МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВПО «Пензенская ГСХА»

И.Д. Минина
Н.В. Королькова

СТАТИСТИКА

Пенза 2013

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВПО «Пензенская ГСХА»

Кафедра «Анализ и аудит»

И.Д. Минина
Н.В. Королькова

СТАТИСТИКА

Учебное пособие

Часть 1 Теория статистики

Для студентов высших учебных заведений, обучающихся
по направлениям 080100 «Экономика» (профиль
«Бухгалтерский учет, анализ и аудит» и «Финансы и кредит»)
и 080200 «Менеджмент»

Пенза 2013

УДК 311 (075)
ББК 65.051 (я7)
М 61

Рецензент − О.А. Столярова, кандидат экономических наук, доцент, зав. кафедрой «Экономика АПК» Пензенской ГСХА

Печатается по решению методической комиссии экономического факультета от 5 марта 2012 г. протокол № 33.

Минина, Инна Дмитриевна
М 61 Статистика: учебное пособие. – Часть 1 Теория статистики/ И.Д. Минина, Н.В. Королькова. − Пенза: РИО ПГСХА, 2013. – 225 с.

Учебное пособие предназначено для оказания помощи студентам в освоении курса общей теории статистики. В методической разработке дано теоретическое описание содержания методов, категорий и понятий общей теории статистики, приведены решения типовых задач, рассмотрены практические возможности использования статистических методов для изучения общественных явлений. С целью выработки навыков по применению статистических методов для исследования и анализа экономических процессов приведены задачи для самостоятельного решения.

Учебное пособие предназначено для студентов, обучающихся по направлениям бакалавриата.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Данное учебное пособие предназначено для оказания помощи студентам в осмыслении категорий статистической науки, умении увидеть конкретное содержание в каждом статистическом показателе и выработке навыков по применению научных методов статистического исследования при решении задач различного типа в разных областях экономики.

Учебное пособие состоит из 10 глав. Каждая глава состоит из методических указаний с решением типовых задач и задач, предназначенных для самостоятельного решения студентами.

В первой части каждой главы даются методические указания, где раскрывается содержание каждого статистического метода и показываются методические исчисления показателей, которые используются в аналитической работе, а также приводятся решения типовых задач.

Во второй части глав представлен набор задач для проведения практических занятий и самостоятельных заданий студентам, на аналитических данных, взятых из статистических сборников и периодической печати. Незначительная часть задач построена по условным данным.

В приложениях содержатся необходимые данные для решения задач и математико-статистические таблицы. В приложении приведены также Интернет-ресурсы, содержащие статистическую информацию и аналитические обзоры.

В результате овладения знаниями и навыками по изучаемой дисциплине у студентов должны быть сформированы следующие профессиональные компетенции:

1. Способность понимать сущность и значение информации в развитии современного информационного общества, сознавать опасности и угрозы, возникающие в этом процессе, соблюдать основные требования информационной безопасности, в том числе защиты государственной тайны.

2. Способность собрать и проанализировать исходные данные, необходимые для расчета экономических и социально-экономических показателей, характеризующих деятельность хозяйствующего субъекта;

3. Способность на основе типовых методик и действующей нормативно-правовой базы рассчитать экономические и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов;

4. Способность выполнять необходимые для составления экономических планов расчеты, обосновать их и представлять результаты работы в соответствии с принятыми в организации стандартами;

5. Способность выбрать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, проанализировать результаты расчетов и обосновать полученные выводы;

6. Способность на основе описания экономических процессов и явлений строить стандартные теоретические и эконометрические модели, анализировать и содержательно интерпретировать полученные результаты;

7. Способность анализировать и интерпретировать данные отечественной и зарубежной статистики о социально-экономических процессах и явлениях, выявлять тенденции изменения социально-экономических показателей;

8. Способность владеть методами количественного анализа и моделирования, теоретического и экспериментального исследований.

Бакалавры по направлению 080100 «Экономика» по профилю «Финансы и кредит» и профилю «Бухгалтерский учет, анализ и аудит» по результатам обучения статистике должны обладать знаниями в разрезе компетенций ОК-12, ПК-1, ПК-2, ПК-3, ПК-4, ПК-5, ПК-6, ПК-8 и ПК-9, а по направлению 080200 «Менеджмент» − ОК-15.

1 СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

Статистическое наблюдение – это массовое, планомерное, научно организованное наблюдение за явлениями социальной и экономической жизни, которое заключается в регистрации отобранных признаков у каждой единицы совокупности.

Проведение статистического наблюдения предлагает решение широкого круга разнообразных вопросов.

Сначала необходимо решить программно-методологические вопросы его проведения. Это определение цели и объема наблюдения, состав признаков, подлежащих регистрации, разработка документов для сбора данных, выбор отчетной единицы и единицы, относительно которой будет проводиться наблюдение. Затем решаются вопросы организационного характера. Сюда относятся: определение состава органов, проводящих наблюдение; подборка кадров для проведения наблюдения; составление календарного плана работ по подготовке, проведению и обработке материалов наблюдения; проведение тиражирования документов для сбора данных.

Статистические наблюдения классифицируются *по форме, виду и способу проведения* статистического наблюдения.

Формы статистического наблюдения выделяются на основе их наиболее общих организационных особенностей. В отечественной статистике выделяют три основные формы: отчетность, специально организованное статистическое наблюдение и регистры.

Виды статистического наблюдения классифицируются по следующим признакам:

- охвату наблюдением единиц совокупности, подлежащих статистическому исследованию;
- систематичности наблюдения;
- источнику сведений.

По первому признаку выделяют *сплошное наблюдение*, когда наблюдению подвергаются все без исключения единицы совокупности и *несплошное*, при котором сведения собирают не о всех единицах совокупности, а только некоторой их части, отобранной определенным способом. Несплошное наблюдение, в свою очередь, подразделяют на *выборочное, основного массива, монографическое*. Различие между этими видами заключается в способе отбора тех единиц, которые должны быть подвергнуты наблюдению.

По признаку систематичности наблюдения делятся на *непрерывное, или текущее*, и *прерывное наблюдение*. Последнее подразделяется на периодическое и единовременное. *Текущее* – это наблюдение, которое проводится постоянно, по мере возникновения фактов (например, регистрация браков и разводов). *Прерывное* проводится с перерывами, время от времени. Если оно проводится регулярно, т.е. через равные промежутки времени, оно называется *периодическим*, если же такой регулярности нет, оно называется *единовременным*.

Программно-методологические вопросы наблюдения. Каждый объект состоит из многих элементов или единиц. Элемент объекта, который является носителем признаков, подлежащих регистрации, называется единицей наблюдения.

Программа наблюдения получает свое воплощение в перечне вопросов, ответы на которые нужно получить в процессе наблюдения. Вопросы программы наблюдения фиксируются в формуляре (бланке наблюдения). Конструируя формуляр наблюдения, нужно учитывать объем программы наблюдения, способ проведения наблюдения и способ обработки данных.

К проектируемому наблюдению составляются инструкции, которые представляют собой разъяснения программно-методологических вопросов наблюдения, в частности, как следует записать ответ на тот или иной вопрос и на основании чего (источник сведений). Инструкция может быть оформлена в виде отдельного документа или записана на самом бланке наблюдения.

Организационные вопросы наблюдения. Организационный план наблюдения – это документ, в котором зафиксированы все важнейшие организационные мероприятия, проведение которых необходимо для успешного осуществления наблюдения.

В организационном плане указываются цель, объект, единицы, место, время, орган наблюдения, а также перечисляются подготовительные мероприятия и наблюдения: подбор и обучение кадров, разбивка территории на части, в которых проведение наблюдения поручается различным лицам и т.д.

По источнику сведений различают наблюдение *непосредственное*, когда факты, подлежащие регистрации, устанавливаются лицами путем замера, подсчета числа каких-либо предметов и т.п., *документированное*, при котором необходимые сведения берутся из соответствующих документов, и *опрос*, особенность которого состоит в том, что сведения финансируются со слов опрашиваемого. При этом различаются следующие виды опроса: *экспедиционный, саморегистрация, явочный способ, корреспондентский способ, анкетный.*

В ходе наблюдения могут возникнуть погрешности. Погрешности, проявляющиеся в процессе наблюдения, называются *ошибками наблюдения*. Все погрешности, возникающие при сплошном наблюдении, называются *ошибками регистрации*. Для предупреждения или уменьшения этих погрешностей следует предусматривать специальные контрольные мероприятия (например, проведение повторного наблюдения единиц наблюдения, отобранных в порядке выборки). При несплошном наблюдении, в частности, выборочном, могут возникать специфические ошибки, называемые *ошибками репрезентативности*. Они появляются в силу того, что наблюдение является несплошным.

После получения статистических формуляров следует провести проверку полноты и качеств собранных данных. *Контроль полноты* – это проверка того, насколько полно охвачен объект наблюдением, иначе говоря, о всех ли единицах наблюдения собраны сведения. *Контроль качества материала* осуществляется с помощью логического и арифметического контроля.

Задача 1

Рассмотрев формуляры Всесоюзной переписи населения 1989 г., а также Всероссийских переписей населения 2002 и 2010 г., ответьте на следующие вопросы:

1) к какому виду относится каждый из них?

2) дайте определение объекта каждой из переписей;

3) в чем заключаются различия в программах этих переписей?

4) в чем заключаются различия в формулировках вопросов о возрасте?

5) укажите различия в постановке вопросов о семейном положении;

6) в чем заключаются и чем обусловлены различия в постановке вопросов о занятиях населения в этих переписях?

7) чем еще между собой различаются формуляры переписей?

8) имеются ли в переписных листах подсказы? Если есть, то в каких вопросах, и какого содержания (полные, неполные)?

Задача 2

Составьте перечень наиболее существенных признаков следующих единиц статистического наблюдения:

1) фермерское хозяйство;
2) жилой дом (для жилищной переписи);
3) вуз;
4) библиотека;
5) театр;
6) совместное предприятие.

Задача 3

Какие вы наметите признаки, которые следует регистрировать при проведении:

1) обследования промышленной фирмы с целью изучения текучести рабочей силы;
2) обследования работы городского транспорта с целью изучения роли различных его видов в перевозках пассажиров;
3) обследования студентов вуза с целью изучения бюджета времени.

Задача 4

Сформулируйте объект, единицу и цель наблюдения и разработайте программу обследования:

1) детских садов;
2) фирм, выпускающих детское питание;
3) автозаправочных станций;
4) гостиничного комплекса региона.

Задача 5

Сформулируйте вопросы для включения их в формуляр наблюдения по следующим признакам объектов наблюдения:

1) количество работников на фирме;
2) численный состав семьи:
3) родственные связи членов семьи;
4) пол и возраст человека.

Задача 6

Сформулируйте вопросы программы наблюдения и составьте макет статистического формуляра, а также краткую инструкцию к его заполнению, для изучения зависимости успеваемости от пола, возраста, семейного положения, жилищных условий и общественной активности студента вуза при проведении специального статистического обследования по состоянию на февраль 2011 г. Укажите, к какому виду относится данное наблюдение по времени, охвату и способу получения данных.

Задача 7

С целью изучения мнения студентов об организации учебного процесса в вузе, в котором вы учитесь, необходимо провести специальное обследование. Требуется определить:

1) объект и единицу наблюдения;
2) признаки, подлежащие регистрации;
3) вид и способ наблюдения;
4) разработать формуляр и написать краткую инструкцию по его заполнению;
5) составить организационный план обследования;
6) произвести наблюдение в вашей студенческой группе и результаты его представить в виде таблиц.

Задача 8

Разработайте программу и формуляр единовременного обследования жилищных условий студентов вузов своего города по состоянию на 01.01.2011 г., а также организационный план этого наблюдения.

Задача 9

Определите место, время и органы проведения статистических наблюдений:

1) учета валютных операций коммерческих банков;

2) выборочного обследования бюджетов семей пенсионеров;

3) учета доходов граждан и источников их поступлений, который осуществляется налоговыми инспекциями по итогам календарного года.

Задача 10

Перепись населения проводилась в период с 14 октября по 25 октября 2010 г. Критическим моментом было 0 часов ночи с 14 на 15 октября.

Счетчик пришел:

1) в семью № 1 15 октября. В этой семье 10 октября умер человек. Как должен поступить счетчик:

а) не вносить сведения об умершем в переписной лист;

б) внести с отметкой о смерти;

в) внести без отметки о смерти;

2) в семью № 2 15 октября и попал на свадьбу. Два часа назад молодожены возвратились из загса после регистрации брака (до этого в зарегистрированном браке они не состояли). Что должен записать счетчик в ответ на вопрос: «Состоите ли вы в браке в настоящее время» о каждом из супругов – состоит или не состоит?

3) в семью № 3 16 октября. В семье 15 октября родился ребенок. Как должен поступить счетчик относительно этого ребенка:

а) внести в переписной лист,

б) не вносить в переписной лист;

4) в семью № 4 20 октября. Один из членов семьи на вопрос: «Состоит ли он в браке в настоящее время», ответил, что не состоит, и показал счетчику свидетельство о расторжении брака, в котором указано, что брак расторгнут в первый день переписи – 15 октября. Несмотря на возражения опрашиваемого, счетчик зарегистрировал его состоящим в браке. Правильно ли поступил счетчик?

2 СТАТИСТИЧЕСКАЯ СВОДКА И ГРУППИРОВКА

Важнейшим этапом исследования социально-экономических явлений и процессов является систематизация первичных данных и получение на этой основе сводной характеристики объекта в целом при помощи обобщающих показателей, что достигается путем сводки и группировки первичного статистического материала.

Метод группировок в единстве с другими статистическими методами является важным средством социально-экономического познания, а также ведущим звеном в статистическом исследовании. Можно собрать прекрасный статистический материал, но испортить его неумелой сводкой и группировкой. Рассмотрим основные понятия и категории.

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

По глубине и точности обработки материала 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 сложной 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.

Метод группировки и его место в системе статистических методов. Группировкой называется разделение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам. Группировка в статистическом анализе выполняет следующие определенные функции:

− выделение социально-экономических типов явлений;

− изучение структуры и структурных сдвигов, происходящих в социально-экономических явлениях;

− анализ взаимосвязей между явлениями.

Виды статистических группировок. В соответствии с функциями группировки, различают следующие ее виды: типологическая, структурная, аналитическая.

Типологическая группировка – это разделение качественно неоднородной совокупности на отдельные качественно однородные группы и выявление на этой основе экономических типов явлений. Таким образом, основная задача такой группировки – это идентификация типов социально-экономических явлений, поэтому важное значение при ее построении должно уделяться выбору группировочного признака.

Структурная группировка – это выявление закономерностей распределения единиц однородной совокупности по варьирующим значениям исследуемого признака. Она позволяет изучить структуру совокупности и происходящих в ней сдвигов. Надобность в таких группировках возникает потому, что однородность однокачественных явлений, элементов, входящих в статистическую совокупность, отнюдь не означает их тождественности.

Структурные группировки отличаются от типологических не столько по внешнему виду, сколько по целям, т.е. отличаются по уровню качественных различий между группами.

Аналитическая группировка – это исследование взаимосвязей варьирующих признаков в пределах однородной совокупности. При ее построении можно установить взаимосвязи между двумя признаками и более. При этом один признак будет результативным, а другой – факторным.

Факторными называются признаки, оказывающие влияние на изменение результативных. Результативными называются признаки, изменяющиеся под влиянием факторных.

Основные этапы построения аналитической группировки следующие:

− обоснование и выбор факторного и результативного признаков;

− группировка единиц совокупности по факторному признаку;

− подсчет числа единиц в каждой из образованных групп и определение объема варьирующих признаков в созданных группах;

− исчисление средних размеров результативного показателя (признака) по каждой из образованных групп;

− оформление результатов группировки в таблице;

− сопоставление изменения значений факторного и результативного признаков, определяющее характер связи между ними, т.е. выявление взаимосвязи между признаками, когда с возрастанием значения факторного признака систематически возрастает или убывает значение результативного признака.

Принципы построения статистических группировок и классификаций. Построение группировки начинается с определения состава группировочных признаков.

Выбор группировочного признака, т.е. признака, по которому производится объединение единиц исследуемой совокупности в группы – один из самых существенных и сложных вопросов теории группировки и статистического исследования. От правильного выбора группировочного признака зависят выводы статистического исследования. В качестве основания группировки необходимо использовать существенные обоснованные признаки.

Группированным 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 (объем торгов, возраст человека, доход семьи и т.д.), а вторые отражают состояние единицы совокупности (пол человека, семейное положение, отраслевую принадлежность предприятия, его форму собственности и т.д.).

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

Если группировка строится по атрибутивному признаку, то число групп, как правило, будет таким, сколько имеется градаций, видов состояний у этого признака. Например, группировка предприятий по формам собственности учитывает муниципальную, федеральную и собственность субъектов Федерации.

Если группировка проводится по количественному признаку, то число групп зависит от числа единиц исследуемого объекта и степени колеблемости группировочного признака, в каждом отдельном случае его необходимо обосновать.

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.

Согласно формуле, выбор числа групп зависит от объема совокупности.

Недостаток формулы состоит в том, что ее применение дает хорошие результаты, если совокупность состоит из большого числа единиц и распределение единиц по признаку, положенному в основание группировки, близко к нормальному.

Для определения числа групп можно воспользоваться данными таблицы 1, где количество групп задается, не так жестко, как в предыдущей формуле.

Таблица 1 – Определение числа групп в зависимости от численности совокупности

Численность совокупности, ед.	Рекомендуемое число групп
До 40	3-4
40-60	4-5
60-100	5-6
100-300	6-8
Свыше 300	8-10

Когда определено число групп, то следует определить интервалы группировки.

Интервал – это значение варьирующего признака, лежащее в определенных границах. Каждый интервал имеет свою величину, верхнюю и нижнюю границы или хотя бы одну из них. Нижней границей интервала называется наименьшее значение признака в интервале, а верхней границей – наибольшее значение признака в интервале. Величина интервала представляет собой разность между верхней и нижней границами интервала.

Интервалы группировки в зависимости от их величины бывают *равные и неравные*.

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:

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

где X_{max} и X_{min} − максимальное и минимальное значения признака в совокупности;

n – число групп.

Если максимальное и минимальное значения сильно отличаются от смежных с ними значений вариантов в упорядоченном ряду значений группировочного признака, то для определения величины интервала следует использовать не максимальное и минимальное значения, а значения, несколько превышающие минимум и несколько меньшие, чем максимум.

Существуют нижеследующие правила записи числа шага интервала. Если величина интервала, рассчитанная по формуле, представляет собой величину, которая имеет один знак до запятой (например, 0,88; 1,585; 4,71), то в этом случае полученные значения целесообразно округлить до десятых и их использовать в качестве шага интервала. В приведенном выше примере это будут соответственно значения: 0,9; 1,6; 4,7. Если рассчитанная величина интервала имеет две значащие цифры до запятой и несколько знаков после запятой (например, 15,985), то это значение необходимо округлить до целого числа (до 16). В случае, когда рассчитанная величина интервала представляет собой трехзначное, четырехзначное и так далее число, эту величину следует округлить до ближайшего числа, кратного 100 или 50. Например, 557 следует округлить до 600.

Если размах вариации признака велик и его значения варьируются неравномерно, то необходимо использовать группировку с неравными интервалами.

Неравные интервалы могут быть прогрессивно-возрастающими или прогрессивно-убывающими в арифметической или геометрической прогрессии. Величина интервалов, изменяющихся в арифметической прогрессии, определяется следующим образом:

$$h_{i+1} = h_i + a,$$

а в геометрической прогрессии:

$$h_{i+1} = h_i \times q,$$

где a – константа, имеющая для прогрессивно-возрастающих интервалов знак «+», а для прогрессивно-убывающих интервалов знак «–»;

q – константа (для прогрессивно-убывающих интервалов $q > 1$; в другом случае – $q < 1$).

Применение неравных интервалов обусловлено тем, что в первых группах небольшая разница в показателях имеет большое значение, а в последних группах эта разница несущественна.

Например, при построении группировки промышленных предприятий строительного комплекса по показателю численности работающих, который варьирует от 400 до 2800 чел., целесообразно рассматривать неравные интервалы. Поэтому следует образовывать неравные интервалы: 400–800; 800–1600; 1600–2800, т.е. величина каждого последующего интервала больше предыдущего на 400 чел. и увеличивается в арифметической профессии.

Интервалы группировок могут быть закрытыми и открытыми.

Закрытыми называются интервалы, у которых имеются верхняя и нижняя границы. У *открытых* интервалов указана только одна граница: верхняя – у первого, нижняя – у последнего.

Например, группы коммерческих банков по уровню дохода работающих в них сотрудников (тыс. руб.): до 10; 10–20; 20–30; 30–40; 40 и более.

При группировке единиц совокупности по количественному признаку границы интервалов могут быть обозначены по разному, в зависимости от того, непрерывный это признак или прерывный.

Если основанием группировки служит непрерывный признак (например, группы строительных фирм по объему работ (млн. руб.): 12–14, 14–16, 16–18, 18–20), то одно и то же значение признака выступает и верхней, и нижней границами двух смежных интервалов. В данном случае объем работ 14 млн. руб. составляет верхнюю границу первого и нижнюю границу второго интервалов; 16 млн. руб. – соответственно второго и третьего и т.д., т.е. верхняя граница i-го интервала равна нижней границе $(i + 1)$ интервала.

При таком обозначении границ может возникнуть вопрос, в какую группу включать единицы объекта, значения признака у которых совпадают с границами интервалов. Например, во вторую или третью группу должна войти строительная фирма с объемом работ 16 млн. руб. Если нижняя граница формируется по принципу «включительно», а верхняя – по принципу «исключительно», то фирма должна быть отнесена к третьей группе, в противном случае – ко второй.

Разновидностью группировок являются ряды распределения.

Ряды распределения – это упорядоченное распределение единиц совокупности по определенному признаку.

В зависимости от признака, положенного в основу ряда распределения, различают атрибутивные и вариационные ряды распределения.

Атрибутивным называют ряд распределения, построенный по качественным признакам, например, распределение студентов группы по полу.

Вариационным рядом называют ряд распределения, построенный по количественному признаку. В зависимости от характера вариации признака различают дискретные и интервальные ряды.

Дискретный вариационный ряд характеризует распределение единиц совокупности по признаку, принимающему только целые значения.

Построение *интервального ряда* целесообразно при непрерывной вариации признака.

Анализ рядов распределения проводится на основе их графического изображения.

Полигон используется при изображении дискретных вариационных рядов, *гистограмма* применяется для изображения интервального вариационного ряда.

Кумулятивная кривая строится по накопленным частотам, которые откладываются по оси ординат, а по оси абсцисс откладываются варианты ряда.

Задача 11

Определите, к какому виду группировки относится статистическая таблица 2.

Таблица 2 – Группировка предприятий по размеру основных фондов (данные условные)

Группа предприятий по размеру основных фондов	Число предприятий	Объем выпускной продукции, млн руб.		Численность занятых, чел.	
		всего	на одном предприятии	всего	на одном предприятии
Мелкие	20	1500	75	2000	100
Средние	20	2000	100	3000	150
Крупные	10	4500	450	5000	500
Итого	50	8000	160	10000	200

Задача 12

Какие из указанных ниже группировок являются типологическими:

1) группировка населения по полу;

2) группировка населения, занятого в народном хозяйстве по отраслям;

3) группировка капитальных вложений на строительство объектов производственного и непроизводственного назначения;

4) группировка предприятий общественного питания по формам собственности.

Задача 13

Известны следующие данные о численности населения Центрального федерального округа РФ на 01.01. 2010 г. в разрезе областей (млн чел.):

1,5	1,2	2,2	1,6
1,9	1,1	0,9	1,8
1,6	0,8	1,3	2,1
2,4	1,3	1,1	1,2

Используя эти данные, постройте интервальный вариационный ряд распределения областей Центрального федерального округа РФ, выделив три группы областей с равными открытыми интервалами. По какому признаку построен ряд распределения: качественному или количественному?

Задача 14

Имеются следующие данные об успеваемости 30 студентов группы по теории статистики в летнюю сессию 2010 г.: 5, 4, 3, 3, 5, 4, 4, 4, 3, 4, 4, 5, 4, 4, 3, 2, 5, 3, 4, 4, 4, 3, 2, 5, 2, 5, 5, 2, 3, 3.

Постройте:

1) ряд распределения студентов по оценкам, полученным в сессию, и изобразите его графически;

2) ряд распределения студентов по уровню успеваемости, выделив две группы студентов: неуспевающих (2 балла), успевающих (3 балла и выше);

3) укажите, каким видом ряда распределения (вариационным или атрибутивным) является каждый из этих двух рядов.

Задача 15

Данные о деятельности банков, представленные в таблице 3.

1. Постройте группировку коммерческих банков по величине собственного капитала, выделив четыре группы с равными интервалами. Рассчитайте по каждой группе сумму активов, собственный капитал, привлеченные ресурсы, прибыль. Результаты

группировки представьте в табличной форме и сформулируйте выводы.

2. Постройте полигон и гистограмму распределения банков по величине собственного капитала.

Таблица 3 – Основные показатели деятельности банков, млн. руб. (данные условные)

№ п/п	Сумма активов	Собственный капитал	Привлеченные ресурсы	Прибыль	Объем вложений в государственные ценные бумаги	Ссудная задолженность
1	645,6	12,0	27,1	8,1	3,5	30,8
2	636,9	70,4	56,3	9,5	12,6	25,7
3	629,0	41,0	95,7	38,4	13,3	26,4
4	619,6	120,8	44,8	38,4	4,4	25,3
5	616,4	49,4	108,7	13,4	15,0	20,9
6	614,4	50,3	108,1	30,1	19,1	47,3
7	608,6	70,0	76,1	37,8	19,2	43,7
8	601,1	52,4	26,3	41,1	3,7	29,1
9	600,2	42,0	46,0	9,3	5,2	56,1
10	600,0	27,3	24,4	39,3	13,1	24,9
11	592,9	72,0	65,5	8,6	16,7	39,6
12	591,7	22,4	76,0	40,5	7,5	59,6
13	585,5	39,3	106,9	45,3	6,7	44,9
14	578,6	70,0	89,5	8,4	11,2	32,2
15	577,5	22,9	84,0	12,8	19,3	45,1
16	553,7	119,3	89,4	44,7	19,4	24,5
17	543,6	49,6	93,8	8,8	5,7	31,1
18	542,0	88,6	26,7	32,2	7,8	37,1
19	517,0	43,7	108,1	20,3	8,3	23,1
20	516,7	90,5	25,2	12,2	9,7	15,8

Задача 16

Постройте аналитическую группировку коммерческих банков, перечисленных в задаче 15, по величине прибыли, выделив четыре группы. Рассчитайте по каждой группе два-три показателя, взаимосвязанных с прибылью. Результаты группировки изложите в табличной форме и сделайте выводы о взаимосвязи показателей.

Задача 17

Используя данные задачи 15, постройте группировку коммерческих банков в целях выявления взаимосвязи между показателями привлеченных ресурсов, объемом вложений в государственные ценные бумаги и ссудной задолженностью от результатов деятельности банков (показатель, выражающий результаты деятельности банков определите самостоятельно).

Задача 18

По данным приложения 2 постройте аналитическую группировку для выявления зависимости между урожайностью зерновых культур и себестоимостью 1 ц зерна.

Задача 19

По данным приложения 2 постройте группировку для выявления зависимости между уровнем рентабельности производства зерна, урожайностью зерновых культур и себестоимостью 1 ц зерна.

Задача 20

По данным приложения 2 постройте типологическую группировку с целью выявления зависимости эффективности производства зерна от организационно-правовой формы хозяйствования.

3 СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

Статистической называется *таблица,* которая содержит сводную числовую характеристику исследуемой совокупности по одному или нескольким существенным признакам, взаимосвязанным логикой экономического анализа. Прежде чем переходить к рассмотрению видов и правил построения статистических таблиц, необходимо иметь представление об основных элементах, ее формирующих.

Основные элементы статистической таблицы, составляющие ее основу, показаны на рисунке 1.

Название таблицы

<table>
<tr><td rowspan="1">Содержание строк</td><td colspan="3">Наименование граф (верхние заголовки)</td><td rowspan="2">Итоговая графа</td></tr>
<tr><td>А</td><td>1</td><td>2</td><td>…</td></tr>
<tr><td rowspan="3">Наименование строк</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Итоговая строка</td><td></td><td></td><td></td><td></td></tr>
</table>

Рисунок 1 – Основные элементы статистической таблицы

Виды таблиц по характеру подлежащего. Подлежащим статистической таблицы называется объект, который в ней характеризуется цифрами. Это могут быть совокупность, отдельные единицы совокупности в порядке их перечня или сгруппированные по одному или нескольким признакам территориальные единицы, временные периоды и т.д.

В соответствии с этим в зависимости от структуры подлежащего, различают статистические таблицы простые, в подлежащем которых дается простой перечень единиц совокупности (перечневые) или только одна какая-либо из них единица, выделенная по определенному признаку (монографические), и сложные, подлежащее которых содержит группы единиц совокупности по одному (групповые) или нескольким (комбинационные) количественным или атрибутивным признакам. При этом подлежащее простой таблицы может быть сформировано по видовому, территориальному и временному принципам.

Приведем примеры разработки подлежащего таблицы.

В таблице 4 приведен пример простой перечневой таблицы, где подлежащим является марка бензина.

Таблица 4 – Цены на бензин ООО «Автобез» на 01.01.2010 г.

Марка бензина	Средневзвешенная цена, руб. за 1 л	Суммарный объем предложений, т	Минимальный объем партии, т
А-92	20	5000	2000
А-95	22	4000	2000
А-98	23	2000	1000

В таблице 5 приведен пример простой территориальной таблицы.

Таблица 5 – Структура инвестиций в основной капитал по видам основных фондов в некоторых федеральных округах РФ в 2010 г. (% от общего объема инвестиций)

Федеральный округ	Всего инвестиций	В том числе:			
		жилища	здания и сооружения	машины, оборудование, инструмент, инвентарь	прочие
Центральный	100	17,8	40,8	35,7	5,7
Северо-Западный	100	6,8	43,6	44,5	5,1
Южный	100	10,0	58,8	24,3	6,9
Приволжский	100	11,4	38,5	41,8	8,3
Уральский	100	4,7	44,3	29,1	21,9
Сибирский	100	8,9	44,6	42,1	4,4
Дальневосточный	100	9,8	43,6	37,2	9,4

Подлежащим является перечень федеральных округов.

В таблице 6 приведен пример простой хронологической таблицы.

Таблица 6 – Динамика инвестиций в основной капитала в одном из регионов РФ (данные условные)

Год	Инвестиции в основной капитал, млн. руб.	Индекс физического объема инвестиций в основной капитал, % к предыдущему году	Инвестиции в основной капитал на душу населения, тыс. руб.
2006	43,9	92	5,1
2007	53,4	106	6,2
2008	65,2	124	7,6
2009	96,1	104	11,3
2010	150,6	115	17,6

Подлежащим в этой таблице являются годы.

В таблице 7 приведен пример групповой таблицы по одному признаку, где подлежащим выступают группы несовершеннолетних, совершивших правонарушение по возрасту.

Таблица 7 – Распределение несовершеннолетних, совершивших правонарушения и преступления в одном из регионов РФ в 2010 (по возрасту) (данные условные)

Группа несовершеннолетних по возрасту, лет	Всего	В том числе		
		имели привод в полицию	состоят в полиции на учете	совершили преступления
До 13 лет	250,2	168,6	81,6	-
14-15	401,2	206,2	128,1	66,9
16-17	584,5	281,5	166,1	136,9
Итого	1235,9	656,3	375,8	203,8

В таблице 8 приведен пример сложной таблицы, где подлежащим являются группы эмитентов фондового рынка, распределенные по величине котировки банковских долгов и средневзвешенной ставке.

Виды таблиц по характеру сказуемого. Система показателей, которыми характеризуется объект изучения, т.е. подлежащее таблицы, образует сказуемое статистической таблицы. Сказуемое формирует заголовки граф и составляет их содержание.

По структурному строению сказуемого различают статистические таблицы с простой и сложной его разработками.

Таблица 8 – Распределение эмитентов фондового рынка по величине котировки банковских долгов и средневзвешенной ставке, выставленных на продажу в одном из вексельных центров в 2010 г. (данные условные)

Группа эмитентов по величине котировки банковского долга, млн. руб.	Подгруппа	Число эмитентов
97-1745	50-75 75-100	6 9
Итого по группе		15
1745-3393	50-75 75-100	2 2
Итого по группе		4
3393-5041	50-75 75-100	3 2
Итого по группе		5
Итого по подгруппам	50-75 75-100	11 13
Всего		24

По простой разработке сказуемого показатель, его определяющий, получается путем простого суммирования значений по каждому признаку отдельно независимо друг от друга.

Сложная разработка сказуемого предполагает деление признака, его формирующего, на группы.

Статистическая таблица со сложной комбинированной разработкой сказуемого содержит два связанных между собой признака: атрибутивный – качественный – категории застрахованных и количественный – страховая сумма.

Пример представлен в таблице 9.

Соблюдение правил построения и оформления статистических таблиц делает их основным средством представления, обработки и обобщения статистической информации о состоянии и развитии анализируемых социально-экономических явлений.

Таблица 9 – Распределение клиентов страховых компаний по категориям и страховым суммам в I квартале 2010 г. (данные условные)

Страховая компания	Всего клиентов, чел.	В том числе распределение клиентов по категориям и страховым суммам на одного застрахованного, тыс. руб.					
		руководители коммерческих структур		сотрудники предприятий, работающие в офисе		охранники, инкассаторы	
		5-15	свыше 15	5-15	свыше 15	5-15	свыше 15
1	444	195	180	13	12	23	21
2	390	150	180	12	15	15	18
3	595	210	300	26	10	21	28
4	352	125	175	10	12	14	16
5	522	200	250	10	15	22	25
6	320	110	110	28	28	22	22
7	480	200	200	15	20	20	25
Итого	3103	1190	1395	117	112	137	155

Задача 19

По данным статистических ежегодников и периодической печати подберите примеры статистических таблиц с перечисленными вариантами разработки сказуемого:

1) с простой разработкой сказуемого;

2) со сложной разработкой сказуемого по двум признакам.

Задача 20

Составьте макеты статистических таблиц, в которых разработка сказуемого будет произведена:

1) в статике;

2) в динамике;

3) в территориальном аспекте;

4) в пространственно-временном аспекте.

По данным статистических ежегодников и периодической печати подтвердите примерами каждый из видов таблиц.

Задача 21

Разработайте макет таблиц, характеризующих распределение численности занятого населения и безработных по семейному положению, и дайте заголовок таблицы. Укажите:

1) к какому виду таблицы относится макет;

2) его подлежащее и сказуемое;

3) признак группировки подлежащего.

Задача 22

Разработайте макет статистической таблицы, характеризующей зависимость успеваемости студентов вашей группы от посещаемости учебных занятий и занятости внеучебной деятельностью. Сформулируйте заголовок таблицы.

Укажите:

1) к какому виду таблицы относится макет;

2) название и вид разработки подлежащего и сказуемого;

3) группировочные признаки.

Задача 23

Разработайте макеты таблиц для характеристики:

1) населения РФ по полу и возрасту;

2) наиболее ликвидных акций на внебиржевом рынке;

3) предприятий какой-либо отрасли;

4) деятельности коммерческих банков;

5) деятельности страховых компаний России;

6) рынка государственных ценных бумаг.

Задача 24

Укажите подлежащее и сказуемое в таблицах 10, 11, 12, 13. Определите вид таблицы по характеру разработки ее подлежащего и сказуемого.

Таблица 10 – Иностранные инвестиции в экономику одной из стран мира (данные условные)

Год	Поступило инвестиций, всего, млрд. долл.	В том числе			В общем объеме инвестиций, %		
		прямые	портфельные	прочие	прямые	портфельные	прочие
2006	6,97	2,44	0,13	4,40	35,0	1,9	63,1
2007	12,29	5,33	0,68	6,28	43,4	5,5	51,1
2008	11,77	3,36	0,19	8,22	28,5	1,7	69,8
2009	9,56	4,26	0,03	5,27	44,6	0,3	55,1
2010	10,96	4,47	0,01	6,48	40,8	0,1	59,1

Таблица 11 – Распределение объема работ, выполненных по договорам строительного подряда, по формам собственности, в 2009-2010 гг. (данные условные)

Форма собственности строительных организаций и предприятий	2009 г.		2010 г.	
	всего, млрд. руб.	удельный вес в общем объеме работ, выполненных по договорам строительного подряда, %	всего, млрд. руб.	удельный вес в общем объеме работ, выполненных по договорам строительного подряда, %
Всего:	52,9	100,0	79,8	100,0
в том числе:				
государственная	5,0	9,5	8,1	10,1
муниципальная	0,1	0,1	0,1	0,1
частная	32,9	62,1	44,3	55,5
смешанная российская	14,9	28,3	27,3	34,3

Таблица 12 – Динамика основных экономических показателей промышленности в регионе (данные условные)

Показатель	2007 г.	2008 г.	2009 г.	2010 г.
Объем промышленной продукции, млрд. руб.	76,3	81,4	161,9	224,8
Уровень рентабельности активов предприятий и организаций, %	4,3	-2,7	5,0	7,5
Уровень рентабельности реализованной продукции предприятий и организаций, %	13,2	16,1	20,8	21,0
Изменение затрат на один рубль продукции, % к предыдущему году	15,0	-2,8	-3,0	2,5
Индекс промышленного производства, % к предыдущему году	99,0	103,0	114,0	115,0

Таблица 13 – Структура безработных в регионе по полу и возрасту

Группа безработных по возрасту, лет	2008 г.			2009 г.			2009 г.		
	всего	в т.ч.		всего	в т.ч.		всего	в т.ч.	
		женщины	мужчины		женщины	мужчины		женщины	мужчины
15-19	15,5	16,4	14,7	13,6	15,3	11,8	11,8	14,2	9,5
20-24	15,8	14,3	17,1	16,8	14,9	18,6	16,2	15,2	17,2
25-29	11,6	11,0	12,1	11,9	11,4	12,4	11,3	10,9	11,8
30-49	39,5	39,7	39,4	43,4	43,7	43,2	48,5	48,1	48,8
50-54	6,5	7,2	5,9	5,6	6,0	5,3	5,2	5,3	5,0
55-59	5,3	5,6	5,1	5,1	5,3	4,9	4,9	4,2	5,5
60 и старше	5,8	5,8	5,7	3,6	3,4	3,8	2,1	2,1	2,2
Итого	100	100	100	100	100	100	100	100	100

Задача 25

Известны следующие данные, представленные в таблице 14.

Определите и исправьте ошибки и недостатки, которые допущены в этой таблице.

Таблица 14 – Распределение численности занятого населения и безработных по семейному положению на конец 2010 г.

Категория населения	Состоят в браке	Холосты, не замужем	Вдовцы, вдовы	Разведены	Всего
Занято население – всего:	74,0	13,6	4,0	8,4	100
в т.ч.:					
мужчины	77,9	15,6	1,3	5,2	100
женщины	69,9	11,6	6,8	11,7	100
Безработные – всего:	54,7	30,0	3,2	12,1	100
в т.ч.:					
мужчины	52,6	34,6	1,3	11,5	100
женщины	57,0	25,1	5,2	12,7	100

Задача 26

Известны следующие данные представленные в таблице 15.

Таблица 15 – Воспроизводство капитальных вложений по объектам производственного назначения в 2010г., проц. к итогу

Направление капитальных вложений	Техническое перевооружение и реконструкция	Расширение действующих предприятий	Новое строительство	Отдельные объекты действующих предприятий	Всего
Капитальные вложения	63	9	15	10	97

Определите:

1) содержат ли данные таблицы ошибку и в чем она выражается;

2) логическим или арифметическим способом контроля можно установить ошибку.

Задача 27

Разработан следующий макет таблицы, который представлен в таблице 16.

Таблица 16 – Распределение населения по категориям занятости и по полу

Группа населения по категориям занятости	Группа населения по полу	Численность населения	
		всего, тыс. чел	% к итогу
Занятое население			
	Мужчины		
	Женщины		
Итого			
Безработные			
	Мужчины		
	Женщины		
Итого			
Всего населения по группам	Мужчины		
	Женщины		
Всего			

Укажите недостатки данного макета таблицы. Переработайте макет с учетом выявленных недостатков и укажите по нему подлежащее, сказуемое и вид таблицы по характеру их разработки.

4 ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ

Несмотря на многообразие видов графических изображений, каждый график должен включать следующие элементы: графический образ, поле графика, масштабные ориентиры и систему координат.

Графический образ – геометрические знаки, совокупность точек, линий, фигур, с помощью которых изображаются статистические величины.

Поле графика представляет собой пространство, в котором размещаются геометрические знаки.

Масштабные ориентиры определяются масштабом и масштабной шкалой.

Масштаб – это мера перевода числовой величины в графическую, а масштабная шкала – линия, определенные точки которой могут быть прочитаны как определенные числа.

Шкала состоит из линии (носителя шкалы) и ряда намеченных на ней точек, расположенных в определенном порядке. Носитель шкалы может быть представлен прямой или кривой линией.

Шкалы могут быть равномерными и неравномерными.

Для размещения геометрических знаков в поле графика необходима система координат. Наиболее распространенной является система прямоугольных координат.

Масштабом равномерной шкалы называется длина отрезка, принятого за единицу и измеренного в каких-либо мерах.

Рассмотрим построение основных видов диаграмм на конкретных числовых примерах.

На столбиковых диаграммах статистические данные изображаются в виде вытянутых по вертикали прямоугольников.

При построении столбиковых диаграмм необходимо выполнять следующие требования:

1) шкала, по которой устанавливается высота столбика, должна начинаться с нуля;

2) шкала должна быть непрерывной;

3) основания столбиков должны быть равны между собой.

Пример. Изобразим графически данные о числе дошкольных учреждений в Пензенской области (на конец года), ед.: 1991 – 904, 1998 – 502, 2005 – 360, 2009 г. – 355.

На горизонтальной оси откладываем произвольные основания четырех столбиков с произвольным расстоянием между основанием. Масштаб на вертикальной оси – 1 деление = 200 единиц (рисунок 2).

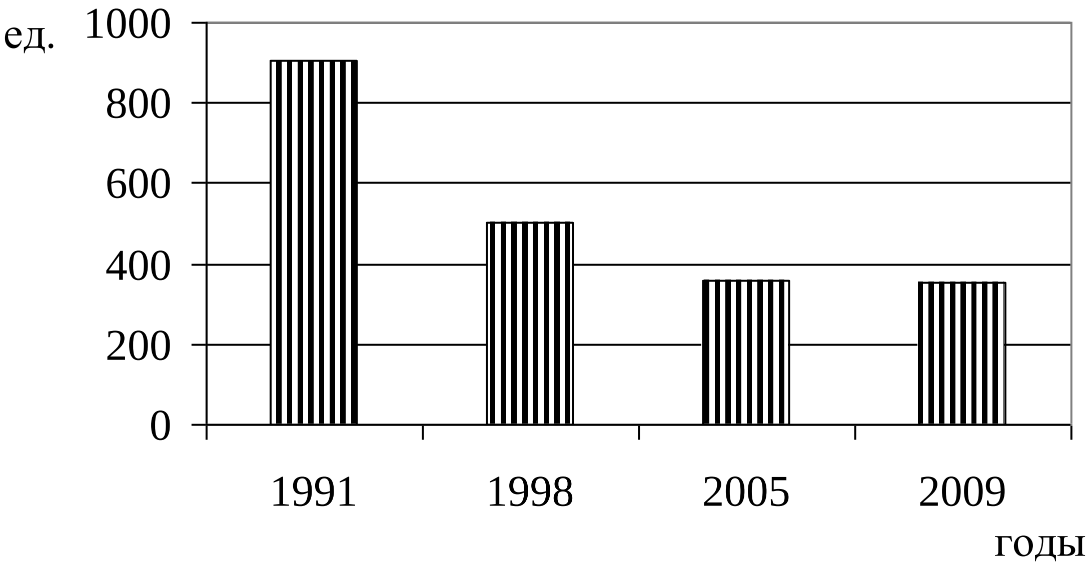

Рисунок 2 – Динамика дошкольных учреждений в Пензенской области (на конец года)

На столбиковой диаграмме изображаемые величины пропорциональны высоте столбиков.

Принципы построения полосовых или ленточных диаграмм абсолютно аналогичен. Разница заключается в том, что масштабной шкалой будет горизонтальная ось.

В круговых и квадратных диаграммах величина изображаемого явления выражается размером площади. Чтобы построить квадратную диаграмму, нужно из сравниваемых статистических величин извлечь квадратные корни, а затем построить квадраты со сторонами, пропорциональными полученным результатам.

Пример. Построим квадратную диаграмму о численности студентов в негосударственных средних специальных учебных заведениях в Пензенской области на начало учебного года: 2005/2006 – 411 чел., 2009/2010 – 112 чел.

