the lazy dog») содержит все 26 букв английского алфавита. – *Прим. ред.*

Участник	A	B	C	D
Место	1	4	3	2

Коэффициент корреляции рангов между оценками двух судей равен –1. Напишите, как распределил места второй судья.

23. Взяв значение показателя в 1984 г. за 100, мы получим значения индекса показателя в 1985 и 1986 гг. соответственно 120 и 156. Вычислите индекс показателя в 1986 г., взяв за 100 значение показателя в 1985 г.

24. Мальчик поднимался в гору в течение 2 ч со скоростью 3 км/ч, а затем спускался с горы 1 ч со скоростью 6 км/ч. Найдите среднюю скорость в пути.

25. Аудиторный тест оценивался из 80 баллов. Стандартное отклонение исходных оценок составило 12.
- (i) Какой была дисперсия оценок?
- (ii) Оценки за тест позже были пересчитаны в нормированные. Найдите стандартное отклонение нормированных оценок.

26. Заявлено, что размеры ковра составляют 5 м и 4 м. Оба результата даны с точностью до метра. Вычислите:
- (i) наименьший возможный периметр ковра;
- (ii) наибольшую возможную площадь ковра.

27. В таблице даны результаты измерения 80 бамбуковых палок с точностью до 10 см:

Длина	20	30	40	50	60	70	80	90
Число палок	0	5	12	19	20	16	8	0

- (i) Запишите длину самой короткой палки.
- (ii) Запишите длину самой длинной палки.
- (iii) Запишите границы модального интервала.
- (iv) Найдите модальную длину бамбуковых палок.

28. Девочка достала карту из стандартной колоды в 52 карты и держит ее в руках. Затем ее сестра вытянула из той же колоды вторую карту. Найдите вероятность того, что карты оказались:

(a) одинаковыми;
(b) одного достоинства;
(c) одного цвета;
(d) разных цветов;
(e) обе королями.

29. В таблице даны результаты измерения длины 80 цилиндров с точностью до 1 см:

Длина	12	13	14	15	16	17	18	19
Число цилиндров	0	5	12	19	20	16	8	0

(i) Запишите средние значения интервалов.
(ii) Вычислите среднюю длину цилиндра (среднюю арифметическую).
(iii) Вычислите дисперсию и стандартное отклонение длины цилиндров.

30. Вступительный взнос в игре составляет \$3. Игрок бросает стандартную шестигранную кость.

Если выпадает 6, то игрок получает \$6.
Если выпадает другое четное число, то игрок получает \$3.
В остальных случаях игрок не получает ничего.

Вычислите:

(i) Ожидаемый выигрыш или проигрыш игрока за один бросок.
(ii) Вступительный взнос, который надо установить, чтобы игра была честной.

Упражнения В

1. Назовите меру средней, которая проще (легче) всего определяется по гистограмме.

2. (i) Проиллюстрируйте следующие данные с помощью гистограммы:

Рост (см)	150-	160-	165-	170-	175-	180-	190-
Число детей	6	4	8	14	12	8	0

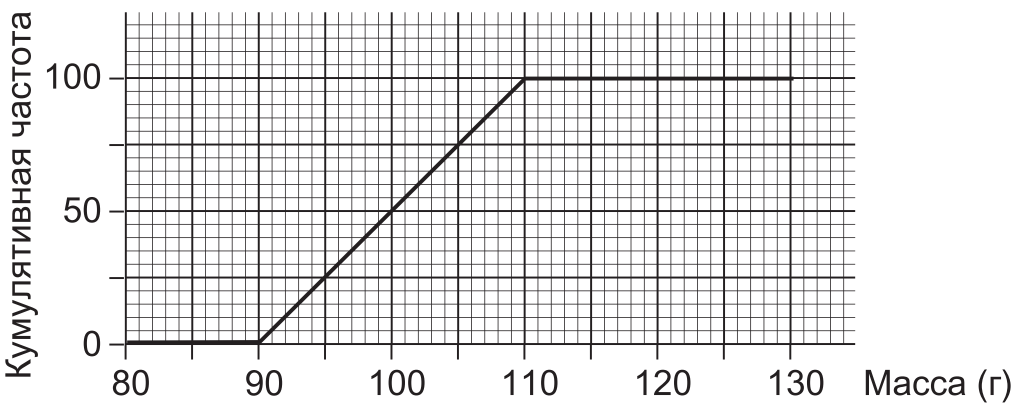
(ii) По гистограмме определите модальный рост.

3. (i) Объясните понятие *«Случайная выборка»*.

(ii) Опишите алгоритм (схему) метода, с помощью которого можно сформировать случайную выборку в 5 учеников из 30 учеников класса.

4. Нарисуйте схематично гистограмму, соответствующую приведенному ниже полигону кумулятивных частот. Укажите названия осей.

5. Назовите два преимущества моды как меры средней.



6. Дайте краткое определение следующих понятий:

(i) временной ряд;
(ii) тренд;
(iii) сезонные колебания;
(iv) периодические (циклические) колебания.

7. Средняя геометрическая чисел x и 4 равна 16. Найдите число x.

8. Найдите: (i) медиану, (ii) моду, (iii) размах, (iv) межквартильный размах приведенного ниже ряда температур, измеренных в °C:

5, 12, 13, –2, 7, 12, 15, 0, 12, 13, 8.

9. В таблице представлены оценки учащихся, полученные за небольшой тест.

