

СТАТИСТИЧЕСКОЕ ВВЕДЕНИЕ
В РАСПРЕДЕЛЕННЫЕ
ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ:
СТАТИСТИКА



◆ ИЗДАТЕЛЬСТВО ТГТУ ◆

СТАТИСТИЧЕСКОЕ ВВЕДЕНИЕ В РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ: СТАТИСТИКА

Контрольные задания и методические указания
по выполнению контрольной работы для студентов
специальностей 080105 «Финансы и кредит» и
080109 «Бухгалтерский учет, анализ и аудит»
очной и заочной форм обучения



УДК 657(075)
ББК У052я73-5
П18

Рекомендовано Редакционно-издательским советом университета

Рецензент

Доктор педагогических наук, профессор,
заведующий кафедрой ТиОКД ТГТУ
Н.В. Молоткова

Составитель

А.В. Пархоменко

П18 Статистическое введение в распределенные вычислительные системы: статистика : контрольные задания и метод. указания / сост. : А.В. Пархоменко. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. – 48 с. – 100 экз.

Приведены методические указания по выполнению контрольной работы, даны варианты заданий, требования по ее оформлению. Предназначены студентам специальностей 080105 «Финансы и кредит» и 080109 «Бухгалтерский учет, анализ и аудит» очной и заочной форм обучения.

УДК 657(075)
ББК У052я73-5

© ГОУ ВПО «Тамбовский государственный
технический университет» (ТГТУ), 2008

Учебное издание

СТАТИСТИЧЕСКОЕ ВВЕДЕНИЕ В РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ: СТАТИСТИКА

Контрольные задания и методические указания

Составитель
ПАРХОМЕНКО Алла Вячеславовна

Редактор О.М. Ярцева
Инженер по компьютерному макетированию Т.А. Сынкова

Подписано в печать 27.05.2008.
Формат 60 × 84 / 16. 2,79 усл. печ. л.
Тираж 100 экз. Заказ № 266

Издательско-полиграфический центр
Тамбовского государственного технического университета,
392000, Тамбов, Советская, 106, к. 14

ВВЕДЕНИЕ

В современном обществе статистика выполняет важную роль в механизме управления экономикой. Она осуществляет сбор, научную обработку, обобщение и анализ информации, характеризующей развитие экономики страны, культуры и уровня жизни населения.

Статистический анализ показателей, отражающих деятельность отдельных предприятий и организаций, служит исходной базой для рассмотрения всей совокупности социально-экономических процессов, протекающих на уровне макроэкономики, и является введением в распределенные вычислительные системы.

Предлагаемые задания и методические указания по выполнению практических заданий подготовлены в соответствии с программой курса «Статистика», разработанной в соответствии с требованиями государственного образовательного стандарта по специальностям 080105 «Финансы и кредит» и 080109 «Бухгалтерский учет, анализ и аудит».

Содержание заданий предполагает, что студенты освоили курс общей теории статистики, владеют методологией расчета таких обобщающих статистических показателей, как относительные и средние величины, структурные характеристики рядов распределения, методами измерения вариации, корреляции, построения уравнений регрессии, анализа динамики явлений, оценки структурных изменений и их влияния на динамику абсолютных и средних показателей.

В практическом задании рассматривается методология статистических расчетов, адаптированная к современной статистической практике. Представлены наиболее важные темы курса статистики: статистическое наблюдение, статистическая сводка и группировка, формы выражения статистических показателей, средние величины, показатели вариации и анализ частотных распределений, закономерности распределения, выборочное наблюдение, статистическое изучение взаимосвязи социально-экономических явлений, статистическое изучение динамики социально-экономических явлений, экономические индексы.

Каждая тема практического задания построена по единой схеме. Сначала приводятся краткие методические указания и пояснения к соответствующим разделам статистики, приводятся основные формулы расчета показателей. Далее, в соответствии со своими вариантами, предлагаются задания для самостоятельного выполнения, представляющие собой совокупность контрольных задач, выполнение которых дает возможность студенту последовательно работать над программой курса и обеспечить хорошее усвоение материала.

1. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

1.1. ПОРЯДОК ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ

В соответствии с учебным планом студенты 2 курса выполняют письменную контрольную работу по статистике, цель которой – изучить важнейшие методологические вопросы, проверить умение применять на практике основные положения курса, приобрести практические навыки в расчетах статистических показателей в области социально-экономической статистики, построении и оформлении статистических таблиц и графиков, научиться понимать экономический смысл исчисленных показателей, проанализировать их.

При подготовке к написанию контрольной работы студенту необходимо руководствоваться программой курса и методическими указаниями по отдельным темам.

Задания к контрольной работе составлены в тридцати вариантах. Каждый вариант состоит из двух теоретических вопросов и трех задач (табл. 1).

Таблица 1

№ варианта	№ теоретического вопроса	№ задач	№ варианта	№ теоретического вопроса	№ задач
1	28, 35	3, 18, 58	16	24, 31	29, 56, 60
2	11, 30	2, 69, 27	17	13, 27	32, 61, 20
3	5, 23	5, 47, 14	18	10, 28	33, 57, 62
4	19, 17	1, 12, 25	19	7, 22	35, 59, 24
5	10, 16	4, 11, 65	20	15, 29	36, 21, 64
6	13, 20	6, 13, 30	21	3, 14	37, 44, 66
7	22, 32	7, 15, 43	22	5, 16	38, 45, 67
8	6, 21	8, 21, 34	23	1, 12	39, 46, 68
9	18, 24	9, 23, 31	24	18, 26	40, 18, 70
10	1, 7	10, 28, 55	25	4, 11	48, 38, 63
11	2, 8	16, 38, 50	26	7, 33	49, 39, 15
12	3, 12	17, 35, 51	27	14, 25	15, 33, 69
13	4, 9	22, 41, 52	28	30, 34	5, 17, 54
14	14, 17	19, 42, 53	29	13, 35	10, 22, 43
15	4, 21	26, 43, 54	30	12, 26	7, 27, 70

Приступая к выполнению контрольной работы, необходимо ознакомиться с соответствующими разделами программы курса и методическими указаниями, изучить рекомендуемую литературу. Особое внимание следует обратить на методы построения, технику расчета и экономический смысл статистических показателей.

Далее следует предварительно наметить схему решения каждой задачи, составить макет статистической таблицы, куда будут занесены исчисленные показатели. Каждая таблица должна иметь заголовок, отражающий краткое содержание таблицы, заголовки по строкам и графам, а также указаны единицы измерения, итоговые показатели.

При выполнении контрольной работы необходимо руководствоваться следующими требованиями:

1. Записи должны быть выполнены аккуратно и разборчиво, допускаются только общепринятые сокращения.
2. В начале работы обязательно должен быть указан номер варианта работы.
3. Страницы должны быть пронумерованы, следует оставить поля для замечаний и внесения исправлений.
4. Реферат и задачи нужно выполнять в том порядке, в каком они даны в задании.
5. Перед каждым вопросом и задачей должно быть полностью приведено его условие. (Решение задачи от ее условия отделите некоторым интервалом.)
6. Решение задач следует сопровождать необходимыми формулами, развернутыми расчетами и краткими пояснениями. Если име-

ются несколько методов расчета того или иного показателя, следует применить наиболее простой из них, указав при этом и другие возможные способы решения.

В процессе решения задач нужно проверять производимые расчеты, пользуясь взаимосвязью между исчисляемыми показателями и обращая внимание на экономическое содержание последних. Задачи, к которым даны ответы без развернутых расчетов, пояснений и кратких выводов, будут считаться нерешенными.

Решение задач следует по возможности оформлять в виде таблиц.

В конце решения каждой задачи необходимо четко сформулировать выводы, раскрывающие экономическое содержание и значение исчисленных показателей.

Все расчеты относительных показателей нужно производить с принятой в статистике точностью до 0,001, а проценты – до 0,1.

Реферат должен быть изложен на 3–4 страницах в виде краткого описания всех рекомендуемых вопросов.

7. Изложение теоретического материала должно сопровождаться конкретными ссылками на соответствующие страницы литературного источника, таблицы исходных данных, иллюстрироваться числовыми примерами и т.п. Умение содержательно, лаконично и точно излагать материал выражается в широком использовании табличного, графического и других методов наглядного изображения. Содержание статистических методов, расчетные формулы, этапы расчетов, система взаимосвязи показателей и т.д. обязательны; целесообразно все вопросы освещать на конкретном числовом примере с последующей экономической и статистической интерпретацией полученных результатов; желательны использование статистических сборников и увязка рассматриваемых вопросов с действующей статистической отчетностью.

8. В конце контрольной работы следует привести список использованной литературы (автор, название учебника, издательство, год издания, страницы). Работа должна быть подписана студентом с указанием даты ее выполнения.

1.2. Методические указания по решению задач

Статистическая сводка и группировка

Сводка – это комплекс последовательных операций по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Различают сводку простую и сложную.

Простая сводка – это операция по подсчету общих итогов по совокупности единиц наблюдения.

Сложная сводка – это комплекс операций, включающих группировку единиц наблюдения, подсчет итогов по каждой группе и по всему объекту и представление результатов группировки и сводки в виде статистических таблиц.

Проведение сводки необходимо осуществлять по следующим этапам:

- выбор группировочного признака;
- определение порядка формирования групп;
- разработка системы статистических показателей для характеристики групп и объекта в целом;
- разработка макетов статистических таблиц для представления результатов сводки.

Группировкой называется расчленение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам. С помощью метода группировки решаются следующие задачи:

- выделение социально-экономических типов явлений;
- изучение структуры явления и структурных сдвигов, происходящих в нем;
- выявление связи и зависимости между явлениями.

Группировка, в которой группы образованы по одному признаку, называется *простой*.

Сложной называется группировка, в которой расчленение совокупности на группы производится по двум и более признакам.

Построение группировки начинается с определения состава группировочных признаков.

Группировочным признаком называется признак, по которому проводится разбиение единиц совокупности на отдельные группы.

В основание группировки могут быть положены как количественные, так и атрибутивные признаки. Первые имеют числовое выражение, а вторые отражают состояние единицы совокупности.

После того как определено основание группировки, следует решить вопрос о количестве групп, на которые надо разбить исследуемую совокупность.

Определение числа групп можно осуществить и математическим путем с использованием формулы Стерджесса:

$$n = 1 + 3,322 \lg N,$$

где n – число групп; N – число единиц совокупности.

Когда определено число групп, то следует определить интервалы группировки.

Интервал – это значение варьирующего признака, лежащее в определенных границах.

Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с *равными интервалами*.

Величина равного интервала определяется по следующей формуле:

$$h = \frac{R}{n},$$

где $R = X_{\max} - X_{\min}$, т.е. размах вариации; X_{\max} , X_{\min} – максимальное и минимальное значения признаков совокупности, n – число групп.

Формы выражения статистических показателей

Статистический показатель представляет собой количественную характеристику социально-экономических явлений и процессов в условиях качественной определенности. Качественная определенность показателя заключается в том, что он непосредственно связан с внутренним содержанием изучаемого явления или процесса, его сущностью.

Все используемые в статистической практике показатели по форме выражения классифицируются на абсолютные, относительные и средние.

Результаты статистического наблюдения регистрируются, прежде всего, в форме первичных *абсолютных величин*. Абсолютная величина отражает уровень развития явления.

В статистике все абсолютные величины являются именованными, измеряются в конкретных единицах (человеках, рублях, штуках, киловатт-часах, человеко-днях и человеко-часах и т.д.).

Относительная величина в статистике – это обобщающий показатель, который дает числовую меру соотношения двух сопоставляемых абсолютных величин.

Правильность расчета относительной величины зависит от сопоставимости сравниваемых показателей и наличия реальных связей между ними. По способу получения относительные показатели – всегда величины производные, определяемые в форме коэффициентов, процентов, промилле, продецимилле и т.п.

Относительные величины образуют систему взаимосвязанных статистических показателей. По содержанию выражаемых количественных соотношений наиболее часто выделяют следующие типы относительных величин:

- динамики;
- плана;
- реализации плана;
- структуры;
- координации;
- интенсивности и уровня экономического развития;
- сравнения.

Средняя величина – это обобщающий показатель, характеризующий типический уровень явления. Он определяется делением усредняемой величины на сумму единиц совокупности.

В средних величинах погашаются индивидуальные различия единиц совокупности, обусловленные случайными обстоятельствами.

Средние величины делятся на два больших класса:

- степенные средние;
- структурные средние.

В качестве структурных средних рассматриваются мода и медиана.

Степенные средние подразделяются на:

- среднюю арифметическую;
- среднюю гармоническую;
- среднюю геометрическую;
- среднюю квадратическую, кубическую и т.д.

Перечисленные средние объединяются в общей формуле (при различной величине k):

$$\bar{X} = \sqrt[k]{\frac{\sum X_i^k f_i}{\sum f_i}},$$

где \bar{X} – средняя величина исследуемого явления; X_i – i -й вариант осредняемого признака ($i = 1, n$); f_i – вес i -го варианта.

Наиболее распространенным видом средних величин является средняя арифметическая, которая, как и все средние, в зависимости от характера имеющихся данных может быть простой или взвешенной.

Средняя арифметическая простая (невзвешенная).

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum X_i}{n}.$$

При расчете средних величин отдельные значения осредняемого признака могут повторяться, встречаться по несколько раз. В подобных случаях расчет средней производится по сгруппированным данным или вариационным рядам, которые могут быть дискретными или интервальными.

Средняя арифметическая взвешенная определяется по формуле:

$$\bar{X} = \frac{\sum X_i f_i}{\sum f_i}.$$

Средняя гармоническая взвешенная имеет более сложную конструкцию, чем средняя арифметическая. Среднюю гармоническую применяют для расчетов тогда, когда в качестве весов используются не единицы совокупности – носители признака, а произведения этих единиц на значения признака (т.е. $w = Xf$).

$$\bar{X} = \frac{\sum W_i}{\sum \frac{W_i}{X_i}}.$$

Показатели вариации и анализ частотных распределений

Колеблемость, многообразие, изменяемость величины признака у единиц совокупности называются *вариацией*. Вариация существует в пространстве и во времени.

Показатели вариации делятся на две группы: абсолютные и относительные.

К *абсолютным* относятся размах вариации, среднее линейное отклонение, дисперсия и среднее квадратическое отклонение. Вторая группа показателей вычисляется как отношение абсолютных показателей вариации к средней арифметической (или медиане). *Относительными* показателями вариации являются коэффициенты осцилляции, вариации, относительное линейное отклонение и др.

Самым простым абсолютным показателем является *размах вариации* R .

Размах показывает, насколько велико различие между единицами совокупности, имеющими самое маленькое и самое большое значение признака.

$$R = X_{\max} - X_{\min}.$$

Среднее линейное отклонение (d) представляет собой среднюю величину из отклонений вариантов признака от их средней. Оно вычисляется как средняя арифметическая из абсолютных значений отклонений вариантов x_i и x (взвешенная или простая в зависимости от исходных условий) по следующим формулам:

$$\bar{d} = \frac{\sum |X_i - \bar{X}|}{n} \quad (\text{простая}); \quad \bar{d} = \frac{\sum |X_i - \bar{X}| f_i}{\sum f_i} \quad (\text{взвешенная}).$$

Дисперсия представляет собой средний квадрат отклонений индивидуальных значений признака от их средней величины и вычисляется по формулам простой и взвешенной дисперсий (в зависимости от исходных данных):

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{n} \quad (\text{простая дисперсия});$$

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2 f_i}{\sum f_i} \quad (\text{взвешенная дисперсия}).$$

Корень квадратный из дисперсии называется *средним квадратическим отклонением* (σ).

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} \quad \text{или} \quad \sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2 f_i}{\sum f_i}}.$$

Среднее квадратическое отклонение – это обобщающая характеристика размеров вариации признака в совокупности. Оно выражается в тех же единицах измерения, что и признак (в метрах, тоннах, рублях, процентах и т.д.).

Для целей сравнения колеблемости различных признаков в одной и той же совокупности или же при сравнении колеблемости одного и того же признака в нескольких совокупностях используются показатели вариации, приведенные в *относительных величинах*. Базой для сравнения служит средняя арифметическая. Эти показатели вычисляются как отношение размаха вариации, среднего линейного отклонения или среднего квадратического отклонения к средней арифметической или медиане. Чаще всего они выражаются в процентах и определяют не только сравнительную оценку вариации, но и дают характеристику однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33 % (для распределений, близких к нормальному). Различают следующие относительные показатели вариации (V).

Коэффициент осцилляции (V_R):

$$V_R = \frac{R}{\bar{x}} \cdot 100 \ %.$$

Линейный коэффициент вариации ($V_{\bar{d}}$):

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100 \ % \quad \text{или} \quad V_{\bar{d}} = \frac{\bar{d}}{M_e} \cdot 100 \ %.$$

Коэффициент вариации (V_{σ}):

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100 \ %.$$

Правило сложения дисперсий. Если данные представлены в виде аналитической группировки, то можно вычислить дисперсию общую, межгрупповую и внутригрупповую.

Общая дисперсия σ^2 измеряет вариацию признака во всей совокупности под влиянием всех факторов, обусловивших эту вариацию:

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

Межгрупповая дисперсия (δ_x^2) характеризует систематическую вариацию, т.е. различия в величине изучаемого признака, возникающие под влиянием признака-фактора, положенного в основание группировки.

$$\delta_x^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 n_i}{\sum n_i}.$$

где \bar{x}_i и n_i – соответственно групповые средние и численности по отдельным группам.

Внутригрупповая дисперсия (y_i^2) отражает случайную вариацию, т.е. часть вариации, происходящую под влиянием неучтенных факторов и не зависящую от признака-фактора, положенного в основание группировки.

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

Средняя из внутригрупповых дисперсий ($\bar{\sigma}_i^2$):

$$\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i}.$$

Существует закон, связывающий три вида дисперсии. Общая дисперсия равна сумме средней из внутригрупповых и межгрупповой дисперсий:

$$\sigma^2 = \bar{\sigma}_i^2 + \delta_x^2.$$

Данное соотношение называют правилом сложения дисперсий.

В статистическом анализе широко используется показатель, представляющий собой долю межгрупповой дисперсии в общей дисперсии. Он носит название *эмпирического коэффициента детерминации* (η^2):

$$\eta^2 = \frac{\delta_x^2}{\sigma^2}.$$

Этот коэффициент показывает долю общей вариации изучаемого признака, обусловленную вариацией группировочного признака.

Корень квадратный из эмпирического коэффициента детерминации носит название *эмпирического корреляционного отношения* (η):

$$\eta = \sqrt{\frac{\delta_x^2}{\sigma^2}}.$$

Оно характеризует влияние признака, положенного в основание группировки, на вариацию результативного признака. Эмпирическое корреляционное отношение изменяется в пределах от 0 до 1.

Наряду с вариацией индивидуальных значений признака вокруг средней может наблюдаться и вариация индивидуальных долей признака вокруг средней доли.

Внутригрупповая дисперсия доли определяется по формуле

$$\sigma_{p_i}^2 = p_i(1 - p_i).$$

Средняя из внутригрупповых дисперсий:

$$\bar{\sigma}_{p_i}^2 = \frac{\sum p_i(1 - p_i)n_i}{\sum n_i}.$$

Формула *межгрупповой дисперсии* имеет следующий вид:

$$\delta_{p_i}^2 = \frac{\sum (p_i - \bar{p})^2 n_i}{\sum n_i},$$

где n_i – численность единиц в отдельных группах; p – доля изучаемого признака во всей совокупности, которая определяется по формуле

$$\bar{p} = \frac{\sum p_i n_i}{\sum n_i}.$$

Общая дисперсия определяется по формуле

$$\sigma_p^2 = \bar{p}(1 - \bar{p}).$$

Три вида дисперсии связаны между собой следующим образом:

$$\sigma_p^2 = \bar{\sigma}_{p_i}^2 + \delta_{p_i}^2.$$

Это соотношение дисперсий называется теоремой сложения дисперсии доли признака.

Мода (M_o) – значение признака, наиболее часто встречающееся в исследуемой совокупности.

Медианой (Me) называется значение признака, приходящееся на середину ранжированной (упорядоченной) совокупности.

Свойство медианы заключается в том, что сумма абсолютных отклонений значений признака от медианы меньше, чем от любой другой величины:

$$\sum |x_i - Me| = \min.$$

Для определения медианы необходимо провести ранжирование. Если ранжированный ряд включает четное число единиц, то медиана определяется как средняя из двух центральных значений.

Для определения медианного значения признака по следующей формуле находят номер медианной единицы ряда (N_{Me}):

$$N_{Me} = \frac{n+1}{2},$$

где n – объем совокупности.