Извлечем квадратные корни из показателей численности студентов. Это составит соответственно 20,2 и 10,6 чел. Выбираем масштаб. Примем 1 см за четыре человека. Тогда стороны первого квадрата составят 5,1 (20,2 : 4), а второго − 2,7 (10,6 : 4) (рисунок 3).

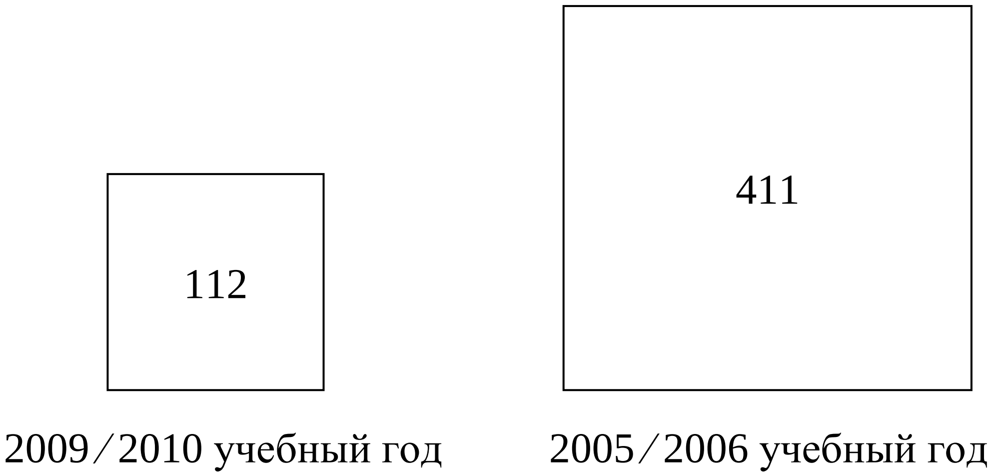

Рисунок 3 – Численность студентов в негосударственных средних специальных учебных заведениях в Пензенской области (на конец года, чел.)

Круговые диаграммы строятся аналогично. Разница в том, что на графике вычеркиваются круги, площади которых пропорциональны квадратным корням из изображаемых величин. Основываясь на данных предыдущего примера, необходимо построить два круга с радиусами 5,16 и 2,7 см.

Секторные диаграммы удобно стоить следующим образом: вся величина явления принимается за 100 %, рассчитываются доли отдельных его частей в процентах. Круг разбивается на секторы пропорционально частям изображаемого целого. Таким образом, на 1 % приходится 3,6°. Для получения центральных углов секторов, изображающих доли частей целого, необходимо их процентное выражение умножить на 3,6°.

Пример. Изобразим с помощью секторной диаграммы состав студентов государственных и муниципальных высших и средних специальных учебных заведений в Пензенской области в начале 2009/2010 учебного года по формам обучения. По очной форме обучения обучалось 51 %, по очно-заочной форме – 3 %, по заочной – 44 %, экстернат – 2 %. Построим круг произвольно-го радиуса. По данным о доле студентов по каждой форме обучения определим величину центральных углов: для очной формы – 183,6° (51:3,6) для очно-заочной – 10,8° (3:3,6), для заочной – 158,4° (44:3,6), для экстерната – 7,2° (2:3,6).

При помощи транспортира разделим круг на соответствующие сектора (рисунок 4).

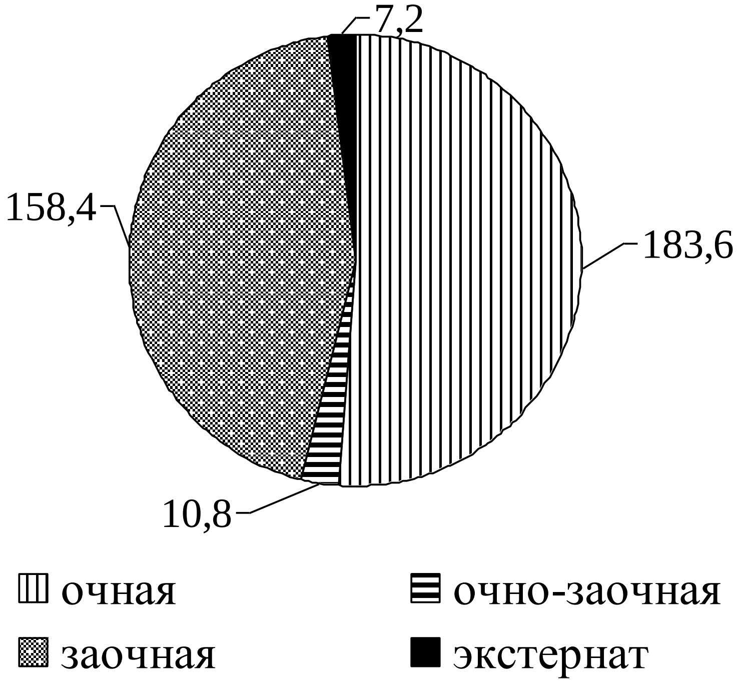

Рисунок 4 – Структура форм обучения студентов в государственных и муниципальных высших учебных заведениях в Пензенской области на начало 2007/2008 учебного года, проц.

Для одновременного сопоставления трех величин, связанных между собой таким образом, что одна величина является произведением двух других применяют диаграммы, называемые «знак Варзара».

Знак Варзара представляет собой прямоугольник, у которого один сомножитель принят за основание, другой – за высоту, а вся площадь равна произведению.

Пример. Имеются данные по Пензенской области о производстве зерновых культур в 2009 г. Урожайность составила 20,3 ц с 1 га, а посевная площадь 771,0 тыс. гектар.

Положим в основание прямоугольника урожайность, а в качестве высоты возьмем показатель посевной площади. Тогда площадь полученного прямоугольника будет пропорциональна валовому сбору зерновых культур (рисунок 5).

Линейные диаграммы широко применяются для характеристики изменений явлений во времени, выполнения плановых заданий, а также для изучения рядов распределения, выявления связи между явлениями. Линейные диаграммы строятся на координатной сетке. Геометрическими знаками в линейных диаграммах служат точки и последовательно соединяющие их отрезки прямой, которые складываются в ломанные кривые.

Рисунок 5 – Зависимость валового сбора зерновых культур в Пензенской области от урожайности и посевной площади в 2009 г.

Пример. Изобразим динамику урожайности зерновых культур в Пензенской области по данным таблицы 17.

Таблица 17 – Динамика урожайности зерновых культур за 2000-2009 гг. в хозяйствах всех категорий в Пензенской области, ц с 1 га

2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
10,0	13,8	14,8	13,7	13,4	13,5	15,0	12,8	19,3	20,3

В прямоугольной системе координат на ось ординат нанесем данные об урожайности зерновых культур, на ось абсцисс – данные о периодах времени (рисунок 6).

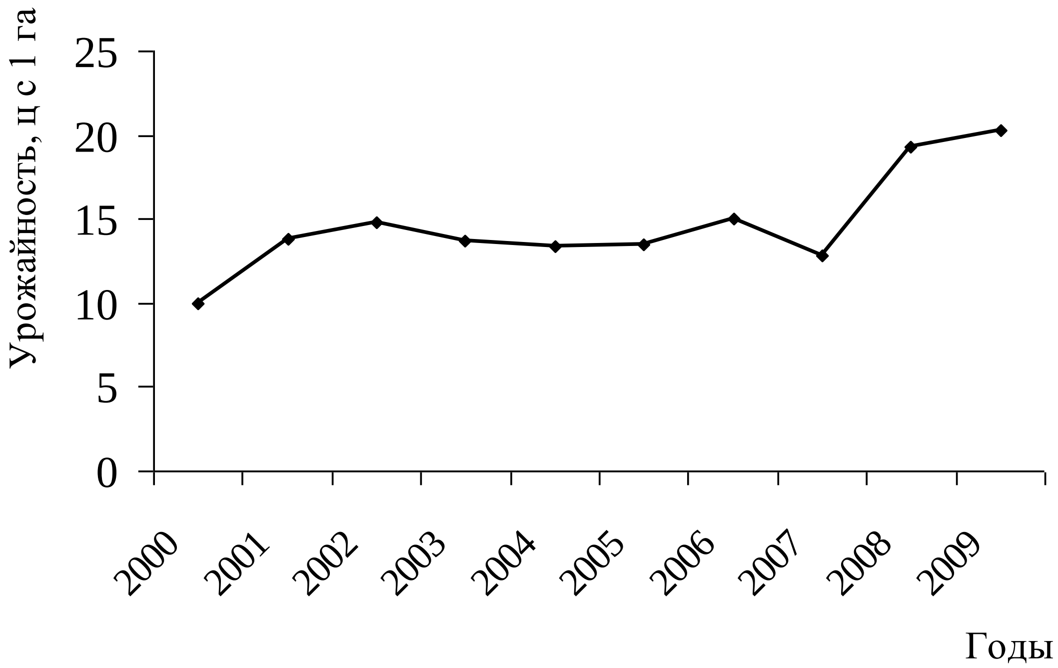

Рисунок 6 – Динамика урожайности зерновых культур в Пензенской области в хозяйствах всех категорий

Ряды распределения чаще всего изображаются в виде полигона или гистограммы. Полигон строят в основном для изображения дискретных рядов (рисунок 7). При его построении на оси абсцисс откладываются значения варьирующего признака, а на оси ординат – абсолютные или относительные численности единиц совокупности (частоты или частости).

Полигон на рисунке 7 построен на основании (условных) данных о распределении семей по числу детей.

Гистограмма распределения применяется чаще всего для изображения интервальных рядов. Для ее построения по оси абсцисс откладываются интервалы признака, а по оси ординат – численности единиц совокупности. На отрезках, изображающих интервалы, стоят прямоугольники, площади которых пропорциональны численностям единиц (рисунок 8).

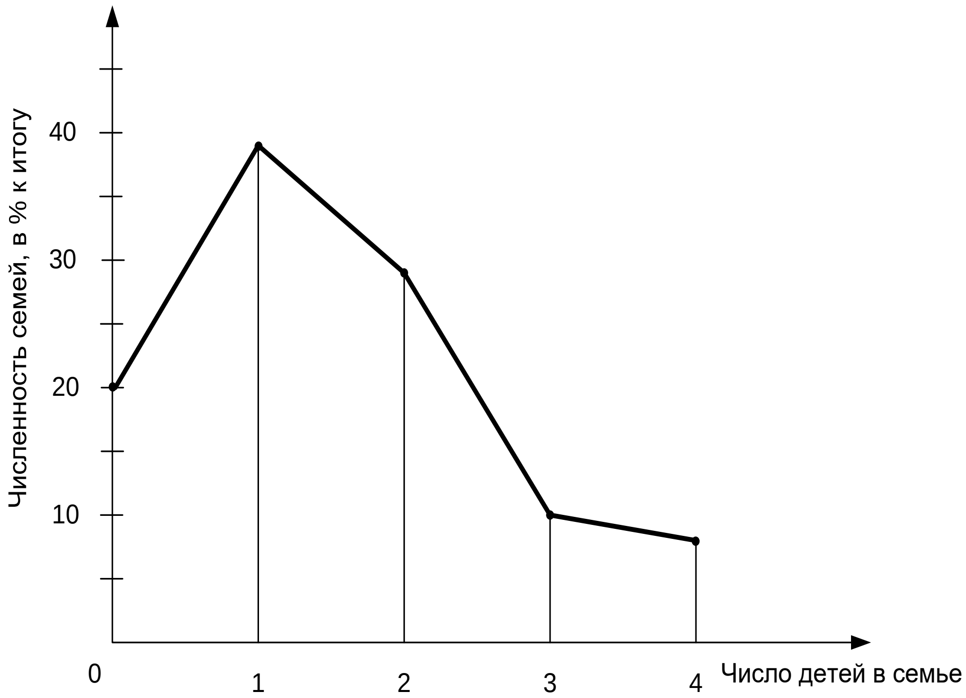

Рисунок 7 – Полигон распределения семей по числу детей в одном из регионов в 2010 г.

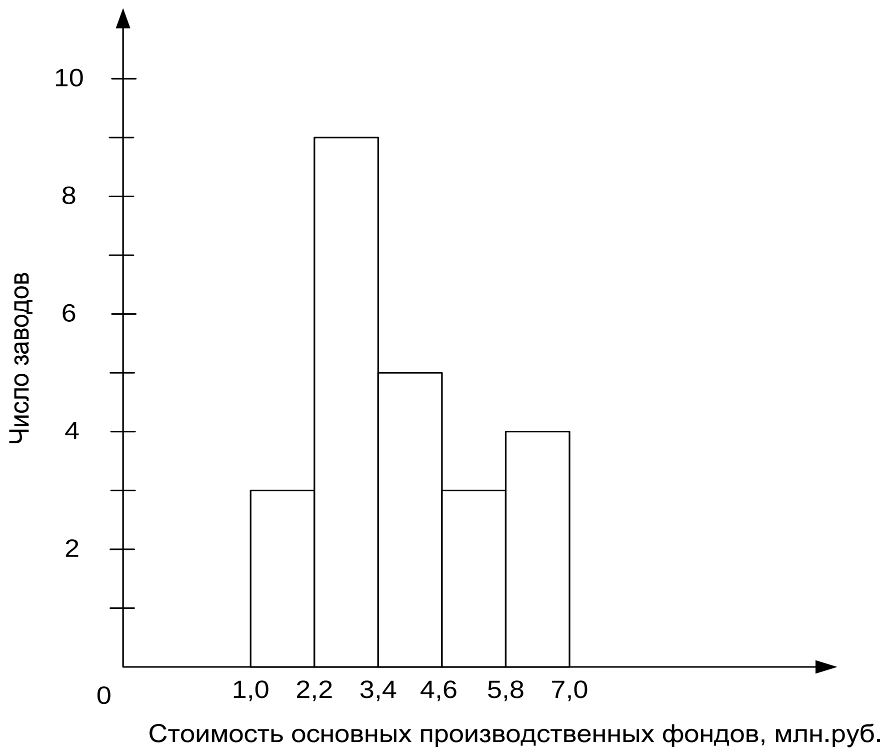

Рисунок 8 – Гистограмма распределения фирм в одной из отраслей по стоимости основных производственных фондов в 2010 г.

Пример. Имеются следующие условные данные.

Таблица 18 – Продажа моркови на рынках сельхозпродуктов города в 2010 г., ц

Месяц	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	Итого за год
Морковь	36	42	44	54	43	70	41	43	39	37	37	34	520

Определяем среднемесячную продажу моркови. Она составляет 43,3 ц. Вычертим круг радиусом, равным среднемесячному показателю (R = 43,3 ц). На горизонтальном диаметре построим шкалу, взяв длину радиуса, равную 2,7 см. Следовательно, 1 см = 43,3 / 2,7 ≈ 16 ц. Затем весь круг разделим на 12 радиусов (соответственно числу месяцев в году). На радиусе сделаем отметку согласно масштабу исходя из приведенных данных за каждый месяц. Данные, которые превысили среднемесячный уровень, отмечаются за пределами окружности на продолжении радиуса. Отметки различных месяцев соединяются между собой (рисунок 9).

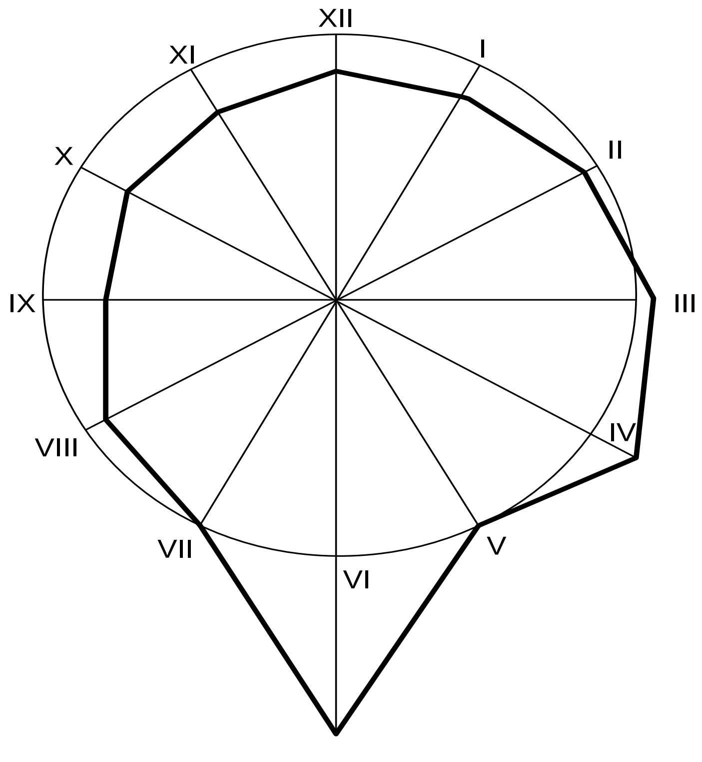

Рисунок 9 – Продажа моркови на рынках сельхозпродуктов в одном из городов в 2010 г., ц

Статистические карты представляют собой вид графических изображений статистических данных на схематический географической карте и характеризуют уровень или степень распространения того или иного явления на определенной территории. При этом различают картограммы и картодиаграммы.

Картограмма – это схематическая карта, на которой штриховкой различной густоты, точками или окраской определенной степени насыщенности показывается сравнительная интенсивность какого-либо показателя в пределах каждой единицы нанесенного на карту территориального деления (например, плотность населения по области, распределение районов по урожайности и т.п.).

Картодиаграмма представляет собой сочетание диаграммы с географической картой. В качестве изобразительных знаков в картограмме используют диаграммные фигуры (столбики, квадраты, круги, фигуры, полосы), которые размещаются на контуре географической карты. Картодиаграммы дают возможность географически отразить более сложные статистико-географические построения, чем картограммы.

Задача 28

По данным таблицы 19 отразите с помощью полосовой диаграммы данные о грузообороте автомобильного транспорта.

Таблица 19 – Перевезено грузов и грузооборот автомобильного транспорта организаций всех видов деятельности в Пензенской области (включая объемы по малым предприятиям, и без учета неформальной деятельности)

Год	Перевезено грузов – всего, млн. т	Грузооборот – всего, млн. т/км
2004	15,4	619,5
2005	15,1	668,5
2006	16,0	804,3
2007	19,9	989,3
2008	27,8	1526,5
2009	14,8	1144,5

Задача 29

По данным таблицы 20 отразите валовые сборы сельскохозяйственных культур столбиковыми диаграммами.

Таблица 20 – Валовые сборы основных сельскохозяйственных культур в хозяйствах всех категорий в Пензенской области, тыс.т

	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
Зерновые культуры (в весе после доработки)	1046	1170	837	907	993	1108	932	1422	1461
Сахарная свекла (фабричная)	392	355	698	536	678	1147	923	1091	929
Семена подсолнечника	20	24	32	28	41	38	39	46	67
Картофель	448	405	513	496	488	432	410	427	479
Овощи (включая овощи закрытого грунта)	156	153	155	146	148	133	130	139	142

Задача 30

По данным таблицы 21 отразите динамику в области с помощью квадратных, круговых или фигурных диаграмм.

Таблица 21 – Число построенных квартир в Пензенской области, ед.

2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
2994	2705	3539	3836	4884	6451	6902	7101

Задача 31

По данным таблицы 22 постройте диаграммы, отражающие структуру расходов бюджета Пензенской области на социальные нужды.

Таблица 22 – Расходы бюджета Пензенской области на социально-культурные мероприятия в 2009 г.

Показатель	Млн. руб.	В % к итогу
Расходы – всего,	21103,3	100,0
из них: образование	8957,6	43
культура, искусство и кинематография, средства массовой информации	1975,5	9
здравоохранение и спорт	4688,8	22
социальная политика	5481,4	26

Задача 32

По данным таблицы 23 отразите динамику и структуру земельного фонда графически.

Таблица 23 – Общий земельный фонд Пензенской области

Показатель	1991 г.	1998 г.	2000 г.	2005 г.	2007 г.	2009 г.
Общая площадь земли	4335,2	4335,2	4335,2	4335,2	4335,2	4335,2
Сельскохозяйственные угодья	3067,3	3058,8	3047,6	3046,4	3044,4	3042,9
в т.ч.: пашня	2502,3	2376,5	2213,3	2235,8	2243,7	2253,4
залежь	3,4	92,0	214,2	185,6	176,1	165,4
многолетние насаждения	21,7	22,8	22,9	22,5	22,4	22,4
кормовые угодья	540,3	567,5	597,2	602,5	602,2	601,7
Несельскохозяйственные угодья (леса, кустарники, болота, земли под водой и др.)	1267,9	1276,4	1287,6	1288,8	1290,8	1292,3
Структура сельскохозяйственных угодий, %	100	100	100	100	100	100
пашня	81,6	77,7	72,6	73,4	73,7	74,1
залежь	0,1	3,0	7,0	8,1	5,8	5,4
многолетние насаждения	0,7	0,8	0,8	0,7	0,7	0,7
кормовые угодья	17,6	18,5	19,6	19,8	19,8	19,8

Задача 33

Отразите данные о динамике цен на отдельные товары народного потребления с помощью линейного графика по данным таблицы 24.

Таблицы 24 – Индексы цен на отдельные товары народного потребления в Пензенской области (декабрь к декабрю предыдущего года; в процентах)

Показатель	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2209 г.
Все товары	119,0	118,5	115,1	109,3	113,9	109,1	108,6	113,0	114,8	107,6
Продовольственные товары	120,7	115,1	110,5	109,6	116,2	109,6	108,4	115,4	117,7	105,9
Мясопродукты	132,2	123,8	102,9	108,7	124,8	116,4	106,7	106,6	123,8	105,1
Рыбопродукты	121,2	126,6	110,2	105,0	113,7	115,3	105,6	103,0	117,0	110,3
Молоко и молочная продукция	124,5	110,4	108,3	114,6	121,6	111,5	107,7	126,0	113,9	104,3
Непродовольственные товары	126,3	100,6	109,5	108,4	109,0	108,2	109,1	109,3	110,5	109,9
Автомобили легковые	102,3	105,7	106,9	108,1	105,5	104,4	108,0	103,3	104,8	108,2
Меха и меховые изделия	119,7	111,9	106,0	108,4	103,8	107,8	109,7	109,4	112,2	104,2
Строительные материалы	123,3	115,6	112,6	109,7	108,1	113,5	115,1	128,9	120,5	97,9

Задача 34

По данным таблицы 25 постройте линейные графики.

Таблица 25 – Динамика численности постоянного населения в отдельных областях, тыс. чел. на начало года

	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
Пензенская область	1449,2	1436,0	1422,7	1408,0	1396,0	1388,0	1379,8	1373,2
Ульяновская область	1378,9	1364,5	1350,7	1335,9	1321,7	1312,2	1305,0	1298,6
Самарская область	3235,7	3217,6	3201,2	3189,0	3178,6	3172,8	3171,4	3170,1

Задача 35

По данным таблицы 26 отразите структуру затрат графически.

Таблица 26 – Структура затрат на производство продукции сельского хозяйства в крупных и средних сельскохозяйственных организациях Пензенской области (в процентах)

Год	Затраты, всего	Из них:			
		оплата труда с начислениями	материальные затраты	амортизация основных средств	прочие
1995	100	17	54	18	11
2000	100	20	66	8	6
2001	100	21	67	6	6
2002	100	23	67	5	5
2003	100	23	66	5	6
2004	100	20	69	4	7
2005	100	19	70	5	6
2006	100	19	70	5	6
2007	100	18	70	6	5
2008	100	17	71	6	6

Задача 36

По данным таблицы 27 постройте радиальную диаграмму по данным о производстве шоколада и шоколадных изделий в кондитерском объединении.

Таблица 27 – Производство шоколада и шоколадных изделий в кондитерском объединении по месяцам в 2010 г., тонн

I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII

970	880	974	1010	850	930	460	730	947	965	880	920

Задача 37

Постройте знак Варзара по данным таблицы 28.

Таблица 28 – Расход кормов скоту и птице в сельскохозяйственных организациях Пензенской области

	1991 г.	1995 г.	2000 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
Всего кормов*, тыс. тонн	2426	1634,9	681,1	620,6	588,5	584,3	563,7	580,0	606,3	594,8	626,2
Расход кормов в расчете на одну голову условного скота, ц к. ед.	31,1	32,3	28,2	29,4	27,5	30,1	30,3	29,3	28,8	28,0	27,6

* в пересчете на кормовые единицы

Задача 38

В таблице 29 отражена продажа основных продуктов на рынках города по месяцам 2010 г. Постройте радиальные диаграммы по каждому виду продуктов питания.

Таблица 29 – Продажа основных продуктов питания на рынке города

Продукты	Месяц											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
А	1	2	3	4	5	6	7	8	9	10	11	12
Картофель, тыс. т	2,4	3,7	3,8	5,5	5,2	9,7	15	12	14	11	6,6	6,5
Овощи, тыс. т	2,2	2,9	3,3	4,1	8,4	7,9	20	16	16	6,5	3,6	2,9

Окончание таблицы 29

А	1	2	3	4	5	6	7	8	9	10	11	12
Мясо, т	186	168	175	215	216	167	125	146	154	246	317	234
Плоды, ягоды, виноград, тыс. т	30	40	43	54	67	29	35	34	45	35	29	29

Задача 39

По данным таблицы 30 постройте спиральную диаграмму.

Таблица 30 – Продажа кондитерских изделий в магазинах города (данные условные)

Месяц	2008 г.	2009 г.	2010 г.
Январь	365	373	420
Февраль	412	305	450
Март	346	336	416
Апрель	405	457	479
Май	475	517	506
Июнь	504	543	601
Июль	407	438	501
Август	367	440	520
Сентябрь	448	427	459
Октябрь	443	388	525
Ноябрь	415	401	498
Декабрь	379	387	481

Задача 40

В таблице 31 представлены данные по 16 районам области о посевных площадях озимого ячменя и его урожайности.

Постройте по этим данным:

1) картограмму с помощью штриховки для характеристики изменения урожайности в районах области;

2) точечную картограмму для характеристики размещения посевов ячменя в районах.

Указание. Схематическую карту области и размещение на ней районов сделайте по собственному усмотрении.

Таблица 31 – Посевные площади и урожайность озимого ячменя в районах области в 2010 г.

Номер района	Посевная площадь, тыс. га	Урожайность озимого ячменя, ц с 1 га	Номер района	Посевная площадь, тыс. га	Урожайность озимого ячменя, ц с 1 га
1	14,1	17,5	9	15,9	31,6
2	9,2	20,1	10	2,6	18,1
3	10,2	36,1	11	9,3	24,3
4	3,1	27,2	12	17,4	26,3
5	3,3	28,1	13	19,9	28,2
6	2,4	16,1	14	21,7	22,5
7	11,1	16,4	15	12,1	19,5
8	9,9	32,3	16	4,1	16,9

Задача 41

В таблице 32 отражены данные о производстве сельскохозяйственной продукции в области.

Таблица 32 – Производство сельскохозяйственной продукции в области в 2010 г., тыс. ц

Номер района	Молоко	Мясо в живом весе скота	
		крупного рогатого	свиней
1	14,8	1,7	13,9
2	14,5	1,6	13,8
3	58,0	7,7	10,3
4	40,1	4,5	5,5
5	15,0	1,6	0,7
6	14,5	0,8	0,9
7	37,7	4,5	8,4
8	38,9	3,4	9,2
9	46,8	5,4	15,5
10	44,8	4,4	11,5

Постройте картодиаграмму, изобразив:

1) производство зерна с помощью столбиковых диаграмм;

2) производство молока при помощи квадратных диаграмм;

3) производство мяса в живом весе с помощью круговых диаграмм.

Указание. Схематическую карту области постройте произвольно.

5 СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

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.

Абсолютные показатели. Данные показатели отражают физические размеры изучаемых статистикой процессов и явлений, а именно их массу, площадь, объем, протяженность, временные характеристики, а также могут представлять объем совокупности, т.е. число составляющих единиц. К абсолютным показателям, например, относятся площадь территории страны, объем промышленного производства, эксплуатационная длина железнодорожных путей сообщения, число предприятия отрасли и т.п.

Абсолютные статистические показатели всегда являются именованными числами. В зависимости от социально-экономической сущности исследуемых явлений, их физических свойств они выражаются в натуральных, стоимостных или трудовых единицах измерения.

В международной практике используются такие *натуральные единицы измерения*, как тонны, килограммы, квадратные, кубические и простые метры, километры, мили, литры, баррели, штуки и т.д.

В группу натуральных также входят условно-натуральные измерители, которые используются в тех случаях, когда какой-либо продукт имеет несколько разновидностей и общий объем можно определить только исходя из общего для всех разновидностей потребительского свойства.

В условиях рыночной экономики особое значение имеют *стоимостные единицы измерения*, позволяющие дать денежную оценку социально-экономическим объектам и явлениям.

К *трудовым единицам измерения*, позволяющим учитывать как общие затраты труда на предприятии, так и трудоемкость отдельных операций технологического процесса, относятся человеко-дни и человеко-часы.

В статистической практике для аналитических целей широко применяются относительные показатели.

Относительные показатели. Они представляют собой результат деления одного абсолютного показателя на другой и выражают соотношение между количественными характеристиками социально-экономических процессов и явлений. Поэтому по отношению к абсолютным показателям относительные показатели, или показатели в форме относительных величин, являются производными, вторичными.

При расчете относительного показателя абсолютный показатель, находящийся в числителе получаемого отношения, называется текущим, или сравнительным. Показатель же, с которым производится сравнение и который находится в знаменателе, называется основанием, или базой сравнения. Таким образом, рассчитываемый относительный показатель указывает, во сколько раз сравниваемый абсолютный показатель больше базисного, или какую долю он составляет от базисного показателя, или сколько единиц первого приходится на 1, 100, 1000 и т.д. единиц второго. Относительный показатель может выражаться в коэффициентах, процентах, промилле, продецимилле или быть именованным числом.

Все используемые на практике относительные статистические показатели можно подразделить на следующие виды: показатели динамики, плана, реализации плана, структуры, координации, интенсивности, уровня экономического развития, сравнения.

Относительный показатель динамики (ОПД) представляет собой отношение уровня исследуемого процесса или явления за данный период времени (по состоянию на данный момент времени) и уровня этого же процесса или явления в прошлом:

$$ОПД = \frac{\text{Текущий уровень}}{\text{Предшествующий или базисный уровень}}$$

Рассчитанная таким образом величина показывает, во сколько раз текущий уровень превышает предшествующий (базисный) или какую долю от последнего он составляет. Данный показатель может быть выражен кратным отношением или переведен в проценты.

Различают относительные показатели динамики с постоянной и переменной базой сравнения. Если сравнение осуществляется с одним и тем же базисным уровнем, например первым годом рассматриваемого периода, получают относительные показатели динамики с постоянной базой (базисные). При расчете относительных показателей динамики с переменной базой (цепных) сравнение осуществляется с предшествующим уровнем, т.е. основание относительной величины последовательно меняется.

В таблице 33 приведено производство сахара-песка в Пензенской области.

Таблица 33 – Производство сахара-песка в Пензенской области, тыс. т

2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
287	212	224	207	181

Рассчитаем относительные показатели динамики с переменной и постоянной базой сравнения и отразим результаты в таблице 34.

Таблица 34 – Расчет показателей динамики производства сахара-песка в Пензенской области

Переменная база сравнения (цепные показатели)	Постоянная база сравнения (базисные показатели)
$\frac{212}{287}\times100=73{,}9\ \%$ $\frac{224}{212}\times100=105{,}7\ \%$ $\frac{207}{224}\times100=92{,}4\ \%$ $\frac{181}{207}\times100=87{,}4\ \%$	$\frac{212}{287}\times100=73{,}9\ \%$ $\frac{224}{287}\times100=78{,}0\ \%$ $\frac{207}{287}\times100=72{,}1\ \%$ $\frac{181}{287}\times100=63{,}1\ \%$

Относительные показатели динамики с переменной и постоянной базой сравнения взаимосвязаны между собой следующим образом: произведение всех относительных показателей с переменной базой равно относительному показателю с постоянной базой за исследуемый период. Так, для рассчитанных показателей получим:

$$0{,}739 \times 1{,}057 \times 0{,}924 \times 0{,}874 = 0{,}631$$

Относительные показатели выполнения планового задания. Все субъекты финансово-хозяйственной деятельности (от малых предприятий и до крупных корпораций) в той или иной степени осуществляют как текущее, так и стратегическое планирование, а также сравнивают реально достигнутые результаты с ранее намеченными. Для этой цели используются относительные показатели выполнения плана (ОПВП) и планового задания (ОППЗ).

$$ОПВП = \frac{Фактический\ уровень\ текущего\ периода}{Плановый\ уровень\ на\ этот\ же\ период}$$

$$ОППЗ = \frac{Плановый\ уровень\ на\ текущий\ период}{Фактически\ достигнутый\ уровень\ в\ базисном\ периоде}$$

Первый показатель отражает фактический объем производства в процентах или коэффициентах по сравнению с плановым уровнем, второй – какой объем производства был запланирован от уровня, фактически произведенного в сравниваемом периоде. Например, оборот торговой фирмы в 2009 г. составил 2,0 млн. руб., на следующий год руководство фирмы считает реальным довести оборот до 2,8 млн. рублей. Фактический оборот фирмы в 2010 г. составил 2,6 млн. рублей. Тогда относительный показатель выполнения плана составил 92,9 % $\left(\frac{2{,}6}{2{,}8}\times 100\right)$, а относительный показатель планового задания 140 % $\left(\frac{2{,}8}{2{,}0}\times 100\right)$.

Между рассмотренными относительными величинами существует следующая зависимость:

$$ОПД = ОПВД \times ОППЗ$$

Относительный показатель структуры (ОПС) представляет собой соотношение структурных частей изучаемого объекта и их целого:

$$ОПС = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель по всей совокупности в целом}}$$

Выражается относительный показатель структуры в долях единицы или в процентах. Рассчитанные величины, соответственно называемые долями или удельными весами, показывают, какой долей обладает или какой удельный вес имеет та или иная часть в общем итоге.

В таблице 35 представлены исходные данные и расчет относительных показателей структуры.

Таблица 35 – Поголовье крупного рогатого скота по формам собственности в Пензенской области (на 1 января 2010 г.)

Показатель	Тыс. гол.	% к итогу
Все формы собственности	297,9	100
в том числе: государственная	9,0	$\frac{9{,}0}{297{,}9}\times 100 = 3{,}02$
муниципальная	2,4	$\frac{2{,}4}{297{,}9}\times 100 = 0{,}81$
частная	286,5	$\frac{286{,}5}{297{,}2}\times 100 = 96{,}17$

Относительный показатель координации (*ОПК*) представляет собой отношение одной части к другой части этой же совокупности:

$$ОПК = \frac{\begin{array}{c}\textit{Показатель, характеризующий} \\ \textit{i – ую часть совокупности}\end{array}}{\begin{array}{l}\textit{Показатель, характеризующий часть совокупности,} \\ \textit{выбранную в качестве базы сравнения}\end{array}}$$

При этом в качестве базы сравнения выбирается та часть, которая имеет наибольший удельный вес или является приоритетной с экономической, социальной или какой-либо другой точки зрения. В результате получают, во сколько раз данная часть больше базисной, или сколько процентов от нее составляет, или сколько единиц данной структурной части приходится на одну единицу (иногда – на 100, 1000 и т.д. единиц) базисной структурной части.

По данным предыдущей таблицы вычислим, что на каждую голову скота, находящуюся в государственной форме собственности приходилось 32 головы скота частной формы собственности (286,5: 9).

Относительный показатель интенсивности (*ОПИ*) характеризует степень распространения изучаемого процесса или явления и представляет собой отношение исследуемого показателя к размеру присущей ему среды:

$$ОПИ = \frac{\textit{Показатель, характеризующий явление А}}{\begin{array}{l}\textit{Показатель, характеризующий} \\ \textit{среду распространения явлений А}\end{array}}$$

Данный показатель получают сопоставлением разноименных, но взаимосвязанных в своем развитии величин. Поэтому наиболее часто он представляет собой именованную величину, но может быть выражен и в процентах, промилле, продецимилле.

Например, для определения уровня обеспеченности населения легковыми автомобилями рассчитывается число автомашин, приходящихся на 100 семей, для определения плотности населения рассчитывается число людей, приходящихся на 1 км^2.

Разновидностью относительных показателей интенсивности являются относительные показатели уровня экономического развития, характеризующие производство продукции в расчете на душу населения и играющие важную роль в оценке развития экономики государства. Так как объемные показатели производства по своей природе являются интервальными, а показатель численности населения – моментным, в расчете используют среднюю за период численности населения (например, среднегодовую).

К относительным показателям уровня экономического развития относятся показатели, характеризующие производство угля, стали, нефти, газа на душу населения, производство зерна, молока, мяса на душу населения и т.п.

Относительный показатель сравнения (*ОПС*) представляет собой соотношение одного и того же абсолютного показателя, характеризующего разные объекты (предприятия, фирмы, районы, области, страны и т.п.):

$$ОПС = \frac{Показатель, характеризующий\ объект\ А}{Показатель, характеризующий\ объект\ Б}$$

Например, имеются данные о среднедушевых ежемесячных доходах населения за 2009 г., руб.: Пензенская область – 11535, Самарская область – 18175, Саратовская область – 10356, Ульяновская область – 7450. Таким образом, среднедушевой доход населения в Самарской области превышал аналогичный показатель по Пензенской области – в 1,58 раза $\left(\frac{18175}{11535}\right)$, в Саратовской области – в 1,76 раза $\left(\frac{18175}{10356}\right)$, в Ульяновской области – в 2,44 раза $\left(\frac{18175}{7450}\right)$.

Средние показатели. Наиболее распространенной формой статистических показателей, используемой в социально-экономических исследованиях, является средняя величина, представляющая собой обобщенную количественную характеристику признака в статистической совокупности в конкретных условиях места и времени. Показатель в форме средней величины выражает типичные черты и дает обобщающую характеристику однотипных явлений по одному из варьирующих признаков. Он отражает уровень этого признака, отнесенный к единице совокупности. Широкое применение средних объясняется тем, что они имеют ряд положительных свойств, делающих их незаменимыми в анализе явлений и процессов общественной жизни.

Важнейшее свойство средней заключается в том, что она отражает то общее, что присуще всем единицам исследуемой совокупности. Значения признака отдельных единиц совокупности варьируют под влиянием множества факторов, среди которых могут быть как основные, так и случайные.

Сущность средней в том и заключается, что в ней взаимопогашаются те отклонения значений признака, которые обусловлены действием случайных факторов, и учитываются изменения, вызванные действием факторов основных. Это позволяет средней отражать типичный уровень признака и абстрагироваться от индивидуальных особенностей, присущих отдельным единицам.

Для вычисления средних величин используется формула степенной средней:

$$\bar{x} = \sqrt[k]{\frac{\sum x_i^k \times f_i}{\sum f_i}},$$

где $\bar{x}$ – средняя величина исследуемого явления;

x_i – i-й вариант осредняемого признака ($i = 1n$);

f_i – вес i-го варианта.

Формулы средних величин, используемых в статистике, представлены в таблице 36.

Для вычисления среднего значения качественного признака (средняя урожайность, средняя себестоимость и т.д.), как правило используется средняя арифметическая или средняя гармоническая. Чтобы правильно выбрать формулу расчета, необходимо составить логическую формулу средней или исходное соотношение средней (ИСС):

$$ИСС = \frac{Объем\ осредняемого\ признака}{Объем\ совокупности}$$

Таблица 36 – Формулы средних величин, используемых в статистике

Значение k	Наименование средней	Формула средней	
		простая	взвешенная
-1	Средняя гармоническая	$\bar{x}=\dfrac{n}{\sum\frac{1}{x}}$	$\bar{x}=\dfrac{\sum f}{\sum\frac{f}{x}}$
0	Средняя геометрическая	$\bar{x}=\sqrt[n]{x_1\times x_2\times x_3\times\ldots x_n}$	$\bar{x}=\sqrt[\sum f]{x_1^{f1}\times x_2^{f2}\times x_3^{f3}\times\ldots x_n^{fn}}$
1	Средняя арифметическая	$\bar{x}=\dfrac{\sum x}{n}$	$\bar{x}=\dfrac{\sum xf}{\sum f}$
2	Средняя квадратическая	$\bar{x}=\dfrac{\sum x^2}{n}$	$\bar{x}=\sqrt{\dfrac{\sum x^2 f}{\sum f}}$

Средняя геометрическая используется для расчета среднего темпа роста в анализе динамики, а ряд статистических показателей, характеризующих вариацию и взаимосвязь, базируется на средней квадратической и степенных средних более высоких порядков.

Рассмотрим последовательность вычисления средней заработной платы по данным таблицы 37.

Таблица 37 – Заработная плата предприятий АО

Предприятие	Численность промышленно-производственного персонала, чел.	Месячный фонд заработной платы, тыс. руб.	Средняя заработная плата, руб.
А	1	2	3
1	270	564,84	2092
2	121	332,75	2750
3	229	517,54	2260
Итого	620	1415,13	?

Определим исходное соотношение средней для показателя «средняя заработная плата». Независимо от имеющихся в нашем распоряжении данных средняя заработная плата может быть получена только через следующее отношение:

$$ИСС = \frac{Совокупный\ фонд\ заработной\ платы}{Общая\ численность\ персонала}.$$

Если мы располагаем только данными о средней заработной плате и численности работников (гр. 1 и 3), то нам известен знаменатель исходного соотношения, но не известен его числитель. Однако фонд заработной платы можно получить умножением средней заработной платы на численность *ППП*. Поэтому общая средняя может быть рассчитана по формуле средней арифметической взвешенной:

$$\bar{x} = \sqrt{\frac{\sum x_i f_i}{\sum f_i}} = \frac{2092 \times 270 + 2750 \times 121 + 2260 \times 229}{270 + 121 + 229} = 2282\ руб.$$

Необходимо учитывать, что вес *(f)* в отдельных случаях может представлять собой произведение двух или даже трех значений.

Допустим теперь, что в нашем распоряжении только данные о фонде заработной платы и средней заработной плате персонала (гр. 2 и 3, табл. 37), т.е. нам известен числитель исходного соотношения, но не известен его знаменатель. Численность работников по каждому предприятию можно получить делением фонда заработной платы на среднюю заработной платы. Тогда расчет средней заработной платы в целом по трем предприятиям будет произведен по формуле средней гармонической взвешенной:

$$\bar{x} = \sqrt{\frac{\sum w_i}{\sum \frac{w_i}{x_i}}} = \frac{564840 + 332750 + 517540}{\frac{564840}{2092} + \frac{332750}{2750} + \frac{517540}{2260}} = 2282\ руб.$$

В нашем примере мы использовали разные формы средних, но получили один и тот же ответ. Это обусловлено тем, что для конкретных данных каждый раз реализовывалось одно и то же исходное соотношение средней.

Средние показатели могут рассчитываться по дискретным и интервальным вариационным рядам. При этом расчет производится по средней арифметической взвешенной. Для дискретного ряда данная формула используется так же, как и в приведенном выше примере. В интервальном же ряду для расчета определяются середины интервалов.

По данным таблицы 38 определим величину среднедушевого дохода в Пензенской области.

Таблица 38 – Распределение населения Пензенской области по размеру среднедушевых денежных доходов в 2008 г.

Группа населения по уровню доходов в месяц, руб.	Численность населения, % к итогу
До 1000	0,1
1000–1500	0,4
1500–2000	1,1
2000–3000	4,3
3000–4000	6,5
4000–5000	7,8
5000–7000	15,9
7000–12000	29,9
Свыше 12000	34,0
Итого	100

Последовательность вычисления будет следующая:

1. Запишем исходное логическое соотношение для данной средней:

$$\text{Среднедушевой денежный доход} = \frac{\text{Совокупные денежные доходы всего населения}}{\text{Общая численность населения}}$$

2. Обозначим значение осредняемого признака (среднедушевой денежный доход в среднем за месяц) через X, а частоту повторения данного признака (численность населения, % к итогу) через f.

3. Так как значения осредняемого признака заданы в виде интервалов, то найдем их середины, т.е. $X^1 = \frac{X_H - X_B}{2}$. При этом величину первого интервала условно приравниваем к величине второго, тогда его нижняя граница будет равна 750. Величину последнего интервала условно приравниваем к величине предпоследнего, тогда его верхняя граница составит 14500 руб. В результате получаем следующие середины интервалов (X^1):

750; 1250; 1750; 2500; 3500; 4500; 6000; 9500; 14500

4. Роль численности населения в данном случае выполняет его доля в общем итоге, выраженная в процентах. Для расчета воспользуемся средней арифметической взвешенной:

$$\bar{x} = \frac{\sum x_i^1 f_i}{\sum f_i} = \frac{750\times 0{,}1 + 1250\times 0{,}4 + 1750\times 1{,}1 + 2500\times 4{,}3 + 3500\times 6{,}5 +}{0{,}1 + 0{,}7 + 1{,}8 + 5{,}5 +}$$

$$\frac{+ 4500\times 7{,}8 + 6000\times 15{,}9 + 9500\times 29{,}9 + 14500\times 34{,}0}{+ 7{,}8 + 8{,}9 + 17{,}3 + 29{,}7 + 28{,}4} =$$

$$= \frac{943550}{100} = 9435{,}5 \text{ } \textit{руб.}$$

Следовательно, среднедушевой денежный доход населения в Пензенской области в 2009 г. составил 9435,5 рублей.

