Министерство науки и высшего
образования Российской Федерации

Уральский федеральный университет имени первого Президента России Б.Н.Ельцина

О.С. Кеткина

СТАТИСТИКА: СБОРНИК ЗАДАНИЙ С ПРИМЕРАМИ РЕШЕНИЙ И ПОЯСНЕНИЯМИ

Учебное электронное текстовое издание

Практикум содержит примеры решения типовых задач с пояснениями, задачи для самостоятельного разбора с ответами, а также краткий теоретический материал по следующим разделам статистки: описательная статистика, аналитическая статистика, и статистика в прикладных исследованиях. Материалы предназначены для самостоятельной подготовки студентов к тестовым и контрольным мероприятиям по дисциплине «Статистика».

Издание подготовлено для студентов, обучающихся по программе бакалавриата по направлению подготовки 38.03.05 «Бизнес-информатика».

Подготовлено кафедрой Эконометрики и статистики

Екатеринбург
2019

Оглавление

Глава 1. ОПИСАТЕЛЬНАЯ СТАТИСТИКА

§ 1. Виды средних величин

Средняя величина – это обобщающий показатель, характеризующий типический уровень явления в конкретных условиях места и времени. Он отражает величину признака, отнесенную к единице совокупности.

Средняя всегда обобщает количественную вариацию признака, т.е. в средних величинах погашаются индивидуальные различия единиц совокупности, обусловленные случайными обстоятельствами[1].

В статистике выделяют степенные и структурные средние. Под структурными средними понимают моду и медиану. Степенные, в свою очередь, бывают простыми и взвешенными.

Простая средняя исчисляется по несгруппированным данным:

$$\bar{X} = \sqrt[m]{\frac{\sum_{i=1}^{n} x_i^m}{n}},$$

где x_i – значение осредняемого признака;

m – показатель степени средней;

n – число значений осредняемого признака.

Взвешенная средняя исчисляется по сгруппированным данным, представленным в виде дискретных или интервальных рядов распределения:

$$\bar{X} = \sqrt[m]{\frac{\sum_{i=1}^{k} x_i^m f_i}{\sum_{i=1}^{k} f_i}},$$

где x_i – значение осредняемого признака или серединное значение интервала, в котором измеряется признак;

m – показатель степени средней;

f_i – частота, показывающая сколько раз в совокупности встречается i-е значение осредняемого признака.

---

[1] См.: Харченко Л.П., Долженкова В.Г., Ионин В.Г. и др. Статистика : учебное пособие. – М. : ИНФРА-М, 2002. С. 90.

В зависимости от показателя степени средней (m) получают соответствующую степенную среднюю:

при m = –1 получаем среднюю гармоническую;

при m = 0 среднюю геометрическую;

при m = 1 среднюю арифметическую;

при m = 2 среднюю квадратическую, и т.д. (см. табл. 1.1).

Таблица 1.1[2]

Вид степенной средней	Показатель степени средней (m)	Формула расчета	
		Простая	Взвешенная
Гармоническая	–1	$\bar{X} = \frac{n}{\sum \frac{1}{x}}$	$\bar{X} = \frac{\sum k}{\sum \frac{k}{x}}$, где $k = xf$
Геометрическая	→0	$\bar{X} = \sqrt[n]{\prod x} = \sqrt[n]{x_1 x_2 \dots x_n}$	$\bar{X} = \sqrt[\sum f]{\prod x^f} = \sqrt[\sum f]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}}$
Арифметическая	1	$\bar{X} = \frac{\sum x}{n}$	$\bar{X} = \frac{\sum xf}{\sum f}$
Квадратическая	2	$\bar{X} = \sqrt{\frac{\sum x^2}{n}}$	$\bar{X} = \sqrt{\frac{\sum x^2 f}{\sum f}}$

Вычислив все виды средних для одних и тех же данных, заметим, что значения средних могут отличаться. Здесь действует правило мажорантности средних: с увеличением показателя степени средней (m) может увеличиваться соответствующая средняя величина. Таким образом, будет выполняться соотношение:

$$\bar{X}_{\text{гарм}} \leq \bar{X}_{\text{геом}} \leq \bar{X}_{\text{арифм}} \leq \bar{X}_{\text{квадр}}.$$

Примечание: из свойств мажорантности следует, что выбор формулы для расчета средней не может быть произвольным. Он должен основываться на смысловом содержании исходных данных и на условиях применения конкретной формулы для вычисления средней.

[2] См.: Харченко Л.П., Долженкова В.Г., Ионин В.Г. и др. Статистика : учебное пособие. – М. : ИНФРА-М, 2002. С. 92.

Рассмотрим расчет средних на конкретных примерах.

1.1. Измерив рост всех студентов в группе, получили следующие данные: 1,64 м, 1,86 м, 1,72 м, 1,95 м, 1,76 м, 1,65 м, 1,79 м, 1,82 м, 1,92 м. Найти средний рост студентов в группе.

Решение. Для определения среднего роста студентов в группе необходимо суммарный рост всех студентов в группе разделить на количество студентов. Всего в группе 9 студентов, обозначим рост каждого студента x_i, где i принимает значения от 1-го до 9-ти. Тогда, средний рост определяется по формуле средней арифметической простой:

$$\bar{X} = \frac{\sum_{i=1}^{9} x_i}{9} = \frac{1{,}64 + 1{,}86 + 1{,}72 + 1{,}95 + 1{,}76 + 1{,}65 + 1{,}79 + 1{,}82 + 1{,}92}{9}$$

$$= \frac{16{,}11}{9} = 1{,}79 \text{ м}$$

1.2. На экзамене по предмету «математика» студентами были получены следующие баллы: 46, 49, 55, 63, 74, 76, 84, 87, 92, 94. Найти средний балл за экзамен по математике группы студентов.

1.3. В течение учебного года первые четыре месяца студент не получал стипендию, следующие шесть месяцев размер стипендии составил 2,5 тыс. руб., в оставшиеся два месяца – 3,3 тыс. руб. Найти среднюю стипендию студента в рассматриваемом году.

Решение. Запишем данные о размере стипендии студента в виде таблицы:

Размер стипендии в месяц, тыс. руб. (x_i – значения осредняемого признака)	Число месяцев, в течение которых стипендия составляла данную сумму (f_i - частота, показывающая сколько раз в рассматриваемом периоде (год) встречается i-е значение осредняемого признака)
0	4
2,5	6
3,3	2

Тогда, средний за год размер стипендии студента определяется по формуле средней арифметической взвешенной и составляет:

$$\bar{X} = \frac{\sum xf}{\sum f} = \frac{0 \cdot 4 + 2{,}5 \cdot 6 + 3{,}3 \cdot 2}{4 + 6 + 2} = \frac{21{,}6}{12} = 1{,}8 \text{ тыс. руб.}$$

1.4. Используя данные о продаже билетов по направлению Москва – Екатеринбург за февраль 2018 г.:

Тип билета	Количество проданных билетов, тыс. шт.	Средняя стоимость одного билета, руб.
Плацкарт	360	1 650
Купе	310	3 200
СВ	90	4 300

Найти среднюю стоимость билета на данном направлении в феврале 2018 г.

1.5. Найти средний размер депозита в банке, если банк принимает к размещению следующие виды депозитов:

Тип депозита	Средний размер депозита данного вида, в тыс. руб.	Всего в банке размещено депозитов данного вида, в шт.
Закрытый (без возможности пополнения и снятия)	902	45 200
С возможностью пополнения	390	38 700
С возможностью снятия	630	29 500
Открытый (с возможностью пополнения и снятия)	510	52 600

1.6. Найти среднюю ставку банка по кредитному портфелю в 2017 г., используя данные о выданных банком в 2017 г. кредитах:

Вид кредита	Всего банком выдано кредитов данного вида, в млн. руб.	Средняя ставка по данному виду кредитов, в %
Потребительские кредиты	6 900	15,2
Автокредиты	25 000	12

Ипотека	28 000	9,5

1.7. Были проведены испытания точности спортивной винтовки, а именно, произведены 5 выстрелов, в каждом из них пуля отклонилась от цели на:

Номер испытания по порядку	Отклонение от цели, в мм
1	0 (точно в цель)
2	20
3	–16
4	0 (точно в цель)
5	–4

Найти среднюю величину отклонения от цели при стрельбе из спортивной винтовки.

Решение. В данном случае суммируя все значения осредняемого признака (x_i – отклонение от цели при стрельбе) получаем нулевую сумму:

$$\sum_{i=1}^{5} x_i = 0 + 20 + (-16) + 0 - 4 = 0.$$

В этом случае, используя простую среднюю арифметическую для расчета среднего значения отклонения от цели, мы получили бы нулевой результат. Что неверно, так как по данным испытай нельзя сказать, что винтовка бьет без промаха.

Потому, для расчета средней величины отклонения от цели воспользуемся формулой средней квадратической простой:

$$\bar{X} = \sqrt{\frac{\sum x^2}{n}} = \sqrt{\frac{\sum_{i=1}^{5} x_i^2}{n}} = \sqrt{\frac{0^2 + 20^2 + (-16)^2 + 0^2 + (-4)^2}{5}} \approx 11{,}59 \text{ мм}$$

Т.е. в результате пяти выстрелов среднее отклонение от цели составило 11,59 мм (в обе стороны, т.е. + / –).

1.8. Как и в предыдущем примере, были проведены испытания на точность спортивной винтовки, всего 15 испытаний. В первом, седьмом и девятом испытаниях пуля отклонилась от цели на 12 мм, в четвертом и пятнадцатом попала точно в цель, и т.д. Результаты всех испытаний представлены в таблице:

Номер испытания	Отклонение от цели, в мм	Количество попыток с соответствующим исходом
1, 7, 9	12	3
2, 5, 11, 13	–15	4
3, 6, 8, 10, 12, 14	4	6
4, 15	0 (точно в цель)	2

Найти среднюю величину отклонения от цели при стрельбе из спортивной винтовки.

Решение. Отклонение от цели, измеренное в мм – это значение осредняемого признака (x_i), количество попыток с соответствующим исходом – это частота (f_i).

Тогда, $\sum xf = \sum_{i=1}^{15} x_i f_i = 12 \cdot 3 + (-15) \cdot 4 + 4 \cdot 6 + 0 \cdot 2 = 0.$ Что при использовании средней арифметической взвешенной даст нам нулевой результат средней величины отклонения от цели. Такой результат средней величины отклонения от цели не будет объективным, так как лишь в двух испытаниях из пятнадцати пуля попала точно в цель. Потому, для нахождения средней величины отклонения от цели воспользоваться формулой средней квадратической взвешенной:

$$\bar{X} = \sqrt{\frac{\sum x^2 f}{\sum f}} = \sqrt{\frac{12^2 \cdot 3 + (-15)^2 \cdot 4 + 4^2 \cdot 6 + 0^2 \cdot 2}{3+4+6+2}} = \sqrt{\frac{1428}{15}} = \sqrt{95{,}2}$$

$$= 9{,}76 \text{ мм}$$

Таким образом, среднее отклонение от цели в результате 15-ти испытаний составило 9,76 мм. (отклонение как +, так и –).

1.9. Предположим вы пытаетесь угадать сумму очков, выпадающих на двух игральных костях. У вас 10 попыток. В первой из них вы ошиблись на +4 (т.е. ваша догадка была на 4 очка больше, чем выпало на костях), во второй вы ошиблись на –7 (т.е. ваша догадка была на 7 очков меньше, чем выпало на костях), в третьей вы ошиблись на –6, в четвертой на +10, в пятой на +1, в шестой на –8, в седьмой на +3, в восьмой на –4, в девятой вы угадали, в десятой ошиблись на +7. Найти среднюю величину ошибку вашей догадки.

1.10. В течение недели вы добираетесь до работы на одном и том же трамвае, что по расписанию пребывает на вашу остановку в 09:05. При этом, в понедельник трамвай приходит вовремя, во вторник опаздывает на 7 минут, в среду приходит раньше на 5 минут, в четверг задерживается на 3 минуты, в пятницу вновь пребывает раньше на 5 минут, в субботу трамвай приходит вовремя. Найти среднюю величину отклонения трамвая от расписания.

1.11. Группа из пяти студентов сдает экзамен (решает задачи) который длится 3 часа. Первый студент тратит на решение одной задачи 24 мин, второй – 30, третий – 40, четвертый – 45 мин и пятый – 60 мин. Найти средние затраты времени на решение одной задачи группой студентов (всеми пятью студентами) при условии, что каждый решает задачи самостоятельно.

Решение. Обозначим T_i – продолжительность экзамена (3 часа · 60 мин = 180 мин); x_i – время, затрачиваемое каждым студентом на решение одной задачи (именно для этого параметра нам и нужно определить среднее значение $\bar{X}$); q_i – количество задач решаемых за экзамен каждым студентом.

Тогда, количество задач, решенных за экзамен каждым из студентов (q_i) определим из соотношения: $q_i = \frac{T_i}{x_i}$.

q_1 = 180 мин / 24 мин = 7,5 задач за экзамен, т.е. первый студент за 3 часа экзамена (или 180 мин) решат 7,5 задач;

второй q_2 = 180 / 30 = 6 задач за экзамен;

третий q_3 = 180 / 40 = 4,5 задачи за экзамен;

четвертый q_4 = 180 / 45 = 4 задачи за экзамен;

пятый q_5 = 180 / 60 = 3 задачи за экзамен.

Теперь, зная суммарное время, потраченное всеми студентами на решение задач (= 180 мин · 5 = 900 мин) и суммарное количество решенных ими за экзамен задач (= 7,5 + 6 + 4,5 + 4 + 3 = 25), мы можем определить средние затраты времени на решение одной задачи группой студентов (при условии, что студенты решают задачи самостоятельно) по формуле простой средней арифметической:

$$\bar{X} = \frac{\sum T_i}{\sum q_i} = \frac{\sum_{i=1}^{5} T_i}{\sum_{i=1}^{5} q_i} = \frac{180 \cdot 5}{7{,}5 + 6 + 4{,}5 + 4 + 3} = \frac{900}{25} = 36 \text{ мин}$$

т.е. в среднем студенты группы тратят на решение одной задачи 36 минут.

Перепишем данную формулу используя только исходные данные задачи (т.е. только T_i и x_i):

$$\bar{X} = \frac{\sum_{i=1}^{5} T_i}{\sum_{i=1}^{5} q_i} = \frac{\sum_{i=1}^{5} T_i}{\sum_{i=1}^{5} \frac{T_i}{x_i}} = \frac{180 \cdot 5}{\frac{180}{24} + \frac{180}{30} + \frac{180}{40} + \frac{180}{45} + \frac{180}{60}} = \frac{900}{25} = 36 \text{ мин.}$$

Обратим внимание, что это формула расчета средней гармонической взвешенной, где $T = k$, а $q = f$.

Таким образом, используя формулу средней гармонической взвешенной и только исходные данные (только T_i и x_i, не вычисляя дополнительно количество задач решаемых за экзамен каждым студентом (q_i)) мы могли бы сразу вычислить искомую величину – сколько времени в среднем тратят все студенты группы на решение одной задачи.

Кроме того, если мы сократим в данной формуле параметр T_i (т.к. продолжительность экзамена у всех студентов одинакова и равна 180 мин), то получим:

$$\bar{X} = \frac{180 \cdot 5}{\frac{180}{24} + \frac{180}{30} + \frac{180}{40} + \frac{180}{45} + \frac{180}{60}} = \frac{5}{\frac{1}{24} + \frac{1}{30} + \frac{1}{40} + \frac{1}{45} + \frac{1}{60}} = \frac{5}{\sum_{i=1}^{5} \frac{1}{x_i}}$$

$$= \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\sum \frac{1}{x}} = 36 \text{ мин.}$$

А это в свою очередь формула средней гармонической простой.