Количество баллов	0	1	2	3	4	5	6	7	8	9
Число учащихся	2	6	7	9	13	14	10	7	5	2

Найдите: (i) медиану, (ii) нижний квартиль, (iii) верхний квартиль, (iv) межквартильный размах.

10. В таблице представлены оценки, полученные учащимися по французскому, немецкому языкам и физике.

Учащийся	A	B	C	D	E
Французский язык	79	67	55	43	31
Немецкий язык	91	82	73	64	55
Физика	13	34	44	76	97

Не производя вычислений, запишите

(i) коэффициент ранговой корреляции между оценками по французскому языку и немецкому языку;

(ii) коэффициент ранговой корреляции между оценками по немецкому языку и физике.

11. Вычислите (i) среднюю арифметическую, (ii) дисперсию, (iii) стандартное отклонение следующего ряда чисел (значений):

2, 32, 26, 8, 14, 20.

12. Вычислите (i) среднюю арифметическую, (ii) среднее отклонение (от средней) следующего ряда чисел (значений):

40, 22, 38, 24, 35, 27, 31.

13. Приведите примеры:
- (i) качественной переменной;
- (ii) дискретной количественной переменной;
- (iii) непрерывной количественной переменной.

14. В таблице приведены данные о сроке службы электронной детали двух типов (марок):

Срок службы (ч)	100-	110-	120-	130-	140-	150-	160-	170-
Марка A (количество)	5	13	18	37	54	47	34	22
Марка B (количество)	0	21	55	86	47	21	0	0

- (i) Проиллюстрируйте оба распределения с помощью двух «синхронизированных» гистограмм.
- (ii) Опишите статистическими терминами различия в гистограммах.

15. Куб с ребром 3 см представляет массу в 30 т. Сколько тонн представляет куб с ребром 6 см?

16. (i) Схематично изобразите распределение 50 мальчиков, возраст половины которых 13 лет, а остальных 17 лет.

(ii) Дайте статистическое название такого распределения.

17. На экзаменах по математике и английскому языку 8 студентов получили следующие оценки:

Студент	A	B	C	D	E	F	G	H
Математика	29	38	76	45	50	5	82	91
Английский язык	68	68	34	63	59	57	31	25

- (i) Ранжируйте студентов по оценкам за экзамен по математике.
- (ii) Ранжируйте студентов по оценкам за экзамен по английскому языку.
- (iii) Рассчитайте коэффициент корреляции рангов Спирмена между оценками по математике и английскому языку.
- (iv) Прокомментируйте корреляцию между двумя рядами оценок.

18. Ниже даны индексы цен некоторого товара, рассчитанные по методу (системе) цепных индексов:

1985	1986	1987
100	90	120

Вычислите простой индекс цен в 1987 г., взяв за базу 1985 г.

19. Объясните, что значит *коэффициент регрессии*?

20. Вычислите (i) среднюю арифметическую, (ii) среднюю геометрическую трех чисел 750, 300 и 15.

21. В борьбе за режим экономии компания фиксировала продолжительность телефонных разговоров сотрудников (в сек). Результаты приведены в таблице.

Продолжительность разговора (сек)	0-	20-	40-	60-	80-	100-
Количество звонков	10	16	40	24	14	0

(i) Проиллюстрируйте распределение с помощью гистограммы.
(ii) Найдите модальное значение по гистограмме.
(iii) Проверьте ответ вычислениями (расчетами).

22. Мода некоторого распределения равна 8, среднее значение 10, а медиана 12. Дайте статистическое описание формы распределения.

23. Средний балл, полученный на экзамене, равен 30, а стандартное отклонение 5. Оценки были нормированы таким образом, что новое стандартное отклонение оказалось равным 8. Оценке в 40 баллов соответствует нормированная оценка 56. Вычислите среднее значение нормированных оценок.

24. Подсчитайте

(i) количество букв в каждом слове в предложении

«Теперь самое время для всех надежных, кредитоспособных людей прийти на помощь партии».

(ii) Постройте распределение частот числа букв в слове;

(iii) Найдите медианное значение числа букв в слове.

25. Найдите значение *третьей* четырехчленной скользящей средней данного временного ряда:

17, 21, 23, 31, 33, 33, 27, 35, 37, 33.

26. Требуется построить две секторные диаграммы, иллюстрирующие посевные площади, отведенные под зерновые культуры в двух годах.

Зерновая культура	Площадь (га)	
	1985	1986
Ячмень	66	80
Овес	75	104
Пшеница	129	136

(i) Рассчитайте углы для построения секторной диаграммы за 1985 г.

(ii) Рассчитайте углы для построения секторной диаграммы за 1986 г.

(iii) Радиус диаграммы за 1985 г. положите равным точно 5 см. Вычислите радиус сопоставимой секторной диаграммы за 1986 г.

(iv) Нарисуйте обе секторные диаграммы.

27. Взяв значение показателя в 1985 г. за 100, мы получим значения индекса показателя в 1986 г. и 1987 г. соответственно равными 110 и 132. Вычислите индекс показателя в 1987 г., взяв за 100 значение показателя в 1986 г.

28. Девочка поднималась в гору в течение 1½ ч со скоростью 2,8 км/ч, а затем спускалась с горы 1 ч со скоростью 5 км/ч. Найдите среднюю скорость в пути.

29. Аудиторный тест оценивался из 75 баллов. Дисперсия нормированных оценок составила 64.

(i) Найдите стандартное отклонение нормированных оценок.

