О.С. Балаш, Е.В. Чистопольская

Статистика

Учебное пособие

для студентов, обучающихся по направлению «Экономика» и «Менеджмент» по учебной дисциплине «Статистика»

Саратов
2018

УДК 31:33
ББК 65.051.я73
Б20

Рецензенты:

Заведующий кафедрой статистики институт РЭУ им. Плеханова, д.э.н, доцент Саратовского социально-экономического института (филиал) ФГБОУ ВПО «Российского экономического университета имени Г.В. Плеханова» Толмачев М.Н.

Милованов Д.И., к.э.н., доцент кафедры менеджмента и маркетинга Саратовского государственного университета им. Н.Г. Чернышевского

Одобрено учебно-методической комиссией

Балаш О.С., Чистопольская. Е.В.
Б 20 Статистика: Учебное пособие / О.С. Балаш, Е.В. Чистопольская. - Саратов: Изд-во СГУ, 2018. – 79 с.
ISBN

Учебное пособие подготовлено для студентов, обучающихся по направлению «Экономика» и «Менеджмент» по учебной дисциплине «Статистика». Содержит краткий обзор основных понятий общей теории статистики по всем главам представлено решение типовых примеров. Учебное пособие включает задачи по дисциплине, приведены пояснения наиболее трудных задач. В приложениях даны математико-статистические таблицы.

УДК 31:33
ББК 65.051.я73
Пособие издано в авторской редакции

ISBN

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

В настоящее время статистика стала одним из важнейших инструментов анализа и управления экономикой страны. Статистические данные отображают развитие отдельных сторон жизни общества, служат информационной базой принятия решений и в результате дают возможность увидеть систему взаимосвязей в экономике, прогнозировать динамику развития, делать международные сопоставления. Правильное восприятие и использование статистической информации невозможно без знаний такой отрасли статистической науки как общая теория статистики.

Теория статистики является первой частью единого цикла статистических дисциплин, обеспечивающих теоретическую подготовку экономистов, менеджеров, коммерсантов. В курсе изучаются общие категории, методология сбора, обработки статистической информации, основные методы анализа статистического исследования: группировки, выборочный, индексный, корреляционный и регрессионный.

Изучение общей теории статистики требует последовательной и систематической работы. При освоении курса важно овладеть теоретическими положениями и решить задачи, что способствует более глубокому усвоению материала. Изучение каждой темы необходимо осуществлять последовательно после проработки соответствующей литературы.

ПРОГРАММА ДИСЦИПЛИНЫ «СТАТИСТИКА»
РАЗДЕЛ «ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ»

Тема 1. Статистическая сводка, группировка, вариационные ряды, таблицы

Статистическая таблица и ее элементы. Виды таблиц по характеру подлежащего. Виды таблиц по разработке сказуемого. Основные правила построения таблиц. Чтение и анализ таблицы. Таблицы сопряженности.

Статистический график и его элементы. Классификация графиков. Диаграммы сравнения. Структурные диаграммы. Диаграммы динамики. Статистические карты.

Сущность показателя и его атрибуты. Требования, предъявляемые к построению статистических показателей. Классификация статистических показателей. Понятие о системе показателей.

Абсолютные величины, их значение в статистическом исследовании. Виды абсолютных величин и способы их получения. Единицы измерения абсолютных величин.

Относительные величины в статистике. Виды относительных величин, способы их расчета и формы выражения. Взаимосвязи между отдельными видами относительных величин.

Тема 2. Средние величины. Показатели вариации

Средняя, ее сущность. Определение статистической средней. Основные научные положения теории средней. Взаимосвязь метода средних и группировок. Виды и формы средних. Статистические распределения и их основные характеристики.

Средняя арифметическая и ее свойства. Методы вычисления средней арифметической: 1) определение общей средней на основе групповых (частных) средних; 2) вычисление средней арифметической по данным интервального вариационного ряда; 3) вычисление средней арифметической по методу моментов.

Средняя гармоническая. Критерий выбора средней арифметической или средней гармонической. Исходное соотношение средней. Описательные средние (мода и медиана), способы их вычисления для дискретного и интервального вариационных рядов.

Задачи вариационного анализа. Абсолютные и относительные показатели вариации и их значение в статистике. Размах вариации (вариационный размах). Среднее абсолютное (линейное) отклонение. Дисперсия и среднее квадратическое отклонение, методы их вычисления. Свойства среднего квадратического отклонения. Вычисление среднего квадратического отклонения по преобразованной формуле, по методу моментов. Коэффициенты колеблемости и вариации.

Показатели вариации для альтернативного признака.

Виды дисперсий: общая, внутригрупповая и межгрупповая. Правило сложения дисперсий.

Тема 3. Выборочное наблюдение

Понятие о выборочном наблюдении. Причины его применения. Теоретические основы выборочного метода. Ошибки выборочного наблюдения.

Генеральная и выборочная совокупности и их обобщающие характеристики. Виды и способы отбора единиц из генеральной совокупности. Повторная и бесповторная выборка. Собственно случайная выборка. Механическая выборка. Типическая выборка. Серийная выборка.

Определение средней ошибки выборки. Определение предельной ошибки выборки и пределов для средней и доли. Определение необходимой численности выборки. Способы распространения выборочных данных на генеральную совокупность. Малые выборки.

Тема 4. Индексы

Понятие об индексах. Значение индексов в анализе социально-экономических явлений. Классификация индексов. Классификация показателей при построении индексов и их символика. Индивидуальные индексы (базисные и цепные) и их свойства.

Агрегатная форма общих индексов (цен, физического объема товарооборота, товарооборота, физического объема продукции, себестоимости, затрат на производство). Индексируемые величины и проблема их соизмерения. Веса индексов. Индексы Ласпейреса и Пааше. Идеальный индекс цен Фишера. Средняя арифметическая и гармоническая формы общих индексов (цен, физического объема товарооборота, товарооборота). Взаимосвязь между общими индексами и условие ее осуществления. Ряды общих индексов с постоянной и переменной базой сравнения, с переменными и постоянными весами.

Индексный метод анализа динамики среднего уровня. Индексы средней величины переменного состава. Индексы постоянного состава. Индексы структурных сдвигов. Взаимосвязь этих индексов. Индексный метод выявления роли отдельных факторов. Важнейшие экономические индексы, применяемые в статистике. Территориальные индексы.

Тема 5. Ряды динамики

Понятие о рядах динамики. Виды рядов динамики. Основные правила построения рядов динамики и их анализа. Сопоставимость данных рядов динамики. Смыкание рядов динамики.

Графическое изображение рядов динамики.

Аналитические показатели ряда динамики. Средний уровень ряда (средняя хронологическая) и способы его вычисления. Абсолютные приросты (базисные и цепные) и средний абсолютный прирост. Темпы роста (снижения) и прироста. Средний темп роста (снижения). Абсолютное содержание 1% прироста.

Приведение рядов динамики к одному основанию. Коэффициенты опережения.

Компоненты уровня ряда динамики. Основная тенденция ряда и методы ее выявления. Способ скользящей средней. Аналитическое выравнивание.

Анализ рядов динамики. Элементы интерполяции и экстраполяции. Прогнозирование социально-экономических явлений и процессов на основе рядов динамики и регрессионных моделей.

Тема 6. Статистическое изучение связи

Изучение связи – важнейшая задача научного анализа. Виды и формы связей. Важнейшие методы статистики, применяемые в анализе связи между явлениями: метод группировок, балансовый метод и корреляционно-регрессионный метод.

Причинность, регрессия, корреляция. Основные задачи и предпосылки применения корреляционно-регрессионного анализа. Методы изучения корреляционной связи. Уравнение регрессии как форма аналитического выражения статистической связи. Выбор уравнения связи. Показатели тесноты связи. Линейный коэффициент корреляции. Корреляционное отношение.

ОСНОВЫ ОБЩЕЙ ТЕОРИИ СТАТИСТИКИ.

1. СТАТИСТИЧЕСКАЯ СВОДКА, ГРУППИРОВКА, ВАРИАЦИОННЫЕ РЯДЫ, ТАБЛИЦЫ

Статистическое исследование делится на три этапа: 1) статистическое наблюдение; 2) статистическая сводка и группировка; 3) статистический анализ и исчисление обобщающих показателей.

Статистическая сводка – процесс рациональной обработки данных наблюдения с целью приведения их в стройную систему, удобную для статистического анализа и практического использования.

Группировка – метод, позволяющий расчленить совокупность на однородные группы по определенным существенным для них признакам (например, группировки предприятий по организационно-правовым формам, размерам и т.д.).

В зависимости от характера признака различают качественные и количественные группировки. *Качественная* группировка – распределение совокупности по атрибутивному признаку, который характеризует свойство, качество совокупности и не имеет количественного выражения (например, группировки населения по полу, образованию и т.д.). Разновидностью качественных группировок является группировка по альтернативному признаку[1]. *Количественная* группировка – распределение совокупности по величине значений количественного группировочного признака (например, группировка населения по возрасту, группировка рабочих по стажу работы и др.).

Определение числа групп и величины интервала группировки. При построении *качественной группировки* выделяется столько групп, сколько разновидностей имеет качественный признак. Данный подход используется в том случае, если число состояний качественного признака невелико (например, при построении группировок населения по полу, источникам средств существования; группировки предприятий по формам собственности и др.). Если качественный признак имеет большое количество разновидностей, то органами статистики разрабатываются номенклатуры и классификации.

В результате статистического наблюдения исследователь получает данные, в которых содержатся сведения о том, какие значения принимал изучаемый признак. Если значения признака у отдельных элементов изменяются, то говорят, что признак колеблется, варьирует (так называемая вариация признака).

---

[1] *Альтернативным* называется такой качественный признак, варианты которого могут принимать только одно из двух значений (например, работник предприятия имеет высшее образование или не имеет, изделие может быть годным или бракованным и т.д.).

Значения, которые принимает изучаемый признак, называются *вариантами*, а число, показывающее, сколько раз наблюдается в исследуемой совокупности тот или иной вариант, называется *частотой*. Вариант обычно обозначается через x_i, а соответствующая частота - m_i, причем $\Sigma m_i = n$, где n - общее число значений вариантов или объем совокупности.

Отношение частоты к объему совокупности называется *относительной частотой (частостью)* появления значения данного варианта:

$$p_i = \frac{m_i}{n}, \qquad \Sigma p_i = 1.$$

Если расположить все варианты признака в порядке их возрастания или убывания $x_1, x_2,...,x_k$ и указать для каждого варианта x_i его частоту m_i: $m_1,m_2,...,m_k$, то получим вариационный ряд, или ***ряд распределения***. Таким образом, ***вариационным рядом*** называется совокупность вариантов ранжированного признака и соответствующие им частоты [1,3-60].

Различают *дискретные* (в случае, если признак имеет дискретный характер) *и интервальные* (когда признак непрерывен) вариационные ряды. Обычно дискретные ряды применяются в случае, когда число вариантов невелико. Если же число вариантов значительно или признак имеет непрерывный характер, то рекомендуется строить интервальный вариационный ряд, основанный на объединении отдельных вариант в группы принадлежащие соответствующим интервалам. Возможен переход от интервального ряда к дискретному, в этом случае в качестве вариантов используются середины данных интервалов. В зависимости от характера распределения статистической совокупности могут устанавливаться интервалы равной и неравной длины. В статистике *неравные* интервалы применяются чаще равных интервалов и делятся на *возрастающие* (например, группировка городов по численности населения) и *убывающие* (например, группировка рабочих по степени выполнения норм выработки).

Построение интервального вариационного ряда с постоянными интервалами равной длины может быть осуществлено следующим способом.

1. Находится наименьшее x_{min} и наибольшее x_{max} значения признака в исследуемой совокупности.

2. Определяется размах вариации как разность между максимальным и минимальным значениями признака, т.е. $R = x_{max} - x_{min}$.

3. Длина постоянного интервала может быть вычислена *двумя методами*: а) заранее установлено число выделяемых групп; d) число выделяемых групп неизвестно:

а) как правило, при построении количественных группировок выделяют 4-7 групп с равным (постоянным) интервалом, величина (h) которого определяется по формуле:

$$h = \frac{x_{max} - x_{min}}{n},$$

где x_{min} и x_{max} - соответственно минимальное и максимальное значения группировочного признака;

n - число выделяемых групп;

b) если число выделяемых групп заранее не установлено и распределение единиц совокупности приближается к нормальному, то величину постоянного интервала можно определить по формуле американского ученого Г. Стерджесса:

$$h = \frac{x_{max} - x_{min}}{1 + 3{,}322 \lg N},$$

где N - число единиц совокупности.

4. Устанавливаются границы интервалов вида $(a_k, b_k]$. При нахождении нижней границы первого интервала рекомендуется поступать следующим образом: от наименьшего значения признака отступают на половину величины интервала h, т.е. $a_1 = x_{min} - h/2$, тогда верхняя граница первого интервала $b_1 = a_1 + h$, далее полагают $a_2 = b_1$, $b_2 = a_2 + h$ и т.д. Построение интервалов осуществляется до тех пор, пока в соответствующий интервал не попадет максимальное значение признака.

5. После того, как установлена полная шкала интервалов, определяют число вариантов, попавших в каждый интервал, т.е. находят частоты m_k. При этом следует помнить, что в случае совпадения значения варианта с границей интервала его следует относить к соответствующему интервалу.

Различают два способа построения интервалов группировки: закрытые и открытые интервалы.

Закрытыми называются интервалы, в которых верхняя граница предшествующего интервала по величине совпадает с нижней границей последующего интервала (см. табл. 1.1).

Таблица 1.1

Распределение рабочих предприятия по заработной плате

Группы рабочих по заработной плате[2], руб	Численность рабочих, чел
3300-3900	20
3900-4500	30
4500-5100	50
5100-5700	40
5700-6300	10
Итого	150

В случае открытых интервалах их границы не повторяются (см. табл. 1.2).

[2] В какую группу (первую или вторую) отнести рабочего, заработная плата которого составляет 3 900 руб? Обычно, при закрытых интервалах руководствуются следующим правилом: нижняя граница показывает включительно, а верхняя - исключительно.

Таблица 1.2

Распределение рабочих предприятия по заработной плате

Группы рабочих по заработной плате-, руб	Численность рабочих, чел
3301-3900	20
3901-4500	30
4501-5100	50
5101-5700	40
5701-6300	10
Итого	150

Вариационные ряды представляются в табличной форме. *Статистическая таблица* – форма наиболее рационального и систематизированного изложения результатов сводки и группировки. В таблице наглядно проявляется связь между признаками изучаемой совокупности. Она служат основой для сравнения и статистического анализа. Не всякая таблица является статистической. Так, например, следуют различать математические таблицы (таблицы умножения, квадратов, логарифмов и т.д.) и другие виды нестатистических таблиц.

Статистическая таблица представляет собой комбинацию горизонтальных строк и вертикальных граф (колонок, столбцов). *Макет таблицы* – незаполненная цифрами таблица, которая имеет общий, боковые и верхние заголовки, т.е. в нем приведены наименования граф и строк.

Элементы статистической таблицы. Элементами таблицы являются статистическое подлежащее и сказуемое.

Подлежащее статистической таблицы – объект, составные части или единицы объекта изучения (например, предприятия, территории, виды продукции и т.п.). Оно показывает, о чем идет речь в таблице. Обычно подлежащее таблицы дается в левой ее части, в наименовании строк.

Сказуемое статистической таблицы – совокупность показателей, которыми характеризуется подлежащее. Оно записывается в верхних заголовках и составляет содержание граф.

Виды статистических таблиц. В зависимости от строения подлежащего статистические таблицы подразделяются на три вида: простые, групповые и комбинационные.

Простыми называется таблицы, подлежащее которых не содержит группировки. Они могут быть *перечневыми, территориальными и хронологическими (динамическими).* Это зависит от того, что отражается в подлежащем: перечень единиц совокупности, территорий, моментов (периодов) времени.

[3] С точки зрения статистического анализа, удобнее использовать группировки с закрытыми интервалами.

Групповыми называются таблицы, подлежащее которых содержит группировку только по одному признаку (например, группировку предприятий по численности работников).

В *комбинационных* таблицах подлежащее разбивается на группы не по одному, а по двум и более признакам, взятым в комбинации (например, группировка населения по полу и возрасту). Различие между групповой и комбинационной таблицами состоит в разном числе признаков, по которым происходит расчленение совокупности.

Графически вариационные ряды могут представляться в виде полигона распределения (многоугольника), гистограммы, кумуляты и огивы распределения [1, 50-53].

ЗАДАЧИ

Задача 1.

При изучении покупательского спроса населения на обувь зафиксирована продажа следующих размеров женской обуви:

35	31	32	35	37	38	39	32	35	36	36
36	37	38	40	33	35	37	38	39	39	39
39	39	40	35	37	37	38	36	37	35	37
37	38	39	36	35	37	38	38	39	40	40
39	40	38	37	38	39	37	35	33	33	35
36	40	40	35	33	34	35	37	38	39	35

1. Для обобщения данных реализованного спроса постройте ряд распределения и проанализируйте полученные результаты.
2. Данные распределения изобразите графически.
3. Результаты разработок изложите в виде таблицы и сделайте выводы.

Задача 2.

Имеются следующие данные о дневном поступлении денежных средств во вклады по 30-ти учреждениям сберегательного банка (млн руб):

205,2	209,6	222,6	236,7	62,0	53,1	172,1	56,5
52,5	172,1	56,5	52,6	46,6	53,2	30,1	146,4
18,1	13,6	89,8	62,5	46,3	103,5	73,3	76,6
73,0	32,3	199,6	59,1	71,2	90,8		

1. Постройте интервальный ряд распределения.
2. Изобразите полученный ряд графически в виде полигона и гистограммы распределения.
3. Полученные результаты представьте в таблице и сделайте выводы.