1.12. Используя данные о продаже билетов по направлению Екатеринбург – Санкт-Петербург за январь 2018 г.:

Тип билета	Выручка от продажи билетов, млн руб.	Средняя стоимость одного билета, руб.
Плацкарт	399	2 375
Купе	450	3 600
СВ	66	4 400

Найти среднюю стоимость билета на данном направлении в январе 2018 г.

1.13. Найти средний размер депозита в банке, если банк принимает к размещению следующие виды депозитов:

Тип депозита	Средний размер депозита данного вида, в тыс. руб.	Всего в банке размещено депозитов данного вида на общую сумму, в тыс. руб.
Закрытый (без возможности пополнения и снятия)	905	6 516 000
С возможностью пополнения	420	1 671 600
С возможностью снятия	740	2 146 000
Открытый (с возможностью пополнения и снятия)	810	5 950 260

1.14. Банк в течение дня трижды менял курс продажи евро. По данным об объеме продаж евро (в рублевом эквиваленте) и о курсе продажи евро найти средний за день курс продажи евро в банке.

Объем продаж евро, в руб.	Курс продажи евро, руб. за 1 евро
2 035 175	63,5
839 680	65,6
1 790 845	66,5

1.15. Найти среднюю производительность труда на одного рабочего по трем цехам предприятия вместе, если по каждому из цехов известна стоимость выпущенной продукции и средняя производительность труда из расчета на одного рабочего:

Цех	Стоимость выпущенной продукции, тыс. руб.	Средняя производительность труда одного рабочего, тыс. руб.
1	2 760	24
2	2 016	19,2
3	6 939	27

1.16. Даны объемы производства предприятия А за 5 лет (см. табл. 1.2). Найти средний темп роста объемов производства предприятия за рассматриваемый период.

Таблица 1.2

Период (год)	Объем производства в тыс. тонн
2001	45,000
2002	58,500
2003	70,200
2004	77,220
2005	88,803

Решение. Обозначим темп роста объема производства i_n (это наш осредняемый признак, наше x_i). n – рассматриваемый период, i_n определяется как отношение текущего объема выпуска к уровню выпуска предыдущего года.

Тогда, темп роста объема производства предприятия А за первый период (с 2001 по 2002 год) составит $i_1 = 58,5 / 45 = 1,3$;

за второй – $i_2 = 70,2 / 58,5 = 1,2$;

за третий – $i_3 = 77,27 / 70,2 = 1,1$;

за последний четвертый период – $i_4 = 80,803 / 77,22 = 1,15$.

Если обозначить объем производства в начальный период времени (2001) – q_0, а объем производства в конечный период (2005) – q_n, то

$q_n = q_0 \cdot i_1 \cdot i_2 \cdot i_3 \cdot i_4$.

Для нахождения среднего значения темпа роста за рассматриваемый период (обозначим его $\bar{i}$) заменим в вышеприведенной формуле индивидуальные значения показателей темпов роса (i_1, i_2, i_3, i_4) средним ($\bar{i}$), это даст соотношение: $q_n = q_0 \cdot \bar{\imath} \cdot \bar{\imath} \cdot \bar{\imath} \cdot \bar{\imath} = q_0 \cdot \bar{\imath}^4$. Т.е. предположим, что если объем производства ежегодно прирастает (изменяется) с одинаковым темпом роста $\bar{\imath}$, то через 4 периода предприятие с объема производства q_0 выйдет на уровень производства q_n.

Тогда, $\bar{\imath}$ (искомое $\bar{X}$) определяется соотношением:

$$\bar{\imath} = \left(\frac{q_n}{q_0}\right)^{\frac{1}{n}} = \sqrt[n]{i_1 \cdot i_2 \cdot i_3 \cdot i_4}.$$

Обратим внимание, что это формула расчета средней геометрической простой, где $\bar{\imath} = \bar{X}$, $i_i = x_i$.

Для данных задачи

$$\bar{\imath} = \left(\frac{q_n}{q_0}\right)^{\frac{1}{n}} = \sqrt[n]{i_1 \cdot i_2 \cdot i_3 \cdot i_4} = \left(\frac{88{,}803}{45}\right)^{\frac{1}{4}} = \sqrt[4]{1{,}3 \cdot 1{,}2 \cdot 1{,}1 \cdot 1{,}15}$$
$$= \sqrt[4]{1{,}9734} \approx 1{,}185.$$

Т.е. средний темп роста производства на предприятии А за рассматриваемый период составил приблизительно 1,185.

Проверка.

$$88{,}803 = q_n = q_0 \cdot \bar{\imath}^4 = 45 \cdot (1{,}185)^4 = 45 \cdot \left(\sqrt[4]{1{,}9734}\right)^4 = 45 \cdot 1{,}9734$$
$$= 88{,}803.$$

1.17. На основании данных по двум операторам мобильной связи необходимо определить, у кого из них и насколько выше средний, за рассматриваемый период, темп роста стоимости минуты разговора (звонки на номера любого оператора, не пакетные тарифы).

Период (год)	Стоимость минуты разговора, руб.	
	Оператор мобильной связи А	Оператор мобильной связи Б
2011	1,00	0,64
2012	1,20	0,72
2013	1,08	0,81
2014	1,62	1,62

Решение. Вычислим ежегодные темпы роста стоимости минуты разговора двух операторов. Как и в предыдущей задаче обозначим темп роста – i_n (это наш осредняемый признак, наше x_i). n – рассматриваемый период; i_n определяется как отношение текущей стоимости минуты разговора к стоимости минуты разговора в предыдущем году. Результаты вычислений запишем в виде таблицы:

Период (год)	Ежегодный темп роста стоимости минуты разговора (i_n)	
	Оператор мобильной связи А	Оператор мобильной связи Б
2011	–	–
2012	i_1 = 1,20 / 1,00 = 1,2	i_1 = 0,72 / 0,64 = 1,125
2013	i_2 = 1,08 / 1,20 = 0,9	i_2 = 0,81 / 0,72 = 1,125
2014	i_3 = 1,62 / 1,08 = 1,5	i_3 = 1,62 / 0,81 = 2

Средний за период 2011–2014 темп роста стоимости минуты разговора оператора А вычислим по формуле средней геометрической простой:

$$\bar{\imath} = \bar{X} = \sqrt{\Pi\, x} = \sqrt[n]{x_1 x_2 \dots x_n} = \sqrt[3]{i_1 \cdot i_2 \cdot i_3} = \sqrt[3]{1{,}2\ \cdot 0{,}9\ \cdot 1{,}5} = \sqrt[3]{1{,}62}$$
$$= 1{,}174$$

т.е. в среднем за три года стоимость минуты разговора ежегодно увеличивалась в 1,174 раза.

Для расчета среднего темпа роста оператора Б воспользуемся формулой средней геометрической взвешенной, т.к. среди ежегодных темпов роста присутствуют повторяющиеся значения (1,125):

$$\bar{\imath} = \bar{X} = \sqrt[\Sigma f]{\Pi\, x^f} = \sqrt[\Sigma f]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}} = \sqrt[3]{i_1^2\ i_3^1} = \sqrt[3]{1{,}125^2\ \cdot 2} = \sqrt[3]{2{,}53125}$$
$$= 1{,}363$$

т.е. в среднем за три года стоимость минуты разговора ежегодно увеличивалась в 1,363 раза, что на 0,189 выше чем у оператора А. В результате стоимость минуты разговора двух операторов сравнялась, в то время как в начале рассматриваемого периода данный показатель оператора Б был в 1,5625 раза (1 / 0,64 = 1,5625) меньше, чем у конкурента.

1.18. По данным о ВВП страны А за период с 2005 по 2010 год найти средний темп роста ВВП за рассматриваемый период.

Период (год)	ВВП, млн. ден. ед.
2005	12 500
2006	14 500
2007	11 600
2008	10 440
2009	9 918
2010	14 877

При анализе рядов динамики[3] единицей совокупности выступает момент времени (дата) или интервал времени (период), что влияет на формулы расчета среднего значения – появляются формулы средних хронологических величин.

В *моментных рядах динамики* уровни ряда характеризуют значение показателя по состоянию на определенные моменты времени (на даты). Например, численность населения города А на 01 января 2000 г. составила 1,49 млн. чел.

В *интервальных рядах динамики* уровни ряда характеризуют значение показателя, накопленного за определенный интервал времени (период). Например, за 2014 г. заводом было выпущено 2 тыс. автомобилей. Или за 3 квартал 2015 г. объем потребления мяса птицы в стране Б составил 3,56 тыс. тонн.

Для интервального ряда в любой точке внутри изучаемого интервала значение признака принимается постоянным.

Средний уровень ряда для интервальных рядов с равными временными интервалами вычисляется по формуле:

---

[3] Рядами динамики (или хронологическими, динамическими, временными рядами) называют числовой ряд значений некоторого статистического показателя зафиксированного в последовательные моменты или периоды времени.

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n},$$

где n – количество уровней ряда (количество равных временных отрезков), Y_i – значение уровня ряда в интервале i (i = 1,2, …n).

Средний уровень ряда для интервальных рядов с неравными временными интервалами вычисляется по формуле средней арифметической взвешенной:

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i \cdot t_i}{\sum_{i=1}^{n} t_i},$$

где t_i – длина временного интервала.

1.19. По нижеприведенным данным найти среднюю урожайность чечевицы за рассматриваемый период:

Динамика урожайности чечевицы по годам					
Год	2014	2015	2016	2017	2018
Урожайность, в ц/га	98	76	106	116	129

Решение. В данном случае у нас интервальный ряд с равными временными интервалами – годами (i). Показатель урожайности это Y_i, т.е. значение уровня ряда в соответствующем интервале (году i). В любой точке внутри изучаемого года значение признака принимается постоянным. Соответственно, средняя урожайность чечевицы за 5 рассматриваемых лет составила:

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{98 + 76 + 106 + 116 + 129}{5} = 105 \text{ ц/га.}$$

1.20. По нижеприведенным данным найти средний размер ключевой ставки Банка России за 2018 г.:

Данные о динамике ключевой ставки Банка России за 2018 г.				
период	01.01–11.02	12.02–25.03	26.03–16.09	17.09–31.12
размер ключевой ставки, в %	7,75	7,50	7,25	7,50

Решение. В данном случае у нас интервальный ряд с неравными временными интервалами определяемыми количеством дней (t_i). Значение ключевой ставки это Y_i, т.е. значение уровня ряда в соответствующем интервале (например, с 01 января по 11 февраля). В любой точке внутри изучаемого интервала (t_i) значение ключевой ставки принимается постоянным. Соответственно, средняя величина ключевой ставки Банка России в 2018 году составила:

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i \cdot t_i}{\sum_{i=1}^{n} t_i} = \frac{7{,}75 \cdot 42 + 7{,}50 \cdot 42 + 7{,}25 \cdot 175 + 7{,}50 \cdot 106}{42 + 42 + 175 + 106} = \frac{2\,704{,}25}{365} = 7{,}41\ \%.$$

Несколько иначе рассчитывается средний уровень ряда для моментных временных рядов[4]. Например, если имеется моментный ряд, содержащий n уровней (Y_1, Y_2, …, Y_n) с равными (n – 1) интервалами между уровнями (моментами времени), то формула для расчета среднего уровня ряда принимает вид:

$$\bar{Y} = \frac{\frac{1}{2} Y_1 + Y_2 + \cdots + Y_{n-1} + \frac{1}{2} Y_n}{n-1}$$

Эта средняя носит название средняя хронологическая[5].

Если имеется моментный ряд, содержащий n уровней (Y_1, Y_2, …, Y_n) с неравными (n – 1) интервалами между уровнями (моментами времени), то

---

[4] Примером моментного ряда может выступать, например, численность сотрудников предприятия на начало каждого месяца. Моментом для данного ряда является первое число каждого месяца, т.е. момент времени на который и определяется численность сотрудников предприятия.

[5] В данной формуле предполагается, что в пределах каждого периода, разделяющего моментные наблюдения, развитие происходит по линейному закону, т.е. значение показателя между двумя моментами времени либо монотонно возрастает, либо монотонно убывает.

для расчета среднего уровня ряда предварительно вычисляют значения уровней ряда в серединах интервалов:

$$\bar{Y}_1 = \frac{Y_1 + Y_2}{2}, \bar{Y}_2 = \frac{Y_2 + Y_3}{2}, \dots, \bar{Y}_{n-1} = \frac{Y_{n-1} + Y_n}{2},$$

а затем вычисляется общий средний уровень ряда (как средняя арифметическая взвешенная):

$$\bar{Y} = \frac{\frac{Y_1 + Y_2}{2} \cdot t_1 + \frac{Y_2 + Y_3}{2} \cdot t_2 + \dots + \frac{Y_{n-1} + Y_n}{2} \cdot t_{n-1}}{t_1 + t_2 + \dots + t_{n-1}} = \frac{\sum_{i=1}^{n-1} \bar{Y}_i \cdot t_i}{\sum_{i=1}^{n-1} t_i},$$

где t_i – время, в течение которого среднее значение интервала принимается постоянным.

1.21. Найти среднесписочную численность сотрудников предприятия за пять месяцев, если известна численность сотрудников на первое число каждого месяца с марта по август:

Показатель	март	апрель	май	июнь	июль	август
Численность сотрудников на 01 число месяца, чел.	769	1 729	676	512	429	299

Решение. I способ. В данном случае у нас моментный временной ряд, где моментом «замера» показателя является первое число каждого месяца. Мы примем промежутки между датами (моментами) равными друг другу (не будем делать различия между продолжительностью месяца в 30 и 31 день). В данном временном ряду показатель на начало каждого периода одновременно является показателем на конец предыдущего периода. Например, на 01 марта численность сотрудников составила 769 чел., а на конец марта – 1 729 чел. Очевидно, что значение среднего показателя внутри каждого интервала зависит от того, как шло развитие в рамках интервалов, происходило это постепенно или скачкообразно. Мы этого не знаем. Принято предполагать, что в пределах каждого периода, разделяющего моментные наблюдения, развитие происходило по линейному закону. А значит, среднее значение показателя внутри каждого интервала может быть вычислено как полусумма значений на начало и конец данного интервала.

Например, в марте численность сотрудников в среднем была равна (769 + 1 729) / 2 = 1 249 чел.

Количество таких средних будет равно количеству интервалов между моментами «замера» показателя. В нашем случае 6 уровней ряда, интервалов между уровнями будет 6 – 1 = 5.

Как результат, среднее значение показателя за все пять периодов можно вычислить как среднее арифметическое средних в каждом интервале.

Для нашего примера среднесписочная численность сотрудников предприятия за пять месяцев равна:

$$\bar{Y} = \frac{\frac{769+1\ 729}{2}+\frac{1\ 729+676}{2}+\frac{676+512}{2}+\frac{512+429}{2}+\frac{429+299}{2}}{6-1}$$

$$= \frac{1\ 249+1\ 202{,}5+594+470{,}5+364}{5} = 776 \text{ чел.}$$

II способ – воспользоваться для расчетов формулой средней хронологической:

$$\bar{Y} = \frac{\frac{Y_1}{2}+Y_2+Y_3+\dots+Y_{n-1}+\frac{Y_n}{2}}{n-1},$$

где n = 6 (количество моментов ряда – 01 число каждого месяца).

В нашем случае

$$\bar{Y} = \frac{\frac{Y_1}{2}+Y_2+Y_3+Y_4+Y_5+\frac{Y_6}{2}}{5} = \frac{\frac{769}{2}+1\ 729+676+512+429+\frac{299}{2}}{5} = 776 \text{ чел.}$$

т.е. среднесписочная численность сотрудников за пять месяцев (с начала марта по начало августа) составила 776 человек.