(ii) Найдите стандартное отклонение исходных оценок.

30. (i) Объясните кратко, что понимается под *тенденциозной выборкой*;

(ii) Приведите пример формирования тенденциозной выборки.

31. Автомобиль преодолел 140 км за 2 ч. Известно, что расстояние указано с точностью до 10 км, а время – до 10 мин.

Вычислите:

(i) наибольшую возможную среднюю скорость автомобиля в км/мин;

(ii) наименьшую возможную среднюю скорость автомобиля в км/ч.

32. Дайте определения следующих понятий:

a) случайная выборка;

b) систематическая выборка;

c) стратифицированная выборка;

d) долевая выборка .

33. Приведите примеры:

(i) качественной переменной;

(ii) дискретной количественной переменной;

(iii) непрерывной количественной переменной.

34. Изобразите схематично следующие непрерывные распределения:

(a) с правосторонней асимметрией;

(b) с левосторонней асимметрией;

(c) бимодальное;

(d) гиперболическое;

(e) экспоненциальное.

35. Ниже приведены данные за 1986 г. о количестве студентов технического колледжа, изучавших следующие дисциплины: математика – 16, статистика – 8, физика – 14, химия – 10, биология – 10, геология – 4.

Проиллюстрируйте данные с помощью столбиковой диаграммы.

36. В таблице приведены данные опроса учеников о расстоянии от дома до школы. Результаты представлены с точностью до километра.

Расстояние, км	1	2	3	4	5
Число учеников	6	9	12	3	0

Проиллюстрируйте распределение с помощью полигона частот.

37. Мальчик достал карту из стандартной колоды в 52 карты. Карту вернули в колоду. Затем карту вытянул его брат. Найдите вероятность того, что карты оказались:

(a) одинаковыми;

(b) одного достоинства;

(c) одного цвета;

(d) разных цветов;

(e) обе королями червей.

38. В таблице даны результаты измерения длины 160 цилиндров с точностью до сантиметра:

Длина, см	22	23	24	25	26	27	28	29
Число цилиндров	0	15	22	39	40	26	18	0

(i) Запишите минимальную из возможных длину цилиндра.

(ii) Запишите длину средней партии.

(iii) Найдите арифметическую среднюю длины цилиндров.

(iv) Найдите дисперсию и стандартное отклонение.

39. Вступительный взнос в игре составляет \$2. Игрок вытягивает карту из стандартной колоды в 52 карты.

Если игрок достает туз, то получает \$6,50.

Если вытягивает картинку, то получает \$3,25.

В остальных случаях игрок не получает ничего.

Вычислите:

(i) Ожидаемый выигрыш или проигрыш игрока за одну карту.

(ii) Вступительный взнос, который надо установить, чтобы игра была честной.

ПОДРОБНЫЕ ОБЗОРНЫЕ ВОПРОСЫ

1. В таблице приведены данные о сроке службы электронной детали двух типов (марок):

Срок службы (ч)	110-	120-	130-	140-	150-	160-	170-	180-
Марка А (количество)	22	34	47	54	37	18	13	5
Марка В (количество)	0	0	21	47	86	55	21	0

(i) Проиллюстрируйте оба распределения с помощью полигонов кумулятивных частот, построив их на одной диаграмме.
(ii) Опишите статистическими терминами различия в распределениях.
(iii) Дайте рекомендации потенциальным покупателям.

2. В таблице приведены значения двух переменных x и y.

x	7,0	10,8	15,0	9,0	13,0	11,2
y	20,0	32,5	44,0	26,0	38,0	31,5

(i) Вычислите коэффициент корреляции рангов Спирмена между переменными x и y.
(ii) Постройте диаграмму рассеивания y от x.
(iii) Вычислите среднюю арифметическую значений x и y, а также полусредние значения.
(iv) Проведите линию наилучшего соответствия.
(v) Найдите уравнение прямой $y = Ax + B$, где A и B определите с точностью до 2 значащих цифр.
(vi) По уравнению найдите значение y при $x = 10,0$.
(vii) Проверьте по графику ответ, полученный в п. (vi).

3. В таблице приведены результаты измерения 100 металлических стержней, отобранных из партии по подозрению в несоответствии требованиям.

Длина стержня (см)	26	27	28	29	30	31	32
Число стержней	1	4	23	48	20	3	1

(i) Постройте огиву (кривую кумулятивных частот).

(ii) По огиве найдите

a) медиану;

b) нижний квартиль;

c) верхний квартиль;

d) межквартильный размах.

(iii) Стержень может использоваться по назначению, если только его длина находится в границах допустимого интервала от 27,5 до 30,5 см. По огиве оцените долю бракованных стержней, которые должны пойти на переработку.

4. В таблице приведены данные о еженедельных продажах (в $) для двух магазинов L и M.

	1984			1985			1986		
	Январь – Апрель	Май – Август	Сен-тябрь–	Январь – Апрель	Май – Август	Сен-тябрь–	Январь – Апрель	Май – Август	Сен-тябрь –
Магазин L	650	600	700	1100	1050	1150	1550	1500	1600
Магазин M	900	630	720	1050	780	870	1200	930	1020

(i) На одной диаграмме постройте два графика, иллюстрирующие данные (масштаб 2 см $\Rightarrow$ \$200 по одной оси и 2 см $\Rightarrow$ 4 месяца по другой).

(ii) Определите число звеньев скользящей средней для выделения тенденции во временных рядах.