В отличие от дискретных вариационных рядов определение моды и медианы по интервальным рядам требует проведения определенных расчетов по следующим формулам:

$$M_o = x_0 + i \frac{(f_{M_o} - f_{M_o-1})}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

где x_0 – нижняя граница модального интервала (*модальным* называется интервал, имеющий наибольшую частоту); i – величина модального интервала; f_{M_o} – частота модального интервала; f_{M_o-1} – частота интервала, предшествующего модальному; f_{M_o+1} – частота интервала, следующего за модальным.

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

где x_0 – нижняя граница медианного интервала (*медианным* называется первый интервал, накопленная частота которого превышает половину общей суммы частот); i – величина медианного интервала; S_{Me-1} – накопленная частота интервала, предшествующего медианному; f_{Me} – частота медианного интервала.

Моду и медиану в интервальном ряду можно определить графически. Мода определяется по гистограмме распределения. Для этого выбирается самый высокий прямоугольник, который является в данном случае модальным. Затем правую вершину модального прямоугольника соединяем с правым верхним углом предыдущего прямоугольника, а левую вершину модального прямоугольника – с левым верхним углом последующего прямоугольника. Далее из точки их пересечения опускают перпендикуляр на ось абсцисс.

Абсцисса точки пересечения этих прямых и будет модой распределения. Медиана рассчитывается по кумуляте. Для ее определения из точки на шкале накопленных частот (частостей), соответствующей 50 %, проводится прямая, параллельная оси абсцисс, до пересечения с кумулятой. Затем из точки пересечения указанной прямой с кумулятой опускается перпендикуляр на ось абсцисс. Абсцисса точки пересечения является медианой.

Аналогично с нахождением медианы в вариационных рядах распределения можно отыскать квартили.

Квартили представляют собой значения признака, делящие ранжированную совокупность на четыре равновеликие части. Различают квартиль нижний (Q_1), отделяющий j часть совокупности с наименьшими значениями признака, и квартиль верхний (Q_3), отсекающий j часть с наибольшими значениями признака. Это означает, что 25 % единиц совокупности будут меньше по величине Q_1 ; 25 % единиц будут заключены между Q_1 и Q_2 ; 25 % – между Q_2 и Q_3 и остальные 25 % превосходят Q_3 . Средним квартилем Q_2 является медиана.

Для расчета квартилей по интервальному вариационному ряду используются формулы:

$$Q_1 = X_{Q_1} + i \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}}; \quad Q_3 = X_{Q_3} + i \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}},$$

где X_{Q_1} – нижняя граница интервала, содержащего нижний квартиль (интервал определяется по накопленной частоте, первой превышающей 25 %); X_{Q_3} – нижняя граница интервала, содержащего верхний квартиль (интервал определяется по накопленной частоте, первой превышающей 75 %); i – величина интервала; S_{Q_1-1} – накопленная частота интервала, предшествующего интервалу, содержащему нижний квартиль; S_{Q_3-1} – то же для верхнего квартиля; f_{Q_1} – частота интервала, содержащего нижний квартиль; f_{Q_3} – то же для верхнего квартиля.

Понятие о закономерности распределения

В статистике существует определенная зависимость между изменением значений варьирующего признака и частот. Частоты в этих рядах с увеличением значения варьирующего признака первоначально увеличиваются, а затем после достижения какой-то максимальной величины в середине ряда уменьшаются. Это свидетельствует о том, что частоты в вариационных рядах изменяются закономерно в связи с изменением варьирующего признака. Такие закономерности изменения частот в вариационных рядах называются *закономерностями распределения*.

Имея дело с эмпирическим распределением, можно предположить, что данному распределению соответствует определенная, характерная для него теоретическая кривая. Выдвинув гипотезу о той или иной форме распределения, стремятся описать эмпирический ряд с помощью математической модели, выражающей некоторый теоретический закон распределения. Среди различных кривых распределения особое место занимает *нормальное распределение*.

Нормальное распределение выражается следующей стандартизированной кривой нормального распределения:

$$y_t = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2},$$

где y_t – ордината кривой нормального распределения; $t = \frac{x - \bar{x}}{\sigma}$ – стандартизованное отклонение; e и π – математические постоянные; x – варианты вариационного ряда; \bar{x} – их средняя величина; y – среднее квадратическое отклонение.

Выяснение общего характера распределения предполагает оценку степени его однородности, оценку его симметричности, остро- или плосковершинности.

Для однородных совокупностей характерны одновершинные распределения. Многовершинность свидетельствует о неоднородности изучаемой совокупности. Появление двух и более вершин делает необходимой перегруппировку данных с целью выделения более однородных групп.

Симметричным называется распределение, в котором частоты любых двух вариантов, равноотстоящих в обе стороны от центра распределения, равны между собой. Рассчитанные для таких распределений средняя, мода и медиана также равны.

В статистике для характеристики асимметрии используются следующие показатели:

- относительный показатель асимметрии (A_s):

$$A_s = \frac{\bar{x} - Mo}{\sigma} \quad \text{или} \quad A_s = \frac{\bar{x} - Me}{\sigma}.$$

Его величина может быть положительной и отрицательной. В первом случае речь идет о правосторонней асимметрии, а во втором – о левосторонней.

При правосторонней асимметрии $Mo > Me > x$.

Коэффициент асимметрии может изменяться от -3 до $+3$;

- показатель эксцесса (Ek), определяемый по формуле с использованием центрального момента четвертого порядка
- $$\left(\mu_4 = \frac{\sum (x_i - \bar{x})^4}{n} \right)$$

$$E_k = \frac{\mu_4}{\sigma^4} - 3.$$

При симметричном распределении $Ek = 0$. Если $Ek > 0$, распределение является островершинным; если $Ek < 0$ – плосковершинным.

Количественная характеристика соответствия может быть получена с помощью особых статистических показателей критериев согласия.

Для выполнения курсовой работы целесообразно использовать критерий согласия К. Пирсона (χ^2).

$$\chi^2 = \sum \frac{(f_3 - f_1)^2}{f_1},$$

где f_3 и f_1 – эмпирические и теоретические частоты, соответственно.

$$f_1 = \frac{h}{\sigma} \sum f \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

С помощью величины χ^2 и числа степеней свободы $\gamma = n - 1$ по специальным таблицам определяется вероятность P (χ^2). На основе P выносится суждение о существенности или несущественности расхождения между эмпирическим и теоретическим распределениями. При $P > 0,5$ считается, что эмпирическое и теоретическое распределения близки, при $P \in [0,2; 0,5]$ совпадение между ними удовлетворительное, в остальных случаях – недостаточное.

Если число степеней свободы большое, то применяется соотношение, равное $\sqrt{2\chi^2} - \sqrt{2\gamma - 1}$. Расхождение между эмпирическим и теоретическим распределениями существенно при значениях этой разности, превосходящих 2.

Выборочное наблюдение

Выборочное наблюдение – это такое несплошное наблюдение, при котором отбор подлежащих обследованию единиц осуществляется в случайном порядке, отобранная часть изучается, а результаты распространяются на всю исходную совокупность. Наблюдение организуется таким образом, что эта часть отобранных единиц в уменьшенном масштабе *репрезентирует* (представляет) всю совокупность.

Преимущество выборочного наблюдения по сравнению со сплошным можно реализовать, если оно организовано и проведено в строгом соответствии с научными принципами *теории выборочного метода*. Такими принципами являются:

- обеспечение случайности отбора единиц;
- достаточного их числа.

По виду различают индивидуальный, групповой и комбинированный отбор.

По методу отбора различают повторную и бесповторную выборки.

Способ отбора определяет конкретный механизм или процедуру выборки единиц из генеральной совокупности.

По степени охвата единиц совокупности различают большие и малые ($n < 30$) выборки.

В практике выборочных исследований наибольшее распространение получили следующие виды выборки: *собственно-случайная, механическая, типическая, серийная, комбинированная*.

Применяя выборочный метод в статистике, обычно используют два основных вида обобщающих показателей: *среднюю величину количественного признака и относительную величину альтернативного признака.*

Выборочная доля (w), или частота, определяется отношением числа единиц, обладающих изучаемым признаком m , к общему числу единиц выборочной совокупности n .

Для характеристики надежности выборочных показателей различают *среднюю и предельную ошибки выборки.*

Ошибка выборочного наблюдения – это разность между величиной параметра в генеральной совокупности и его величиной, вычисленной по результатам выборочного наблюдения:

- для среднего значения ошибка будет определяться так:

$$\Delta_{\bar{x}} = |\bar{x} - \tilde{x}|;$$

- для доли (альтернативного признака)

$$\Delta_w = |w - p|.$$

Предельная ошибка выборки ($\Delta_{\tilde{x}}$) и средняя ошибка ($\mu_{\tilde{x}}$) связаны следующим соотношением:

$$\Delta_{\tilde{x}} = t \mu_{\tilde{x}}.$$

Расчет средней и предельной ошибок выборки позволяет определить возможные пределы, в которых будут находиться характеристики генеральной совокупности

- для средней $\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}$ или $\bar{x} = \tilde{x} \pm \Delta_{\tilde{x}}$;
- для доли $w - \Delta_w \leq p \leq w + \Delta_w$ или $p = w \pm \Delta_w$.

Это означает, что с заданной вероятностью можно утверждать, что значение генеральной средней следует ожидать в пределах от $\tilde{x} - \Delta_{\tilde{x}}$ до $\tilde{x} + \Delta_{\tilde{x}}$.

Аналогичным образом может быть записан доверительный интервал генеральной доли: $w - \Delta_w$ до $w + \Delta_w$.

Статистическое изучение взаимосвязи социально-экономических явлений

Наиболее разработанной в теории статистики является методология так называемой парной корреляции, рассматривающая влияние вариации факторного признака X на результативный Y . Овладение теорией и практикой парной корреляции представляет исходный этап познания других приемов и методов изучения корреляционной связи.

В основу выявления и установления аналитической формы связи положено применение в анализе исходной информации математических функций. При изучении связи показателей применяются различного вида уравнения прямолинейной и криволинейной связи:

а) линейная регрессия, выражается уравнением прямой вида: $\bar{Y}_x = a_0 + a_1 x$;

б) нелинейная регрессия, выражается уравнениями вида:

парабола – $\bar{Y}_x = a_0 + a_1 x + a_2 x^2$;

гипербола – $\bar{Y}_x = a_0 + \frac{a_1}{x}$ и т.д.,

где \bar{Y}_x – теоретические значения результативного признака, полученные по уравнению регрессии; a_0, a_1 – коэффициенты (параметры) уравнения регрессии.