Задача 42

По данным таблицы 39 рассчитайте возможные относительные величины.

Таблица 39 – Реализация продукции на сельскохозяйственном предприятии (данные условные)

Вид продукции	2009 г.		2010 г.	
	план	факт	план	факт
Картофель, ц	45000	51300	54000	53050
Овощи открытого грунта, ц	10000	8450	6150	4980
Молоко, ц	44400	42860	45500	45365

Задача 43

По данным таблицы 40 рассчитайте возможные относительные величины.

Таблица 40 – Производство основных продуктов животноводства в хозяйствах всех категорий в Пензенской области

	1998 г.	2000 г.	2006 г.	2007 г.	2008 г.
Мясо (в убойном весе) тыс. т	53,4	54,0	68,6	91,6	99,8
в т.ч.: говядина и телятина	29,7	26,9	26,7	24,6	26,4
свинина	16,0	18,7	19,1	22,1	22,3
баранина и козлятина	1,4	1,1	1,0	1,2	1,3
мясо птицы	5,9	7,0	21,2	43,3	49,4
Молоко, тыс. т.	460,9	435,9	503,8	537,1	549,2
Яйца, млн. шт.	355,3	330,4	229,2	269,9	275,3
Шерсть (в физическом весе), т	252	192	226	318	277

Задача 44

Объем продаж в АО в 2010 г. в сопоставимых ценах вырос по сравнению с предшествующим годом на 5 % и составил 146 млн. руб. Определите объем продаж в 2009 г.

Задача 45

Торговая фирма планировала в 2010 г. по сравнению с 2009 г. увеличить оборот на 14,5 %. Выполнение установленного плана составило 102,7 %. Определите относительный показатель динамики оборота.

Задача 46

По данным таблицы 41 рассчитайте возможные относительные величины.

Таблица 41 – Производство продукции растениеводства в Пензенской области, тыс. т.

Вид продукции	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.
Зерновые культуры (в весе после доработки)	1170,4	836,9	906,6	992,5	1107,9	932,1	1422,3	1461,4
Сахарная свекла, (фабричная)	354,6	697,9	536,1	678,4	1146,7	923,2	1091,2	928,6
Семена подсолнечника	23,7	32,3	27,9	41,1	38,2	38,6	46,3	67,7
Картофель	405,1	513,3	496,0	487,6	432,4	409,8	426,5	479,1
Овощи (включая овощи закрытого грунта)	152,7	155,0	146,4	148,0	133,1	129,6	139,4	151,7
Справочно: среднегодовая численность населения, тыс. чел.	1451,0	1429,4	1414,4	1402,0	1392,0	1383,9	1379,8	1376,5

Задача 47

Объем продаж компании в первом полугодии 2010 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.

Задача 48

По данным таблицы 42 рассчитайте структуру и динамику производства продукции животноводства по формам собственности.

Таблица 42 – Производство продукции животноводства по формам собственности в Пензенской области

Вид продукции	Все формы собственности		В том числе:					
			государственная		частная		муниципальная	
	2008 г.	2009 г.	2008 г.	2009 г.	2008 г.	2009 г.	2008 г.	2009 г.
Мясо (в убойном весе), тыс. т	99,8	105,3	1,3	1,0	93,0	99,0	0,3	0,1
Молоко, тыс. т	549,2	511,2	17,7	13,7	514,7	487,7	9,9	3,8
Производство яиц, млн шт.	275,3	232,4	10,9	3,5	262,4	317,5	-	-
Шерсть, т	277,0	236,0	3,0	1,0	262,4	215,0	-	-

Задача 49

По данным таблицы 43 восстановите отсутствующие показатели, помеченные многоточием.

Таблица 43 – Динамика и структура продукции животноводства на сельскохозяйственном предприятии (данные условные)

Вид продукции	Товарная продукция, млн. руб.			2010 г. в % к 2009 г.	% выполнения плана в 2010 г.	Доля к итогу 2010 г., %
	2009 г.	2010 г.				
		план	факт			
Молоко	…	…	290	105	102	…
Крупный рогатый скот в живом весе	80	…	130	…	105	28
Прочая продукция животноводства	…	45	…	110	…	…
Итого	…	…	…	…	…	…

Задача 50

Предприятие планировало увеличить выпуск продукции в 2010 г. по сравнению с 2009 г. на 18 %. Фактический же объем продукции составил 112,3 % от прошлогоднего уровня. Определите относительный показатель реализации плана.

Задача 51

По данным таблицы 44 рассчитайте структуру экспорта и импорта Российской федерации.

Таблица 44 – Экспорт Российской Федерации, млрд. долл. США

	1995 г.	2005 г.	2008 г.
Экспорт – всего:	78,2	241,0	397,0
в том числе:			
продовольственные товары и сельскохозяйственное сырье	1,4	4,5	6,6
минеральные продукты	33,3	156,0	271,0
продукция химической промышленности	7,8	14,4	24,5
кожевенное сырье, пушнина	0,3	0,3	0,3
древесина и целлюлозно-бумажные изделия	4,4	8,3	9,6
текстиль, текстильные изделия и обувь	1,1	1,0	0,8
металлы, драгоценные камни и изделия из них	20,91	40,6	50,5
машины, оборудование и транспортные средства	8,0	13,5	21,5
прочие товары	1,0	2,5	9,6

Задача 52

По данным таблицы 45 вычислите возможные относительные величины.

Таблица 45 – Численность постоянного населения

Пензенской области
(на начало года, тыс. человек)

Год	Все население	В т.ч.		Городское население	В т.ч.		Сельское население	В т.ч.	
		мужчины	женщины		мужчины	женщины		мужчины	женщины
2000	1500,2	690,2	810,0	969,1	443,7	525,4	531,1	246,6	284,5
2001	1484,1	681,6	802,5	961,2	438,7	522,5	522,9	242,9	280,0
2002	1466,2	671,7	794,5	952,2	433,0	519,2	514,0	238,6	275,4
2003	1449,2	662,6	786,6	943,7	427,8	515,9	505,5	234,8	270,7
2004	1436,0	655,5	780,5	938,1	424,2	613,9	497,9	231,3	266,6
2005	1422,7	648,4	774,3	933,4	420,9	612,5	489,3	227,5	261,8
2006	1408,0	640,4	767,6	926,9	416,7	510,2	481,1	223,7	257,4
2007	1396,0	634,3	761,7	923,0	414,2	508,8	473,0	220,1	252,9
2008	1388,0	630,3	757,7	920,1	412,2	507,9	467,9	218,1	249,8
2009	1379,8	626,3	753,5	917,9	410,8	507,1	461,9	215,5	246,4

Задача 53

По данным таблицы 46 рассчитайте средний курс акций по всем площадкам.

Таблица 46 – Результаты торговой сессии по акциям АО «Возрождение»

Торговая площадка	Средний курс, руб.	Объем продаж, шт.
Российская торговая система	446	138626
Московская межбанковская валютная биржа	449	175535
Московская фондовая биржа	455	200

Задача 54

По данным таблицы 47 определите средний уровень квалификации рабочих предприятия.

Таблица 47 – Распределение рабочих по тарифному разряду

Тарифный разряд	1	2	3	4	5	6
Число рабочих, чел.	2	3	26	74	18	4

Задача 55

По данным таблицы 48 о реализации товара на рынках города определите среднюю цену товара за I, II квартала и за полугодие.

Таблица 48 – Реализация продукции на рынках города

Рынок	I квартал		II квартал	
	цена за 1 кг, руб.	продано, т	цена за 1 кг, руб.	продано, т
1	85	24	95	1900
2	75	37	80	2800
3	80	29	90	2070

Задача 56

По данным таблицы 49 рассчитайте среднюю отраслевую загрузку производственных мощностей по каждому виду продукции.

Таблица 49 – Производственные мощности металлургических комбинатов и уровень их использования

Комбинат	Мощность, млн т в год			Загрузка, %		
	чугун	сталь	прокат	чугун	сталь	прокат
Магнитогорский	10,5	18,5	12,0	41,3	63,4	53,4
Череповецкий	9,5	13,5	11,5	60,5	70,4	58,5
Новолипецкий	9,5	9,9	7,0	71,4	73,7	89,0
Нижнетагильский	7,0	8,0	4,5	64,2	70,6	82,9
Западно-сибирский	6,0	6,9	4,3	69,3	75,4	82,5
Челябинский	4,0	7,0	4,0	36,4	44,9	43,7
Кузнецкий	3,7	4,8	3,5	74,2	67,0	76,7

Задача 57

По данным таблицы 50 рассчитайте среднюю цену 1 м^2 жилья.

Таблица 50 – Распределение жилья, предлагаемого к продаже, по уровню цен

Цена 1 $м^2$, долл. США	Общая площадь, тыс. $м^2$
200−400	20,4
400−500	29,5
500−600	7,3
600−700	7,0
700−800	4,0

Задача 58

По данным таблицы 51 определите средний размер земельного участка крестьянских (фермерских) хозяйств.

Таблица 51 – Распределение крестьянских (фермерских) хозяйств по размеру земельного участка

Размер земельного участка, га	Удельный вес в общем числе хозяйств, %
До 3	18,0
4−5	9,7
6−10	13,9
11−20	15,5
21−50	18,7
51−70	6,0
71−100	5,7
101−200	7,0
Свыше 200	5,5
Итого	100,0

Задача 59

По данным таблицы 52 определите среднюю себестоимость 1 ц свеклы в целом по фермерским хозяйствам

Таблицы 52 – Производство сахарной свеклы в фермерских хозяйствах

Группа хозяйств по себестоимости 1 ц, руб.	Число хозяйств	Валовой сбор в среднем на 1 хозяйство, ц
До 60	32	111,3
60−70	58	89,7
70−80	124	113,5
80 и более	17	130,1

Задача 60

По данным таблицы 53 определите средние затраты на 1 руб. произведенной продукции.

Таблица 53 – Затраты на производство продукции в корпорации

Предприятие	Общие затраты на производство, тыс. руб.	Затраты на 1 руб. произведенной продукции, руб.
1	2323,4	75
2	8215,9	71
3	4420,6	73
4	3525,3	78

Задача 61

По данным таблицы 54 определите средний процент брака в целом по предприятию.

Таблица 54 – Качество произведенной продукции на предприятии

Вид продукции	Процент брака	Стоимость бракованной продукции, руб.
А	1,3	2135
В	0,9	3560
С	2,4	980

Задача 62

По данным таблицы 55 определите в целом по акционерному обществу средний удельный вес продукции I сорта в 2009 и 2010 гг.

Таблица 55 – Качество произведенной продукции в АО

Предприятие	2009 г.		2010 г.	
	удельный вес продукции I сорта, %	стоимость продукции I сорта, млн. руб.	удельный вес продукции I сорта, %	стоимость всей произведенной продукции, млн. руб.
I	92	130,2	95	153,7
II	80	67,5	82	65,4

Задача 63

По данным таблицы 56 определите средний надой молока на одну корову обычным методом и методом моментов, а также среднюю жирность молока.

Таблица 56 – Группировка сельскохозяйственных предприятий района по среднегодовому надою молока

Группа сельхозпредприятий по среднему годовому надою молока от одной коровы, кг	Число сельхозпредприятий	Среднегодовое поголовье коров (на 1 сельхозпредприятие)	Процент жира в молоке
До 2000	4	417	3,0
2000−2200	9	350	3,3
2200−2400	15	483	3,8
2400 и более	8	389	2,9

7 ПОКАЗАТЕЛИ ВАРИАЦИИ И АНАЛИЗ

ЧАСТОТНЫХ РАСПРЕДЕЛЕНИЙ

Исследование вариации является составным элементом статистического анализа, позволяющим оценить колебания значений изучаемого признака, взаимосвязь его с другими признаками. Статистические показатели, характеризующие вариацию, служат критерием типичности рассчитанных по совокупности средних величин, используются в определении ошибок выборочных характеристик.

Совокупность значений изучаемого признака с указанием числа их различных значений называется распределением признака. Распределение представляют в форме вариационного ряда. В соотношении значений признака (вариантов) и числа единиц (частот) проявляется закономерность распределения. Она описывается следующими статистическими показателями:

- частотные показатели;
- показатели центра распределения;
- показатели степени вариации;
- показатели формы распределения.

Частотными показателями любого ряда являются абсолютная численность i-й группы – частота f_1 и относительная частота – частность d_i.

Кумулятивная (накопленная) частота, S_i (частность S_d) характеризует объем совокупности со значением вариантов, не превышающих X_i. Кумулятивные частотные показатели образуются последовательным суммированием абсолютных или относительных частот.

Плотность частоты (частности) представляет собой частоту, приходящуюся на единицу интервала, т.е.:

$$q_i = f_i/h_i \text{ или } q_i = d_i/h_i;$$

где h_i – величина i-го интервала.

Данный показатель используют, если интервалы вариационного ряда неравные и необходимо графически изобразить этот ряд в виде гистограммы.

К показателям центра распределения относят среднюю, моду и медиану. *Средняя величина* характеризует типичный уровень признака в совокупности. По данным вариационного ряда распределения рассчитывается как арифметическая взвешенная:

на основе частот $\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$;

на основе частостей $\bar{x} = \sum x_i d_i$, где $\sum d_i = 1$

Мода (M_o) – это значение признака, наиболее часто встречающегося в совокупности. В дискретном ряду мода находится по определению, а в интервальных рассчитывается по формуле

$$M_0 = x_0 + i\frac{(f_2 - f_1)}{(f_2 - f_1) + (f_2 - f_3)}$$

где x_0 и i – соответственно нижняя граница и величина модального интервала;

f_1, f_2, f_3 – частоты предмодального, модального и послемодального интервала.

Медиана (Me) – значение признака, приходящееся на середину ранжированной совокупности, т.е. эта варианта, которая делит ряд распределения на две равные части.

В дискретном ряду распределения медиана находится непосредственно по накопленной частоте, соответствующей номеру медианы.

Для определения медианы в ранжированном ряду необходимо вначале найти номер медианы:

$$N = \frac{n+1}{2}$$

Затем используют кумулятивные частоты S_1 или частоты S_d.

В интервальном ряду распределения медиана вычисляется по формуле

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

где x_0 и i – соответственно нижняя граница и величина медианного интервала;

f_{Me} – частота медианного интервала;

S_{Me-1} – накопленная частота предмедианного интервала.

Пример. Для расчета показателей центра распределения воспользуемся данными, представленными в таблице 57, о содержании влаги в поступившей партии товара в магазин.

Таблица 57 – Распределение образцов по влажности

Влажность, % (x)	Число образцов (f)	Накопленная частота (S)	Середина интервала (x')	$x'f$
А	1	2	3	4
До 14	20	20	13	260
14-16	30	50	15	450
16-18	25	75	17	425
18-20	15	90	19	285
20 и более	10	100	21	210
Итого	100	-	-	1630

Так как ряд распределения интервальный, заменяем значения графы А на среднее значение интервалов (x') (графа 3). Средний процент влажности находим по формуле

$$\overline{x'} = \frac{\sum x_i f_i}{\sum f_i}$$

$$\overline{x'} = \frac{1630}{100} = 16{,}3\,\%$$

Следовательно, для данных образцов средний процент влажности составляет 16,3 %.

По данным таблицы 57 рассчитаем моду по формуле

$$M_0 = x_0 + i\frac{(f_2 - f_1)}{(f_2 - f_1) + (f_2 - f_3)}$$

Модальным является интервал 14−16, так как у него наибольшая частота.

Нижняя граница 14, величина интервала (i) = 2, предмодальная частота (f_1) = 20, модальная частота (f_2) = 30, послемодальныя частота (f_3) = 25:

$$M_0 = 14 + 2\frac{30 - 20}{(30 - 20) + (30 - 25)} = 15{,}3\,\%$$

Таким образом, наибольший процент влажности в партии поступивших товаров составил 15,3 %.

Для нахождения медианы находим предварительно ее порядковый номер (графа 1):

$$N = \frac{100}{2} = 50$$

Затем воспользуемся накопленной частотой, отраженной в графе 3. Все образцы, начиная с 21 и по 50 включительно, имеют влажность 14−16 %. Значит, этот интервал медианный.

Вычисляем значение медианы:

$$Me = 14 + 2\frac{50 - 20}{30} = 16\,\%.$$

Таким образом, половина партии товара имеет влажность менее 16 %, и половина имеет содержание влажности в партии товара выше 16 %.

В симметричных рядах распределения значения медианы совпадают со средней величиной *(x = Me = M_o)*, а в умеренно асимметричных они соотносятся таким образом:

$$3 \times (\bar{x} - M_e) \approx \bar{x} - M_0$$

В приведенном примере соотношение характеристики центра распределения содержания влаги в партии товара свидетельствует об умеренной асимметрии:

$$3 \times (16{,}3 - 16) \approx 16{,}3 - 15{,}3.$$

Порядковые характеристики. Рассмотренные обобщающие показатели центра распределения не вскрывают характера последовательного изменения частот, поэтому в анализе закономерностей распределения используются также ранговые (порядковые) показатели: квартили и децили.

Квартили Q – это значения вариантов, которые делят упорядоченный ряд по объему на четыре равновеликие части. Следовательно, в ряду распределения выделяют три квартиля. Медиана является одновременно вторым квартилем. Расчет квартилей основывается на кумулятивных частотах (частостях), и определяются первый и третий квартили по формулам:

Первый квартиль $$Q_1 = x_{Q1} + i\frac{0{,}25\sum_1^m f_i - S_{Q1-1}}{f_{Q_1}}.$$

Третий квартиль $$Q_3 = x_{Q3} + i\frac{0{,}75\sum_1^m f_i - S_{Q3-1}}{f_{Q3}}.$$

По данным таблицы 57 определим первый и третий квартили:

$$Q_e = 14 + 2\frac{0{,}25 \times 100 - 20}{30} = 14{,}3\ \%,$$

$$Q_3 = 18 + 2\frac{0{,}75 \times 100 - 75}{15} = 18{,}0\ \%.$$

Следовательно, в ряду распределения по данным о содержании влаги в поступившей партии товара в магазин первый квартиль составляет 14,3 %, а третий – 18,0, т.е. 25 % товаров содержат процент влажности, не превышающий 14,3 , а у 75 % товаров процент влажности не превышает 18 %.

Децили – это значения вариантов, которые делят упорядоченный ряд по объему на десять равных частей. В ряду распределения выделяют девять децилей, так как медиана является одновременно пятым децилем. Расчет децилей также основан на кумулятивных частотах (частостях) и определяется по формулам:

$$D_1 = x_{D_1} + i\frac{0{,}1\sum_{1}^{m} f_i - S_{D_1 - 1}}{f_{D_1}};$$

$$D_2 = x_{D_2} + i\frac{0{,}2\sum_{1}^{m} f_i - S_{D_2 - 1}}{f_{D_2}}$$

и т.д.

$$D_9 = x_{D_9} + i\frac{0{,}9\sum_{1}^{m} f_i - S_{D_9 - 1}}{f_{D_9}}.$$

Пример. По данным таблицы 57 определим 1-й и 9-й децили:

$$D_1 = 12 + 2\frac{0{,}1 \times 100 - 0}{20} = 13\,\%;$$

$$D_9 = 20 + 2\frac{0{,}9 \times 100 - 90}{10} = 20\,\%.$$

Таким образом, значения децилей указывает на то, что среди 10 % партии товара с минимальным процентом влажности, максимальный процент ее составляет 13 %, а среди 10 % партии товара с наибольшим процентом влажности минимальный процент ее составил 20 %, т.е. в 1,54 раза больше.

Показатели вариации и способы их расчета. Для измерения и оценки вариации используют абсолютные и относительные характеристики.

В практическом анализе оценка рассеяния значений признака имеет такое же большое значение, как и определение характеристик центра распределения. Для измерения и оценки вариации пользуются абсолютными и относительными характеристиками.

Самая предварительная оценка рассеяния (вариации) по данным рядов распределения определяется с помощью *вариационного размаха R*, который показывает, насколько велико различие между единицами совокупности, имеющими самое маленькое и самое большое значение признака, т.е. он равен: $R = X_{max} - X_{min}$.

Среднее линейное отклонение $\bar{d}$ является обобщающей мерой вариации индивидуальных значений признака от средней арифметической величины. Она дает абсолютную меру вариации.

Если данные не сгруппированы, то расчет среднего линейного отклонения осуществляется по принципу невзвешенной средней, т.е.

$$\bar{d} = \frac{\sum_{1}^{n} |x_i - \bar{x}|}{n}.$$

Если данные вариации представлены вариационными рядами распределения, то расчет производится по принципу взвешенной средней, т.е.

$$\bar{d} = \frac{\sum_{1}^{n} |x_i - \bar{x}| f_i}{\sum_{1}^{m} f_i}.$$

Дисперсия $(\sigma)^2$ – это средний квадрат отклонений индивидуальных значений признака от средней величины. Дисперсию используют не только для оценки вариации, но и при измерении взаимосвязей, для проверки статистических гипотез.

Она вычисляется по формулам:

для несгруппированных данных $\sigma = \dfrac{\sum_{1}^{n} (x - \bar{x})^2}{n}$;

для сгруппированных данных $\sigma = \frac{\sum_{1}^{n}(x-\bar{x})^2 f_i}{\sum_{1}^{m} f_i}$.

Среднее квадратическое отклонение σ представляет собой корень второй степени из среднего квадрата отклонений отдельных значений признака от их средней, т.е. оно исчисляется путем извлечения квадратного корня из дисперсии и измеряется в тех же единицах, что и варьирующий признак.

Формулы расчета следующие:

для несгруппированных данных $\sigma^2 = \sqrt{\frac{\sum_{1}^{n}(x-\bar{x})^2}{n}}$;

для сгруппированных данных $\sigma^2 = \sqrt{\frac{\sum_{1}^{n}(x-\bar{x})^2 f_i}{\sum_{1}^{m} f_i}}$.

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

Для сравнения колеблемости различных признаков в одной и той же совокупности, или же при сравнении колеблемости одного и того же признака в нескольких совокупностях, вычисляются относительные показатели вариации. Базой для сравнения служит средняя арифметическая. Эти показатели вычисляются как отношение размаха, или среднего линейного отклонения, или среднего квадратического отклонения к средней арифметической. Чаще всего они выражаются в процентах и характеризуют 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 нижеследующие относительные показатели вариации (V).

Коэффициент осцилляции: $V_R = \frac{R}{\bar{x}} \times 100\,\%.$

Линейный коэффициент вариации: $V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \times 100\,\%.$

Коэффициент вариации: $V_\sigma = \frac{\sigma}{\bar{x}} \times 100\,\%.$

Для расчета показателей вариации возьмем за основу рассмотренный ранее ряд распределения о содержании влаги в поступившей партии товара в магазин (таблица 57).

Исходные данные и промежуточные вычисления для расчета показателей вариации оформляем в таблице 58.

Размах вариации по данным ряда распределения составляет 8 % (21 %–13 %).

Таблица 58 – Исходные и расчетные данные для вычисления показателей вариации

Влажность, % (х)	Число образцов (f)	Середина интервала X^1	$\lvert x_i - \bar{x}^1 \rvert f_1$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2 f$
А	1	2	3	4	5
До 14	20	13	66,0	-3,3	217,80
14−16	30	15	39,0	-1,3	50,70
16−18	25	17	17,5	0,7	12,25
18−20	15	19	40,5	2,7	109,35
20 и более	10	21	47,0	4,7	220,9
Итого	100	-	210,0	-	611,0

Для расчета среднего линейного отклонения ($\bar{d}$) необходимые вычисления оформим в графе 3.

$$\bar{d} = \frac{210}{100} = 2{,}1\%$$

Значение среднего возьмем из предыдущих расчетов ($\bar{x} = 16{,}3\,\%$).

Следовательно, отклонение вариантов признана от их средней величины составляет 2,1 %.

Для вычисления дисперсии (δ^2) дополнительные расчеты оформим в графах 4 и 5 (таблице 58):

$$\delta^2 = \frac{611}{100} = 6{,}1\,\%$$

Извлечем квадратный корень из дисперсии и получим величину среднего квадратического отклонения:

$$\delta = \sqrt{\delta^2} = \sqrt{6{,}1} = 2{,}47\,\%.$$

Значение показателя говорит о том, что варианты отклонялись от средней величины в ту или другую в среднем на 2,47 %.

Рассчитаем коэффициент вариации:

$$V = \frac{2{,}47}{16{,}3} \times 100 = 15{,}16\,\%.$$

Значение коэффициента вариации говорит о том, что анализируемая совокупность является качественно однородной, так как отклоняются от средней величины на 15, 16 % в ту или иную сторону.

Вариация альтернативного признака. В ряде случаев возникает необходимость в измерении дисперсии альтернативных признаков, т.е. признаков, которыми одни единицы обладают и не обладают другие.

Обозначим наличие данного признака 1, отсутствие 0, долю вариантов, обладающих данным признаком *p,* а не обладающих им q. Так как p + q = 1, то средняя $\bar{x}$ = p, а дисперсия альтернативного признака $q^2 = pq$, где $p = m/n$, n – число наблюдений, m – число единиц совокупности, обладающие данным признаком, q =1 – p.

Пример. Определим дисперсию альтернативного признака по следующим данным: налоговой инспекцией одного из районов города проверено 200 коммерческих киосков и в 150 обнаружены финансовые нарушения. Тогда:

$n = 200$, $m = 150$
$p = 150/200 = 0{,}75$
$q = 1 – 0{,}75 = 0{,}25$
$\sigma^2 = 0{,}75 \times 0{,}25 = 0{,}1875$.

Наряду с изучением вариации признака по всей совокупности в целом часто бывает необходимо проследить количественные изменения признака по группам, на которые разделяется совокупность, а также между группами. Такое изучение вариации достигается посредством вычисления и анализа различных видов дисперсии.

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

Общая дисперсия измеряет вариацию признака во всей совокупности под влиянием всех факторов, обусловливающих эту вариацию:

$$\sigma^2 = \frac{\sum_1^m (x_i - \bar{x})^2 f_i}{\sum f_i}.$$

Межгрупповая дисперсия характеризует систематическую вариацию, т.е. различия в величине изучаемого признака, возникающие под влиянием признака-фактора, положенного в основание группировки. Она рассчитывается по формуле

$$\delta^2 = \frac{\sum_1^m (x_i - \bar{x})^2 f_i}{\sum_1^m f_i},$$

где x_i и $\bar{x}$ – соответственно средняя i-й группы и общая средняя варьирующего признаках;

f_i – частота i-й группы.

Внутригрупповая дисперсия отражает случайную вариацию, т.е. часть вариации, происходящую под влиянием неучтенных факторов, и не зависящую от признака-фактора, положенного в основание группировки. Данная дисперсия рассчитывается отдельно для каждой i-й группы по формуле

$$\sigma_i^2 = \frac{\sum_{1}^{fi}(x_i - \overline{x_i})^2{}_i}{f_i},$$

где x_i – значение признака у отдельных элементов совокупности;

f_i – число единиц в группе i.

Для всех групп в целом вычисляется *средняя из внутригрупповых дисперсий,* взвешенных на частоты соответствующих групп по формуле

$$\overline{\sigma^2} = \frac{\sum_{1}^{m}\sigma_i^2 f_i}{\sum_{1}^{m} f_i}.$$

Взаимосвязь между тремя дисперсиями получила название *правила сложения дисперсий,* в соответствии с которым: $\sigma^2 = \delta^2 + \overline{\sigma}^2$, т.е. согласно этому правилу общая дисперсия, возникающая под влиянием всех факторов, равна сумме дисперсий, возникающих под влиянием всех прочих факторов, и дисперсии, возникающей за счет группировочного признака.

Зная любые два вида дисперсий можно определить или проверить правильность расчета третьего вида.

Пример. О рабочих одной из бригад известны нижеследующие данные (таблица 59).

Таблица 59 – Распределение рабочих по тарифному разряду

Тарифный разряд	Число рабочих	Дневная выработка деталей одним рабочим, шт.
3	2	100, 120

4	4	120, 120, 140, 160
5	5	140, 160, 170, 180, 200

Определить по этим данным:

1) внутригрупповую дисперсию по выработке деталей одним рабочим, имеющим данный разряд;

2) среднюю внутригрупповых дисперсий по трем группам рабочих;

3) межгрупповую дисперсию;

4) общую дисперсию выработки рабочих этой бригады.

Решение.

1. Для расчета внутригрупповых дисперсий вычислим средние по каждой группе:

$$\overline{x_1} = \frac{100+120}{2} = 110 \text{ шт.,}$$

$$\overline{x_2} = \frac{100+120+140+160}{4} = 135 \text{ шт.,}$$

$$\overline{x_3} = \frac{140+160+170+180+200}{5} = 170 \text{ шт.,}$$

2. Рассчитаем внутригрупповые дисперсии:

$$\sigma_1^2 = \frac{\sum_{1}^{f_i}(x_i - \overline{x})^2}{f_i} = \frac{(100-110)^2 + (120-110)^2}{2} = \frac{200}{2} = 100$$

$$\sigma_2^2 = \frac{\begin{matrix}(120-135)^2 + (120-135)^2 + \\ + (140-135)^2 + (160-135)^2\end{matrix}}{4} = \frac{100}{4} = 275$$

$$\sigma_3^2 = \frac{\begin{matrix}(140-170)^2 + (160-170)^2 \\ + (170-170)^2 + (180-170)^2 \\ + (200-170)^2\end{matrix}}{5} = \frac{2000}{5} = 400;$$

3. Определим среднюю из внутригрупповых дисперсий:

$$\overline{\sigma}^2 = \frac{100\times 2 + 275\times 4 + 400\times 5}{11} = \frac{3300}{11} = 300 \text{ шт.}$$

4. Определим общую среднюю величину для расчета межгрупповой дисперсии:

$$\bar{x}_{об} = \frac{100\times 2 + 135\times 4 + 175\times 5}{11} = \frac{1610}{11} = 146{,}4 \text{ шт.}$$

5. Теперь определим межгрупповую дисперсию:

$$\overline{\delta}^2 = \frac{(110-146{,}4)^2\times 2 + (135-146{,}4)\times 4 + (170-146{,}4)^2\times 5}{11} =$$

$$= \frac{5954{,}56}{11} = 541{,}3$$

6. Определим общую дисперсию обычным способом:

$$\sigma^2 = \frac{\begin{array}{l}(110-146{,}4)^2 + (120-146{,}4)^2 + (\ldots-146{,}4)^2 + \ldots + \\ (200-146{,}4)^2\end{array}}{11} =$$

$$= \frac{9254{,}56}{11} = 841{,}3$$

7. Проверим полученный результат, вычислив общую дисперсию по правилу сложения дисперсий:

$$\sigma^2 = \delta^2 + \overline{\sigma}^2 = 541{,}3 + 300 = 841{,}3.$$

Таким образом, общая дисперсия, вычисленная по правило сложения дисперсий, в точности совпадает по числовому значению с результатом вычисления ее непосредственно на основе данных по всей совокупности рабочих.

На основании правила сложения дисперсий можно определить показатель тесноты связи между группировочным (факторным) и результативным признаками. Он называется *эмпирическим корреляционным отношением,* обозначается η («эта») и рассчитывается по формуле

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

Для нашего примера эмпирическое, корреляциооное отношение

$$\eta = \sqrt{\frac{541{,}3}{841{,}3}} = 0{,}64.$$

Таким образом, можно сделать вывод о том, что между дневной выработкой деталей и квалификацией рабочих существует средняя статистическая связь, так как корреляционное отношение равно 0,64.

Показатели асимметрий и эксцесса. Выявление общего характера распределения предполагает оценку не только степени его однородности, но и его симметричности, остро- или плосковершинности. *Симметричным* называется *распределение,* в котором частоты любых двух вариантов, равноотстоящих в обе стороны от центра распределения, равны между собой.

Простейшей мерой асимметричности распределения является отклонение между характеристиками центра распределения. Поскольку в симметричном распределений $\bar{x} = Me = Mo$, то чем заметнее асимметрия, тем больше отклонение $(\bar{x} - Mo)$.

Стандартное отклонение называют *коэффициентом асимметрии Пирсона*

$$K_a = \frac{\bar{x} - Mo}{\sigma}$$

где $\bar{x}$ – средняя арифметическая ряда распределения;

Mo – мода;

σ – среднее квадратическое отклонение.

$K_a = 0$, если ряд распределения симметричен, т. е. $\bar{x} = Mo$, если $K_a > 0$ скошенность ряда правосторонняя (т.е. $\bar{x} > Mo$), и при $K_a < 0$ скошенность ряда левосторонняя (т.е. $\bar{x} < Mo$).

В практических расчетах часто в качестве асимметрии применяется *нормированный коэффициент асимметрии третьего порядка,* который не зависит от масштаба, выбранного при измерении варианта, так как является отвлеченной величиной и определяется по формуле

$$A_3 = \frac{\mu_3}{\sigma^3}$$

Если $A_3 < 0$, то в ряду распределения преобладают варианты, которые меньше, чем средняя, т.е. ряд отрицательно асимметричен (или с левосторонней скошенностью – более длинная ветвь влево). Если $A_3 > 0$, то для ряда распределения характерна положительная асимметрия (правосторонняя скошенность – более длинная ветвь вправо), $A_3 = 0$ при симметричном распределении, так как варианты равноудалены от $\bar{x}$ и имеют одинаковую частоту, поэтому $\mu_3 = 0$.

Для определения *крутизны* (заостренности) графика распределения вычисляется центральный момент четвертого порядка и определяется *нормированный момент четвертого порядка* по формуле

$$A_4 = \frac{\mu_4}{\sigma^4}$$

Для нормального распределения $A_4 = 3$. При измерении *асимметрии эталоном* служит симметричное (нормальное) распределение, для которого $A_3 = 0$.

Аналогично при оценке *крутизны* в качестве *эталонного* выбирается *нормальное распределение,* которое сравнивается с фактическим и вычисляется *показатель эксцесса распределения:*

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

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

Оценка существенности показателей асимметрии и эксцесса позволяет сделать вывод о том, можно ли отнести данное эмпирическое распределение к типу нормального распределения.

Пример. Определим коэффициент асимметрии и эксцесса, а также нормированные моменты третьего и четвертого порядка по данным о распределении магазинов по размеру товарооборота.

Необходимые для расчета нормированных моментов показатели представлены в таблице 60.

Таблица 60 – Расчет нормированных моментов

Группа магазинов по размеру товарооборота, млн. руб. x	Число магазинов, f_i	Середина интервала x_i	$x_i f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2 f_i$	$(x_i - \bar{x})^3 f_i$	$(x_i - \bar{x})^4 f_i$
А	1	2	3	4	5	6	7
50-60	7	55	385	-12,2	1041,88	12710,95	155073,45
60-70	15	65	975	-2,2	72,6	159,75	351,38
70-80	6	75	450	7,8	365,04	2847,3	22209,03
80-90	4	85	340	17,8	1267,36	22559	401550,32
Итого	32	-	2150	-	2746,88	38277	579184,18

Решение.

1. Найдем среднюю арифметическую (графа 3):

$$\bar{x} = \frac{\sum_1^m x_i f_i}{\sum_1^m f_i} = \frac{2150}{32} = 67{,}19 \approx 67{,}2 \text{ млн. руб.}$$

2. Найдем дисперсию, т.е. центральный момент второго порядка (графа 5):

$$\sigma^2 = \frac{\sum_1^m (x-\bar{x})^2 f_i}{\sum_1^m f_i} = \frac{2746{,}88}{32} = 85{,}84.$$

2. Найдем среднее квадратическое отклонение (стандарт):

$$\sigma = \sqrt{\sigma^2} = \sqrt{85{,}84} = 9{,}265.$$

4. Определим центральный момент третьего порядка (графа 6):

$$\mu_3 = \frac{\sum_1^m (x_i-\bar{x})^3 f_i}{\sum_1^m f_i} = \frac{38277}{32} = 1196{,}16.$$

5. Определим нормированный момент третьего порядка:

$$A_3 = \frac{\mu_3}{\sigma^3} = \frac{1196{,}16}{9{,}265^3} = 15{,}04.$$

6. Определим моду для того, чтобы найти коэффициент асимметрии Пирсона:

$$Mo = 60 + 10\frac{15-7}{(15-7)+(15-6)} = 64{,}71 \text{ млн. руб.}$$

7. Определим коэффициент асимметрии Пирсона:

$$K_a = \frac{\bar{x} - Mo}{\sigma} = \frac{67{,}19-64{,}71}{9{,}265} = 0{,}268.$$

В данном случае асимметрия небольшая и скошенность правосторонняя.

8. Найдем центральный момент четвертого порядка (графа 7):

$$M_4 = \frac{\sum_1^m (x_i - \bar{x})^4 f_i}{\sum_1^m f_i} = \frac{579184{,}18}{32} = 18099{,}51.$$

9. Определим нормированный момент четвертого порядка:

$$A_4 = \frac{M_4}{\sigma^4} = \frac{18099{,}51}{9{,}265^3} = 2{,}456.$$

10. Определим эксцесс распределения:

$$E_k = A_4 - 3 = 2{,}456 - 3 = -0{,}544.$$

Так как $E_k < 0$ распределение низковершинное.

Построение нормального распределения по эмпирическим данным. Имея дело с эмпирическим распределением, можно предположить, что данному распределению соответствует 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.

Рассмотрим на примере расчет теоретических частот ряда распределения по данным таблицы 60.

Нормальным $N(\bar{x}, \sigma)$ называют распределение непрерывной случайной величины x, если соответствующая ее плотность распределения выражается формулой

$$f_{(x)} = \varphi_{(x,\bar{x},\sigma^2)} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-x)^2}{2\sigma^2}}$$

или

$$\varphi_{(t)} = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}},$$

где x – значение изучаемого признака;

$t\frac{x-\bar{x}}{\sigma}$ – нормированное отклонение;

σ – среднее квадратическое отклонение изучаемого признака;

$\bar{x}$ – средняя арифметическая ряда распределения;

σ^2 – дисперсия значений изучаемого признака;

π – постоянное число, которое равно 3,1415;

e – основание натурального логарифма, равное 2,7182.

Случайные величины, распределенные по нормальному закону, различаются значениями параметров $\bar{x}$ и σ, поэтому важно выяснить, как эти параметры влияют на вид нормальной кривой.

В зависимости от их значений, она может иметь разный центр группирования, т.е. быть более удлиненной или сжатой.

1. Сначала необходимо найти среднюю влажность образцов. Она была найдена в таблице 57: $\bar{x}$ = 16,3 %.

2. Затем определяется среднее квадратическое отклонение. Оно также было рассчитано ранее в таблице 60: σ = 2,55 %.

3. Определим нормированное отклонение t для каждого варианта (таблица 60, графа 4).

4. По таблице распределения функция φ_i (см. приложение 2), определим ее значение (таблица 60, графа 5).

5. Определим теоретические значения частоты f^1 по формуле

$$f^1 = \frac{k\sum f}{\sigma}\varphi_{(t)},$$

где k – длина интервала.

Если вариационный ряд распределения имеет равные интервалы, то

$$\frac{k\sum f}{\sigma}const.$$

Для нашего примера эта величина равна 2 × 100/2,55 = 78.

Затем это значение *(const)* умножим на величину $\varphi(t)$ при данном t и получим искомую теоретическую частоту (графа 6).

Таблица 61 – Расчет нормированного отклонения по данным о влажности образцов

Влажность, %	Число образцов, (f_i)	Середина интервала, x_i	$(x_i - \bar{x})$	$\frac{x_i - \bar{x}}{\sigma}$	$\varphi(i)$	Теоретическая частота, f_i'
А	1	2	3	4	5	6
До 14	20	13	-3,3	1,294	0,1849	14
14–16	30	15	-1,3	0,510	0,3503	27
16–18	25	17	0,7	0,275	0,3847	30
18–20	15	19	2,7	0,059	0,2275	18
20 и более	10	21	4,7	1,843	0,0748	6
Итого	100	-	0,0	-	-	5

6. Сравним на графике эмпирическое f и теоретические f_i частоты, полученные на основе данных таблицы (рисунок 10).

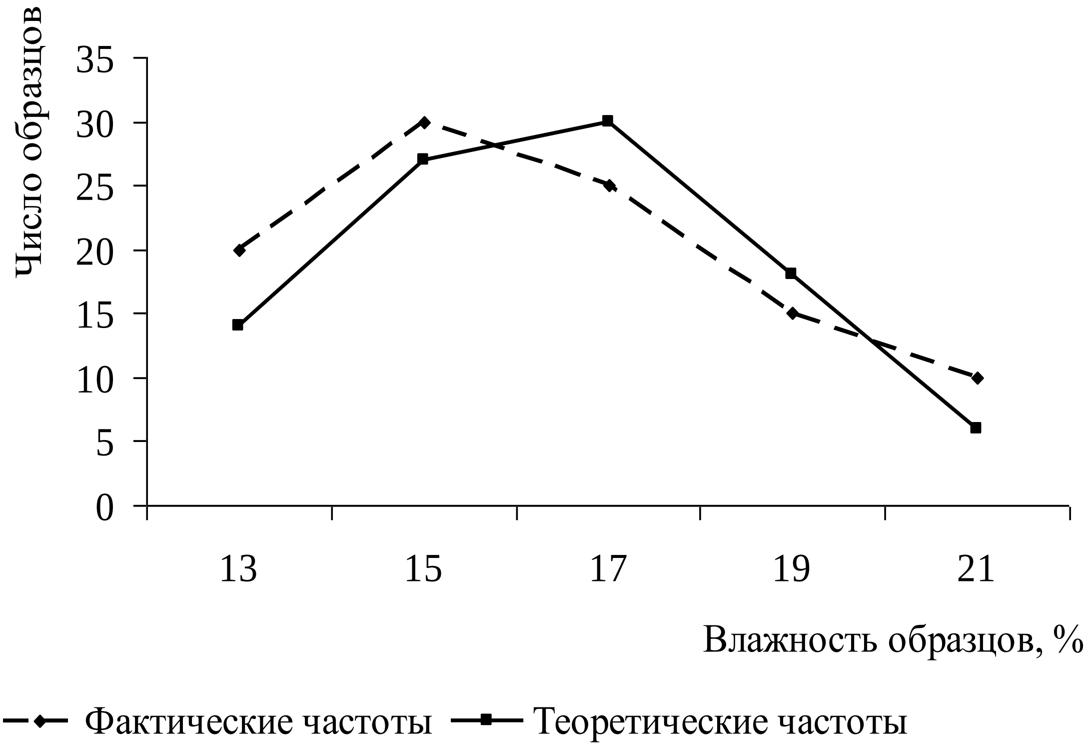

Рисунок 10 – Распределение образцов по влажности

Критерии согласия. Поскольку все предположения о характере того или иного распределения – это гипотезы, то они должны быть подвергнуты статистической проверке с помощью показателей, которые называются *критериями согласия*.

Критерии согласия дают возможность установить, когда расхождения между теоретическими и эмпирическими частотами следует признать случайными (несущественными). Существует ряд критериев согласия. Наиболее часто на практике применяются критерии Пирсона, Романовского, Колмогорова. Рассмотрим их.

Критерий согласия Пирсона (χ^2) вычисляется по формуле

$$\chi^2 = \sum_1^m \frac{(f' - f_i')^2}{f_i'},$$

где f_i' и f' – эмпирические и теоретические частоты, соответственно;

m – число групп, на которые разбито эмпирическое разделение.

Для распределения χ^2 составлены таблицы, где указано критическое значение критерия, согласно χ^2, для выбранного уровня значимости α и данного числа степеней свободы $\nu = n - 1$ (см. приложение 4). Для оценки существенности расчетное значение $\chi^2_{\text{расч}}$ сравнивается с табличным $\chi^2_{\text{табл}}$.

При полном совпадении теоретического и эмпирического распределений $\chi^2 = 0$, в противном случае $\chi^2 > 0$. Если $\chi^2_{\text{расч}} > \chi^2_{\text{табл}}$, то при заданном уровне значимости α число степеней свободы и гипотезу ν о случайности расхождения отклоняют.

В случае, если $\chi^2_{\text{расч}} \le \chi^2_{\text{табл}}$, делаются заключения о том, что эмпирический ряд хорошо согласуется с гипотезой о предполагаемом распределении и с вероятностью $(1-\alpha)$ можно утверждать, что расхождения между теоретическими и эмпирическими частотами случайны.

Критерий Романовского (C) основан на использовании критерия χ^2 Пирсона, т.е. уже найденных значениях и числа степеней свободы ν:

$$C = \frac{\left|x^2 - \nu\right|}{\sqrt{2\nu}}$$

При $C < 3$ расхождения между теоретическим и эмпирическим распределением считаются случайными, если же $C > 3$, то неслучайными, и, следовательно, теоретическое распределение не может служить моделью для изучаемого эмпирического распределения.