(iii) Для каждого магазина найдите тренд и постройте линию тренда на графике.

(iv) Установите, у какого из магазинов дела идут лучше. Обоснуйте ответ.

(v) Деятельность какого магазина больше подвержена сезонным колебаниям? Обоснуйте ответ.

(vi) Оцените средний недельный объем продаж в магазине L в период январь – апрель 1987 г.

5. В таблице приведены данные по двум районам R и S.

Возрастная группа	Район R		Район S		Стандартное население
	Население	Количество смертей	Население	Смертность на 1000 человек	
0-	500	4	30%	9	35%
25-	750	3	35%	7	30%
45-	1000	5	25%	6	25%
65-	750	12	10%	11	10%

Рассчитайте:

(i) коэффициент смертности в районе R;

(ii) коэффициент смертности для каждой возрастной группы в районе R;

(iii) стандартизованный коэффициент смертности в районе R;

(iv) количество смертей в каждой возрастной группе в районе S;

(v) коэффициент смертности в районе S;

(vi) стандартизованный коэффициент смертности в районе S.

Дайте обоснованный ответ на вопрос, какой район обеспечивает лучшие условия для долгожительства.

6. На грани красной стандартной шестигранной кости нанесены цифры 0, 1, 1, 2, 2 и 2. Грани синей кости пронумерованы аналогично. Обе кости бросают одновременно. Счет в игре определяется суммой цифр на выпавших гранях.

(i) Составьте таблицу размером 6×6, иллюстрирующую все 36 возможных исходов.

(ii) Просчитайте вероятность всех возможных исходов.

(iii) Запишите вероятность результата 0, 1 и 2 при бросании одной кости.

(iv) Постройте граф, на котором проиллюстрируйте все возможные исходы при бросании двух костей.

(v) Проверьте по графу ответ на вопрос п. (ii).

ЧИСЛЕННЫЕ ОТВЕТЫ К УПРАЖНЕНИЯМ, ВОПРОСАМ, ТЕСТАМ

Хотя все ответы тщательно проверялись, автор не может гарантировать точность каждого ответа.

Стр. 21 **Упражнение 2А**

1	114°, 105°, 53°, 88°
2	180°, 71°, 108°, 1°
3	38°, 94°, 43°, 29°, 22°, 121°, 13°
4	123°, 7°, 68°, 162°
5	22°, 97°, 43°, 14°, 29°, 155°

Стр. 22 **Упражнение 2В**

1	40°, 100°, 60°, 80°, 60°, 20°
2	28°, 15°, 2°, 83°, 8°, 224°
3	50°, 75°, 25°, 42°, 9°, 159°
4	40°, 32°, 76°, 36°, 72°, 104°
5	72°, 72°, 90°, 108°, 18°

Стр. 23 **Упражнение 3А**

1. Суша 33°, 234°, 53°, 13°, 27°
Океан 60°, 276°, 24°
R : r = 1,34 : 1

2.
(i) 81°, 141°, 30°, 108°
(ii) 423 000, 243 000, 324 000, 90 000
(iii) 1,9 см
(iv) 1,6 см
(v) 90°, 149°, 40°, 81°
(vi) значительный рост (80%) для остальных стран

Стр. 28 **Упражнение 3B**

1. (i) 1,9 см

(ii)

60°	120 млн
45°	90 млн
180°	360 млн
50°	100 млн
25°	50 млн

(iv) 2,15 см

(v) 70°, 63°, 149°, 55°, 23°

2. (i) 160°, 130°, 70°

(ii) (a) 400, (b) 325, (c) 176

(iii) 2 см

(iv) 2,3 см

(v) 234°, 54°, 72°

(vi) одна и та же часть учащихся ездит в школу на автобусе;
больше ходит пешком;
меньше добирается на велосипеде.

Стр. 39 **Упражнение 7**

1.	Количество голов	0	1	2	3	4	5
2.	Частота	7	13	6	3	0	1

Стр. 47 **Упражнение 8A**

1. Высота столбца 1; 2; 4; 3,2; 2,4; 0,6

2. Высота столбца 1; 2; 3; 6; 4,5; 0,5

Стр. 47 **Упражнение 8B**

1. Высота столбца 60, 90, 100, 175, 75

2. Высота столбца 9, 14, 7, 4, 1

Стр. 58 **Упражнение 9A**

1. Кумулятивная частота 0, 0, 6, 14, 27, 46, 80

2. P1: 0, 12, 32, 60, 96, 116, 132, 154, 174, 200, 200

P2: 0, 0, 0, 0, 0, 0, 40, 88, 152, 188, 200.

Стр. 59 **Упражнение 9B**

1.

«До», или «меньше чем» (км/ч)	0	10	12	14	15	16	18	20	25
Кумулятивная частота	0	0	4	14	21	25	31	35	40

2. «До», или

«меньше чем» (км/ч)	3500	4000	4200	4400	4600	5000
Кумулятивная частота	0	12	32	56	92	100

Стр. 75 Упражнение 10А

1. 6

2.

Значение	7	8	9	10	11	12
Частота	2	3	5	4	3	1

Мода = 9

3.

Число букв в слове	1	2	3	4	5	6	7	8	9	10
Частота	1	0	0	0	1	0	0	3	0	2

Мода = 8

Стр. 75 Упражнение 10В

1. 8

2.

Значение	6	7	8	9	10	11	12
Частота	1	3	4	5	7	6	4

Мода = 10

3.