В уравнениях регрессии параметр a_0 показывает усредненное влияние на результативный признак неучтенных (не выделенных для исследования) факторов; параметр a_1 (а в уравнении параболы и a_2) – коэффициент регрессии показывает, насколько изменяется в среднем значение результативного признака при увеличении факторного на единицу собственного измерения.

Параметры уравнения a_0, a_1 находят методом наименьших квадратов (метод решения систем уравнений, при котором в качестве решения принимается точка минимума суммы квадратов отклонений), т.е. в основу этого метода положено требование минимальности сумм квадратов отклонений эмпирических данных y_i от выровненных \bar{y} :

$$S = \sum (Y_i - \bar{Y}_x)^2 \rightarrow \min.$$

Для прямой зависимости:

$$S = \sum (y_i - a_0 - a_1 x_i)^2 \rightarrow \min.$$

Для нахождения минимума данной функции приравняем к нулю ее частные производные и получим систему двух линейных уравнений, которая называется *системой нормальных уравнений*:

$$\begin{cases} na_0 + a_1 \sum x = \sum y; \\ a_0 \sum x + a_1 \sum x^2 = \sum xy, \end{cases}$$

где n – объем исследуемой совокупности (число единиц наблюдений).

Значимость коэффициентов простой линейной регрессии (применительно к совокупностям, у которых $n < 30$) осуществляют с помощью *t-критерия Стьюдента*. При этом вычисляют расчетные (фактические) значения *t-критерия*:

- для параметра a_0 $t_{a_0} = |a_0| \frac{\sqrt{n-2}}{\sigma_{\text{ост}}}$;
- для параметра a_1 $t_{a_1} = |a_1| \frac{\sqrt{n-2}}{\sigma_{\text{ост}}} \sigma_x$,

где n – объем выборки; $\sigma_{\text{ост}} = \sqrt{\sum (y - \bar{y})^2 / n}$ – среднее квадратическое отклонение результативного признака y от выровненных значений \bar{y} ; $\sigma_x = \sqrt{\sum (x - \bar{x})^2 / n}$ или $\sigma_x = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$ – среднее квадратическое отклонение факторного признака x от общей средней \bar{x} .

Вычисленные по формулам значения сравнивают с критическими t , которые определяют по таблице Стьюдента с учетом принятого уровня значимости α и числом степеней свободы вариации $\gamma = n - 2$.

Параметр признается значимым при условии, если $t_{\text{расч}} > t_{\text{табл}}$.

Проверка адекватности регрессионной модели может быть дополнена корреляционным анализом. Для этого необходимо определить *тесноту* корреляционной связи между переменными x и y . Теснота корреляционной связи, как и любой другой, может быть измерена эмпирическим корреляционным отношением η , когда δ^2 (межгрупповая дисперсия) характеризует отклонения групповых средних результативного признака от общей средней: $\eta = \sqrt{\frac{\delta^2}{\sigma^2}}$.

Теоретическое корреляционное отношение η представляет собой относительную величину, получающуюся в результате сравнения среднего квадратического отклонения выровненных значений результативного признака \bar{y}_x , т.е. рассчитанных по уравнению регрессии, со средним квадратическим отклонением эмпирических (фактических) значений результативного признака y :

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}} = \sqrt{\frac{\sigma^2 - \sigma_{\text{ост}}^2}{\sigma^2}} = \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma^2}} \quad \text{или} \quad \eta = \sqrt{1 - \frac{\sum (y - \bar{y}_x)^2}{\sum (y - \bar{y})^2}}.$$

Теоретическое корреляционное отношение применяется для измерения тесноты связи при линейной и криволинейной зависимостях между результативным и факторным признаком. При криволинейных связях теоретическое корреляционное отношение называют *индексом корреляции R* . Корреляционное отношение может находиться в пределах от 0 до 1, т.е. ($0 < \eta < 1$).

Кроме того, при линейной форме уравнения применяется другой показатель тесноты связи – *линейный коэффициент корреляции*:

$$r_{xy} = \frac{\bar{y}_x - \bar{y} \bar{x}}{\sigma_x \sigma_y} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n \sigma_x \sigma_y} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}},$$

где n – число наблюдений.

Для практических вычислений при малом числе наблюдений, $n < (20 - 30)$, линейный коэффициент корреляции удобнее определить по следующей формуле:

$$r = \frac{\sum xy - \sum x \frac{\sum y}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \left[\sum y^2 - \frac{(\sum y)^2}{n} \right]}}.$$

Он принимает значения в интервале: $-1 < r < +1$.

Отрицательные значения указывают на обратную связь, положительные – на прямую.

Для оценки *значимости коэффициента корреляции r* используют t -критерий Стьюдента, который применяется при t -распределении, отличном от нормального.

$$t_r = r \sqrt{\frac{n-2}{1-r^2}},$$

где $(n - 2)$ – число степеней свободы при заданном уровне значимости α и объеме выборки n .

После проверки адекватности, установления точности и надежности построенной модели ее необходимо проанализировать. Прежде всего, нужно проверить, согласуются ли знаки параметров с теоретическими представлениями о направлении влияния признака-фактора на результативный признак (показатель).

Для удобства интерпретации параметра a_1 используют *коэффициент эластичности*. Он показывает средние изменения результативного признака при изменении факторного признака на 1 % и вычисляется по формуле, %:

$$\varepsilon_{x_i} = a_1 \frac{\bar{x}_i}{\bar{y}}.$$

Проверка адекватности всей модели осуществляется с помощью расчета F -критерия и величины средней ошибки аппроксимации ($\bar{\varepsilon}$).

Значение средней ошибки аппроксимации ($\bar{\varepsilon}$), определяемой по формуле

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \bar{y}_{1,2,\dots,k}|}{y_{1,2,\dots,k}} \cdot 100,$$

не должно превышать 12 – 15 %.

Проверку значимости уравнения регрессии производят на основе вычисления F -критерия Фишера:

$$F = \frac{\sigma_y^2}{\sigma_{\text{ост}}^2} \frac{n-m}{m-1},$$

где m – число параметров в уравнении регрессии.

Полученное значение критерия $F_{\text{расч}}$ сравнивают с критическим (табличным) для принятого уровня значимости 0,05 или 0,01 и чисел степеней свободы $\nu_1 = m - 1$ и $\nu_2 = n - m$. Если оно окажется больше соответствующего табличного значения, то данное уравнение регрессии статистически значимо, т.е. доля вариации, обусловленная регрессией, намного превышает случайную ошибку.

Считается, что уравнение регрессии пригодно для практического использования в том случае, если $F_{\text{расч}} > F_{\text{табл}}$ не менее чем в 4 раза.

Множественный коэффициент корреляции вычисляется при наличии линейной связи между результативным и несколькими факторными признаками, а также между каждой парой факторных признаков:

$$R_{y/x_1x_2\dots x_k} = \sqrt{\frac{\delta^2}{\sigma^2}} = \sqrt{1 - \frac{\sigma_{\text{ост}}^2}{\sigma^2}}.$$

В случае оценки тесноты связи между Y и X_1, X_2 множественный коэффициент корреляции определяется по следующей формуле:

$$R_{y/x_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}}.$$

Проверка значимости коэффициента множественной корреляции осуществляется на основе F -критерия Фишера

$$F_p = \frac{\frac{1}{2} R_{y/x_1x_2}^2}{\frac{1}{n-3} (1 - R_{y/x_1x_2}^2)}.$$

Гипотеза о незначимости коэффициента множественной корреляции отвергается, если

$$F_p > F_{kp} (\alpha; \nu_1 = 2; \nu_2 = n - 3).$$

Статистическое изучение динамики социально-экономических явлений

Процесс развития, движения социально-экономических явлений во времени в статистике принято называть *динамикой*. Для отображения динамики строят *ряды динамики* (хронологические, временные), которые представляют собой ряды изменяющихся во времени значений статистического показателя, расположенных в хронологическом порядке.

Уровни ряда обычно обозначаются через «у», моменты или периоды времени, к которым относятся уровни, – через «t».

Показатели анализа ряда динамики. Анализ интенсивности изменения во времени осуществляется с помощью показателей, получаемых в результате сравнения уровней. К таким показателям относятся: *абсолютный прирост, темп роста, темп прироста, абсолютное значение одного процента прироста.*

Система средних показателей включает *средний уровень ряда, средний абсолютный прирост, средний темп роста, средний темп прироста.*

Важнейшим статистическим показателем анализа динамики является *абсолютный прирост (сокращение)*, т.е. *абсолютное изменение*, характеризующее увеличение или уменьшение уровня ряда за определенный промежуток времени. Абсолютный прирост с переменной базой называют *скоростью роста*.

Абсолютный прирост (цепной): *Абсолютный прирост (базисный):*

$$\Delta y^{\text{ц}} = y_i - y_{i-1} \qquad \Delta y^{\text{б}} = y_i - y_0,$$

где y_i – уровень сравниваемого периода; y_{i-1} – уровень предшествующего периода; y_0 – уровень базисного периода.

Для оценки интенсивности, т.е. относительного изменения уровня динамического ряда за какой-либо период времени исчисляют *темпы роста (снижения)*.

Темп роста (цепной): *Темп роста (базисный):*

$$T_p^{\text{ц}} = \frac{y_i}{y_{i-1}} \cdot 100 \qquad \text{или} \qquad T_p^{\text{б}} = \frac{y_i}{y_0} \cdot 100$$

$T_p = K_p \cdot 100.$

Темп прироста (сокращения) показывает, на сколько процентов сравниваемый уровень больше или меньше уровня, принятого за базу сравнения, и вычисляется как отношение абсолютного прироста к абсолютному уровню, принятому за базу сравнения.

Темп прироста может быть положительным, отрицательным или равным нулю, выражается он в процентах и долях единицы (коэффициенты прироста).

Темп прироста (цепной): *Темп прироста (базисный):*

$$T_{\text{пр}}^{\text{н}} = \frac{\Delta y_{\text{н}}}{y_{i-1}} \cdot 100 \qquad T_{\text{пр}}^{\text{б}} = \frac{\Delta y_{\text{б}}}{y_0} \cdot 100.$$

Темп прироста (сокращения) можно получить и из темпа роста, выраженного в процентах, если из него вычесть 100 %. Коэффициент прироста получается вычитанием единицы из коэффициента роста:

$$T_{\text{пр}} = T_{\text{р}} - 100; \quad K_{\text{пр}} = K_{\text{р}} - 1.$$

Сравнение абсолютного прироста и темпа прироста за одни и те же периоды времени показывает, что при снижении (замедлении) темпов прироста абсолютный прирост не всегда уменьшается, иногда он может возрасти. Поэтому, чтобы правильно оценить значение полученного темпа прироста, его рассматривают в сопоставлении с показателем абсолютного прироста. Результат выражают показателем, который называют *абсолютным значением (содержанием) одного процента прироста* и рассчитывают как отношение абсолютного прироста к темпу прироста за тот же период времени, %:

$$|\%| = \frac{\Delta y_{\text{н}}}{T_{\text{пр}}^{\text{н}}} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100} = \frac{y_{i-1}}{100} = 0,01 y_{i-1}.$$

Абсолютное значение одного процента прироста равно сотой части предыдущего (или базисного) уровня. Оно показывает, какое абсолютное значение скрывается за относительным показателем – одним процентом прироста.