Задача 3.

Имеются следующие данные о размере посевных площадей 40 фермерских хозяйств (га):

123,5	164,3	276,5	254,0	56,3	64,8	67,9	50,0
205,2	209,6	222,6	236,7	62,0	53,1	172,1	56,5
52,5	172,1	56,5	52,6	46,6	53,2	30,1	146,4
18,1	13,6	89,8	62,5	46,3	103,5	73,3	76,6
73,0	32,3	199,6	59,1	71,2	90,8	125,0	90,0

1. Постройте интервальный ряд распределения.
2. Изобразите полученный ряд графически в виде полигона и гистограммы распределения.
3. Полученные результаты представьте в таблице и сделайте выводы.

Задача 4.

Имеются следующие данные по группе предприятий отрасли промышленности:

Номер предприятия	Средняя годовая стоимость основных производственных фондов, млн руб	Объем валовой продукции, млн руб	Номер предприятия	Средняя годовая стоимость основных производственных фондов, млн руб	Объем валовой продукции, млн руб
1	7,0	12,9	11	3,0	1,4
2	1,0	1,6	12	6,1	9,6
3	3,5	2,5	13	3,9	4,2
4	4,5	5,6	14	3,8	4,4
5	4,9	4,7	15	5,6	8,9
6	2,3	2,8	16	3,3	4,3
7	6,6	11,9	17	4,5	7,9
8	2,0	2,5	18	3,0	1,4
9	4,7	3,5	19	4,1	5,0
10	2,7	2,3	20	3,1	3,2

Для изучения зависимости между стоимостью основных производственных фондов и выпуском продукции произведите группировку предприятий по стоимости основных производственных фондов, образовав пять групп с равными интервалами.

1. По каждой группе предприятий и по всем предприятиям в целом подсчитайте:

а) число предприятий;

b) стоимость основных производственных фондов - всего и в среднем на одно предприятие;
c) стоимость продукции - всего и в среднем на одно предприятие.
d) фондоотдачу (выпуск валовой продукции на один рубль основных производственных фондов).

2. Результаты расчетов представьте в виде групповой таблицы.

3. Проведите анализ данных таблицы по выявлению зависимости степени использования основных производственных фондов (фондоотдачи) от размеров предприятий по стоимости основных фондов.

4. Сделайте выводы.

Задача 5.

Получены следующие данные о работе магазинов ассоциации в 2008 г.:

Магазины	Товарооборот, млн руб			Магазины	Товарооборот, млн руб	
	по договору	фактически			по договору	фактически
1	68,7	99,0		11	123,5	100,8
2	45,7	33,4		12	87,5	98,5
3	65,8	98,5		13	130,0	129,0
4	125,7	143,1		14	50,6	60,0
5	88,5	88,6		15	90,0	99,0
6	190,5	191,5		16	60,5	60,0
7	200,0	198,0		17	190,5	200,0
8	130,0	139,0		18	78,5	80,0
9	80,0	78,0		19	120,0	120,0
10	98,0	100,1		20	100,0	100,5

1. На основе приведенных данных произведите группировку магазинов по уровню выполнения договорных условий, выделив следующие группы магазинов (%):
a) до 100,0;
b) 100,0 – 110,0;
c) 110,0 и выше.

2. По каждой группе и в целом по всем магазинам подсчитайте:
a) число магазинов;
b) товарооборот по договору и фактически;
a) средний процент выполнения договорных условий.

3. Полученные результаты представьте в таблице. Сделайте выводы.

Задача 6.

Разработан следующий макет таблицы:

Распределение рабочих фирмы по полу и степени выполнения норм выработки

Группы рабочих по степени выполнения норм выработки	Мужчины	Женщины	*Итого*	Численность рабочих	Выполнение норм выработки, %
до 90 90-100 110-120 120-140 140 и выше					
Всего					

Правильно ли разработан макет таблицы? КАКОВЫ его недостатки? РАЗРАБОТАЙТЕ правильный макет таблицы и СФОРМУЛИРУЙТЕ его заголовок.

Задача 7.

Разработан следующий макет таблицы:

Распределение населения области по полу и типу поселения

Группы населения по типу поселения	Группы населения по полу	Численность населения	
		тыс. чел	процент к итогу
Городское	мужчины женщины		
Итого Сельское	мужчины женщины		
Итого Всего населения	мужчины женщины		
Всего			

УДАЧЕН ли макет таблицы? Если нет, в чем его недостатки? ПЕРЕРАБОТАЙТЕ макет этой таблицы с учетом выявленных недостатков. ПОСТРОЙТЕ оптимальный по компактности макет.

Задача 8.

ПРОВЕРЬТЕ данные приведенных ниже таблиц и УСТАНОВИТЕ: 1) содержится ли в них ошибка; 2) можно ли указать, какое число в каждой из этих таблиц содержит ошибку (итог или одно из слагаемых) и исправить его.

1. Объем розничного товарооборота фирмы в фактических ценах за 1 квартал 2003 г. (тыс. руб):

Месяцы	Январь	Февраль	Март	Всего
Тыс. руб	3186	3098	3350	9934

2. Валовой сбор, посевная площадь и урожайность зерновых культур по сельскохозяйственному предприятию:

Валовой сбор, ц	58900
Посевная площадь, га	3100
Урожайность, ц/га	21,8

Задача 9.

Известны следующие данные о распределении объема продукции определенного вида по сортам:

Сорт	1	2	3	Итого
Объем продукции, процент к итогу	63,4	31,9	5,5	100,8

СОДЕРЖАТ ли данные этой таблицы ошибку? Если да, то в чем она выражается? КАКИМ СПОСОБОМ контроля – арифметическим или логическим – можно установить эту ошибку?

Задача 10.

Ниже приведена таблица, характеризующая распределение работников трех цехов предприятия на 1 января 2004 г.:

Цехи	Мужчины	Женщины	Всего
1	160	130	290
2	220	138	398
3	180	120	300
Итого	560	328	988

НАЙДИТЕ и ИСПРАВЬТЕ наиболее вероятные ошибки, допущенные в этой таблице.

Задача 11.

Получена информация по предприятиям химической промышленности:

1. На основе приведенных данных постройте групповую таблицу по признаку относительного уровня рентабельности, образовав при этом три группы с равными интервалами.
2. Дайте характеристику каждой группы в целом по числу предприятий, уровню рентабельности и фондовооруженности.
3. Полученные результаты оформите в виде таблицы.

Сделайте выводы.

Предприятия	Уровень рентабельности, %	Фондовооруженность, млн руб		Предприятия	Уровень рентабельности, %	Фондовооруженность, млн руб
1	68,7	9,0		11	12,5	30,8
2	45,7	3,4		12	87,5	8,5
3	65,8	8,5		13	13,0	29,0
4	25,7	43,1		14	50,6	4,0
5	88,5	8,6		15	90,0	9,0
6	19,5	1,5		16	60,5	7,0
7	20,0	8,0		17	19,5	9,0
8	13,0	39,0		18	78,5	10,0
9	80,0	8,0		19	12,0	29,0
10	98,0	50,1		20	10,0	9,5

Задача 12.

Численность населения России с денежными доходами ниже величины прожиточного минимума выросло с 19,8 (2016 г.) до 22 млн чел (2017 г.). За этот же период удельный вес (в процентах от общей численности населения) населения России с денежными доходами ниже величины прожиточного минимума выросло с 13,5 до 15,0%.

ПРЕДСТАВЬТЕ эти данные в виде статистической таблицы, СФОРМУЛИРУЙТЕ ее заголовок, УКАЖИТЕ ее подлежащее и сказуемое. К КАКОМУ виду таблиц она относится?

2. СРЕДНИЕ ВЕЛИЧИНЫ, ПОКАЗАТЕЛИ ВАРИАЦИИ

2.1. СРЕДНИЕ ВЕЛИЧИНЫ

Средняя величина – обобщающий показатель, выражающий типичный уровень (размер) варьирующего признака в расчете на единицу однородной совокупности. Она отражает то общее, что складывается в каждой единице совокупности, улавливает общие черты, общие закономерности, проявляющиеся в статистической совокупности в конкретных условиях места и времени. Метод средних – один из главных в статистике.

Виды и формы средней величины. Различают следующие *основные виды* средней величины: 1) средняя арифметическая; 2) средняя гармоническая; 3) средняя квадратическая; 4) средняя геометрическая; 5) описательные средние (медиана и мода).

Каждый из этих видов имеет *формы*: простая и взвешенная средняя. Простые средние используются для *несгруппированных* данных, взвешенные – для *сгруппированных* данных. Все виды и формы средних обозначаются в статистике – $\bar{x}$.

Средняя арифметическая и методы ее вычисления

Средняя арифметическая - наиболее распространенный вид средних величин, используемый при изучении социально-экономических явлений. Когда речь идет о средней величине без указания ее вида, подразумевается именно средняя арифметическая.

Средняя арифметическая простая. Она используется в тех случаях, когда расчет осуществляется по несгруппированным данным. Средняя арифметическая простая равна частному от деления суммы индивидуальных значений признака на их количество. Например, имеются данные о стаже работы четверых рабочих (лет): 6, 9, 11, 14. Для получения среднего стажа работы необходимо общий стаж работы всех рабочих разделить на общую численность рабочих, т.е.

$$\bar{x}=\frac{6+9+11+14}{4}=10 \text{ } \textit{лет.}$$

Итак, формула средней арифметической простой выразится в следующем виде[4]:

$$\bar{x}=\frac{x_1+x_2+...+x_n}{n}=\frac{\sum_{i=1}^{n}x_i}{n},$$

или в более простой записи:

[4] В дальнейшем изложении для простоты написания формул не будем пользоваться надписными значками ниже и выше знака суммы (подразумевая во всех случаях, что речь идет о сумме значений от 1-го члена до последнего) и подписными значками при *x, m, f* и *w*.

$$\bar{x}=\frac{\sum x}{n},$$

где x – значения варьирующего признака;

n – число этих значений (число единиц совокупности);

$\sum$ – знак суммирования.

Средняя арифметическая простая применяется в тех случаях, когда каждое индивидуальное значение признака встречается один или одинаковое число раз. База для вычисления этой средней – первичные записи результатов наблюдений.

Средняя арифметическая взвешенная. На практике чаще используются сгруппированные данные: отдельные значения исследуемой совокупности встречаются не один, а много, причем неодинаковое число раз, т.е. представляют собой ряд распределения.

Пример 2.1. По данным табл. 2.1 вычислим средний стаж работы рабочих предприятия.

Таблица 2.1

Распределение рабочих предприятия по стажу работы

Стаж работы, лет *(x)*	Численность рабочих, чел *(m)*
6	90
9	70
11	30
14	10
Итого	200

В этом случае средний стаж работы ($\bar{x}$) может быть получен путем деления общей суммы стажа работы всех рабочих (общего объема варьирующего признака) на общую численность рабочих (общий объем совокупности):

$$\bar{x}=\frac{6\cdot 90+9\cdot 70+11\cdot 30+14\cdot 10}{200}=8{,}2\ \textit{лет}.$$

Формула средней арифметической взвешенной имеет вид:

$$\bar{x}=\frac{x_1m_1+x_2m_2+...+x_nm_n}{m_1+m_2+...+m_n},\ \text{или}\ \bar{x}=\frac{\sum xm}{\sum m},$$

где m – частоты, веса (абсолютные показатели).

Средняя арифметическая взвешенная зависит не только от величины вариантов, но и от соотношения весов. Чем больше веса имеют малые значения вариантов, тем меньше величина средней и наоборот.

Средняя арифметическая простая рассматривается как частный случай средней арифметической взвешенной, в которой все веса либо равны между собой, либо равны единице (m=1).

Методы вычисления средней арифметической

На практике используются два метода расчета средней арифметической: 1) определение общей средней на основе групповых (частных) средних; 2) вычисление средней арифметической по данным интервального вариационного ряда.

1. Определение общей средней на основе групповых (частных) средних.

Пример 2.2. Имеем следующие данные о заработной плате работников двух предприятий (табл. 2.2).

Таблица 2.2

Заработная плата работников предприятий за январь 2008 г.

Предприятия	Средняя заработная, руб (*x*)	Численность работников, в % к итогу (*f*)
1	3500	30,0
2	3900	70,0
Итого	х	100,0

Как исчислить среднюю заработную плату ($\bar{x}$), приходящуюся на одного работника, в целом по двум предприятиям? Подсчитаем среднюю заработную плату по формуле средней арифметической простой (4.1):

$$\bar{x} = \frac{\sum x}{n} = \frac{3500+3900}{2} = 3700 \text{ руб.}$$

Очевидно, что такой расчет средней заработной платы был бы верен лишь в случае, если бы на предприятиях было по одному работнику, что абсурдно, или по одинаковому их числу, чего тоже практически не бывает. Правильным будет такое решение, которое учтет различия в численности работников предприятий. Это значит, что групповую (частную) заработную плату одного работника надо взвесить по их численности, т.е. используется формула:

$$\bar{x} = \frac{\sum xm}{\sum m} = \frac{3500 \cdot 30{,}0 + 3900 \cdot 70{,}0}{100{,}0} = 3780 \text{ руб.}$$

Применение весов дало возможность учесть все значения осредняемого признака.

2. Вычисление средней арифметической по данным интервального вариационного ряда. Для вычисления средней необходимо от интервалов перейти к их серединам**.**

Пример 2.3. В табл. 2.3 приведено распределение рабочих предприятия по стажу работы.

Таблица 2.3

Распределение рабочих предприятия по стажу работы

Группы рабочих по стажу работы, лет	Численность рабочих, чел (m)	Серединное значение интервала (x)
до 2	228	1,0
2-5	100	3,5
5-10	50	7,5
10-20	22	15,0
Итого	400	х

Определим средний стаж работы по формуле средней арифметической взвешенной:

$$\overline{x} = \frac{\sum \mathrm{xm}}{\sum m} = \frac{1{,}0 \cdot 228 + 3{,}5 \cdot 100 + 7{,}5 \cdot 50 + 15{,}0 \cdot 22}{400} = 3{,}2 \text{ года.}$$

Принимая серединные значения интервалов в качестве *x*, условно предполагаем, что варианты в каждом интервале располагаются равномерно, хотя в действительности этого может и не оказаться. Однако практика показывает, что такие допущения близки к действительности, особенно если численность явлений достаточно велика.

Средняя гармоническая

Средняя гармоническая представляет собой обратную величину средней арифметической из обратных значений варьирующего признака.

Пример 2.4. Трое рабочих изготавливают детали одного и того же вида. Затраты рабочего времени на изготовление одной детали составили:

Номер рабочего	Затраты рабочего времени на изготовление одной детали, ч
1	0,5
2	0,33
3	0,2

Рабочие проработали по одному часу. Определим средние затраты рабочего времени на изготовление одной детали ($\overline{x}$) по формуле средней арифметической простой:

$$\overline{x} = \frac{\sum x}{n} = \cancel{\frac{0{,}5+0{,}33+0{,}2}{3}} = 0{,}34 \text{ ч,}$$

т.е. в целом по трем рабочим средние затраты рабочего времени на изготовление одной детали составили 0,34 ч.

Верно ли это? Покажем, что этот расчет является неправильным. Составим *исходное соотношение* (базу, основу) расчета средних затрат рабочего времени на изготовление одной детали:[5]

[5] *Исходное соотношение* – это логическая формула средней. Оно составляется на основе теоретического и экономического анализа.

на изготовление одной детали
Средние затраты
рабочего времени
Общие затраты рабочего времени на производство
всего объема продукции (всех деталей)
= –Количество произведенной продукции (деталей).

Тогда средние затраты рабочего времени на изготовление одной детали составят:

$$\bar{x} = \frac{1+1+1}{\frac{1}{0{,}5}+\frac{1}{0{,}33}+\frac{1}{0{,}2}} = \frac{3}{10} = 0{,}30 \text{ ч},$$

где в знаменателе дроби сумма обратных значений усредняемого признака (затраты рабочего времени на изготовление одной детали), а в числителе - число единиц совокупности.

Итак, применение в данном примере формулы средней арифметической является неправомерным действием, искажающим (точнее, завышающим) результат и не обоснованным логической сущностью этого показателя.

Такой результат получен по формуле средней гармонической простой:

$$\bar{x} = \frac{1+1+...+1}{\frac{1}{x_1}+\frac{1}{x_2}+...+\frac{1}{x_n}}, \quad \text{или} \quad \bar{x} = \frac{n}{\sum \frac{1}{x}},$$

где x, n – то же, что в формулах средней арифметической.

Пример 2.5. Изменим условие примера 2.4. Допустим, что рабочие проработали различное число часов, затраты рабочего времени на изготовление одной детали не изменились:

Номер рабочего	Затраты рабочего времени на изготовление одной детали, ч	Время работы рабочего, ч
1	0,5	3
2	0,33	5
3	0,2	8

Для расчета средних затрат рабочего времени на изготовление одной детали по трем рабочим вместе используем то же самое исходное соотношение данного показателя, т.е. получим:

$$\bar{x} = \frac{3+5+8}{\frac{3}{0{,}5}+\frac{5}{0{,}33}+\frac{8}{0{,}2}} = \frac{16}{61} = 0{,}26 \text{ ч}.$$

Расчет произведен по формуле средней гармонической взвешенной:

$$\bar{x} = \frac{w_1+w_2+...+w_n}{\frac{w_1}{x_1}+\frac{w_2}{x_2}+...+\frac{w_n}{x_n}}, \quad \text{или} \quad \bar{x} = \frac{\sum w}{\sum \frac{w}{x}},$$

где w – веса (частоты) вариационного ряда.