Критерий Колмогорова (λ) основан на определении максимального расхождения между накопленными частотами эмпирических и теоретических распределений:

$$\lambda = \frac{D}{\sqrt{\sum f}},$$

где D – максимальное значение разности между накопленными эмпирическими и теоретическими частотами;

$\sum f$ – сумма эмпирических частот.

Необходимым условием использования этого критерия является достаточно большее число наблюдений (не меньше ста).

Рассчитав значение λ, по таблице $P(\lambda)$ (см. приложение 3) определяют вероятность, с которой можно утверждать, что отклонение эмпирических частот от теоретических случайны. Вероятность $P(\lambda)$ может изменяться от 0 до 1. При $P(\lambda) = 0$ – полное расхождение частот. Если λ принимает значение до 0,3, то $P(\lambda) = 1$.

Пример. Используя данные таблицы 62, проверим правильность выдвинутой гипотезы о распределении влажности образцов по закону нормального распределения. Для оценки близости эмпирических и теоретических частот воспользуемся критерием Пирсона и Романовского.

Определим критерий Пирсона. Все расчеты показаны в таблице 62.

Таблица 62 – Исходные и расчетные данные *для определения критерия Пирсона*

Влажность, %	Частоты ряда		$f' - f_i'$	$(f' - f_i')^2$	$\frac{(f' - f_i')^2}{f_i'}$
	эмпирические, (f')	теоретические (f_i')			
А	1	2	3	4	5
До 14	20	14	6	36	2,571
14−16	30	27	3	9	0,333
16−18	25	30	-5	25	0,833
18−20	15	18	-3	9	0,500
20 и более	10	6	4	16	2,667
Итого	100	95	-	-	6,904

Решение

1. Фактическое значение $\chi^2_{\text{факт}}$ = 6,904.

Если частота первых двух групп меньше 5, то их необходимо объединить с третьей группой.

2. Находим критическое (табличное) значение $\chi^2_{\text{табл}}$. В примере выделено пять групп (вариантов), следовательно, и пять групп частот. Число степеней свободы $\nu = 5 - 1 = 4$. Примем наиболее часто используемый уровень значимости α = 0,05 и обратимся к приложению 4.

По таблице значения χ^2 критерия Пирсона для степеней свободы $\nu = 4$ и уровня значимости α = 0,05 определяем, что $\chi^2_{\text{табл}}$ = 9,488. Так как фактическое значение $\chi^2_{\text{факт}}$ = 6,904, т.е. меньше табличного, можно считать случайными расхождения между эмпирическими и теоретическими частотами и выдвинутая гипотеза о близости эмпирического распределения и нормальному не опровергается.

3. Определим критерий Романовского:

$$C = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}} = \frac{6{,}904 - 4}{\sqrt{2 \times 4}} = \frac{2{,}904}{2{,}828} = 1{,}027$$

Так как C<3, гипотеза не отвергается.

Критерий Романовского также подтверждает, что расхождения между эмпирическими и теоретическими частотами несущественны, т.е. случайны.

Задача 64

По данным приведенным в таблице 63 определите показатели центра распределения и сделайте выводы о характере данного распределения.

Таблица 63 – Результаты экзаменационной сессии

Балл оценки знаний студентов	2	3	4	5	Итого
Число оценок, полученных студентами	6	75	120	99	300

Задача 65

Имеются данные о распределении безработных по длительности перерыва в работе, представленные в таблице 64.

Определите:

− медианные и квартильные значения продолжительности перерыва в работе;

− объясните их содержание и проведите сравнительный анализ.

Таблица 64 – Распределение безработанных по длительности перерыва в работе N региона

Длительность перерыва в работе, месяцев	В % к общей численности мужчин, женщин	
	мужчин	женщин
До 3	27,4	20,4
3−6	38,3	47,1
6−9	14,6	13,5
9−12	10,7	10,4
12 и более	9,0	8,6
Итого	100,0	100,0

Задача 66

По данным таблицы 65 определите квартили и децили уровня кредитных вложений, объясните их содержание.

Таблица 65 – Распределение коммерческих банков по величине кредитных вложений

Величина кредитных вложений, млн. руб.	До 200	200–400	400–600	600–800	800–1000	1000 и более	Итого
Число банков	5	10	8	7	4	2	36

Задача 67

Для оценки степени децильной дифференциации населения определите децили среднедушевого дохода по данным таблицы 66. Объясните их содержание.

Таблица 66 – Распределение населения по величине среднедушевого денежного дохода в России в 2000 г.

Среднедушевой доход, руб. в месяц	До 500	500–750	750–1000	1000–1500	1500–2000	2000–3000	3000–4000	Свыше 4000	Итого
Численность населения, млн. чел.	4,5	10,5	14,3	30,1	24,7	30,7	14,9	15,9	145,6

Задача 68

В таблице 67 представлены следующие данные.

Таблица 67 – Распределение длины пробега автофургона торговой фирмы

Длина пробега за один рейс, км	30–40	40–50	50–60	60–70	70–80	80 и свыше	Итого
Число рейсов за 1 месяц	20	25	14	18	8	5	90

Определите:

1) среднюю длину пробега за один рейс;
2) среднее квадратическое отклонение;

3) коэффициент вариации.

Оцените количественную однородность совокупности.

Задача 69

Информация о распределении безработных по возрасту представлена в таблице 68.

Таблица 68 – Распределение численности безработных по возрасту

Возраст безработных, лет	В % к общей численности безработных	
	2008 г.	2009 г.
До 20	7,9	8,6
20-24	18,3	17,7
25-29	13,3	12,4
30-34	12,0	12,0
35-39	14,7	13,0
40-44	13,0	13,8
45-49	10,5	10,7
50-54	5,4	6,7
55-59	3,1	2,6
60-72	1,8	2,5
Итого	100,0	100,0

Определите:

1) для каждого года средний возраст безработного;

2) среднее квадратическое отклонение;

3) коэффициент вариации.

Сравните вариацию возраста безработных за два года.

Задача 70

Распределение фермерских хозяйств по посевной площади представлено в таблице 69.

Определите дисперсию и среднее квадратическое отклонение посевных площадей, применив для расчета средний арифмической и дисперсии − способ моментов.

Таблица 69 – Распределение фермерских хозяйств *по посевной площади*

Посевные площади, га	До 100	100–200	200–300	300–400	400–500	500 и более	Итого
Удельный вес хозяйств, % к итогу	17	20	28	25	7	3	100

Задача 71

1. Средняя величина признака в совокупности равна 20, а средний квадрат отклонений индивидуальных значений этого признака от средней величины – 400. Определите коэффициент вариации.

2. Дисперсия признака равна 10, средний квадрат его индивидуальных значений − 140. Чему равна средняя?

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

4. 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.

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

6. Средний квадрат отклонений вариантов признака от произвольной величины равен 300, а сама произвольная величина равна 70 единицам. Определите дисперсию признака, если известно, что средняя величина его варианта равна 80.

7. Средний квадрат отклонений вариантов признака от произвольной величины равен 61. Средняя величина признака больше произвольной величины на шесть единиц и равна 10. Найдите коэффициент вариации.

Задача 72

По данным таблицы 70 сравните вариацию производительности труда по цехам и сделайте выводы.

Таблица 70 – Производительность труда на предприятии

Цех	Средняя часовая производительность труда, $м^2$	Среднее квадратическое отклонение в группе
1	29,20	2,40
2	18,22	2,27
3	28,36	3,47

Задача 73

Имеется нижеследующая информация о семьях сотрудников, представленная в таблице 71.

Таблица 71 – Распределение семей сотрудников корпорации по количеству детей

Количество детей в семье	Число сотрудников по подразделениям		
	1-е	2-е	3-е
0	4	7	5
1	6	10	13
2	3	3	3
3	2	1	-

Определите:

1) внутригрупповые дисперсии;
2) среднюю из внутригрупповых дисперсий;
3) межгрупповую дисперсию;
4) общую дисперсию.

Проверьте правильность произведенных расчетов с помощью правила сложения дисперсий и рассчитайте эмпирические корреляционное отношение.

Задача 74

В таблице 72 представлена информация о строительных фирмах.

Таблица 72 – Распределение строительных фирм по объему инвестиций

Объем инвестиций, млн. руб.	6–8	8–10	10–12	12–14	14–16	16–18	18–20	Итого

Число фирм	4	6	32	34	27	10	7	120

Определите характеристики распределения:

1) среднюю;

2) моду;

3) среднее квадратическое отклонение;

4) коэффициент вариации и ассиметрии.

Сделайте вывод о характере распределения строительных фирм.

Задача 75

По данным предыдущей задачи определите показатели ассиметрии и эксцесса распределения строительных форм по объему инвестиций. Сделайте выводы.

Задача 76

По данным задачи 69 проверьте близость эмпирического и теоретического распределений численности безработных за 2008 г. с помощью критериев согласия Романовского и Колмогорова.

Задача 77

По данным задачи 69 определите критерий согласия Пирсона (χ^2) и проверьте близость эмпирического и теоретического распределений численности безработных за 2008 г.

7 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Выборочное наблюдение – это вид несплошного наблюдения, при котором признаки регистрируются у отдельных единиц изучаемой статистической совокупности, отобранных с использованием специальных методов, а полученные в процессе обследования результаты с определенным уровнем вероятности распространяются на всю исходную совокупность.

Этапы подготовки и проведения выборочного наблюдения:

1 этап – определение цели обследования;

2 этап – установление границ генеральной совокупности;

3 этап – составление программы наблюдения и программы разработки данных;

4 этап – определение вида выборки, процента отбора и метода отбора;

5 этап – отбор и регистрация наблюдаемых признаков у отобранных единиц;

6 этап – расчет выборочных характеристик и их ошибок;

7 этап – распространение полученных результатов на генеральную совокупность.

Генеральная совокупность – это совокупность, из которой производится отбор единиц совокупности.

Выборочная совокупность – это совокупность отобранных в определенном порядке единиц, по которым собирается информация.

Основные характеристики генеральной и выборочной совокупностей приведены в таблице 73.

Отбор единиц в выборочную совокупность может быть повторным или бесповторным.

При *повторном отборе* попавшая в выборку единица подвергается обследованию, возвращается в генеральную совокупность и наравне с другими единицами участвует в дальнейшей процедуре отбора. Некоторые единицы могут попадать в выборку дважды, трижды или даже большее число раз. И при изучении выборочной совокупности они будут рассматриваться как отдельные независимые наблюдения. Число единиц, участвующих в отборе, при таком подходе остается постоянным. Поэтому вероятность попадания в выборку для всех единиц совокупности на протяжении всего процесса отбора также не меняется.

Таблица 73 − Основные характеристики генеральной и выборочной совокупностей

Характеристика	Генеральная совокупность	Выборочная совокупность
Объем совокупности (число единиц)	N	n
Численность единиц, обладающих обследуемых признаком	M	m
Средний размер признака	$\bar{x} = \dfrac{\sum x_i}{N}$	$\tilde{x} = \dfrac{\sum x_i}{n}$
Доля единиц, обладающих обследуемым признаком	$p = \dfrac{M}{N}$	$\omega = \dfrac{m}{n}$
Дисперсия количественного признака	$\sigma_{\bar{x}}^2 = \dfrac{\left[\sum\left(x_i - \bar{x}\right)^2\right]}{N}$	$\sigma_{\tilde{x}}^2 = \dfrac{\left[\sum\left(x_i - \tilde{x}\right)^2\right]}{n}$
Дисперсия доли	$\sigma_p^2 = p(1-p)$	$\sigma_\omega^2 = \omega(1-\omega)$

При *бесповторном отборе* попавшая в выборку единица подвергается обследованию и в дальнейшей процедуре отбора не участвует. Такой отбор целесообразен и практически возможен в тех случаях, когда объем генеральной совокупности четко определен. Получаемые при этом результаты являются более точными по сравнению с результатами, основанными на повторной выборке.

В зависимости от состава и структуры генеральной совокупности, выбирается вид выборки или способ отбора. К наиболее распространенным на практике видам относятся:

– собственно случайная (простая случайная) выборка;
– механическая (систематическая) выборка;
– типическая (стратифицированная, расслоенная) выборка;
– серийная (гнездовая) выборка;
– комбинированная.

Собственно случайная выборка заключается в отборе единиц из генеральной совокупности в целом, без разделения ее на группы, подгруппы или серии отдельных единиц. При этом единицы отбираются в случайном порядке, не зависящем ни от последовательности расположения единиц в совокупности, ни от значений их признаков.

После проведения отбора с использованием какого-либо алгоритма, реализующего принцип случайности, или на основе таблицы случайных чисел, необходимо определить границы генеральных характеристик.

Механическая (систематическая) выборка применяется в тех случаях, когда генеральная совокупность каким-либо образом упорядочена, т.е. имеется определенная последовательность в расположении единиц (например, номера домов, списки избирателей).

При проведении механического отбора устанавливается *шаг отсчета* (h), т.е. расстояние между отбираемыми единицами ($h = \frac{N}{n}$ − величина, обратная доле выборки), и *начало отсчета* – номер единицы, которая должна быть обследована первой.

Механический отбор всегда бывает *бесповторным*. При этом отборе используются формулы, применяемые при собственно случайном бесповторном отборе. Данный вид отбора имеет преимущество перед случайным отбором, его не только легче организовать, но при нем единицы выборочной совокупности равномернее распределяются в генеральной совокупности.

Типическая выборка. Используется в случаях, когда все единицы генеральной совокупности объединены в несколько крупных типических групп (страт). Отбор единиц в выборочную совокупность из каждой типической группы осуществляется собственно случайным или механическим способом.

Отбор единиц в типическую выборку может быть организован либо пропорционально объему типических групп, либо пропорционально внутригрупповой вариации (дифференциации) признака.

Объем выборки из типической группы при отборе, пропорциональном численности единиц типических групп, определяется по формуле

$$n_i = n \times \frac{N_i}{N},$$

где n_i − объем выборки из i-й типической группы;

N_i – объем i-й типической группы в генеральной совокупности.

Так как в типическую выборку должны попасть представители всех групп, средняя ошибка типической выборки зависит только от средней из внутригрупповых дисперсий $\bar{\sigma}_i^2$ или $\overline{\omega(1-\omega)}$, а не от общей дисперсии $\sigma_{\tilde{x}}^2$ или $\omega \times (1-\omega)$.

Серийная выборка. Заключается в собственно случайном либо механическом отборе групп единиц (серий), внутри которых производится сплошное обследование. Единицей отбора при этой выборке является группа или серия, а не отдельная единица генеральной совокупности. При серийном отборе внутри отобранных групп обследуются все без исключения единицы.

Поскольку внутри серий обследуются все без исключения единицы, средняя ошибка выборки при отборе равновеликих серий зависит от величины только *межгрупповой дисперсии* ($\delta_{\tilde{x}}^2$ или δ_{ω}^2):

межгрупповая дисперсия средних $\delta_{\tilde{x}}^2 = \dfrac{\sum(\tilde{x}_i - \tilde{x})^2}{r}$,

где $\tilde{x}_i$ – средняя i-й серии;

$\tilde{x}$ – средняя по всей выборочной совокупности;

r – число отобранных серий;

межгрупповая дисперсия доли $\delta_{\omega}^2 = \dfrac{\sum(\omega_i - \bar{\omega})^2}{r}$,

где ω_i – доля признака i-й серии;

$\bar{\omega}$ – общая доля признака во всей выборочной совокупности.

Комбинированная выборка. Предполагает использование нескольких способов выборки.

При проведении выборочного наблюдения неизбежна некоторая свойственная ему погрешность (ошибки). Классификация ошибок выборочного наблюдения представлена на рисунке 11.

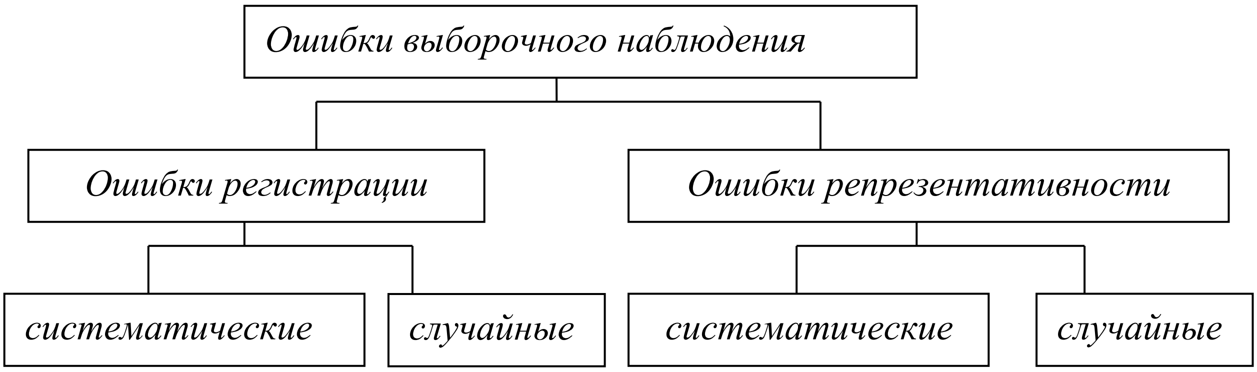

Рисунок 11 – Классификация ошибок выборочного наблюдения

Ошибки регистрации являются следствием неправильного установления значения наблюдаемого признака или неправильной записи. Случайные ошибки регистрации – это результат действия различных случайных факторов (например, перепутаны строки или графы при заполнении статистического формуляра). Такие ошибки имеют разную направленность: они могут и повышать, и понижать значения показателей.

Систематические ошибки регистрации всегда имеют одинаковую тенденцию либо к увеличению, либо к уменьшению значения показателей по каждой единице наблюдения, и поэтому величина показателя по совокупности в целом будет включать в себя накопленную ошибку.

Ошибки репрезентативности характерны только для несплошного наблюдения и обусловлены тем, что выборочная совокупность не может по всем параметрам в точности воспроизвести генеральную совокупность. Получаемые расхождения отражают, в какой степени попавшие в выборку единицы могут представлять всю генеральную совокупность.

Систематические ошибки репрезентативности связаны с нарушением принципов формирования выборочной совокупности.

Случайные ошибки репрезентативности обусловлены действием случайных факторов, не содержащих каких-либо элементов системности в направлении воздействия на рассчитываемые выборочные характеристики.

Так как случайная ошибка выборки возникает вследствие случайных различий между границами выборочной и генеральной совокупностей, то при достаточно большом объеме выборки эта ошибка будет сколь угодно мала. Случайные ошибки могут быть доведены до незначительных размеров, и их размеры и пределы можно определить с достаточной точностью на основании закона больших чисел.

Средняя (стандартная) ошибка выборки представляет собой такое расхождение между средним выборочной и генеральной совокупностей $(\tilde{x}-\bar{x})$, которое не превышает $\pm\sigma$.

Средняя ошибка выборки *зависит от объема выборки* (чем больше численность при прочих равных условиях, тем меньше величина средней ошибки выборки) и *степени варьирования признака* (чем меньше вариация признака, следовательно, и дисперсия, тем меньше ошибка выборки, и наоборот). Формулы расчета средней ошибки выборки представлены в таблице 74.

Таблица 74 – Средняя ошибка в зависимости от способа формирования выборочной совокупности

Вид выборки	Повторный отбор		Бесповторный отбор	
	для средней ($\mu_{\tilde{x}}$)	для доли (μ_{ω})	для средней ($\mu_{\tilde{x}}$)	для доли (μ_{ω})
Собственно-случайная и механическая	$\frac{\sigma}{\sqrt{n}}$	$\sqrt{\frac{w(1-w)}{n}}$	$\sqrt{\frac{\sigma^2}{n}\left(1-\frac{n}{N}\right)}$	$\sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}$
Типическая (при отборе пропорциональном объему групп)	$\sqrt{\frac{\bar{\sigma}^2}{n}}$	$\sqrt{\frac{\overline{w(1-w)}}{n}}$	$\sqrt{\frac{\bar{\sigma}^2}{n}\left(1-\frac{n}{N}\right)}$	$\sqrt{\frac{\overline{w(1-w)}}{n}\left(1-\frac{n}{N}\right)}$
Серийная (гнездовая)	$\sqrt{\frac{\delta_{\tilde{x}}^2}{r}}$	$\sqrt{\frac{\delta_{\omega}^2}{r}}$	$\sqrt{\frac{\delta_{\tilde{x}}^2}{r}\left(\frac{R-r}{R-1}\right)}$	$\sqrt{\frac{\delta_{\omega}^2}{r}\left(\frac{R-r}{R-1}\right)}$

Предельная ошибка выборки – максимально возможное расхождение выборочной и генеральной средних $(\tilde{x}-\bar{x})$, т.е. максимум ошибки при заданной вероятности ее появления.

О величине предельной ошибки можно судить с определенной вероятностью, на величину которой указывает коэффициент доверия (t).

Предельная ошибка выборки (Δ) определяется по формуле

$$\Delta_{\tilde{x}} = t\mu_{\tilde{x}} \text{ или } \Delta_{\omega} = t\mu_{\omega},$$

где Δ – предельная ошибка выборки;

t – коэффициент доверия, зависящий от вероятности, с которой гарантируется предельная ошибка выборки (см. приложение 5).

Предельная ошибка выборки позволяет определять предельные значения характеристик генеральной совокупности при заданной вероятности и их доверительные интервалы.

Интервальная оценка генеральной совокупности

$$\tilde{x} - \Delta_{\tilde{x}} \le \bar{x} \le \tilde{x} + \Delta_{\tilde{x}},$$

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

Интервальная оценка генеральной доли

$$\omega - \Delta_{\omega} \le p \le \omega + \Delta_{\omega}.$$

Пример. Для определения среднего возраста рабочих предприятия была проведена 5%-я механическая выборка. В результате обследования получены данные, приведенные в таблице 75.

Таблица 75 – Данные обследования для определения среднего возраста рабочих предприятия

Возраст рабочего, лет	18-28	28-38	38-48	Свыше 48
Число рабочих, человек	10	50	35	5

С вероятностью 0,954 нужно определить:

1) пределы, в которых находится средний возраст рабочих предприятия;

2) пределы, в которых находится доля рабочих старше 48 лет.

Решение.

1. Рассчитаем среднюю арифметическую и выборочную дисперсию по данным таблицы 76.

Таблица 76 – Данные для расчета

Границы интервалов, лет	Число рабочих, чел. (n_i)	Середина интервала, лет (x_i)	$x_i n_i$	$x_i^2 n_i$
18−28	10	23	230	5290
28−38	50	33	1650	54450
38−48	35	43	1505	64715
Свыше 48	5	53	265	14045
Итого	100	-	3650	138500

Так как численность выборочной совокупности $n = \sum n_i$, значение выборочной средней равно:

$$\tilde{x} = \frac{\sum x_i n_i}{n} = \frac{3650}{100} = 36{,}5 \text{ лет}$$

Выборочная дисперсия равна:

$$\sigma_{\tilde{x}}^2 = \frac{\sum x_i^2 n_i}{\sum n_i} - \left(\tilde{x}\right)^2 = \frac{138500}{100} - (36{,}5)^2 = 52{,}75$$

Стандартная ошибка средней равна:

$$\mu_{\tilde{x}} = \sqrt{\sigma_{\tilde{x}}^2 \div n \times \left(1 - n \div N\right)} = \sqrt{52{,}75 \div 100 \times (1 - 0{,}05)} = 0{,}7 \text{ года}$$

Значение предельной ошибки выборки для средней при $F(t) = 0{,}954$ и коэффициенте доверия $t = 2$ составит (см. приложение 5):

$$\Delta_{\tilde{x}} = \mu_{\tilde{x}} \times t = 0{,}7 \times 2 = 1{,}4 \text{ года}$$

Пределы, в которых находится средний возраст рабочих предприятия, следующие:

$$\tilde{x} - \Delta_{\tilde{x}} \le \bar{x} \le \tilde{x} + \Delta_{\tilde{x}}$$
$$36{,}5 - 1{,}4 \le \bar{x} \le 36{,}5 + 1{,}4$$
$$35{,}1 \le \bar{x} \le 37{,}9$$

Значит, с вероятностью 0,954 можно утверждать, что средний возраст рабочих предприятия, находится в интервале от 35,1 лет до 37,9 лет.

2. Выборочная доля составляет:

$$\omega = m \div n = 5 \div 100 = 0{,}05$$

Стандартная ошибка доли равна:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0{,}05(1-0{,}05)}{100}(1-0{,}05)} = 0{,}021$$

Значение предельной ошибки выборки для средней при $F(t) = 0{,}954$ и коэффициенте доверия $t = 2$ (см. приложение 5) составит:

$$\Delta_{\omega} = \mu_{\omega} \times t = 0{,}021 \times 2 = 0{,}042$$

Границы, в которых находится доля рабочих предприятия следующие:

$$\omega - \Delta_{\omega} \le p \le \omega + \Delta_{\omega}$$
$$0{,}05 - 0{,}042 \le p \le 0{,}05 + 0{,}042$$
$$0{,}008 \le p \le 0{,}092$$

С вероятностью 0,954 можно утверждать, что доля рабочих старше 48 лет будет находиться в пределах от 0,8 до 9,2 %.

При подготовке выборочного наблюдения с заранее заданным значением допустимой ошибки выборки важно правильно определить объем (численность) выборочной совокупности. Объем выборки (n) должен быть достаточным для того, чтобы обеспечить репрезентативность выборки.

Формула необходимой численности выборки для различных способов отбора выводится из формулы предельной ошибки выборки и представлена в таблице 77.

Таблица 77 – Необходимый объем (численность) выборки

Вид выборки	Повторный отбор		Бесповторный отбор	
	при определении среднего размера признака	при определении доли признака	при определении среднего размера признака	при определении доли признака
Собственно случайная и механическая	$\frac{t^2\sigma_{\tilde{x}}^{\,2}}{\Delta_{\tilde{x}}^2}$	$\frac{t^2\omega(1-\omega)}{\Delta_{\omega}^2}$	$\frac{t^2\sigma_{\tilde{x}}^2 N}{\Delta_{\tilde{x}}^2 N + t^2\sigma_{\tilde{x}}^2}$	$\frac{t^2\omega(1-\omega)N}{\Delta_{\omega}^2 N + t^2\omega(1-\omega)}$
Типическая	$\frac{t^2\overline{\sigma}_i^{\,2}}{\Delta_{\tilde{x}}^2}$	$\frac{t^2\overline{\omega(1-\omega)}}{\Delta_{\omega}^2}$	$\frac{t^2\overline{\sigma}_i^2 N}{\Delta_{\tilde{x}}^2 N + t^2\overline{\sigma}_i^{\,2}}$	$\frac{t^2\overline{\omega(1-\omega)}N}{\Delta_{\omega}^2 N + t^2\overline{\omega(1-\omega)}}$
Серийная	$\frac{t^2\delta_{\tilde{x}}^{\,2}}{\Delta_{\tilde{x}}^2}$	$\frac{t^2\delta_{\omega}^{\,2}}{\Delta_{\omega}^2}$	$\frac{t^2\delta_{\tilde{x}}^{\,2}R}{\Delta_{\tilde{x}}^2 R + t^2\delta_{\tilde{x}}^2}$	$\frac{t^2\delta_{\omega}^{\,2}R}{\Delta_{\omega}^2 R + t^2\delta_{\omega}^2}$

Пример. Для определения среднего дохода на душу населения необходимо провести выборочное наблюдение способом случайного отбора. В городе проживает 15 тыс. семей, среднее квадратическое отклонение составляет 60 руб. Определим необходимую численность выборки при условии, что с вероятностью 0,954 ошибка выборки не превысит 10 руб.

Численность выборки при случайном отборе определяется по формуле

$$n = \frac{t^2\sigma_{\tilde{x}}^2 N}{\Delta_{\tilde{x}}^2 N + t^2\sigma_{\tilde{x}}^2}\,.$$

Коэффициент доверия, зависящий от вероятности, с которой гарантируется предельная ошибка выборки, устанавливается по приложению 5: при $F(t) = 0{,}954$, $t = 2{,}0$.

Подставим данные в формулу:

$$n = \frac{t^2 \sigma_{\tilde{x}}^2 N}{\Delta_{\tilde{x}}^2 N + t^2 \sigma_{\tilde{x}}^2} = \frac{2^2 \times 60^2 \times 15000}{10^2 \times 15000 + 2^2 \times 60^2} = 142{,}6 \approx 143 \text{ семьи}$$

Таким образом, для наблюдения необходимо отобрать 143 семьи. Это позволит определить средний доход на душу населения с ошибкой не более 10 руб. Вероятность, что ошибка не превысит 10 руб., равна 0,954.

Конечная цель любого выборочного наблюдения – распространение его характеристик на генеральную совокупность. Существуют два способа распространения данных выборочного наблюдения на генеральную совокупность: прямого пересчета и поправочных коэффициентов.

Способ прямого пересчета применяется в том случае, если цель выборочного наблюдения заключается в определении объема признака генеральной совокупности, когда известна лишь численность ее единиц.

Пример. При выборочном обследовании партии нарезных батонов в 2000 ед. доля нестандартных изделий в выборке составляет: $\omega = 0{,}1$ (10:100) при установленной с вероятностью $F(t) = 0{,}954$ предельной ошибке выборки $\Delta_\omega = \pm 0{,}06$.

На основе этих данных доля нестандартных изделий во всей партии составит: $p = 0{,}1 \pm 0{,}06$, или от 0,04 до 0,16.

Способом прямого пересчета можно определить пределы абсолютной численности нестандартных изделий во всей партии:

минимальная численность, шт.: $2000 \times 0{,}04 = 80$;

максимальная численность, шт.: $2000 \times 0{,}16 = 320$.

Способ поправочных коэффициентов применяется в тех случаях, когда цель выборочного метода состоит в уточнении результатов сплошного наблюдения.

Пример. По данным годовых отчетов в сельскохозяйственных организациях имеется 127500 голов коров. В целях проверки данных сплошного учета провели контрольные обходы части обследованных предприятий и выявили, что если данные сплошного учета в организациях, попавших в выборку, показали 550 голов коров, то данные выборки в этих же организациях – 560 голов коров. Необходимо определить фактическую численность коров в сельскохозяйственных организациях.

Определим «процент недоучета» (коэффициент) при сплошном наблюдении:

$$\frac{560-550}{550}\times 100 = 1{,}82\,\%$$

Количество голов коров необходимо умножить на этот коэффициент:

$$127500\times 1{,}0182 = 129821 \text{ голов}$$

Это значит, что при сплошном учете недоучтена 2321 голова коров.

Задача 78

Планируется 25%-е собственно случайное выборочное обследование населения области. Определите, на сколько процентов ошибка такой выборки при бесповторном отборе будет меньше ошибки повторной выборки.

Задача 79

В таблице 78 представлены результаты выборочного обследования покупателей магазина (собственно случайная повторная выборка).

С вероятностью 0,997 определите границы среднего размера покупки, границы удельного веса покупок на сумму до 100 руб.

Таблица 78 − Результаты выборочного обследования покупателей магазина (данные условные)

Стоимость покупки, руб.	До 100	100−200	200−300	300 и более
Число покупателей	17	58	89	43

Задача 80

Из партии готовой продукции с целью проверки ее соответствия технологическим требованиям произведена 10%-я собственно случайная бесповторная выборка, результаты которой представлены в таблице 79.

Таблица 79 − Результаты выборочного обследования соответствия яиц технологическим требованиям

Вес яйца, г	46	47	48	49	50	51	52
Число яиц, шт.	46	123	158	97	36	18	12

Можно ли принять всю партию при условии, что доля яиц с весом 51 г и более с вероятностью 0,997 не должна превышать 8 %?

Задача 81

В результате выборочного обследования населения региона установлено, что с вероятностью 0,997 среднедушевые доходы находятся в интервале от 5380 до 6620 руб. в месяц. Определите границы среднедушевых доходов с вероятностью 0,954.

Задача 82

Определите, сколько семей необходимо охватить собственно случайной выборкой для определения доли семей, не имеющих детей, с вероятностью 0,954 и предельной ошибкой 2 %. Известно, что в регионе проживают 600 тыс. семей, а дисперсия изучаемого признака по результатам ранее проведенных обследований не превышала 0,19.

Задача 83

В организации в порядке случайной бесповторной выборки было опрошено 100 служащих из 1000, и получены данные об их доходах за октябрь. Они представлены в таблице 80.

Определите:

1) пределы, в которых находится среднемесячный доход служащих, с вероятностью 0,997;

2) долю служащих, имеющих месячный доход 6600 руб. и выше, гарантируя результат с вероятностью 0,954;

3) необходимую численность выборки при определении среднемесячного дохода служащих, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 70 руб.;

4) необходимую численность выборки при определении доли служащих с размером месячного дохода 6600 и выше, чтобы с вероятностью 0,954 предельная ошибка не превышала 7 %.

Таблица 80 − Группировка служащих по размеру дохода

Месячный доход, руб.	5800−6200	6200−6600	6600−7000	7000−7400
Число служащих	12	60	20	8

Задача 84

Для изучения дифференциации процентных ставок по вкладам населения в отделении банка проведена 10%-я механическая выборка. В таблице 81 представлено распределение вкладов по срокам хранения.

Таблица 81 − Группировка вкладов по срокам хранения

Группа вкладов по сроку хранения, дней	До 30	30−60	60−90	90−180	180−360	360 и свыше
Число вкладов	31	130	209	82	43	15

Определите:

1) средний срок хранения вкладов и долю вкладов со сроком хранения более 180 дней;

2) с вероятностью 0,954 пределы, в которых можно ожидать среднюю продолжительность хранения вклада и доли вкладов со сроком хранения более 180 дней;

3) необходимую численность выборки при определении доли вкладов со сроком хранения более 180 дней, чтобы с вероятностью 0,997 предельная ошибка выборки не превысила 5 %.

Задача 85

Какова должна быть численность выборки, чтобы с вероятностью 0,954 гарантировать, что размер ошибки выборки не превысит 0,21? При этом установлено, что дисперсия (средний квадрат отклонений) равна 1,85.

Задача 86

В поселке проживает 4200 человек. Предполагается путем случайного бесповторного отбора определить долю пенсионеров. Какова должна быть численность выборки, чтобы с вероятностью 0,997 ошибка выборки не превысила 0,02, если дисперсия равна 0,14.

Задача 87

В результате выборочного обследования сбора томатов в каждой 8-й теплице агрофирмы получены следующие предварительные данные об урожайности, представленные в таблице 82.

Таблица 82 − Результаты обследования сбора томатов в теплицах для получения предварительных данных об урожайности

Номер теплицы	1	2	3	4	5	6	7	8	9	10
Урожайность, кг на 1 кв.м	9,2	8,2	8,7	8,1	8,0	9,0	8,5	9,3	8,6	8,4

С вероятностью 0,997 определите среднюю урожайность томатов по агрофирме в целом; виды на урожай с учетом того, что площадь каждой теплицы составляет 200 кв. м.

Задача 87

Площадь яровой пшеницы в хозяйстве составляет 300 га. Какую площадь необходимо обследовать для определения урожайности, если среднее квадратическое отклонение составляет 1,2 ц с 1 га, а ошибка выборки не должна превышать 0,5 ц с 1 га при вероятности 0,997.

8 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Социально-экономические явления представляют собой результат одновременного воздействия большого числа разнообразных и взаимосвязанных факторов. Понять и изучить какое-либо явление можно, исследуя его во взаимосвязи с окружающими признаками.

В статистике различают факторные и результативные признаки. *Факторные признаки* обусловливают изменения других, связанных с ними признаков. *Результативные признаки* изменяются под действием факторных признаков.

В статистике различают:

– *функциональную связь*, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака. Функциональная связь проявляется во всех случаях наблюдения и для каждой конкретной единицы исследуемой совокупности;

– *стохастическую зависимость*, при которой причинная зависимость проявляется не в каждом отдельном случае, а в общем, среднем при большом числе наблюдений. Частным случаем стохастической зависимости является корреляционная связь, при которой изменение среднего значения результативного признака обусловлено изменением факторных признаков.

Для *функциональной связи* характерны следующие особенности:

1) каждому значению величины факторного признака соответствует только *одно* или *несколько точно определенных значений* результативного признака;

2) эта связь обычно *выражается формулами*, что в большей степени присуще точным наукам;

3) функциональная зависимость *с одинаковой силой* проявляется у всех единиц совокупности;

4) она является *полной и точной*, так как обычно известны перечень всех факторов и механизм их воздействия на результативный признак (в виде уравнения).

Корреляционные связи имеют следующие особенности:

1) средняя величина результативного признака меняется под влиянием изменения *многих факторных признаков*, ряд из которых может быть неизвестен;

2) разнообразие факторов, их взаимосвязи и противоречивое действие вызывают *широкое варьирование результативного признака*;

3) корреляционные связи обнаруживаются не в единичных случаях, а в массе, для их исследования требуются *массовые наблюдения*;

4) связь между признаками-факторами и результативным признаком *неполная*, а проявляется в общем, среднем.

В *зависимости от направления действия* функциональные и корреляционные связи делят на прямые и обратные; *по аналитическому выражению* – на прямолинейные и криволинейные.

При наличии *прямых связей* направление изменения результативного признака совпадает с направлением изменения признака-фактора. С увеличением (уменьшением) значений факторного признака происходит увеличение (уменьшение) результативного признака.

Обратные связи характеризуются тем, что направление изменения результативного признака не совпадает с направлением изменения признака-фактора. С увеличением (уменьшением) значений факторного признака происходит уменьшение (увеличение) результативного признака.

Прямолинейные связи выражаются уравнением прямой линии. При наличии этих связей с возрастанием величины факторного признака происходит непрерывное возрастание (или убывание) величин результативного признака.

Криволинейные связи выражаются уравнениями кривых линий – гиперболой, параболой и другими, поскольку с возрастанием величины факторного признака возрастание (или убывание) результативного признака происходит неравномерно, или направление его изменения меняется на обратное.

Корреляционные связи *в зависимости от количества признаков, включенных в модель*, делятся на однофакторные (парные) и многофакторные (множественные).

Однофакторные (парные) связи отражают зависимость между одним признаком-фактором и результативным признаком (при абстрагировании от влияния других признаков).

Многофакторные (множественные) связи характеризуются зависимостью между несколькими факторными признаками и результативным признаком (факторы действуют комплексно, т.е. одновременно и во взаимосвязи).

Для выражения функциональных связей применяют балансовый и индексный методы.

Метод балансовых построений широко используют для анализа связей и пропорций в экономике. Статистический баланс представляет собой систему показателей, которая состоит из двух сумм абсолютных величин, связанных знаком равенства:

$$A + C = D + E$$

Посредством балансов связывают в единую систему абсолютные величины, показывающие движение ресурсов. Суммы показателей образуют систему величин, характеризующих размер ресурсов на начало периода, поступление и выбытие, размер ресурсов на конец периода.

При рассмотрении корреляционных связей используют:

– методы взаимной сопряженности – для изучения связи между атрибутивными (качественными) признаками;

– метод параллельных рядов;

– графический метод (корреляционного поля);

– табличный метод (корреляционной таблицы);

– метод аналитических группировок;

– корреляционно-регрессионный анализ и другие методы – для выявления связей между количественными (варьирующими) признаками.

Изучение и оценка связей между атрибутивными (качественными) признаками в статистике осуществляется с использованием методов взаимной сопряженности (непараметрических методов оценки связи).

Методы взаимной сопряженности строятся на применении следующих показателей:

– коэффициента ассоциации;
– коэффициента контингенции;
– биссериального коэффициента корреляции;
– коэффициента взаимной сопряженности А.А. Чупрова;
– коэффициента взаимной сопряженности Пирсона.

Коэффициент ассоциации ($K_а$) применяется для изучения связи между качественными признаками в том случае, когда каждый признак принимает только два значения (состоит только из двух групп), т.е. позволяет изучить связь между альтернативными признаками. Коэффициент вычисляется по формуле

$$K_а = \frac{ad - bc}{ad + bc},$$

где a, b, c, d − частоты «таблицы четырех полей» (таблица 83).

Коэффициент ассоциации изменяется в пределах от (−1) до (+1). Чем ближе этот показатель к 1 или (−1), тем сильнее связаны между собой изучаемые признаки. Если коэффициент ассоциации не ниже 0,3, можно говорить о наличии существенной связи между признаками.

Коэффициент контингенции ($K_к$) определяется по формуле

$$K_к = \frac{ad - bc}{\sqrt{(a+b)\cdot(b+d)\cdot(a+c)\cdot(c+d)}}$$

Коэффициент контингенции применяются в том случае, когда хотя бы одно значение из четырех показателей в «таблице четырех полей» отсутствует.

Таблица 83 − Таблица для вычисления коэффициентов ассоциации и контингенции

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	$a+b+c+d$

По абсолютной величине коэффициент контингенции всегда меньше коэффициента ассоциации и изменяется от (−1) до (+1). Чем ближе к 1 или (−1), тем сильнее связаны между собой изучаемые признаки.

Пример. В результате исследования крупнейших страховых компаний получены нижеследующие данные (таблица 84).

Таблица 84 – Группировка страховых компаний по классу устойчивости

Группа страховых компаний по классу устойчивости	По величине выплат, %		
	< 30	> 30	Итого
1−3	6	22	28
4−5	11	14	25
Итого	17	36	53

Определите степень тесноты связи меду признаками с использованием коэффициентов ассоциации и контингенции.

1. Коэффициент ассоциации определяется по формуле

$$K_{\text{а}} = \frac{ad - bc}{ad + bc} = \frac{6 \times 14 - 22 \times 11}{6 \times 14 + 22 \times 11} = \frac{-158}{326} = -0{,}48$$

Связь между признаками разнонаправленная, и степень тесноты существенная.

2. Коэффициент контингенции определяется по формуле

$$K_{\text{к}} = \frac{ad - bc}{\sqrt{(a+b)\cdot(b+d)\cdot(a+c)\cdot(c+d)}} =$$
$$= \frac{6 \times 14 - 22 \times 11}{\sqrt{(6+22)\times(6+11)\times(14+22)\times(14+11)}} = -0{,}24$$

Связь между признаками присутствует, и она разнонаправленная.

Биссериальный коэффициент корреляции (r) позволяет изучить связь между качественным альтернативным и количественным варьирующим признаками и определяется по формуле

$$r = \frac{\left|\overline{Y}_2 - \overline{Y}_1\right|}{\sigma_Y} \times \frac{pq}{Z},$$

где $\overline{Y}_2, \overline{Y}_1$ – средние значения признака в группах;

σ_Y – среднее квадратическое отклонение фактических значений признака от среднего уровня;

p – доля первой группы в совокупности;

q – доля второй группы;

Z – табличные значения Z-распределения в зависимости от p.

Величина коэффициента варьирует в пределах от 0 до 1.

Коэффициент взаимной сопряженности А.А.Чупрова применяется для измерения тесноты связи между варьированием двух атрибутивных признаков, когда это варьирование образует несколько (три и более) групп и определяется по формуле

$$K_{\text{ч}} = \sqrt{\frac{\chi^2}{\sqrt{(m_1 - 1) \times (m_2 - 1)}}} = \sqrt{\frac{\chi^2}{n \times \sqrt{(m_1 - 1) \times (m_2 - 1)}}},$$

где χ^2 – хи-квадрат;

$$\chi^2 = \left(\sum_1^{m_1} \sum_1^{m_2} \frac{f_{ij}^2}{f_i \times f_j} \right) - 1,$$

где f_i f_j – эмпирические частоты в i-й строке j-м столбце;

m – число групп по каждому признаку;

n – количество наблюдений.

Числовое значение коэффициента изменяется от 0 до 1, но уже при значении 0,3 можно говорить о тесной связи между вариацией изучаемых качественных признаков.

Пример. В результате исследования крупнейших страховых компаний получены нижеследующие данные (таблица 85).

Определите связь между признаками с использованием коэффициента взаимной сопряженности А.А. Чупрова.

Таблица 85 – Группировка страховых компаний по классу устойчивости

Группа страховых компаний по классу устойчивости	По величине выплат, %			
	< 30	30−40	> 40	Итого
1–2	2	6	10	18
3–4	14	9	7	30
5	1	2	2	5
Итого	17	17	19	53

Коэффициент взаимной сопряженности А.А. Чупрова определяется по формуле

$$K_ч = \sqrt{\frac{\chi^2}{\sqrt{(m_1-1)\times(m_2-1)}}}.$$

Для его расчета необходимо определить хи-квадрат по формуле:

$$\chi^2 = \left(\sum_1^{m_1}\sum_1^{m_2}\frac{f_{ij}^2}{f_i \times f_j}\right) - 1 =$$

$$= \left(\begin{array}{l} \frac{10^2}{19\times18} + \frac{14^2}{17\times30} + \frac{9^2}{17\times30} + \frac{2^2}{17\times18} + \frac{6^2}{17\times18} + \\ + \frac{7^2}{19\times30} + \frac{1^2}{17\times5} + \frac{2^2}{17\times5} + \frac{2^2}{19\times5} - 1 \end{array}\right) = 0{,}153$$

Следовательно, коэффициент А.А. Чупрова равен:

$$K_ч = \sqrt{\frac{0{,}153}{\sqrt{(3-1)\times(3-1)}}} = 0{,}277.$$

Связь между признаками средней степени тесноты.