Число букв в слове	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Частота	1	0	0	1	2	2	3	0	0	0	0	1	1	1

Модальное число букв в слове = 7

Стр. 81 Упражнение 11А

1. (a) 3800, 3900 Мода = 3840
(b) 3841

2. (a) 0; 9,5; 19,5; 24,5; 29,5; 30,5; 34,5; 39,5; 50,5
(c) 25,7 мин
(d) 25,75 мин

Стр. 81 Упражнение 11В

1. (b) 40,50
(c) 44 мин
(d) 44,3 мин

2. (c) 17,5; 20,5
(d) 19,1 м
(e) 19,14 м

Стр. 98 Упражнение 12A

1. (a) Кумулятивные частоты:
0, 480, 930, 1350, 1950, 2100
(c) 34
(d) 34,29

2. (a) 4,5
(b) Кумулятивные частоты:
0, 6, 16, 29, 45, 71, 89, 100
(d) 20,5 км
(e) 20,46 км

Стр. 99 Упражнение 12B

1. (a)

«До», или «меньше чем»	20	22	24	25	26	2830
Кумулятивная частота	0	3	8	15	24	3032

(c) 25,1 см
(d) 25,11 см

2. (a) 4,5
(b)

«До», или «меньше чем»	4,5	9,5	14,5	19,5	24,5	29,5
Кумулятивная частота	3	11	24	39	46	50

(d) 4,5
(e) 14,83

Стр. 115 Упражнение 13A

1. (i) 2,5 7,5 12,5 17,5 22,5
(ii) 12,1

2. (i) 0 3,5 1,75
(ii) 3,5 6,5 5
(iii) 4,07

Стр. 116 **Упражнение 13В**

1.	(i)	90	100	110	120	130
	(ii)	107,73				
2.		7,23				

Стр. 122 **Упражнение 14А**

1.	63	**2.**	44	**3.**	3,5	**4.**	49,91
5.	89,99	**6.**	49,96	**7.**	40,93	**8.**	56,91
9.	84,90	**10.**	64	**11**.	36	**12**.	36,59
13.	38,26	**14.**	30	**15.**	22,89		

Стр. 122 **Упражнение 14В**

1.	42	**2.**	45	**3.**	2,7	**4.**	33,87
5.	68,97	**6.**	69,82	**7.**	43,93	**8.**	57,86
9.	69,52	**10.**	81	**11**.	75	**12**.	43,09
13.	46,42	**14.**	40	**15.**	38,26		

Стр. 130 **Упражнение 15А**

1.	8	**2.**	5	**3.**	3	**4.**	3½
5.	4½	**6.**	4¼	**7.**	3½	**8.**	8
9.	11	**10.**	10				

Стр. 130 **Упражнение 15В**

1.	9	**2.**	7	**3.**	5	**4.**	6½
5.	3	**6.**	7	**7.**	6	**8.**	6½
9.	6¾	**10.**	12				

Стр. 136 **Упражнение 16А**

1.	(b)	22,9 года	53,5 года	30,6 года
2.	(b)	13,5 км	26,4 км	12,9 км

Стр. 137 **Упражнение 16В**

1.	(b)	7,13 см	9,14 см	2,01 см
2.	(b)	7,2 см	8,9 см	1,7 см

Стр. 141 Упражнение 17A

1.	10,4	19,1	49
2.	15,3	22,9	28,1

Стр. 141 Упражнение 17B

1.	7,8	8,9	9,9
2.	6,4	7,4	9,2

Стр. 144 Упражнение 18A

1.	5	3,14	**2.**	8	3,2	**3.**	7	1,2
4.	3	1,6	**5.**	6	2,89	**6.**	8	2,2
7.	6	2,25	**8.**	10	5,5	**9.**	7	6,67
10.	10	6,57						

Стр. 144 Упражнение 18B

1.	4	2,67	**2.**	6	3,56	**3.**	7	3
4.	3	1,71	**5.**	7	1,8	**6.**	6	4,8
7.	8	3	**8.**	7	3,78	**9.**	9	3,33
10.	11	6,8						

Стр. 147 Упражнение 19A

1.	5	0,83	**2.**	7	0,87	**3.**	3	0,92
4.	5,5	0,83						

Стр. 148 Упражнение 19B

1.	7	0,86	**2.**	6	0,90	**3.**	4	0,91
4.	6,5	0,86						

Стр. 153 Упражнение 20А

1.	5	6,67	2,58
2.	6	18	
3.	5,25	2,29	
4.	8,85	2,97	
5.	12,86	3,59	
6.	9,14	3,02	
7.	7,80	2,79	

Стр. 154 Упражнение 20B

1.	5	6,67	2,58
2.	7	16,8	4,10
3.	6,75	2,60	
4.	10,25	3,20	
5.	7,65	2,77	
6.	1,43	1,20	
7.	13,43	3,66	

Стр. 158 Упражнение 21А

1.	6	1,25	1,12
2.	4	1,38	1,18
3.	0,7	0,02	0,12
4.	7,5	1,03	1,01
5.	1,33	1,15	
6.	1,62	1,27	
7.	0,017	0,13	
8.	1,54	1,24	
9.	1,58	1,26	
10.	1,75	1,32	
11.	0,014	0,12	
12.	1,36	1,17	
13.	1,36	1,17	
14.	1,44	1,20	
15.	0,014	0,12	
16.	1,36	1,17	