Средняя гармоническая применяется также при расчетах индекса покупательной способности рубля, который обратно пропорционален индексу цен на товары и услуги.

Описательные средние (медиана и мода)

Помимо перечисленных в статистике употребляются еще две особые разновидности средних величин, которые вытекают из характеристики статистических рядов и не являются результатом каких-либо алгебраических действий; условно их можно назвать описательными (структурными) средними. Эти средние – медиана и мода. Особенность описательных средних состоит в том, что они, будучи вполне определенными значениями признака, могут быть использованы, как и другие средние, в качестве обобщающей характеристики всей совокупности в целом.

Медиана. Название *"медиана"* взято из геометрии, где им именуется отрезок, соединяющий одну из вершин треугольника с серединой противоположной стороны и разделяющий, таким образом, сторону треугольника на две равные части. В статистике медиана также делит на две равные части, но не сторону треугольника, а ранжированный (построенный в порядке возрастания или убывания) вариационный ряд. Итак, *медиана* – это варианта, делящая ранжированный вариационный ряд на две равные части. Значения вариант одной из этих частей меньше медианы, а вторая часть содержит варианты, значения которых больше медианы. Медиана не зависит ни от амплитуды колебаний ряда, ни от распределения частот в пределах двух равных частей ряда, поэтому ее применение позволяет получить более точные результаты, чем при использовании других форм средних. Обозначают медиану символом *Me*.

Пример 2.6. Допустим, имеется группа рабочих из 7 человек. Требуется охарактеризовать эту группу с точки зрения стажа работы. Если расположить стаж работы рабочих в порядке возрастания варьирующего признака, т.е. стажа работы, получим следующий ранжированный ряд:

Порядковые номера рабочих по стажу работы	1	2	3	4	5	6	7
Стаж работы, лет	7	8	8	9	10	11	14

Для характеристики среднего стажа работы можно найти среднюю арифметическую:

$$\bar{x} = \frac{7+8+8+9+10+11+14}{7} = 9{,}6 \text{ лет.}$$

Но для этих же целей можно определить медиану. Это значит найти такого рабочего, стаж работы которого разделяет всю совокупность на две равные части: трое имеют стаж работы меньше его, а трое – больше. Такой рабочий имеет порядковый номер № 4 и стаж работы его 9 лет. Следовательно, медианой для данной совокупности по признаку стажа работы будет 9 лет (*Me* = 9 лет).

Если всем единицам ранжированного вариационного ряда придать порядковые номера, то номер медианы в ряду с нечетном числом членов "n" определится как $\frac{n+1}{2}$. Так, в нашем примере из 7 членов ряда номер медианы $\frac{7+1}{2}=4,$ т.е. медианой является варианта, стоящая в ряду 4-ой по порядку.

Если же вариант – четное число, то медиану приходится определять как среднюю из двух центральных вариант, порядковые номера которых $\frac{n}{2}$ и $\frac{n}{2}+1.$ Изменим условие предыдущего примера: группа рабочих состоит из 8 человек, рабочий с порядковым номером № 8 в ранжированном ряду имеет стаж работы 16 лет:

Порядковые номера рабочих по стажу работы	1	2	3	4	5	6	7	8
Стаж работы, лет	7	8	8	9	10	11	14	16

В данном случае $\mathrm{Me} = \frac{9+10}{2} = 9{,}5$ лет. Однако, если число единиц в совокупности достаточно большое и различия между величинами рядом стоящих вариант незначительные, то можно считать медианой одну из центральных вариант с порядковым номером $\frac{n}{2}$. Так обычно поступают, определяя медиану при четном числе членов ряда.

Мода. *Модой* в статистике именуется понятие, имеющее некоторое сходство с тем, что понимается под модой в обычном смысле слова. Если в быту называют модной одеждой тот фасон, который в данный момент приобретает широкое распространение, то в теории статистики *модой* называют такое *значение варьирующего признака,* которое *встречается чаще,* чем какие-либо другие. Итак, *мода* – это варианта, обладающая наибольшей частотой. Мода отражает некоторый наиболее типичный уровень признака, характерный для данной совокупности. Обозначают моду символом *Mo.*

Отыскание медианы и моды для *дискретного и интервального вариационных рядов* производится различно.

Определение медианы и моды по данным *дискретного ряда* не представляет трудностей.

Пример 2.7. В табл. 2.4 представлено распределение рабочих предприятия по тарифному разряду.

Таблица 2.4

Распределение рабочих предприятия по тарифному разряду

Тарифный разряд рабочих	Численность рабочих, чел	Накопленные частоты
1	15	15
2	35	50
3 ***(Mo)***	**45**	95
4 ***(Me)***	40	135
5	35	170
6	30	200
Итого	200	х

Чтобы определить медиану тарифного разряда рабочих, необходимо найти ее порядковый номер, а затем по накопленным частотам (частостям) определить величину варианты, обладающей таким номером. В нашем примере, где число членов ряда 200 (четное число), принимаем за медиану одну из центральных вариант, номер которой $\frac{200}{2} = 100.$ Затем по накопленным частотам определяем, что 100-й член ряда имеет 4 тарифный разряд, т.е. *Me* = 4-му тарифному разряду рабочих.

В данном примере наибольшая численность рабочих имеет третий тарифный разряд, т.е. *мода равна третьему тарифному разряду*: *Mo* = 3.

Пример 2.8. Покажем, как найти медиану и моду для *интервального ряда* на примере распределения работников предприятия по возрасту (табл. 2.5):

Таблица 2.5

Распределение работников предприятия по возрасту

Группы работников по возрасту, лет	Численность работников, чел	Накопленные частоты
До 20	45	45
20-25	120	165
25-30	225	**390**
30-35 ***(Me)***	260	650
35-40 ***(Mo)***	**280**	930
40-45	105	1035
45-50	85	1120
50-55	70	1190
55-60	50	1240
Итого	1240	х

Прежде всего определяют интервалы, в которых находятся медиана и мода. Их называют соответственно медианным и модальным интервалами.

Медианный интервал определяется по накопленным частотам. Внутри интервала медиана вычисляется по формуле[6]:

$$Me = x_0 + h\frac{\frac{\sum m}{2} - S_{Me-1}}{m_{Me}},$$

где x_0 – нижняя граница медианного интервала;

h – величина медианного интервала;

$\sum n$ – сумма частот или число членов ряда, тогда $\frac{\sum m}{2}$ – номер медианы;

S_{Me-1} – сумма накопленных частот интервалов, предшествующих медианному;

m_{Me} – частота медианного интервала.

Если в интервальном вариационном ряде в качестве весов используются частости (относительные частоты), то как и определение медианного интервала, так и вычисление медианы осуществляются аналогично. Однако в формуле в качестве весов вместо частот (m) используются частости (f):

$$Me = x_0 + h\frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}},$$

где $\sum f$ – сумма частостей, тогда $\frac{\sum f}{2}$ – номер медианы;

S_{Me-1} – сумма накопленных частостей интервалов, предшествующих медианному;

f_{Me} – частость медианного интервала.

Исчислим медиану по данным табл. 2.5. Медианный интервал находится в пределах 30-35 лет, т.к. в пределах этого интервала находится варианта, которая делит совокупность на две равные части $\frac{1240}{2} = 620.$ Подставляя в формулу медианы необходимые численные значения, получим:

$$\mathrm{Me} = 30 + 5\frac{\frac{1240}{2} - 390}{260} = 34{,}4 \text{ года.}$$

Это значит, одна половина работников имеет возраст до 34,4 лет, а другая – свыше 34,4.

Модальный интервал - это интервал с наибольшей частотой, т.е. интервал 35-40 лет в нашем примере. Внутри интервала мода определяется по формуле[7]:

---

6 Формула медианы *(M_e)* выведена исходя из допущения о равномерности нарастания накопленной частоты внутри интервала.

7 Формула моды (*Mo)* выведена математически, исходя из допущения, что в модальном и двух соседних интервалах кривая распределения представляет собой параболу второго порядка. Тогда *Mo* находится как абсцисса вершины параболы.

$$Mo = x_0 + h\frac{d_1}{d_1 + d_2},$$

где x_0 – нижняя граница модального интервала;

h – величина модального интервала;

d_1 – разность между частотой модального интервала (m_{Mo}) и частотой интервала,

предшествующему модальному (m_{Mo-1}), т.е.: $d_1 = m_{Mo} - m_{Mo-1}$;

d_2 – разность между частотой модального интервала (m_{Mo}) и частотой интервала,

следующего за модальным (m_{Mo+1}), т.е.: $d_2 = m_{Mo} - m_{Mo+1}$.

При использовании в интервальном вариационном ряде в качестве весов частостей (относительных частот) – f мода также определяется по вышеприведенной формуле, однако:

d_1 – разность между частостью модального интервала (f_{Mo}) и частостью

интервала, предшествующему модальному (f_{Mo-1}), т.е.: $d_1 = f_{Mo} - f_{Mo-1}$;

d_2 – разность между частостью модального интервала (f_{Mo}) и частостью

интервала, следующего за модальным (f_{Mo+1}), т.е.: $d_2 = f_{Mo} - f_{Mo+1}$.

Рассмотрим исчисление моды из интервального вариационного ряда на примере распределения работников по возрасту (табл. 4.13). В данном примере модальный интервал находится в пределах 35-40 лет, т.к. на этот интервал приходится наибольшая частота (280). Моду определим по формуле:

$$Mo = 35 + 5\frac{280 - 260}{(280 - 260) + (280 - 105)} = 35{,}5 \text{ лет}$$

Это значит, что модальный возраст работников равен 35,5 лет. Таким образом, мода является наиболее распространенной и в этом смысле наиболее типичной величиной в распределении. Но мода и средняя величина по-разному характеризуют совокупность. Мода определяет непосредственно размер признака, свойственный хотя и значительной части, но все же не всей совокупности. В нашем примере, если даже предположить, что возраст всех работников пятой группы (35-40 лет) составляет 35,5 лет, то и в этом случае мода соответствует только 22,6% общей численности всех работников. Поэтому мода по своему обобщающему значению уступает средней, которая характеризует совокупность в целом, т.к. складывается под воздействием всех без исключения элементов совокупности.

2.2. ПОКАЗАТЕЛИ ВАРИАЦИИ

Понятие о вариации признаков

При изучении явлений и процессов статистика встречается с разнообразной вариацией признаков, характеризующих отдельные единицы совокупностей. Величины признаков варьируют под воздействием различных причин и условий. Чем разнообразнее условия, влияющие на размер данного признака, тем больше его вариация. *Вариацией признака называется колеблемость его значений при переходе от одной единицы совокупности к другой.*

Пример 2.9. Имеются следующие данные о стаже работы рабочих двух бригад (табл. 2.6).

Таблица 2.6

Стаж работы рабочих двух бригад

Бригады	Численность рабочих (чел)	Стаж работы отдельных рабочих, лет			
1	4	8	9	11	12
2	4	4	8	12	16

Средняя арифметическая в обоих случаях будет одинаковой, хотя в первой бригаде средняя значительно меньше отличается от индивидуальных значений признака, чем во второй:

$$\bar{x}_2 = \frac{\sum x}{n} = \frac{4+8+12+16}{4} = 10 \quad \textit{лет.}$$

Следовательно, с помощью средней величины нельзя охарактеризовать колеблемость признака. Для этого необходимы специальные показатели, которые называются показателями вариации. Они могут быть: 1) абсолютными; 2) относительными.

К *абсолютным показателям вариации* относятся: вариационный размах, среднее линейное (абсолютное) отклонение, дисперсия, среднее квадратическое отклонение. Эти показатели характеризуют абсолютную колеблемость (вариацию) признаков и выражаются в тех же единицах измерения, что и изучаемый признак (у дисперсии единицы измерения не записываются).

К *относительным показателям вариации* относятся коэффициенты осцилляции, колеблемости и вариации. Эти показатели характеризуют относительную колеблемость (вариацию) признаков и выражаются чаще в процентах.

Значение показателей вариации заключается в следующем: 1) показатели вариации дополняют средние величины, за которыми скрываются индивидуальные различия; 2) показатели вариации характеризуют степень однородности статистической совокупности по данному признаку; 3) показатели вариации характеризуют границы вариации признака.

Показатели вариации и способы их расчета

Абсолютные показатели вариации

Вариационный размах *(R)* или размах вариации - это разность между максимальным и минимальным значениями варьирующего признака в изучаемой совокупности:

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

Вариационный размах характеризует пределы изменения варьирующего признака. Его величина зависит только от двух крайних значений признака, в то время как колеблемость в целом складывается из всех его значений. В приведенном выше примере (табл. 2.6) размах вариации равен в первой бригаде: 12-8=4 года, во второй: 16-4=12 лет.

Среднее линейное (абсолютное) отклонение $\bar{d}$ представляет собой среднюю арифметическую из абсолютных отклонений конкретных вариант от их средней арифметической. При исчислении среднего линейного отклонения принимаются во внимание только абсолютные значения отклонений (без учета знаков: "+" или "-"), т.к. алгебраическая сумма отклонений индивидуальных значений признака от средней равна нулю, т.е. $\sum (x - \bar{x}) m = 0$. Различают простое и взвешенное среднее линейное отклонение. Простое отклонение используют для несгруппированных данных, а взвешенное – для сгруппированных:

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n}$$ – простое отклонение;

$$\bar{d} = \frac{\sum |x - \bar{x}| m}{\sum m}$$ – взвешенное отклонение.

Символы x, $\bar{x}$, n, m имеют те же значения, что и в предыдущей главе. Отклонения берутся по модулю, на что указывают прямые скобки в числителе формул.

Рассмотрим вычисление среднего линейного отклонения по данным табл. 2.6, результаты расчетов представим в табл. 2.7.

Таблица 2.7

К расчету показателей вариации стажа работы рабочих двух бригад

Номер рабочего	Стаж работы, лет		Отклонение индивидуальных значений признака от средней $(x-\bar{x})$		Абсолютные индивидуальные отклонения от средней $\lvert x-\bar{x}\rvert$		Квадраты отклонений индивидуальных значений от средней $(x-\bar{x})^2$	
	бригада № 1	бригада № 2	бригада № 1	бригада № 2	бригада № 1	бригада № 2	бригада № 1	бригада № 2
1	8	4	-2	-6	2	6	4	36
2	9	8	-1	-2	1	2	1	4
3	11	12	+1	+2	1	2	1	4
4	12	16	+2	+6	2	6	4	36
Итого	40	40	–	–	6	16	10	80

Среднее линейное отклонение:

$$\bar{d}=\frac{\sum\lvert x-\bar{x}\rvert}{n}; \quad \bar{d}_1=\frac{6}{4}=1{,}5 \text{ года}; \ \bar{d}_2=\frac{16}{4}=4{,}0 \text{ года}.$$

Как видно из расчетов, среднее линейное отклонение стажа рабочих во второй бригаде почти в 2,7 раза больше, чем в первой. Среднее линейное отклонение используется в настоящее время редко из-за трудностей в применении математических методов анализа вариации, которые возникают в результате абстрагирования от знака отклонения. Поэтому в статистике для характеристики колеблемости признака чаще всего пользуются дисперсией и средним квадратическим отклонением.

Дисперсия – средний квадрат отклонений индивидуальных значений признака от средней арифметической. Дисперсия обозначается греческой буквой σ (сигма) в квадрате и вычисляется по формулам простой и взвешенной дисперсий (в зависимости от исходных данных):

$$\sigma^2=\frac{\sum(x-\bar{x})^2}{n}$$ – простая дисперсия;

$$\sigma^2=\frac{\sum(x-\bar{x})^2 m}{\sum m}$$ – взвешенная дисперсия.

Дисперсия является фундаментальным показателем вариации. Однако ее применение в ряде случаев бывает не совсем удобным, т.к. размерность дисперсии равна квадрату размерности изучаемого признака. Поэтому в таких случаях для измерения вариации признака вычисляют среднее квадратическое отклонение.

Среднее квадратическое отклонение равно корню квадратному из суммы квадратов отклонений индивидуальных значений признака от их средней, т.е. из дисперсии:

$$\sigma=\sqrt{\frac{\sum(x-\bar{x})^2}{n}}$$ – простое отклонение;

$$\sigma=\sqrt{\frac{\sum(x-\bar{x})^2 m}{\sum m}}$$ – взвешенное отклонение.

Среднее квадратическое отклонение показывает, на сколько единиц в среднем отклоняются варианты от их средней арифметической.

Рассмотрим вычисление дисперсии и среднего квадратического отклонения по данным о стаже рабочих в двух бригадах, для чего воспользуемся расчетами, приведенными в табл. 2.7.

Дисперсия: $\sigma^2=\frac{\sum(x-\bar{x})^2}{n}$; $\sigma_1^2=\frac{10}{4}=2{,}5;$ $\sigma_2^2=\frac{80}{4}=20{,}0.$

Среднее квадратическое отклонение:

$$\sigma=\sqrt{\frac{\sum(x-\bar{x})^2}{n}};\quad \sigma_1=\sqrt{\frac{10}{4}}=\sqrt{2{,}5}=1{,}58 \text{ года};\ \sigma_2=\sqrt{\frac{80}{4}}=\sqrt{20{,}0}=4{,}47 \text{ года}.$$

Исчисленные показатели наглядно показывают различную колеблемость стажа отдельных рабочих в различных бригадах. В этой связи дисперсия и среднее квадратическое отклонение являются наиболее распространенными на практике при изучении колеблемости.