Примечание. Заметим, что наш первый и второй способы эквивалентны:

$$\bar{Y} = \frac{\frac{Y_1+Y_2}{2}+\frac{Y_2+Y_3}{2}+\frac{Y_3+Y_4}{2}+\frac{Y_4+Y_5}{2}+\frac{Y_5+Y_6}{2}}{6-1} = \frac{\frac{Y_1}{2}+Y_2+Y_3+Y_4+Y_5+\frac{Y_6}{2}}{5}.$$

1.22. По нижеприведенным данным вычислите среднюю стоимость основных фондов (ОФ) предприятия за первый квартал:

Показатель	декабрь	январь	февраль	март
Стоимость ОФ (здания, оборудование и пр.) на конец месяца, млн. руб.	92	102	149	180

§ 2. Структурные средние и анализ формы распределения

2.1. Структурные средние.

Важной характеристикой центра распределения, помимо средней, являются, так называемые структурные средние – *мода* (*Мо*) и *медиана (Ме)*. В отличие от средней арифметической, на которую влияют все значения изучаемого признака x_i, структурные средние не зависят от крайних значений признака. Потому они являются лучшей, чем среднее арифметическое, характеристикой центра распределения для рядов с неопределенными границами (например, для рядов с открытыми крайними границами интервалов).

По определению *мода* это значение признака наиболее часто встречающееся в вариационном ряду. А *медианой* называют значение признака, которое делит упорядоченную последовательность x_i на две равные по численности части. В итоге у одной половины единиц совокупности значение признака не превышает медианный, а у другого – превышает медианный уровень.

Для дискретного ряда мода и медиана находятся непосредственно по определению.

Так, по данным табл. 1.3 определим, что типичным исходом футбольного матча во время Чемпионата мира 2018 г. был один забитый гол, т.е. мода равна 1 (наибольшее число матчей – 46, это матчи с одним забитым голом, 1 – значение признака которое встречается чаще всего, это и есть мода).

Таблица 1.3

Распределение футбольных матчей по числу забитых за матч мячей обеими командами (Чемпионат мира 2018)

Число забитых мячей в матче	Число матчей	
0	21	67 (= 21 + 46)
1	46	
2	34	
3	14	
4	9	

Для списка банков Санкт-Петерурга (табл. 1.4), упорядоченных по размеру собственного капитала, медианным банком будет банк «Петровский», а медианой – 268 млн руб., именно это значение признака делит упорядоченную (в данном случае по возрастанию) последовательность значений (169, 237, 268, 290, 1 007) на две равные по численности части (169, 237) и (290, 1 007).

Таблица 1.4

Банки Санкт-Петербруга, ранжированы по размеру собственного капитала

Название	Собственный капитал, млн. руб.
Балтонэксимбанк	169
Банк «Санкт – Петербург»	237
Петровский	268
Балтийский	290
Промстройбанк	1 007

При четном числе элементов вариационного ряда за медиану принимают среднюю арифметическую из двух центральных вариант упорядоченного ряда. Так, по данным табл. 1.3 из 124 элементов (21 + 46 +

34 + 14 + 9 = 124) центральными являются 62 и 63. Оба они равны 1 и потому медиана равна (1+1) / 2 = 1.

На практике встречаются многовершинные распределения, т.е. распределения в которых несколько максимумов частот, несколько мод. Наличие нескольких вершин (нескольких мод) является признаком того, что изучаемая совокупность состоит из неоднородных, с точки зрения изучаемого признака, единиц. Например, изучая спрос на мероприятия, проводимые на детских праздниках, было получено многовершинное распределение, распределение с двумя модами (см. табл. 1.5). Так, наиболее востребованными оказались фокусы и участие в дне рождения слона (у этих опций развлечений максимальная частота – 57). Что наталкивает на мысль о том, что изучаемая совокупность семей неоднородна (в выборке присутствуют более обеспеченные семьи и семьи со средним достатком). Как результат, напрашивается вывод, что такую совокупность имеет смысл разделить на две и рассматривать каждую из них отдельно.

Таблица 1.5

Изучение спроса на мероприятия, проводимые на детских праздниках

Наиболее предпочитаемая опция	Количество респондентов, чел.
Фокусы	57
Клоуны	26
Театр кукол	19
Костюмированный балл	22
Квест по мотивам мультфильма / фильма	21
Конкурсы и музыкальная программа	14
Мастер классы (рисование, кулинария)	34
Участие в дне рождения слона	57

Если данные признака x_i представлены в виде интервального ряда, то расчет моды и медианы ведется по формулам. При этом, для расчета моды используют две формулы – первая для случая интервального ряда с равными интервалами, вторая для случая интервального ряда с неравными интервалами.

Для случая равных интервалов моду определяют по формуле:

$$Mo = x_{Mo} + h_{Mo} \cdot \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}, \quad (1.1)$$

где x_{Mo} – нижняя граница модального интервала;

h_{Mo} – ширина модального интервала;

f_{Mo} – частота в модальном интервале;

f_{Mo-1} – частота интервала, предшествующего модальному;

f_{Mo+1} – частота интервала, следующего за модальным.

Данная интерполяционная формула была предложена Р.М. Орженцким. Для выведения формулы принято допущение, что в модальном и двух соседних с ним интервалах кривая распределения признака x_i представляет собой параболу 2-го порядка, и мода находится как абсцисса точки максимума кривой распределения[6].

Для случая неравных интервалов моду вычисляют по формуле:

$$Mo = x_{Mo} + h_{Mo} \cdot \frac{f'_{Mo} - f'_{Mo-1}}{(f'_{Mo} - f'_{Mo-1}) + (f'_{Mo} - f'_{Mo+1})}, \quad (1.2)$$

где x_{Mo} – нижняя граница модального интервала;

h_{Mo} – ширина модального интервала;

f'_{Mo} – плотность в модальном интервале;

f'_{Mo-1} – плотность интервала, предшествующего модальному;

f'_{Mo+1} – плотность интервала, следующего за модальным;

а плотность это отношение частоты к ширине соответствующего интервала $\left(f' = \frac{f}{h}\right)$.

Для расчета медианы неважно являются интервалы равными или нет, при расчете используется одна универсальная формула:

---

6 См.: Громыко Г.Л., Воробьев А.Н., Карасева Л.А. и др. Теория статистики : Учебник. – М. : ИНФРА-М, 2005. С. 100.

$$Me = x_{Me} + h_{Me} \cdot \frac{0{,}5 \cdot \sum f_i - S_{Me-1}}{f_{Me}}, \qquad (1.3)$$

где x_{Me} – нижняя граница медианного интервала;

h_{Me} – ширина медианного интервала;

Σf_i – число единиц изучаемого множества;

S_{Me-1} – накопленная чистота в интервале, предшествующем медианному;

f_{Me} – частота в медианном интервале.

При расчете медианы важно не забывать соблюдать условие ранжированности (упорядоченности по возрастанию или убыванию) элементов ряда.

1.23. Дано распределение фермерских хозяйств Рязанской области по урожайности зерновых культур. Найти моду и медиану данного распределения.

Урожайность, ц/га	Число фермерских хозяйств
5 – 15	8
15 – 25	29
25 – 35	39
35 – 45	59
45 – 55	9
55 – 65	5

Решение. В данном случае у нас интервальный ряд с равными интервалами (10 = 15 – 5 = 25 – 15 = 35 – 25 = 45 – 35 = 55 – 45 = 65 – 55). А потому, для вычисления моды воспользуемся формулой (1.1).

Сначала необходимо определить модальный интервал, т.е. тот интервал, в котором содержится мода.

Как мы рассуждаем? В равных интервалах содержится 8, 29, 39, 59, 9 и 5 значений признака x_i (x_i – урожайность ц/га). Сами значения признака мы не знаем, знаем только их количество. В каком из равных интервалов содержится Мода?

Вероятность встретить повторяющиеся значения признака наибольшая в том интервале, где содержится большее количество значений признака. Т.е. для интервального ряда с равными интервалами модальным является интервал, которому соответствует максимальная частота или частость[7]. В данном случае наибольшая частота по числу фермерских хозяйств – f_{max} = 59. Таким образом, модальным является интервал 35–45. Его нижняя граница равна x_{Mo} = 35, ширина h_{Mo} = 45 – 35 = 10, частота в модальном интервале равна f_{Mo} = 59 = f_{max}, частота в интервале, предшествующем модальному равна f_{Mo-1} = 39, частота интервала, следующего за модальным f_{Mo+1} = 9.

$$Mo = x_{Mo} + h_{Mo} \cdot \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})} = 39 + 10 \cdot \frac{59-39}{(59-39)+(59-9)} \approx 41{,}9 \text{ ц/га}$$

Таким образом, в Рязанской области чаще всего встречаются хозяйства с урожайностью 41,9 ц/га[8].

Следующим шагом определим медиану. Сначала убедимся, что наш интервальный ряд упорядочен. Действительно, наш ряд упорядочен по возрастанию относительно изучаемого признака x_i (урожайность), все интервалы ряда идут по нарастающей с 5 до 65 ц/га.

Для вычисления медианы воспользуемся формулой (1.3). Но прежде, определим медианный интервал. Для этой цели воспользуемся накопленными частотами (или частостями). Вычисления приведены в табл. 1.6.

---

[7] Частость (w_i) – относительное выражение частоты (частости – частоты, выраженные в долях, $\sum w_i = 1$).

[8] После вычислений полезно проверить попадает ли значение моды в модальный интервал. Если нет, значит при расчетах допущена ошибка. В нашем случае *Mo* = 41,9 ц/га это значение принадлежит модальному интервалу (35 – 45), значит мы верно вычислили показатель *Mo*.

Таблица 1.6

Урожайность, ц/га (x_i)	Число фермерских хозяйств (f_i)	Частости (w_i)	Накопленные частоты (S_i)	Накопленные частости (W_i)
5 – 15	8	= 8 / 149 = 0,05	8	0,05
15 – 25	29	= 29 /149 = 0,19	8 + 29 = 37	0,25
25 – 35	39	= 39 / 149 = 0,26	8 + 29 + 39 = **76**	**0,51**
35 – 45	59	= 59 / 149 = 0,40	8 + 29 + 39 + 59 = 135	0,91
45 – 55	9	= 9 /149 = 0,06	8 + 29 + 39 +59 +12 =144	0,97
55 – 65	5	= 5 /149 = 0,03	8 + 29 + 39 + 59 + 9 + 5 = 149	1,00
Итого:	149	1,00	–	–

Для наших 149 фермерских хозяйств медианным, стоящим в середине, является (149 + 1) / 2 = 75-е фермерское хозяйство. Накапливаем частоты до тех пор, пока не будет превзойден номер медианы. Так, 8 фермерских хозяйств имеют урожайность 5–15 ц/га, 8 + 29 = 37 фермерских хозяйств имеют урожайность 15–25 ц /га, а 37 + 39 = 76 фермерских хозяйств имеют урожайность 25–35 ц/га, т.е. 75-е фермерское хозяйство имеет урожайность 25–35 ц/га. Таким образом, медианным является интервал (25–35). Это видно и из ряда накопленных частот (S_i), где значение 75 попадает в третий интервал (25−35).

Если исходить из частостей, то медианным является интервал в который попадает значение 0,5 (половина всей изучаемой совокупности), это также третий интервал (25–35).

Нижняя граница медианного интервала x_{Me} = 25, ширина медианного интервала h_{Me} = 10, число единиц изучаемого множества Σf_i = 149, накопленная чистота в интервале, предшествующем медианному S_{Me-1} = 37, частота в медианном интервале f_{Me} = 39. Подставляем значения в формулу (1.3), получаем:

$$Me = x_{Me} + h_{Me} \cdot \frac{0{,}5 \cdot \sum f_i - S_{Me-1}}{f_{Me}} = 25 + 10 \cdot \frac{0{,}5 \cdot 149 - 37}{39} = 34{,}6 \text{ ц/га}^{9},$$

С использованием частостей, получаем аналогичный результат:

$$Me = x_{Me} + h_{Me} \cdot \frac{0{,}5 \cdot \sum f_i - S_{Me-1}}{f_{Me}} = 25 + 10 \cdot \frac{0{,}5 \cdot 1 - 0{,}25}{0{,}26} = 34{,}6 \text{ ц/га}.$$

Таким образом, у половины хозяйств урожайность превышает 34,6 ц/га, а у другой половины урожайность меньше 34,6 ц/га.

1.24. Рассмотрим распределение малых городов и поселков городского типа (ПГТ) по числу жителей в некоторой стране А:

Число жителей, тыс. чел.	Число городов и ПГТ, шт.
0 – 5	26
5 – 10	84
10 – 20	269
20 – 50	371
50 – 100	176

Необходимо определить какова численность, наиболее часто встречающихся в данной стране А, городов и поселков городского типа.

Решение. По смыслу, в задаче требуется определить моду данного распределения. Начнем с вычисления ширины интервалов данного распределения. Вычисления представлены в табл. 1.7.

Таблица 1.7

Число жителей, тыс. чел. (x_i)	Число городов и ПГТ, шт. (f_i)	Ширина интервала, тыс. чел. (h_i)	Плотность распределения, (f'_i)
0 – 5	23	5	23 / 5 = 4,6
5 – 10	45	5	45 / 5 = 9
10 – 20	90	10	90 / 10 = 9
20 – 50	375	30	375 / 30 = 12,5

[9] После вычислений, как и в случае моды, полезно проверить попадает ли значение медианы в медианный интервал. Если нет, значит в вычислениях допущена ошибка. В нашем случае *Me* = 34,6 ц/га это значение принадлежит медианному интервалу (25–35), значит мы верно вычислили показатель *Me*.

50 – 100	176	50	176 / 50 = 3,52

В данном случае интервалы разной величины. И выбрать интервал с наибольшей частотой в качестве модального мы не можем. Так как важна частота на единицу длины интервала.

Например, второй интервал ширины h = 5 содержит 45 значений признака, а третий, в два раза большей ширины h = 10, содержит в два раза больше единиц признака – 90. Т.о. частота на единицу длины интервала у них одинаковая (45 / 5 = 90 / 10 = 9) несмотря на то, что количество значений признака во втором интервале меньше, чем в третьем. Потому нельзя сказать, что вероятность встретить повторяющиеся значения признака в третьем интервале больше, чем во втором.

В данном случае, когда расчет модального значения выполняется по рядам распределения с неравными интервалами, необходимо вычислить плотности и уже по ним, найдя максимальное значение плотности, определить модальный интервал. А затем, используя формулу (1.2), определить моду.

В нашем случае наибольшая плотность $f\,'_{i\_max}$ = 12,5. Соответственно, модальным является интервал (20–50). Подставив соответствующие значения в форуму (1.2) получаем:

$$Mo = x_{Mo} + h_{Mo} \cdot \frac{f'_{Mo} - f'_{Mo-1}}{(f'_{Mo} - f'_{Mo-1}) + (f'_{Mo} - f'_{Mo+1})}$$
$$= 20 + 30 \cdot \frac{12{,}5 - 9}{(12{,}5 - 9) + (12{,}5 - 3{,}52)} = 28{,}4 \text{ тыс. чел}$$[10].

Таким образом, в стране А чаще всего встречаются города и поселки городского типа с численностью населения 28,4 тыс. чел.

---

[10] Проверка. *Mo* = 28,4 тыс. чел. принадлежит модальному интервалу (20–50), значит мы верно вычислили показатель *Mo*.

1.25. Определить моду и медиану для следующего распределения – ежедневное число поездок на общественном транспорте:

Число поездок в день	Количество граждан, совершающих в день данное количество поездок
0	221
1	146
2	434
3	405
4	302
5	112

2.2 Анализ формы распределения.

Степень асимметричности распределения характеризуется коэффициентом асимметрии (As):

$$As = \frac{\mu_3}{\sigma^3} = \frac{\frac{\sum_i (x_i - \bar{x})^3}{n}}{\sigma^3} \text{ (для несгруппированных)}$$

или

$$As = \frac{\mu_3}{\sigma^3} = \frac{\frac{\sum_i (x_i - \bar{x})^3 \cdot f_i}{\sum f_i}}{\sigma^3} \text{ (для сгруппированных данных),}$$

а также коэффициентом асимметрии Пирсона ($As_\text{П}$):

$$As_\text{П} = \frac{\bar{x} - Mo}{\sigma},$$

где σ – среднее квадратическое отклонение;

μ_3 – третий центральный момент;

x_i – значение признака (для интервального ряда x_i заменяют x'_i – центральным значением интервала);

$\bar{x}$ – среднее значение признака;

n – количество наблюдений в распределении;

f_i – частоты распределения;

Mo – мода.