Методы анализа основной тенденции развития в рядах динамики

Основной тенденцией развития (трендом) называется плавное и устойчивое изменение уровня явления во времени, свободное от случайных колебаний.

С этой целью ряды динамики подвергаются обработке методами укрупнения интервалов, скользящей средней и аналитического выравнивания.

Основным содержанием *метода аналитического выравнивания* в рядах динамики является то, что общая тенденция развития рассчитывается как функция времени:

$$\bar{y}_t = f(t),$$

где \bar{y}_t – уровни динамического ряда, вычисленные по соответствующему аналитическому уравнению на момент времени t .

Определение теоретических (расчетных) уровней \bar{y}_t производится на основе адекватной математической модели, которая отображает (аппроксимирует) основную тенденцию ряда динамики.

Простейшими моделями (формулами), выражающими тенденцию развития, являются:

- *линейная функция* – прямая $\bar{y}_t = a_0 + a_1 t$,

где a_0 и a_1 – параметры уравнения; t – время;

- *показательная функция* – $\bar{y}_t = a_0 a_1^t$;
- *степенная функция* – кривая второго порядка (парабола)

$$\bar{y}_t = a_0 + a_1 t + a_2 t^2.$$

Расчет параметров функции производится *методом наименьших квадратов*, в котором в качестве решения принимается точка минимума суммы квадратов отклонений между теоретическими и эмпирическими уровнями:

$$S = \sum (y_i - \bar{y}_t)^2 \rightarrow \min,$$

где \bar{y}_t – выравненные (расчетные) уровни; y_i – фактические уровни.

Параметры уравнения a_i , удовлетворяющие этому условию, могут быть найдены решением *системы нормальных уравнений*.

$$\begin{aligned} n a_0 + a_1 \sum t &= \sum y; \\ a_0 \sum t + a_1 \sum t^2 &= \sum yt, \end{aligned}$$

где y – фактические (эмпирические) уровни ряда; t – время (порядковый номер периода или момента времени).

Расчет параметров упрощается, если за начало отсчета времени ($t = 0$) принять центральный интервал (момент).

При четном числе уровней (например, 4) значения t – условного обозначения времени будут выглядеть следующим образом:

$$-3 \qquad -1 \qquad +1 \qquad +3$$

При нечетном числе уровней (например, 5) значения устанавливаются по-другому:

$$-2 \qquad -1 \qquad 0 \qquad +1 \qquad +2$$

Тогда система нормальных уравнений принимает вид:

$$\begin{cases} a_0 n = \sum y; \\ a_1 \sum t^2 = \sum ty. \end{cases}$$

Из первого уравнения $a_0 = \frac{\sum y}{n}$.

Из второго уравнения $a_1 = \frac{\sum yt}{\sum t^2}$.

Основная тенденция (тренд) показывает, как воздействуют систематические факторы на уровень ряда динамики, а колеблемость уровней около тренда служит мерой воздействия остаточных факторов. Ее можно измерить по формуле:

$$\sigma_t = \sqrt{\sum (y - \bar{y}_t)^2 / n}.$$

Относительной мерой колеблемости является коэффициент вариации

$$V = \frac{\sigma_t}{\bar{y}}.$$

Экономические индексы

В статистике под *индексом* понимается относительный показатель, который выражает соотношение величин какого-либо явления во времени, в пространстве или сравнение фактических данных с любым эталоном.

Индивидуальные индексы получают в результате сравнения однотоварных явлений.

В зависимости от экономического назначения индивидуальные индексы бывают: физического объема продукции, себестоимости, цен, трудоемкости и т.д.

Индекс физического объема продукции i_q рассчитывается по формуле

$$i_q = \frac{q_1}{q_0}.$$

Этот индекс показывает, во сколько раз возрос (уменьшился) выпуск какого-либо одного товара в отчетном периоде по сравнению с базисным, или сколько процентов составляет рост (снижение) выпуска товара. Если из значения индекса, выраженного в процентах, вычесть 100 %, то полученная величина покажет, на сколько процентов возрос (уменьшился) выпуск продукции.

Индексы других показателей строятся аналогично.

Агрегатный индекс – сложный относительный показатель, который характеризует среднее изменение социально-экономического явления, состоящего из несоизмеримых элементов.

Числитель и знаменатель агрегатного индекса представляют собой сумму произведений двух величин, одна из которых меняется (индексируемая величина), а другая остается неизменной в числителе и знаменателе (вес индекса).

Методика построения агрегатного индекса предусматривает решение трех вопросов:

- 1) какая величина будет индексируемой;
- 2) по какому составу разнородных элементов явления необходимо исчислить индекс;
- 3) что будет служить весом при расчете индекса.

Индекс стоимости продукции, или товарооборота (I_{pq}), представляет собой отношение стоимости продукции текущего периода ($Y_{p_1q_1}$) к стоимости продукции в базисном периоде ($Y_{p_0q_0}$) и определяется по формуле

$$I_{pq} = \frac{\sum p_1q_1}{\sum p_0q_0}.$$

Такой индекс показывает, во сколько раз возросла (уменьшилась) стоимость продукции (товарооборота) отчетного периода по сравнению с базисным, или сколько процентов составляет рост (снижение) стоимости продукции. Если из значения индекса стоимости вычесть 100 % ($I_{pq} - 100$), то разность покажет, на сколько процентов возросла (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным. Разность числителя и знаменателя показывает, на сколько рублей увеличилась (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным.

Аналогично строятся индексы для других показателей, которые являются произведением двух множителей.

Значение индекса стоимости продукции (товарооборота) зависит от двух факторов: изменения количества продукции и цен, что обуславливает возможность построения еще двух индексов: физического объема продукции и цен.

Индекс физического объема продукции – это индекс количественного показателя. В этом индексе индексируемой величиной будет количество продукции в натуральном выражении, а весом – цена. Только умножив несоизмеримые между собой количества разнородной продукции на их цены, можно перейти к стоимостям продукции, которые будут уже величинами соизмеримыми. Так как индекс физического объема – индекс количественного показателя, то весами будут цены базисного периода.

$$I_q = \frac{\sum q_1p_0}{\sum q_0p_0}.$$

Индекс цен – это индекс качественного показателя. Индексируемой величиной будет цена товара, так как этот индекс характеризует изменение цен. Весом будет выступать количество произведенных товаров.

$$I_p = \frac{\sum p_1q_1}{\sum p_0q_1},$$

где в числителе дроби – фактическая стоимость продукции текущего периода, а в знаменателе – условная стоимость тех же товаров в ценах базисного периода.

Стоимость продукции можно представить как произведение количества товара на его цену. Точно такая же связь существует и между индексами стоимости, физического объема и цен, т.е.

$$I_{pq} = I_p I_q \quad \text{или} \quad \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \frac{\sum q_1 p_0}{\sum q_0 p_0}.$$

Разность числителя и знаменателя каждого индекса-сомножителя выражает размер изменения общей абсолютной величины под влиянием изменения одного фактора.

$$(Yq1p0 - Yq0p0) + (Yq1p1 - Yq1p0) = Yq1p1 - Yq0p0.$$

Средний индекс – это индекс, вычисленный как средняя величина из индивидуальных индексов. При исчислении средних индексов используются две формы средних: арифметическая и гармоническая.

Средний арифметический индекс физического объема продукции вычисляется по формуле

$$I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}.$$

Средний гармонический индекс тождествен агрегатному, если индивидуальные индексы будут взвешены с помощью слагаемых числителя агрегатного индекса. Например, индекс себестоимости можно исчислить так:

$$I_z = \frac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}.$$

Аналогично рассчитываются индексы других показателей.

2. ВАРИАНТЫ КОНТРОЛЬНОЙ РАБОТЫ

2.1. Теоретические вопросы

1. Предмет, метод и задачи статистики. Источники статистической информации. Связь статистики с другими науками.
2. Основные понятия статистики: статистическая совокупность, единицы и признаки совокупности, статистический показатель, система показателей.
3. Метод статистики. Статистическая закономерность. Закон больших чисел. Задачи статистики в условиях формирования рыночных отношений.
4. Виды статистического наблюдения. Статистическая сводка, ее содержание и задачи.
5. Группировка, виды группировок, их роль в анализе социально-экономических процессов. Группировочные признаки, их обоснование и выбор. Определение числа групп. Классификация.
6. Статистические ряды распределения, их виды. Статистические таблицы, их виды.
7. Виды абсолютных величин, единицы измерения. Относительные величины, их виды и способы расчета.
8. Понятие средней величины. Виды средних величин.
9. Понятие вариации и ее значение. Показатели вариации: размах вариации, дисперсия, среднее квадратическое отклонение, коэффициент вариации.
10. Виды дисперсий, правило сложения дисперсий, свойства дисперсии.
11. Коэффициент детерминации. Эмпирическое корреляционное отношение.
12. Показатели вариации качественных признаков. Теорема сложения дисперсии доли признака.
13. Теоретическая кривая распределения, ее виды. Относительный показатель асимметрии.
14. Структурные характеристики вариационного ряда распределения.
15. Понятие выборочного наблюдения и его значение. Генеральная и выборочная совокупности, их характеристики.
16. Средняя и предельная ошибки выборки. Теорема П.Л. Чебышева.
17. Метод и виды выборки. Определение необходимого объема выборки.
18. Основные этапы статистического изучения социально-экономических явлений. Понятие факторного и результативного признаков. Классификация связи между явлениями и их признаками.
19. Понятие корреляции. Корреляционный анализ. Регрессионный анализ. Виды уравнений регрессии.
20. Парная регрессия. Метод наименьших квадратов.
21. Многофакторный корреляционный и регрессионный анализ. Уравнение множественной линейной двухфакторной регрессии.