Показатели вариации для альтернативного признака

Выше были рассмотрены показатели вариации для количественных признаков. Однако может наблюдаться и вариация качественных признаков. При наличии двух взаимно исключающих друг друга вариантов говорят о наличии альтернативной изменчивости качественных признаков. *Альтернативным* называется такой качественный признак, варианты которого могут принимать только одно из двух значений (например, изделие может быть годным или бракованным, преподаватель вуза имеет ученую степень или не имеет, студент сдал экзамен или не сдал и т.д.). Вариация альтернативного признака выражается в том, что одни единицы совокупности обладают данным признаком, другие - не обладают.

Альтернативный признак принимает два значения:

1 - наличие альтернативного признака у единицы совокупности;

0 - отсутствие признака.

Обозначим:

N - общее число единиц совокупности;

p и M - соответственно доля и число единиц совокупности, обладающих альтернативным признаком: $p=\frac{M}{N}$;

q и (N-M) - соответственно доля и число единиц совокупности, не обладающих альтернативным признаком:

$$q=\frac{N-M}{N}=1-\frac{M}{N}=1-p, \quad где \quad p+q=1$$

Построим ряд распределения по качественному признаку (табл. 2.10).

Таблица 2.10

Вариационный ряд распределения
качественного (альтернативного) признака

Варьирующий признак *(x)*	1	0
Частости ($f=\frac{m}{\sum m}$)	*p*	*q*

Вычислим среднюю арифметическую альтернативного признака:

$$\bar{x}=\frac{\sum xm}{\sum m}=\sum xf=1\cdot p+0\cdot q=p, \quad где \quad \bar{x}=p$$

Тогда дисперсию альтернативного признака определим по следующей формуле:

$$\sigma^2=\frac{\sum (x-\bar{x})^2 m}{\sum m}=\sum (x-\bar{x})^2 f=(1-p)^2 p+(0-p)^2 q=q^2 p+p^2 q=pq(q+p)=pq$$

Итак:

$$\sigma^2=pq$$

В свою очередь, среднее квадратическое отклонение альтернативного признака составит:

$$\sigma=\sqrt{pq}$$

Предельное значение дисперсии альтернативного признака равно 0,25. Оно получается при $p=q=0{,}5$

ЗАДАЧИ

Задача 13.

Имеются данные об установленной мощности 20 сахарных заводов, т:

1550 1600 1700 1300 1400 1200 1570 1423 1600 1800
2550 1660 1080 300 900 1000 1500 420 600 2800

1. Вычислите среднюю мощность сахарного завода:
   a) на основе индивидуальных данных;
   b) на основе построенного ряда распределения.
2. Изобразите полученный ряд графически.
3. Найдите среднее квадратическое отклонение и коэффициент вариации, моду и медиану.

Задача 14.

Имеются следующие данные по цехам предприятия за январь и февраль:

Номер цеха	Январь		Март	
	средняя заработная плата рабочих, тыс. руб	численность рабочих, чел	средняя заработная плата рабочих, тыс. руб	фонд заработной платы рабочих, тыс. руб
1	30,2	105	32,4	4007,8
2	35,5	90	33,7	3920,5
3	37,9	98	41,5	4250,6

1. Вычислите среднюю месячную заработную плату рабочих по заводу в целом: за январь и за март.
2. Дайте обоснование применения формул для исчисления среднего значения.
3. Определите за январь показатели вариации; моду и медиану.
1. Сделайте выводы.

Задача 15.

Имеются данные о посевной площади и урожайности по 4 фермерским хозяйствам области:

Хозяйства области	2017 г.		2018 г.	
	урожайность, ц/га	посевная площадь, га	урожайность, ц/га	валовой сбор, ц
1	14	200	15	4000
2	16	500	20	6 500
3	20	300	17	6000
4	15	100	16	9000

1. Определите среднюю урожайность по хозяйствам для каждого года.
2. Как изменилась средняя урожайность в 2018 г. по сравнению с 2017 г.
3. Дайте обоснование применения формул для исчисления среднего значения. Сделайте выводы.

Задача 16.

По трём предприятиям города имеются следующие данные:

Пред-приятия	I квартал		II квартал	
	изготовлено продукции, тыс. шт.	выполнение задания, %	задание по договорам, тыс. шт.	ожидаемое выполнение, %
1	560	100	700	106
2	750	105	800	103
3	850	98	900	100

Определите по трем промышленным предприятиям города:
1) средний процент выполнения задания в 1 квартале;
2) средний процент ожидаемого выполнения задания во 2 квартале;
3) средний процент ожидаемого выполнения задания в 1 полугодии.

Сделайте выводы.

Задача 17.

Используя условие задачи 2, рассчитайте средний размер поступления средств во вклады, а также показатели вариации, моду и медиану.

Сделайте выводы.

Задача 18.

Используя условие задачи 3, рассчитайте средний размер посевных площадей и показатели вариации, моду и медиану.

Сделайте выводы.

Задача 19.

Используя условие задачи 4, рассчитайте средний размер стоимости основных производственных фондов, среднее квадратическое отклонение и коэффициент вариации, моду и медиану.

Сделайте выводы.

Задача 20.

Используя условие задачи 4, рассчитайте средний размер валовой продукции, среднее квадратическое отклонение и коэффициент вариации, моду и медиану.

Сделайте выводы.

Задача 21.

Используя условие задачи 5, рассчитайте товарооборот по договору и фактически:
а) в целом по всем магазинам ассоциации;
б) в среднем на один магазин.

Сделайте выводы.

Задача 23.

Известны следующие данные о среднем дневном товарообороте предпринимателей, имеющих различный стаж предпринимательской деятельности:

Стаж предпринимательской деятельности, лет	Численность предпринимателей, чел	Средний дневной товарооборот, тыс. руб
до 2	2	10, 12
2 – 5	4	12, 12, 14, 16
5 и более	5	14, 16, 17, 18, 20

ОПРЕДЕЛИТЕ: 1) внутригрупповые дисперсии по среднему дневному товарообороту предпринимателей, имеющих различный стаж предпринимательской деятельности; 2) среднюю из внутригрупповых дисперсий по трем группам предпринимателей; 3) межгрупповую дисперсию; 4) общую дисперсию среднего дневного товарооборота предпринимателей.

Проверьте правильность произведенных расчетов с помощью правила (теоремы) сложения дисперсии.

Задача 24.

Имеются следующие данные о численности и заработной плате работников предприятия торговли за месяц:

Группы товаров	Численность работников, чел	Внутригрупповая дисперсия заработной платы
Продовольственные товары	80	6,5
Непродовольственные товары	30	20,4

ОПРЕДЕЛИТЕ:
1) среднюю из внутригрупповых дисперсий заработной платы работников;
2) межгрупповую дисперсию заработной платы работников предприятия (если известно, что общая дисперсия заработной платы работников этого предприятия равна 33,6);

3) эмпирический коэффициент детерминации и эмпирическое корреляционное отношение.

СДЕЛАЙТЕ выводы.

3. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

3.1. Сущность, ошибки и виды выборочного наблюдения

Выборочным наблюдением (методом) называется такой вид несплошного наблюдения, который основан на теории вероятности и математической статистике.

Основные понятия выборочного наблюдения:

1) *генеральная совокупность* – это совокупность, подлежащая изучению или наблюдению *(N)*;
2) *выборочная совокупность* – это совокупность, подвергшаяся непосредственному наблюдению. Она выбирается из генеральной совокупности в порядке случайного, типического, серийного и механического отборов *(n)*;
3) *генеральная средняя* $(\bar{x})$, т.е. это средняя, рассчитанная для всей изучаемой генеральной совокупности;
4) *выборочная средняя* $(\tilde{x})$ – это средняя, исчисленная для выборочной совокупности;
5) *генеральная доля* (p) – это доля единиц совокупности, обладающих данным признаком во всей генеральной совокупности, например, доля брака во всей произведенной продукции (генеральной совокупности);
6) *выборочная доля* (w) – доля признака в выборочной совокупности.

Средняя (генеральной или выборочной совокупности) и *доля* (генеральной или выборочной совокупности) являются *сводными характеристиками* соответственно генеральной или выборочной совокупности.

В статистике принято выборочную долю *(w)* называть частностью – это отношение единиц, обладающих данным признаком, в общем объеме выборочной совокупности.

Ошибки выборочного наблюдения. При проведении любого статистического наблюдения (как сплошного, так и выборочного) могут происходить ошибки при регистрации единиц. Ошибки регистрации, в свою очередь, делятся на случайные и систематические. Однако только выборочному наблюдению кроме ошибок регистрации присущи *ошибки репрезентативности (представительности)*. Эти ошибки возникают в связи с тем, что состав выборочной совокупности по изучаемому признаку имеет иную структуру, чем состав генеральной совокупности.

Ошибки репрезентативности возникают даже при соблюдении всех правил проведения выборочного наблюдения. Однако при должной организации выборки они могут быть доведены до сколько угодно малых размеров. Ошибки репрезентативности характеризуют размер отклонения между выборочными и генеральными сводными (обобщающими) характеристиками:

а) для среднего значения признака – $|\tilde{x} - \bar{x}|$;

б) для доли признака – $|w - p|$.

Размер ошибок следует рассматривать как в абсолютном, так и в относительном выражении. Например, процент утайки скота при его переписи может оказаться одинаковым при сплошном и выборочном наблюдении, но в первом случае было неучтено 1000 голов, а во втором случае - 100 голов.

Соотношение ошибок сплошного и выборочного наблюдения:

1. При наличии систематических ошибок сплошное наблюдение дает менее точный результат. В этом случае выборочное наблюдение можно использовать в качестве *контроля* данных сплошного наблюдения (например, контрольные обходы при переписи скота).
2. Если систематические ошибки не возникают, то более точный результат дает сплошное наблюдение. В этом случае выборочное наблюдение может быть использовано с целью *экономии затрат труда и средств* при условии, что сможем определить с достаточной точностью ошибку репрезентативности.

Виды выборочного наблюдения. Организация выборочного наблюдения может быть разнообразна как по схемам, так и по методам проведения. Различают две *схемы проведения* выборочного наблюдения:

1) повторная выборка;
2) бесповторная выборка.

1. *Повторной* называется выборка, когда единицы генеральной совокупности, отобранные в выборочную совокупность, после их изучения или обследования *возвращаются* обратно в генеральную совокупность. При таком способе объем генеральной совокупности не изменяется. При этом одна и та же единица генеральной совокупности может повторно попасть под обследование.

Повторная выборка производится по схеме *возвращенного шара* в урну. Например, имеется урна с белыми и красными шарами, общее количество разноцветных шаров неизвестно. Необходимо с помощью повторной выборки определить долю белых и красных шаров. Последовательно вытаскиваем 10 шаров. Первый белый – обозначим через 1 данный цвет и опять опускаем в урну, второй – тоже белый – опять проставим 1 и возвращаем его, третий – красный шар – обозначим его через 0 и т.д. Допустим, для десяти шаров получим следующее распределение:

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

Тогда доля белых шаров в выборочной совокупности (n=10) составит: $w = \frac{6}{10} = 0{,}6$.

Можно привести другой пример: необходимо определить средний вес одной овцы – после взвешивания ее опять выпускают в загон.

2. *Бесповторный* называется выборка, при которой единицы генеральной совокупности, отобранные в состав выборочной совокупности, обратно в генеральную после обследования *не возвращаются*. При этом способе объем генеральной совокупности сокращается и, в конечном счете, он будет меньше на объем выборки: $(N - n)$. Каждая единица генеральной совокупности может быть отобрана в выборочную только один раз.

В теории вероятности бесповторная выборка называется выборкой, проводимой по схеме *невозвращенного шара*, т.е. в нашем примере после обследования цвета шар не возвращается в урну.

И повторная, и бесповторная выборки могут производиться следующими способами:

1. Случайный отбор (случайная выборка).
2. Типический отбор.
3. Серийный (гнездовой) отбор.
4. Только бесповторной выборке присущ способ механического отбора.

1. С*лучайный отбор* единиц в выборочную совокупность характеризуется следующим:

a) отбор единиц производится из всей генеральной совокупности в целом;

b) отбор единиц носит случайный характер и производится либо с использованием способа жеребьевки, либо с использованием таблиц случайных чисел. При использовании жеребьевки все единицы генеральной совокупности нумеруются, и на каждую единицу заводится жребий. Жребии тщательно перемешиваются и наудачу отбираются в выборочную совокупность. Случайный отбор может также осуществляться по таблице случайных чисел. Она составлена с помощью ЭВМ на специальном датчике случайных чисел (см. приложение 1).

2. *Типический отбор* (типическая выборка) характеризуется:

а) вся генеральная совокупность делится на типические группы по какому-либо типическому признаку (например, при нахождении среднего роста студентов их предварительно по полу разбивают на две группы: юноши и девушки);

б) отбор типических единиц в выборочную совокупность производится не из всей генеральной совокупности, а из типических групп.

Например, на потоке II курса университета учится 200 студентов, из них 60 юношей и 140 девушек. Необходимо определить с помощью типической выборки средний рост студентов, причем объем выборочной совокупности равен 30 человек, т.е. 15% ($\frac{30}{200} \cdot 100$), тогда в выборочную совокупность попадут:

- юношей: $15 \cdot \frac{60}{100} = 9$ *чел*;
- девушек: $15 \cdot \frac{140}{100} = 21$ *чел*;

итого: *n* =30 *чел* (9+21);

c) отбор единиц в выборочную совокупность из типических групп производится с помощью случайной выборки.

3. *Серийный (гнездовой)* отбор:

a) вся генеральная совокупность предварительно разбивается на гнезда, серии или части;

b) отбор единиц в выборочную совокупность производится не отдельными единицами, а целыми сериями;

c) отобранные серии (гнезда) обследуются сплошным наблюдением.

Например, необходимо провести анализ брака продукции. На складе находятся 100 ящиков с деталями, из них берем наугад 10 ящиков и производим обследование всех деталей в этих ящиках.

Серийный отбор является наиболее простым в организации, но он наименее точен по результатам (ящик может быть вообще со всеми бракованными деталями).

Самым точным из всех способов является типическая выборка, но она наиболее сложна в организации (например, при изучении средней продолжительности телефонных разговоров их необходимо разделить на служебные и личные, однако это сделать трудно). Случайная выборка занимает промежуточное положение и по точности, и по сложности организации между серийным и типическим отборами.

4. *Механический отбор* (механическая выборка) характеризуется следующим:

a) отбор производится из всей совокупности в целом, т.е. из генеральной совокупности;

b) отбор единиц производится не в случайном порядке, а с использованием механических приемов – либо по алфавиту, либо по списку, либо по определенному интервалу. Например, при изучении брака мы отбираем в выборочную совокупность каждую пятую, десятую и т.д. деталь.

Промежуток, через который попадают единицы в выборку, зависит от принятой пропорции отбора. Пропорция отбора устанавливается делением численности генеральной совокупности на объем выборки. Если эта пропорция оказывается дробной, то берется ближайшее целое число. Например, если надо провести 10%–ную механическую выборку студентов потока (всего 200 чел), то составляется список их фамилий по алфавиту и механически отбирается каждый десятый студент:

1, 11, 21,..., 191,

или можно начать отсчет с любого другого числа, например:

7, 17, 27,..., 197.

Всего в выборку попадет 20 чел.

Механическая выборка – разновидность случайной выборки, однако, она производится бесповторным отбором.

По точности эти четыре метода располагаются следующим образом:

серийная < *механическая* < *случайная* < *типическая.*

На практике часто применяют комбинированные способы отбора, которые включают в себя черты из разных способов (например, способ, сочетающий серийный и механический отбор).

3.2. Определение средней и предельной ошибки выборки

3.2.1. Определение средней ошибки выборки

Средняя ошибка выборки μ показывает: на сколько в среднем возможные варианты выборочных средних отклоняются от их генеральной средней.

Существуют специальные формулы расчета средней ошибки *случайной выборки* (приведем 4 формулы).

1. Средняя ошибка выборки для среднего значения признака при повторном отборе:

$$m_x = \sqrt{\frac{s^2}{n}}, \tag{3.1}$$

где σ^2– дисперсия признака в выборочной совокупности, которая определяется по

данным выборочного наблюдения по формулам, изложенным в главе

«Показатели вариации»;

n – объем выборочной совокупности.

Итак, средняя ошибка для среднего значения признака при повторном отборе зависит от степени колеблемости значений признака в выборочной совокупности и от численности выборки (*n)*.

2. Средняя ошибка для среднего значения признака при бесповоротном отборе:

$$m_x = \sqrt{\frac{s^2}{n}\left(1 - \frac{n}{N}\right)}, \tag{3.2}$$

где $\frac{n}{N}$– доля выборки в генеральной совокупности.

Сравнительная оценка ошибок для повторного и бесповторного отбора:

$$m = m \cdot \sqrt{1 - \frac{n}{N}}. \tag{3.3}$$

При бесповторном отборе мы получаем более точный результат. Отсюда вытекает *следствие*: практически при небольшой доле выборки, т.е.

при малом n, лучше пользоваться формулой ошибки для повторного отбора, даже в том случае, когда выборка бесповторная, т.к. второй сомножитель $\sqrt{1-\frac{n}{N}}$ не имеет практического значения.

Пример 3.1. Вычислите среднюю ошибку выборки при определении средней продолжительности горения электроламп, если из всей продукции 10 000 ламп обследуется 100 ламп, из которых 7 электроламп оказались бракованными. Выборочная дисперсия продолжительности горения электроламп по данным прошлых выборочных наблюдений составляет 14 400.