Знак коэффициента асимметрии и коэффициента асимметрии Пирсона совпадает.

Для симметричных распределений (вкл. нормальное распределение) варианты (x_i), равноудаленные от $\bar{x}$, имеют одинаковую частоту, потому μ_3 и соответственно $As = 0$.

Если $As > 0$, то в вариационном ряду преобладают варианты, которые больше средней ($\bar{x}$) и распределение имеет более длинную ветвь справа (скошено вправо). Такое распределение называют правосторонне ассиметричным, в этом случае $\bar{x} > Mo$.

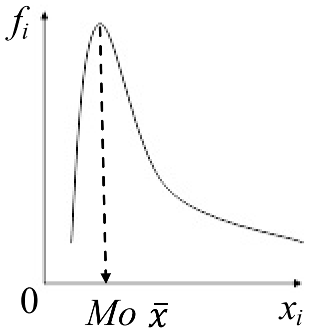

Рис. 1. Правосторонняя асимметрия ($As > 0$)

Примером такого распределения может служить средний уровень заработной платы (з/п) в стране. Когда большинство людей получают зарплату меньше среднего значения ($Mo < \bar{x}$), но незначительная группа людей получает з/п много выше среднего уровня (скошено вправо), что и выводит среднюю заработную плату на уровень выше Mo.

Если $As < 0$, то в вариационном ряду преобладают варианты, которые меньше средней ($\bar{x}$) и распределение имеет более длинную ветвь слева (скошено влево). Такое распределение называют левосторонне ассиметричным, в этом случае $Mo > \bar{x}$.

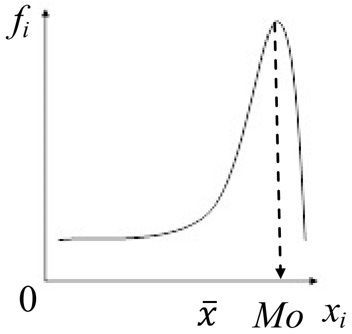

Рис. 2. Левосторонняя асимметрия ($As < 0$)

Примером такого распределения может служить средний по стране уровень оплаты за коммунальные услуги. Когда большинство людей платят больше среднего значения ($Mo > \bar{x}$), но за счет того, что незначительная группа людей платит значительно меньше среднего уровня (скошено влево), средний уровень расходов на ЖКХ смещается в меньшую сторону, становиться меньше Mo.

Для распределений по форме близких к нормальному закону, медиана находится между модой и средней величиной. Для таких распределений при правосторонней асимметрии $\bar{x} > Me > Mo$, при левосторонней асимметрии $\bar{x} < Me < Mo$.

Для оценки крутизны (заостренности) распределения вычисляют эксцесс (Ex):

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{\frac{\sum_i (x_i - \bar{x})^4}{n}}{\sigma^4} - 3 \text{ (для несгруппированных)}$$

или

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{\frac{\sum_i (x_i - \bar{x})^4 \cdot f_i}{\sum f_i}}{\sigma^4} - 3 \text{ (для сгруппированных данных)},$$

где σ – среднее квадратическое отклонение;

μ_4 – четвертый центральный момент;

x_i – значение признака (для интервального ряда x_i заменяют x'_i – центральным значением интервала);

$\bar{x}$ – среднее значение признака;

n – количество наблюдений в распределении;

f_i – частоты распределения.

Для нормального распределения $\frac{\mu_4}{\sigma^4} = 3$ (и $Ex = 0$), оно и служит эталоном для оценки крутизны распределений. Если $Ex > 0$ распределение называют островершинным, если $Ex < 0$ – плосковершинным.

1.26. По данным табл. 1.8 найти показатели асимметрии и крутизны распределения.

Таблица 1.8

Заработная плата, тыс. руб.	Количество рабочих, чел.
20,260 – 21,510	6
21,510 – 22,760	12
22,760 – 24,010	29
24,010 – 25,260	24
25,260 – 26,510	6
26,510 – 27,760	5
27,760 – 29,010	4

Решение. Так как распределение задано интервальным рядом (равные интервалы с шагом h = 1 250 руб.) показатели асимметрии и крутизны определим по формулам:

$$As = \frac{\mu_3}{\sigma^3} = \frac{\frac{\sum_i (x_i - \bar{x})^3 \cdot f_i}{\sum f_i}}{\sigma^3} \text{ (для сгруппированных данных),}$$

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{\frac{\sum_i (x_i - \bar{x})^4 \cdot f_i}{\sum f_i}}{\sigma^4} - 3 \text{ (для сгруппированных данных).}$$

Для нахождения As и Ex сначала найдем среднее и среднее квадратическое отклонение распределения. Необходимые для расчета величины вычислены в таблице:

Заработная плата, тыс. руб. $x_i - x_{i+1}$	Количество рабочих, чел. f_i	Центр интервала x'_i	$x'_i \cdot f_i$	$(x'_i - \bar{x})^2 \cdot f_i$	$(x'_i - \bar{x})^3 \cdot f_i$	$(x'_i - \bar{x})^4 \cdot f_i$
20,260–21,510	6	20,885	125,310	58,594	–183,105	572,205
21,510–22,760	12	22,135	265,620	42,188	–79,102	148,315
22,760–24,010	29	23,385	678,165	11,328	–7,080	4,425
24,010–25,260	24	24,635	591,240	9,375	5,859	3,662
25,260–26,510	6	25,885	155,310	21,094	39,551	74,158
26,510–27,760	5	27,135	135,675	48,828	152,588	476,837
27,760–29,010	4	28,385	113,540	76,563	334,961	1 465,454
Итого	86	–	2 064,860	267,969	263,672	2 745,056

Среднее значение найдем по формуле средней арифметической взвешенной, в качестве x_i возьмем центральные значения интервалов x'_i:

$\bar{x} = \frac{\sum x'_i \cdot f_i}{\sum f_i} = \frac{2\ 064{,}860}{86} = 24{,}010$ тыс. руб. (т.е. средний уровень заработной платы 86 сотрудников предприятия составил 24 010 руб.).

В формуле для вычисления среднего квадратического отклонения в качестве x_i также возьмем центральные значения интервалов x'_i:

$$\sigma = \sqrt{\frac{(x'_i - \bar{x})^2 \cdot f_i}{\sum f_i}} = \sqrt{\frac{267{,}969}{86}} = \sqrt{3{,}116} = 1{,}765 \text{ тыс. руб.}$$

Коэффициент асимметрии

$$As = \frac{\mu_3}{\sigma^3} = \frac{\frac{\sum_i (x_i - \bar{x})^3 \cdot f_i}{\sum f_i}}{\sigma^3} = \frac{\frac{263{,}672}{86}}{(1{,}765)^3} = \frac{3{,}066}{(1{,}765)^3} = 0{,}577$$

Асимметрия заметная. Т.к. $As > 0$, распределение скошено вправо (правосторонняя асимметрия).

Эксцесс распределения

$$Ex = \frac{\mu_4}{\sigma^4} - 3 = \frac{\frac{\sum_i (x_i - \bar{x})^4 \cdot f_i}{\sum f_i}}{\sigma^4} - 3 = \frac{\frac{2\ 745{,}056}{86}}{(1{,}765)^4} - 3 = \frac{31{,}919}{(1{,}765)^4} - 3 = 0{,}29$$

Коэффициент крутизны (эксцесс) $Ex > 0$, т.е. распределение островершинное.

1.27. По данным табл. 1.9 найти показатели асимметрии и крутизны распределения.

Таблица 1.9

Заработная плата, руб. в час	Количество рабочих, чел.
300 – 400	3
400 – 500	5
500 – 600	6
600 – 700	9
700 – 800	18
800 – 900	19
900 – 1 000	2

1.28. Найдите значение Mo и $\bar{X}$ для следующего дискретного ряда:

Признак (xi)	Частота признака (fi)
1	22
2	28
3	36
4	14

Что можно сказать о коэффициенте асимметрии (As) данного ряда? Изобразите схематично данное распределение.

Решение. В данном распределении наиболее часто встречается значение признака $x_i = 3$ (его частота $f_i = 36 = f_{max}$ – максимальная в данном распределении). Таким образом, Мода = 3.

Среднее значение найдем по формуле средней взвешенной:

$$\bar{X} = \frac{\sum xf}{\sum f} = \frac{1 \cdot 22 + 2 \cdot 28 + 3 \cdot 36 + 4 \cdot 14}{22 + 28 + 36 + 14} = \frac{242}{100} = 2{,}42.$$

Тогда мы получаем: $\overline{\mathrm{X}} = 2{,}42 < Mo = 3, т.е. \overline{\mathrm{X}} < Mo$. Т.о. в данном случае мы имеем левостороннюю асимметрию, т.е. показатель $As < 0$. См. рис. 2.

1.29. Вычислите значение Mo и $\bar{X}$ для следующего ряда:

Признак	Частота признака
1	9
2	15
3	14
4	10

Что можно сказать о коэффициенте асимметрии (As) данного ряда? Изобразите схематично данное распределение.

Решение. В данном распределении наиболее часто встречается значение признака $x_i = 2$ (его частота $f_i = 15 = f_{max}$ – максимальная в данном распределении). Таким образом, Мода = 2.

$$\bar{X} = \frac{\sum xf}{\sum f} = \frac{1 \cdot 9 + 2 \cdot 15 + 3 \cdot 14 + 4 \cdot 10}{9 + 15 + 14 + 10} = \frac{121}{48} \approx 2{,}52.$$

Получаем: $\overline{\mathrm{X}} = 2{,}52 > Mo = 2,$ т. е. $\overline{\mathrm{X}} > Mo.$ В данном случае мы имеем правостороннюю асимметрию, т.е. $As > 0$ (см. рис. 1).

1.30. Вычислите значение Mo и $\bar{X}$ для следующего ряда:

Признак	Частота признака
1	7
2	21
3	14
4	14

Что можно сказать о коэффициенте асимметрии (As) данного ряда? Изобразите схематично данное распределение.

1.31. Вычислите значение Mo и $\bar{X}$ для следующего ряда

Признак	Частота признака
1	5
2	7
3	10
4	2

Что можно сказать о коэффициенте асимметрии (As) данного ряда? Изобразите схематично данное распределение.

1.32. Вычислите значение Mo и $\bar{X}$ для следующего ряда:

Признак	Частота признака
1	2
2	14
3	10
4	6

Что можно сказать о коэффициенте асимметрии (As) данного ряда? Изобразите схематично данное распределение.

2.3 Другие структурные характеристики вариационного ряда.

Медиану используют если необходимо разделить совокупность на две равные по численности части. Но если необходимо разделить совокупность на две неравные по численности части или на большее количество частей, тогда рассчитывают соответствующее *квартильное*, *квантильное*, *децильное*, или *перцентильное* значения признака. Квартиль – значение признака, делящее совокупность на четыре равные части, квантиль – на пять равных частей, дециль – на десять равных частей, и перцентиль – на сто равных частей.

Логика расчета и формулы для вычисления данных показателей аналогичны таковым для медианы. Рассмотрим их на примере квартилей. Для этого упорядочим совокупность (например, по возрастанию) и обозначим значения x_i, делящие совокупность на четыре равные части – Q_1, Q_2 и Q_3. Первый квартиль (Q_1) делит элементы совокупности на две части – 1/4 значений x_i лежит ниже Q_1, 3/4 значений x_i лежат выше Q_1. Второй квартиль (Q_2) делит совокупность x_i на две равные части (т.е. Q_2 совпадает с медианой). Выше третьего квартиля (Q_3) располагаются 3/4 всех значений x_i, а ниже оставшаяся 1/4 часть совокупности x_i (см. рис. 3).

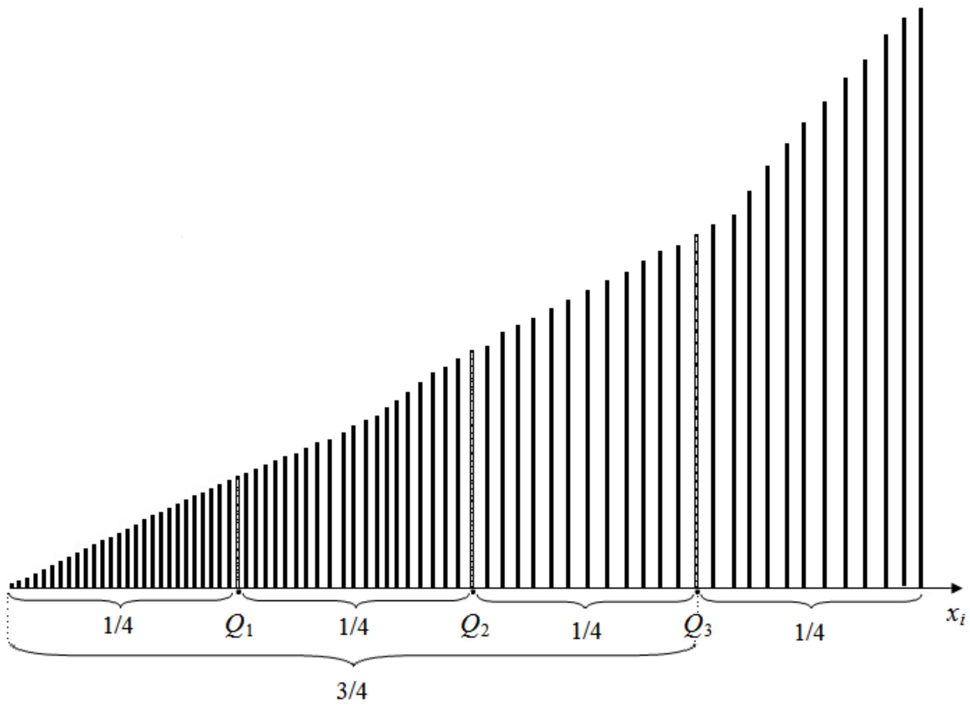

Рис. 3. Квартили.

Например, чтобы определить какой уровень признака является границей, отделяющей первые 75 % наблюдений с самыми меньшими значениями признака от остальных 25 % наблюдений с большими значениями, для анализируемой совокупности следует рассчитать 75-% *квартиль* (или 3-й квартиль, Q_3).

В случае дискретного ряда Q_3 (75-% квартильное значение признака x_i) определяется на основании накопленных частот. В задаче 1.25. из всей совокупности 1 620 граждан, упорядоченных по количеству совершаемых ими поездок, 75 % будет соответствовать 1 215 гражданин (1 620 · 0,75 = 1 215). По накопленным частотам найдем, что этот гражданин совершает 4 поездки в день, т.е. 75-% квартильное значение признака $x_i = 4$ ($Q_3 = 4$).

В случае интервального ряда Q_3 (или 75%-й квартиль) определяется по формуле:

$$Q_3 = x_{Q_3} + h_{Q_3} \cdot \frac{0{,}75 \cdot \sum f_i - S_{Q_3 - 1}}{f_{Q_3}}, \qquad (1.4)$$

где x_{Q_3} – нижняя граница интервала, в котором находится 3-й квартиль;

h_{Q_3} – ширина 3-го квартильного интервала;

Σf_i – число единиц изучаемой совокупности;

S_{Q_3-1} – накопленная частота в интервале, предшествующем 3-му квартильному;

f_{Q_3} – частота в 3-м квартильном интервале.

При вычислении квартильных, квантильных, децильных, или перцентильных значений признака в интервальном ряду также как и при нахождении медианы неважно имеем ли мы дело с равноинтервальным или неравноинтервальным рядом, формула расчета не меняется.