Коэффициент взаимной сопряженности Пирсона ($K_п$) определяется по формуле

$$K_n = \sqrt{\frac{\chi^2}{\chi^2 + n}},$$

где n – количество наблюдений.

Коэффициент изменяется от 0 до 1. Чем ближе к единице, тем теснее связь между атрибутивными признаками.

Метод приведения параллельных данных. Метод основан на сопоставлении двух или нескольких рядов статистических величин. Такое сопоставление позволяет установить наличие связи и получить представление о ее характере.

Взаимосвязь двух признаков изображается графически с помощью поля корреляции. В системе координат на оси абсцисс откладываются значения факторного признака, а на оси ординат – результативного. Каждое пересечение линий, проводимых через эти оси, обозначается точкой. При отсутствии тесных связей наблюдается беспорядочное расположение точек на графике. Чем сильнее связь между признаками, тем теснее будут группироваться точки вокруг определенной линии, выражающей форму связи.

На основе сравнения параллельных рядов можно применить элементарные показатели, характеризующие направление и тесноту связи:

– коэффициент Фехнера (коэффициент корреляции знаков);
– коэффициент Спирмена (коэффициент корреляции рангов);
– множественный коэффициент ранговой корреляции (коэффициент конкордации).

Коэффициент Фехнера ($K_ф$) основан на степени согласованности направлений отклонений индивидуальных значений факторного и результативного признаков от соответствующих средних величин. Для расчета этого показателя исчисляют средние значения факторного и результативного признаков (по арифметической простой), а затем проставляют знаки отклонений для значений взаимосвязанных пар признаков (если фактическое значение признака больше средней величины, то ставится знак «+», если меньше, то знак «−»). Коэффициент определяется по формуле следующего вида:

$$K_{\phi} = \frac{C - H}{C + H},$$

где C – число совпадений знаков отклонений;

H – число несовпадений знаков отклонений.

Значение коэффициента Фехнера находится в пределах от −1 до +1, т.е. $-1 \le К_{ф} \le +1$. Если: $К_{ф} = \pm 1$, то связь между между признаками функциональная; $К_{ф} = 0$, то связь отсутствует.

Данный показатель позволяет уловить направление вариации, но не учесть точно ее величину.

Пример. Имеются данные по совокупности стран о размере ВВП на душу населения по полной первоначальной стоимости основных фондов и фактическом конечном потреблении домашних хозяйств по полной первоначальной стоимости основных фондов (таблица 86).

Определите степень тесноты связи между признаками с использованием коэффициента Фехнера.

Решение.

1. Определим средние значения по факторному и результативному признакам (по арифметической простой):

$$\bar{x} = \frac{\sum x}{n} = \frac{58800}{8} = 7350$$

$$\bar{y} = \frac{\sum y}{n} = \frac{454}{8} = 56{,}75$$

2. Сопоставляя фактические значения с их средними, расставим знаки, например, 14400 больше 7350, значит, знак «+», 2390 меньше 7350, значит, знак «−». В результате получаем семь совпадений знаков и одно несовпадение.

3. Коэффициент Фехнера определяется по формуле

$$K_{\text{ф}} = \frac{C - H}{C + H} = \frac{7 - 1}{7 + 1} = 0{,}75$$

Можно предположить наличие тесной и однонаправленной связи между признаками.

Таблица 86 – Размер ВВП на душу населения по полной первоначальной стоимости основных фондов и фактическом конечном потреблении домашних хозяйств по полной первоначальной стоимости основных фондов

Страна	Исходные данные		Расчетные данные	
	ВВП на душу населения по полной первоначальной стоимости основных фондов, долл. (*x*)	фактическом конечном потреблении домашних хозяйств по полной первоначальной стоимости основных фондов к РФ, % (*y*)	(*x*)	(*y*)
РФ	14400	100	+	+
Белоруссия	10740	85	+-	+
Молдавия	2930	37	-	-
Украина	6810	59	-	+
Армения	5900	34	-	-
Азербайджан	6370	46	-	-
Казахстан	9700	69	+	+
Киргизия	1950	24	-	-

Коэффициент корреляции рангов (коэффициент Спирмэна) (ρ) применяют для анализа связи двух признаков (*x, y*). Он учитывает согласованность рангов, т.е. номеров, которые занимают единицы совокупности по каждому из этих признаков, и определяется по формуле

$$\rho = 1 - \frac{6\sum d^2}{n^3 - n} \text{ или } \rho = 1 - \frac{6\sum d^2}{n(n^2 - 1)},$$

где d – разность рангов x и y;

n – число наблюдаемых пар значений x и y.

Коэффициент Спирмэна принимает любые значения в интервале от (+1) (полная корреляция рангов, в этом случае $\sum d = 0$) до (−1) (полная обратная корреляция рангов, в этом случае $\frac{6\sum d^2}{n^3 - n} = 2$. При ρ = 0, когда $\frac{6\sum d^2}{n^3 - n} = 1$, корреляция рангов отсутствует.

Значимость коэффициента корреляции рангов Спирмэна проверяется на основе t-критерия Стьюдента. Расчетное значение критерия определяется по формуле

$$t_{\text{р}} = \rho_{x/y} \sqrt{\frac{n-2}{1-\rho^2_{X/Y}}}$$

Значение коэффициента корреляции считается статистически существенным, если $t_{\text{р}} > t_{\text{кр}}$ *(α; k = n−2)*.

Пример. Определите степень тесноты связи между признаками с использованием коэффициента Спирмена на основании данных по совокупности стран ближнего зарубежья о размере ВВП на душу населения по полной первоначальной стоимости основных фондов и уровню безработицы, представленных в таблице 87.

Решение.

1. Расставим ранги в порядке возрастания величин факторного и результативного признаков и рассчитаем $\sum d^2$:

$$\sum d^2 = \sum (x-y)^2 = 148.$$

2. Определим коэффициент Спирмена:

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \times 148}{8 \times (8^2 - 1)} = -0{,}76.$$

Связь между признаками довольно тесная и разнонаправленная.

Таблица 87 − Размер ВВП на душу населения по полной первоначальной стоимости основных фондов и уровень безработицы в странах ближнего зарубежья

Страна	Исходные данные		Расчетные данные		$\sum(x-y)^2$
	ВВП на душу населения по полной первоначальной стоимости основных фондов, долл.	уровень безработицы, %	ранг по ВВП на душу населения по полной первоначальной стоимости основных фондов, долл. (x)	ранг по уровню безработицы, % (y)	
РФ	14400	13,4	8	3	25
Белоруссия	10740	7,7	7	1	36
Молдавия	2930	25,8	2	5	9
Украина	6810	29,3	5	7	4
Армения	5900	26,5	3	6	9
Азербайджан	6370	15,8	4	4	0
Казахстан	9700	12,7	6	2	16
Киргизия	1950	35,0	1	8	49
Итого	-	-	-	-	148

Множественный коэффициент ранговой корреляции (коэффициент конкордации) (ω) применяется для оценки тесноты связи между несколькими (тремя и более) признаками при использовании ранговой корреляции. Коэффициент изменяется в пределах от 0 до 1 и характеризует степень тесноты связи, но уже при значении 0,5 можно говорить о тесной связи между вариацией изучаемых признаков.

Множественный коэффициент ранговой корреляции (коэффициент конкордации) определяется по формуле

$$\omega = \frac{12S}{m^2 \cdot (n^3 - n)},$$

где m – количество факторов;

n – число наблюдений;

S – отклонение суммы квадратов рангов от средней квадратов рангов.

$$S = \sum_{1}^{n}\left(\sum_{1}^{m} r_{ij}\right)^2 - \frac{\left(\sum_{1}^{n}\sum_{1}^{m} r_{ij}\right)^2}{n},$$

где r_{ij} – ранг i-го фактора у j-й единицы.

Значимость коэффициента конкордации проверяется на основе χ^2-критерия Пирсона:

$$\chi^2 = \frac{12S}{m \cdot n(n-1)}.$$

Если фактическое значение хи-квадрата больше табличного значения (приложение 4), при вероятности $L = 0{,}05 (0{,}01; 0{,}10)$ и числе степеней свободы $\nu = \eta - 1$, то это подтверждает значимость коэффициента конкордации.

В случае наличия связанных рангов коэффициент конкордации определяется по формуле

$$\omega = \frac{12S}{m^2 \cdot \left(n^3 - n\right) - m \cdot \sum_{j=1}^{m} Tj},$$

где $Tj = \frac{1}{12}\sum_{j=1}^{m}\left(t_j^3 - t_i\right);$

t – количество связанных рангов по отдельным показателям.

Проверка значимости осуществляется по формуле

$$\chi^2 = \frac{12S}{m \cdot n(n+1) - \frac{1}{n-1} \cdot \sum_{j=1}^{m} Tj}.$$

Ранговые коэффициенты корреляции Спирмэна и конкордации имеют преимущество, что с помощью их можно измерять и оценивать связи как между количественными, так и между атрибутивными признаками, которые поддаются ранжированию.

Связь между признаками можно наглядно увидеть, если построить *график*, отложив на оси абсцисс значения факторного признака (x), а на оси ординат – значения результативного признака (y). Нанеся на графике точки, соответствующие значениям x и y, можно получить корреляционное поле, по характеру расположения точек в котором можно судить о направлении и силе связи. Если точки беспорядочно разбросаны по всему полю, это говорит об отсутствии зависимости между двумя признаками. Если точки концентрируются вокруг оси, идущей от нижнего левого угла в верхний правый, то имеется прямая зависимость между варьирующими признаками. Если точки будут концентрироваться вокруг оси, идущей от верхнего левого угла в нижний правый, то существует обратная зависимость.

Табличный метод (корреляционной таблицы) применяется при наличии большого числа значений результативного признака, соответствующих одному значению факторного признака. Для построения корреляционной таблицы сначала осуществляют группировку совокупности по факторному и результативному признакам (комбинированную группировку), а затем в графах таблицы располагают группы по факторному признаку. Числа, расположенные на пересечении граф и строк, показывают частоту повторения данного сочетания значений x и y.

Направление расположения частот корреляционной таблицы дает возможность предполагать наличие или отсутствие связи, а также ее направление.

Если частоты располагаются по диагонали из левого верхнего угла в правый нижний угол таблицы (т.е. с ростом значений факторного признака растут значения результативного), то можно предполагать наличие прямой однонаправленной связи. Если частоты располагаются по другой диагонали, то предполагают наличие прямой обратной связи.

Пример. Постройте корреляционную таблицу на основании данных о численности врачей (на 10000 человек) и заболеваемости населения (на 1000 человек) по совокупности регионов РФ в 2007 г. представленных в таблице 88.

Таблица 88 − Численность врачей и заболеваемость по совокупности регионов РФ в 2007 г.

Регион	Численность врачей на 10000 человек населения, чел. (x)	Заболеваемость на 1000 человек населения, чел. (y)
1	8,2	506,9
2	8,4	820,3
3	11,8	558,0
4	13,8	684,0
5	14,2	824,5
6	14,7	990,3
7	15,7	886,0
8	17,0	1010,8
9	17,5	1056,9
10	17,8	904,8
11	18,2	793,2
12	18,9	914,0
13	19,0	910,7
14	20,9	1214,9

Решение.

1. Проведем группировку по факторному признаку:

$$n = 4,$$

$$h = \frac{20{,}9 - 8{,}2}{4} = 3{,}175.$$

Границы групп:
1-я группа: 8,200 − 11,375
2-я группа: 11,375 – 14,550
3-я группа: 14,550 – 17,725
4-я группа: 17,725 – 20,900

2. Проведем группировку по результативному признаку:

$$n = 4$$

$$h = \frac{1214{,}9 - 506{,}9}{4} = 177.$$

Границы групп:
1-я группа: 506,9 – 683,9
2-я группа: 683,9 – 860,9
3-я группа:860,9 – 1037,9
4-я группа:1037,9 – 1214,9
3. Сформируем корреляционную таблицу.

Таблица 89 – Корреляционная таблица

	8,2 – 11,375	11,375 – 14,550	14,550 – 17,725	17,725 – 20,900	$\sum f_y$
506,9 – 683,9	1	1			2
683,9 – 860,9	1	2		1	4
860,9 – 1037,9			3	3	6
1037,9 – 1214,9			1	1	2
$\sum f_x$	2	3	4	5	14

Частоты в таблице 89 расположились по диагонали из левого верхнего угла в правый нижний угол таблицы (с ростом значений факторного признака растут значения результативного), т.е. имеет место быть прямая однонаправленная связь.

Существуют ситуации, когда все клетки корреляционной таблицы заполнены. Однако это не говорит об отсутствии связи. В этих ситуациях необходимо установить, как расположена в таблице основная масса случаев. Для этого в каждой графе рассчитывают средние значения результативного признака, соответствующие определенному значению фактора (по формуле арифметической средней взвешенной). По изменению размера значений результативного признака судят о форме и направлении связи.

Корреляционно-регрессионный анализ *предполагает установление аналитической формы связи (регрессионный анализ) и измерение тесноты, направление связи (корреляционный анализ)*.

Регрессионный анализ заключается в определении аналитического выражения связи, в котором изменение одной величины (называемой зависимой или результативным признаком) обусловлено влиянием одной или нескольких независимых величин (факторов), а множество всех прочих факторов, также оказывающих влияние на зависимую величину, принимается за постоянные и средние значения. Регрессия может быть однофакторной (парной) и многофакторной (множественной).

По форме зависимости различают: линейную регрессию, нелинейную регрессию.

По направлениям связи различают:

– прямую регрессию (положительную), возникающую при условии, что если с увеличением или уменьшением независимой величины значения зависимой также соответственно увеличиваются или уменьшаются;

– обратную (отрицательную) регрессию, проявляющуюся при условии, что с увеличением или уменьшением независимой величины зависимая соответственно уменьшается или увеличивается.

Парная регрессия характеризует связь между двумя признаками – результативным и факторным. Аналитическая связь между ними описывается следующими уравнениями:

прямой $\bar{y}_x = a_0 + a_1 x$;

параболой $\bar{y}_x = a_0 + a_1 x + a_2 x^2$;

гиперболой $\bar{y}_x = a_0 + \dfrac{a_1}{x}$ и т.д.

Оценка параметров уравнений регрессии (a_0, a_1 и a_2 в уравнении параболы второго порядка) осуществляется методом наименьших квадратов, в основе которого лежит предположение о независимости наблюдений исследуемой совокупности.

Если результативный и факторный признаки возрастают одинаково, примерно в арифметической прогрессии, то это свидетельствует о том, что связь между ними линейная, а при обратной связи – гиперболическая. Если факторный признак увеличивается в арифметической прогрессии, а результативный – значительно быстрее, то используется связь параболическая или степенная.

В уравнениях регрессии параметр:

a_0 – показывает усредненное влияние на результативный признак неучтенных (не выделенных для исследования) факторов;

a_1 (а в уравнениях параболы a_2) – коэффициент регрессии который показывает, насколько изменяется в среднем значение результативного признака при увеличении факторного на единицу собственного измерения.

Система нормальных уравнений для нахождения параметров линейной парной регрессии методом наименьших квадратов имеет следующий вид:

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

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

Если связь между признаками X и Y криволинейная и описывается уравнением параболы второго порядка, то система нормальных уравнений имеет вид

$$\begin{cases} na_0 + a_1\sum x + a_2\sum x^2 = \sum y \\ a_0\sum x + a_1\sum x^2 + a_2\sum x^3 = \sum xy \\ a_0\sum x^2 + a_1\sum x^3 + a_2\sum x^4 = \sum x^2 y \end{cases}$$

Система нормальных уравнений для нахождения параметров гиперболы

$$\begin{cases} na_0 + a_1\sum \frac{1}{x} = \sum y, \\ a_0\sum \frac{1}{x} + a_1\sum \frac{1}{x^2} = \sum \frac{y}{x} \end{cases}$$

Интерпретация моделей регрессии осуществляется методами той отрасли знаний, к которой относятся исследуемые явления. Но всякая интерпретация начинается со статистической оценки уравнения регрессии в целом и оценки значимости входящих в модель факторных признаков, т.е. выяснения, как они влияют на величину результативного признака. Чем больше величина коэффициента регрессии, тем значительнее влияние данного признака на моделируемый. Особое значение при этом имеет знак перед коэффициентом регрессии. Знаки коэффициентов регрессии говорят о характере влияния на результативный признак. Если факторный признак имеет знак плюс, то с увеличением данного фактора результативный признак возрастает. Если факторный признак имеет знак минус, то с его увеличением результативный признак уменьшается.

Множественная (многофакторная) регрессия – это изучение связи между тремя и более связанными между собой признаками, она описывается функцией вида

$$\bar{y}_{1,2,...,k} = f\left(x_1, x_2, ..., x_k\right).$$

Построение моделей множественной регрессии включает несколько этапов:

– выбор формы связи (уравнения регрессии);

– отбор факторных признаков;

– обеспечение достаточного объема совокупности для получения несмещенных оценок.

Зависимости между социально-экономическими явлениями описываются следующими типами моделей:

линейной: $\bar{y}_{1,2,...k} = a_0 + a_1x_1 + a_2x_2 + ... + a_kx_k$;

степенной: $\bar{y}_{1,2,...,k} = a_0x_1^{a_1} \times x_2^{a_2}, ..., x_k^{a_k}$

показательной: $\bar{y}_{1,2,...,k} = e^{a_0 + a_1x\ + a_2x_2 + ... + a_kx_k}$

параболической: $\bar{y}_{1,2,...k} = a_0 + a_1x_1^2 + a_2x_2^2 + ... + a_kx_k^2$

гиперболической: $\bar{y}_{1,2,...k} = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + ... + \frac{a_k}{x_k}$

Наиболее приемлемым способом отбора факторных признаков является метод шаговой регрессии. Сущность метода заключается в последовательном включении факторов в уравнение регрессии и последующей проверке их значимости.

Система нормальных уравнений для нахождения параметров линейной множественной регрессии методом наименьших квадратов имеет следующий вид:

$$\begin{cases} na_0 + a_1\sum x_1 + a_2\sum x_2 + \ldots + a_i\sum x_i + \ldots + a_k\sum x_k = \sum Y, \\ a_0\sum x_1 + a_1\sum x_1^2 + a_2\sum x_2 x_1 + \ldots + a_i\sum x_i x_1 + \ldots + a_k\sum x_k x_1 = \sum Y x_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\sum x_k + a_1\sum x_1 x_k + a_2\sum x_2 x_k + \ldots + a_i\sum x_i x_k + \ldots + a_k\sum x_k^2 = \sum Y x_k \end{cases}$$

Фактор является незначимым, если его включение в уравнение регрессии только изменяет значение коэффициентов регрессии, не уменьшая суммы квадратов остатков и не увеличивая их значения. Если при включении в модель соответствующего факторного признака величина множественного коэффициента корреляции увеличивается, а коэффициент регрессии не изменяется (или меняется несущественно), то данный признак существенен, и его включение в уравнение регрессии необходимо.

Если же при включении в модель факторного признака коэффициенты регрессии меняют не только величину, но и знаки, а множественный коэффициент корреляции не возрастает, то данный факторный признак признается нецелесообразным для включения в модель связи.

Сложность и взаимное переплетение отдельных факторов, обуславливающих исследуемое экономическое явление, могут проявляться в *мультиколлинеарности* – тесной связи между факторными признаками, включенными в модель. Одним из индикаторов определения наличия мультиколлинеарности между признаками является превышение парным коэффициентом корреляции величины 0,8 ($r_{x_i x_j}$) и др.

Устранение мультиколлениарности осуществляется путем исключения из корреляционной модели одного или нескольких линейно-связанных факторных признаков или преобразование исходных факторных признаков в новые, укрупненные факторы.

Проверка адекватности моделей, построенных на основе уравнений регрессии, начинается с проверки значимости каждого коэффициента регрессии.

Значимость коэффициентов регрессии осуществляется с помощью t-критерия Стьюдента:

$$t_{\text{р}} = \frac{|a_i|}{\sqrt{\sigma_{a_i}^2}},$$

где $\sigma_{a_i}^2$ – дисперсия коэффициента регрессии, определяемая по формуле

$$\sigma_{a_i}^2 = \frac{\sigma_Y^2}{k},$$

где σ_Y^2 – дисперсия результативного признака;

k – число факторных признаков в уравнении.

Параметр модели признается статистически значимым, если

$$t_p > t_{kp}\ (\alpha;\ v = n - k - 1),$$

где α – уровень значимости;

$v = n - k - 1$ – число степеней свободы.

Табличное значение (t_{kp}) находится по приложению 6.

Проверка *адекватности всей модели* осуществляется с помощью расчета *F*-критерия Фишера:

$$F = \frac{\delta_{yx}^2}{\sigma_{ocm}^2} \div \frac{k}{n-k-1} \text{ или } F = \frac{r_{yx}^2}{1-r_{yx}^2} \div \frac{k}{n-k-1}$$

входные параметры $\alpha, \nu_1 = k+1, \nu_2 = n-k-1,$

где k – число факторных признаков в уравнении.

n – количество единиц совокупности;

r_{yx}^2 – коэффициенты детерминации.

Если $F_{факт} > F_{табл}$, то можно говорить о надежности построенного уравнения. $F_{табл}$ определяется по приложению 7.

Корреляционный анализ имеет задачей количественное определение тесноты связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи).

В статистике различают следующие варианты зависимостей:

– парная корреляция – связь между двумя признаками (результативным и факторным или двумя факторными);

– частная корреляция – зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков;

– множественная корреляция – зависимость результативного и двух или более факторных признаков, включенных в исследование.

На практике различают следующие показатели корреляции:

– линейный коэффициент корреляции;

– эмпирическое корреляционное отношение;

– индекс корреляции (теоретическое корреляционное отношение);

– частные коэффициенты корреляции;

– коэффициент множественной корреляции;

– коэффициент детерминации.

Линейный коэффициент корреляции используется для изучения связи между двумя признаками при наличии между ними линейной зависимости и определяется по формуле

$$r_{yx} = \frac{\overline{yx} - \bar{y} \times \bar{x}}{\sigma_x \sigma_y},$$

где $\bar{x}, \bar{y}, \overline{yx}$ – средние значения факторного, результативного признаков и произведения факторного и результативного признаков;

σ_x, σ_y − средние квадратические отклонения факторного и результативного признаков.

Линейный коэффициент корреляции изменяется в пределах от (−1) до (+1): $-1 \leq r \leq 1$. Интерпретация выходных значений представлена в таблице 90.

Таблица 90 − Оценка линейного коэффициента корреляции

Значение коэффициента связи	Характер связи	Интерпретация связи
$r = 0$	Отсутствует	-
$0 < r < 1$	Прямая, однонаправленная	С увеличением x увеличивается y
$-1 < r < 0$	Обратная	С увеличением x уменьшается y, и наоборот
$r = 1$	Функциональная	Каждому значению факторного признака строго соответствует одно значение результативного признака

Производя расчет по итоговым значениям исходных переменных, линейный коэффициент корреляции можно вычислить по формуле

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

В том случае, когда исходная информация представлена в виде корреляционной таблицы, необходимо учесть частоты повторений индивидуальных значений факторного и результативного признаков, а также число повторений данного сочетания их значений. В этом случае формула коэффициента корреляции будет иметь следующий вид:

$$r_{xy}=\frac{n\cdot\sum xyf_{xy}-\sum xf_x\cdot\sum yf_y}{\sqrt{\left[n\sum x^2f_x-\left(\sum xf_x\right)^2\right]\cdot\left[n\sum y^2f_y-\left(\sum yf_y\right)^2\right]}}.$$

Линейный коэффициент корреляции может быть также выражен через дисперсии слагаемых:

$$r_{xy}=\frac{\sigma_x^2+\sigma_y^2-\sigma_{x-y}^2}{2\sigma_x\sigma_y}.$$

Эмпирическое корреляционное отношение (η) рассчитывается по данным группировки в случае нелинейной зависимости между признаками и определяется по формуле

$$\eta=\sqrt{\frac{\delta^2}{\sigma^2}},$$

где δ^2 – межгрупповая дисперсия, характеризующая вариацию результативного признака, обусловленную группировочным признаком;

σ^2 – общая дисперсия результативного признака.

Показатель изменяется в пределах от 0 до 1. Интерпретация значений коэффициента представлена в таблице 91.

Таблица 91 − Оценка эмпирического корреляционного отклонения

Значение коэффициента	Интерпретация связи
0,1–0,3	Слабая
0,3–0,5	Умеренная
0,5–0,7	Заметная
0,7–0,9	Тесная
0,9–0,99	Весьма тесная

Индекс корреляции (теоретическое корреляционное отношение) (R) используют для измерения связи при любой ее форме и определяется по формуле

$$R = \sqrt{\frac{\sigma_y^2 - \sigma_{y-y_x}^2}{\sigma_y^2}},$$

где $\sigma_{y-y_x}^2 = \frac{\sum (y - y_x)^2}{n}$ – дисперсия отклонений;

σ_y^2 – дисперсия фактических значений результативного признака;

$\sigma_{y_x}^2 = \sigma_y^2 - \sigma_{y-y_x}^2$ – дисперсия теоретических значений результативного признака.

Индекс корреляции изменяется в пределах от 0 до 1. Если он равен нулю, то связи между вариацией признаков x и y нет. Чем он ближе к 1, тем связь между признаками теснее.

Частные коэффициенты корреляции (r) применяются для характеристики тесноты связи между двумя признаками при фиксированном значении других признаков и определяется по формуле

$$r_{1,2,3,\ldots,n} = \frac{r_{1,2,3,\ldots,n-1} - r_{1,n,3,\ldots,n-1} \times r_{2,n,3,\ldots,n-1}}{\sqrt{\left(1 - r_{1,n,3,\ldots,n-1}^2\right) \times \left(1 - r_{2,n,3,\ldots,n-1}^2\right)}}.$$

Коэффициент множественной корреляции (R) применяется в случае оценки связи между результативным (Y) и двумя факторными (X_1, X_2) признаками. Множественный коэффициент корреляции имеет вид

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2 \times r_{yx_1} \times r_{yx_2} \times r_{x_1x_2}}{1 - r_{x_1x_2}^2}},$$

где r – парные коэффициенты корреляции между признаками.

Значения коэффициента находятся в пределах от 0 до 1. Чем ближе значение коэффициента к единице, тем теснее связь между признаками. Множественный коэффициент корреляции можно рассчитать, используя парные коэффициенты корреляции *(r_{yx})* и β-коэффициенты:

$$R_{x_1,x_2,\ldots x_n\ldots} = \sqrt{\beta_1 \times r_{yx_1} + \beta_2 \times r_{yx_2} + \ldots + \beta_n \times r_{yx_n}}\ .$$

Коэффициент детерминации (d) показывает, какая доля вариации изучаемого результативного признака объясняется влиянием факторов, включенных в уравнение множественной регрессии, и представляет собой квадрат коэффициента корреляции:

$$d = r^2 \times 100\,\%\,,$$

$$d = R^2 \times 100\,\%$$

Коэффициент изменяется в пределах от 0 до 100 и показывает, на сколько процентов изменение результативного признака зависит от выбранных в модель факторных признаков. Остальные проценты (до 100) отражают влияние других, не учтенных в модели признаков.

Значимость линейного коэффициента корреляции проверяется на основе t-критерия Стьюдента (приложение 6).

Если объем совокупности (n) <50 единиц, то формула критерия имеет вид

$$t_{\text{расч}} = \sqrt{\frac{r^2}{1-r^2} \times (n-2)} = \frac{|r|}{\sqrt{1-r^2}} \times \sqrt{n-2}$$

входные параметры $\alpha, k = n-2$.

Если расчетное значение $t_{\text{расч}} > t_{\text{табл}}$, то коэффициент корреляции принято считать значимым.

Если объем совокупности (n) более 100 единиц, то используется формула

$$t_{\text{расч}} = \frac{|r|}{\sqrt{1-r^2}} \times \sqrt{n}\ .$$

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

$$F_p = \frac{\frac{1}{2} R^2_{yx_1x_2}}{\frac{1}{n-3} \cdot \left(1 - R^2_{yx_1x_2}\right)}$$

входные параметры $\alpha, \nu_1 = 2, \nu_2 = n-3$.

Если $F_{\text{расч}} > F_{\text{табл}}$, то коэффициент множественной корреляции считается значимым. $F_{\text{табл}}$ определяется по приложению 7.

Для выявления влияния каждого отдельного фактора на результативный признак вычисляют *стандартизированные коэффициенты:*

– коэффициенты эластичности (α-коэффициент);

– β-коэффициенты.

Коэффициенты эластичности (Э) (α-коэффициент), определяется по формуле

$$Э_{x_i} = a_i \times \frac{\bar{x}_i}{\bar{y}},$$

где $\bar{x}_i$ – среднее значение соответствующего факторного признака;

$\bar{y}$ – среднее значение результативного признака;

a_i – коэффициент регрессии при соответствующем факторном признаке.

Коэффициент эластичности показывает, на сколько процентов в среднем изменяется значение результативного признака при изменении факторного признака на 1 %.

β-коэффициент определяется по формуле

$$\beta_i = a_i \times \frac{\sigma_{x_i}}{\sigma_y}$$

где σ_{x_i} – среднее квадратическое отклонение i-го фактора;

σ_y – среднее квадратическое отклонение результативного признака.

β-коэффициент показывает, на какую часть среднего квадратического отклонения изменится результативный признак при изменении соответствующего факторного признака на свое среднее квадратическое отклонение.

Пример. По данным о сумме активов и заемных средств по семи сельскохозяйственным организациям Пензенской области определим:

1) зависимость между признаками с использованием уравнения парной линейной регрессии;

2) линейный коэффициент корреляции и детерминации;

3) коэффициент эластичности и β-коэффициент.

Исходные данные и расчетные показатели представим в таблице 92.

Таблица 92 − Уровень активов и заемных средств сельскохозяйственных организаций в Пензенской области на 01.01.2010 г., (млн. руб.)

№ предприятия	Заемные средства, (*x*)	Сумма активов, (*y*)	x^2	xy	y^2	$\bar{y}_x$
1	12,8	81,1	163,84	1038,08	6577,21	64,63
2	43,2	85,0	1866,24	3672	7225,00	81,17
3	7,3	48,3	53,29	352,59	2332,89	61,64
4	61,9	76,6	3831,61	4766,3	5867,56	91,34
5	13,8	58,1	190,44	801,78	3375,61	65,18
6	25,3	64,3	640,09	1626,79	4134,49	71,43
7	40,3	101,6	1624,09	4094,48	10322,56	79,59
Итого	204,6	515	8369,6	16352,02	39835,32	515

1. Определим параметры уравнения с помощью системы нормальных уравнений:

$$\begin{cases} 7a_0 + a_1 204{,}6 = 515 \\ a_0 204{,}6 + a_1 8369{,}6 = 16352{,}02 \end{cases}$$

a_0 = 57,67; a_1 = 0,544

Подставим полученные значения в уравнение линейной регрессии $\bar{y}_x = a_0 + a_1 x$:

$$\bar{y}_x = 57{,}67 + 0{,}544x$$

С увеличением заемных средств на 1 млн. руб. сумма активов возрастет в среднем на 0,544 млн руб.

2. Для расчета линейного коэффициента корреляции исчислим необходимые составляющие:

$$\bar{x} = \frac{204{,}6}{7} = 29{,}23$$

$$\bar{y} = \frac{515}{7} = 73{,}57$$

$$\overline{xy} = \frac{16352{,}02}{7} = 2336{,}00$$

$$\sigma_x = \sqrt{\frac{8369{,}6}{7} - 29{,}23^2} = 18{,}47$$

$$\sigma_y = \sqrt{\frac{39835{,}32}{7} - 73{,}57^2} = 16{,}68$$

Следовательно,

$$r_{XY} = \frac{\overline{yx} - \bar{y} \times \bar{x}}{\sigma_x \sigma_y} = \frac{2336{,}00 - 73{,}57 \times 29{,}23}{18{,}47 \times 16{,}68} = 0{,}602$$

Связь между признаками заметная и однонаправленная.

Определим коэффициент детерминации:

$$d = r^2 \times 100\% = 0{,}602^2 \times 100\% = 36{,}24\ \%$$

Вариация размера активов на 36,24 % зависит от величины заемных средств сельскохозяйственных организаций, на 63,76 % − от других факторов.

2. Определим коэффициент эластичности:

$$Э_{x_i} = a_i \times \frac{\bar{x}_i}{\bar{y}} = 0{,}544 \times \frac{29{,}23}{73{,}57} = 0{,}216.$$

Если размер заемных средств увеличить на 1 %, то размер активов вырастет на 0,216 %.

Найдем значение β-коэффициента:

$$\beta_i = a_i \times \frac{\sigma_{X_i}}{\sigma_Y} = 0{,}544 \times \frac{18{,}47}{16{,}68} = 0{,}602.$$

Если размер заемных средств увеличится на свое среднее квадратическое отклонение, то активы организаций вырастут на 0,602 своего среднего квадратического отклонения.

Задача 89

В таблице 93 представлены размеры товарной продукции и фондообеспеченность в расчете на 1 га сельскохозяйственных угодий по 10 предприятиям области.

1. Используя метод приведения параллельных данных, установите направление и характер связи между товарной продукцией, полученной на 1 га, и фондообеспеченностью на 1 га.

2. Вычислите линейный коэффициент корреляции и коэффициент детерминации, β-коэффициент, коэффициент эластичности.

Сделайте выводы.

Таблица 93 − Показатели развития предприятий области

Номер предприятия	Товарная продукция на 1 га, тыс. руб.	Фондообеспеченость на 1 га, тыс. руб.
1	0,98	3,32
2	1,97	3,28
3	2,34	6,26
4	4,30	8,05
5	2,12	5,74
6	2,90	3,27
7	3,11	6,44
8	1,81	3,14
9	3,73	6,93
10	2,76	2,88

Задача 90

В таблице 94 представлены данные по объему внесенных минеральных удобрений на 1 га и урожайности зерновых культур по 10 предприятиям области.

Таблица 94 − Объем внесенных минеральных удобрений на 1 га и урожайность зерновых культур по предприятиям области

Номер предприятия	Внесено минеральных удобрений на 1 га, кг	Урожайность зерновых, ц с 1 га
1	60	28
2	30	16
3	42	22
4	44	24
5	36	22
6	35	18
7	40	22
8	34	18
9	32	24
10	38	22

1. Определите вид корреляционной зависимости, постройте уравнение регрессии, рассчитайте параметры уравнения, вычислите тесноту связи.

2. Вычислите линейный коэффициент корреляции и коэффициент детерминации, β-коэффициент, коэффициент эластичности.

Объясните полученные статистические характеристики.

Задача 91

По данным, приведенным в таблице 95, исследуйте зависимость между яловостью коров и среднегодовым надоем молока на одну корову:

1) найдите параметры уравнения регрессии;

2) рассчитайте коэффициенты корреляции и детерминации;

3) вычислите индекс корреляции;

4) проверьте гипотезу о значимости (существенности) уравнения регрессии и показателей тесноты корреляционной связи.

Таблица 95 – Исходные данные по уровню яловости коров и среднегодовому надою молока на одну корову

Яловость коров, %	22	28	24	15	16	12	11	12	18	19
Среднегодовой надой молока на 1 корову, ц	17	25	25	37	33	40	18	15	23	21

Задача 92

В таблице 96 представлены данные о посевной площади зерновых культур, валовом сборе и внесении минеральных удобрений на 1 га посевной площади.

1. Используя метод приведения параллельных данных, установите направление и характер связи между факторами.

2. Постройте множественное уравнение регрессии, предварительно сформулировав и обосновав выбор результативного и факторных признаков, рассчитайте параметры уравнения.

3. Вычислите множественный и частный коэффициенты корреляции.

Проанализируйте полученные результаты.

Таблица 96 − Посевная площадь зерновых культур, валовой сбор и внесение минеральных удобрений на 1 га посевной площади по предприятиям области

Номер района	Посевная площадь зерновых культур, га	Валовой сбор, тыс. т	Внесено минеральных удобрений на 1 га посевной площади, кг
1	35776	91,1	26
2	22374	44,3	22
3	32871	100	26
4	12222	28,9	18
5	8922	17,8	24
6	5699	17,8	24
7	21349	47,2	20
8	23298	58,3	26
9	46084	109,7	28
10	19743	30,1	26
11	49985	106,6	26
12	15282	27,9	20
13	10812	21,2	22
14	22430	60,3	26
15	19339	33,1	18

Задача 93

По данным, приведенным в таблице 97, исследуйте зависимость между расходом кормов на одну корову и среднегодовым надоем молока на одну корову.

Таблица 97 – Исходные данные по расходу кормов на одну корову и среднегодовому надою молока на одну корову

Расход кормов на 1 корову, ц к. ед.	48	40	34	42	47	39	55	54	41	35
Среднегодовой надой молока на 1 корову, ц	35	25	22	31	31	32	50	42	28	32

Определите параметры уравнения регрессии, коэффициенты корреляции и детерминации, индекс корреляции. Проверьте гипотезу о значимости (существенности) уравнения регрессии и показателей тесноты корреляционной связи.

Задача 94

Результаты распределения основных категорий потенциальных мигрантов по уровню образования представлены в таблице 98.

Таблица 98 − Распределение основных категорий потенциальных мигрантов по уровню образования

Образование	Основные категории потенциальных мигрантов				Итого
	руководители	специалисты	служащие	рабочие	
Высшее	55	48	12	7	122
Неполное высшее	5	3	3	5	16
Среднее специальное	36	44	51	39	170
Среднее общее	4	4	33	39	80
Неполное среднее	0	1	1	10	12
Итого	100	100	100	100	400

Рассчитайте все возможные модификации коэффициентов взаимной сопряженности. Сформулируйте выводы, вытекающие из анализа полученных коэффициентов.

Задача 95

Используя следующие данные:

1. $\overline{xy} = 100,\ \bar{x} = 10,\ \bar{y} = 8,\ \overline{x^2} = 136,\ \overline{y^2} = 100,\ a_0 = 4{,}8$, постройте линейное уравнение регрессии, вычислите линейный коэффициент корреляции;

2. $\bar{x} = 20,\ \bar{y} = 10,\ Э_x = 0{,}8$, определите параметры линейного уравнения (a_0 и a_1) регрессии;

3. $a_0 = 3{,}5,\ r_{xy} = 0{,}85,\ \sigma_y^2 = 36,\ \sigma_x^2 = 49$, постройте линейное уравнение регрессии.

Задача 96

В ходе проведенного обследования оценки уровня жизни работающих на предприятиях различной формы собственности было опрошено 100 респондентов.

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

Результаты опроса представлены в таблице 99.

Таблица 99 − Результаты обследования уровня жизни работающих на предприятиях различной формы собственности

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

9 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНАМИКИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Для характеристики и анализа социально-экономических явлений за некоторый период применяют показатели и методы, характеризующие эти процессы во времени (динамике).

Ряд динамики – это ряд последовательно расположенных статистических показателей (в хронологическом порядке), изменение которых показывает определенную тенденцию развития изучаемого явления.

Элементы ряда динамики:

− показатель времени (t);

− уровень развития изучаемого явления (y).

Ряды динамики классифицируются по нижеследующим признакам (таблица 100).

Таблица 100 − Классификация рядов динамики

Классификационный признак	Вид рядов динамики
Способ выражения уровней ряда	Ряд абсолютных величин Ряд средних величин Ряд относительных величин
Способ представления хронологии	Моментный ряд Интервальный ряд
Расстояние между периодами или датами	Равноотстоящие уровни во времени Неравноотстоящие уровни во времени
Наличие основной тенденции в ряду	Стационарный ряд Нестационарный ряд
Число показателей	Изолированный ряд Многомерный ряд

Примером интервального ряда динамики абсолютных величин с равноотстоящими уровнями во времени служат данные о производстве хлопчатобумажных тканей в РФ (таблица 101).

Особенностью интервального ряда динамики является то, что каждый его уровень складывается из данных за более короткие интервалы (субпериоды) времени. Свойство суммирования уровней за последовательные интервалы времени позволяет получать ряды динамики более укрупненных периодов.

Уровни интервального ряда динамики могут быть суммированы, так как не содержат повторного счета.

Таблица 101 − Динамика производства хлопчатобумажных тканей в РФ, млрд. м*^{2}

2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
2264	2329	2149	2225	2222	2108	1915	1427	1484

* Россия в цифрах. 2011: Крат. стат. сб. / Росстат. – М., 2011. – С. 236

Примером моментного ряда абсолютных величин с равноотстоящими уровнями во времени можно назвать ряд динамики, показывающий число предприятий и организаций в РФ на 1 января (таблица 102):

Таблица 102 − Динамика численности предприятий и организаций в РФ (на 1 января), тыс. единиц

2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.	2011 г.
4417,1	4767,3	4506,6	4674,9	4771,9	4907,8	4823,3

Уровни моментного ряда – обобщенные итоги учета числа предприятий и организаций в РФ по состоянию на определенную дату (1 января каждого года). Отдельные уровни моментного ряда содержат элементы повторного счета, и это делает бессмысленным суммирование уровней моментных рядов динамики.

Примером интервального ряда динамики средних величин с неравноотстоящими уровнями во времени может служить ряд динамики среднегодовой численности занятых в сельском хозяйстве, охоте и лесном хозяйстве в РФ (таблица 103).

Таблица 103 − Динамика среднегодовой численности занятых в сельском хозяйстве, охоте и лесном хозяйстве в РФ, тыс. чел.

1995 г.	2005 г.	2007 г.	2009 г.	2010 г.
6700	7381	6925	6580	6465

Примером интервального ряда динамики относительных величин с равноотстоящими уровнями во времени служит ряд динамики, характеризующий удельный вес численности занятых в экономике и безработных в общей численности экономически активного населения России (%):

Таблица 104 − Удельный вес численности занятых в экономике и безработных в общей численности экономически активного населения России, проц.*

Год	Численность экономически активного населения – всего	в том числе	
		занятые в экономике	безработные
2005	100	92,8	7,2
2006	100	92,8	7,2
2007	100	93,9	6,1
2008	100	93,7	6,3
2009	100	91,6	8,4
2010	100	92,5	7,5

* Россия в цифрах. 2011: Крат. стат. сб. / Росстат. – М., 2011. – С. 94

По содержанию показатели динамических рядов различают на состоящие из частных показателей и агрегированных показателей. *Частные показатели* характеризуют явления изолированно, односторонне (например, динамика показателей среднесуточного объема потребленной воды). *Агрегированные показатели* – производные из частных и характеризуют изучаемое явление комплексно.

С помощью рядов динамики изучение закономерностей развития социально-экономических явлений осуществляется в следующих основных направлениях:

1) характеристика уровней развития изучаемых явлений во времени;

2) измерение динамики изучаемых явлений посредством системы статистических показателей;

3) выявление и количественная оценка основной тенденции развития (тренда);

4) изучение периодических колебаний;

5) экстраполяция и прогнозирование.

Основные требования к построению ряда динамики:

1) сопоставимость по территории − несопоставимость возникает в результате изменения границ стран, регионов и т.д.;

2) сопоставимость по кругу охватываемых объектов – объекты с одинаковой полнотой должны быть охвачены обследованием;

3) сопоставимость по времени регистрации – обследование должно быть проведено с учетом сезонности явления;

4) сопоставимость по стоимостным показателям – различие может возникнуть вследствие изменения цен;

5) сопоставимость по методологии расчета – методология расчета и обследования должна быть единой;

6) сопоставимость по единицам измерения – несопоставимость возникает в случае, если показатель может быть представлен в разных единицах измерения;

7) достоверность – несопоставимость возникают вследствие неодинаковой репрезентативности выборки по различным периодам.

Когда имеются уровни ряда, исчисленные по разной методологии или в разных границах, такой ряд динамики приводят к сопоставимому виду с помощью следующих методов:

– смыкание рядов динамики;
– приведение рядов к одному основанию.

Под смыканием рядов динамики понимают объединение в один ряд (более длинный) двух или нескольких рядов, уровни которых исчислены по разным методологиям или в разных границах. Для осуществления такого смыкания необходимо, чтобы данные для одного из периодов (переходного) были исчислены по двум методологиям. Смыкание рядов осуществляется в абсолютных величинах и относительных величинах.

Пример. В таблице 105 представлены данные, характеризующие общий объем продукции сельского хозяйства в одном из регионов (данные условные).

Для приведения ряда динамики к сопоставимому виду для 2006 г. определим коэффициент соотношения уровней двух рядов:

$$\frac{42{,}8}{37{,}1}=1{,}15$$

Умножая на этот коэффициент уровни первого ряда, получаем их сопоставимость с уровнями второго ряда:

2004 г.: $43{,}6\times1{,}15=50{,}1$ млн. руб.