Решение:

а) вычислим среднюю ошибку выборки для средней продолжительности горения электроламп при повторном отборе:

$$\mu_x = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{14400}{100}} = 12ч.$$

Допустим, выборочная средняя ($\bar{x}$) составит 800 ч, тогда может оказаться, что генеральная средняя будет находиться в пределах:

$\bar{x} = 800ч \pm 12$;

б) вычислим среднюю ошибку выборки для бесповторного отбора:

$$m_x = \sqrt{\frac{s^2}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{14400}{100}\left(1-\frac{100}{10000}\right)} = 11{,}94ч.$$

Из данного примера видно, что при незначительной доле выборочной совокупности в объеме генеральной совокупности ($\frac{n}{N} = \frac{100}{10000} = 0{,}01$, или 1%), средние ошибки при повторном и бесповторном отборах отличаются несущественно.

Определение средней ошибки выборки при измерении доли признака в генеральной совокупности. Дисперсия для альтернативного признака в генеральной совокупности равна:

$$\sigma^2 = \rho \cdot q, \tag{3.4}$$

(см. главу «Показатели вариации»). В этой связи в выборочной совокупности она составит:

$$\sigma^2 = w(1-w), \tag{3.5}$$

где w – доля единиц, обладающих изучаемым признаком, в выборочной совокупности.

Заменим σ^2 в формулах (3.1 и 3.2) на *w(1-w)*. Тогда, средняя ошибка для доли признака при повторном отборе:

$$m_w = \sqrt{\frac{w(1-w)}{n}}. \tag{3.6}$$

Средняя ошибка для доли признака при бесповторном отборе:

$$m_w = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}. \tag{3.7}$$

Пример 3.2. По данным примера 3.1 определим среднюю ошибку выборки доли бракованных электроламп (m_w).

Решение:

1. Доля (удельный вес) бракованных изделий по данным выборочного наблюдения составит: $w = \frac{7}{100} = 0{,}070$, или 7,0%. Соответственно доля (удельный вес) годных электроламп по выборке будет равна: $1 - w = \frac{100-7}{100} = 0{,}930$, или 93,0%.
2. Вычислим среднюю ошибку выборки для доли бракованных изделий при повторном отборе:

$$m_w = \sqrt{\frac{w(1-w)}{n}} = \sqrt{\frac{0{,}07(1-0{,}07)}{100}} = 0{,}0255, \textit{ или } 2{,}55\%.$$

   Тогда может оказаться, что доля (удельный вес) бракованных изделий в генеральной совокупности будет находиться в пределах: $p = 0{,}07 \pm 0{,}0255, \qquad p = 7{,}0 \pm 2{,}55.$
3. Вычислим среднюю ошибку выборки для доли бракованных изделий при бесповторном отборе:

$$m_w = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)} = \sqrt{\frac{0{,}07(1-0{,}07)}{100}\left(1-\frac{100}{10000}\right)} = 0{,}0254, \textit{ или } 2{,}54\%.$$

ЗАДАЧИ

Задача 25.

Произведено выборочное 10% обследование магазинов города.

Имеются следующие данные о величине товарооборота для 40 магазинов города:

Товарооборот, тыс. руб	Число магазинов		Товарооборот, тыс. руб	Число магазинов
100,0-150,0	1		250,0-300,0	7
150,0-200,0	9		300,0-350,0	5
200,0-250,0	15		350,0 и более	3

1. По данным ряда распределения определите: средний товарооборот для магазина; среднее квадратическое отклонение; коэффициент вариации; моду и медиану.
2. Изобразите полученный ряд графически в виде гистограммы и полигона распределения.

3. С вероятностью до 0,996 определите возможные пределы величины среднего товарооборота для всех магазинов.
4. С вероятностью до 0,993 установите возможные пределы удельного веса магазинов, имеющих товарооборот менее 200 тыс. руб. Сделайте выводы.

Задача 26.

Произведено выборочное 12% обследование магазинов города. Получены следующие данные о величине товарооборота для 50 магазинов города.

Товарооборот, тыс. руб	Число магазинов		Товарооборот, тыс. руб	Число магазинов
менее 100,0	15		300,0-400,0	7
100,0-200,0	12		400,0 –500,0	4
200,0-300,0	9		500,0 и более	3

1. По данным ряда распределения определите: средний товарооборот для магазина; среднее квадратическое отклонение; коэффициент вариации; моду, медиану.
2. Изобразите полученный ряд графически в виде гистограммы и полигона распределения.
3. С вероятностью 0,993 определите возможные пределы величины среднего товарооборота для всех магазинов.
4. С вероятностью 0,996 установите возможные пределы удельного веса магазинов, имеющих товарооборот более 200 тыс. руб.

Сделайте выводы.

Задача 27.

В целях изучения норм расходования сырья при изготовлении продукции проведена 15% механическая выборка, в результате которой получено следующее распределение изделий по массе:

Масса изделия, г	Число изделий, шт		Масса изделия, г	Число изделий, шт
до 20	1		22-23	25
20-21	20		23-24	10
21-22	40		24 и более	4

1. По данным ряда распределения определите: среднюю массу изделия; среднее квадратическое отклонение; коэффициент вариации; моду, медиану.

2. Изобразите полученный ряд графически в виде гистограммы и полигона распределения.
3. С вероятностью 0,993 определите предельную ошибку выборочной средней и возможные границы, в которых ожидается средняя масса изделия для всей партии изготовленных изделий.
4. С вероятностью 0,996 установите предельную ошибку выборочной доли и границы удельного веса изделий с массой веса от 22 до 24 г.

Сделайте выводы.

Задача28.

В целях изучения урожайности подсолнечника проведено 6% выборочное обследование 100 га посевов, в результате которого получены данные:

Урожайность, ц/га	Посевная площадь, га		Урожайность, ц/га	Посевная площадь, га
до 13	10		17-19	20
13-15	40		свыше 19	5
15-17	25			

Вычислите:

1. среднюю урожайность подсолнечника с 1 га; среднее квадратическое отклонение; коэффициент вариации; моду, медиану;
2. изобразите полученный ряд графически в виде гистограммы и полигона распределения;
3. с вероятностью 0,993 определите предельную ошибку выборочной средней и возможные границы, в которых ожидается средняя урожайность подсолнечника;
4. с вероятностью 0,996 установите предельную ошибку выборочной доли и границы удельного веса посевных площадей с урожайностью от 15 до 19 ц с га.

Сделайте выводы.

Задача 29.

Проведено 7% обследование обувных магазинов города. На основании данных продажи пар обуви, приведенных в задаче 1:

1. постройте дискретный ряд распределения;
2. найдите средний размер реализованной обуви и показатели вариации: среднее квадратическое отклонение, коэффициент вариации, моду и медиану;
3. с вероятностью 0,993 определите предельную ошибку выборочной средней и возможные границы, в которых ожидается средний размер обуви;

4. с вероятностью 0,996 установите предельную ошибку выборочной доли и границы удельного веса обуви, имеющей размеры от 36 до 37;
5. полученный ряд изобразите графически в виде полигона распределения.

Сделайте выводы.

Задача 30.

Проведено 10% обследование фермерских хозяйств. Данные приведены в задаче 3.

1. По данным о посевной площади для 40 фермерских хозяйств постройте интервальный ряд распределения.
2. Полученный ряд изобразите графически в виде гистограммы и полигона распределения.
3. Найдите средний размер посевной площади и показатели вариации: среднее квадратическое отклонение, коэффициент вариации, моду и медиану.
4. С вероятностью 0,993 определите предельную ошибку выборочной средней и возможные границы, в которых ожидается средний размер посевной площади.
5. С вероятностью 0,996 установите предельную ошибку выборочной доли и границы удельного веса участков, имеющих размеры менее 50 га.

Сделайте выводы.

Задача 31.

Для определения средней продолжительности телефонных разговоров по сотовой системе связи в порядке случайной бесповторной выборки обследовано 150 телефонных разговоров. При этом было установлено, что среднее квадратическое отклонение по длительности разговора равно 55 с.

КАКОВА вероятность того, что предельная ошибка средней продолжительности телефонных разговоров не превышает 9 с.?

Задача 32.

Для изучения общественного мнения о работе правоохранительных органов в порядке механического отбора опрошено 4 500 чел, или 2% общей численности городского населения. Из числа опрошенных положительно оценили работу правоохранительных органов 1 872 чел.

С вероятностью 0,954 ОПРЕДЕЛИТЕ:

1) выборочную долю (удельный вес) опрошенных, положительно оценивших работу правоохранительных органов;

2) предельную ошибку выборочной доли и возможные пределы, в которых будет находиться доля всего городского населения, положительно оценивающего работу правоохранительных органов;
3) возможные пределы (доверительный интервал), в которых будет находиться численность всего городского населения, положительно оценивающего работу правоохранительных органов.

Задача 33.

Для изучения уровня душевого дохода проведено методом механического отбора 1%-ное обследование 800 семей города. Оказалось, что 276 семей из числа обследованных относились к малообеспеченным.

С вероятностью 0,954 ОПРЕДЕЛИТЕ:
1) выборочную долю (удельный вес) малообеспеченных семей по данным выборки;
2) предельную ошибку выборочной доли и возможные пределы, в которых будет находиться доля малообеспеченных семей среди всех семей города.

4. ИНДЕКСЫ

4.1. Понятие об индексах и их классификация

Понятие об индексах. Слово индекс *(index)* в переводе с латинского языка означает показатель, указатель. *Индексами* в статистике называют относительные величины, характеризующие уровень планового задания, степень выполнения плана и изменение уровня явления во времени, а также результат сопоставления изучаемого явления в пространстве. Чаще индексы применяются при изучении динамики общественных явлений. Индексы выражаются в коэффициентах (когда база для сравнения принимается за единицу) или в процентах (когда база для сравнения принимается за 100).

При построении индексов различают два вида совокупностей: однородные и неоднородные. *Однородные (соизмеримые)* совокупности состоят из сходных по натурально-вещественной форме единиц (например, товары в магазине одного вида – хлеб, молоко и т.п.). *Неоднородные (несоизмеримые)* совокупности – совокупности, состоящие из различных по натурально-вещественной форме единиц (например, различные товары в магазине многих видов). Единицы однородной совокупности можно суммировать. Отличительной особенностью неоднородных совокупностей является невозможность непосредственного суммирования единиц этих совокупностей (например, нельзя складывать различные виды продовольственных товаров).

Обозначения индексируемых показателей. При помощи индексного метода определяется роль отдельных факторов в изменении сложного явления во времени или пространстве. Для решения этой задачи индексируемые показатели могут быть подразделены на: факторные (количественные и качественные) и результативные. Факторные показатели выступают в качестве аргумента; результативные – в качестве функции, т.е. они изменяются под влиянием факторных показателей.

Количественные показатели характеризуют размер (объем) того или иного явления и выражаются абсолютными величинами, например:

q – количество проданного товара (физический объем товарооборота) или количество произведенной продукции (физический объем продукции) в натуральном выражении;

Качественные показатели характеризуют уровень явления в расчете на ту или иную единицу совокупности, например:

p – цена единицы товара (продукции, работ, услуг);

Результативные (или количественно-качественные, сложные, производные) показатели отражают результат взаимодействия количественных и качественных показателей, например:

pq – товарооборот (стоимость проданных товаров) или стоимость произведенной продукции;

Обозначения периодов. Для того, чтобы различать, к какому периоду относится индексируемый показатель возле его символа внизу справа ставят знаки:

- *"1"* – знак отчетного (текущего) периода (т.е. периода, уровень которого сравнивается), например, q_1 - количество проданного товара в натуральном выражении в отчетном периоде и т.д.;
- *"0"* – знак базисного (предыдущего) периода (т.е. того периода, с уровнем которого производится сравнение), например, p_0 - цена единицы товара в базисном периоде и т.д.

Если производится расчет индексов планового задания или выполнения плана, то период, для которого устанавливается уровень планового задания, обозначается подстрочным знаком *"пл"*.

Если изменение социально-экономических явлений изучается за ряд периодов, то каждый период обозначается соответственно подстрочным знаком: *"0", "1", "2", "3"* и т.д.

Обозначения индексов:

i – индивидуальные индексы;

I – общие индексы.

Внизу справа символа индекса принято ставить обозначение индексируемого показателя (например, i_p – индивидуальный индекс цен).

4.2. Индивидуальные индексы

Индивидуальные индексы рассчитываются для однородных совокупностей, они характеризуют изменение только одного элемента совокупности.

Построим индивидуальные (в дальнейшем и общие) индексы применительно к решению задачи изучения *динамики* социально-экономических явлений. В общем виде индивидуальные индексы могут быть представлены формулой:

$$i_x = \frac{x_1}{x_0}, \tag{4.1}$$

где x_1 и x_0 – величина индексируемого показателя (количественного, качественного, количественно-качественного) соответственно в отчетном и базисном периодах.

Если в результате вычислений полученный индекс больше единицы или 100%, то это указывает на рост уровня изучаемого явления, если же меньше единицы или 100%, то - на снижение уровня. Если индекс выражен в процентах, то разность между его величиной и 100% $(\Delta i_x = i_x - 100)$ покажет, на сколько процентов увеличился (+) или уменьшился (-) индексируемый показатель.

Исходя из принятых условных обозначений индексируемых показателей, несложно записать формулы *конкретных* индивидуальных индексов.

Индивидуальный индекс *физического объема товарооборота*:

$$i_q = \frac{q_1}{q_0}, \tag{4.2}$$

где q_1 и q_0 – количество проданного товара в натуральном выражении соответственно в отчетном и базисном периодах.

Расчет индивидуального индекса *физического объема продукции* производится по формуле (4.2) (где q - количество произведенной продукции в натуральном выражении).

Индивидуальный индекс *цен:*

$$i_p = \frac{p_1}{p_0}, \tag{4.3}$$

где p_1 и p_0 – цена единицы товара (продукции) соответственно в отчетном и базисном периодах.

Индивидуальный *индекс товарооборота (товарооборота в фактических ценах или стоимости проданных товаров):*

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0}, \tag{4.4}$$

где p_1q_1 и p_0q_0 – товарооборот в фактических ценах (стоимость проданных товаров) соответственно отчетного и базисного периодов.

Если $i_{pq} = 0{,}947$ или 94,7%, то индивидуальный индекс товарооборота показывает снижение стоимости проданных товаров на 5,3%. Между индивидуальными индексами товарооборота (товарооборота в фактических ценах или стоимости проданных товаров), цен и физического объема товарооборота имеет место следующая взаимосвязь:

$$i_{\text{pq}} = \frac{p_1 q_1}{p_0 q_0} = \frac{P_1}{P_0} \cdot \frac{q_1}{q_0} = i_P \cdot i_q, \text{итак } i_{\text{pq}} = i_P \cdot i_q. \qquad (4.5)$$

Это свойство позволяет вычислить любой из трех взаимосвязанных индексов (если известны два других). Оно распространяется также для любого числа сомножителей.

4.3. Общие индексы в агрегатной форме

Общие индексы вычисляют для сложных совокупностей, состоящих из различных по натурально-вещественной форме единиц (например, для набора различных потребительских товаров). Основной, исходной и наиболее распространенной формой любого общего индекса является *агрегатная*. Латинское слово "агрегат" (*aggregatus*) означает складываемый, суммируемый несоизмеримых и не поддающихся суммированию элементов. Методология построения общих индексов в агрегатной форме составляет главное содержание теории индексов. Их строят различно для количественных, качественных и количественно-качественных показателей.

Агрегатная форма общего индекса физического объема товарооборота. С помощью индивидуальных индексов физического объема товарооборота выявляется динамика физического объема товарооборота, т.е. количества проданных товаров *конкретных* видов, например, продовольственных товаров (хлебобулочные изделия, мясопродукты, молочные продукты, овощи, фрукты, яйца и т.д.). По одним видам товаров, возможно, будет наблюдаться увеличение физического объема товарооборота, по другим – снижение. Как получить представление об изменении физического объема товарооборота по всем, например, продовольственным товарам (или в целом как по продовольственным, так и непродовольственным товарам)?

Для определения общего количества проданных товаров в отчетном и базисном периодах непосредственно суммировать эти товары нельзя: *во-первых,* они имеют различные единицы измерения (*кг, л, шт* и т.д.), более того, суммирование товаров, имеющих одинаковые единицы измерения, также экономически бессмысленно – "результат" суммирования литров молока с литрами винно-водочных изделий; *во-вторых,* товары (продукты) не равноценны по количеству затраченного на них общественного труда и, следовательно, стоимости, денежным измерителем которой выступает цена (p). Таким образом, для того чтобы определить, как изменилось количество проданных товаров в отчетном периоде по сравнению с базисным, необходимо найти соизмеритель, который бы привел несоизмеримые элементы признака "количество" в соизмеримый вид. Таким соизмерителем, или *весами* выступают цены, т.е. в качестве весов берутся неизменные цены как для базисного, так и для отчетного периодов. Применение неизменных цен устраняет (элиминирует) влияние динамики цен на динамику проданной (произведенной или потребленной) продукции.

Пользуясь принятыми обозначениями, агрегатную форму общего индекса физического объема товарооборота запишем в следующем виде:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}, \qquad (4.6)$$

где в числителе дроби – условная стоимость проданных в отчетном периоде товаров в ценах базисного периода, а в знаменателе – фактическая стоимость проданных товаров в базисном периоде. Знак Σ означает, что суммируются стоимости (pq) различных товаров (количество слагаемых зависит от количества видов товаров).

В формулах общих индексов следует различать: 1) *индексируемую,* т.е. переменную величину (в формуле (4.6) индексируемой величиной является количество проданного товара – q); 2) *веса* – постоянная величина (в формуле (4.6) в качестве весов используются цены базисного периода - p_0).