22. Проверка адекватности регрессионной модели. Теоретическое корреляционное отношение.
23. Линейный коэффициент корреляции. Эмпирическое и теоретическое корреляционное отношение. Множественный коэффициент корреляции.
24. Многошаговый регрессионный анализ. Проверка значимости уравнения регрессии.
25. Непараметрические методы. Коэффициенты ассоциации и контингенции.
26. Понятие и классификация рядов динамики. Показатели ряда динамики: абсолютный прирост, темп роста, абсолютное значение одного процента прироста.
27. Средние показатели рядов динамики.
28. Методы анализа основной тенденции развития в рядах динамики. (Трендовый анализ).
29. Методы изучения сезонных колебаний.
30. Корреляция рядов динамики.
31. Понятие автокорреляции. Нециклический коэффициент автокорреляции.
32. Понятие экономических индексов. Основные обозначения. Классификация индексов.
33. Индивидуальные и общие индексы. Формула Пааше. Формула Ласпейреса.
34. Понятие средних индексов. Индекс Струмилина.
35. Индексы структурных сдвигов. Понятие индекса-дефлятора.

2.2. Задачи

1. Имеются следующие данные об успеваемости 20 студентов группы по теории статистики в летнюю сессию : 5, 4, 4, 4, 3, 2, 5, 3, 4, 4, 4, 3, 2, 5, 2, 5, 5, 2, 3, 3. Постройте: а) ряд распределения студентов по баллам оценок, полученных в сессию; б) ряд распределения студентов по уровню успеваемости, выделив в нем две группы студентов: неуспевающие (2 балла), успевающие (3 балла и выше); в) укажите, каким видом ряда распределения (вариационным или атрибутивным) является каждый из этих двух рядов.

2. Известны следующие данные о результатах сдачи абитуриентами вступительных экзаменов на I курс вуза (баллов):

18	16	20	17	19	20	17
17	12	15	20	18	19	18
18	16	18	14	14	17	19
16	14	19	12	15	16	20

Постройте: а) ряд распределения абитуриентов по результатам сдачи ими вступительных экзаменов, выделив четыре группы абитуриентов с равными интервалами; б) ряд, делящий абитуриентов на поступивших и не поступивших в вуз, учитывая, что проходной балл составил 15 баллов. Укажите, по какому группировочному признаку построен каждый из этих рядов распределения: атрибутивному или количественному.

3. Имеются следующие данные о производстве бумаги в РФ:

	2004 г.	2005 г.	2006 г.	2007 г.
Произведено бумаги, тыс. т	3603	2882	2215	2771

Вычислите относительные показатели динамики с переменной и постоянной базой сравнения. Проверьте их взаимосвязь.

4. Производство автомобилей в РФ в январе – мае характеризуется следующими данными, тыс. шт.:

	Январь	Февраль	Март	Апрель	Май
Всего	65,0	83,2	79,3	89,9	76,6
В том числе:					
грузовые	11,0	11,5	12,0	11,0	9,3
легковые	54,0	71,7	67,3	78,9	67,3

Рассчитайте относительные показатели динамики с постоянной базой сравнения. Сделайте выводы.

5. Объем продаж АО «ЛОМО» в 2007 г. в сопоставимых ценах вырос по сравнению с предшествующим годом на 5 % и составил 146 млрд. р. Определите объем продаж в 2006 г.

6. Торговая фирма планировала в 2007 г. по сравнению с 2006 г. увеличить оборот на 14,5 %. Выполнение установленного плана составило 102,7 %. Определите относительный показатель динамики оборота.

7. Объем продаж компании Samsung в странах СНГ в первом полугодии 2007 г. составил 250 млн. долл. В целом же за год компания планировала реализовать товаров на 600 млн. долл. Вычислите относительный показатель плана на второе полугодие.

8. Имеются следующие данные о внешнеторговом обороте России со странами дальнего зарубежья и СНГ, млн. долл.:

	IV квартал 2006 г.	I квартал 2007 г.
Экспорт	22 761	20 972
Импорт	18 274	13 954

Вычислите относительные показатели структуры и координации.

9. Имеются следующие данные по остановкам предприятий и производств в промышленности Центрально-Черноземного района в сентябре:

Область	Число предприятий, на	Потери рабочего времени,

	которых имели место про- стои более одной смены	тыс. человеко-дней
Белгородская	61	124
Воронежская	103	286
Курская	57	211
Липецкая	67	131
Тамбовская	60	202

Определите средние потери рабочего времени на одно предприятие в целом по району.

10. Производственная деятельность одного из отделений корпорации за месяц характеризуется следующими данными:

Предприятие	Общие затраты на производст- во, тыс. р.	Затраты на 1 р. произведенной продукции, к.
1	2323,4	75
2	8215,9	71
3	4420,6	73
4	3525,3	78

Определите средние затраты на 1 р. произведенной продукции в целом по отделению.

11. По трем районам города имеются следующие данные (на конец года):

Район	Число отделений Сбербанка	Среднее число вкла- дов в отделение	Средний размер вкла- да, р.
1	4	1376	275
2	9	1559	293
3	5	1315	268

Определите средний размер вклада в Сбербанке в целом по городу.

12. Качество продукции предприятия характеризуется следующими данными (за месяц):

Вид продукции	Процент брака	Стоимость бракованной продукции, р.
A	1,3	2135
B	0,9	3560
C	2,4	980

Определите средний процент брака в целом по предприятию.

13. Распределение студентов одного из факультетов по вузу характеризуется следующими данными:

Возраст студентов, лет	17	18	19	20	21	22	23	24	Всего
Число студентов	20	80	90	110	130	170	90	60	750

Вычислите: а) размах вариации; б) среднее линейное отклонение; в) дисперсию; г) среднее квадратическое отклонение; д) относительные показатели вариации возраста студентов.

14. Определите среднюю длину пробега автофургона торгово-посреднической фирмы и вычислите все показатели вариации, если известны:

Длина пробега за один рейс, км	Число рейсов за квартал
30 – 50	20
50 – 70	25
70 – 90	14
90 – 110	18
110 – 130	9
130 – 150	6
Всего	92

15. Имеется следующий ряд распределения телеграмм, принятых отделением связи, по числу слов:

Количество слов в телеграмме	Число телеграмм
12	18
13	22
14	34
15	26
16	20
17	13
18	7
Итого	140

Рассчитайте абсолютные и относительные показатели вариации.

16. Средняя урожайность зерновых культур в двух районах за ряд лет характеризуется следующими данными, ц/га:

	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.
1-й район	30	20	23	16	22
2-й район	25	34	30	28	29

Рассчитайте все показатели вариации. Определите, в каком районе урожайность зерновых культур более устойчива.

17. Имеются следующие данные выборочного обследования студентов одного из вузов:

Затраты времени на дорогу до института, ч	Число студентов, % к итогу
До 0,5	7
0,5 – 1,0	18
1,0 – 1,5	32
1,5 – 2,0	37
Свыше 2,0	6
Всего	100

Вычислите абсолютные и относительные показатели вариации.

18. Имеются следующие данные о распределении скважин в одном из районов бурения по глубине:

Группы скважин по глубине, м	Число скважин
До 500	4
500 – 1000	9
1000 – 1500	17
1500 – 2000	8
Свыше 2000	2
Итого	40

Определите дисперсию и среднее квадратическое отклонение глубины скважин.

19. Акционерные общества области по среднесписочной численности работающих на 1 января 2008 г. распределялись следующим образом:

Группы АО по среднесписочной численности работающих	До 400	400 – 600	600 – 800	800 – 1000	1000 – 1200	1200 – 1400	1400 – 1600	1600 – 1800	Итого
Количество АО	11	23	36	42	28	17	9	4	170

Рассчитайте: а) среднее линейное отклонение; б) дисперсию; в) среднее квадратическое отклонение; г) коэффициент вариации.

20. По данным о распределении сельских населенных пунктов по числу дворов вычислите общую дисперсию тремя способами: а) обычным; б) упрощенным; в) по формуле $a = x^2 - (x)^2$.

Населенные пункты по числу дворов	Число населенных пунктов, % к итогу
До 100	15,5
101 – 200	28,6
201 – 300	21,7
301 – 400	20,3
Свыше 400	13,9
Итого	100,0

21. Средняя величина признака в совокупности равна 19, а средний квадрат индивидуальных значений этого признака – 397. Определите коэффициент вариации.

22. Дисперсия признака равна 9, средний квадрат индивидуальных его значений – 130. Чему равна средняя?

23. Средняя величина в совокупности равна 16, среднее квадратическое отклонение – 8. Определите средний квадрат индивидуальных значений этого признака.

24. Средний квадрат отклонений индивидуальных значений признака от их средней величины равен 100, а средняя – 15. Определите, чему равен средний квадрат отклонений индивидуальных значений признака от величины, равной 10 и 25.

25. Средняя величина признака равна 14, а дисперсия – 60. Определите средний квадрат отклонений вариантов признака от 19.

26. Имеются данные о распределении семей сотрудников финансовой корпорации по количеству детей:

Число детей в семье	Число семей сотрудников по подразделениям		
	Первое	Второе	Третье
0	4	7	5
1	6	10	13

2	3	3	3
3	2	1	–

Вычислите: а) внутригрупповые дисперсии; б) среднюю из внутригрупповых дисперсий; в) межгрупповую дисперсию; г) общую дисперсию. Проверьте правильность произведения расчетов с помощью правила сложения дисперсий.

27. Распределение основных фондов по малым предприятиям отрасли характеризуется следующими данными:

Группы предприятий по стоимости основных фондов, тыс. р.	Число предприятий	Основные фонды в среднем на предприятии, тыс. р.	Групповые дисперсии
12 – 27	18	18	1,14
27 – 12	40	32	1,09
42 – 57	26	48	1,69
57 – 72	12	69	1,84

Рассчитайте коэффициент детерминации и эмпирическое корреляционное отношение. Сделайте выводы.

28. Каким должен быть объем случайной бесповторной выборки из генеральной совокупности численностью 10 000 единиц при среднем квадратическом отклонении не более 20, предельной ошибке, не превышающей 5 %, и вероятности 0,997?

29. Из партии импортируемой продукции на посту Московской региональной таможни было взято в порядке случайной повторной выборки 20 проб продукта А. В результате проверки установлена средняя влажность продукта А в выборке, которая оказалась равной 6 % при среднем квадратическом отклонении 1 %. С вероятностью 0,683 определите пределы средней влажности продукта во всей партии импортируемой продукции.