Принцип нахождения 75%-го квартильного интервала аналогичен поиску медианного интервала. По накопленным частотам определяют в какой из интервалов попадает 75 % изучаемой совокупности, этот интервал и будет 75-% квартильный интервал.

По своей сути формула определения 3-го квартиля отличается от формулы расчета медианы только коэффициентом при Σf_i. При расчете медианы берут половину совокупности – $0{,}5 \cdot \Sigma f_i$, при определении 3-го квартиля $0{,}75 \cdot \Sigma f_i$. При вычислении 10%-го децильного значения в формулу подставим $0{,}10 \cdot \Sigma f_i$, а для расчета 29%-го перцентиля – $0{,}29 \cdot \Sigma f_i$, и т.д.

При вычислении квартильного, квантильного, децильного, или перцентильного значения признака важно не забывать соблюдать условие ранжированности (упорядоченности по возрастанию или убыванию) элементов изучаемой совокупности.

Рассмотрим задачу на нахождение Q_3 (75-% квартильного значения признака x_i) для интервального ряда. Пусть у нас есть холдинг, включающий 100 предприятий, производящих один и тот же вид продукции. При этом себестоимость производства продукции на предприятиях различна:

Номер группы предприятий	Себестоимость производства одного изделия, руб. x_i	Число предприятий f_i	Накопленные частоты S_i
1	79 – 90	12	12

2	90 – 110	15	27
3	110 – 130	49	76
4	130 и выше	24	100
–	–	100	–

Необходимо определить уровень себестоимости производства продукции, который отделяет 75 % предприятий с относительно невысокой себестоимостью от 25 % предприятий с более высокой себестоимостью. Впоследствии на 25 % предприятий с более высокой себестоимостью будут проведены оптимизационные мероприятия с целью снижения себестоимости производства продукции.

В задаче требуется найти 75-% квартильное значение признака x_i – себестоимости производства продукции. Так как признак x_i задан интервальным рядом, определять Q_3 будем по формуле (1.4). Первоначально определим интервал в который попадает Q_3. У нас есть 100 значений признака x_i (для каждого из предприятий указана себестоимость производства продукции), x_i упорядочены по возрастанию. Судя по столбцу накопленных частот признак, соответствующий 75 % всей совокупности (признак под номером 75), попадает в интервал 110 – 130, этот интервал и будет 3-м квартильным интервалом. По формуле (1.4)

$$Q_3 = x_{Q_3} + h_{Q_3} \cdot \frac{0{,}75 \cdot \sum f_i - S_{Q_3 - 1}}{f_{Q_3}} = 110 + 20 \cdot \frac{0{,}75 \cdot 100 - 27}{49} \approx 129{,}59 \text{ руб.}$$

Следовательно, у 75 % предприятий себестоимость не превышает 129,59 руб. за единицу продукции. Оставшиеся 25 % предприятий производят продукцию по более высокой себестоимости, на этих предприятиях проведем оптимизационные мероприятия с целью снижения себестоимости производства единицы продукции.

Задачи для самостоятельного решения по данной теме смотрите во втором параграфе третьей главы (задачи 3.7. – 3.12.).

Глава 2. АНАЛИТИЧЕСКАЯ СТАТИСТИКА

§ 1. Показатели вариации и способы их расчета

Вариация – несовпадение (или различие) уровней одного и того же показателя у разных объектов.

Например, рост Володи 176 см, Ольги 164 см, Ивана 193 см. Т.е. показатель «рост» у разных людей отличается. И это отличие нам необходимо измерить, охарактеризовать силу вариации, ее величину.

Самая приблизительная оценка вариации может быть дана с помощью *размаха (амплитуды) вариации* (R):

$$R = x_{max} - x_{min},$$

где x_{max} и x_{min} – наибольшее и наименьшее значения варьирующего признака.

Размах (амплитуда) вариации характеризует лишь максимальное различие значений признака, не учитывает вариацию между всеми значениями совокупности.

Более строгие показатели вариации – показатели колеблемости относительно среднего уровня признака. К ним относятся:

- среднее линейное отклонение;
- дисперсия;
- среднее квадратическое отклонение;
- коэффициент вариации.

Среднее линейное отклонение (a) – простейший показатель колеблемости относительно среднего уровня признака, вычисляется как среднее арифметическое значение абсолютных отклонений признака от его среднего уровня:

$$a = \frac{\sum|x_i - \bar{x}|}{n},$$

при повторяемости отдельных значений x_i используют формулу:

$$a = \frac{\sum|x_i - \bar{x}|f_i}{\sum f_i},$$

где x_i – значение варьирующего признака в дискретном ряду или середина интервала в интервальном распределении;

f_i – частота варьирующего признака;

$\bar{x}$ – среднее значение варьирующего признака.

Среднее линейное отклонение просто вычислить, удобно интерпретировать, но трудно использовать в математической статистике. Так как модули нельзя поставить в соответствие с каким-либо вероятностным законом.

Потому, в статистике для измерения вариации чаще применяются среднее квадратическое отклонение (параметр нормального распределения) и показатель дисперсии.

Дисперсия признака (σ^2) определяется на основе квадратической степенной средней[11]:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n},$$

при повторяемости отдельных значений x_i используют формулу:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i}.$$

Показатель σ, равный $\sqrt{\sigma^2}$ (квадратному корню из дисперсии), называется *средним квадратическим отклонением*, и вычисляется по формуле:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}},$$

при повторяемости отдельных значений x_i используют формулу:

[11] Формулу для расчета дисперсии можно преобразовать *методом моментов*, приведя к виду: $\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \overline{x^2} - (\bar{x})^2$ или $\sigma^2 = \frac{\sum x_i^2 f_i}{\sum f_i} - \left(\frac{\sum x_i f_i}{\sum f_i}\right)^2 = \overline{x^2} - (\bar{x})^2$ (при повторяемости значений признака x_i). Здесь $\overline{x^2}$ – среднее значение квадратов признака (или начальный момент второго порядка), $\bar{x}$ – среднее значение признака (или начальный момент первого порядка).

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}}.$$

Среднее квадратическое отклонение и среднее линейное отклонение показывают на сколько в среднем отклоняются конкретные варианты признака от его среднего значения.

Коэффициент вариации (V) – дает относительную оценку вариации, рассчитывается как отношение среднего квадратического отклонения к среднему значению признака. Или как отношение среднего линейного отклонения к медиане. Результат записывается в виде коэффициента или в процентах:

$$V = \frac{\sigma}{\bar{x}} \left(\text{или } V = \frac{\sigma}{\bar{x}} \cdot 100\,\%\right) \text{ или } V = \frac{a}{Me} \left(\text{или } V = \frac{a}{Me} \cdot 100\,\%\right).$$

Рассмотрим расчет показателей вариации на примере.

2.1. По данным об объемах сбора урожая кукурузы (тыс. тонн) по 12 компаниям агрохолдинга:

18	18	18	29	29	29
37	37	37	37	46	46

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

Решение.

1) дискретный ряд распределения имеет вид:

x_i – значение признака (объем сбора урожая, в тыс. тонн)	f_i - частота признака
18	3
29	3
37	4
46	2

2) среднее квадратическое отклонение (σ) определим по формуле:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i}},$$

где $\bar{x}$ – среднее значение варьирующего признака.

$\bar{x}$ вычислим по формуле средней арифметической взвешенной:

$$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f_i}.$$

Расчеты представлены в таблице:

x_i – значение признака (объем сбора урожая)	f_i - частота признака	$x_i \cdot f_i$	$(x_i - \bar{x})^2 \cdot f_i$
18	3	18 · 3	$(18 - 31{,}75)^2 \cdot 3 = 567{,}1875$
29	3	29 · 3	$(29 - 31{,}75)^2 \cdot 3 = 22{,}6875$
37	4	37 · 4	$(37 - 31{,}75)^2 \cdot 4 = 110{,}25$
46	2	46 · 2	$(46 - 31{,}75)^2 \cdot 2 = 406{,}125$
Итого	$\sum f_i = \mathbf{12}$	$\sum x_i \cdot f_i = \mathbf{381}$	$\sum(x_i - \bar{x})^2 \cdot f_i = \mathbf{1\ 106{,}25}$

$$\bar{x} = \frac{381}{12} = 31{,}75 \text{ (тыс. тонн)},$$

т.е. в среднем каждая из 12-ти агрокомпаний собирает 31,75 тыс. тонн кукурузы.

$$\sigma = \sqrt{\frac{1\ 106{,}25}{12}} = 9{,}601 \text{ (тыс. тонн)}$$

т.е. среднее квадратическое отклонение составило 9,601 тыс. тонн.

Коэффициент вариации определим по формуле:

$$V = \frac{\sigma}{\bar{x}}.$$

Таким образом, коэффициент вариации

$$V = \frac{9{,}601}{31{,}75} = 0{,}302 \text{ или } 30{,}2\ \%.$$

2.2. По данным об объемах производства продукции (тыс. тонн) по 14 компаниям холдинга:

54	54	54	75	75	75	75
64	64	76	76	76	76	76

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

Решение.

1) дискретный ряд распределения имеет вид:

x_i – значение признака (объем производства)	f_i - частота признака
54	3
64	2
75	4
76	5

2) Расчеты представлены в таблице:

x_i – значение признака (объем производства)	f_i - частота признака	$x_i \cdot f_i$	$(x_i - \bar{x})^2 \cdot f_i$
54	3	54 · 3	(54 – 69,286)2 · 3 = 700,9592
64	2	64 · 2	(64 – 69,286)2 · 2 = 55, 8776
75	4	75 · 4	(75 – 69,286)2 · 4 = 130,6122
76	5	76 · 5	(76 – 69,286)2 · 5 = 225,4082
Итого	$\sum f_i =$ **14**	$\sum x_i \cdot f_i =$ **970**	$\sum (x_i - \bar{x})^2 \cdot f_i =$ **1 112,875**

$$\bar{x} = \frac{970}{14} = 69{,}286 \text{ (тыс.тонн)},$$

т.е. в среднем каждая из 14-ти компаний холдинга производит 69,286 тыс. тонн продукции.

$$\sigma = \sqrt{\frac{1\,112{,}875}{14}} = 8{,}916 \text{ (тыс.тонн)}$$

т.е. среднее квадратическое отклонение составило 8,916 тыс. тонн.

Коэффициент вариации

$$V = \frac{8{,}916,}{69{,}2686} = 0{,}129 \text{ или } 12{,}9\ \%.$$

2.3. По данным об объемах сбора урожая кукурузы (тыс. тонн) по 10 компаниям агрохолдинга:

5	5	5	12	12
18	18	18	18	18

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

2.4. По данным об объемах производства продукции (тыс. тонн) по 8 компаниям холдинга:

10	10	10	12
8	8	8	12

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

2.5. По данным об объемах сбора урожая кукурузы (тыс. тонн) по 10 компаниям агрохолдинга:

5	5	10	10	10
12	12	12	12	12

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

2.6. По данным об объемах производства продукции (тыс. тонн) по 8 компаниям холдинга:

7	7	7	12
9	9	9	12

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

2.7. По данным об объемах сбора урожая кукурузы (тыс. тонн) по 10 компаниям агрохолдинга:

7	7	7	6	6
9	9	9	9	9

1) необходимо построить дискретный ряд распределения;

2) определить среднее квадратическое отклонение и коэффициент вариации.

§ 2. Ряды динамики. Основные показатели изменения уровней ряда.

Ряды динамики (или временные, хронологические ряды) – это последовательность упорядоченных во времени числовых значений определенного статистического показателя.

Например, курс обыкновенных акций ОАО «Газпром» на начало рабочего дня за следующие дни:

Дата	27.11.2017	28.11.2017	29.11.2017	30.11.2017
Курс обыкн. акций (руб.)	133,07	132,88	134,85	133,00

будет являться рядом динамики, так как включает два обязательных элемента – время и конкретное значение показателя (курс акций) на этот момент времени. Числовые значения показателя принято называть уровнями ряда и обозначать через Y. Так, первый член ряда Y_0 (или Y_1) называют начальным уровнем, а последний Y_n – конечным уровнем ряда.

Для анализа направления и размера изменения уровней ряда во времени рассчитывают следующие показатели:

- абсолютный прирост;
- темп роста;
- темп прироста;
- абсолютное значение одного процента прироста.

При этом вычисляют базисные (в этом случае текущей уровень ряда сравнивается с начальным Y_0) и цепные показатели (в это случае сравниваются соседние уровни ряда). Формулы для расчета цепных и базисных показателей приведены в табл. 2.1.

Таблица 2.1[12]

Показатель	Базисный	Цепной
Абсолютный прирост	$\Delta_i^{баз} = Y_i - Y_0$ $\Delta_i^{баз} = \Sigma\, \Delta_i^{цеп}$	$\Delta_i^{цеп} = Y_i - Y_{i-1}$
Коэффициент роста	$K_р^{баз} = Y_i / Y_0$ $K_р^{баз} = \Pi K_р^{цеп}$	$K_р^{цеп} = Y_i / Y_{i-1}$
Темп роста	$Т_р^{баз} = (Y_i / Y_0) \cdot 100\%$	$Т_р^{цеп} = (Y_i / Y_{i-1}) \cdot 100\%$
Коэффициент прироста	$K_{пр}^{баз} = K_р^{баз} - 1$ $K_{пр}^{баз} = (Y_i - Y_0) / Y_0$ $K_{пр}^{баз} = \Delta_i^{баз} / Y_0$	$K_{пр}^{цеп} = K_р^{цеп} - 1$ $K_{пр}^{цеп} = (Y_i - Y_{i-1}) / Y_{i-1}$ $K_{пр}^{цеп} = \Delta_i^{цеп} / Y_{i-1}$
Темп прироста	$Т_{пр}^{баз} = K_{пр}^{баз} \cdot 100\%$ $Т_{пр}^{баз} = Т_р^{баз} - 100\%$ $Т_{пр}^{баз} = K_р^{баз} \cdot 100\% - 100\%$	$Т_{пр}^{цеп} = K_{пр}^{цеп} \cdot 100\%$ $Т_{пр}^{цеп} = Т_р^{цеп} - 100\%$ $Т_{пр}^{цеп} = K_р^{цеп} \cdot 100\% - 100\%$
Абсолютное значение одного процента прироста	*	$A^{цеп} = \dfrac{Y_i - Y_{i-1}}{\frac{Y_i - Y_{i-1}}{Y_{i-1}} \cdot 100\%}$ $A^{цеп} = Y_{i-1} / 100$ $A^{цеп} = \Delta_i^{цеп} / Т_{пр}^{цеп}$ $A^{цеп} = (Y_i - Y_0) / (Т_р^{цеп} - 100\%)$

[12] См.: Харченко Л.П., Долженкова В.Г., Ионин В.Г. и др. Статистика : учебное пособие. – М. : ИНФРА-М, 2002. С. 124.

*Базисный показатель «абсолютное значение одного процента прироста» это постоянная величина равная одной сотой Y_0 ($A^{баз} = \frac{Y_i - Y_0}{\frac{Y_i - Y_0}{Y_0} \cdot 100\%} = \frac{Y_0}{100\%}$), его расчет не имеет особого смысла.