В факторном индексном анализе разность между числителем и знаменателем общего индекса физического объема товарооборота называют *абсолютным приростом товарооборота* за счет изменения количества проданных товаров (в денежных единицах):

$$\Delta \mathrm{pq}_q = \sum p_0 q_1 - \sum p_0 q_0. \qquad (4.7)$$

На практике общие индексы *количественных* показателей в агрегатной форме взвешиваются по весам *базисного* периода.

Агрегатная форма общего индекса цен. Общий индекс цен характеризует изменение уровня цен в среднем в совокупности товаров, т.е. *индексируемой величиной* является *цена* товара (p). В качестве *весов* индекса используется *количество* проданных товаров (q) одного и того же периода (либо базисного периода, либо отчетного периода). В этой связи возможны две схемы построения общего индекса цен.

Впервые агрегатная форма общего индекса цен была предложена в 1864 г. немецким ученым Э. Ласпейресом, в которой в качестве весов использовалось количество проданных товаров базисного периода (q_0):

$$I_p^{\text{Л}} = \frac{\sum p_1 q_0}{\sum p_0 q_0}, \qquad (4.8)$$

где в числителе дроби – условная стоимость товаров, проданных в базисном периоде по ценам отчетного периода, а в знаменателе – фактическая стоимость товаров, проданных в базисном периоде.

Вторая схема построения общего индекса цен с использованием в качестве весов количество проданных товаров отчетного периода (q_1) предложена в 1874 г. немецким экономистом Г. Пааше:

$$I_p^{П} = \frac{\sum p_1 q_1}{\sum p_0 q_1}, \tag{4.9}$$

где в числителе дроби – фактическая стоимость товаров, проданных в отчетном периоде, а в знаменателе – условная стоимость товаров, проданных в отчетном периоде по ценам базисного периода.

Индексы цен Ласпейреса и Пааше имеют различное экономическое содержание:

- индекс цен Ласпейреса показывает, на сколько изменились цены в отчетном периоде по сравнению с базисным по тем товарам, которые были реализованы в базисном периоде;
- индекс цен Пааше характеризует изменение цен отчетного периода по сравнению с базисным по товарам, реализованным в отчетном периоде.

Из двух схем построения общего индекса цен предпочтение, в условиях рыночной экономики, роста потребительских цен и изменения уровня и структуры потребления, отдается схеме Э. Ласпейреса.

Экономия (дополнительные затраты) населения в результате изменения цен определяется на основе общего индекса цен в агрегатной форме как разность между числителем и знаменателем:

$$Э^{Л} = \sum p_1 q_0 - \sum p_0 q_0, \tag{4.10}$$

$$\text{или } Э^{П} = \sum p_1 q_1 - \sum p_0 q_1. \tag{4.11}$$

Если эта разность (в денежных единицах) имеет знак «–», то это означает экономию (в результате снижения цен), если «+» - дополнительные затраты населения (в результате роста цен). Причем на основе формулы Ласпейреса получим *условную* сумму экономии (дополнительных затрат) населения от изменения цен (4.10), а на основе формулы Пааше - *фактическую* сумму (4.11). В факторном индексном анализе разность между числителем и знаменателем общего индекса цен называют *абсолютным приростом (снижением) товарооборота* в результате изменения *цен* (т.е. влияния этого, одного фактора).

Общие индексы *качественных* показателей в агрегатной форме взвешиваются по весам *отчетного* периода. Это правило относится ко всем индексам качественных показателей, кроме индекса цен, который может быть взвешен по весам как базисного, так и отчетного периодов. Учитывая данное правило, рассмотрим методы построения других общих индексов качественных показателей.

Общие индексы результативных (количественно–качественных) показателей. В отличие от предыдущих индексов (количественного, затем качественного показателей) общий индекс количественно-качественного показателя – **общий индекс товарооборота** (товарооборота в фактических ценах или стоимости проданных товаров) I_{pq} строится как отношение сумм соответствующих показателей отчетного и базисного периодов и будет иметь следующий вид:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \tag{4.12}$$

где $p_1 q_1$ и $p_0 q_0$ – товарооборот в фактических ценах соответственно отчетного

и базисного периодов по совокупности каких-либо товаров.

Общий индекс *товарооборота* (товарооборота в фактических ценах или стоимости проданных товаров) показывает *изменение в среднем товарооборота по* какой-либо *совокупности* товаров в отчетном периоде по сравнению с базисным. По этой же формуле определяется общий индекс *стоимости продукции.*

В факторном индексном анализе разность между числителем и знаменателем общего индекса товарооборота (4.12) называют *абсолютным приростом товарооборота за счет влияния двух факторов*, т.е. изменения цен и количества проданных товаров.

Общие индексы количественно-качественных показателей в агрегатной форме строятся как отношение сумм соответствующих показателей двух периодов и взвешивания не требуют. Помимо индексов *товарооборота* (товарооборота в фактических ценах или стоимости проданных товаров) в экономико-статистическом анализе применяются другие агрегатные индексы. Рассмотрим некоторые из них.

4.4. Индексы переменного состава, постоянного состава и структурных сдвигов

Качественные индексируемые показатели чаще всего отображаются не индивидуальными, а средними величинами. При этом общая средняя представляет собой среднюю из частных средних. Она складывается, с одной стороны, под влиянием значений показателя и индивидуальных элементов (единиц), из которых состоит совокупность, и, с другой стороны, - под влиянием соотношения их весов (структуры) совокупности. Поэтому при анализе динамики среднего уровня возникает вопрос, в какой мере изменение среднего уровня обусловлено действием каждого фактора в отдельности. Ответ можно получить путем построения системы взаимосвязанных индексов: индексов переменного состава, постоянного состава и структурных сдвигов.

Индекс переменного состава. Этот индекс называют также *индексом среднего уровня*, т.к. он показывает изменение средней величины явления во времени (например, изменение средней цены 1 кг картофеля по совокупности рынков, или средней урожайности изделий по совокупности сельскохозяйственных предприятий и т.п.). В общем виде индекс переменного состава это отношение средней величины качественного показателя в отчетном периоде к средней его величине в базисном периоде:

$$i_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum x_1 m_1}{\sum m_1} \div \frac{\sum x_0 m_0}{\sum m_0}, \qquad (4.19)$$

где x – индексируемая величина (ею могут быть p – цена единицы товара, продукции, z – себестоимость единицы продукции, V – урожайность и т.д.);

m – веса (q – физический объем товарооборота или продукции, $П$ – посевная площадь и т.д.);

$\bar{x}_0$ и $\bar{x}_1$ – средняя величина (средняя арифметическая взвешенная) изучаемого качественного показателя соответственно в базисном и отчетном периодах: $\bar{x} = \frac{\sum xm}{\sum m}$.

Индекс переменного состава показывает *изменение средней* величины ($\bar{x}$) вследствие совместного влияния *двух факторов*: 1) изменения уровня индексируемого показателя в отдельных частях совокупности (x); 2) изменения частей совокупности (m) или доли (удельного веса) этих частей совокупности (f), т.е. изменения структуры совокупности, структурных сдвигов. В силу этого индекс переменного состава не обязательно должен находиться в границах соответствующих индивидуальных индексов. Он может оказаться больше наибольшего индивидуального индекса или меньше наименьшего. Такое положение получило название *парадокса* индекса переменного состава.

Разность между числителем и знаменателем индекса переменного состава ($\Delta\bar{x}_{(x,f)}$) показывает *абсолютное изменение* средней величины изучаемого показателя за счет совместного влияния *двух факторов*:

$$\Delta\bar{x}_{(x,f)} = \bar{x}_1 - \bar{x}_0 = \frac{\sum x_1 m_1}{\sum m_1} - \frac{\sum x_0 m_0}{\sum m_0}. \qquad (4.20)$$

Учитывая, что индекс переменного состава зависит от двух факторов, возникает вопрос: в какой степени изменение средней величины произошло вследствие изменения отдельных этих факторов? Для оценки влияния каждого фактора на изменение средней величины строят два индекса: 1) индекс постоянного состава; 2) индекс структурных сдвигов. При построении этих индексов поочередно устраняется (элиминируется) влияние одного из факторов, зафиксировав его на постоянном уровне.

Индекс постоянного состава. Он показывает *изменение средней величины* показателя только *за счет изменения уровня* показателя в отдельных частях совокупности, или, можно сказать по-другому, индекс постоянного состава показывает *изменение в среднем* показателя за счет изменения усредняемых уровней показателя – это общепринятая точка зрения. Данный индекс взвешивается по весам отчетного периода:

$$\bar{i}_x = \frac{\sum x_1 m_1}{\sum m_1} \div \frac{\sum x_0 m_1}{\sum m_1}. \qquad (4.21)$$

Разность между числителем и знаменателем индекса постоянного состава ($\Delta\bar{x}_x$) показывает *абсолютное изменение* средней величины изучаемого показателя *за счет изменения уровня* показателя в отдельных частях совокупности:

$$\Delta\bar{x}_x = \frac{\sum x_1 m_1}{\sum m_1} - \frac{\sum x_0 m_1}{\sum m_1}. \qquad (4.22)$$

Индекс влияния структурных сдвигов. Этот индекс принято взвешивать по весам базисного периода:

$$i_f = \frac{\sum x_0 m_1}{\sum m_1} \div \frac{\sum x_0 m_0}{\sum m_0}. \qquad (4.23)$$

В индексе структурных сдвигов устраняется влияние первого фактора – x и оценивается влияние изменения второго фактора – m. Индекс структурных сдвигов показывает: как изменилась средняя величина показателя за счет изменения структуры совокупности, или в какой мере влияет изменение состава (структуры) совокупности за изучаемый период на изменение среднего уровня показателя.

Разность между числителем и знаменателем индекса структурных сдвигов ($\Delta\bar{x}_f$) показывает *абсолютное* изменение средней величины изучаемого показателя *за счет изменения структуры совокупности*:

$$\Delta\bar{x}_f = \frac{\sum x_0 m_1}{\sum m_1} - \frac{\sum x_0 m_0}{\sum m_0}. \qquad (4.24)$$

Между индексами переменного состава, постоянного состава и структурных сдвигов имеет место следующая взаимосвязь:

$$i_{\bar{x}} \qquad \bar{i}_x \qquad i_f. \qquad (4.25)$$

Взаимосвязь индексов позволяет также на основе двух известных индексов определить третий - неизвестный.

Аналогичная взаимосвязь имеет место между общим и факторными абсолютными изменениями средней величины:

$$\Delta\bar{x}_{(x,f)} \qquad \Delta\bar{x}_x \ + \ \Delta\bar{x}_f. \qquad (4.54)$$

ЗАДАЧИ

Задача 34.

Имеется информация о реализации продуктов на рынке:

Продукты	Количество реализованных товаров, кг		Цена за 1 кг, руб	
	базисный период	отчетный период	базисный период	отчетный период
А	2 000	2 100	30	50
Б	1 500	1 880	50	65

B	3 500	6 000	80	75

ОПРЕДЕЛИТЕ:

1. Индивидуальные индексы цен, физического объема товарооборота и стоимости реализованных товаров (товарооборота).
2. Общие индексы: а) цен (индексы Пааше и Ласпейреса); б) физического объема товарооборота; в) стоимости реализованных товаров (товарооборота).
3. Сумму экономии (перерасхода) населения от изменения цен на эти товары: а) фактическую сумму – на основе формулы Пааше; б) условную сумму – на основе формулы Ласпейреса.

Сделайте выводы. Покажите взаимосвязь между исчисленными общими индексами.

Задача 35.

Имеется следующая информация о реализации продуктов на рынке:

Продукты	Количество, т		Цена за 1 кг, руб	
	базисный период	отчетный период	базисный период	отчетный период
Картофель	10	60	57,0	59,6
Морковь	15	20	65,5	68,5
Свекла	7	5	49,8	50,0

Определите:

1. Индивидуальные индексы цен, физического объема товарооборота и объема товарооборота в фактических ценах.
2. Общие индексы цен (индекс Пааше), физического объема товарооборота и объема товарооборота в фактических ценах.
3. Абсолютное изменение объема товарооборота в целом и за счет действия отдельных факторов (вследствие изменения цен и физического объема товарооборота).

Сделайте выводы. Покажите взаимосвязь между исчисленными общими индексами.

Задача 36.

Имеется следующая информация о реализации продуктов на рынке:

Продукты	Базисный период		Отчетный период	
	количество, т	цена, руб	количество, т	цена, руб
А	52	55	50	68
Б	15	86	25	95
В	65	37	100	42

Определите:

1. Индивидуальные индексы цен, физического объема товарооборота и стоимости реализованных товаров (товарооборота).
2. Общие индексы: а) цен (индексы Пааше и Ласпейреса); б) физического объема товарооборота; в) стоимости реализованных товаров (товарооборота).
3. Сумму экономии (перерасхода) населения от изменения цен на эти товары: а) фактическую сумму – на основе формулы Пааше; б) условную сумму – на основе формулы Ласпейреса.

Сделайте выводы. Покажите взаимосвязь между исчисленными общими индексами.

Задача 37.

Имеется следующая информация о реализации продукта ***А*** на рынках города:

Рынки	Январь		Март	
	количество, т	цена, руб	количество, т	цена, руб
1	100	40	20	78
2	20	50	50	80
3	15	60	70	100

Определите:

1) индекс цен переменного состава;
2) индекс цен постоянного состава;
3) индекс влияния структурных сдвигов;
4) изменение средней цены единицы продукта ***А*** (в абсолютных величинах) в марте по трем рынкам города в целом и за счет действия отдельных факторов:

   а) изменения цены единицы продукта ***А*** в отдельных рынках города;

   b) изменения структуры продажи продукта ***А*** на рынках города.

Сделайте выводы.

Задача 38.

Имеется информация о реализации мяса говядины на рынках г. Саратова:

Рынок	Январь		Март	
	количество, т	цена 1 кг, руб	количество, т	цена 1 кг, руб
Крытый	100	90	120	95
Центральный	90	80	110	90
Северный	55	70	100	80

Определите:

1) индекс цен переменного состава;
2) индекс цен постоянного состава;
3) индекс влияния структурных сдвигов;
4) изменение средней цены мяса говядины (в абсолютных величинах) в марте по трем рынкам города в целом и за счет действия отдельных факторов:
   a) изменения цены мяса говядины в отдельных рынках города;
   b) изменения структуры продажи мяса говядины на рынках города.

Сделайте выводы.

Задача 39.

Имеется следующая информация о реализации картофеля в продовольственных магазинах города:

Магазины	I квартал		II квартал	
	количество, т	цена 1 кг, руб	количество, т	цена 1 кг, руб
1	20	6,8	60	7,5
2	30	6,5	32	7,3
3	55	6,3	10	7,1

Определите:

1. Индексы цен картофеля по каждому магазину.
2. По двум магазинам вместе:
   a) индекс цен переменного состава;
   b) индекс цен постоянного состава;
   c) индекс влияния структурных сдвигов.
3. Изменение средней цены картофеля (в абсолютных величинах) во II квартале по магазинам города в целом и за счет действия отдельных факторов:
   a) изменения цены картофеля в отдельных продовольственных магазинах;

b) изменения структуры продажи картофеля в магазинах.

Сделайте выводы.

Задача 40.

В отчетном периоде по сравнению с базисным объем товарооборота в фактических ценах увеличился на 30,0% или на 50,0 тыс. руб, цены возросли на 45,0%.

Определите:

1) абсолютное (абсолютный прирост) и относительное (индекс) изменение физического объема товарооборота;
2) дополнительные расходы населения от повышения цен.

Задача 40.

В отчетном периоде по сравнению с базисным объем товарооборота в фактических ценах увеличился на 35,0% или на 100,0 тыс. руб, цены снизились на 15,0%.

Определите:

1) абсолютное (абсолютный прирост) и относительное (индекс) изменение физического объема товарооборота;
2) экономию населения от снижения цен.

Задача 41.

В отчетном периоде по сравнению с базисным цены увеличились на 50,0%, физический объем продаж возрос на 25,0% или на 80,0 тыс. руб.

Определите:

1) абсолютное (абсолютный прирост) и относительное (индекс) изменение объема товарооборота в фактических ценах;
2) дополнительные расходы населения от повышения цен.

Задача 42.

В отчетном периоде по сравнению с базисным цены увеличились на 15,0%, физический объем продаж уменьшился на 25,0% или на 120,0 тыс. руб.

Определите:

1) абсолютное (абсолютный прирост) и относительное (индекс) изменение объема товарооборота в фактических ценах;
2) дополнительные расходы населения от повышения цен.

Задача 43.

В отчетном периоде по сравнению с базисным цены снизились на 20,0%, физический объем продаж увеличился на 15,0% или на 90,0 тыс. руб.

Определите:

1) абсолютное (абсолютный прирост) и относительное (индекс) изменение объема товарооборота в фактических ценах;
2) экономию населения от снижения цен.

Задача 44.

В отчетном периоде по сравнению с базисным товарооборот в фактических ценах увеличился на 40,0% или на 200,0 тыс. руб, физический объем продаж уменьшился на 25,0%.

Определите:

1) абсолютное изменение (абсолютный прирост) физического объема товарооборота;
2) относительное изменение (индекс) цен;
3) экономию (дополнительные затраты) населения от изменения цен.

5. РЯДЫ ДИНАМИКИ

5.1. Определение и виды динамических рядов

Процессы и явления общественной жизни, которые изучаются статистикой, находятся в постоянном движении и развитии. При этом меняются размеры, состав, объем и структура конкретных общественных явлений. Изучая общественную жизнь во времени, приходится иметь дело с динамическими (временными) рядами.