2005 г.: $40{,}8\times1{,}15=46{,}9$ млн. руб.

2006 г.: $45{,}1\times1{,}15=51{,}7$ млн. руб.

Таблица 105 − Объем продукции сельского хозяйства в регионе (данные условные), млн. руб.

Уровень продукции сельского хозяйства	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
В старых границах региона	43,6	40,8	45,1	37,1	-	-	-
В новых границах региона	-	-	-	42,8	44,3	45,5	47,2

В таблице 106 представлен сопоставимый ряд динамики общего объема продукции сельского хозяйства в одном из регионов (в новых границах).

Таблица 106 − Динамика общего объема продукции сельского хозяйства в регионе (в новых границах), млн. руб.

2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
50,1	46,9	51,7	42,8	44,3	45,5	47,2

Прием перехода от абсолютных показателей к относительным именуют в статистике приведением рядов к одному основанию. В таких случаях уровни всех рассматриваемых рядов приводятся в процентах (или коэффициентах) к уровню одного и того же периода или момента времени (либо иной базе сравнения).

Применив этот способ для нашего примера, получим ряд динамики, характеризующий общий объем продукции региона (таблица 107).

Таблица 107 − Общий объем продукции в новых границах региона, (проц. к 2007 г.)

2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
117,5	110,0	121,6	100,0	103,5	106,3	110,3

При изучении рядов динамики возникает необходимость получения сравнительных характеристик направления и интенсивности роста одновременно развивающихся во времени явлений. Это достигается путем *приведения рядов динамики к общему (единому) основанию*. Ряды, приведенные к одному основанию, легче интерпретировать и анализировать.

Пример. Сравним динамику поголовья коров и производства молока в хозяйствах всех категорий в Пензенской области (таблица 108).

Таблица 108 − Динамика поголовья коров и производства молока в хозяйствах всех категорий в Пензенской области

Год	Поголовье коров, тыс. гол.	Производство молока, тыс. т
2001	85,6	161,3
2002	78,1	167,1
2003	73,1	161,1
2004	70,0	158,1
2005	60,0	160,6
2006	60,4	181,6
2007	60,5	194,8
2008	58,2	197,5
2009	50,4	179,2
2010	50,2	161,4

В таблице 109 представлены результаты приведения абсолютных уровней рядов динамики к общему основанию. За постоянную базу сравнения принят 2001 г.

Таблица 109 − Приведение к общему основанию уровней рядов динамики поголовья коров и производства молока в хозяйствах всех категорий в Пензенской области, проц.

Год	Поголовье коров	Производство молока
2001	100	100
2002	92,2	103,6
2003	85,4	99,9
2004	81,8	98,0
2005	70,1	99,6
2006	70,6	112,6
2007	70,7	120,8
2008	68,0	122,4
2009	58,9	111,1
2010	58,6	100,1

Из этих данных видно, что производство молока в хозяйствах всех категорий в Пензенской области возрастает по сравнению со стабильным значительным сокращением поголовья коров.

Для количественной оценки динамики социально-экономических явлений применяются статистические показатели: абсолютный прирост, темп роста, темп прироста, абсолютное значение 1 % прироста, темп наращивания и др.

В основе расчета показателей рядов динамики лежит сравнение его уровней. В зависимости от применяемого способа сопоставления показатели динамики могут вычисляться на постоянной и переменной базах сравнения.

Для расчета показателей динамики на постоянной базе каждый уровень ряда сравнивается с одним и тем же базисным уровнем. Исчисляемые при этом показатели называются *базисными*. Для расчета показателей динамики на переменной базе каждый последующий уровень ряда сравнивается с предыдущим. Вычисленные таким образом показатели динамики называются *цепными*.

Показатели, используемые для анализа уровней ряда динамики, представлены в таблице 110.

Между базисными и цепными абсолютными приростами имеется связь (сумма цепных абсолютных приростов равна базисному абсолютному приросту последнего периода ряда динамики):

$$\Delta^{б}_{y_n} = \sum \Delta^{ц}_{y}$$

Таблица 110 − Аналитические показатели изменения уровней ряда динамики

Показатель	Методика расчета		Пояснение
	цепной	базисный	
Абсолютный прирост (Δy)	$\Delta^{ц}_{y} = y_i - y_{i-1}$	$\Delta^{б}_{y} = y_i - y_1$	Выражает абсолютную скорость роста (+), снижения (−) уровня ряда динамики. $\Delta_y > 0$ – рост $\Delta_y < 0$ – спад $\Delta_y = 0$ – стабильность
Коэффициент роста ($K_р$)	$K^{ц}_{р} = \frac{y_i}{y_{i-1}}$	$K^{б}_{р} = \frac{y_i}{y_1}$	Выражает интенсивность изменения уровней ряда динамики
Темп роста ($T_р$), %	$T^{ц}_{р} = \frac{y_i}{y_{i-1}} \times 100$	$T^{б}_{р} = \frac{y_i}{y_1} \times 100$	Выражает интенсивность изменения уровней ряда динамики
Темп прироста ($T_{пр}$), %	$T^{ц}_{пр} = \frac{y_i - y_{i-1}}{y_{i-1}} \times 100$ или $T^{ц}_{пр} = T_p - 100$	$T^{б}_{пр} = \frac{y_i - y_1}{y_1} \times 100$ или $T^{б}_{пр} = T_p - 100$	Выражает изменение величины абсолютного прироста уровней ряда динамики в относительной величине
Абсолютное значение 1 % прироста (*1 % Δy*)	$1\%\Delta y = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} 100\,\%} = 0{,}01 y_{i-1}$		Расчет данного показателя имеет экономический смысл только на цепной основе. Показывает, сколько абсолютных величин приходится на 1 % прироста (уменьшения)

Если показатели уровня ряда принимают как положительные, так и отрицательные значения (например, прибыль и убыток в организации за ряд лет), то темпы роста и прироста не рассчитываются и не имеют экономической интерпретации.

Между цепными и базисными темпами роста существуют связи:

1) произведение всех цепных коэффициентов роста равно базисному;

2) результат деления двух базисных коэффициентов равен цепному (промежуточному).

Такие взаимосвязи проявляются только в случае, если темпы роста (цепные и базисные) выражены в коэффициентах.

Отношение темпов роста (или прироста) по двум динамическим рядам (в одинаковом промежутке времени) называют коэффициентом опережения (замедления):

$$K_{о} = \frac{T_2}{T_1}\text{, где } T_2 > T_1 \text{ или } K_{о} = \frac{T_1}{T_2}\text{, где } T_1 > T_2$$

Данную формулу удобнее применять для ряда представляющего постоянное повышение. При отсутствии четкой тенденции к росту рекомендуется использовать средние показатели рядов динамики (средние темпы роста).

Для получения обобщающих показателей динамики социально-экономических явлений определяются средние величины: средний уровень ряда динамики, средний абсолютный прирост, средний темп роста, средний темп прироста и др. (таблица 111).

Пример. Требуется провести анализ динамики производства комбикорма за 2004-2010 гг. в РФ. Для удобства и наглядности исходные и рассчитанные показатели представлены в таблице 112.

Таблица 111 − Средние обобщающие показатели ряда динамики

Показатель	Методика расчета
Средний уровень ряда ($\bar{y}$)	$\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$ – для интервального ряда с равноотстоящими уровнями; $\bar{y} = \dfrac{\sum_{i=1}^{n} y_i t_i}{\sum t_i}$ – для интервального ряда с неравноотстоящими уровнями; $\bar{y} = \dfrac{\frac{1}{2} y_1 + y_2 + y_3 + \ldots + \frac{1}{2} y_n}{n-1}$ – для моментного ряда с равноотстоящими уровнями; $\bar{y} = \dfrac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \ldots + (y_{n-1} + y_n)t_{n-1}}{2(t_1 + t_2 + t_3 + \ldots + t_{n-1})} = \dfrac{\sum (y_i + y_{i+1})t_1}{2\sum t_i}$ – для моментного ряда динамики с неравноотстоящими уровнями.
Средний абсолютный прирост ($\overline{\Delta y}$)	$\overline{\Delta y} = \dfrac{\sum_{i=1}^{n-1} \Delta_{i/i-1}}{n-1}$ или $\overline{\Delta y} = \dfrac{y_n - y_1}{n-1}$
Средний темп роста ($\bar{T}_{\text{р}}$)	$\bar{T}_{\text{р}} = \sqrt[n-1]{\dfrac{y_n}{y_1}} \times 100$ или $\bar{T}_{\text{р}} = \sqrt[m]{K_{2/1} \times K_{3/2} \times \ldots \times K_{n/n-1}} = \sqrt[m]{\prod K_{p\, i/i-1}}$ – для ряда с равноотстоящими уровнями; $\bar{T}_{\text{р}} = \sqrt[\sum t]{K_{2/1}^{t_1} \times K_{3/2}^{t_2} \times \ldots \times K_{n/n-1}^{t_n}}$ – для ряда с неравноотстоящими уровнями.
Средний темп прироста ($\bar{T}_{\text{пр}}$)	$\bar{T}_{\text{пр}} = \bar{T}_{\text{р}} - 100$

Таблица 112 − Динамика производства комбикорма в РФ расчет аналитических показателей динамики

Год	Производство комбикорма, млн. т	Абсолютные приросты (снижение), тыс.т		Темпы роста, %		Темпы прироста, %		Абсолютное значение 1 % прироста, тыс. т
		цепной	базисный	цепной	базисный	цепной	базисный	
А	Б	1	2	3	4	5	6	7
2004	9,1	-	-	-	100,0	-	-	-
2005	10,0	+0,9	+0,9	109,9	109,9	+9,9	+9,9	0,091
2006	11,4	+1,4	+2,3	114,0	125,3	+14,0	+25,3	0,100
2007	12,5	+1,1	+3,4	109,6	137,4	+9,6	+37,4	0,114
2008	13,7	+1,2	+4,6	109,6	150,5	+9,6	+50,5	0,125
2009	14,7	+1,0	+5,69	107,3	161,5	+7,3	+61,5	0,137
2010	16,0	+1,3	+6,9	108,8	175,8	+8,8	+75,8	0,147
Итого	87,4	6,9	-	-	-	-	-	-

Абсолютное увеличение производства комбикорма за 2010 г. по сравнению с 2009 г. составило: 16 − 14,7=1,3 млн. т (таблица 112, графа 1), по сравнению с базисным 2004 г. производство в 2010 г. выросло на 6,9 млн. т.

Так, для 2010 г. темпы роста по сравнению с 2004 г. составил $\left(\frac{16{,}0}{9{,}1}\right)\times 100\% = 175{,}8\%$ (таблица 112, графа 4).

Расчет темпа прироста (таблица 112, графа 6) показал, на сколько процентов производство комбикорма в 2010 г. выросло по сравнению с 2004 г.:

$$\left(\frac{6{,}9}{9{,}1}\right)\times 100\ \% = 75{,}8\ \% \text{ или } 175{,}8\,\% - 100\% = 75{,}8\,\%$$

Для 2010 г. абсолютное значение 1 % прироста (таблица 112, графа 7) равно:

$$0{,}01 \times 14{,}7 = 0{,}147$$

Среднее производство комбикорма за 7 лет составило:

$$\bar{y} = \frac{87,4}{7} = 12,5 \text{ тыс. т}$$

Среднегодовой темп роста производства комбикорма за период рассчитаем двумя способами:

$$\bar{T}_p = \sqrt[7-1]{\frac{16,0}{9,1}} = 1,098 \text{ или } 109,8\ \%$$

$$\bar{T}_p = \sqrt[6]{1,099 \times 1,14 \times 1,096 \times 1,096 \times 1,073 \times 1,088} =$$
$$= \sqrt[6]{1,7569} = 1,098 \text{ или } 109,8\ \%$$

Средний темп прироста:

$$\bar{T}_{\text{пр}} = \bar{T}_{\text{р}} - 100 = 109,8\ \% - 100\ \% = 9,8\ \%$$

Важным направлением в исследовании закономерностей динамики социально-экономических процессов является изучение общей тенденции развития (тренда).

Тренд – основная (достаточно устойчивая) тенденция развития явления в ряду динамики.

Это можно осуществить, применяя специальные методы анализа рядов динамики. Конкретное их использование зависит от характера исходной информации и предопределяется задачами анализа. Различные результаты действия постоянных, периодических и разовых причин и факторов на уровне развития социально-экономических явлений во времени обусловливают необходимость изучения *основных компонентов рядов динамики*: *тренда* (долговременное движение), *сезонных колебаний* (кратковременное систематическое движение), *случайных отклонений* (несистематическое случайное движение).

На практике наиболее распространенными методами статистического изучения тренда являются:

– метод укрупнения интервалов;
– метод скользящей средней;

– аналитическое выравнивание.

Данные методы обработки рядов называются сглаживанием или выравниванием рядов динамики.

Метод укрупнения интервалов применяется для выявления тренда в рядах динамики колеблющихся уровней, затушевывающих основную тенденцию развития. Метод основан на укрупнении периодов, к которым относятся уровни ряда динамики. Например, преобразование месячных периодов в квартальные, квартальных в годовые и т.д. метод простого суммирования не используется в моментных рядах, а также если уровни ряда выражены относительной или средней величиной.

Метод укрупнения интервалов эффективен, если первоначальные уровни ряда относятся к коротким промежуткам времени. В ряду с укрупненными интервалами времени закономерность изменения уровней будет более наглядной.

Пример. Имеются данные о динамике объема оборота розничной торговли в регионе по месяцам (таблица 113).

Таблица 113 − Динамика объема оборота розничной торговли

в регионе по месяцам, млрд. руб.

в фактических ценах (данные условные)

Месяц	Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
Объем оборота розничной торговли	6,7	6,4	7,3	7,4	7,7	7,8	8,1	8,4	8,2	8,5	8,0	9,7

Укрупним интервалы до квартала и рассчитаем суммарный и среднемесячный товарообороты по кварталам. Новые данные представлены в таблице 114.

Новые данные более четко выражают закономерность изменения оборота розничной торговли – увеличение из квартала в квартал.

Таблица 114 − Динамика объема оборота

розничной торговли в регионе
по кварталам (данные условные)

Квартал	Объем оборота розничной торговли, млрд. руб.	
	общий	среднемесячный
I	20,4	6,8
II	22,9	7,6
III	24,7	8,2
IV	26,2	8,7

Метод скользящей средней[1]. В основу этого метода положено определение по исходным данным теоретических уровней, в которых случайные колебания погашаются, а основная тенденция развития выражается в виде некоторой плавной линии.

Скользящая средняя – подвижная динамическая средняя, которая рассчитывается по ряду при последовательном передвижении на один интервал, то есть сначала вычисляют средний уровень из определенного числа первых по порядку уровней ряда, затем – средний уровень из такого же числа членов, начиная со второго. Таким образом, средняя как бы скользит по ряду динамики от его начала к концу, каждый раз отбрасывая один уровень в начале и добавляя один следующий.

Нахождение скользящей средней по четному числу членов осложняется тем, что средняя может быть отнесена только к середине между двумя датами. Чтобы ликвидировать этот сдвиг, применяется центрирование, т.е. нахождение средней из средних для отнесения полученного уровня к определенной дате. При центрировании необходимо также находить скользящие суммы, скользящие средние по этим суммам и средние из средних.

Недостатком метода является укорачивание сглаженного ряда по сравнению с фактическим, что ведет к потере информации.

Пример. На основе данных таблицы о производстве яиц в хозяйствах всех категорий за 1999-2010 гг. в Пензенской области требуется провести выравнивание ряда методом скользящей средней.

[1] Данный метод сглаживает (устраняет) лишь случайные колебания. Если ряд содержит сезонную волну, она сохранится и после сглаживания.

Таблица 115 − Динамика производства яиц в хозяйствах всех категорий в Пензенской области и расчет скользящих средних

Год	Производство яиц, млн. шт.	Трехчленные скользящие суммы	Трехчленные скользящие средние	Четырехчленные скользящие суммы	Четырехчленные скользящие средние (нецентрированные)	Четырехчленные скользящие средние (центрированные)
А	Б	1	2	3	4	5
1999	340,0	-	-	-	-	-
2000	330,4	-	311,2	-	300,0	-
2001	263,1	933,5	286,7	-	283,8	291,9
2002	266,6	866,1	268,3	1200,1	270,4	277,1
2003	275,2	804,9	272,8	1135,3	261,9	266,2
2004	276,7	818,5	260,4	1081,6	259,1	260,5
2005	229,2	781,1	253,7	1047,7	257,7	258,4
2006	255,1	761,0	251,4	1036,2	257,4	257,6
2007	269,9	754,2	266,8	1030,9	280,9	269,2
2008	275,3	800,3	289,5	1029,5	300,2	290,6
2009	323,4	868,6	310,3	1123,7	-	-
2010	332,3	931,0	-	1200,9	-	-

$$\bar{y}_1 = \frac{340,+330,4+263,1}{3} = \frac{933,5}{3} = 311,2$$

$$\bar{y}_2 = \frac{330,4+263,1+266,6}{3} = \frac{860,1}{3} = 286,7 \text{ и т.д.}$$

Применение в анализе рядов динамики методов укрупнения интервалов и скользящей средней позволяет выявить тренд для его описания, но получать обобщенную статистическую оценку тренда посредством этих методов невозможно. Решение этой задачи достигается *методом аналитического выравнивания*.

Аналитическое выравнивание. Метод является наиболее совершенным методом обработки рядов динамики в целях устранения случайных колебаний и выявления тренда. Основным содержанием метода аналитического выравнивания в рядах динамики является то, что основная тенденция развития y_t рассчитывается как функция времени: $\bar{y}_t = f(t)$.

Последовательность аналитического выравнивания:

– определение на основе фактических данных вида (формы) гипотетической функции $\bar{y}_t = f(t)$, способной наиболее адекватно отразить тенденцию развития исследуемого показателя;

– нахождение по эмпирическим данным параметров;

– расчет по найденному уравнению теоретических (выравненных) уровней.

В аналитическом выравнивании наиболее часто используются нижеследующие простейшие функции (таблица 116).

В статистической практике параметры функций невысокой степени имеют конкретную интерпретацию характеристик динамического ряда.

Выбор той или иной функции для выравнивания ряда динамики осуществляется, как правило, на основании графического изображения эмпирических данных, дополняемого содержательным анализом особенностей развития исследуемого показателя (явления) и специфики разных функций, их возможности отразить те или иные нюансы развития. Определенную вспомогательную роль при выборе аналитической функции играют механические приемы сглаживания.

Условия использования уравнений:

1. Выравнивание по прямой линии (*линейной функции*) эффективно для рядов, уровни которых изменяются в арифметической прогрессии, т.е. абсолютные приросты более или менее постоянны;

2. Если ускорение более или менее постоянно, то такое развитие хорошо описывается *параболой 2-го порядка*;

Таблица 116 − Функции, используемые в аналитическом выравнивании

Функция	Формула	Условные обозначения параметров
Линейная	$\bar{y}_t = a_0 + a_1 t$	$\hat{y}_t$ – теоретические (выравненные) уровни; t – условное обозначение времени (1, 2, 3,…); a_0 − значение выравненного значения для центрального года в динамическом ряду, принятого за начало отчета при t = 0; a_1 − коэффициент регрессии, определяющий направление развития. Если $a_1>0$, то уровни ряда динамики равномерно возрастают, а при $a_1<0$ происходит их равномерное снижение.
Показательная	$\bar{y}_t = a_0 a_1^t$	a_0 − значение выравненного значения для центрального года в динамическом ряду, принятого за начало отчета при t = 0; a_1, − темп роста (снижения) изучаемого явления в единицу времени, т.е. интенсивность развития.
Парабола 2-го порядка	$\bar{y}_t = a_0 + a_1 t + a_2 t^2$	a_0 − значение выравненного значения для центрального года в динамическом ряду, принятого за начало отчета при t = 0; a_1 − коэффициент регрессии, определяющий направление развития. Может быть как со знаком плюс, так и со знаком минус; a_2 – характеризует постоянное изменение интенсивности развития (в единицу времени). При $a_2>0$ происходит ускорение развития, а при $a_2<0$ идет процесс замедления роста.
Гиперболическая	$\bar{y}_t = a_0 + \frac{a_1}{t}$	a_0 − значение выравненного значения для центрального года в динамическом ряду, принятого за начало отчета при t = 0; a_1, − изменение ускорения.
Степенная	$\bar{y}_t = a_0 t^{a_1}$	a_0 − значение выравненного значения для центрального года в динамическом ряду, принятого за начало отчета при t = 0; a_1, − изменение ускорения.

3. Если при последовательном расположении t значения уровней меняются в геометрической прогрессии, т.е. цепные коэффициенты роста примерно постоянны, то такое развитие можно отразить *показательной функцией*;

4. Если обнаружено замедленное снижение уровней ряда, которые по логике не могут снизиться до нуля, для описания характера тренда выбирают *гиперболическую функцию* или *степенную* и др.

Параметры искомых уравнений при аналитическом выравнивании определяются решением системы нормальных уравнений, полученных методом наименьших квадратов (таблица 117).

Таблица 117 − Расчет параметров функций

Функция	Формула	Расчет параметров
1	2	3
Линейная[1]	$\bar{y}_t = a_0 + a_1 t$	$a_0 = \dfrac{\sum y \sum t^2 - \sum t \sum yt}{n\sum t^2 - \left(\sum t\right)^2}$ $a_1 = \dfrac{n\sum yt - \sum t \sum y}{n\sum t^2 - \left(\sum t\right)^2}$
Показательная[2]	$\bar{y}_t = a_0 a_1^t$	$\lg a_0 = \dfrac{\sum \lg y}{n}$ $\lg a_1 = \dfrac{\sum t \lg y}{\sum t^2}$

[1] Расчет параметров a_0 и a_1 упрощается, если отсчет времени ведется от середины ряда. При нечетном числе уровней серединная точка (год, месяц) принимается за нуль. Тогда предшествующие периоды обозначаются соответственно −1, −2, −3 и т.д., а следующие за средним (центральным) − соответственно 1, 2, 3 и т.д.

При четном числе уровней два серединных момента (периода) времени обозначаются −1 и +1, а все последующие и предыдущие, соответственно, через два интервала: ±3, ±5, ±7 и т.д.

При таком порядке отсчета времени $\sum t = 0$, поэтому параметры рассчитываются по следующим формулам:

$$a_0 = \frac{\sum y}{n} \text{ и } a_1 = \frac{\sum yt}{\sum t^2}.$$

[2] Для расчета выравненных уровней удобнее пользоваться формулой логарифмов: $\lg \hat{y}_t = \lg a_0 + t \lg a_1$.

Окончание таблицы 117

1	2	3
Парабола 2-го порядка	$\bar{y}_t = a_0 + a_1 t + a_2 t^2$	$a_0 = \dfrac{\sum t^4 \sum y - \sum t^2 \sum t^2 y}{n\sum t^4 - \sum t^2 \sum t^2}$ $a_1 = \dfrac{\sum ty}{\sum t^2}$ $a_2 = \dfrac{n\sum t^2 y - \sum t^2 \sum y}{n\sum t^4 - \sum t^2 \sum t^2}$
Гиперболиче-ская	$\bar{y}_t = a_0 + \dfrac{a_1}{t}$	$\begin{cases} na_0 + a_1 \sum \dfrac{1}{t} = \sum y \\ a_0 \sum \dfrac{1}{t} + a_1 \sum \left(\dfrac{1}{t}\right)^2 = \sum \dfrac{y}{t} \end{cases}$

Степень соответствия выравненных значений фактическим характеризуется:

остаточным средним квадратическим отклонением

$$\sigma_{\text{ост}} = \sqrt{\frac{\sum \left(y_i - \bar{y}_{t_i}\right)^2}{n}},$$

где y_i – фактическое значение уровней;

$\bar{y}_{t_i}$ – выравненное значение уровней;

n – число уровней динамического ряда;

средней ошибкой аппроксимации

$$\varepsilon_t = \frac{1}{n} \sum_{t=1}^{n} \frac{\left|y_t - \bar{y}_t\right|}{y_t} \times 100\,\%.$$

Проводя сравнительную оценку моделей тренда можно использовать одну из перечисленных характеристик. Наиболее адекватной считается функция, у которой остаточное среднее квадратическое отклонение или средняя ошибка аппроксимации минимальны.

Пример. В таблице 118 приведены исходные и расчетные данные о динамике валового сбора овощей (включая овощи закрытого грунта) в Пензенской области за 2004-2010 гг.

Для выравнивания ряда динамики по прямой используем уравнение $\bar{y}_t = a_0 + a_1 t$.

Расчет необходимых значений дан в таблице. По итоговым данным определяем параметры уравнения:

$$a_0 = \frac{\sum y}{n} = \frac{965{,}7}{7} = 137{,}96$$

$$a_1 = \frac{\sum ty}{\sum t^2} = \frac{-73}{28} = -2{,}61$$

В результате получаем следующее уравнение основной тенденции производства овощей в Пензенской области за 2004−2010 гг.:

$$\bar{y}_t = 137{,}96 - 2{,}61t$$

Параметр a_1 трендовой модели показывает, что валовой сбор овощей (включая овощи закрытого грунта) в Пензенской области снижался в среднем на 2,61 тыс. т.

Подставляя в уравнение принятые обозначения t, вычислим выравненные уровни ряда динамики:

2004 г. – $\bar{y}_1 = 137{,}96 - 2{,}61 \times (-3) = 145{,}79$

2005 г. – $\bar{y}_2 = 137{,}96 - 2{,}61 \times (-2) = 143{,}18$

и т.д. (см. значения в графе 6 таблицы 118).

Для выравнивания ряда динамики по параболе 2-го порядка используем уравнение $\bar{y}_t = a_0 + a_1 t + a_2 t^2$.

Расчет необходимых значений дан в таблице 118. По итоговым данным определяем параметры уравнения:

Таблица 118 − Расчет теоретических уровней линейного тренда [6]

Год	Валовой сбор овощей (включая овощи закрытого грунта), тыс.т (y)	Условное обозначение времени (t)	t^2	t^4	ty	t^2y	Уравнение прямой			Уравнение параболы 2-го порядка		
							$\bar{y}_t$	$y-\bar{y}_t$	$(y-\bar{y}_t)^2$	$\bar{y}_t$	$y-\bar{y}_t$	$(y-\bar{y}_t)^2$
А	Б	1	2	3	4	5	6	7	8	9	10	11
2004	146,4	-3	9	81	-439,2	1317,6	145,79	0,61	0,3721	144,77	1,63	2,6569
2005	148,0	-2	4	16	-296,0	592	143,18	4,82	23,2324	143,16	4,84	23,4256
2006	133,1	-1	1	1	-133,1	133,1	140,57	-7,47	55,8009	141,15	-8,05	64,8025
2007	129,6	0	0	0	0	0	137,96	-8,36	69,8896	138,74	-9,14	83,5396
2008	139,4	1	1	1	139,4	139,4	135,35	4,05	16,4025	135,93	3,47	12,0409
2009	151,7	2	4	16	303,4	606,8	132,74	18,96	359,4816	132,72	18,98	360,2404
2010	117,5	3	9	81	352,5	1057,5	130,13	-12,63	159,5169	129,11	-11,61	134,7921
Итого	965,7	0	28	196	-73	3846,4	965,72	-0,02	684,6960	965,58	0,12	681,498

$$a_0 = \frac{\sum t^4 \sum y - \sum t^2 \sum t^2 y}{n\sum t^4 - \sum t^2 \sum t^2} = \frac{196 \times 965{,}7 - 28 \times 3846{,}4}{7 \times 196 - 28 \times 28} = \frac{81578}{588} = 138{,}74$$

$$a_1 = \frac{\sum ty}{\sum t^2} = \frac{-73}{28} = -2{,}61$$

$$a_2 = \frac{n\sum t^2 y - \sum t^2 \sum y}{n\sum t^4 - \sum t^2 \sum t^2} = \frac{7 \times 3846{,}4 - 28 \times 965{,}7}{7 \times 196 - 28 \times 28} = \frac{-114{,}8}{588} = -0{,}2$$

В результате получаем следующие уравнение основной тенденции производства овощей (включая овощи закрытого грунта) в Пензенской области за 2004−2010 гг.:

$$\bar{y}_t = 138{,}74 - 2{,}61t - 0{,}2t^2$$

Подставляя в уравнение принятые обозначения *t*, вычислим выравненные уровни ряда динамики:

2004 г. – $\bar{y}_1 = 138{,}74 - 2{,}61 \times (-3) - 0{,}2 \times (-3)^2 = 144{,}77$

2005 г. – $\bar{y}_2 = 138{,}74 - 2{,}61 \times (-2) - 0{,}2(-2)^2 = 143{,}16$

и т.д. (см. значения в графе 9 таблицы 118).

По итоговым данным определяем остаточное среднее квадратическое отклонение:

для уравнения прямой: $\sigma_{ост} = \sqrt{\frac{684{,}696}{7}} = \pm 9{,}89$

для уравнения параболы 2-го порядка: $\sigma_{ост} = \sqrt{\frac{681{,}498}{7}} = \pm 9{,}87$

Из сравнения полученных значений следует, что по критерию минимальности предпочтение следует отдать трендовой модели $\bar{y}_t = 138{,}74 - 2{,}61t - 0{,}2t^2$, синтезированной на основе уравнения параболы 2-го порядка $\bar{y}_t = a_0 + a_1 t + a_2 t^2$.

Основная тенденция (тренд) показывает, как воздействуют систематические факторы на уровень ряда динамики, а колеблемость уровней около тренда служит мерой воздействия остаточных факторов. Её можно измерить показателем *средняя квадратическая ошибка уравнения тренда* (остаточное среднее квадратическое отклонение тренда, скорректированное по числу степеней свободы):

$$\sigma_{\text{ост}} = \sqrt{\frac{\sum\left(y_i - \bar{y}_{t_i}\right)^2}{n-m}},$$

где m – число параметров уравнения тренда;

$(n-m)$ – число степеней свободы.

Относительной мерой колеблемости является коэффициент вариации

$$\nu = \frac{\sigma}{\bar{y}} \times 100\% .$$

На основе данных примера рассчитаем:

– среднее квадратическое отклонение:

для уравнения прямой:

$$\sigma_{\text{ост}} = \sqrt{\frac{684{,}696}{7-3}} = \pm 13{,}08$$

для уравнения параболы 2-го порядка:

$$\sigma_{\text{ост}} = \sqrt{\frac{681{,}498}{7-5}} = \pm 18{,}46;$$

– коэффициент вариации:

для уравнения прямой:

$$\nu = \frac{\pm 13{,}08}{137{,}96} \times 100\,\% = \pm 9{,}48\,\%$$

для уравнения параболы 2-го порядка:

$$\nu = \frac{\pm 18{,}46}{137{,}96} \times 100\,\% = \pm 13{,}38\,\%.$$

Таким образом, колеблемость валового сбора овощей (включая овощи закрытого грунта), обусловленная воздействием остаточных факторов, минимальна по трендовой модели уравнения прямой и составляет ± 13,08 тыс. т. Её вариация по уравнению прямой составляет ± 9,48 %.

В рядах динамики, уровни которых являются месячными или квартальными показателями, наряду со случайными колебаниями часто наблюдаются *сезонные колебания*.

Сезонные колебания – более или менее устойчивые внутригодовые колебания уровней развития социально-экономических явлений. Ярко выраженный сезонный характер имеет сельскохозяйственное производство, особенно растениеводство в условиях открытого грунта. Это вызывает неравномерность использования трудовых ресурсов, напряженность в работе транспорта и т.д.

В качестве иллюстрации рядов с сезонными колебаниями рассмотрим данные об объеме оборота розничной торговли в регионе за 2 года по месяцам (таблица 119) и их графическое изображение (рисунок 12)

Таблица 119 − Объем оборота розничной торговли в регионе, млрд. руб. (данные условные)

Год	Месяц											
	Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
1	4,7	4,5	5,3	5,4	5,7	5,8	6,3	6,6	6,4	7,0	7,2	8,9
2	6,7	6,4	7,3	7,4	7,7	7,8	8,1	8,4	8,2	8,5	8,0	9,7

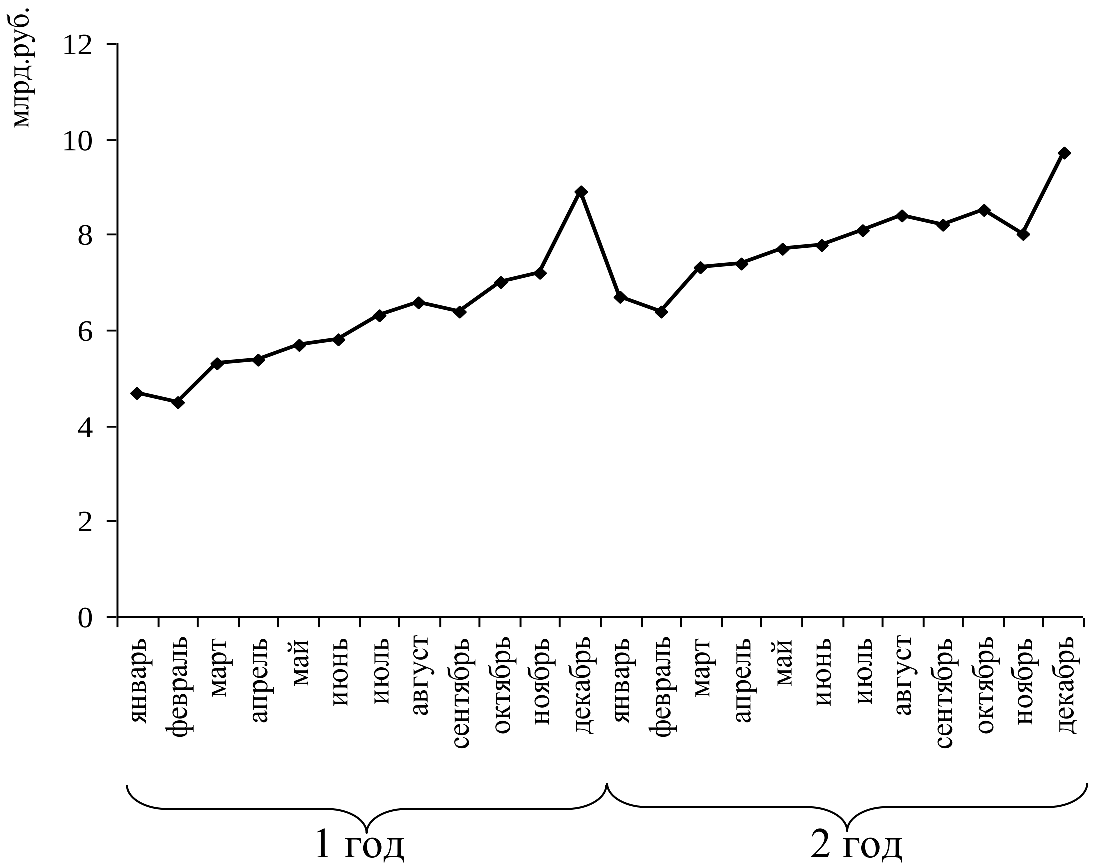

Рисунок 12 – Динамика объема товарооборота розничной торговли в регионе за 2 года по месяцам

Вместо месячных показателей могут быть квартальные (таблица 120). Если колебания не случайны, они сохраняются и в квартальных уровнях.

Таблица 120 − Объем оборота розничной торговли в регионе за 2 года по кварталам, млрд. руб.

Год	1 год				2 год			
Квартал	I	II	III	IV	I	II	III	IV
Объем оборота розничной торговли	14,5	17,0	19,3	23,0	20,4	22,9	24,6	26,1

При статистическом изучении в рядах внутригодовой динамики сезонных колебаний решаются следующие две взаимосвязанные задачи: выявление специфики развития изучаемого явления во внутригодовой динамике; измерение сезонных колебаний изучаемого явления с построением сезонной волны.

Для измерения сезонных колебаний исчисляются индексы сезонности. *Индекс сезонности (* I_s *)* – относительный показатель, который используют для расчета сезонной составляющей. При исчислении индексов применяют различные методы, выбор которых зависит от характера общей тенденции ряда динамики.

Методы расчета индексов сезонности.

1. Если ряд динамики не имеет ярко выраженной тенденции развития, то индексы сезонности исчисляют непосредственно по эмпирическим данным без их предварительного выравнивания. Для расчета индексов сезонности необходимо иметь данные по периодам не менее чем за три года. Сущность метода заключается в расчете средних по периодам $\tilde{y}_i$ и для всего анализируемого ряда $\bar{y}$. В данном варианте индекс сезонности определяется по следующей формуле:

$$I_s = \frac{\tilde{y}_i}{\bar{y}} 100\%,$$

где $\tilde{y}_i$ – фактическое месячное (квартальное) значение уровня;

$\bar{y}$ – среднемесячное (квартальное) значение за год.

В качестве среднего уровня ряда может быть использована также мода или другая структурная средняя.

2. Если ряд динамики имеет тренд (нестационарный ряд динамики), то расчет проходит следующие этапы:

1) определяют по внутригодовым уровням ряда (месячным, квартальным) за несколько лет расчетные (выравненные) уровни по методикам скользящей средней или аналитического выравнивания ($\bar{y}_t$);

2) исчисляют относительную величину фактических значений уровней ряда (y_i) и выравненных (расчетных) значений ($\bar{y}_t$);

3) усредняют полученные показатели сезонности за весь исследуемый период:

$$I_s = \sum \frac{y_i}{\bar{y}_t} 100\%,$$

3. Для получения устойчивой тенденции сезонных колебаний, на которых бы не отражались особенности развития явлений и процессов в конкретные периоды, индекс сезонности находят по формуле

$$\bar{I}_s = \left(\sum I_s\right) \div t,$$

где I_s – индекс сезонности;

t – число лет.

Полученные усредненные индексы сезонности являются искомыми, характеризующими «сезонную волну».

Сезонная волна – графическое изображение полученных индексов сезонности.

На основе полученного индекса сезонности рассчитывают коэффициент сезонности (K_s):

$$K_s = \sqrt{\sum\left(I_s - 1\right)^2 \div n},$$

где n – число периодов.

В таблице 121 показан расчет индексов сезонности и абсолютных отклонений уровней от среднего на примере данных об объеме оборота розничной торговли в регионе.

Для характеристики силы колеблемости уровней динамического ряда из-за сезонной неравномерности предлагается использовать среднее квадратическое отклонение индексов сезонности (в процентах) от 100 %:

$$\sigma_{\text{сез}} = \sqrt{\frac{\sum\left(I_{сез} - 100\,\%\right)^2}{n}} = \sqrt{\frac{1349{,}06}{12}} = 10{,}6\,.$$

В экономических исследованиях приходится изучать динамику нескольких показателей одновременно, т.е. рассматривать параллельно несколько динамических рядов. В этих случаях можно встретить ряды, у которых колебания уровней взаимообусловлены. Например, динамика рыночных цен на продукцию в известной степени зависит от объемов производства этой продукции и т.п.

Таблица 121 − Сезонные колебания объема оборота розничной торговли в регионе (данные условные)

Месяц	Объем оборота розничной торговли, млрд. руб. y_i	Индекс сезонности, % к среднемесячному уровню $\frac{y_i}{\bar{y}}100\%$	Абсолютное отклонение от среднего месячного уровня $y_i-\bar{y}$	Абсолютное отклонение, % к среднему месячному уровню $\frac{(y_i-\bar{y})}{\bar{y}}100\%$	$(I_{сез}-100\%)^2$	$(y_i-\bar{y})^2$
Январь	6,7	85,4	-1,15	-14,6	213,20	1,30
Февраль	6,4	81,5	-1,45	-18,5	342,30	2,10
Март	7,3	93,0	-0,55	-7,0	49,00	0,30
Апрель	7,4	94,3	-0,45	-5,7	32,50	0,20
Май	7,7	98,1	-0,15	-1,9	3,60	0,02
Июнь	7,8	99,4	-0,05	-0,6	0,36	0,003
Июль	8,1	103,2	+0,25	3,2	10,20	0,06
Август	8,4	107,0	+0,55	7,0	49,00	0,30
Сентябрь	8,2	104,4	+0,35	4,4	19,40	0,10
Октябрь	8,5	108,3	+0,65	8,3	68,90	0,40
Ноябрь	8,0	101,9	+0,15	1,9	3,60	0,02
Декабрь	9,7	123,6	+1,85	23,6	557,00	3,40
Итого	94,2	-	-	-	1349,06	8,20

При изучении таких рядов динамики возникает необходимость измерить зависимость между ними, т.е. определить, насколько изменения уровней одного ряда зависят от изменения уровней другого ряда. Эта задача решается путем коррелирования уровней рядов динамики, на основе исчисления коэффициента корреляции по следующей формуле

$$r=\frac{\overline{xy}-\bar{x}\times\bar{y}}{\sigma_x\times\sigma_y} \text{ или } r=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\left[\sum x^2-\frac{\left(\sum x\right)^2}{n}\right]\left[\sum y^2-\frac{\left(\sum y\right)^2}{n}\right]}}.$$

Прежде чем делать вывод о тесноте связи между рассматриваемыми рядами динамики, их необходимо проверить на автокорреляцию.

Автокорреляция – зависимость между последовательными (соседними) уровнями ряда динамики. Например, численность населения за определенный год зависит (при прочих равных условиях) от численности в предшествующие годы; поголовье скота, численность которого в каждый год зависит от численности поголовья в предшествующие годы и т.д.

Наличие автокорреляции устанавливается при помощи коэффициента автокорреляции для парной линейной связи:

$$r = \frac{\overline{xy} - \bar{x} \times \bar{y}}{\sigma_x \times \sigma_y}.$$

Коэффициент автокорреляции можно рассчитывать либо между соседними уровнями, либо между уровнями, сдвинутыми на любое число единиц времени m. Этот сдвиг, именуемый временным лагом, определяет порядок коэффициента автокорреляции: 1-го порядка при $m = 1$, т.е. между соседними уровнями; 2-го порядка при $m = 2$, т.е. при сдвиге уровней на два периода, и т.д.

Если исходные фактические уровни ряда, относящиеся к определенному моменту времени (или периоду) t, обозначить через y_t, то сдвинутые уровни (в зависимости от направления сдвига) соответственно обозначают y_{t+1} или y_{t-1}. Тогда формула коэффициента автокорреляции примет вид

$$r_a = \frac{\overline{y_t \times y_{t+1}} - \bar{y}_t \times \bar{y}_{t+1}}{\sigma_{y_t} \times \sigma_{y_{t+1}}} \text{ или } r_a = \frac{\overline{y_t \times y_{t-1}} - \bar{y}_t \times \bar{y}_{t-1}}{\sigma_{y_t} \times \sigma_{y_{t+1}}}.$$

Если значение последнего уровня мало отличается от первого, то для того чтобы сдвинутый ряд не укорачивался, его можно условно дополнить, принимая, что $y_n = y_1$. Тогда $\bar{y}_t = \bar{y}_{t+1}$ и $\sigma_{y_t} = \sigma_{y_{t+1}}$, поскольку они рассчитываются для одного и того же ряда.

При такой замене формула коэффициента автокорреляции принимает вид

$$r_a = \frac{\overline{y_t \times y_{t+1}} - (\bar{y}_t)^2}{\sigma^2_{y_t}} \text{ или } \frac{\sum y_t \times y_{t+1} - n(\bar{y}_t)^2}{\sum y_t^2 - n(\bar{y}_t)^2}.$$

Для определения величины коэффициента автокорреляции, которая достоверно свидетельствует о наличии или отсутствии автокоррелированности наблюдений, необходимо фактические коэффициенты сравнить с табличными (см. приложение 8). Если фактическая величина больше его критического значения, указанного в таблице, то делается заключение о том, что автокорреляция имеется. Если фактическая величина меньше табличного, то следует отказаться от гипотезы о наличии автокорреляции.

Если между уровнями ряда (при корреляции рядов динамики) существует автокорреляция, она должна быть устранена.

Пример. По данным, представленным в таблице 122, об изменении расхода кормов и валового надоя молока в хозяйствах населения в Пензенской области за 5 лет рассмотрим применение коррелирования уровней для измерения связи между рядами динамики.

Таблица 122 − Исходные и расчетные данные для определения коэффициента корреляции

Год	Расход кормов, тыс. т (x)	Валовой надой молока, тыс. т (y)	x^2	y^2	xy
2006	286,3	324,7	81967,7	105430,1	92961,6
2007	288,0	329,8	82944,0	108768,0	94982,4
2008	292,9	337,5	85790,4	113906,2	98853,8
2009	291,7	315,0	85088,9	99225,0	91885,5
2010	263,1	287,7	69221,6	82771,3	75693,9
Итого	1422	1594,7	405012,6	510100,7	454377,2

Из данных таблицы видно, что

$$\bar{x} = \frac{1422}{5} = 284,4$$

$$\bar{y} = \frac{1594,7}{5} = 318,9$$

$$\overline{xy} = \frac{454377,13}{5} = 90875,4$$

$$\bar{x} \times \bar{y} = 284,4 \times 318,9 = 90695,2$$

$$\sigma_x^2 = \frac{\sum x^2}{n} - (\bar{x})^2 = \frac{405012,6}{5} - (284,4)^2 = 119,2$$

$$\sigma_y^2 = \frac{\sum y^2}{n} - (\bar{y})^2 = \frac{510100,7}{5} - (318,9)^2 = 322,9$$

$r = \frac{90875,4 - 90695,2}{\sqrt{119,2 \times 322,9}} = 0,919$ говорит, что в данном случае имеет место прямая (знак «+») и заметная (величина 0,919) связь между уровнями и рядов расхода кормов и валового надоя молока в хозяйствах населения в Пензенской области.