2.8. Рассмотрим расчет вышеперечисленных показателей изменения уровней ряда на примере производства обуви некоторой компанией. Данные охватывают временной интервал с 2005 по 2008 гг.:

Год	2005	2006	2007	2008
Производство обуви, млн руб.	15,2	16,5	17,0	16,2
Абсолютный прирост базисный (Y_0=2005) цепной	– –	16,5-15,2 = 1,3 16,5-15,2 = 1,3	17,0-15,2 =1,8 17,0-16,5 = 0,5	16,2-15,2 = 1,0 16,2-17,0 = –0,8
Коэффициент роста базисный (Y_0=2005) цепной	1,0 –	16,5 / 15,2 = 1,086 16,5 / 15,2 = 1,086	17,0 / 15,2 = 1,118 17,0 / 16,5 = 1,030	16,2 / 15,2 =1,066 16,2 / 17,0 = 0,953
Темп роста (%) базисный (Y_0=2005) цепной	100% –	1,086 · 100 = 108,6 1,086 · 100 = 108,6	1,118 · 100 = 111,8 1,030 · 100 =103,0	1,066 · 100 = 106,6 0,953 · 100 = 95,3
Коэффициент прироста базисный (Y_0=2005) цепной	– –	1,086 – 1 = 0,086 1,086 – 1 = 0,086	1,118 – 1 = 0,118 1,030 – 1 = 0,030	1,066 – 1 = 0,066 0,095 – 1 = –0,047
Темп прироста (%) базисный (Y_0=2005) цепной	– –	0,086 · 100 = 8,6 0,086 · 100 = 8,6	0,118 · 100 = 11,8 0,030 · 100 = 3,0	0,066 · 100 = 6,6 –0,047 · 100 = –4,7
Абсолютное значение одного процента прироста	–	15,2 / 100 = 0,152	16,5 / 100 = 0,165	17,0 / 100 = 0,170

Также для анализа временных рядов вычисляют коэффициент опережения ($K_{оп}$). Данный показатель равен отношению темпов роста (или прироста) двух динамических рядов (для одинаковых отрезков времени). Коэффициент опережения применяют при сопоставлении темпов роста (или прироста) одного и того же показателя, но для разных территорий (объектов) или по разным показателям, при этом в рассмотрение принимается один и тот же отрезок времени.

2.9. Например, рассмотрим данные о вводе в действие жилой площади двух стран и сравним их темпы роста по данному показателю.

Решение. Вычислим цепные темпы роста для каждой из стран, результаты представим в виде коэффициентов (формула для расчета: $Т_р^{цеп} = (Y_i / Y_{i-1})$). На основании полученных данных рассчитаем коэффициент опережения:

$K_{оп} = Т_{р\,А}^{цеп} / Т_{р\,Б}^{цеп}$, где $Т_{р\,А}^{цеп}$ – темп роста (цепной) страны А, $Т_{р\,Б}^{цеп}$ – темп роста (цепной) страны Б. Результаты представлены в таблице:

год	1992	1993	1994	1995	1996	1997	1998	1999
Страна А. Введено млн м2	27,8	29,3	30,9	32,2	34,5	34,2	36,2	37,4
Страна Б. Введено млн м2	54,2	55,4	56,2	57,1	58,2	57,4	59,2	59,9
Страна А. Темп роста (цепной).	–	1,054	1,055	1,042	1,071	0,991	1,058	1,033
Страна Б. Темп роста (цепной).	–	1,022	1,014	1,016	1,019	0,986	1,031	1,012
Коэффициент опережения (Коп)	–	1,031	1,040	1,026	1,051	1,005	1,026	1,021

На основании приведенных расчетов можно заключить, что страна А вводит в действие жилую площадь темпами опережающими страну Б.

Обобщающей характеристикой динамического ряда служит средняя величина. Средние значения показателей абсолютный прирост, коэффициент роста, темп роста и темп прироста вычисляются по формулам:

- средний абсолютный прирост

$\bar{\Delta}= \Delta_i^{\text{баз}}/n$ или $\bar{\Delta}= \Delta_i^{\text{цеп}}/n$, где n – количество периодов;

- средний коэффициент роста (геометрический)

$\bar{K}_{\text{р}} = \sqrt[n]{\prod K_{\text{р}}^{\text{цеп}}} = \sqrt[n]{K_{\text{р}}^{\text{баз}}} = \sqrt[n]{\frac{Y_n}{Y_0}}$, где n – число коэффициентов (или число периодов за которые определяется средний коэффициент), П – знак произведения;

- средний коэффициент роста (параболический)[13]

$$\frac{\sum_{i=1}^{n} Y_i}{Y_0} = \bar{K}_{\text{р п}} + \bar{K}_{\text{р п}}^2 + \bar{K}_{\text{р п}}^3 + \cdots + \bar{K}_{\text{р п}}^n$$

- средний темп роста

$$\bar{\text{T}}_{\text{р}} = \bar{K}_{\text{р}} \cdot 100\%$$

- средний темп прироста

$$\bar{\text{T}}_{\text{пр}} = \bar{\text{T}}_{\text{р}} - 100\%.$$

Средний коэффициент роста (геометрический) зависит от значений крайних уровней ряда Y_n и Y_0. Потому, при анализе неоднородных по характеру изменения периодов их принято разбивать на более однородные (с похожей динамикой уровней) и рассчитывать средние для каждого из них. Иначе можно получить бессмысленный результат.

2.10. Например, рассмотрим данные об остатках вкладов населения в некотором банке А за первое полугодие 2019 года (на начало месяца):

Месяц	янв.	фев.	март	апр.	май	июнь	июль
Сумма вкладов,	27,6	28,9	29,7	30,5	30,7	30,9	27,4

[13] См.: Громыко Г.Л., Воробьев А.Н., Карасева Л.А. и др. Теория статистики : Учебник. – М. : ИНФРА-М, 2005. С. 302.

млрд руб.							

Необходимо определить средний за месяц коэффициент роста вкладов. Решение. Вычислим средний коэффициент роста по геометрической формуле: $\overline{K}_{\text{р}} = \sqrt[n]{\frac{Y_n}{Y_0}} = \sqrt[6]{\frac{27{,}2}{27{,}4}} = 0{,}999$ (или 99,9 %) что говорит о ежемесячном на 0,1 % снижении суммы вкладов. Вместе с тем каждый месяц, кроме последнего, сумма вкладов увеличивалась. Очевидно, что в данном случае расчет среднего коэффициента роста по геометрической формуле дает бессмысленный результат и целесообразно или 1) разделить рассматриваемый период на интервалы с похожей динамикой уровней и рассчитать среднее для каждого из них или 2) воспользоваться для нахождения среднего коэффициента роста параболической формулой[14]. В данном примере мы воспользуемся 1) подходом.

Рассматриваемое полугодие удобно разделить на две части – с января по июнь, и с июня по июль. Значение среднего коэффициента роста с января по июнь составляет $\overline{K}_{\text{р}} = \sqrt[n]{\frac{Y_n}{Y_0}} = \sqrt[5]{\frac{30{,}9}{27{,}6}} = 1{,}023$ что говорит о ежемесячном росте суммы вкладов на 2,3 % в течение первых пяти месяцев первого полугодия 2019 г. А затем, в течение последнего месяца полугодия, сумма вкладов сократилась на 11,3 % $\left(\overline{K}_{\text{р}} = \sqrt[n]{\frac{Y_n}{Y_0}} = \sqrt[1]{\frac{27{,}4}{30{,}9}} = 0{,}887 \text{ (или 88,7 \%)}\right)$.

2.11. Имеются данные о производстве и распределении сельскохозяйственной продукции РФ за 2005 и 2009 гг.:

Показатель	млрд руб.			коэффицент роста	темп прироста (%)	среднегодовой коэфф. роста	абсолютное значение 1 % прироста
	2005	2009	абсолютн. прирост				

[14] См.: Громыко Г.Л., Воробьев А.Н., Карасева Л.А. и др. Теория статистики : Учебник. – М. : ИНФРА-М, 2005. С. 303.

Объем производства	980			1,250			
Экспорт	150	255				1,142	
Внутреннее потребление		760	-40		-5,000		
Запасы	30				600,000		

Необходимо заполнить недостающие значения.

Решение. В нашей таблице присутствует среднегодовой показатель – «среднегодовой коэффициент роста» для которого размер интервала между уровнями равен 1 году. Потому, все остальные показатели, с интервалом в четыре года, будут являться базисными, т.е. мы будем соотносить текущее значение Y_{2009} с базисным Y_{2005}.

Абсолютный прирост (базисный) = $\Delta_i^{баз} = Y_i - Y_0 = Y_{2009} - Y_{2005}$, откуда для показателя «Экспорт» $\Delta_i^{баз} = 255 - 150 = 105$ млрд. руб.;
для показателя «Внутреннее потребление» $Y_{2005} = Y_{2009} - \Delta_i$ баз $= 760 - (-40) = 800$ млрд. руб.

Коэффициент роста (базисный) = $K_р^{баз} = Y_i / Y_0 = Y_{2009} / Y_{2005}$, откуда для показателя «Экспорт» $K_р^{баз} = Y_{2009} / Y_{2005} = 255 / 150 = 1{,}7$;
для показателя «Объем производства» $K_р^{баз} = Y_{2009} / Y_{2005} = 760 / 800 = 0{,}95$;
для показателя «Внутреннее потребление» $Y_{2009} = Y_{2005} \cdot K_р^{баз} = 1{,}25 \cdot 980 = 1\,225$ млрд. руб.

Темп прироста (базисный) = $T_{пр}^{баз} = K_р^{баз} \cdot 100\% - 100\%$, откуда для показателя «Объем производства» $T_{пр}^{баз} = 1{,}25 \cdot 100\% - 100\% = 25\ \%$;

для показателя «Экспорт» $T_{пр}^{баз} = 1{,}7 \cdot 100\% - 100\% = 70\%$;

для показателя «Запасы» $K_р^{баз} = (T_{пр}^{баз} + 100\%) / 100\% = (600 + 100) / 100 = 7$.

Средний коэффициент роста (в нашем случае средний за год) = $\overline{K}_р = \sqrt[n]{K_р^{баз}}$, где n – число периодов за которые определяется средний коэффициент, в нашем случае $n = 2009 - 2005 = 4$. Откуда для показателя

«Объем производства» $\overline{K}_{р} = \sqrt[n]{K_{р}^{баз}} = \sqrt[4]{1{,}25} = 1{,}057$ (округляем до третьего знака после запятой, придерживаемся исходного формата таблицы);

для показателя «Внутреннее потребление» $\overline{K}_{р} = \sqrt[n]{K_{р}^{баз}} = \sqrt[4]{0{,}95} = 0{,}987$;

для показателя «Запасы» $\overline{K}_{р} = \sqrt[n]{K_{р}^{баз}} = \sqrt[4]{7} = 1{,}627$.

Далее, используя вышеприведенные формулы, вычисляем оставшиеся показатели:

абсолютный прирост (базисный) для показателя «Объем производства» = $\Delta_i^{баз}$ = 1 225 – 980 = 245 млрд. руб.;

для показателя «Запасы» $Y_{2009} = Y_{2005} \cdot K_{р}^{баз} = 7 \cdot 30 = 210$ млрд. руб. и абсолютный прирост (базисный) = $\Delta_i^{баз}$ = 210 – 30 = 180 млрд. руб.

Результаты вычислений представлены в таблице ниже, выделены подчеркиванием:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2005	2009	абсолютн. прирост			
Объем производства	980	<u>1 225</u>	<u>245</u>	1,250	<u>25,000</u>	<u>1,057</u>
Экспорт	150	255	<u>105</u>	<u>1,700</u>	<u>70,000</u>	1,142
Внутреннее потребление	<u>800</u>	760	-40	<u>0,950</u>	-5,000	<u>0,987</u>
Запасы	30	<u>210</u>	<u>180</u>	<u>7,000</u>	600,000	<u>1,627</u>

2.12. Заполните недостающие показатели о производстве и распределении сельскохозяйственной продукции РФ за 2012 и 2016 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2012	2016	абсолютн. прирост			
Объем производства	740	962				1,068

Экспорт	250				68,000	
Внутреннее потребление		462		1,050		
Запасы	50				60,000	1,125

2.13. Заполните недостающие показатели о производстве и распределении наукоемкой продукции РФ за 2005 и 2010 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2012	2016	абсолютн. прирост			
Объем производства	1 770	2 124				1,018
Экспорт	420		63		15,000	
Внутреннее потребление		1 625		1,300		
Запасы	100				–84,000	

2.14. Заполните недостающие показатели о производстве и распределении продукции химической отрасли РФ за 1996 и 2000 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	1996	2000	абсолютн. прирост			
Объем производства	740			1,250		
Экспорт		435			50,000	
Внутреннее потребление	400	420				1,012
Запасы	50		20		40,000	

2.15. Заполните недостающие показатели о производстве и распределении сельскохозяйственной продукции РФ за 2009 и 2015 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2009	2015	абсолютн. прирост			

Объем производства	990			1,300		
Экспорт	150	255				1,092
Внутреннее потребление		760	–40		–5,000	
Запасы	40				580,000	

2.16. Заполните недостающие показатели о производстве и распределении наукоемкой продукции РФ за 1992 и 1996 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	1992	1996	абсолютн. прирост			
Объем производства	980			1,250		
Экспорт	150	240				1,125
Внутреннее потребление		934	134		16,750	
Запасы	30				70,000	

2.17. Заполните недостающие показатели о производстве и распределении продукции химической отрасли РФ за 1996 и 2002 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	1996	2002	абсолютн. прирост			
Объем производства	950			0,920		
Экспорт	160	180				1,020
Внутреннее потребление		646	–114		–15,000	
Запасы	30				60,000	

2.18. Заполните недостающие показатели о производстве и распределении продукции лесной промышленности РФ за 2001 и 2006 гг.:

Показатель	млрд руб.	коэффициент	темп	среднегодовой

	2001	2006	абсолютн. прирост	роста	прироста (%)	коэфф. роста
Объем производства	760		190		25,000	
Экспорт		470		2,000		
Внутреннее потребление	360	414				1,028
Запасы	165				–60,000	

2.19. Заполните недостающие показатели о производстве и распределении сельскохозяйственной продукции РФ за 2010 и 2018 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2010	2018	абсолютн. прирост			
Объем производства	1 690			1,200		
Экспорт		574			40,000	
Внутреннее потребление	1 200	1 200				1,000
Запасы	80		174		217,500	

2.20. Заполните недостающие показатели о производстве и распределении продукции лесной промышленности РФ за 2001 и 2006 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2005	2009	абсолютн. прирост			
Объем производства	740	962				1,030
Экспорт	250				-10,000	
Внутреннее потребление		462		1,050		
Запасы	50				450,000	1,209

2.21. Заполните недостающие показатели о производстве и распределении наукоемкой продукции РФ за 2009 и 2018 гг.:

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2009	2018	абсолютн. прирост			
Объем производства	1800			1,200		
Экспорт		555			20,652	
Внутреннее потребление	1200	1500				1,025
Запасы	140		–35		-25,000	

Глава 3. СТАТИСТИКА В ПРИКЛАДНЫХ ИССЛЕДОВАНИЯХ

§ 1. Индексы.

Индекс – это относительная величина, показывающая, во сколько раз уровень изучаемого явления в данных условиях отличается от уровня того же явления в других условиях. Различие условий может проявляться во времени (тогда говорят об индексах динамики), в пространстве (территориальные индексы) и т.п.[15]

Рассмотрим три основных типа индексов:

- индивидуальные индексы;
- общие (сводные) индексы;
- индексы средних величин.

Наиболее часто понятие «индекс» используется для обозначения относительного изменения какого-либо показателя во времени. При этом индексы вычисляются путем деления значения показателя в текущий (отчетный) период времени на значение этого же показателя в базисный

[15] См.: Харченко Л.П., Долженкова В.Г., Ионин В.Г. и др. Статистика : учебное пособие. – М. : ИНФРА-М, 2002. С. 138.

период времени (база сравнения). Для удобства используют подстрочные обозначения: «0» для базисного периода и «1» для текущего периода.

Показатель, изменение которого характеризует индекс, называют индексируемой величиной.

Индивидуальные индексы принято обозначать символом *i*, они применяются при сравнении уровней отдельных элементов.

Например, индивидуальный индекс цены товара марки А (см. табл. 3.1) составит $i_p = \frac{p_1}{p_0} = \frac{2,1}{1,2} = 1,75,$

где p_1 – цена товара марки А в текущем периоде (2019 г.);

p_0 – цена товара марки А в базисном периоде (2010 г.).