30. В одном из лесничеств Рязанской области методом случайной выборки обследованы 1000 деревьев с целью установления их среднего диаметра, который оказался равным 210 мм при $y = 126,5$ мм. С вероятностью 0,683 определите пределы среднего диаметра деревьев в генеральной совокупности.

31. Из партии в 1 млн. шт. мелкокалиберных патронов путем случайного отбора взято для определения дальности боя 1000 шт. Результаты испытаний представлены в следующей таблице:

Дальность боя, м	25	30	35	40	45	50	Итого
Число патронов, шт.	120	180	280	170	140	110	1000

С вероятностью 0,954 определите среднюю дальность боя по выборке, ошибку выборки и возможные пределы средней дальности боя для всей партии патронов.

32. В порядке механической выборки обследован возраст 100 студентов вуза из общего числа 2000 человек. Результаты обработки материалов наблюдения приведены в таблице:

Возраст, лет	17	18	19	20	21	22	23
Число студентов, человек	11	13	18	23	17	10	8

Установите: а) средний возраст студентов вуза по выборке; б) величину ошибки при определении возраста студентов на основе выборки; в) вероятные пределы колебания возраста для всех студентов при вероятности 0,997.

33. В процессе технического контроля из партии готовой продукции методом случайного бесповторного отбора было проверено 70 изделий, из которых четыре оказались бракованными. Можно ли с вероятностью 0,954 утверждать, что доля бракованных изделий во всей партии не превышает 7 %, если процент отбора равен 10?

34. Сколько телефонных разговоров необходимо обследовать на основе случайной бесповторной выборки, чтобы ошибка при определении доли телефонных разговоров с длительностью более 5 мин не превышала 10 % с вероятностью 0,954?

35. С целью определения средней месячной заработной платы персонала гостиниц города было проведено 25%-ное выборочное бесповторное обследование с отбором единиц пропорционально численности типических групп. Для отбора сотрудников внутри каждого типа гостиниц использовался механический отбор. Результаты обследования представлены в следующей таблице:

Тип гостиницы	Средняя месячная заработная плата, р.	Среднее квадратическое отклонение, р.	Число сотрудников, человек
1	870	40	30
2	1040	160	80
3	1260	190	140
4	1530	215	190

С вероятностью 0,954 определите пределы средней месячной заработной платы всех сотрудников гостиниц.

36. Имеются данные о связи между средней заработной платой и прожиточным минимумом. Составьте линейное уравнение регрессии. Вычислите параметры и линейный коэффициент корреляции. Сформулируйте выводы.

№ района	Средняя заработная плата, тыс. р.	Прожиточный минимум на душу населения, тыс. р./месяц
1	1,08	0,49
2	1,63	0,49
3	1,04	0,46
4	1,49	0,52
5	0,97	0,38
6	0,90	0,33
7	0,77	0,34
8	0,69	0,34
9	0,57	0,33

10	0,52	0,28
----	------	------

37. С помощью поля корреляции изобразите графически следующие данные о зависимости объема продаж облигаций на ММВБ и доходности к погашению:

Группы серий по объему продаж, млн. р., x	Группы серий по доходности к погашению, %, y				Всего серий
	43 – 50	50 – 57	57 – 64	64 – 71	
3 – 59	3				3
59 – 115	2	3			5
115 – 171	3	1			4
171 – 227		4	5		9
227 – 283			3	2	3
283 – 339					2
Итого	8	8	8	2	26

Рассчитайте \bar{y}_x (средние групп) и постройте теоретическую линию регрессии. Сформулируйте вывод.

38. Взаимосвязь между стоимостью активной части основных фондов и затратами на производство работ по 35 строительным фирмам представлена следующей таблицей:

Затраты на производство строительно-монтажных работ, % к стоимости активной части основных фондов	Стоимость активной части основных фондов, тыс. р.				Всего фирм
	50 – 100	100 – 150	150 – 200	200 – 250	
1 – 5			2	4	6
5 – 9		2	6	4	12
9 – 13		5	3		8
13 – 17	2	2			4
17 – 21	5				5
Итого	7	9	11	8	35

Постройте поле корреляции и теоретическую линию регрессии. Сформулируйте вывод.

39. Имеются данные о связи между средней взвешенной ценой и объемом продаж облигаций на ММВБ:

№ серии	Средняя взвешенная цена X	Объем продаж, млн. р. Y
22 041	84,42	79,5
22 042	82,46	279,7
22 043	80,13	71,4
22 044	63,42	242,8
22 045	76,17	76,3
22 046	75,13	74,7
22 047	74,84	210,7
22 048	73,03	75,1
22 049	73,41	75,5
22 050	71,34	335,3

Составьте линейное уравнение регрессии. Вычислите параметры и рассчитайте линейный коэффициент корреляции и корреляционное отношение. Сравните величину коэффициента корреляции и корреляционного отношения. Сформулируйте выводы.

40. Зависимость между объемом произведенной продукции и балансовой прибылью по 10 предприятиям одной из отраслей промышленности характеризуется следующими данными:

№ предприятия	Объем реализованной продукции, млн. р.	Балансовая прибыль, млн. р.
1	491,8	133,8
2	483,0	124,1
3	481,7	62,4
4	478,7	62,9
5	476,9	51,4
6	475,2	72,4
7	474,4	99,3
8	459,5	40,9
9	452,9	104,0
10	446,5	116,1

Определите вид корреляционной зависимости, постройте уравнение регрессии, рассчитайте параметры уравнения, вычислите тесно-твязи. Объясните полученные статистические характеристики.

41. Имеются следующие данные о посевной площади зерновых культур, валовом сборе и внесении минеральных удобрений на 1 га посевной площади:

№ фермерского хозяйства	Посевная площадь зерновых культур, тыс. га	Валовой сбор, тыс. т	Внесено минеральных удобрений на 1 га посевной площади, кг
1	4,0	6,0	30
2	2,0	4,6	33

3	3,1	4,4	20
4	3,2	4,5	25
5	3,4	5,5	29
6	3,5	4,8	20
7	3,7	5,1	21
8	3,2	5,2	20
9	3,9	7,0	35
10	3,5	5,3	30
11	5,0	7,5	35
12	3,7	7,7	30
13	5,0	7,3	40
14	3,8	7,0	42
15	5,0	6,7	39

Используя метод приведения параллельных данных, установите направление и характер связи между факторами. Постройте множественное уравнение регрессии, предварительно сформулировав и обосновав выбор результативного и факторных признаков, рассчитайте параметры уравнения, вычислите множественный и частный коэффициенты корреляции. Проанализируйте полученные результаты.

42. В ходе проведенного обследования оценки уровня жизни работающих на предприятиях различной формы собственности было опрошено 100 респондентов. Результаты опроса представлены в следующей таблице:

Форма собственности предприятия	Удовлетворенность уровнем жизни		Итого
	Вполне удовлетворен	Не удовлетворен	
Государственное	30	55	85
Частное	10	5	15
Итого	40	60	100

Рассчитайте коэффициенты ассоциации и контингенции. Сформулируйте выводы, вытекающие из анализа полученных коэффициентов.

43. Остатки вкладов населения в сбербанках города характеризуются следующими данными на 1-е число месяца, тыс. р.:

Январь	Февраль	Март	Апрель	Май	Июнь	Июль
910,5	920,0	915,4	920,8	917,0	921,3	925,9

Определите: а) среднемесячные остатки вкладов населения за первый и второй кварталы; б) абсолютный прирост изменения среднего остатка вклада во втором квартале по сравнению с первым.

44. Списочная численность работников фирмы составила на 1-е число месяца, человек:

январь	– 347
февраль	– 350
март	– 349
апрель	– 351
май	– 345
июнь	– 349
июль	– 357
август	– 359
сентябрь	– 351
октябрь	– 352
ноябрь	– 359
декабрь	– 353
январь (посл. года)	– 360

Определите: а) среднемесячную численность работников в первом и втором полугодиях; б) среднегодовую численность работников фирмы; в) абсолютный прирост численности работников фирмы во втором полугодии по сравнению с первым.

45. Имеются следующие данные по объединению о производстве промышленной продукции за ряд лет (в сопоставимых ценах), млн. р.

04	05	06	07	08	
67,7	73,2	75,7	77,9	81,9	84,4

Для анализа ряда динамики определите: а) средний уровень ряда динамики; б) цепные и базисные темпы роста и прироста; в) для каждого года абсолютное значение 1 % прироста. Результаты расчетов изложите в табличной форме.

46. Имеются следующие данные о производстве молока в России за ряд лет, млн. т:

2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
13,3	13,5	14,8	16,1	16,6	16,4

Установите начальный, конечный и базисный уровни ряда динамики для определения: а) среднего уровня ряда; б) цепных и базисных абсолютных приростов; в) цепных и базисных темпов роста. Определите для каждого года абсолютное значение 1 % прироста. Результаты расчетов изложите в табличной форме.

47. Ввод в действие жилых домов предприятиями всех форм собственности в одном из регионов за ряд лет характеризуется следующими данными, млн. м² общей площади:

2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
17	18	19	20	21	20	22	23

Для анализа ряда динамики: 1) определите: цепные и базисные: а) абсолютные приросты; б) темпы роста; в) темпы прироста; 2) среднегодовой темп прироста; 3) определите для каждого года абсолютное значение 1 % прироста; 4) в целом за весь период рассчитайте среднегодовой абсолютный прирост. Результаты расчетов оформите в таблице и сделайте выводы.

48. Производство электроэнергии в регионе за ряд лет характеризуется следующими данными, млрд. кВт/ч:

2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
915	976	1038	1111	1150	1202	1239	1294

Для анализа ряда динамики: 1) определите показатели, характеризующие динамику производства электроэнергии по годам к базисному 2002 г.: а) темпы роста; б) темпы прироста; в) абсолютные приросты; 2) рассчитайте для каждого года абсолютное значение 1 % прироста. Результаты расчетов изложите в табличной форме.

49. Темпы роста объема продукции текстильной промышленности в области за ряд лет характеризуются следующими данными (в процентах к предыдущему году):

2002 г.	2003 г.	2004 г.	2005 г.	2006 г.
106,3	105,2	106,1	106,3	105,9

Определите среднегодовой темп роста и прироста объема продукции за пятилетие.