Для расчета коэффициента автокорреляции по первому ряду построим таблицу.

Таблица 123 − Расчетная таблица для определения коэффициента автокорреляции

Год	Расход кормов, тыс. т (x_t)	Расход кормов со сдвигом на один год, тыс. т (x_{t+1})	$x_t x_{t+1}$	x_t^2
2006	286,3	288,0	82454,4	81967,7
2007	288,0	292,9	84355,2	82944,0
2008	292,9	291,7	85438,9	85790,4
2009	291,7	263,1	76746,3	85088,9
2010	263,1	286,3	75325,5	69221,6
Итого	1422	1422	404320,3	405012,6

$$\overline{x_t \times x_{t+1}} = \frac{\sum x_t \times x_{t+1}}{n} = \frac{404320,3}{5} = 80864,1$$

$$\bar{x}_t = \frac{\sum x_t}{n} = \frac{1422}{5} = 284{,}4$$

$$\left(\bar{x}_t\right)^2 = 284{,}4^2 = 80883{,}4$$

$$\sigma^2_{x_t} = \frac{\sum x_t^{\ 2}}{n} - \left(\bar{x}_t\right)^2 = \frac{405012{,}6}{5} - 80883{,}4 = 119{,}1$$

$$r_a = \frac{\overline{x_t \times x_{t+1}} - \left(\bar{x}_t\right)^2}{\sigma^2_{x_t}} = \frac{80864{,}1 - 80883{,}4}{119{,}1} = -0{,}162$$

Далее возникает вопрос о величине коэффициента актокорреляции, которая свидетельствует о наличии или отсутствии автокоррелированности наблюдений. Поэтому необходимо фактические коэффициенты, полученные расчетным путем, сравнить с табличными (см. приложение 8).

Проведем сопоставление полученного коэффициента автокорреляции с их табличной величиной при численности n = 5. При уровне значимости *P = 0,05* (5%-й уровень) величина r_a может только в пяти случаях из ста превысить 0,253. В нашем примере r_a = *−0,162*, т.е. он ниже табличного значения при уровне существенности 0,05. Следовательно, автокорреляция отсутствует.

Затем проводим аналогичные вычисления для второго ряда (y − валовой надой молока) (таблица 124).

Таблица 124 − Расчетная таблица для определения коэффициента автокорреляции

Год	Валовой надой молока, тыс.т (y)	Валовой надой молока со сдвигом на один год, тыс. т (y_{t+1})	$y_t y_{t+1}$	y_t^2
2006	324,7	329,8	107086,1	105430,1
2007	329,8	337,5	111307,5	108768,0
2008	337,5	315,0	106312,5	113906,3
2009	315,0	287,7	90625,5	99225,0
2010	287,7	324,7	93416,2	82771,3
Итого	1594,7	1594,7	508747,8	510100,7

По итоговым данным таблицы рассчитаем необходимые величины:

$$\bar{y}_t = \frac{\sum y_t}{n} = \frac{1594,7}{5} = 318,9$$

$$(\bar{y}_t)^2 = 318,9^2 = 101697,2$$

$$\sigma^2_{y_t} = \frac{\sum y_t^{\ 2}}{n} - (\bar{y}_t)^2 = \frac{510100,7}{5} - (318,9)^2 = 322,9$$

$$\overline{y_t \times y_{t+1}} = \frac{\sum y_t \times y_{t+1}}{n} = \frac{508747,8}{5} = 101749,6$$

Значение подставим в формулу коэффициента автокорреляции:

$$r_a = \frac{\overline{y_t \times y_{t+1}} - (\bar{y}_t)^2}{\sigma^2_{y_t}} = \frac{101749,6 - 101697,2}{322,9} = \frac{52,4}{322,9} = 0,162$$

0,162 < 0,253 – автокорреляция отсутствует.

Существует несколько *способов исключения автокорреляции в рядах динамики*: *коррелирование отклонений от выравненных уровней, коррелирование последовательных разностей*.

При *коррелировании отклонений фактических уровней от выравненных* необходимо:

1) произвести аналитическое выравнивание сравниваемых рядов по любому рациональному многочлену;

2) определить величину отклонения каждого фактического уровня ряда динамики от соответствующего ему выравненного значения;

3) произвести коррелирование полученных отклонений.

Формула коэффициента корреляции между отклонениями:

$$r_{d_x d_y} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^{\ 2} \sum d_y^{\ 2}}},$$

где $d_x = x_i - \bar{x}_t$;

$d_y = y_i - \bar{y}_t$.

Данный коэффициент характеризует степень связи между отклонениями фактических уровней сравниваемых рядов от соответствующих им выравненных уровней коррелируемых рядов динамики.

При *коррелировании разностей* измеряется теснота связи между разностями последовательных величин уровней в каждом динамическом ряду. В этом случае показателем тесноты связи между изучаемыми рядами является коэффициент корреляции разностей:

$$r_{\Delta_x \Delta_y} = \frac{\sum \Delta_x \Delta_y}{\sqrt{\sum \Delta_x^2 \sum \Delta_y^2}},$$

где $\Delta_x = x_i - x_{i-1}$;

$\Delta_y = y_i - y_{y-1}$.

Если недостаточно точно подобрано уравнение тренда (или по другим причинам), остаточные величины могут содержать автокорреляцию.

В целях проверки используют коэффициент автокорреляции для рядов с нулевым значением среднего уровня:

$$r_a = \frac{\sum_{t=1}^{n-1} \varepsilon_t \varepsilon_{t+1}}{\sum_{t=1}^{n} \varepsilon_t^2},$$

где $\varepsilon_t = y_i - \bar{y}_t$;

$\varepsilon_{t+1} = y_i - \bar{y}_{t+1}$,

или критерий Дурбина-Ватсона (в иной транскрипции Дурбина-Уотсона, DW):

$$d = \frac{\sum_{t=1}^{n-1} \left(\varepsilon_t - \varepsilon_{t+1}\right)^2}{\sum_{t=1}^{n} \varepsilon_t^2}.$$

Для проверки нулевой гипотезы об отсутствии автокорреляции в остаточных величинах рассчитанное фактическое значение d сравнивают с табличными d_1 и d_2 (см. приложение 9):

1) если $d > d_2$ (до $4 - d_2$), гипотеза об отсутствии автокорреляции принимается;

2) если $d < d_1$, гипотеза об отсутствии автокорреляции отвергается;

3) если $d_1 \leq d \leq d_2$ и d_2 или $(4 - d_2) \leq d \leq (4 - d_1)$, ничего определенного сказать нельзя, и требуется дальнейшее исследование (уточнение тренда, увеличение числа наблюдений и пр.);

4) если $d > (4 - d_1)$, имеет место отрицательная автокорреляция.

Определяемые в анализе рядов динамики показатели изменения уровней, тренда, сезонной волны имеют широкое применение при прогнозировании, т.е. при получении статистической оценки возможной меры развития социально экономических явлений на будущее.

Под *экстраполяцией* понимается распространение выявленных в анализе рядов динамики закономерностей развития изучаемого явления на будущее.

Основой прогнозирования является предположение, что закономерность, действующая внутри анализируемого ряда динамики, выступающего в качестве базы прогнозирования, сохраняется и в дальнейшем. Точность прогноза зависит от того, насколько обоснованными окажутся предположения о сохранении на будущее действий тех факторов, которые сформировали в базисном ряду динамики его основные компоненты.

Важное значение при экстраполяции имеет продолжительность базисного ряда динамики и сроков прогнозирования. На каждые 7–10 периодов фактических данных осуществляют один период прогноза.

Экстраполяцию ряда динамики можно осуществить различными способами. Простейшие *приемы экстраполяции рядов динамики*:

1. Если при анализе ряда динамики обнаруживается, что абсолютные приросты уровней примерно постоянные, можно рассчитать средний абсолютный прирост и последовательно приба-

вить его к последнему уровню ряда столько раз, на сколько периодов экстраполируется ряд:

$$\bar{y}_{i+1} = \bar{y}_i + \bar{\Delta}_t,$$

где $\bar{y}_{i+1}$ – экстраполируемый уровень;

$(i+t)$ – номер этого уровня (года);

i – номер последнего уровня (года) исследуемого периода, за который рассчитан $\bar{\Delta}$;

$\bar{\Delta}$ – средний абсолютный прирост.

По данным примера прогнозируемый объем производства комбикорма в РФ составит в 2011 г. 16 + 1,2 = 17,2 млн. т;

2. Если за исследуемый ряд лет годовые коэффициенты роста остаются более-менее постоянными, можно рассчитать средний коэффициент роста и умножить последний уровень ряда на средний коэффициент роста в степени, соответствующей периоду экстраполяции:

$$\bar{y}_{i+1} = y_i \times \bar{k}_p^t,$$

где y_i – последний уровень ряда динамики;

t – срок прогноза;

$\bar{k}_p$ – средний коэффициент роста.

По данным примера прогнозируемый объем производства комбикорма в РФ в 2011 г. составит $16{,}0 \times 1{,}098^1 = 17{,}6$ млн. т;

3. Можно экстраполировать ряды на основе выравнивания их по определенной аналитической формуле. Зная уравнение для теоретических уровней и подставляя в него значения t за пределами исследованного ряда, можно рассчитать для данных t вероятностные уровни $\bar{y}_t$:

$$\bar{y}_t \pm t_\alpha \cdot \sigma_{\bar{y}_t},$$

где $\bar{y}_t$ – расчетное значение уровня ряда по уравнению тренда;

$\sigma_{\bar{y}_t}$ – средняя квадратическая ошибка тренда, скорректированная по числу степеней свободы;

t_α – доверительная величина.

Применение формулы проиллюстрируем на данных экстраполяции производства овощей (включая овощей закрытого грунта) в Пензенской области в 2011 г.

Число степеней свободы при n = 5 и m = 3 составляет 2. При уровне значимости α = 0,05 коэффициент доверия t_α по таблице Стьюдента равен 4,3. Значение остаточного среднего квадратического отклонения, скорректированное на число степеней свободы равно ±13,08.

Прогнозное значение валового сбора овощей (включая овощи закрытого грунта) в Пензенской области на 2011 г. составит:

$$\bar{y}_t = 137{,}96 - 2{,}61 \times 4 = 127{,}52 \text{ тыс.т.}$$

Значение вероятностных границ интервала составляет: $127{,}52 \pm 4{,}3 \times 13{,}08$. Следовательно, с вероятностью 0,95 верхняя граница валового сбора овощей (включая овощи закрытого грунта) в Пензенской области составит $127{,}52 - 56{,}24 = 71{,}28$ тыс. т, а нижняя граница – $127{,}52 + 56{,}24 = 183{,}76$ тыс. т.

Пользуясь этим методом, следует помнить, что экстраполяция динамического ряда на основе уравнения, полученного при выравнивании, только тогда может дать оценки, близкие к реальным значениям, когда в эмпирическом ряду исключаются случайные колебания, измеряемые средним квадратическим отклонением разности $(y_t - \bar{y}_t)$, и между случайными отклонениями отсутствует автокорреляция.

В ходе прогнозирования следует учитывать, что прогнозирование основанное только на обработке данных наблюдения – рискованно, если оно не учитывает множества взаимосвязанных фактов и моментов, которые способны изменить тенденцию развития в будущем. Экономическое прогнозирование невозможно без хорошего знания изучаемого явления и владения различными методами обработки динамических рядов, которые в каждом отдельном случае позволяют обнаружить общую закономерность изменения, периодичность в повышении или снижении уровней, случайные колебания, автокорреляцию и корреляцию между отдельными рядами.

Задача 97

Определите вид рядов динамики, характеризующих изменение следующих статистических показателей:

1) поголовье птицы в хозяйствах всех категорий (по состоянию на конец каждого года);

2) производство обуви по субъектам РФ по годам;

3) число умерших по годам;

4) среднемесячная заработная плата работников по отраслям экономики по годам;

5) задолженность по налоговым платежам в бюджетную систему РФ (на начало каждого года);

6) индексы производства продукции сельского хозяйства по категориям хозяйств (за ряд лет);

7) удельный вес товарной продукции легкой промышленности в общем объеме продукции по годам.

Задача 98

Известны нижеследующие данные о производстве стали в двух странах за 2005-2010 гг. (таблица 125).

Таблица 125 − Динамика производства стали, млн т (данные условные)

Страна	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
А	9,5	12,8	14,5	16,9	19,1	20,2
Б	20,6	28,3	35,7	43,2	45,8	55,8

С целью анализа производства стали в двух странах необходимо: привести ряды динамики к общему основанию; изобразить относительные величины динамики в виде линейной диаграммы;

рассчитать коэффициент опережения производства стали в стране Б по сравнению со страной А. Сделайте выводы.

Задача 99

В таблице 126 представлены данные о численности населения и производстве молока в России.

Таблица 126 − Данные о численности населения и производстве молока в России за 2001-2010 гг.

Показатель	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
Производство молока, млн. т	32,9	33,5	33,4	32,0	31,1	31,3	32,0	32,4	32,6	31,9
Численность населения, млн. чел.	146,3	145,2	145,0	144,2	143,5	142,8	142,2	142,0	141,9	

Определите:

1) средние уровни представленных рядов динамики;

2) производный ряд динамики производства молока на душу населения для каждого года и рассчитайте средний уровень ряда динамики; цепные и базисные темпы роста и прироста; абсолютные значения 1 % прироста;

3) результаты расчетов представьте в табличной форме и сделайте выводы.

Задача 100

Производство зерна в хозяйстве возрастало следующими темпами к уровню предыдущего года: 2005 г. – 130 %, 2006 – 109 %, 2007 г. – 63 %, 2008 г. – 266 %, 2009 г. – 103 %, 2010 г. – 124 %. Производство молока в 2004 г. составило 28 тыс. ц, а в 2010 г. – 33 тыс. ц. Сравните динамику производства зерна и молока в хозяйстве. Сделайте выводы.

Задача 101

На основании данных таблицы восстановите недостающие показатели динамики и уровни динамического ряда, помеченные многоточием.

Таблица 127 – Динамика производства зерна в Пензенской области в хозяйствах всех категорий в весе после доработки

Год	Валовой сбор зерна, тыс. т	Базисный показатель динамики		
		абсолютный прирост, тыс. т	темп роста, %	темп прироста, %
2005	992,5	-	100,0	-
2006	…	115,4	…	…
2007	…	…	93,9	…
2008	…	…	…	43,3
2009	…	…	…	47,2
2010	…	-585,2	…	…

Задача 102

Данные о валовом сборе кормовых корнеплодов (включая сахарную свеклу на корм скоту) за 2001-2010 гг. в хозяйствах всех категорий в Пензенской области представлены в таблице 128.

Таблица 128 − Динамика валового сбора кормовых корнеплодов (включая сахарную свеклу на корм скоту) в хозяйствах всех категорий в Пензенской области за 2001-2010 гг., тыс. т

2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
18,5	18,8	33,5	28	42,6	28,2	30,6	58,7	69,1	24,6

Определите:

1) цепные и базисные темпы роста и прироста; абсолютные значения 1 % прироста. Результаты расчетов представьте в табличной форме;

2) средний уровень ряда динамики; среднегодовые абсолютные приросты, среднегодовые темпы роста и прироста производства за 2001-2010 гг.;

3) проведите подбор математической функции для определения тренда методом аналитического выравнивания и оцените ее адекватность.

Сделайте выводы по данным проведенных расчетов.

Задача 103

На основании данных таблицы 129 восстановите недостающие показатели динамики и уровни динамического ряда, помеченные многоточием.

Таблица 129 − Динамика производства молока в Пензенской области в хозяйствах всех категорий

Год	Производство молока, тыс. т	Цепной показатель динамики			
		абсолютный прирост, тыс. т	темп роста, %	темп прироста, %	абсолютное значение 1 % прироста, тыс. т
2003	488,6	-	-	-	-
2004	…	9,7	…	…	…
2005	…	…	101,1	…	…
2006	…	…	…	2,0	…
2007	…	…	…	…	…
2008	…	12,1	…	…	5,371
2009	…	-38	…	…	…
2010	…	…	91,3	…	…

Задача 104

На основании данных таблицы 130 о производстве продукции животноводства в хозяйствах всех категорий в Пензенской области произведите изучение общей тенденции производства продукции животноводства на основе:

1) сглаживания уровней рядов динамики с помощью механических приемов выравнивания (укрупнения рядов, скользящей средней);

2) аналитического выравнивания. Выразите общую тенденцию развития каждого вида продукции за 1999-2010 гг. соответствующими математическими уравнениями;

3) определения выравненных (теоретических) уровней рядов динамики.

Сделайте выводы по результатам расчетов.

Таблица 130 − Динамика производства некоторых видов продукции животноводства в Пензенской области

Год	Молоко, тыс. т	Яйца, млн. шт.	Шерсть, т	Мясо (в убойном весе), тыс. т
1999	473,3	340,0	216	52,8
2000	435,9	330,4	192	54,0
2001	435,3	263,1	205	47,9
2002	466,5	266,6	218	50,6
2003	488,6	275,2	221	64,0
2004	498,3	276,7	234	64,7
2005	503,8	229,2	226	68,6
2006	514,1	255,1	261	77,8
2007	537,1	269,9	318	91,6
2008	549,2	275,3	277	99,8
2009	511,2	323,4	236	105,3
2010	466,6	332,3	262	114,2

Задача 105

В таблице 131 представлены данные о ценах на огурцы свежие в регионе за 2008-2010 гг. по месяцам. Определите индекс сезонности, а результаты расчетов представьте в табличной форме. Постройте сезонную волну и сделайте выводы.

Таблица 131 – Динамика цен на огурцы свежие в регионе за 2008-2010 гг. по месяцам, руб. за 1 кг

Год	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2008	83,6	114,8	113,9	87,4	59,4	41,6	34,3	29,0	42,6	68,0	72,4	78,8
2009	92,9	113,4	103,3	90,2	71,7	41,8	31,0	25,4	26,9	55,8	62,5	71,8
2010	84,5	113,0	101,9	83,8	70,3	42,9	32,3	31,1	43,6	80,3	77,3	79,3

Задача 106

По данным о яйценоскости кур-несушек за 2008-2010 гг. (таблица 132) проведите анализ внутригодовой динамики яйценоскости при помощи определения индекса сезонности:

1) методом постоянной средней;

2) методом аналитического выравнивания по прямой. Представьте графически сезонную волну развития изучаемого явления по месяцам года.

Таблица 132 – Динамика яйценоскости кур-несушек за 2008-2010 гг., шт. (данные условные)

Месяц	2008 г.	2009 г.	2010 г.
Январь	12	13	14
Февраль	15	17	16
Март	18	20	22
Апрель	20	23	25
Май	22	26	26
Июнь	23	28	27
Июль	27	27	28
Август	26	25	24
Сентябрь	22	24	23
Октябрь	17	19	20
Ноябрь	12	16	17
Декабрь	10	13	14

Задача 107

В таблице 133 представлены данные о валовом сборе сена в хозяйствах всех категорий в Пензенской области

Таблица 133 – Валовой сбор сена в хозяйствах всех категорий в Пензенской области за 2008-2010 гг., тыс. т

Год	Сено		
	многолетних трав	однолетних трав	естественных сенокосов (включая улучшенные)
2001	203,0	31,8	285,8
2002	157,6	51,0	320,9
2003	145,4	51,1	384,7
2004	197,4	43,8	430,2
2005	162,4	34,5	507,9
2006	160,3	31,9	507,9
2007	149,3	24,9	508,0
2008	170,0	49,6	514,6
2009	137,2	54,1	507,6
2010	90,6	24,8	489,0

Проведите экстраполяцию на ближайшие годы на основе: среднего абсолютного прироста; среднего темпа роста; аналитического выравнивания уровней ряда динамики. Сравните полученные результаты и выберите наилучший прогноз.

10 ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ

Индекс – это относительная величина сравнения сложных совокупностей и отдельных единиц. При этом под сложной понимается такая статистическая совокупность, отдельные элементы которой непосредственно не подлежат суммированию.

Основные задачи индексного метода:

– оценка динамики обобщающих показателей, характеризующих сложные, непосредственно несоизмеримые совокупности;

– анализ влияния отдельных факторов на изменение результативных обобщающих показателей;

– анализ влияния структурных сдвигов на изменение средних показателей однородной совокупности;

– оценка территориальных, в том числе международных, сравнений.

Индексы классифицируют по степени охвата единиц совокупности, базе сравнения, виду весов, форме построения и составу явления (таблица 134).

Таблица 134 – Классификация статистических индексов

Признак	Вид индекса
По степени охвата единиц Совокупности	Индивидуальные Общие (сводные)
По базе сравнения	Динамические (цепные, базисные) Индексы выполнения плана Территориальные
По виду индексов	С постоянными весами С переменными весами
По форме построения	Агрегатные Средние взвешенные
По составу явления	Индексы постоянного состава Индексы переменного состава

Условные обозначения, используемые в теории индексного метода:

p – цена за единицу товара (услуги);

q – количество (объем) какого-либо продукта (товара) в натуральном выражении;

pq – общая стоимость продукции данного вида (товарооборот);

z – себестоимость единицы продукции (изделия);

zq – общая себестоимость продукции данного вида (денежные затраты на ее производство);

T – общие затраты времени на производство продукции или общая численность работников;

$w = \dfrac{q}{T}$ – производство продукции данного вида в единицу времени (либо выработка продукции на одного работника, т.е. производительность труда);

$t = \dfrac{T}{q}$ – затраты рабочего времени на единицу продукции (трудоемкость единицы продукции);

1 – подстрочный символ показателя текущего (отчетного) периода;

0 – подстрочный символ показателя предшествующего (базисного) периода.

Индивидуальный индекс (i) характеризует динамику уровня изучаемого явления во времени за два сравниваемых периода или отражает соотношение элементов совокупности.

Основной элемент индексного соотношения – индексируемая величина. *Индексируемая величина* – признак, изменение которого характеризует индекс.

Основные формулы вычисления индивидуальных индексов:

$i_p = \dfrac{p_1}{p_0}$ – индекс цены;

$i_q = \dfrac{q_1}{q_0}$ – индекс физического объема реализации;

$i_{pq} = \dfrac{p_1 q_1}{p_0 q_0}$ – индивидуальный индекс товарооборота;

$$i_z = \frac{z_1}{z_0}$$ – индекс себестоимости;

$$i_{zq} = \frac{z_1 q_1}{z_0 q_0}$$ – индекс затрат на производство продукции;

$$i_w = \frac{w_0}{w_1} = (q_1 \div T_1) \div (q_0 \div T_0)$$ – индекс количества продукции, произведенной в единицу времени;

$$i_t = \frac{t_1}{t_0} = (T_1 \div q_1) \div (T_0 \div q_0)$$ – индекс трудоемкости;

$$i_{1/t} = \frac{t_0}{t_1} = (q_1 \div T_1) \div (q_0 \div T_0)$$ – индекс производительности труда (по трудоемкости).

Приведенные индивидуальные индексы взаимосвязаны:

$$i_{pq} = i_p \times i_q$$

$$i_{zq} = i_z \times i_q$$

$$i_w = 1 \div i_t = t_0 \div t_1$$

$$i_w = i_q \times i_T$$

Пример. Определим индивидуальные индексы цен и физического объема продукции за III и IV кварталы.

Таблица 135 – Исходная информация для расчета индивидуальных индексов цен и физического объема продукции

Товар	Ед. изм.	III квартал		IV квартал	
		Цена за 1 ед., руб.	Количество	Цена за 1 ед., руб.	Количество
Молоко	л	9,8	7500	10,1	6800
Яйца	шт.	26,7	1690	27,0	1830
Картофель	кг	4,85	14750	5,2	10050

При определении по данным таблицы статистических индексов Ш квартал принимается за базисный период, в котором цена единицы товара обозначается p_0, а количество – q_0. IV квартал принимается за текущий или отчетный период, в котором цена единицы товара обозначается p_1, количество – q_1.

Тогда индивидуальные индексы составят:

1) цен

по молоку: $i_p^м = 10{,}1 / 9{,}8 = 1{,}0306$, или $103{,}06\ \%$

по яйцам: $i_p^я = 27{,}0 / 26{,}7 = 1{,}0112$, или $101{,}12\ \%$

по картофелю: $i_p^к = 5{,}2 / 4{,}85 = 1{,}0722$, или $107{,}22\ \%$

В IV квартале цены увеличились по всем видам представленной продукции: на молоко – на 3,06 %, на яйца – на 1,12 %, на картофель – на 7,22 %;

2) физического объема

по молоку: $i_q^м = 6800 / 7500 = 0{,}9067$ или $90{,}67\ \%$

по яйцам: $i_q^я = 1830 / 1690 = 1{,}0828$ или $108{,}28\ \%$

по картофелю:

$i_q^к = 10050 / 14750 = 0{,}6814$ или $68{,}14\ \%$

Количество проданного молока в IV квартале снизилось по сравнению с III на 9,33 % (90,67 % – 100 %), объем продажи яиц повысился на 8,28 %, а объем реализации картофеля сократился на 31,86 %;

Определим взаимосвязь индексов:

по молоку: $i^м{}_{pq} = i_p \times i_q = 1{,}0306 \times 0{,}9067 = 0{,}9344$

по яйцам: $i^я{}_{pq} = i_p \times i_q = 1{,}0112 \times 1{,}0828 = 1{,}0949$

по картофелю: $i^к{}_{pq} = i_p \times i_q = 1{,}0722 \times 0{,}6814 = 0{,}7306$

В IV квартале по сравнению с III выручка от реализации молока снизилась на 6,56 %, яиц – выросла на 9,5 %, картофеля – снизилась на 26,94 %.

Общий индекс (*I*) характеризует обобщающие результаты совместного изменения всех единиц, образующих статистическую совокупность. Общие индексы могут быть представлены в трех формах: агрегатной, средней арифметической, средней гармонической.

Исходной формой сводного индекса является *агрегатная форма*.

Основные функции агрегатных индексов:

– синтетическая (в индексе обобщаются (агрегируются) непосредственно несоизмеримые явления);

– аналитическая (посредством индексного метода измеряется влияние отдельных факторов на совокупное изменение изучаемого показателя).

Основными элементами агрегатного индекса выступают индексируемая величина и вес индекса.

Индексируемая величина – признак, изменение которого характеризует индекс (цена товаров, затраты рабочего времени на производство продукции, количество проданных товаров и т.д.).

Вес индекса – величина, тесно связанная с индексируемой величиной и служащая для целей соизмерения индексируемых величин. При выборе веса индекса необходимо учитывать: если строится индекс количественного показателя, то веса берутся за базисный период; при построении индекса качественного показателя – веса отчетного периода.

При расчете агрегатного индекса для разнородной совокупности находят такой общий показатель, в котором можно объединить все ее элементы.

Представим основные формулы для расчета общих индексов в таблице 136.

Пример. Продажа винограда на рынках города характеризуется нижеследующими данными (данные условные) (таблица 137).

Таблица 137 − Исходная информация для расчета агрегатных индексов

Рынок	2009 г.		2010 г.		Расчетные графы, тыс. руб.		
	Продано, тыс. кг	Цена за 1 кг., руб.	Продано, тыс. кг	Цена за 1 кг., руб.	p_0q_0	p_1q_1	p_0q_1
Центральный	19,3	45	18,9	49	868,5	926,1	850,5
Северный	18,5	50	17,2	53	925,0	911,6	860,0
Итого	37,8	х	36,1	х	1793,5	1837,7	1710,5

Определите:

Таблица 136 – Основные формулы исчисления сводных, или общих, индексов

Наименование	Формула расчета	Интерпретация индекса	Интерпретация значения индекса, уменьшенное на 100 % (I-100)	Интерпретация разности числителя и знаменателя
А	1	2	3	4
Индекс физического объема продукции	$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$	Во сколько раз изменилась стоимость продукции в результате изменения ее объема	На сколько процентов изменилась стоимость продукции в результате изменения ее объема	На сколько рублей изменилась стоимость продукции в результате роста (уменьшения) ее объема
Индекс цен Пааше (по отчетным весам)	$I^{П}{}_{p} = \frac{\sum p_1 q_1}{\sum p_0 q_1}$	Во сколько раз изменилась стоимость продукции, произведенной в отчетном периоде по сравнению с базисным периодом в результате изменения цен	На сколько процентов изменилась стоимость продукции, произведенной в отчетном периоде по сравнению с базисным периодом в результате изменения цен	На сколько рублей изменилась стоимость продукции в результате роста (уменьшения) цен или на сколько рублей продукция в отчетном периоде стала дороже (дешевле), чем в базисном периоде
Индекс цен Ласпейреса (по базисным весам)	$I^{Л}{}_{p} = \frac{\sum p_1 q_0}{\sum p_0 q_0}$	Во сколько раз изменилась стоимость продукции, произведенной в базисном периоде в результате изменения цен	На сколько процентов изменилась стоимость продукции, произведенной в базисном периоде в результате изменения цен	На сколько рублей продукция в базисном периоде стала дороже (дешевле) из-за изменения цен на нее в отчетном периоде

Продолжение таблицы 136

A	1	2	3	4
Индекс стоимости продукции (товарооборота)	$I_{pq}=\frac{\sum p_1q_1}{\sum p_0q_0}$	Во сколько раз возросла (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным	На сколько процентов возросла (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным	На сколько рублей увеличилась (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным
Индекс физического объема	$I_q=\frac{\sum q_1z_0}{\sum q_0z_0}$	Во сколько раз изменились издержки производства продукции в результате изменения объема ее производства	На сколько процентов изменились издержки производства продукции в результате изменения объема ее производства	На сколько рублей изменились издержки производства продукции в результате роста (уменьшения) объема ее производства
Индекс себестоимости продукции	$I_z=\frac{\sum z_1q_1}{\sum z_0q_1}$	Во сколько раз изменились издержки производства продукции в результате изменения себестоимости продукции	На сколько процентов изменились издержки производства продукции в результате изменения ее себестоимости	На сколько рублей изменились издержки производства в результате роста (уменьшения) себестоимости продукции
Индекс издержек производства	$I_z=\frac{\sum z_1q_1}{\sum z_0q_0}$	Во сколько раз возросли (уменьшились) издержки производства продукции, или сколько процентов составил рост (снижение) издержек производства продукции в текущем периоде по сравнению с базисным	На сколько процентов возросли (уменьшились) издержки производства продукции в текущем периоде по сравнению с базисным	На сколько рублей увеличились (уменьшились) издержки производства продукции в текущем периоде по сравнению с базисным

Окончание таблицы 136

А	1	2	3	4
Индекс физического объема продукции	$I_q = \dfrac{\sum q_1 t_0}{\sum q_0 t_0}$	Во сколько раз изменились затраты времени на производство продукции в результате изменения объема ее производства	На сколько процентов изменились затраты времени на производство продукции в результате изменения объема ее производства	На сколько человеко-часов возросли (уменьшились) затраты времени на производство продукции в результате роста (уменьшения) объема производства продукции
Индекс производительности труда по трудовым затратам	$I_t = \dfrac{\sum t_0 q_1}{\sum t_1 q_1}$	Во сколько раз увеличилась (уменьшилась) производительность труда	На сколько процентов изменилась производительность труда в текущем периоде по сравнению с базисным	Абсолютный размер экономии (перерасхода) затрат живого труда в связи с ростом (уменьшением) его производительности
Индекс затрат времени на производство продукции	$I_t = \dfrac{\sum t_1 q_1}{\sum t_0 q_0}$	Во сколько раз изменились затраты времени на производство продукции	На сколько процентов изменились затраты времени на производство продукции в текущем периоде по сравнению с базисным	На сколько человеко-часов увеличились (уменьшились) затраты на производство продукции в текущем периоде по сравнению с базисным

1) агрегатные индексы цен, физического объема и стоимости;

2) абсолютный прирост стоимости продажи винограда, вследствие изменения цен и физического объема продажи винограда;

3) индексы цен Паше и Ласпейреса.

Решение.

1. Исчислим индексы:

цен

$$I^{П}{}_{p} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{926{,}1 + 911{,}6}{850{,}5 + 860{,}0} = \frac{1837{,}7}{1710{,}5} = 1{,}074 \text{ или } 107{,}4\ \%$$

физического объема

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{850{,}5 + 860{,}0}{868{,}5 + 925{,}0} = \frac{1710{,}5}{1793{,}5} = 0{,}954 \text{ или } 95{,}4\ \%$$

стоимости (товарооборота)

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{926{,}1 + 911{,}6}{868{,}5 + 925{,}0} = \frac{1837{,}7}{1793{,}5} = 1{,}025 \text{ или } 102{,}5\ \%$$

Объем продаж винограда (товарооборот) на рынках города в 2010 г. по сравнению с 2009 г. вырос на 2,5 %, в том числе вследствие роста цен на виноград (на 7,4 %) и сокращения количества проданного винограда (на −4,6 %).

Установим взаимосвязь индексов:

$$I_{pq} = I_p \times I_q = 1{,}074 \times 0{,}954 = 1{,}025$$

2. Рассчитаем абсолютное изменение каждого фактора:

цен

$$\sum p_1 q_1 - \sum p_0 q_1 = 1837{,}7 - 1710{,}5 = 127{,}2 \text{ тыс. руб.,}$$

т.е. изменение цены привело к увеличению расходов покупателей на 127,2 тыс. руб.;

физического объема

$$\sum p_0q_1 - \sum p_0q_0 = 1710{,}5 - 1793{,}5 = -83 \text{ тыс. руб.,}$$

т.е. сокращение физического объема реализации привело к экономии расходов покупателей на 83 тыс. руб.;

стоимости (товарооборота)

$$\sum p_1q_1 - \sum p_0q_0 = 1837{,}7 - 1793{,}5 = 44{,}2 \text{ тыс. руб.,}$$

т.е. в целом товарооборот вырос на 44,2 тыс. руб.

Взаимосвязь приростов

$$(\sum p_1q_1 - \sum p_0q_1) + (\sum p_0q_1 - \sum p_0q_0) = 127{,}2 + (-83) = 44{,}2 \text{ тыс. руб.}$$

Простые индексы не дают полной картины изменения параметров, так как предполагается, что все составляющие индексируемой величины имеют равное влияние на общий результат, поэтому применяют средние взвешенные индексы.

*Средний инд*екс – это индекс, вычисленный как средняя величина из индивидуальных индексов.

При исчислении средних индексов используют две формы:

средние арифметическиеиндексы – $\bar{I}_{\text{ариф}} = \dfrac{\sum if}{\sum f}$;

средние гармонические индексы – $\bar{I}_{\text{гарм}} = \dfrac{\sum M}{\sum \frac{M}{i}}$,

где i – индивидуальные индексы изучаемого показателя (индексируемой величины);

M и f – веса соответственно в среднем гармоническом и среднем арифметическом индексах.

Веса для среднего арифметического и среднего гармонического индексов определяются исходя из тождества их агрегатному, который является основной формой общего индекса. При этом для каждого конкретного индекса веса особые.

При исчислении индекса физического объема ($I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$) можно использовать среднюю арифметическую форму, если в числителе производится замена $q_1 = i_q q_0$, тогда индекс примет вид

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

Аналогично, выражая продукцию базисного периода как $q_0 = q_1 \div i_q$, осуществляем замену в знаменателе агрегатного индекса физического объема. В результате получаем общий индекс физического объема в форме среднего гармонического индекса

$$I_q = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{i_q}}.$$

Весами индивидуальных индексов объема в среднем арифметическом индексе служит стоимость продукции базисного периода в базисных ценах $q_0 p_0$, а в среднем гармоническом индексе – стоимость продукции отчетного периода в базисных ценах $q_1 p_0$.

Если исходить из агрегатной формулы индекса цен $I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$, то тождественной ему средний арифметический индекс

$$\bar{I}_p = \frac{\sum i_p p_0 q_1}{\sum p_0 q_1}.$$

В этом случае весами для индивидуальных индексов должна служить стоимость продукции отчетного периода в базисных ценах p_0q_1, что не совсем удобно для расчетов на практике.

Средний гармонический индекс цен, исходя из $p_0 = p_1 \div i_p$, будет иметь вид:

$$\bar{I}_p = \frac{\sum p_1q_1}{\sum \frac{p_1q_1}{i_p}}.$$

В данной формуле весами служит стоимость продукции текущего периода в текущих ценах p_1q_1. Данная формула более предпочтительна для практического использования.

При решении конкретных задач выбор той или иной формы среднего индекса определяется прежде всего тем, какие исходные данные имеются в распоряжении.

Пример. В таблице представлены экономические показатели работы магазина «Детский мир» за два периода (данные условные).

Таблица 138 − Экономические показатели работы магазина «Детский мир»

Отдел	Стоимость товаров, млн руб.		Индивидуальный индекс	
	2009 г. (p_0q_0)	2010 г. (p_1q_1)	цены (i_p)	физического объема (i_q)
Обувь	236	250	1,05	0,975
Одежда	547	535	1,03	1,08
Игрушки	720	736	1,00	1,12
Итого	1503	1521	х	х

Необходимо определить средние индексы цен и физического объема.

1. Определим средний индекс цен:

$$\bar{I}_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}} = \frac{250 + 535 + 736}{\frac{250}{1{,}05} + \frac{535}{1{,}03} + \frac{736}{1{,}00}} = \frac{1521}{1493{,}51} = 1{,}018$$ или 101,8 %,

т.е. в среднем цены на товары в магазине выросли на 1,8 %.

2. Рассчитаем средний индекс физического объема:

$$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0} = \frac{236 \times 0{,}975 + 547 \times 1{,}08 + 720 \times 1{,}12}{236 + 547 + 720} = \frac{1627{,}26}{1503} = 1{,}083$$

или 108,3 %, т.е. в среднем объем реализованных товаров по отделам магазина вырос на 8,3 %.

Данный пример рассматривался на основе исходной формулы Пааше.

В случае если за исходную принимается формула Ласпейреса ($I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$), заменяя в числителе p_1 на $i_p p_0$ ($i_p = \frac{p_1}{p_0}$), получим средний арифметический индекс цен

$$I_p^{Л} = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0},$$

т.е. весами служит стоимость отдельных групп продукции базисного периода $p_0 q_0$.

Средний гармонический индекс цен Ласпейреса выражается следующей формулой

$$\bar{I}^{Л}{}_p = \frac{\sum p_1 q_0}{\sum \frac{p_1 q_0}{i_p}}.$$

Вместо абсолютных данных о стоимости отдельных изделий в базисном периоде можно принимать их долю (удельный вес) в общей стоимости

$$d_0 = \frac{q_0 p_0}{\sum q_0 p_0}.$$

Тогда формула среднего арифметического индекса из индивидуальных будет иметь вид $\bar{I}_q = \sum i_q d_0$, поскольку $\sum d_0 = 1$.

Если d_0 выражено в процентах, формула среднего арифметического индекса

$$\bar{I}_q = \frac{\sum i_q d_0}{100\%}.$$

При изучении динамики качественных показателей (себестоимости, цены, производительности труда и т.д.) приходится определять изменение средней величины индексируемого показателя, которое обусловлено взаимодействием двух факторов – изменением значения индексируемого показателя у отдельных групп единиц и изменением структуры явления. Под изменением структуры явления понимается изменение доли отдельных групп единиц совокупности в общей их численности.

Эта задача решается с помощью индексного метода, т.е. путем построения системы взаимосвязанных индексов, в которую включаются три индекса: переменного состава, постоянного состава и структурных сдвигов.

Индекс переменного состава отражает соотношение средних уровней изучаемого явления, относящихся к разным периодам времени. Индекс характеризует изменение среднего уровня признака за счет влияния двух факторов. Индекс переменного состава рассчитывается по формуле

$$I_{\text{перем}} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} = I_{\text{пост}} \times I_{\text{стр}}.$$

Этот индекс разлагается на два индекса-сомножителя: индекс постоянного (фиксированного) состава и индекс структурных сдвигов.

Индекс постоянного (фиксированного) состава – это индекс, исчисленный с весами, зафиксированными на уровне одного какого-либо периода, и показывающий изменение только индексируемой величины. Индекс фиксированного состава определяется как агрегатный индекс

$$I_{\text{пост}} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_1}{\sum f_1} = \frac{\sum x_1 f_1}{\sum x_0 f_1}.$$

Индекс структурных сдвигов показывает структурные изменения, а именно изменение доли отдельных единиц совокупности в общей их численности

$$I_{\text{стр}} = \frac{\sum x_0 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0}.$$

Если в индексах средних уровней в качестве весов используются удельные веса единиц совокупности в общей численности совокупности (показатели доли $d = f \div \sum f$), то система индексов может быть записана в следующем виде

$$I_{\text{перем}} = \frac{\sum x_1 d_1}{\sum x_0 d_0}; \qquad I_{\text{пост}} = \frac{\sum x_1 d_1}{\sum x_0 d_1}; \qquad I_{\text{стр}} = \frac{\sum x_0 d_1}{\sum x_0 d_0}$$

Абсолютный прирост (уменьшение) среднего уровня признака в целом по совокупности ($\Delta\bar{x}$) находят как разность числителя и знаменателя индекса переменного состава

$$\Delta\bar{x} = \bar{x}_1 - \bar{x}_0 = \frac{\sum x_1 f_1}{\sum f_1} - \frac{\sum x_0 f_0}{\sum f_0} \text{ или } \Delta\bar{x} = \sum x_1 d_1 - \sum x_0 d_0$$

Абсолютный прирост (уменьшение) среднего уровня признака в целом по совокупности за счет изменения значений изучаемого признака у отдельных единиц совокупности и за счет структурных изменений рассчитывается соответственно как разность числителей и знаменателей индексов постоянного состава и структурных сдвигов:

а) за счет изменения значений изучаемого признака у отдельных единиц совокупности

$$\Delta\bar{x}_{(x)} = \frac{\sum x_1 f_1}{\sum f_1} - \frac{\sum x_0 f_1}{\sum f_1} \text{ или } \Delta\bar{x} = \sum x_1 d_1 - \sum x_0 d_1$$

б) за счет структурных изменений

$$\Delta\bar{x}_{(f)} = \frac{\sum x_1 f_1}{\sum f_1} - \frac{\sum x_1 f_0}{\sum f_0} \text{ или } \Delta\bar{x} = \sum x_1 d_1 - \sum x_1 d_0 .$$

В общем виде: $\Delta\bar{x} = \Delta_{\bar{x}_{(x)}} + \Delta_{\bar{x}_{(f)}}$

Пример. По данным таблицы 139 определите индексы средних цен переменного состава, постоянного состава и структурных сдвигов.

Таблица 139 − Исходная информация для расчета (данные условные)

Рынок	2009 г.		2010 г.		Расчетные графы, тыс.руб.		
	Продано, тыс. кг	Цена за 1 кг, руб.	Продано, тыс. кг	Цена за 1 кг, руб.	p_0q_0	p_1q_1	p_0q_1
Центральный	19,3	45	18,9	49	868,5	926,1	850,5
Северный	18,5	50	17,2	53	925,0	911,6	860,0
Итого	37,8	х	36,1	х	1793,5	1837,7	1710,5

Решение

1. Исчислим индекс товарооборота винограда переменного состава на рынках города:

$$I_{\text{р}}^{\text{перем}} = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} = \frac{1837{,}7}{36{,}1} : \frac{1793{,}5}{37{,}8} = 50{,}91 : 47{,}45 = 1{,}0729$$

или 107,29 %, т.е. цена винограда на каждом рынке в среднем в текущем периоде по сравнению с предыдущим периодом выросла на 7,29 %

2. Рассчитаем индекс средних цен постоянного состава:

$$I_{\text{стр}} = \frac{\sum p_0 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} = \frac{1710{,}5}{36{,}1} : \frac{1793{,}5}{37{,}8} = 47{,}38 : 47{,}45 = 0{,}9985$$

или 99,85 %.

Первая часть этого выражения позволяет ответить на вопрос, какой была бы средняя цена в текущем периоде, если бы цены на каждом рынке сохранились на уровне предыдущего периода. Вторая часть отражает фактическую среднюю цену предыдущего периода. В целом по полученному значению индекса можно сделать вывод, что за счет структурных сдвигов цены снизились на 0,15 %.