Таблица 3.1

Наименование товара	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	2010	2019	2010	2019
марка А	1,2	2,1	8	30
марка В	2,4	2	16	18

Индивидуальные индексы показывают соотношение между отчетным и базисным показателями, обычно выраженное в процентах, или характеризуют во сколько раз уменьшился или увеличился показатель за рассматриваемый период.

Так, индивидуальный индекс цены товара марки А i_p = 1,75 свидетельствует о росте цены товара марки А за 10 лет на 75 % (1,75 · 100 % – 100 % = 75 %) или в 1,75 раза.

Общие (сводные) индексы, обозначаемые символом *I*, используются при сравнении сложных неоднородных совокупностей, отдельные элементы которых можно сопоставить только после приведения их к общей мере.

Например, рассмотрим предприятие, выпускавшее в 1999 г. резиновые лодки и палатки. К 2009 г. ассортимент выпускаемой продукции полностью изменился, виды продукции, выпускаемые предприятием – сапоги, калоши и тенты (см. табл. 3.2).

Таблица 3.2

Наименование товара	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	1999	2009	1999	2009
резиновые лодки	1,8		10	
палатки	1,2		26	
сапоги		0,321		20
калоши		0,14		12
тенты		1,6		18

Вычисление индивидуальных индексов цены или количества выпускаемой продукции не представляется возможным, так как для каждого вида продукции есть только один показатель, базисного (1999 г.) или отчетного (2009 г.) периодов. Сравнить объемы выпускаемой продукции отчетного и базисного периодов в физических единицах невозможно, они несопоставимы. Проанализировать можно показатели выручки (Q) отчетного и базисного периодов, т.е. приведенные к общей мере (стоимости в руб.) агрегатные показатели.

Обозначим объем товаров через q, цены – через p, тогда выручка в базисном периоде равна $\sum q_0 p_0$, а в отчетном – $\sum q_1 p_1$. Отношение агрегатов, выручки отчетного периода к выручке базисного, дает общий (сводный) индекс выручки:

$$I_Q = \frac{Q_1}{Q_0} = \frac{\sum q_1 p_1}{\sum q_0 p_0} = \frac{20 \cdot 0{,}321 + 12 \cdot 0{,}14 + 18 \cdot 1{,}6}{10 \cdot 1{,}8 + 26 \cdot 1{,}2} = \frac{36{,}9}{49{,}2} = 0{,}75.$$

Таким образом, объем выручки предприятия сократился на 25% (0,75 · 100 % – 100 % = –25 %) в отчетном 2009 г. по сравнению с базисным 1999 г. В абсолютном выражении сокращение выручки составило $\sum q_1 p_1 -$ *q0p0= 36,9–49,2= –12,3 млн руб.*

Индексы средних величин применяются при изучении изменения во времени (или пространстве) среднего показателя. Например, изменение во времени средней по стране цены квадратного метра жилья в новостройках, или сравнение средней заработной платной в разных регионах, и т.п.

При этом среднее значение анализируемого показателя определяется по формуле средней арифметической взвешенной:

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i},$$

где x_i – значение индексируемой величины у отдельных элементов (например, цена квадратного метра жилья в новостройках города i);

f_i – веса, по которым взвешиваются отдельные значения x.

Сравнение средних значений анализируемого показателя в отчетном и базисном периодах ведет к формуле ***индекса переменного состава***:

$$I_{\text{пер. сост}} = \frac{\bar{X}_1}{\bar{X}_0} = \frac{\sum x_1 f_1}{\sum f_1} \bigg/ \frac{\sum x_0 f_0}{\sum f_0}$$

он характеризует динамику изменения среднего показателя за счет изменения двух факторов (x и f) вместе, т.е. за счет изменения значений индексируемой величины у отдельных элементов (x_i) и за счет изменения весов (f_i), по которым взвешиваются отдельные значения x.

Если в индексе переменного состава зафиксировать веса на уровне отчетного периода (f_1), то мы получим ***индекс фиксированного состава***:

$$I_{\text{фикс. сост}} = \frac{\sum x_1 f_1}{\sum f_1} \bigg/ \frac{\sum x_0 f_1}{\sum f_1} = \frac{\sum x_1 f_1}{\sum x_0 f_1}$$

он характеризует динамику изменения среднего показателя только за счет изменения значений индексируемой величины у отдельных элементов (x_i).

Если в индексе переменного состава зафиксировать значения индексируемой величины у отдельных элементов на уровне базисного периода (x_0), то мы получим ***индекс структурных сдвигов***:

$$I_{\text{стр. сдв}} = \frac{\sum x_0 f_1}{\sum f_1} \bigg/ \frac{\sum x_0 f_0}{\sum f_0}$$

он характеризует динамику изменения среднего показателя только за счет изменения весов (f_i), по которым взвешиваются отдельные значения x.

Индексы переменного состава, фиксированного состава и структурных сдвигов связаны соотношением:

$$I_{\text{пер. сост}} = I_{\text{пост. сост}} \cdot I_{\text{стр. сдв}}.$$

Рассмотрим индексы средних величин на примере конкретного показателя – средней цены определенного вида продукции. Обозначим стоимость единицы продукции через p, объем продаж продукции отдельных предприятий (веса) через q. И рассмотрим среднюю цену некоторого товара (например, общая тетрадь, 48 листов) выпускаемого тремя предприятиями холдинга:

Предприятие	Цена за ед., в руб.		Продано, тыс. шт.	
	2001 p_0	2005 p_1	2001 q_0	2005 q_1
А	10,5	12,6	45	47,2
Б	21,5	24,2	15	17,4
В	24,3	22,7	12,5	7,8

Необходимо определить изменение цены единицы продукции и физического объема продаж на каждом предприятии, а также оценить изменение средней цены единицы продукции по трем предприятиям холдинга вместе. При этом требуется оценить изменение средней цены за счет изменения цен на отдельных предприятиях, за счет изменения долей продаж отдельными предприятиями, а также за счет совместного влияния этих двух эффектов.

Решение. Для ответа на вопрос задачи необходимо рассчитать индивидуальные индексы цены и объема продаж по каждому предприятию, а также индексы цены: переменного состава, фиксированного состава и структурных сдвигов.

Индивидуальный индекс цены на каждом предприятии определим по формуле: $i_p = \frac{p_1}{p_0}$.

по предприятию А $i_p^{\text{А}} = \frac{p_1}{p_0} = \frac{12,6}{10,5} = 1,2$ (рост цены на 20 %);

по предприятию Б $i_p^{\text{Б}} = \frac{p_1}{p_0} = \frac{24,2}{21,5} \approx 1,126$ (рост цены на 12,6 %);

по предприятию В $i_p^{\text{В}} = \frac{p_1}{p_0} = \frac{22,7}{24,3} \approx 0,934$ (снижение цены на 6,6 %).

Индивидуальный индекс физического объема продаж на каждом предприятии вычислим по формуле: $i_q = \frac{q_1}{q_0}$.

по предприятию А $i_q^{\text{А}} = \frac{q_1}{q_0} = \frac{47,2}{45} \approx 1,049$ (рост физического объема продаж на 4,9 %);

по предприятию Б $i_q^{\text{Б}} = \frac{q_1}{q_0} = \frac{17,4}{15} = 1,16$ (рост физического объема продаж на 16 %);

по предприятию В $i_q^{\text{В}} = \frac{q_1}{q_0} = \frac{7,8}{12,5} = 0,624$ (снижение физического объема продаж на 37,6 %).

Для определения индекса цены переменного состава ($I_{\text{пер.сост}} = \bar{p}_1/\bar{p}_0$), рассчитаем среднюю цену единицы продукции по трем предприятиям вместе в базисном ($\bar{p}_0$) и отчетном ($\bar{p}_1$) периодах как среднюю арифметическую взвешенную. Так, средняя по трем предприятиям цена в базисном периоде

$$\bar{p}_0 = \frac{\sum p_0 q_0}{\sum q_0} = \frac{10,5 \cdot 45 + 21,5 \cdot 15 + 24,3 \cdot 12,5}{45 + 15 + 12,5} = \frac{1\,098,75}{72,5} \approx 15,155 \text{ руб.,}$$

а средняя цена в отчетном периоде

$$\bar{p}_1 = \frac{\sum p_1 q_1}{\sum q_1} = \frac{12,6 \cdot 47,2 + 24,2 \cdot 17,4 + 22,7 \cdot 7,8}{47,2 + 17,4 + 7,8} = \frac{1\,192,86}{72,4} \approx 16,476 \text{ руб.}$$

Откуда получаем индекс цены переменного состава:

$$I_{\text{пер. сост}} = \bar{p}_1/\bar{p}_0 = \frac{\sum p_1 q_1}{\sum q_1} \Big/ \frac{\sum p_0 q_0}{\sum q_0} = \frac{16,476}{15,155} = 1,087$$

таким образом, средняя по трем предприятиям цена возросла на 8,7 % за рассматриваемый период, с 2001 по 2005 г.

Изменение средней цены могло произойти за счет двух факторов: 1) изменение цен на отдельных предприятиях, 2) изменение долей продаж отдельных предприятий в общем объеме продаж.

Чтобы исключить влияние на среднюю цену изменения долей продаж отдельных предприятий в общем объеме продаж, рассчитаем индекс фиксированного состава (зафиксировав объем продаж на уровне отчетного периода q_1):

$$I_{\text{фикс. сост}} = \frac{\sum p_1 q_1}{\sum q_1} \Big/ \frac{\sum p_0 q_1}{\sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{1\,192{,}86}{10{,}5 \cdot 47{,}2 + 21{,}5 \cdot 17{,}4 + 24{,}3 \cdot 7{,}8}$$
$$= \frac{1\,192{,}86}{1\,059{,}24} = 1{,}126$$

т.о., средняя по трем предприятиям цена единицы продукции выросла на 12,6 % только за счет роста цен на отдельных предприятиях, в нашем случае за счет роста цен единицы продукции на предприятиях А и Б.

Для оценки влияния на среднюю цену изменения долей продаж отдельных предприятий в общем объеме продаж, рассчитаем индекс структурных сдвигов (зафиксировав цены на предприятиях на уровне базисного периода p_0):

$$I_{\text{стр. сдв}} = \frac{\sum p_0 q_1}{\sum q_1} \Big/ \frac{\sum p_0 q_0}{\sum q_0} = \frac{1\,059{,}24}{72{,}4} \Big/ \frac{1\,098{,}75}{72{,}5} = 14{,}630/15{,}155 = 0{,}965$$

таким образом, средняя по трем предприятиям цена единицы продукции снизилась на 3,5 % только за счет структурного сдвига. А именно, за счет увеличения доли продаж (с 0,62 = 45 / 72,5 до 0,65 = 47,2 / 72,4) продукции предприятия А с более низкой ценой и за счет уменьшения доли продаж (с 0,17 = 12,5 / 72,5 до 0,11 = 7,8 / 72,4) продукции предприятия В с более высокой ценой.

Проверим вычисление индексов средней цены:

$$I_{\text{пер. сост}} = 1{,}087 = I_{\text{пост. сост}} \cdot I_{\text{стр. сдв}} = 1{,}126 \cdot 0{,}965 = 1{,}087.$$

3.1. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2018 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	I квартал	II квартал	I квартал	II квартал
А	1,2	2,1	8	30
Б	2,4	2	16	18

Необходимо определить: 1) индивидуальные индексы цены и физического объема продаж на каждом предприятии; 2) индекс цены переменного состава, фиксированного состава и структурных сдвигов.

3.2. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2016 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	I квартал	II квартал	I квартал	II квартал
А	2,4	2,88	8	24
Б	1,2	3,6	16	16

1) определите индивидуальные индексы цены и физического объема продаж на каждом предприятии;

2) введя обозначения p – для цены товара и q – для объема продаж, запишите формулу для расчета индекса цен структурных сдвигов. Вычислите индекс и поясните результаты.

3.3. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2019 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	I квартал	II квартал	I квартал	II квартал
А	3	2,7	15	12
Б	2,4	3	10	15

1) определите индивидуальные индексы цены и физического объема продаж на каждом предприятии;

2) введя обозначения p – для цены товара и q – для объема продаж, запишите формулу для расчета индекса цен фиксированного (постоянного) состава. Вычислите индекс и поясните результаты.

3.4. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2017 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	II квартал	III квартал	II квартал	III квартал
А	1,2	2,4	8	30
Б	2,4	1,6	16	18

1) определите индивидуальные индексы цены и физического объема продаж на каждом предприятии;

2) введя обозначения p – для цены товара и q – для объема продаж, запишите формулу для расчета индекса цен переменного состава. Вычислите индекс и поясните результаты.

3.5. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2015 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	I квартал	IV квартал	I квартал	IV квартал
А	1,5	2,5	25	15
Б	2,5	3,2	25	10

1) определите индивидуальные индексы цены и физического объема продаж на каждом предприятии;

2) введя обозначения p – для цены товара и q – для объема продаж, запишите формулу для расчета индекса цен структурных сдвигов. Вычислите индекс и поясните результаты.

3.6. Имеется информация о цене и объемах продаж одного вида продукции на двух предприятиях холдинга за два квартала 2019 г.:

Предприятие	Цена за ед., в тыс. руб.		Продано, тыс. шт.	
	I квартал	III квартал	I квартал	III квартал
А	1,8	2,6	9	18
Б	3,6	3	18	27

Необходимо определить: 1) индивидуальные индексы цены и физического объема продаж на каждом предприятии; 2) индекс цены переменного состава, фиксированного состава и структурных сдвигов.

§ 2. Коэффициенты дифференциации.

Существует множество коэффициентов дифференциации, мы рассмотрим те из них, что строятся с применением структурных характеристик вариационного ряда.

Структурные характеристики вариационного ряда, квартили, квантили, децили, и перцентили применяют при изучении дифференциации. Так, при изучении дифференциации доходов населения вычисляют *децильный коэффициент дифференциации* (Кд). Кд равен отношению девятого дециля к первому децилю и показывает во сколько раз минимальные доходы 10 % наиболее обеспеченного населения превышают максимальные доходы 10 % наименее обеспеченного населения страны (области, края, и др.).

Рассмотрим вычисление децильного коэффициента дифференциации на примере.

3.7. Для данных таблицы вычислите показатель, характеризующий, во сколько раз минимальные доходы 10 % наиболее состоятельных граждан превышают максимальные доходы 10 % наименее обеспеченных граждан некоторой страны А.

Среднедушевой совокупный доход,	Количество граждан в соответствующей группе дохода, в	Накопленные частоты,

\$ в мес. x_i	млн чел. f_i	млн чел. S_i
до 7 000	20	20
7 000 - 21 000	30	50
21 000 - 41 000	26	76
41 000 - 101 000	16	92
свыше 101 000	8	100

Решение. В задаче требуется найти децильный коэффициент дифференциации. Все граждане страны у нас упорядочены по доходу (по возрастанию) – от «до 7 000» до «свыше 101 000» $ в месяц, все граждане в совокупности 100 млн чел. составляют 100 %.

Возьмем первую 1 / 10 часть наименее обеспеченного населения страны (от 0 % до 10 % населения), в ней гражданин соответствующий 10 % (самый крайний по величине дохода в этой группе) будет иметь максимальный доход из всех граждан данной группы. Для нахождения уровня его дохода необходимо вычислить 10-% дециль. Формулы для определения децилей в интервальном ряду схожи с формулой для нахождения медианы и имеют вид:

$$\text{Первый дециль} = \text{Д}_1 = x_{\text{Д}_1} + h_{\text{Д}_1} \cdot \frac{0{,}10 \cdot \sum f_i - S_{\text{Д}_1 - 1}}{f_{\text{Д}_1}}$$

Первый дециль (или 10-% дециль) показывает уровень самого высокого дохода у 1 / 10 части (0–10 %) самого малообеспеченного населения страны)

$$\text{Девятый дециль} = \text{Д}_9 = x_{\text{Д}_9} + h_{\text{Д}_9} \cdot \frac{0{,}90 \cdot \sum f_i - S_{\text{Д}_9 - 1}}{f_{\text{Д}_9}}$$

Девятый дециль (или 90-% дециль) показывает уровень самого низкого дохода у 90–100 % самого зажиточного населения страны). В формулах

$x_{\text{Д}_1}$ / $x_{\text{Д}_9}$– нижняя граница интервала, в котором находится 1-й / 9-й дециль;

$h_{\text{Д}_1}$/ $h_{\text{Д}_9}$ – ширина 1-го / 9-го децильного интервала;

Σf_i – число единиц изучаемой совокупности;

$S_{Д_1-1}$ / $S_{Д_9-1}$ – накопленная чистота в интервале, предшествующем 1-му / 9-му децильному интервалу;

$f_{Д_1}$ / $f_{Д_9}$ – частота в 1-м / 9-м децильном интервале.