Стр. 159 Упражнение 21В

1.	9	1,33	1,15
2.	9	1,38	1,17
3.	0,3	0,014	0,12
4.	5,5	1,04	1,02
5.	1,25	1,12	
6.	1,38	1,18	
7.	0,015	0,12	
8.	1,00	1,00	
9.	1,05	1,03	
10.	1,38	1,17	
11.	0,016	0,13	
12.	1,04	1,02	
13.	1,19	1,09	
14.	1,78	1,34	
15.	0,019	0,14	
16.	1,32	1,15	

Стр. 165 Упражнение 22А

1.	1039,35	32,24
2.	0,097	0,31
3.	4142,42	64,36
4.	0,050	0,22

Стр. 165 Упражнение 22В

1.	0,016	0,13
2.	0,016	0, 13
3.	28,67	5,35
4.	0,011	0,10

Стр. 175 Упражнение 23А

1.	56	30
2.	38	74
3.	8	

Стр. 176 Упражнение 23В

1.	44	55
2.	28	60
3.	12	

Стр. 0189 Упражнение 24А

1.	(ii)	5,9	6,3	6,7	7,1	7,5	7,9	8,3
	(iv)	+0,1	–0,3	+0,2				
	(v)	9,2 (1000 $м^3$)						
2.	(ii)	2840	2760	2680	2520	2440	2360	2280
	(iv)	–15	–25	+10	+30			
	(v)	495						

Стр. 190 Упражнение 24В

1.	(ii)	180	213	243	273	300	330	360	
	(iv)	–21	+8	–6					
2.	(ii)	680	720	760	800	840	880	920	960
	(iv)	+7	+3	–2	–8				
	(v)	277							

Стр. 193 Упражнение 25А

1. 2,7 км/ч
2. 54
3. 8,2 $г/см^3$
4. 18 ф.ст. за кг

Стр. 194 Упражнение 25В

1. 15,7 км/ч
2. 56,4
3. 8,16 $г/см^3$
4. 25 ф.ст. за кг

Стр. 198 Упражнение 26A

1.	125	80		
2.	120	108	120	90
3.	$9,60	$10,56	88	

Стр. 199 Упражнение 26B

1.	120	83		
2.	120	80	120	50
3.	$7,80	$3,90	308	

Стр. 202 Упражнение 27A

1.	105,10	
2.	91,6	109,03
3.	X = 107	

Стр. 203 Упражнение 27B

1.	104,6	
2.	96,8	103,06
3.	X = 100	

Стр. 210 Упражнение 28A

1.	7,4 на 1000	8, 6,10	7,7 на 1000
	7,5 на 1000	6,7 на 1000	E: более низкий стандартизированный коэффициент смертности
2.	20, 45, 39	10,4 на 1000	10,79 на 1000
	10,3 на 1000	10, 9, 14	10,65

Лучшие условия жизни в регионе S: более низкий стандартизированный коэффициент смертности.

Стр. 211 Упражнение 28B

1.	8,15 на 1000	8,08 на 1000	8,3 на 1000
	7, 6, 10	7,30 на 1000	N: более низкий стандартизированный коэффициент смертности
2.	9, 8, 11	8,9 на 1000	9,45 на 1000
	18, 52, 18	8,8 на 1000	9,80 на 1000

Немного лучше условия жизни в городе J: более низкий стандартизированный коэффициент смертности.

Стр. 223 Упражнение 29A

1. (21,7; 35,6) (17,3; 29) (26; 42,2)
$y = 1{,}5x+2{,}5$ 29,5

2. B, F, G, I, J, L A, C, E, H, K D
(3,8; 10) 10,6 г 6 см^3 2,7 г/см^3

Стр. 224 Упражнение 29B

1.	(ii)	(2,2; 7,9)	(1,9; 6,9)	(2,4; 8,6)
	(iii)	Не отведена земля под культуру – нет и урожая		
	(v)	$y = 3{,}6x$		
	(vi)	Средняя урожайность с 1 га (в т)		
2.	(ii)	(8,3; 11,7)	(5,7)	(12,5; 17,5)
	(iii)	$y = 1{,}4x$		
	(v)	дискретные данные		
	(vi)	Да, так как y возрастает с ростом х		
	(vii)	y		

Стр. 242 Упражнение 30A

1.	1/15	7/15	7/15				
2.	4/15	19/30	1/10				
3.	P(2) = 1/36	P(6)=5/36	P(9)=4/36	7			
4.	первое вытягивание карты			1/13	3/13	5/13	1/2
	повторное вытягивание карты			1/169	1/2	1/13	1/2

Стр. 243 **Упражнение 30B**

1.	25/144	35/72	49/144				
2.	1/7	4/7	2/7				
3.	(i)						

	1	**2**	**3**	**4**	**5**	**6**
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

	(ii)	1/18	1/9	1/9	6 или 12	
4.	(a)	1/13	5/13	3/13	1/2	
	(b)	1/221	25/51	11/221	95/663	26/51

Стр. 255 **Упражнение 31A**

1. 5408 6300
2. 51,6 км/ч 39,9 км/ч
3. 7%

Стр. 255 **Упражнение 31B**

1. 46,75 м^2 33,75 м^2
2. 6%
3. 2% 4% 6%

Стр. 271 **Упражнение 33A**

1. 0,3212 имеется положительная корреляция
2. –0,8393 сильная отрицательная корреляция