3. Найдем индекс структурных сдвигов:

$$I_{\text{р}}^{\text{пост}} = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_1}{\sum q_1} = \frac{1837{,}7}{36{,}1} : \frac{1710{,}5}{36{,}1} = 50{,}91 : 4738 = 1{,}0745$$

или 107,45 %, т.е. если бы структура реализации винограда по рынкам не изменилась, средняя цена выросла на 7,45 %. Влияние на среднюю цену первого фактора оказалось сильнее, что отражается в следующей взаимосвязи:

$$I_{\text{р}}^{\text{перем}} = I_{\text{р}}^{\text{пост}} \times I_{\text{стр}} = 1{,}0745 \times 0{,}9985 = 1{,}0729$$

Аналогично построение индексов переменного, постоянного составов и структурных сдвигов для анализа изменения себестоимости, урожайности и пр. (таблица 140).

В ряде случаев для анализа социально-экономических явлений применяют систему индексов.

Таблица 140 − Система взаимосвязанных индексов

Индексируемая величина	Индекс переменного состава	Индекс постоянного состава	Индекс структурных сдвигов
Себестоимость	$\frac{\sum z_1q_1}{\sum q_1}:\frac{\sum z_0q_0}{\sum q_0}$	$\frac{\sum z_1q_1}{\sum z_0q_1}$	$\frac{\sum z_0q_1}{\sum q_1}:\frac{\sum z_0q_0}{\sum q_0}$
Урожайность	$\frac{\sum y_1П_1}{\sum П_1}:\frac{\sum y_0П_0}{\sum П_0}$	$\frac{\sum y_1П_1}{\sum y_0П_1}$	$\frac{\sum y_0П_1}{\sum П_1}:\frac{\sum y_0П_0}{\sum П_0}$
Производительность труда	$\frac{\sum T_0}{\sum q_0}:\frac{\sum T_1}{\sum q_1}$	$\frac{\sum t_0q_1}{\sum t_1q_1}$	$\frac{\sum t_0q_0}{\sum t_0q_1}:\frac{\sum q_0}{\sum q_1}= =\left[\frac{\sum T_0}{\sum q_0}:\frac{\sum T_1}{\sum q_1}\right]:\frac{\sum t_0q_1}{t_1q_1}$

Ряд индексов, каждый из которых рассчитан по отношению к предыдущему периоду называют цепными индексами, а ряд индексов с постоянной базой сравнения – базисными (таблица 141).

Таблица 141 − Цепные и базисные индивидуальные индексы

Наименование индивидуального индекса	Цепные индексы	Базисные индексы
Индекс цен	$\frac{p_1}{p_0};\frac{p_2}{p_1};...;\frac{p_t}{p_{t-1}}$	$\frac{p_1}{p_0};\frac{p_2}{p_0};...;\frac{p_t}{p_0}$
Индекс физического объема	$\frac{q_1}{q_0};\frac{q_2}{q_1};...;\frac{q_t}{q_{t-1}}$	$\frac{q_1}{q_0};\frac{q_2}{q_0};...;\frac{q_t}{q_0}$
Индекс товарооборота (стоимости)	$\frac{p_1q_1}{p_0q_0};\frac{p_2q_2}{p_1q_1};...;\frac{p_tq_t}{p_{t-1}q_{t-1}}$	$\frac{p_1q_1}{p_0q_0};\frac{p_2q_2}{p_0q_0};...;\frac{p_tq_t}{p_0q_0}$

Цепные и базисные индексы могут быть построены и для общих индексов, которые могут иметь постоянные и переменные веса (таблица 142).

Таблица 142 − Цепные и базисные сводные (общие) индексы

Веса	Цепные индексы	Базисные индексы
1	2	3
Индекс цен		
Постоянные	$I_{p\,1/0}=\frac{\sum p_1q_0}{\sum p_0q_0};I_{p\,2/1}=\frac{\sum p_2q_0}{\sum p_1q_0};$ $I_{p\,3/2}=\frac{\sum p_3q_0}{\sum p_2q_0};...;I_{p\,t/t-1}=\frac{\sum p_tq_0}{\sum p_{t-1}q_0}$	$I_{p\,1/0}=\frac{\sum p_1q_0}{\sum p_0q_0};I_{p\,2/0}=\frac{\sum p_2q_0}{\sum p_0q_0};$ $I_{p\,3/0}=\frac{\sum p_3q_0}{\sum p_0q_0};...;I_{p\,t/0}=\frac{\sum p_tq_0}{\sum p_0q_0}$
Переменные	$I_{p\,1/0}=\frac{\sum p_1q_1}{\sum p_0q_1};I_{p\,2/1}=\frac{\sum p_2q_2}{\sum p_1q_2};$ $I_{p\,3/2}=\frac{\sum p_3q_3}{\sum p_2q_3};...;I_{p\,t/t-1}=\frac{\sum p_tq_t}{\sum p_{t-1}q_t}$	$I_{p\,1/0}=\frac{\sum p_1q_1}{\sum p_0q_1};I_{p\,2/0}=\frac{\sum p_2q_2}{\sum p_0q_2};$ $I_{p\,3/0}=\frac{\sum p_3q_3}{\sum p_0q_3};...;I_{p\,t/0}=\frac{\sum p_tq_t}{\sum p_0q_t}$
Индекс физического объема		
Постоянные	$I_{q\,1/0}=\frac{\sum q_1p_0}{\sum q_0p_0};I_{q\,2/1}=\frac{\sum q_2p_0}{\sum q_1p_0};$ $I_{q\,3/2}=\frac{\sum q_3p_0}{\sum q_2p_0};...;I_{q\,t/t-1}=\frac{\sum q_tp_0}{\sum q_{t-1}p_0}$	$I_{q\,1/0}=\frac{\sum q_1p_0}{\sum q_0p_0};I_{q\,2/0}=\frac{\sum q_2p_0}{\sum q_0p_0};$ $I_{q\,3/0}=\frac{\sum q_3p_0}{\sum q_0p_0};...;I_{q\,t/0}=\frac{\sum q_tp_0}{\sum q_0p_0}$
Переменные	$I_{q\,1/0}=\frac{\sum q_1p_1}{\sum q_0p_1};I_{q\,2/1}=\frac{\sum q_2p_2}{\sum q_1p_2};$ $I_{q\,3/2}=\frac{\sum q_3p_3}{\sum q_2p_3};...;I_{q\,t/t-1}=\frac{\sum q_tp_t}{\sum q_{t-1}p_t}$	$I_{q\,1/0}=\frac{\sum q_1p_1}{\sum q_0p_1};I_{q\,2/0}=\frac{\sum q_2p_2}{\sum q_0p_2};$ $I_{q\,3/0}=\frac{\sum q_3p_3}{\sum q_0p_3};...;I_{q\,t/0}=\frac{\sum q_tp_t}{\sum q_0p_t}$

Территориальные индексы служат для сравнения показателей в пространстве, т.е. по странам, экономическим районам, областям и т.д.

Построение территориальных индексов определяется выбором базы сравнения и весов или уровня, на котором фиксируются веса. При двусторонних сравнениях каждая территория может быть и сравниваемой (числитель индекса), и базой сравнения (знаменатель). Веса как первой, так и второй территорий в принципе также имеют равные основания использоваться при расчете индекса. Это может привести к различным и даже противоречивым результатам. Исключить подобную неопределенность можно следующими способами:

1) в качестве весов принимаются объемы проданных товаров по двум регионам вместе взятым

$$Q = q_a + q_b$$

Территориальный индекс цен в этом случае рассчитывается по следующей формуле

$$I_{pb/a} = \frac{\sum p_b Q}{\sum p_a Q}.$$

В формуле данного территориального индекса вместо суммарных используются стандартизированные веса (стандартизированная структура). В качестве таких весов может выступать структура продажи данных видов продукции по более крупному территориальному образованию, например, республика. В этом случае индекс имеет вид

$$I_p = \frac{\sum p_a q_{\text{респ}}}{\sum p_b q_{\text{респ}}};$$

2) расчет территориальных индексов учитывает соотношение весов сравниваемых территорий. При этом способе первый

шаг заключается в расчете средней цены каждого товара по двум территориям, вместе взятым

$$\bar{p}_i = \frac{\sum p_i q_i}{\sum q_i}.$$

После этого непосредственно рассчитывается территориальный индекс

$$I_{pb/a} = \frac{\sum p_b q_b}{\sum \bar{p} q_b} : \frac{\sum p_a q_a}{\sum \bar{p} q_a}.$$

Данный подход к расчету территориального индекса обеспечивает известную взаимосвязь

$$I_p \times I_q = I_{pq}.$$

Индекс физического объема реализации при этом строится следующим образом

$$I_{qb/a} = \frac{\sum q_b \bar{p}}{\sum q_a \bar{p}}.$$

Аналогично строятся индексы для сравнения цен территории А с ценами территории Б.

Задача 108

В таблице 143 представлены данные о ценах на молоко и объемах его производства в Пензенской области за 2006-2010 гг.

При условии 100%-й реализации молока в каждом году определите цепные и базисные индивидуальные индексы цен, физического объема реализации и товарооборота. Проверьте взаимосвязь цепных и базисных индексов.

Таблица 143 − Данные о ценах на молоко и объемах его производства в Пензенской области

Год	Цена за 1 т, руб.	Произведено, тыс. т
2006	5621	514,1
2007	6945	537,1
2008	7924	549,2
2009	7912	511,2
2010	11384	466,6

Задача 109

В таблице 144 представлены данные о трудоемкости продукции предприятия и объемах ее производства (данные условные).

Таблица 144 – Объемы производства и трудоемкость производства продукции на предприятии

Вид продукции	2009 г.		2010 г.	
	Произведено, тыс. шт.	Затраты на 100 изделий, чел.-ч	Произведено, тыс. шт.	Затраты на 100 изделий, чел.-ч
А	275	75	291	72
Б	163	119	174	115

Рассчитайте индексы производительности труда, физического объема продукции, затрат труда. Рассчитайте абсолютную сумму изменения факторов и проанализируйте её.

Задача 110

По данным таблицы 145 вычислите:

1) агрегатный индекс себестоимости и абсолютную сумму изменения затрат за счет изменения себестоимости;

2) агрегатный индекс физического объема продукции и абсолютную сумму изменения затрат за счет этого фактора;

3) агрегатный индекс затрат и абсолютную сумму изменения затрат за счет двух факторов – изменения себестоимости и физического объема продукции.

Таблица 145 – Себестоимость и производство продукции в ОАО ПТФ «Васильевская»

Вид продукции	Количество продукции, ц		Себестоимость 1 ц, руб.	
	2009 г.	2010 г.	2009 г.	2010 г.
Зерно	28874	11245	224,18	471,68
Молоко	15734	1369	1159,27	1651,57
Прирост живой массы крупного рогатого скота	1175	146	8403,40	7582,19
Прирост живой массы птицы	647865	644332	2622,87	2766,68

Задача 111

Вычислите общие индексы цен, стоимости проданных товаров и физического объема товарооборота по овощному магазину по данным таблицы 146.

Таблица 146 – Продажа овощей в магазине

Продукция	Товарооборот в ценах соответствующего периода, руб.		Цена 1 кг, руб.	
	март	апрель	март	апрель
Картофель	35679	58311	67,96	88,35
Капуста	3653	4761	28,10	31,74
Морковь	1262,80	2337,44	28,70	41,74

Задача 112

В таблице 147 имеются нижеследующие данные о реализации мясных продуктов на рынке (данные условные).

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

Таблица 147 − Данные о реализации мясных продуктов на рынке города за сентябрь-октябрь

Продукт	Сентябрь		Октябрь	
	цена за 1 кг, руб.	продано, ц	цена за 1 кг, руб.	продано, ц
Говядина	370	26,3	380	24,1
Баранина	260	8,8	260	9,2
Свинина	290	14,5	295	12,3

Задача 113

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 %.

Задача 114

Как изменились общие затраты труда на предприятии, если стоимость продукции в сопоставимых ценах возросла на 12,4 %, а производительность труда (расчет по выработке) повысилась на 3,4 %?

Задача 115

В таблице 148 представлены данные о реализации молочных продуктов на рынке. Рассчитайте сводные индексы цен, товарооборота и физического объема реализации.

Таблица 148 − Данные о реализации молочных продуктов на рынке города (данные условные)

Продукт	Товарооборот, тыс. руб.		Изменение цены в декабре по сравнению с ноябрем, %
	ноябрь	декабрь	
Молоко	97	63	+2,1
Сметана	45	40	+3,5
Творог	129	115	+4,2

Задача 116

Производительность труда (расчет по трудоемкости) на предприятии в текущем периоде по сравнению с базисным выросла на 2,5 %, при этом численность рабочих увеличилась на 18 человек и составила 236 человек. Как изменился физический объем продукции?

Задача 117

По данным таблицы 149 проанализируйте изменение валового сбора за счет изменения урожайности, структуры и размера посевных площадей.

Таблица 149 − Площади посева и урожайность овощных культур в сельскохозяйственной организации (данные условные)

Культура	Посевная площадь, га		Урожайность, ц с 1 га	
	2009 г.	2010 г.	2009 г.	2010 г.
Рожь	300	250	14,3	14,1
Пшеница	350	310	14,1	14,3
Ячмень	400	500	13,2	12,9
Всего	1050	1060	-	-

Задача 118

Цены на потребительские товары и услуги в регионе в январе по сравнению с предшествующим месяцем возросли на 3,4 %, а в феврале по сравнению с январем – на 4,5 %.

Как изменились цены в марте по сравнению с февралем, если:

1) общий рост цен за 1 квартал данного года составил 110,7 %;

2) при расчете всех индексов использовались веса декабря предшествующего года?

Задача 119

Уровень рыночных цен на молочные продукты и объем их реализации в двух городах характеризуются данными, представленными в таблице 150.

Рассчитайте двумя способами территориальный индекс цен города А по отношению к городу Б.

Таблица 150 − Уровень рыночных цен на молочные продукты и объем их реализации в городах А и Б

Продукт	Город А		Город Б	
	цена за 1 кг, руб.	продано, т	цена за 1 кг, руб.	продано, т
Молоко	35	76	32	68
Масло	170	45	176	39
Творог	150	60	155	55
Сыр	290	32	284	41

ЛИТЕРАТУРА

1. Ганченко, О.И. Практикум по общей теории статистики / О.И. Ганченко, М.Р. Ефимова, Е.В. Петрова. − 3-е изд. − М.: Финансы и статистика, 2011. − 368 с.
2. Годин, А.М. Статистика / А.М. Годин. − 10-е изд. − М.: Дашков и К°, 2012. − 452 с.
3. Громыко, Г. Теория статистики / Г. Громыко. − М.: Инфра-М, 2010. − 480 с.
4. Елисеева, И.И. Общая теория статистики / И.И. Елисеева, М.М. Юзбашев; под ред. И.И. Елисеевой. − 5-е изд., перераб. и доп. − М.: Финансы и статистика, 2004. – 656 с.
5. Ендронова, В.Н. Общая теория статистики / В.Н. Ендронова, М.В. Малафеева − М.: Магистр, 2009. − 608 с.
6. Ефимова, М.Р. Общая теория статистики / М.Р. Ефимова, Е.В. Петрова, В.Н. Румянцев. − 2-е изд., перераб. и доп. − М.: Инфра-М, 2011. − 416 с.
7. Зинченко, А.П. Статистика / А.П. Зинченко. – М.: КолосС, 2007. – 362 с.
8. Лысенко, С.Н. Общая теория статистики / С.Н. Лысенко, И.А. Дмитриева. − М.: Форум, 2011. − 208 с.
9. Лялин, В.С. Статистика. Теория и практика в Excel / В.С. Лялин, И.Г. Зверева, Н.Г. Никифорова. − М.: Финансы и статистика, 2010. − 448 с.
10. Минашкин, В.Г. Практикум по теории статистики / В.Г. Минашкин, Р.А. Шмойлова, Н.А. Садовникова; под ред. Р.А. Шмойловой. − 3-е изд., перераб. и доп. − М.: Финансы и статистика, 2011. − 416 с.
11. Назарова, М.Г. Общая теория статистики / М.Г. Назарова. − 2-е изд. − М.: Омега-Л, 2011. − 410 с.
12. Общая теория статистики: статистическая методология в изучении коммерческой деятельности / А.И. Харламов, О.Э. Башина, В.Т. Бабурин и др.; под ред. А.А. Спирина, О.Э. Башиной. – М.: Финансы и статистика, 1995. – 296 с.
13. Практикум по общей теории статистики и сельскохозяйственной статистике / А.П. Зинченко, С.С. Сергеев, И.Д. Политова и др.; под ред. А.П. Зинченко. – 4-е изд., перераб. и доп. – М.: Финансы и статистика, 1988. – 328 с.

14. Статистика. Базовый курс / Под ред. И.И. Елисеевой. − М.: Юрайт, 2012. − 558 с.

15. Статистика / Под ред. С.А. Орехова. − М.: Эксмо, 2010. − 448 с.

16. Теория статистики / Р.А. Шмойлова, В.Г. Минашкин, Н.А. Садовникова и др.; под ред. Р.А. Шмойловой. – 5-е изд. – М.: Финансы и статистика, 2009. – 656 с.

17. Улитина, Е.В. Статистика / Е.В. Улитина, О.В. Леднева, О.Л. Жирнова; под ред. Е.Улитиной. − М.: Московская финансово-промышленная Академия, 2011. − 312 с.

18. Эверитт, Б.С. Большой словарь по статистике / Б.С. Эверитт. − 3-е изд. − М.: Проспект, 2010. − 736 с.

Приложение 1

Интернет-ресурсы, содержащие статистическую информацию и аналитические обзоры

Организация	Адрес сайта
Государственный комитет Российской Федерации по статистике	http://www.gks.ru
Пензенский областной комитет по статистике	http://www.pnz.gks.ru
Информационное агентство AK&M	http://www.akm.ru
Центральный банк РФ	http://www.cbr.ru
Московская межбанковская валютная биржа	http://www.micex.ru
Федеральная комиссия по рынку ценных бумаг	http://www.fedcom.ru
Министерство финансов России	http://www.minfin.ru
Федеральная служба по налогам и сборам	http://www.nalog.ru
Фондовая биржа РТС	http://www.rtsnet.ru
Московский городской комитет государственной статистики	http://www.mosstat.ru
Экономика и жизнь: агентство консультаций и деловой информации	http://www.akdi.ru
Росбизнесконсалтинг Информационные системы	http://www.rbc.ru

Приложение 2

Основные показатели производства и реализации зерна
в Иссинском, Мокшанском и Пензенском районах
Пензенской области в 2009 году

	Посевная площадь, га	Валовой сбор в весе после доработки, ц	Сбор с 1 га, ц	Количество реализованной продукции, ц	Выручено, руб.
1	2	3	4	5	6
Иссинский					
1. СПК «Салмовка»	1200	43736	36,4	45429	17525780
2. ООО «Русь»	232	3806	16,4	1990	662051
3. СПК РАО «Иссинское»	3003	53005	17,7	67778	16258153
4. СПК «Победа»	690	12152	17,6	6796	1928262
5. Иссинский филиал ООО РАО «Пензенская зерновая компания»	9307	216881	23,3	151122	44477505
6. ООО «Пенза Золотая Нива»	6997	154541	22,1	92997	22464918
7. СХПК РАО «Булычевское»	270	5108	18,9	480	127617
Мокшанский					
8. ФГУП УЧХОЗ «Рамзай ПГСХА»	1719	49544	28,8	7092	5158049
9. ТнВ «Пугачевское»	2936	121491	41,4	64115	23385620
10. ЗАО АПК «Нечаевский»	5353	121675	22,7	54013	23037092
11. ООО «Луговое»	1723	34566	20,1	23697	5372825
12. ОАО «Сервис»	4650	133940	28,8	123305	48858319
13. ООО «Колос»	1813	21661	11,9	25687	7335001
14. ООО «Шукша»	770	12547	16,3	11629	2863334
15. ООО «Агросервис»	1456	43304	29,7	20719	11100202
16. ООО «Труженик»	3100	99773	32,2	57294	20473930
17. ООО «Агрофирма Биокор-С»	1050	20057	19,1	5491	5773772
18. ООО «Полевое»	1300	32956	25,4	21097	5293417
19. ООО «Керская ферма»	391	9416	24,1	7036	1407200
20. ООО «Ресурс»	380	5815	15,3	2938	919086
21. ООО «Контактагро»	1204	17195	14,3	12970	2624200
22. ООО «АПК Лидер»	1140	16549	14,5	10066	3340590
23. ООО «Агро-Лайн Пенза»	2590	64992	25,1	39175	9012872

Окончание приложения 2

1	2	3	4	5	6
24. Мокшанский филиал ООО РАО «Пензенская зерновая компания»	4177	62217	14,9	53722	15349774
25. ООО «Фаэтон-Агропенза»	3366	63630	18,9	86441	26069087
26. ООО «Раздолье»	480	7650	15,9	7650	1506750
27. ООО НПП «Иннаучагро-центр»	378	7680	20,3	1570	1709000
Пензенский					
28. СПК «Широкополье»	2238	51917	23,2	34489	7556666
29. ОАО «Тепличный»	1957	30120	15,4	21955	6659059
30. СПК «Краснополье»	2041	50422	24,7	37882	9779749
31. ОАО «Племзавод Еланский»	3691	87712	23,8	35275	12314158
32. ЗАО «Истоки Хопра»	1855	31165	16,8	6727	2338184
33. ООО РАО «Кондольское»	1959	23911	12,2	29913	7007179
34. ООО «Агрокомплекс Пензенский»	1352	14453	10,7	10789	2664937
35. ЗАО «Петровский хлеб»	9333	276895	29,7	236874	108094341
36. ЗАО «Константиново»	4230	107905	25,5	18436	8712316
37. ООО «Агрофирма Богословская»	150	3174	21,1	400	80000
38. ООО «Организатор»	2503	24864	9,9	17628	5711002
39. ООО «Нива»	2310	66785	28,9	39708	11662536
40. ООО «Сура-Агро»	1510	46314	30,8	52989	16454949
41. Пензенский филиал ООО РАО «Пензенская зерновая компания»	7064	137298	19,4	125410	36873170
42. Кондольский филиал ООО РАО «Пензенская зерновая компания»	3982	60439	15,2	80847	24190134
43. ООО «Биотехноформ»	650	350	0,5	350	100000
44. ООО «Марьино»	3497	89575	25,6	36982	10162172

Приложение 3

Значения плотности $f(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$ вероятности для нормального закона распределения

$$f(-t) = f(t)$$

Целые и деся-тые до-ли *t*	Сотые доли *t*									
	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3633	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3525	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1972	1849	1826	1804	1781	1758	1736

Целые и десятые доли *t*	Сотые доли *t*									
	0	1	2	3	4	5	6	7	8	9
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0762	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046

Целые и десятые доли *t*	Сотые доли *t*									
	0	1	2	3	4	5	6	7	8	9
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
4,0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
4,1	0,0001338	–	–	–	–	–	–	–	–	–
4,5	0,0000160	–	–	–	–	–	–	–	–	–
5,0	0,0000015	–	–	–	–	–	–	–	–	–

Приложение 4

Распределение Пирсона (χ^2-распределение) (значение $\chi^2_{табл}$ для вероятностей P($\chi^2 \succ \chi^2_{табл}$))

v	Вероятность										
	0,999	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
1	0,05157	0,04393	0,03157	0,03628	0,03982	0,00393	0,0158	0,0642	0,102	0,148	0,455
2	0,00200	0,0100	0,0201	0,0404	0,0506	0,103	0,211	0,446	0,575	0,713	1,386
3	0,0243	0,0717	0,115	0,185	0,216	0,352	0,584	1,005	1,213	1,424	2,366
4	0,0908	0,207	0,297	0,429	0,484	0,711	1,064	1,649	1,923	2,195	3,357
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	2,675	3,000	4,351
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	3,455	3,828	5,348
7	0,598	1,989	1,239	1,564	1,690	2,167	2,833	3,822	4,255	4,671	6,346
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	5,071	5,527	7,344
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	5,899	6,393	8,343
10	1,479	2,156	2,558	3,059	3,247	3,240	4,865	6,179	6,787	7,267	9,342
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	7,584	8,148	10,341
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	8,438	9,034	11,340
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	9,299	9,926	12,340
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	10,165	10,821	13,339
15	3,483	4,601	5,229	5,985	6,262	7,261	8,847	10,307	11,036	11,721	14,339

v	Вероятность										
	0,999	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	11,912	12,624	15,338
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	12,892	13,531	16,338
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	13,675	14,440	17,338
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352	18,338
20	5,921	7,434	8,260	9,237	9,591	10,871	12,443	14,578	15,452	16,266	19,337
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	16,344	17,182	20,337
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	17,240	18,101	21,337
23	7,529	9,260	10,196	11,293	11,688	13,091	14,848	17,187	18,137	19,021	22,337
24	8,035	9,886	10,856	11,992	12,401	13,848	15,659	18,062	19,037	19,943	23,337
25	8,649	10,520	11,524	12,697	13,120	14,611	16,173	18,940	19,939	20,887	24,337
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	20,848	21,792	25,336
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	21,749	22,719	26,136
28	10,391	12,461	13,565	14,547	15,308	16,928	18,937	21,588	22,657	23,617	27,386
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	23,567	24,577	28,336
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	24,478	25,508	29,336

v	Вероятность									
	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
1	1,074	1,323	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	2,408	2,773	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	3,665	4,108	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,268
4	4,878	5,385	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,465
5	6,064	6,626	7,289	9,236	11,070	12,839	13,388	15,086	16,750	20,517
6	7,231	7,841	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	8,383	9,037	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	9,524	10,219	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,125
9	10,656	11,389	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	11,781	12,549	13,412	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	12,899	13,701	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	14,011	14,845	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	15,119	15,984	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	16,222	17,117	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	17,322	18,245	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697

v	Вероятность									
	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
16	18,418	19,369	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	19,511	20,489	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	20,601	21,605	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	21,689	22,718	23,900	27,204	30,144	32,852	33,687	38,191	38,582	43,820
20	22,775	23,628	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	23,858	24,935	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	24,939	26,039	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	26,018	27,141	28,429	32,567	35,172	38,076	38,968	41,638	44,181	49,728
24	27,096	28,241	29,553	33,193	36,415	39,384	40,270	42,980	45,558	51,170
25	28,172	29,339	30,675	34,362	37,652	40,046	41,566	44,314	46,928	52,620
26	29,246	30,434	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	30,319	31,328	32,912	36,741	40,113	43,194	44,140	46,963	49,645	55,476
28	31,391	32,320	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,893
29	32,461	33,711	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,302
30	33,530	34,800	36,250	40,256	43,773	46,979	47,962	50,692	53,672	59,703

Приложение 5

Нормальный закон распределения $\Phi(t) = P(|T| \le t_{\text{табл}})$

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	0797	0876	0955	1034	1113	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2960	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7071	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7994	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882

Целые и десятые доли *t*	Сотые доли *t*									
	0	1	2	3	4	5	6	7	8	9
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9090
1,7	9109	9127	9146	9164	9181	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9426	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9756	9762	9768	9774	9780
2,3	9786	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9841	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9910	9912	9915	9917	9920	9922	9924	9926	9928
2,7	9931	9933	9935	9937	9938	9940	9942	9944	9946	9947
2,8	9949	9951	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	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	9981	9981	9982	9983	9983	9984	9984	9985	9985	9986
3,5	9995	9996	9996	9996	9996	9996	9996	996	9997	9997
3,6	9997	9997	9997	9997	9997	9997	9997	9998	9998	9998

Целые и десятые доли *t*	Сотые доли *t*									
	0	1	2	3	4	5	6	7	8	9
3,7	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3,8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3,9	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
4,0	0,999936	9999	9999	9999	9999	9999	9999	9999	9999	9999
4,5	0,999994	-	-	-	-	-	-	-	-	-
5,0	0,99999994	-	-	-	-	-	-	-	-	-

Приложение 6

Распределение Стьюдента (t-распределение)

ν	Вероятность $\alpha = St(t) = P(\lvert T \rvert \rangle t_{\text{табл}})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,563	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,406	0,559	0,727	0,920	1,156	1,46	2,015	2,71	3,365	4,043	6,859
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,327	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,583
11	0,129	0,260	0,396	0,543	0,697	0,976	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,539	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,888	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965

ν	Вероятность $\alpha = St(t) = P(\|T\|\rangle t_{\text{табл}})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,326	1,789	2,093	2,539	2,861	3,833
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,066	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,868	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,402	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,50	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Приложение 7

Распределение Фишера-Снедекора (F-распределение)

Значение $F_{табл}$ удовлетворяющие условию $P\left(F \geq F_{табл}\right)$. Первое значение соответствует вероятности 0,05; второе − вероятности 0,01 и третье – вероятности 0,001;

ν_1 – число степеней свободы числителя; ν_2 − знаменателя

$\frac{\nu_1}{\nu_2}$	1	2	3	4	5	6	8	12	24	∞	t
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	253,3	12,71
	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366	63,66
	406523	500016	536700	562527	576449	585953	598149	6110598	623432	636535	636,2
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50	4,30
	98,49	99,01	00,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50	9,92
	998,46	999,00	999,20	999,20	999,20	999,20	999,40	999,60	999,40	999,40	31,00
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53	3,18
	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12	5,84
	67,47	148,51	141,10	137,10	134,60	132,90	130,60	128,30	125,90	123,50	12,94
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63	2,78
	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	1346	4,60
	74,13	61,24	56,18	53,43	51,71	50,52	49,00	47,41	45,77	44,05	8,61

	1	2	3	4	5	6	8	12	24	∞	t
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36	2,57
	16,26	13,27	12,06	11,39	10,97	10,67	10,27	9,89	9,47	9,02	4,03
	47,04	36,61	33,20	31,09	20,75	28,83	27,64	26,42	25,14	23,78	6,86
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67	2,45
	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88	3,71
	35,51	26,99	23,70	21,90	20,81	20,03	19,03	17,99	16,89	15,75	5,96
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23	2,36
	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65	3,50
	29,22	21,69	18,77	17,19	16,21	15,52	14,63	13,71	12,73	11,70	5,40
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,99	2,31
	11,26	8,65	7,59	7,10	6,63	6,37	6,03	5,67	5,28	4,86	3,36
	25,42	18,49	15,83	14,39	13,49	12,86	12,04	11,19	10,30	9,35	5,04
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71	2,26
	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31	3,25
	22,86	16,39	13,90	12,56	11,71	11,13	10,37	9,57	8,72	7,81	4,78
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54	2,23
	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91	3,17
	21,04	14,91	12,55	11,28	10,48	9,92	9,20	8,45	7,64	6,77	4,59

	1	2	3	4	5	6	8	12	24	∞	t
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40	2,20
	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60	3,11
	19,69	13,81	11,56	10,35	9,58	9,05	8,35	7,62	6,85	6,00	4,49
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30	2,18
	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36	3,06
	18,64	12,98	10,81	9,63	8,89	8,38	7,71	7,00	6,25	5,42	4,32
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21	2,16
	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16	3,01
	17,81	12,31	10,21	9,07	8,35	7,86	7,21	6,52	5,78	4,97	4,12
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13	2,14
	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00	2,98
	17,14	11,78	9,73	8,62	7,92	7,44	6,80	6,13	5,41	4,60	4,14
15	4,45	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07	2,13
	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87	2,95
	16,59	11,34	9,34	8,25	7,57	7,09	6,47	5,81	5,10	4,31	4,07

	1	2	3	4	5	6	8	12	24	∞	t
16	4,41	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01	2,12
	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75	2,92
	16,12	10,97	9,01	7,94	7,27	6,80	6,20	5,55	4,85	4,06	4,02
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96	2,11
	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65	2,90
	15,72	10,66	8,73	7,68	7,02	6,56	5,96	5,32	4,63	3,85	3,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92	2,10
	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,01	2,57	2,88
	15,38	10,39	8,49	7,46	6,81	6,35	5,76	5,13	4,45	3,67	3,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88	2,09
	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49	2,86
	15,08	10,16	8,28	7,26	6,61	6,18	5,59	4,97	4,29	3,52	3,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84	2,09
	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42	2,84
	14,82	9,95	8,10	7,10	6,46	6,02	5,44	4,82	4,15	3,38	3,85

	1	2	3	4	5	6	8	12	24	∞	t
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,82	2,08
	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36	2,83
	14,62	9,77	7,94	6,95	6,32	5,88	5,31	4,70	4,03	3,26	3,82
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78	2,07
	7,94	5,72	4,82	4,31	3,99	3,75	3,45	3,12	2,75	2,30	2,82
	14,38	9,61	7,80	6,81	6,19	5,76	5,19	4,58	3,92	3,15	3,79
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76	2,07
	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26	2,81
	14,19	5,46	7,67	6,70	6,08	5,56	5,09	4,48	3,82	3,05	3,77
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73	2,06
	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21	2,80
	14,03	9,34	7,55	6,59	5,98	5,55	4,99	4,39	3,84	2,97	3,75
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71	2,06
	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17	2,79
	13,88	9,22	7,45	6,49	5,89	5,46	4,91	4,31	3,66	2,87	3,72
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69	2,06
	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13	2,78
	13,74	9,12	7,36	6,41	5,80	5,38	4,83	4,24	3,59	2,82	3,71

	1	2	3	4	5	6	8	12	24	∞	T
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67	2,05
	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10	2,77
	13,61	9,02	7,27	6,33	5,73	5,31	4,76	4,17	3,52	2,76	3,69
28	4,19	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65	2,05
	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06	2,76
	13,50	8,93	7,18	6,25	5,66	5,24	4,69	4,11	3,46	2,70	3,67
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64	2,05
	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03	2,76
	13,39	8,85	7,12	6,19	5,59	5,18	4,65	4,05	3,41	2,64	3,66
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62	2,04
	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01	2,75
	13,29	8,77	7,05	6,12	5,53	5,12	4,58	4,00	3,36	2,59	3,64
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39	2,00
	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60	2,66
	11,97	7,76	6,17	5,31	4,76	4,37	3,87	3,31	2,76	1,90	3,36
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,03	1,96
	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,04	2,58
	10,83	6,91	5,42	4,62	4,10	3,74	3,27	2,74	2,13	1,05	3,29

Приложение 8

Таблица 5%-го и 1%-го уровней вероятности коэффициентов корреляции (r_a)

Размер выборки	Положительное значение r_a		Отрицательное значение r_a	
	5%-й уровень	1%-й уровень	5%-й уровень	1%-й уровень
5	0,253	0,297	-0,753	-0,798
6	0,354	0,447	-0,708	-0,863
7	0,370	0,510	0,674	-0,799
8	0,371	0,531	-0,625	-0,764
9	0,366	0,533	-0,593	-0,737
10	0,360	0,525	-0,564	-0,705
11	0,353	0,515	-0,539	-0,679
12	0,348	0,505	-0,516	-0,655
13	0,341	0,495	-0,497	-0,634
14	0,335	0,485	-0,479	-0,615
15	0,328	0,475	-0,462	-0,597
20	0,299	0,432	-0,399	-0,524
25	0,276	0,398	-0,356	-0,473
30	0,257	0,370	-0,324	-0,433
35	0,242	0,347	-0,300	-0,401
40	0,229	0,329	-0,279	-0,376
45	0,218	0,313	-0,262	-0,256
50	0,208	0,301	-0,248	-0,339

Приложение 9

Распределение критерия Дарбина-Уотсона
для положительной автокорреляции
(для 5%-го уровня значимости)

n	*V=1*		*V=2*		*V=3*		*V=4*		*V=5*	
	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,89
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,63	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77

Окончание приложения 9

n	$V=1$		$V=2$		$V=3$		$V=4$		$V=5$	
	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

КРАТКИЙ СЛОВАРЬ ТЕРМИНОВ И ОПРЕДЕЛЕНИЙ

Абсолютное значение 1 % прироста – определяется либо по цепным темпам роста, либо как сотая часть от предыдущего уровня ряда.

Абсолютные величины – это суммарные обобщающие показатели, характеризующие размеры (уровни) общественных явлений в конкретных условиях места и времени.

Абсолютный прирост – разность между двумя уровнями ряда динамики, имеет ту же размерность, что и уровни самого ряда динамики. Абсолютные приросты могут быть цепными и базисными.

База сравнения – абсолютный показатель, находящийся в знаменателе относительной величины.

Варианта – единица варьирующего признака, каждая из которых имеет определенное числовое значение.

Вариация – это колеблемость, многообразие, изменяемость величины признака у единиц совокупности.

Вариация признака – степень количественного отличия индивидуальных значений признака у различных единиц совокупности.

Весами (частотами) называют показатели повторяемости вариант.

Внутригрупповая дисперсия – отражает случайную вариацию, происходящую под влиянием неучтенных факторов и не зависящую от признака-фактора, положенного в основание группировки.

Время наблюдения – это время, к которому относятся собираемые данные, характеризующие объект наблюдения в состоянии, наиболее отвечающем цели и задачам исследования.

Дискретные ряды – ряды распределения по прерывно варьирующему признаку.

Дисперсия альтернативного признака – равна произведению доли на дополняющую эту долю до единицы число.

Единица наблюдения – это составной неделимый элемент объекта наблюдения, являющийся основой счета и носителем определенного круга признаков, наличие (или отсутствие) которых у каждой единицы изучаемой совокупности должно быть зафиксировано в процессе статистического наблюдения.

Единица совокупности – первичный элемент статистической совокупности, который является носителем признака, подлежащего регистрации, основа ведущегося при обследовании счета.

«Идеальный» индекс И. Фишера – представляет собой среднюю геометрическую из произведений двух агрегатных индексов цен Ласпейреса и Паше.

Индекс переменного состава – характеризует увеличение или уменьшение средней цены по группе товаров в результате изменения цены каждого товара и структуры продукции.

Индекс постоянного состава – характеризует изменение средней цены товара в результате влияния только одного фактора – изменения цен на отдельные товары.

Индекс сезонности – относительный показатель, который используют для расчета сезонной составляющей. При исчислении индексов применяют различные методы, выбор которых зависит от характера общей тенденции ряда динамики.

Индекс цен Г. Пааше – представляет собой сравнение агрегированных цен, которые взвешены по физическим объемам текущего периода.

Индекс цен Э. Ласпейреса – представляет собой сравнение агрегированных цен, взвешенных по физическим объемам базисного периода.

Индекс структурных сдвигов – характеризует влияние изменения структуры продукции на величину средней цены товара.

Индивидуальный индекс цен – характеризует динамику цены товара или услуги.

Индивидуальный уровень цен – это абсолютная величина цены в денежном выражении за единицу конкретного товара на рынке.

Индивидуальные абсолютные величины – это величины, которые характеризуют размеры признака у отдельных единиц совокупности.

Интервальный ряд – ряд, в котором представлены значения признака в виде интервала.

Интервальный ряд динамики – это ряд последовательно расположенных значений признака (за определенный период).

Интерполяция – расчет по имеющимся данным за определенный период времени некоторых недостающих значений внутри этого периода.

Коэффициент вариации – применяется для оценки степени интенсивности вариации признака в совокупности.

Коэффициент детерминации – показывает долю (удельный вес) общей вариации изучаемого признака, обусловленную вариацией группировочного признака.

Критический момент наблюдения – это момент времени, по состоянию на который регистрируются сведения об единицах наблюдения.

Медиана – величина признака, которая находится в середине ряда.

Межгрупповая дисперсия – характеризует вариацию изучаемого признака, возникающую под влиянием признака-фактора, положенного в основание группировки.

Место наблюдения – это место, где производится регистрация данных и заполнение статистических формуляров.

Мода – значение признака (варианты), который чаще всего встречается в данной совокупности.

Моментный ряд динамики – это ряд последовательно расположенных значений признака (на определенную дату).

Несплошное наблюдение – это такое наблюдение, при котором обследованию подвергаются не все единицы изучаемой совокупности, а только их часть.

Общая дисперсия – измеряет вариацию признака во всей совокупности под влиянием всех факторов, обусловивших эту вариацию.

Объект наблюдения – это ограниченное в пространстве и времени определенное целостное множество взаимосвязанных единиц наблюдения, о котором должны быть собраны статистические сведения.

Организационный план наблюдения – это документ, в котором фиксируется решение важнейших вопросов подготовки и проведения статистического наблюдения с указанием конкретных сроков проведения намеченных мероприятий.

Осредняемый (варьирующий) признак – это признак для которого исчисляется средняя.

Относительная величина – это результат деления одного абсолютного показателя на другой и выражение соотношения между количественными характеристиками социально-экономических явлений и процессов.

Отчетность – это организационная форма статистического наблюдения, при которой сведения поступают в статистические органы от предприятий, учреждений и организаций в виде обязательных отчетов об их деятельности в строго установленные сроки и установленном порядке.

Ошибки наблюдения – это расхождения между установленными статистическим наблюдением и действительными значениями изучаемых величин.

Период наблюдения – это календарный промежуток времени, в течение которого осуществляются сбор, проверка статистических данных и оформление их в статистические формуляры.

Правило сложения дисперсий – общая дисперсия, возникающая под действием всех факторов, равна сумме дисперсии, появляющейся под влиянием всех прочих факторов, и дисперсии, возникающей за счет группировочного признака.

Программа наблюдения – совокупность признаков, подлежащих наблюдению и регистрации, где каждый признак логически связан с целью и задачами исследования.

Размах вариации – показывает, на сколько велико различие между единицами совокупности, имеющими самое маленькое и самое большое значения признака.

Ранжированный вариационный ряд – представляет собой перечень отдельных единиц совокупности в порядке возрастания (убывания) значений варьирующего признака.

Ретрополяция – нахождение по имеющимся данным за определенный период времени недостающих значений в начале динамического ряда.

Ряд динамики – это ряд последовательно расположенных статистических показателей (в хронологическом порядке), изучение которых показывает определенную тенденцию развития изучаемого явления.

Ряды распределения – ряды числовых показателей, характеризующих закономерности распределения изучаемой совокупности по значениям того или иного варьирующего признака.

Сезонная волна – графическое изображение полученных индексов сезонности.

«Смыкание рядов» − объединение в один более длинный динамический ряд двух (или нескольких) рядов динамики, уровни которых исчислены по различной методологии или по различным границам территорий. Для смыкания необходимым условием является наличие за один период данных, рассчитанных по разной методологии (или в разных границах).

Сплошное наблюдение – это наблюдение, при котором обследованию подвергаются все без исключения единицы изучаемой совокупности (например, перепись населения страны).

Сравниваемым показателем является абсолютная величина, расположенная в числителе при расчете относительных величин.

Средняя величина – это обобщающий показатель, который дает количественную характеристику признака в статистической совокупности в условиях конкретного места и времени.

Суммарные абсолютные величины – характеризуют итоговое значение признака по определенной совокупности субъектов, охваченных статистическим наблюдением.

Статистика – наука, которая изучает количественную сторону массовых социально-экономических явлений в неразрывной связи с их качественной стороной, а также количественное выражение закономерностей развития процессов в конкретных условиях места и времени.

Статистическая совокупность − представляет собой определенное множество единиц совокупности, которые количественно отличаются друг от друга своими характеристиками, но объединены какой-либо качественной основой. Могут быть однородными и разнородными.

Статистический показатель – категория, которая дает количественную характеристику соотношения признаков общественных явлений. Может быть объемным и расчетным.

Статистический признак – зарегистрированная в ходе сбора первичных данных характеристика единицы совокупности, ее качественная особенность. Признак может быть первичным и вторичным, количественным и атрибутивным.

Суммарные абсолютные величины – характеризуют итоговое значение признака по определенной совокупности субъектов, охваченных статистическим наблюдением.

Темп (коэффициент) прироста – показатель, характеризующий относительную скорость изменения уровня ряда в единицу времени.

Темп (коэффициент) роста – относительный показатель, характеризующий интенсивность изменения уровня ряда. Темпы роста могут рассчитываться как цепные (с предшествующим уровнем ряда), так и базисные (с одним и тем же уровнем).

Тренд – это основная (достаточно устойчивая) тенденция развития явления в ряду динамики.

Формуляр наблюдения – представляет собой особым образом разграфленный лист (листы) бумаги, в котором содержится перечень вопросов программы, свободные места для записи ответов на них, а также для записи шифров (кодов) ответов.

Экстраполяция – расчет прогнозного значения.

Эмпирический коэффициент эластичности – отражает процентные изменения цены в результате увеличения факторного признака на 1 %.

Эмпирическое корреляционное отношение – показывает тесноту связи между группировочным и результативным признаками (связь между причиной и следствием) и изменяется в пределах от 0 до 1.

Инна Дмитриевна Минина
Наталья Владимировна Королькова

СТАТИСТИКА

Учебное пособие

Часть 1 Теория статистики

для студентов высших учебных заведений, обучающихся
по направлениям 080100 «Экономика» (профиль «Бухгалтерский
учет, анализ и аудит» и «Финансы и кредит»)
и 080200 «Менеджмент»

Компьютерная верстка З.Р. Абдуллиной

Корректор Л.А. Артамонова

Сдано в производство 23.01.13 Формат 60 × 84 1/16
Бумага Гознак Print Отпечатано на ризографе
Усл. печ. л. 14,65 Тираж 150 экз.Заказ № 5

РИО ПГСХА
440014, г. Пенза, ул. Ботаническая, 30