Динамический ряд – это ряд чисел, характеризующих изменение того или иного социально-экономического явления во времени. В любом динамическом ряду различаются два элемента:

1) перечень хронологических дат (моментов) или периодов;
2) конкретные количественные значения соответствующего показателя на эти даты или периоды, которые называются уровнями ряда.

Различают крайние и промежуточные уровни. Крайние – это первый и последний уровни ряда. Первый уровень называют также начальным, последний – конечным. Промежуточными уровнями являются все остальные уровни.

Примером рядов динамики могут служить данные табл. 5.1.

Таблица 5.1

Основные показатели
отдельных отраслей промышленности Саратовской области

	2010 г.	2014 г.	2015 г.	2016 г.	2017 г.
Добыча нефти (включая газовый конденсат), тыс. т	1 215	1 535	1 429	1 615	1 632
Добыча естественного газа, млн $м^3$	447	431	506	480	538

В основу классификации динамических рядов следует отнести три признака:

1) способ получения показателя;
2) время, к которому относится показатель;
3) статистическая природа показателя.

1. В зависимости от *способа получения показателя* динамические ряды делятся на:

a) первичные;
b) производные.

Первичные – это динамические ряды показателей, определяемых непосредственно в результате сводки данных статистического наблюдения путем подсчета итогов “по карточкам” или итогов по признакам (например, динамические ряды численности работающих, объема произведенной продукции, стоимости основных средств и т.д.).

Производные – это динамические ряды показателей, которые получаются на основе первичных показателей (например, динамические ряды средней заработной платы, средней выработки, доли мужчин в общей численности населения, доли городского населения и т.д.).

2. В зависимости от *времени*, к которому относится показатель, различают моментные и интервальные динамические ряды.

В *моментных* динамических рядах уровни ряда выражают величину явления на определенную дату (или же момент времени). В них время обозначает момент, к которому относится каждый уровень ряда. Суммировать уровни моментного динамического ряда нельзя, т.к. сумма не имеет экономического смысла.

В *интервальных* динамических рядах каждый уровень ряда относится к определенному интервалу времени (месяцу, кварталу, году и т.д.). Например, производство продукции, фонд заработной платы, средняя урожайность, средняя заработная плата, коэффициенты рождаемости, смертности и т.д.). Уровни интервальных динамических рядов можно суммировать во времени только для абсолютных величин.

3. По *статистической природе показателя* динамические ряды могут быть рядами абсолютных, относительных и средних величин.

Различают два основных способа построения статистических графиков динамики: 1) в виде ломаных “кривых”; 2) в виде столбиковой или полосовой диаграмм. На оси абсцисс откладывается время, на оси ординат – уровни динамического ряда. На одной координатной сетке можно поместить три – четыре кривых, выделяя их особым цветом или рисунком:

сплошная линия, пунктирная, штрихпунктирная (подробно см. главу ”Графические изображения в статистике”).

Важнейшим условием правильного построения динамического ряда является *сравнимость* его уровней. Если уровни динамического ряда не сопоставимы между собой, то данный динамический ряд нельзя анализировать. В противном случае, выводы будут иметь ложный, иллюзорный характер.

5.2. Методы вычисления среднего уровня динамического ряда

Средний уровень динамического ряда является обобщающей характеристикой динамики – это уровень, представляющий собой специфическую среднюю из уровней динамического ряда. Эту среднюю называют также средней хронологической или средней временной.

Для интервальных и моментных динамических рядов способы расчета средней хронологической различны.

Средний уровень *интервального* динамического ряда определяется путем деления суммы уровней этого ряда на число периодов, к которым относится эта сумма, т.е.:

$$\overline{y} = \frac{\sum y_i}{n}, \qquad (5.1)$$

где y_i – i-й уровень динамического ряда $(i = \overline{1,n})$;

n – число периодов.

Пример 5.1. Имеются следующие данные о производстве продукции предприятием (табл.5.2).

Таблица 5.2

Производство продукции предприятием в 2017 г.

Кварталы	Выпуск продукции, тыс. руб
I	350,0
II	400,0
III	450,0
IV	500,0
Итого	1 700,0

Средняя хронологическая интервального динамического ряда вычисляется по формуле средней арифметической простой:

$\overline{y}_{кв} = \frac{1700}{4} = 425,0$ *тыс. руб;* $\overline{y}_{мес} = \frac{1700}{12} = 141,7$ *тыс. руб.*

Иначе рассчитывается средний уровень для моментных динамических рядов, который может быть вычислен только приближенно.

Пример 5.2. Допустим, имеем моментный ряд (табл. 5.3).

Таблица 5.3

Остатки товарно-материальных ценностей торговой базы за I квартал 2018 г.

Даты	1.01.08	1.02.08	1.03.08	1.04.08
Млн руб	12,0	18,0	24,0	8,0

Если воспользоваться полусуммой крайних уровней, то получим средние остатки товарно-материальных ценностей за I квартал:

$$\bar{y} = \frac{12{,}0 + 8{,}0}{2} = 10{,}0 \text{ млн руб.}$$

Результат будет *искаженным,* т.к. остатки товарно-материальных ценностей на начало и конец периода были меньше, чем на другие даты.

Определим средние остатки за каждый месяц:

$$\bar{y}_{ян} = \frac{12{,}0 + 18{,}0}{2} = 15{,}0 \text{ млн руб;}$$

$$\bar{y}_{фев} = \frac{18{,}0 + 24{,}0}{2} = 21{,}0 \text{ млн руб;}$$

$$\bar{y}_{март} = \frac{24{,}0 + 8{,}0}{2} = 16{,}0 \text{ млн руб.}$$

Исходя из средних месячных остатков товарно-материальных ценностей, по формуле средней арифметической простой можно определить средние остатки за I квартал:

$$\bar{y}_{I\ кв} = \frac{15{,}0 + 21{,}0 + 16{,}0}{3} = 17{,}3 \text{ млн руб.}$$

Таким образом, произведено преобразование моментного динамического ряда в интервальный, затем вычислен средний уровень ряда. Однако эти предварительные расчеты можно не производить. Тогда в общем виде формула средней хронологической для моментного динамического ряда запишется так:

$$= \frac{\frac{1}{2} + {}_2 + {}_3 + \ldots + {}_{n-1} + \frac{n}{2}}{n-1}. \qquad (5.2)$$

Итак, средняя хронологическая для моментного динамического ряда равна частному от деления полусуммы крайних уровней плюс сумма полных значений промежуточных уровней на число уровней без одного. В примере 5.2 получим:

$$\bar{y}_{кв} = \frac{\frac{12{,}0}{2} + 18{,}0 + 24{,}0 + \frac{8{,}0}{2}}{4-1} = 17{,}3 \text{ млн руб.}$$

Точнее, для моментных динамических рядов речь идет не о среднем уровне, а о среднем значении функции выравненного динамического ряда.

5.3. Абсолютные показатели динамики

К ним относятся:

1) абсолютный прирост;
2) средний абсолютный прирост.

1. Абсолютный прирост характеризует скорость изменения уровней динамического ряда. Он показывает, на сколько единиц увеличился или уменьшился уровень ряда, и может иметь как положительный, так и отрицательный знак. Абсолютные приросты измеряются в тех же единицах, что и уровни динамического ряда. Они могут быть двух видов:

a) цепными;
b) базисными.

В общем виде покажем примеры их расчета. Имеем n равностоящих во времени уровней динамического ряда:

$$y_1,\ y_2,\ y_3,\ \ldots,\ y_n.$$

Тогда, система цепных абсолютных приростов запишется следующим образом:

$$\begin{gathered}\delta_{2/1} = y_2 - y_1;\\ \delta_{3/2} = y_3 - y_2;\\ \delta_{4/3} = y_4 - y_3;\\ \ldots\\ \delta_{n/n-1} = y_n - y_{n-1}.\end{gathered} \qquad (5.3)$$

Система базисных абсолютных приростов имеет следующий вид[8]:

$$\begin{gathered}\delta_{2/1} = y_2 - y_1;\\ \delta_{3/1} = y_3 - y_1;\\ \delta_{4/1} = y_4 - y_1;\\ \ldots\\ \delta_{n/1} = y_n - y_1.\end{gathered} \qquad (5.4)$$

2. Средний абсолютный прирост характеризует среднюю скорость изменения явления во времени и является обобщающей характеристикой изменения уровней динамического ряда. Существует несколько способов расчета этого показателя, приведем три – основных:

1. Вычисление среднего абсолютного прироста на основе *цепных абсолютных приростов* за все рассматриваемые периоды:

$$\overline{\delta} = \frac{\delta_{2/1} + \delta_{3/2} + \delta_{4/3} + \ldots + \delta_{n/n-1}}{n-1}, \qquad (5.5)$$

где $n-1$ – число цепных абсолютных приростов.

Данная формула раскрывает статистическую природу среднего абсолютного прироста как средней арифметической простой.

2. Вычисление среднего абсолютного прироста на основе базисного абсолютного прироста всего динамического ряда:

---

[8] где y_1 – начальный уровень динамического ряда, принимаемый за базу для сравнения.

$$\overline{\delta} = \frac{\delta_{n/1}}{n-1}. \tag{5.6}$$

Применение этой формулы обусловлено тем, что базисный абсолютный прирост всего динамического ряда равен сумме последовательных цепных абсолютных приростов:

$$\delta_{n/1} = \delta_{2/1} + \delta_{3/2} + \delta_{4/3} + \dots + \delta_{n/n-1}. \tag{5.7}$$

3. Вычисление среднего абсолютного прироста на основе крайних уровней динамического ряда:

$$\overline{\delta} = \frac{y_n - y_1}{n-1}, \tag{5.8}$$

т.к.:

$$\delta_{n/1} = y_n - y_1. \tag{5.9}$$

Итак, в зависимости от имеющихся данных применяется та или иная формула среднего абсолютного прироста.

5.3. Относительные показатели динамики

Они характеризуют интенсивность изменения уровней динамического ряда за определенный промежуток времени и выражаются в виде коэффициентов или процентов. К относительным показателям динамики относятся:

1) темп роста;
2) темп прироста;
3) средний темп роста;
4) средний темп прироста;
5) абсолютное значение 1% прироста;
6) коэффициент (темп) опережения.

1. Темп (коэффициент) роста показывает, во сколько раз увеличился уровень ряда (если этот коэффициент больше единицы) или какую часть базисного уровня составляет уровень текущего периода (если он меньше единицы) за некоторый период времени.

Если коэффициенты роста выражают в процентах, то их называют темпами роста. В дальнейшем эти понятия разграничивать не будем. Имеется в виду, что:

$$T = K \cdot 100. \tag{5.10}$$

Коэффициенты (темпы) роста могут быть: а) цепными; б) базисными. Приведем эти системы коэффициентов роста:

а) система *цепных* коэффициентов роста:

$$K_{2/1} = \frac{y_2}{y_1},\ K_{3/2} = \frac{y_3}{y_2},\ K_{4/3} = \frac{y_4}{y_3}, \dots,\ K_{n/n-1} = \frac{y_n}{y_{n-1}}; \tag{5.11}$$

б) система *базисных* коэффициентов роста:

$$K_{2/1} = \frac{y_2}{y_1}, \quad K_{3/1} = \frac{y_3}{y_1}, \quad K_{4/1} = \frac{y_4}{y_1}, \dots, \quad K_{n/1} = \frac{y_n}{y_1}. \tag{5.12}$$

2. Темп прироста *(T')* характеризует относительную величину прироста и показывает, на сколько процентов увеличился (уменьшился) уровень текущего периода по сравнению с уровнем предшествующего (базисного) периода.

Также как и для темпов роста различают цепные и базисные темпы прироста:

а) система *цепных* темпов прироста (%):

$$T'_{i/i-1} = \frac{\delta_{i/i-1}}{y_{i-1}} \cdot 100, \qquad (i = \overline{2,n}) \tag{5.13}$$

или

$$T'_{i/i-1} = T_{i/i-1} - 100. \tag{5.14}$$

б) система *базисных* темпов прироста (%):

$$T'_{i/1} = \frac{\delta_{i/1}}{y_1} \cdot 100, \qquad (i = \overline{2,n}), \tag{5.15}$$

или

$$T'_{i/1} = T_{i/1} - 100. \tag{5.16}$$

3. Средний темп (коэффициент) роста является обобщающим показателем интенсивности развития за весь анализируемый период. Он характеризует среднюю относительную скорость изменения уровней динамического ряда за анализируемый период.

Существует несколько способов расчета среднего темпа (коэффициента) роста. Рассмотрим три способа:

1. Вычисление среднего коэффициента роста на основе цепных коэффициентов роста за все рассматриваемые периоды:

$$\overline{K} = \sqrt[n-1]{K_{2/1} \cdot K_{3/2} \cdot K_{4/3} \cdot \ldots \cdot K_{n/n-1}}, \tag{5.17}$$

где $n\text{-}1$ – число цепных коэффициентов роста.

Данная формула раскрывает статистическую природу среднего коэффициента роста как среднюю геометрическую из последовательных цепных коэффициентов роста.

2. Вычисление среднего коэффициента роста на основе базисного коэффициента роста всего динамического ряда:

$$\overline{K} = \sqrt[n-1]{K_{n/1}}. \tag{5.18}$$

Применение данной формулы обусловлено тем, что:

$$K_{n/1} = K_{2/1} \cdot K_{3/2} \cdot K_{4/3} \cdot \ldots \cdot K_{n/n-1}. \tag{5.19}$$

3. Вычисление среднего коэффициента роста на основе крайних уровней динамического ряда:

$$\overline{K} = \sqrt[n-1]{\frac{y_n}{y_1}}. \tag{5.20}$$

Формула (5.20) используется в связи с тем, что:

$$K_{n/1} = \frac{y_n}{y_1}.$$

На основе формулы (5.10) средний темп роста составит:

$$\overline{T} = \overline{K} \cdot 100. \tag{5.21}$$

При расчете среднего темпа роста необходимо правильно решить два вопроса:

1. Какой степени извлекать корень? Степень *(n-1)* определяется как разность точки отсчета (номера) конечного уровня динамического ряда и начального.
2. Как извлекать корень степени *n-1*, т.е. 5, 10, 13 и т.д.? Для этого могут быть использованы:

a) специальные таблицы для исчисления средних годовых месячных, квартальных) темпов роста и прироста;

b) логарифмирование:

$$\lg\overline{K} = \frac{1}{n-1}\lg K_{n/1}; \tag{5.22}$$

c) программы в микрокалькуляторах и компьютерах.

4. Средний темп прироста показывает, на сколько долей единицы или процентов изменялся уровень динамического ряда за весь анализируемый период. Он вычисляется как разность среднего темпа роста и 100:

$$\overline{T'} = \overline{T} - 100. \tag{5.23}$$

5. Абсолютное значение 1% прироста показывает, какой абсолютной величине соответствует изменение уровней ряда на 1% по сравнению с предыдущим уровнем. Данный показатель вычисляется как отношение абсолютного прироста за данный период к темпу прироста в процентах за тот же период:

$$A_{i/i-1} = \frac{\delta_{i/i-1}}{T'_{i/i-1}} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100} = \frac{y_{i-1}}{100}. \tag{5.24}$$

6. Коэффициент (темп) опережения используется для сравнительного анализа изменения во времени как одноименных, так и разноименных явлений, определяется делением темпов роста сравниваемых показателей, вычисленных за один и тот же период времени:

$$K_{\text{опер.}} = \frac{T_1}{T_2}, \tag{5.25}$$

причем $T_1 > T_2$.

Коэффициент опережения характеризует, на сколько процентов один из показателей больше другого.

ЗАДАЧИ

Задача 45.

Производство изделия ***А*** на предприятии характеризуется следующими данными:

Квартал	Выпуск, млн шт		Квартал	Выпуск, млн шт

I	2,3		*III*	3,0
II	2,8		*IV*	3,5

Для анализа динамики производства вычислите:

1.а) абсолютные приросты, темпы роста и темпы прироста (цепные и базисные) по кварталам. Полученные данные представьте в таблице;

б) среднее квартальное производство изделий А;

в) средний квартальный темп роста и прироста.

2. Постройте график динамики роста производства изделий.
3. Проведите выравнивание ряда динамики.
4. Сделайте выводы.

Задача 46.

Товарооборот магазина за 2017 г. характеризуется приведенными ниже данными:

Квартал	Товарооборот, млн руб		Квартал	Товарооборот, млн руб
I	2,3		*III*	3,0
II	2,8		*IV*	3,5

Для анализа динамики производства вычислите:

1.а) абсолютные приросты, темпы роста и темпы прироста (цепные и базисные) по кварталам. Полученные данные представьте в таблице;

б) средний квартальный объем товарооборота;

в) средний квартальный темп роста и прироста.

2. Постройте график динамики объема товарооборота.
3. Проведите выравнивание ряда динамики.
4. Сделайте выводы.

Задача 47.

Имеется следующая информация о товарообороте торговой организации до и после укрупнения (млн руб):

Товарооборот	2014	2015	2016	2017
В прежних грани-	560,0	620,0	-	-

цах				
В новых границах	-	810,0	950,0	1100,0

1. Произведите смыкание рядов динамики.
2. Проведите анализ погодовой динамики объема товарооборота и выравнивание ряда динамики.
3. Изобразите интенсивность развития товарооборота графически.
4. Сделайте выводы.

Задача 48.