Стр. 272 **Упражнение 33B**

1. 0,3857 имеется положительная корреляция
2. 0,994 очень сильная, почти полная корреляция

СТР. 273 КРАТКИЕ ОБЗОРНЫЕ ВОПРОСЫ – Упражнения А

1.	Медиана							
2.	45							
3.	7	7	19	2	8	6		
4.	4	3	6	3				
5.	1	–1						
6.	1	7	5					
7.	41	5,7						
8.	16	105	10,25					
9.	тренд, сезонные колебания, случайные колебания, циклические (периодические) колебания							
10.	6250							
11.	A	B	C	D	E	F	G	
	1	4	5	7	6	3	2	
	1	3½	7	5	6	3½	2	
	(iii)	0,8482						
	(iv)	сильная положительная корреляция (очень хорошая согласованность)						
12.	4 см							
13.	86,4							
14.	105	60						
15.	120	12	55%					
16.	1/4	1/13	1/2	3/13	3/26			
17.	35							
18.	7							
19.	с левосторонней асимметрией							
20.	35							
21.	2200 кг	2700 кг	2300 кг					
22.	4-е	1-е	2-е	3-е				
23.	130							
24.	4 км/ч							
25.	144	15						
26.	16 м	24,75 м						
27.	15 см	95 см	55 см	65 см	57 см			
28.	0	1/17	25/51	26/51	1/221			
29.	12	13	14	15	16	17	18	19
	(ii)	15,7 см			(iii)	1,89 см^2		1,38 см
30.	(i)	$1 проигрыш			(ii)	2 ф.ст.		

СТР. 279 КРАТКИЕ ОБЗОРНЫЕ ВОПРОСЫ – Упражнения В

1. Мода
2. 173,8 см
7. 64
8. 12°C 12°C 17°C 7¾°C
9. 5 3 6 3
10. +1 –1
11. 17 105 10,25 (до 2 знаков после запятой)
12. 31 5,71
15. 240 т
16. (ii) бимодальное распределение
17.

A	B	C	D	E	F	G	H
8	7	3	6	4½	4½	2	1
1½	1½	6	3	4	5	7	8

(iii) –0,9643
(iv) Высокие экзаменационные оценки по английскому языку соответствуют низким экзаменационным оценкам по математике
Значительная отрицательная корреляция

18. 133,3
20. 355 150
21. (ii) 52 сек (iii) 52 сек
22. с правосторонней асимметрией
23. 40
24. (i) 3 2 3 4 3 3 4 3 2 4 2 3 3 2 3 5
(ii)

L	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
F	0	1	1	1	3	4	0	1	0	0	0	0	0	0	0	1

(iii) медиана = 6
25. 30
26. (i) 88° 100° 172°
(ii) 90° 117° 153°
(iii) 5,4 см (до 1 знака после запятой)
27. 120
28. 3,68 км/ч
29. 8 6
31. 1,26 км/мин 1,08 км/мин

37. 1/52 1/13 1/2 1/2 1/2704

38. (i) 21,5 см

(ii) 22 23 24 25 26 27 28 29

(iii) 25,6 см (до 2 знаков после запятой)

(iv) 2,08 см^2 1,44 см

39. проигрыш $0,75 $1,25

СТР. 287 ПОДРОБНЫЕ ОБЗОРНЫЕ ВОПРОСЫ

2. (i) 0,9429

(iii) (11; 32) (8,9; 26,2) (13,1; 37,8)

(v) $y = 2{,}8x+1{,}4$

(vi) 29,4

3. (ii) 28,96 см 28,37 см 29,48 см 1,11 см

(iii) 9%

4. 3 точки L M $2000

5. (i) 8 на тысячу

(ii) 8 4 15 16 на тысячу

(iii) 6,85 на тысячу

(iv) Соотношения 27 : 24,5 : 15 : 11

(v) 7,75 на тысячу

(vi) 21,65 на тысячу

Район R, там более низкий стандартизированный коэффициент смертности.

6. (i)

	0	**1**	**1**	**2**	**2**	**2**
0	0	1	1	2	2	2
1	1	2	2	3	3	3
1	1	2	2	3	3	3
2	2	3	3	4	4	4
2	2	3	3	4	4	4
2	2	3	3	4	4	4

(ii)

x	**0**	**1**	**2**	**3**	**4**
P(x)	1/36	4/36	10/36	12/36	9/36

(iii) P(0)=1/36 P(1)=2/6 P(1)=3/6

ОТВЕТЫ К АУДИТОРНЫМ ТЕСТАМ

1. (Стр. 14)

1	C	2	C	3	A
4	D	5	C	6	B
7	C	8	D	9	C
10	D	11	A	12	C
13	D	14	C	15	C
16	D	17	D	18	C
19	A	20	C		

2. (Стр. 46)

1	A	2	D	3	C
4	E	5	C	6	E
7	B	8	C	9	D
10	E	11	C	12	C

3. (Стр. 71)

1	B	2	C	3	A
4	D	5	D	6	B
7	E	8	E	9	D
10	A	11	D	12	E
13	C	14	A	15	C
16	E	17	C	18	D
19	B	20	C		

4. (Стр. 72)

1	A	2	B	3	C
4	B	5	C	6	C
7	D	8	B		

5. (Стр. 85)

1	D	2	C	3	D
4	B	5	D	6	A
7	C	8	B	9	A
10	D	11	C	12	B
13	B	14	B	15	D
16	D	17	E	18	C
19	D	20	A		

6. (Стр. 88)