Для нахождения интервала, содержащего первый дециль, используют накопленные частоты. Номер единицы совокупности (гражданин страны А), соответствующий первому децилю, т.е. 10 %, попадает в первый интервал, т.к. 20 % всего населения страны А получают доход до 7 000 $ в мес.

Ширина открытого интервала «до 7 000 $ в мес.» определяется по ближайшему к нему интервалу 7 000 – 21 000 $ в мес. Т.о. ширина открытого интервала должна была бы быть равна 14 000 $ в мес., и интервал имел бы вид –7 000–7 000 $ в мес.; но т.к. доход не может быть отрицательным, отрытый интервал примет вид 0–7 000 $ в мес., и ширина интервала будет равна 7 000 $ в мес. ($x_{Д_1} = 0$, $h_{Д_1} = 7\,000$ $ в мес.). Частота в 1-м децильном интервале $f_{Д_1} = 20$. А накопленная чистота в интервале, предшествующем 1-му децильному интервалу, $S_{Д_1-1} = 0$, т. к. это первый интервал в ряду и ему ничто не предшествует. Таким образом, первый дециль равен:

$$Д_1 = x_{Д_1} + h_{Д_1} \cdot \frac{0{,}10 \cdot \sum f_i - S_{Д_1-1}}{f_{Д_1}} = 0 + 7\,000 \cdot \frac{0{,}10 \cdot 100 - 0}{20} = 3\,500$$

т.е. максимальный доход 10 % наименее обеспеченных граждан страны А составляет 3 500 $ в месяц.

Для нахождения минимального дохода 10 % наиболее состоятельных граждан страны вычислим 9-й дециль (т.е. самый высокий доход 90 % от всех граждан страны и будет соответствовать самому низкому доходу оставшихся 10 % населения страны А).

9-й (90-%) децильный интервал в нашем примере – интервал с доходом 41 000–101 000 $ в мес. (по накопленным частотам часть населения соответствующая 76–92 % получает доход от 41 000 до 101 000 $ в мес., и соответственно гражданин страны соответствующий 90 % всего населения также попадает в данный интервал). Ширина данного интервала $h_{Д_1}$=101 000 – 41 000 = 60 000 \$ в мес., нижняя граница $x_{Д_1}$= 41 000 \$ в мес., частота в 9-ом децильном интервале $f_{Д_9}$= 16, накопленная чистота в интервале, предшествующем 9-му децильному интервалу, $S_{Д_1-1}$= 76. Таким образом, девятый дециль равен:

$$Д_9 = x_{Д_9} + h_{Д_9} \cdot \frac{0{,}90 \cdot \sum f_i - S_{Д_9-1}}{f_{Д_9}} = 41\,000 + 60\,000 \cdot \frac{0{,}90 \cdot 100 - 76}{16}$$
$$= 93\,500$$

т.е. минимальный доход 10 % наиболее обеспеченных граждан страны А составляет 93 500 \$ в месяц.

И мы получаем децильный коэффициент дифференциации

Кд = $Д_9$ / $Д_1$ = 93 500 / 3 500 = 26,714

т.о., минимальные доходы 10 % наиболее состоятельных граждан страны А превышают максимальные доходы 10 % наименее обеспеченных граждан страны в 26,714 раза.

При изучении городского населения бывает полезным рассчитать квантильный коэффициент показывающий во сколько раз минимальная численность населения 20 % наиболее крупнонаселенных городов страны превышает максимальную численность населения 20 % наименее населенных городов данной страны.

Рассмотрим расчет данного показателя на примере.

3.8. Для данных таблицы вычислите во сколько раз минимальная численность населения 20 % наиболее крупнонаселенных городов страны Б превышает максимальную численность населения 20 % самых малых городов данной страны.

численность населения, чел.	количество населенных пунктов с заданной численностью населения (штук)	накопленные частоты, шт.

x_i	f_i	S_i
до 50 000	9	9
50 000 - 100 000	20	29
100 000 - 250 000	29	58
250 000 - 500 000	21	79
500 000 – 1 000 000	10	90
свыше 1 000 000	11	100
–	100	–

Решение. В задаче требуется найти квантильный коэффициент дифференциации. Все населенные пункты страны упорядочены по численности населения (по возрастанию) – от «до 5 0000» до «свыше 1 000 000», все населенные пункты в совокупности 100 штук составляют 100 %.

Для вычисления квантильного коэффициента необходимо вычислить первый (20%-й) и четвертый (80%-й) квантили.

Формулы для определения квантилей в интервальном ряду схожи с формулой для нахождения медианы и имеют вид:

$$\text{Первый квантиль} = K_1 = x_{K_1} + h_{K_1} \cdot \frac{0{,}20 \cdot \sum f_i - S_{K_1 - 1}}{f_{K_1}}$$

Первый квантиль (или 20-% квантиль) показывает максимальную численность населения в 1 / 20 части (0–20 %) самых малочисленных городов страны)

$$\text{Четвертый квантиль} = K_4 = x_{K_4} + h_{K_4} \cdot \frac{0{,}80 \cdot \sum f_i - S_{K_4 - 1}}{f_{K_4}}$$

Четвертый квантиль (или 80-% квантиль) показывает минимальную численность населения в (80–100 %) самых крупнонаселенных городах страны). В формулах

x_{K_1} / x_{K_4}– нижняя граница интервала, в котором находится 1-й / 4-й квантиль;

h_{K_1} / h_{K_4} – ширина 1-го / 4-го квантильного интервала;

Σf_i – число единиц изучаемой совокупности;

S_{K_1-1} / S_{K_4-1} – накопленная чистота в интервале, предшествующем 1-му / 4-му квантильному интервалу;

f_{K_1} / f_{K_4} – частота в 1-м / 4-м квантильном интервале.

Для нахождения интервала, содержащего первый квантиль, используют накопленные частоты. Номер единицы совокупности (населенный пункт страны Б), соответствующий первому квантилю, т.е. 20 %, попадает во второй интервал (по накопленным частотам 9–29 % всех населенных пунктов страны Б имеют численность населения 50 000–100 000 чел., и соответственно населенный пункт соответствующий 20 % всех городов страны Б также попадает в данный интервал). Ширина данного интервала h_{K_1}=100 000 – 50 000 = 50 000 чел., нижняя граница x_{K_1} = 50 000 чел., частота в 1-ом квантильном интервале f_{K_1}= 20, накопленная чистота в интервале, предшествующем 1-му квантильному интервалу, S_{K_1-1} = 9. Таким образом, первый квантиль равен:

$$K_1 = x_{K_1} + h_{K_1} \cdot \frac{0{,}20 \cdot \sum f_i - S_{K_1-1}}{f_{K_1}} = 50\,000 + 50\,000 \cdot \frac{0{,}20 \cdot 100 - 9}{20}$$

$$= 77\,50$$

т.е. максимальная численность населения 20 % наименее населенных городов страны Б составляет 77 500 чел.

Для нахождения минимальной численности населения 20 % наиболее крупнонаселенных городов страны Б вычислим 4-й квантиль (т.о. максимальная численность 80 % всех городов страны и будет соответствовать минимальной численности оставшихся 20 % городов страны Б).

4-й (80-%) квантильный интервал в нашем примере – интервал с численностью населения 500 000–1 000 000 чел. (по накопленным частотам 79–89 % всех населенных пунктов страны Б имеют численность населения 500 000–1 000 000 чел., и соответственно населенный пункт соответствующий 80 % всех городов страны Б также попадает в данный интервал). Ширина данного интервала $h_{К_4}$ =1 000 000 – 500 000 = 500 000 чел., нижняя граница x_{K_4}= 500 000 чел., частота в 4-ом квантильном интервале f_{K_4}= 10, накопленная чистота в интервале, предшествующем 4-му квантильному интервалу, S_{K_4-1} = 79. Таким образом, четвертый квантиль равен:

$$K_4 = x_{K_4} + h_{K_4} \cdot \frac{0{,}80 \cdot \sum f_i - S_{K_4-1}}{f_{K_4}} = 500\,000 + 500\,000 \cdot \frac{0{,}80 \cdot 100 - 79}{10}$$
$$= 550\,000$$

т.е. минимальная численность населения 20 % самых крупнонаселенных городов страны Б составляет 550 000 чел.

И мы получаем квантильный коэффициент

$$K_K = K_4 \,/\, K_1 = 550\,000 \,/\, 77\,500 = 7{,}097$$

т.о., минимальная численность населения 20 % наиболее крупнонаселенных городов страны Б превышает максимальную численность населения 20 % самых малрнаселенных городов данной страны в 7,097 раза.

3.9. Для данных таблицы вычислите децильный коэффициент дифференциации доходов населения, характеризующий, во сколько раз минимальные доходы 10 % наиболее состоятельных граждан превышают максимальные доходы 10 % наименее обеспеченных граждан страны.

среднедушевой совокупный доход, $ в мес.	количество граждан в соответствующей группе дохода, млн чел.
до 5 000	10
5 000 - 25 000	26
25 000 - 45 000	31
45 000 - 100 000	17
свыше 100 000	16

3.10. Для данных таблицы вычислите во сколько раз минимальные доходы 10 % наиболее состоятельных граждан превышают максимальные доходы 10 % наименее обеспеченных граждан страны.

среднедушевой совокупный доход, $ в мес.	количество граждан в соответствующей группе дохода (в млн чел.)
до 5 000	7
5 000 – 21 000	30
21 000 – 41 000	37
41 000 – 101 000	20
свыше 101 000	6

3.11. Для данных таблицы вычислите во сколько раз минимальная численность населения 20 % наиболее крупнонаселенных городов страны В превышает максимальную численность населения 20 % самых малых городов данной страны.

численность населения, чел.	количество населенных пунктов с заданной численностью населения (штук)
до 50 000	8
50 000 – 150 000	10
150 000 – 250 000	14
250 000 – 450 000	18
450 000 – 750 000	25
свыше 750 000	25

3.12. Для данных таблицы вычислите квантильный коэффициент характеризующий во сколько раз минимальная численность населения 20 % наиболее крупнонаселенных городов страны Г превышает максимальную численность населения 20 % самых малых городов данной страны.

численность населения, чел.	количество населенных пунктов с заданной численностью населения (штук)
до 100 000	25
100 000 – 200 000	22
200 000 – 300 000	19

300 000 – 500 000	16
500 000 – 800 000	12
свыше 800 000	6

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

Громыко, Г.Л. Теория статистики : Учебник / под ред. проф. Г.Л. Громыко.– М. : ИНФРА-М, 2005. – 476 с.

Харченко, Л.П. Статистика : учебное пособие / Л.П. Харченко, В.Г. Долженкова, В.Г. Ионин и др. ; под ред. В.Г. Ионина. – М. : ИНФРА-М, 2002. – 384 с.

ОТВЕТЫ

Глава 1.

1.2. Средний балл по статистике 72 балла.

1.4. Средняя стоимость билета Мск-Екб = 2 596,053 руб. ≈ 2 596,05 руб.

1.5. Средний размер депозита в банке = 610 086,75 руб.

1.6. Средняя ставка по кредитному портфелю 2017 г = 11,2 %.

1.9. Средняя величина ошибки вашей догадки = $\sqrt{34} \approx \pm 5{,}83$ очка.

1.10. Среднее время отклонения трамвая от расписания составит $\sqrt{18} = \pm 4{,}243$.

1.12. Средняя стоимость билета Екатеринбург – Санкт-Петербург 2 970,779 руб. ≈ 2 970,78 руб.

1.13. Средний размер депозита 760 004,67 руб.

1.14. Средний за день курс продажи Евро в банке = 65 руб. за 1 евро.

1.15. Примечание: средняя производительность труда одного рабочего цеха в тыс. руб. это стоимость продукции, в среднем произведенная каждым сотрудником цеха. Средняя производительность труда одного рабочего по всему предприятию (по трем цехам предприятия вместе) ≈ 25,49 тыс. руб.

1.18. Средний темп роста ВВП составит $\sqrt[5]{1{,}19016} \approx 1{,}0354$.

1.22. Средняя стоимость основных фондов (ОФ) предприятия за первый квартал равна 129 млн руб.

1.25. $Mo = 2$, $Me = 3$.

1.27. $As = -0{,}793$; $Ex = -0{,}236$; плосковершинное распределение, скошено влево.

1.30. $Mo = 2$, $\bar{X} = 2{,}625, As > 0$ (правосторонняя асимметрия). О знаке показателя асимметрии несложно догадаться по коэффициенту асимметрии Пирсона ($As_П$). А затем вспомнить, что знак коэффициента асимметрии и коэффициента асимметрии Пирсона совпадает.

1.31. $Mo = 3, \bar{X} = 2{,}375, As < 0$ (левосторонняя асимметрия).

1.32. $Mo = 3, \bar{X} = 2{,}625, As < 0$ (левосторонняя асимметрия).

Глава 2.

2.3. 1)

x_i	f_i
5	3
12	2
18	5

2) 5,647; 0,438.

2.4. 1)

x_i	f_i
10	3
12	2
8	3

2) 1,561; 0,160.

2.5. 2) 2,646; 0,265.

2.6. 2) 1,936; 0,215.

2.7. 2) 1,249; 0,160.

2.21.

Показатель	млрд руб.			коэффициент роста	темп прироста (%)	среднегодовой коэфф. роста
	2009	2018	абсолютн. прирост			
Объем производства	1800	2160	360	1,200	20,000	1,020
Экспорт	460	555	95	1,207	20,652	1,021
Внутреннее потребление	1200	1500	300	1,250	25,000	1,025
Запасы	140	105	-35	0,750	-25,000	0,969

Глава 3.

3.1. 1) 1,75; 0,83(3); 3,75; 1,125; 2) 1,03125; 1,25; 0,825.

3.2. 1) 1,2; 3; 3; 1; 2) 1,2.

3.3. 1) 0,9; 1,25; 0,8; 1,5; 2) 1,075.

3.4. 1) 2; 0,66(6); 3,75; 1,125; 2) 1,05.

3.5. 1) 1,66(6); 1,28; 0,6; 0,4; 2) 0,95.

3.6. 1) 1,44(4); 0,83(3); 2; 1,5; 2) 0,946(6); 0,986(1); 0,96.

3.9. Кд = 24,125 раза.

3.10. Кд ≈ 13,485 раза.

3.11. K_K = 3 раза.

3.12. K_K =5,9375 раза.

Электронный текстовый ресурс

Кеткина Ольга Сергеевна

СТАТИСТИКА: СБОРНИК ЗАДАНИЙ С ПРИМЕРАМИ РЕШЕНИЙ И ПОЯСНЕНИЯМИ

Подготовка к публикации **О.С. Кеткиной**

Компьютерная верстка **О.С. Кеткиной**

Рекомендовано Учебно-методическим советом УрФУ
Разрешено к публикации
Электронный формат – pdf
Объем

Уральский федеральный университет имени первого Президента России Б.Н.Ельцина

620002, Екатеринбург, ул. Гоголя, 25

Портал информационно-образовательных ресурсов УрФУ
http://study.urfu.ru