Имеется следующая информация о величине собственного капитала предприятия до и после укрупнения (млн руб):

Собственный капитал	2015	2016	2017	2018
В прежних границах	600,0	900,0	-	-
В новых границах	-	1500,0	1670,0	2000,0

1. Произведите смыкание рядов динамики.
2. Проведите анализ погодовой динамики собственного капитала предприятия и выравнивание ряда динамики.
3. Сделайте выводы.

Задача 49.

Имеется следующая информация о товарных запасах торговой организации (млн руб):

на 1.01. 18 - 4,5, на 1.02. 18 - 4,8 , на 1.03. 18 - 3,5, на 1.04. 18 - 4,0

1. Вычислите:
   1) цепной и базисный абсолютный прирост;
   2) цепные и базисные темпы роста;
   3) средние товарные запасы за 1 квартал 2018 г.
2. Проведите выравнивание ряда динамики.
3. Сделайте выводы.

Задача 50.

Имеется следующая информация об остатках вкладов в сберегательных банках района (млн руб):

на 1.01.18 - 105, на 1.02. 18 - 108, на 1.03. 18 - 100, на 1.04. 18 - 106.

1. Вычислите:
    1) цепной и базисный абсолютный прирост;
    2) цепные и базисные темпы роста;
    3) средние остатки вкладов на 1 квартал 2018 г.
2. Проведите выравнивание ряда динамики.
3. Сделайте выводы.

5. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СВЯЗИ

Важнейшей задачей экономических исследований является выявление факторов, определяющих уровень и динамику экономического процесса.

Различают два вида зависимостей между экономическими явлениями: функциональную и стохастическую. Функциональная зависимость подразумевает существование однозначного отображения множества значений исследуемых величин, например, зависимость производительности труда от объема произведенной продукции и затрат рабочего времени: $Y=f(x_1,x_2)$.

При изучении реальных явлений сказывается влияние многих незначительных случайных факторов, поэтому каждому значению аргумента соответствует множество значений переменной Y, такая неоднозначность есть проявление стохастической зависимости. Например, при изучении производительности труда Y в зависимости от среднегодовой стоимости основных фондов X каждому значению X соответствует множество значений Y и наоборот. В этом случае говорят о наличии стохастической связи.

Обнаружение и измерение силы стохастической связи чаще всего решается методами корреляционного и регрессионного анализа.

Главной задачей ***корреляционного анали***за является оценка взаимозависимости между переменными величинами на основе выборочных данных. Дополнительно (задачей регрессионного анализа) является обнаружение форм связей между признаками (прямая, гипербола и т.д.).

Различают следующие виды корреляционной связи**:**

1. Относительно числа переменных включенных в модель различают парную корреляционную связь и множественную корреляцию, то есть корреляция между несколькими переменными.
2. Относительно характеристики корреляционную связь: положительную корреляцию – когда с увеличением одного признака второй тоже увеличивается, и отрицательную, когда с увеличением одного признака, значения второго уменьшаются.
3. Относительно формы связи: линейную корреляцию – когда взаимосвязь между переменными характеризуется линейными отношениями, и нелинейную корреляцию – если переменные связанны нелинейным соотношением.

Как характеристику линейной зависимости для определения тесноты связи между признаками используют ***коэффициент корреляции***.

Коэффициент корреляции изменяется в пределах от -1 до +1. Значения $r=\pm 1$ свидетельствуют о наличии линейной функциональной зависимости между признаками, почти наверное, т.е. существует a и b такие, что $P\{Y = aX + b\} = 1$. Если $r=0$, то признаки некоррелируемы (независимы). Положительный знак указывает на положительную корреляцию, то есть с увеличением X признак Y - растет. Отрицательный знак свидетельствует об отрицательной корреляции. Чем ближе $|r|$ к 1, тем зависимость между признаками более существенна, чем ближе к нулю, тем признаки более независимы.

Двумерная корреляционная модель

Объектом изучения при решении задач корреляционного является генеральная совокупность и репрезентативная выборка из нее $(x_1, y_1), (x_2, y_2),..., (x_n, y_n)$.

Корреляционная модель предполагает расчет следующих параметров:

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ - выборочное среднее признаков X и Y;

$\overline{xy} = \frac{1}{n}\sum_{i} x_i y_i$ - среднее XY.

$S_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_{i^2} - (\bar{x})^2}$ - выборочное среднее квадратическое отклонение признака X;

$S_y = \sqrt{\frac{1}{n}\sum_{j=i}^{n} y_{i^2} - (\bar{y})^2}$ - выборочное среднее квадратическое отклонение признака Y;

$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x S_y}$ - выборочный коэффициент линейной корреляции.

Если с помощью корреляционного анализа выявлено наличие статистической связи между признаками, обычно находят вид зависимостей, используя методы ***регрессионного анализа***.

Выборочное уравнение *линейной регрессии* имеет вид:

$$y(x) = b_0 + b_1(x - \bar{x}).$$

Оценки коэффициентов регрессии находят по формулам:

$$b_0 = \bar{y}; \qquad b_1 = r\frac{S_y}{S_x}.$$

Коэффициент детерминации $R^2=r^2$ показывает, какая доля вариации результативного признака вызвана признаком, положенным в основании группировки. Его значения находятся в пределах от 0 до 1.

Корреляция и регрессия связаны между собой. Однако в регрессионном анализе исследуется форма стохастической связи, а в корреляционном анализе ее сила.

Под *регрессионным анализом* понимают метод стохастического анализа зависимости Y от переменных X_j , рассматриваемых как неслучайные величины, независимо от истинного закона распределения. С помощью уравнения регрессии измеряют влияние каждого факторного признака X_j на результативный признак Y.

Построение уравнения регрессии предполагает решение двух задач:

1. Необходимо выбрать независимые перемененные оказывающие влияние на зависимую переменную Y и определить форму связи.
2. Оценить параметры уравнения регрессии. Выбор уравнения зависит от сущности изучаемого явления. Если происходит постоянный рост, то выбирают прямую. Если явление происходит с постоянным ускорением или замедлением, то выбирают параболу. Если происходит рост явления, а затем насыщение и процесс становится постоянным, то выбирают логические кривые.

Чаще всего встречаются следующие уравнения регрессии:

1. Прямая: $y(x) = \beta_0 + \beta_1 x$.
2. Парабола: $y(x) = \beta_0 + \beta_1 x + \beta x^2$.
3. Гипербола: $y(x) = \beta_0 + \beta_1 \frac{1}{x}$.
4. Полиномиальное: $y(x) = \beta_0 + \beta_1 x + \beta x^2 + \dots + \beta_n x^n$.
5. Множественное: $y(x) = \beta_0 + \beta_1 x_1 + \beta x_2 + \dots + \beta_n x_n$.
6. Степенное: $y(x) = \beta_0 x_{1\beta_1} x_{2\beta_2} \dots\dots x_{n\beta_n}$.
7. Логарифмическое: $y(x) = \ln\beta_0 + \beta_1 \ln x_1 + \beta \ln x_2 + \dots + \beta_n \ln x_n$

Для линейных уравнений теория разработана достаточно полно. Поэтому степенные уравнения регрессии логарифмированием могут быть преобразованы в линейные уравнения относительно параметров β_j. Также путем подстановок можно преобразовать гиперболическое и полиномиальное уравнения в линейные.

В общем виде линейная модель регрессионного анализа имеет вид:

$$y = \sum_{k=0}^{k} \beta_j f_j(x_1, x_2, \dots x_k) + \varepsilon,$$

где f_j некоторая функция переменных $x_1, x_2, \dots, x_n$; ε – случайная величина с нулевым математическим ожиданием и дисперсией σ^2.

Линейная парная регрессионная модель

Предположим, что на основании исследуемого явления необходимо оценить линейное уравнение регрессии: $y(x) = M(y/x) = \beta_0 + \beta_1 x$, где β_0, β_1 – неизвестные параметры генеральной совокупности, которые необходимо оценить по результатам репрезентативной выборки.

Пусть из генеральной совокупности (X, Y) произведена выборка объемом n: (x_1,y_1), (x_2,y_2), ….. , (x_n, y_n). Тогда модель регрессионного анализа будет иметь вид:

$$y_i = \beta_0 + \beta_1 x + \varepsilon_i,$$

где ε_i – независимые нормально распределенные случайные величины с нулевым математическим ожиданием и дисперсией σ^2, то есть $M=0$, $D\varepsilon_i = \sigma^2$ для всех $i = 1, 2, \ldots, n$.

В качестве оценок неизвестных параметров β_0, β_1 следует брать такие значения выборочных характеристик, которые минимизируют сумму квадратов отклонений значений результативного признака y_i от условного математического ожидания $\hat{y}_i$. Такой метод оценки называется *методом наименьших квадратов*.

Таким образом,

$$Q = \sum_{i=1}^{n}(y_i - y_i)^2 = \sum_{i=1}^{n}(y_i - b_0 - b_1 x_1)^2 \to \min.$$

Подставив вместо неизвестных параметров β_0, β_1 их оценки b_0 и b_1, найдем частные производные Q по b_0 и b_1, приравняв их к 0.

$$\frac{дQ}{дb_0} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i) = 0;$$

$$\frac{дQ}{дb_1} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)x_i = 0.$$

{

Или

$$\begin{aligned} nb_0 + b_1\sum x_i &= \sum y_i; \\ b_0\sum x_i + b_1\sum x_i^2 &= \sum x_i y_i. \end{aligned} \qquad (9.11)$$

{

Система (9.11) называется *системой нормальных уравнений*. Решая ее, получим оценки коэффициентов регрессии:

$$b_1 = \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}; \qquad (9.12)$$

$$b_0 = \bar{y} - b_1\bar{x} = \frac{1}{n}\sum y_i - b_1\frac{1}{n}\sum x_i\,. \qquad (9.13)$$

Таким образом, получаем оценку уравнения регрессии в виде:

$$y(x) = b_0 + b_1 x. \qquad (9.13)$$

Качество полученной модели оценивают по остаточной дисперсии:

$$\sigma^2 = \frac{1}{n-2}\sum(y_i - y_i)^2 = \frac{1}{n-2}\sum(y_i - b_0 - b_1 x_i)^2, \qquad (9.14)$$

или по *коэффициенту детерминации*:

$$R^2 = 1 - \frac{\sum(y_i - y_i)^2}{\sum(y_i - \bar{y})^2}. \qquad (9.14)$$

Для интерпретации модели можно использовать *коэффициент эластичности*:

$$e_1 = b_1\frac{\bar{x}}{\bar{y}}, \qquad (9.15)$$

который показывает, насколько процентов изменяется результативный признак при увеличении факторного признака на 1%.

ЗАДАЧИ

Задачи 40 – 45.

Результативные и факторные признаки задач (40-45) представлены в табл. 1.

Исследовать на основе корреляционного и регрессионного анализов зависимость одного из результативных признаков от показателей производственно-хозяйственной деятельности предприятий машиностроения (факторные признаки) по данным табл. 2.

Обозначения и наименование показателей:

y_1 - производительность труда;
y_2 - рентабельность;
x_1 – средняя годовая стоимость основных производственных фондов;
x_2 - фондоотдача;
x_3 - трудоемкость единицы продукции.

Таблица 1

Номер задачи	Результативный признак	Факторные признаки
40	y_1	x_1
41	y_1	x_2
42	y_1	x_3
43	y_2	x_1
44	y_2	x_2
45	y_2	x_3

Таблица 2

y_1	y_2	x_1	x_2	x_3
9,26	204,2	167,67	1,45	0,23
9,38	209,6	186,10	1,30	0,24
12,11	222,6	220,45	1,37	0,19
10,81	236,7	169,30	1,65	0,17
9,35	62,0	39,53	1,91	0,23
9,87	53,1	40,41	1,68	0,43
8,17	56,5	102,96	1,94	0,31
9,12	52,6	37,02	1,89	0,26
5,88	46,6	45,74	2,06	0,49
6,30	53,2	40,07	1,96	0,36
6,22	30,1	45,44	1,02	0,37
5,49	18,1	41,08	1,85	0,43
6,50	13,6	136,14	0,88	0,35

6,61	89,7	42,39	0,62	0,38
4,32	63,0	37,39	1,07	0,42
7,37	46,3	101,78	1,60	0,30
7,02	103,5	47,55	1,53	0,32
8,25	73,3	32,61	1,40	0,25
8,15	76,6	103,25	2,22	0,31
8,72	90,0	38,95	1,32	0,26

ЛИТЕРАТУРА

Основная:

1. Общая теория статистики. Практикум : учебное пособие для академического бакалавриата / М. Р. Ефимова, Е. В. Петрова, О. И. Ганченко, М. А. Михайлов ; под ред. М. Р. Ефимовой. — 4-е изд., пер. и доп. — М. : Издательство Юрайт, 2016. — 355 с. — (Серия : Бакалавр. Академический курс).
2. Теория статистики : практикум / Г.Л. Громыко. — 5-е изд., испр. и доп. — М. : ИНФРА-М, 2018. — 238 с. — (Высшее образование: Бакалавриат).
3. Статистика в 2 т. Том 1 : учебник для академического бакалариата / И. И. Елисеева [и др.] ; отв. ред. И. И. Елисеева. — 4-е изд., пер. и доп. — М. : Издательство Юрайт, 2017. — 332 с. — (Серия : Бакалавр. Академический курс).
4. Теория статистики: Учебник / Под ред. проф. *Р.А. Шмойловой.* – 3-е изд., перераб. М.: Финансы и статистика, 2014.
5. Статистика: учебник / коллектив авторов ; под ред. М.Г. Назарова. — 3-е изд., перераб. и доп. — М. : КНОРУС, 2016. — 408 с. — (Бакалаврат).

Дополнительная:

1. Статистика: социально-экономическая статистика: учебное пособие / под общей ред. В.А. Прокофьева / Саратовский государственный социально-экономический университет. – Саратов, 2013. – 120 с.
2. Общая теория статистики: курс лекций по дисциплине «Статистика» Т.А. Крутова [и др.]; РТА.- М.: Изд-во РТА, 2010.
3. Практикум по социально-экономической статистике: учебно-методическое пособие/ под ред. М.Г. Назарова; Академия бюджета и казначейства М-ва Финансов Рос. Федерации. - М.: КноРус, 2009.
4. Статистика [Электронный ресурс]/ Салин В.Н., Чурилова Э.Ю., Шпаковская Е.П.– Электрон. дан. – М.: Кнорус, 2008.
5. Статистика [Электронный ресурс]/ Назаров М.Г., Варагин В.С., Великанова Т.Б.; под ред. Назарова М.Г. – Электрон. дан. – М.: Кнорус, 2009.
6. Теория статистики: Учебник / Под ред. Шмойловой Р.А. – М.: «Финансы и статистика», 2009.
7. *Лысенко С.Н.* Общая теория статистики: Учебное пособие / С.Н. Лысенко, И.А. Дмитриева. - М.: ИД ФОРУМ, НИЦ ИНФРА-М, 2013. - 208с.

ПРИЛОЖЕНИЕ

Приложение 1

Нормальный закон распределения

Значения функции $F(t)=P(\,|T|\leq t_{табл.})=\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx$

Целые и десятые доли t	Сотые доли t									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0,0797	0,0876	0,0955	0,1034	0,1113	0,1192	0,1271	0,1350	0,1428	0,1507
0,2	0,1585	0,1663	0,1741	0,1819	0,1897	0,1974	0,2051	0,2128	0,2205	0,2282
0,3	0,2358	0,2434	0,2510	0,2586	0,2661	0,2737	0,2812	0,2886	0,2961	0,3035
0,4	0,3108	0,3182	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,5	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,6	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,7	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,8	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,9	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6680	0,6729	0,6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	0,7287	0,7330	0,7373	0,7415	0,7457	0,7499	0,7540	0,7580	0,7620	0,7660
1,2	0,7699	0,7737	0,7775	0,7813	0,7850	0,7887	0,7923	0,7959	0,7995	0,8029
1,3	0,8064	0,8098	0,8132	0,8165	0,8198	0,8230	0,8262	0,8293	0,8324	0,8355
1,4	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,5	0,8664	0,8690	0,8715	0,8740	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,6	0,8904	0,8926	0,8948	0,8969	0,8990	0,9011	0,9031	0,9051	0,9070	0,9090
1,7	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,8	0,9281	0,9297	0,9312	0,9328	0,9342	0,9357	0,9371	0,9385	0,9399	0,9412
1,9	0,9426	0,9439	0,9451	0,9464	0,9476	0,9488	0,9500	0,9512	0,9523	0,9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9615	0,9625	0,9634
2,1	0,9643	0,9651	0,9660	0,9668	0,9676	0,9684	0,9692	0,9700	0,9707	0,9715
2,2	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,9780
2,3	0,9786	0,9791	0,9797	0,9802	0,9807	0,9812	0,9817	0,9822	0,9827	0,9832
2,4	0,9836	0,9840	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,5	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,6	0,9907	0,9909	0,9912	0,9915	0,9917	0,9920	0,9922	0,9924	0,9926	0,9929
2,7	0,9931	0,9933	0,9935	0,9937	0,9939	0,9940	0,9942	0,9944	0,9946	0,9947
2,8	0,9949	0,9950	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,9960	0,9961
2,9	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9980
3,1	0,9981	0,9981	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,2	0,9986	0,9987	0,9987	0,9988	0,9988	0,9988	0,9989	0,9989	0,9990	0,9990
3,3	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,4	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	0,9995
3,5	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997
3,6	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	0,9998	0,9998
3,7	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
4,0	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Учебное издание

Балаш Ольга Сергеевна
Чистопольская Елена Владимировна

СТАТИСТИКА
РАЗДЕЛ «ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ»

Учебное пособие

Пособие издано в авторской редакции