1	C	2	B	3	D
4	D	5	D	6	A
7	C	8	D		

7. (Стр. 101)

1	B	2	A	3	C
4	E	5	D	6	C
7	B	8	D	9	B
10	C	11	E	12	A
13	D	14	B	15	B
16	E	17	D	18	D
19	E	20	C		

8. (Стр. 106)

1	C	2	B	3	D
4	E	5	A	6	B
7	D	8	C		

9. (Стр. 114)

1	B	2	A	3	C
4	C	5	B	6	B
7	D	8	C	9	D
10	C	11	B	12	C
13	E	14	B	15	B
16	C	17	B	18	E
19	E	20	B		

10. (Стр. 118)

1	C	2	C	3	C
4	C	5	B	6	B
7	B	8	B	9	B
10	A	11	A	12	A
13	B	14	A	15	A
16	E	17	C	18	C
19	E	20	A		

11. (Стр. 121)

1	B	2	D	3	C
4	A	5	B	6	C
7	B	8	B	9	A
10	A	11	A	12	E
13	E	14	C	15	A
16	A	17	B	18	C
19	D	20	B		

12. (Стр. 126)

1	D	2	D	3	D
4	E	5	E	6	D
7	E	8	C	9	D
10	D	11	B	12	E
13	D	14	C	15	E
16	D	17	B	18	A
19	D	20	E		

13. (Стр. 129)

1	B	2	C	3	A
4	B	5	D	6	C
7	A	8	C	9	E
10	B	11	D	12	C
13	B	14	C	15	E
16	E	17	B	18	A
19	B	20	B		

14. (Стр. 139)

1	C	2	C	3	D
4	B	5	D	6	D
7	D	8	D	9	C
10D	11B	12D			
13	A	14	C	15	C
16	D	17	A	18	A
19	C	20	D		

15. (Стр. 143)

1	B	2	E	3	C
4	C	5	B	6	E
7	D	8	B	9	A
10	A	11	D	12	C
13	A	14	E	15	E
16	C	17	B	18	B
19	C	20	B		

16. (Стр. 174)

1	E	2	B	3	D
4	A	5	C	6	E
7	D	8	A	9	B
10	D	11	E	12	A
13	D	14	A	15	E
16	A	17	E	18	A

17. (Стр. 187)

1	B	2	A	3	C
4	D	5	A	6	B
7	A	8	B	9	A
10	C				

18. (Стр. 196)

1	E	2	D	3	C
4	C	5	C	6	E
7	C	8	D	9	E
10	D	11	E	12	D
13	E	14	E	15	C
16	D	17	B	18	D

19. (Стр. 208)

1	B	2	C	3	D
4	C	5	B	6	B
7	C	8	D	9	D
10	B	11	E	12	E
13	E	14	E	15	C
16	A	17	B	18	D
19	C	20	B		

20. (Стр. 221)

1	A	2	E	3	C
4	D	5	B	6	B
7	A	8	C	9	D
10	E				

21. (Стр. 227)

1	A	2	B	3	C
4	D	5	D	6	A
7	C	8	B	9	C
10	C	11	A	12	D
13	D	14	C	15	B
16	A	17	A	18	C
19	D	20	B		

22. (Стр. 230)

1	A	2	E	3	D
4	E	5	D	6	C
7	C	8	A	9	D
10	C	11	B	12	A
13	E	14	B	15	E
16	E	17	D	18	D

23. (Стр. 232)

1	B	2	D	3	C
4	B	5	C	6	C
7	B	8	E	9	D
10	D	11	A	12	C

24. (Стр. 235)

1	E	2	C	3	C
4	A	5	C	6	D
7	E	8	E	9	C
10	A	11	B	12	D

25. (Стр. 246)

1	B	2	C	3	A
4	D	5	B	6	C
7	E	8	A	9	E
10	C	11	D	12	B
13	A	14	D	15	C
16	E				

26. (Стр. 254)

1	D	2	A	3	E
4	A	5	E	6	A
7	E	8	A	9	E
10	A	11	D	12	A
13	D	14	B	15	D
16	B	17	D	18	D
19	C	20	E		

27. (Стр. 270)

1	C	2	A	3	D
4	B	5	E	6	E
7	B	8	D	9	A
10	C	11	A	12	C
13	B	14	E	15	D
16	B	17	C	18	C
19	E	20	D		

Учебное издание

Алик Хартли

СТАТИСТИКА. ПЕРВАЯ КНИГА

Ответственный за выпуск *А.К. Бурцев*
Редактор *Е.В. Васильевская*
Младший редактор *Т.В. Артемова*
Художественный редактор *Г.Г. Семенова*
Технический редактор *В.Ю. Фотиева*
Корректоры *Н.Б. Вторушина, Г.В. Хлопцева*
Компьютерная верстка *Е.Ф. Тимохиной*
Оформление художника *А.В. Алексеева*

ИБ № 4633

Подписано в печать 16.04.2004
Формат 60x88 $^{1}/_{16}$. Гарнитура «Таймс». Печать офсетная
Усл. п.л. 19,11. Уч.-изд. л. 00,00. Тираж 3000 экз.
Заказ № «С» 119

Издательство «Финансы и статистика»
101000, Москва, ул. Покровка, 7
Телефон (095) 925-35-02. Факс (095) 925-09-57
E-mail: mail@finstat.ru http://www.finstat.ru

Для заметок