50. Приведите уровни следующего ряда динамики, характеризующие численность работников фирмы, к сопоставимому виду, человек:

	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
На 1 января	420	429	427	431	–	–	–	–	–
Среднегодовая численность рабочих	–	–	–	435	442	450	460	465	475

51. Имеются следующие данные о розничном товарообороте во всех каналах реализации в регионе, млрд. р.:

Месяцы	Годы		
	2005	2006	2007
Январь	7,4	7,8	8,3
Февраль	7,9	8,2	8,6
Март	8,7	9,2	9,7
Апрель	8,2	8,6	9,1
Май	7,9	8,3	8,8
Июнь	8,2	8,7	9,1
Июль	8,3	8,8	9,3
Август	8,8	9,3	9,9
Сентябрь	8,7	8,9	9,3
Октябрь	8,8	8,2	9,9
Ноябрь	8,3	8,8	9,8
Декабрь	9,0	9,5	9,3

Для изучения общей тенденции розничного товарооборота региона по месяцам за ряд лет произведите: 1) преобразование исходных данных путем укрупнения периодов времени: а) в квартальные уровни; б) в годовые уровни; 2) сглаживание квартальных уровней розничного товарооборота с помощью скользящей средней. Изобразите графически фактические и сглаженные уровни ряда динамики. Сделайте выводы о характере общей тенденции розничного товарооборота во всех каналах реализации в регионе.

52. Имеются следующие данные об общем объеме розничного товарооборота региона по месяцам, млрд. р.:

1	2	3	4	5	6	7	8	9	10	11	12
22,8	24,9	31,0	29,5	30,5	35,6	36,4	42,6	45,1	47,3	51,0	53,4

Установите, по какой функции – по прямой, параболе второго порядка, показательной кривой – следует произвести выравнивание этого ряда. Найдите тренд, характеризующий динамику общего объема розничного товарооборота региона за 12 месяцев.

53. Имеются следующие данные о среднем размере товарных запасов в универмаге по месяцам года, млн. р.:

1	2	3	4	5	6	7	8	9	10	11	12
21,2	21,3	21,2	21,3	21,2	21,0	21,0	20,2	19,2	20,1	20,8	21,1

Произведите: а) сглаживание ряда товарных запасов универмага методом четырехчленной скользящей средней; б) выравнивание ряда динамики по прямой. Сделайте выводы о характере общей тенденции изучаемого явления.

54. Имеются следующие данные о розничном товарообороте во всех каналах реализации в регионе, млрд. р.:

Месяцы	Годы		
	2005	2006	2007
Январь	7,4	7,8	8,3
Февраль	7,9	8,2	8,6
Март	8,7	9,2	9,7
Апрель	8,2	8,6	9,1
Май	7,9	8,3	8,8

Июнь	8,2	8,7	9,1
Июль	8,3	8,8	9,3
Август	8,8	9,3	9,9
Сентябрь	8,7	8,9	9,3
Октябрь	8,8	8,2	9,9
Ноябрь	8,3	8,8	9,8
Декабрь	9,0	9,5	9,3

Для анализа внутригодовой динамики розничного товарооборота региона определите индексы сезонности, изобразите сезонную волну графически, сделайте выводы.

55. Распределение 1000 семей по уровню среднедушевого дохода за месяц характеризуются следующими данными:

Группы семей по среднедушевому денежному доходу в месяц, р.	Число семей
До 500	50
500 – 600	100
600 – 700	182
700 – 800	162

Продолжение табл.

Группы семей по среднедушевому денежному доходу в месяц, р.	Число семей
800 – 900	150
900 – 1000	120
1000 – 1100	107
1100 – 1200	70
1200 – 1300	48
1300 и более	10
Итого	1000

На основе критерия χ^2 проверьте, согласуется ли распределение семей по среднедушевому доходу с нормальным распределением с вероятностью 0,95.

56. При изучении качества семян было получено следующее распределения семян по проценту всхожести:

Процент всхожести	70	75	80	85	90	92	95	Свыше 95	Итого
Число проб, % к итогу	2	4	7	29	46	8	3	1	100

Рассчитайте моду и медиану.

57. Вычислите моду, медиану, квартили количественного состава семей города на основании следующего их распределения по числу совместно проживающих членов семьи:

Число членов семьи	2	3	4	5	6	7	Итого
Число семей, % к итогу	15	34	25	16	8	2	100

58. С целью исследования качества деталей на предприятии проверена партия из 100 деталей. Результаты представлены в таблице.

Группы деталей по весу, г	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90	90 – 100	100 – 110	110 – 120	Итого
Число деталей	2	4	12	18	21	24	11	8	100

Определите моду, медиану, квартили.

59. Рост цен на молоко в I полугодии в целом по РФ характеризуется следующими данными:

Месяц	Январь	Февраль	Март	Апрель	Май	Июнь
Цена, % к предыдущему периоду	100,8	103,5	98,7	100,1	94,6	95,0

Определите общее изменение цен на молоко за весь рассматриваемый период.

60. По имеющимся в таблице данным о средних оптовых ценах на автомобильный бензин по РФ во II квартале определите недостающие показатели:

Месяц	Цена за 1 т, тыс. р.	Индивидуальные индексы цен	
		Цепные	Базисные
Апрель	?	?	100,0
Май	799	101,9	?
Июнь	?	?	102,8

61. Имеются следующие данные о реализации мясных продуктов на городском рынке:

Продукт	Сентябрь		Октябрь	
	Цена за 1 кг, р.	Продано, ц	Цена за 1 кг, р.	Продано, ц

Говядина	18	26,3	19	24,1
Баранина	15	8,8	15	9,2
Свинина	22	14,5	24	12,3

Рассчитайте сводные индексы цен, физического объема реализации и товарооборота, а также величину перерасхода покупателей от роста цен.

62. Известны следующие данные о реализации фруктов предприятиями розничной торговли округа:

Товар	Цена за 1 кг, р.		Товарооборот, тыс. р.	
	Июль	Август	Июль	Август
Яблоки	8	6	143,5	167,
Груши	11	10	38,9	145,0

Рассчитайте сводные индексы: а) товарооборота; б) цен; в) физического объема реализации. Определите абсолютную величину экономии покупателей от снижения цен.

63. Определите изменение физического объема реализации потребительских товаров предприятиями розничной торговли города в текущем периоде по сравнению с предшествующим, если товарооборот возрос на 42,3 %, а цены повысились на 13,7 %.

64. Объем реализации овощей на рынках города в натуральном выражении в октябре по сравнению с сентябрем возрос на 18,6 %, при этом индекс цен на овощную продукцию составил 92,4 %. Определите изменение товарооборота.

65. Имеются следующие данные о себестоимости и объемах производства продукции промышленного предприятия:

Изделие	2006 г.		2007 г.	
	Себестоимость единицы продукции, тыс. р.	Произведено, тыс. шт.	Себестоимость единицы продукции, тыс. р.	Произведено, тыс. шт.
А	220	63,4	247	52,7
Б	183	41,0	215	38,8
В	67	89,2	70	91,0

Определите: а) индивидуальные и сводный индексы себестоимости; б) сводный индекс физического объема продукции; в) сводный индекс затрат на производство. Покажите взаимосвязь сводных индексов.

66. Имеются следующие данные о реализации молочных продуктов на городском рынке:

Продукт	Товарооборот, тыс. р.		Изменение цены в декабре по сравнению с ноябрем, %
	Ноябрь	Декабрь	
Молоко	9,7	6,3	+2,1
Сметана	4,5	4,0	+3,5
Творог	12,9	11,5	+4,2

Рассчитайте сводные индексы цен, товарооборота и физического объема реализации.

67. Розничный товарооборот РФ в январе характеризуется следующими данными:

Товары	Удельный вес в общем объеме товарооборота, % к итогу	Индивидуальный индекс цен (по сравнению с декабрем), %
Продовольственные	47	123
Непродовольственные	53	112

Определите сводный индекс цен на потребительские товары.

68. По промышленному предприятию имеются следующие данные:

Изделие	Общие затраты на производство в 2007 г., млн. р.	Изменение себестоимости изделия в 2007 г. по сравнению с 2006 г., %
Электромясорубка	1234	+6,0
Кухонный комбайн	5877	+8,4
Миксер	980	+1,6

Определите общее изменение себестоимости продукции в 2007 г. по сравнению с 2006 г. и обусловленный этим изменением размер экономии или дополнительных затрат предприятия.

69. Известны следующие данные по заводу строительных пластмасс:

Вид продукции	Общие затраты на производство в предшествующем году, млн. р.	Изменение объема производства в натуральном выражении, %
Линолеум	2427	+6,5
Винилискожа	985	+4,5
Пенопласт	1365	-2,0
Пленка	771	-11,0

Сделайте сводную оценку увеличения производства продукции (в натуральном выражении).

Определите: а) средний рост цен на данную группу товаров по торговому предприятию; б) перерасход покупателей от роста цен.

70. Имеются следующие данные о реализации картофеля на рынках города:

Рынок	Январь		Февраль	
	Цена за 1 кг, р.	Продано, ц	Цена за 1 кг, р.	Продано, ц
1	2,2	24,5	2,4	21,9
2	2,0	18,7	2,1	18,8
3	1,9	32,0	1,9	37,4

Рассчитайте: а) индекс цен переменного состава; б) индекс цен фиксированного состава; в) индекс структурных сдвигов.

Список литературы

1. Герасимов, Б.И. Статистика / Б.И. Герасимов. – Тамбов : Тамб. гос. техн. ун-та, 1997. – 71 с.
2. Макроэкономическая региональная статистика / Б.И. Герасимов, Л.Г. Гурова, В.В. Дробышева, В.М. Золотухина. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2004. – 112 с.
3. Елисеева, И.И. Общая теория статистики : учебник / И.И. Елисеева, М.М. Юзбашев ; под ред. И.И. Елисеевой. – М. : Финансы и статистика, 1995. – 368 с.
4. Курс социально-экономической статистики : учебник для вузов / под ред. проф. М.Г. Назарова. – М. : Финстатинформ, 2002. – 976 с.
5. Микроэкономическая статистика : учебник / под ред. С.Д. Ильенковой. – М. : Финансы и статистика, 2004. – 544 с.
6. Общая теория статистики : учебник / А.И. Харламов, О.Э. Башина и др. ; под ред. А.А. Спирина, О.Э. Башиной. – М. : Финансы и статистика, 1997. – 296 с.
7. Практикум по теории статистики / Р.А. Шмойлова, и др. ; под ред. Р.А. Шмойловой. – М. : Финансы и статистика, 1998. – 416 с.
8. Социально-экономическая статистика : практикум / под ред. В.Н. Салина, Е.П. Шпаковской. – М. : Финансы и статистика, 2003. – 192 с.
9. Теория статистики : учебник / под ред. проф. Р.А. Шмойловой. – М. : Финансы и статистика, 2002. – 464 с.