

Министерство образования и науки Российской Федерации
**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)**
Кафедра автоматизации обработки информации (АОИ)

СТАТИСТИКА
Учебное пособие
для студентов специальности
«Государственное и муниципальное управление»

Разработчик
доцент каф. АОИ, к.т.н.
З.П.Лепихина

Томск-2013

СОДЕРЖАНИЕ

Введение.....	4
1. Предмет, метод и задачи статистики	6
1.1. Развитие статистической науки	6
1.2. Предмет статистики	9
1.3. Отрасли статистики	11
2. Основные категории статистики	17
3. Статистическое исследование	22
3.1. Этапы статистического исследования	22
3.2. Подготовка статистического наблюдения	22
3.3. Формы, виды и способы статистического наблюдения	25
4. Абсолютные и относительные статистические величины	35
4.1. Абсолютные статистические величины	35
4.2. Относительные статистические величины	35
5. Средние значения и показатели вариации	39
5.1. Вариационные ряды	39
5.2. Виды средних величин	40
5.3. Степенные средние	41
5.4. Структурные средние	45
5.5. Показатели вариации	48
6. Статистические группировки	51
6.1. Значение и сущность группировки	51
6.2. Виды группировок	52
7. Выборочное наблюдение	62
7.1. Понятие о выборочном наблюдении	62
7.2. Простая случайная выборка	64
7.3. Расслоенная (типическая или районированная) выборка ..	67
7.4. Серийная выборка	70
7.5. Механическая выборка	71
7.6. Другие виды	72
7.7. Малые выборки	73
7.8. Решение типовых задач	74
8. Ряды динамики.....	79
8.1. Виды рядов динамики	79
8.2. Показателя рядов динамики.....	79
8.3. Средние показатели.....	85
8.4. Изучение основной тенденции развития.....	88
8.5. Анализ сезонных колебаний.....	98
9. Экономические индексы.....	104
9.1. Основные понятия и определения.....	104
9.2. Индивидуальные и общие индексы.....	106

9.3. Агрегатный индекс как исходная форма индекса.....	111
9.4. Средние индексы.....	115
9.5. Выбор базы и весов индексов.....	118
9.6. Индексы структурных сдвигов.....	121
9.7. Индексы пространственно-территориального сопоставле- ния.....	123
9.8. Индексы-дефляторы.....	125
Литература.....	130

ВВЕДЕНИЕ

В современном обществе статистика является одним из важнейших инструментов управленческой деятельности. В практической деятельности вообще и в практике управления экономикой и социальной сферой, в особенности, используется различная по степени обобщения статистическая информация, различные приемы статистической методологии.

Статистика — наука о массовых стохастических явлениях или процессах, метод их исследования. Массовые явления — непременная принадлежность всякого общества. Каждое из этих явлений содержит элемент случайности, вариацию, индивидуальность, однако, рассматривая совокупность, массу этих явлений, статистика позволяет систематизировать их и обнаруживать определенные регулярности, выражающие причинные связи, закономерности, тенденции. Обобщающие статистические показатели представляют собой выражение закономерностей, являющихся результатом взаимного погашения множества отдельных явлений с отклонениями случайного характера. «При некоторых весьма общих условиях, — писал академик В.С. Немчинов, — результат совместного действия большого числа индивидуальных причин и условий (содержащих в себе элементы случайного характера) почти не зависит от случая». В этом состоит суть закона больших чисел, на котором основана статистика.

В статистике первостепенное значение имеет группировка изучаемых явлений. Именно группировка служит основой научной обработки статистических данных. Группировка позволяет объединить отдельные единицы статистической совокупности в определенные категории, типы, обладающие единой сущностью, выражающие социальную однородность. Вне научной группировки, вне дифференциации в социальном отношении данные статистики не будут выражать сущности явлений и не могут служить делу управления. Наоборот, они дадут искаженную картину социальной действительности.

Своевременное получение полной, ценной, достоверной информации — нелегкая задача, стоящая перед субъектом управления, и решается она разными путями. Здесь и совершенствование системы показателей, описывающих социальную сферу, и совершенствование информационной системы, включая как техники и технологии работы с информацией, так и профессиональной подготовки людей, работающих с информацией.

Задача государственной статистики состоит в том, чтобы организовать наблюдение и анализ процессов, протекающих в общественной

жизни, измерение степени соответствия деятельности органов управления требованиям закономерностей. В частности, статистика призвана определять качество управленческих решений, в особенности программ, планов, факторов, способствующих их выполнению.

Общественные процессы складываются из множества взаимодействующих факторов. Задача статистики состоит в том, чтобы показать место и роль каждого из факторов в общем процессе, раскрыть причинно-следственные зависимости, указать на узкие места. Обобщенные отчеты, группировка и сводные дополнительные аналитические данные, выборочные исследования и дифференцированные территориальные показатели — таковы основные статистические методы исследования, позволяющие в массе конкретных явлений выделять тенденции и закономерности.

Знание статистики необходимо специалисту для принятия решений в условиях стохастики (когда анализируемые явления подвержены влиянию случайностей), для анализа элементов рыночной экономики, в сборе информации, в связи с увеличением числа хозяйственных единиц и их типов, аудите, финансовом менеджменте, прогнозировании.

Курс «Статистика» построен на базе государственного стандарта. Основной задачей курса статистики является овладение знаниями, умениями и навыками в использовании методов получения статистической информации о социально-экономических явлениях и ее обобщения, методологии экономико-статистического анализа и исчисления обобщающих статистических показателей. Изучение данной дисциплины базируется на дисциплинах «Математика», «Теория вероятностей и математическая статистика».

Теоретический материал сопровождается тестами для самоконтроля и упражнениями для самостоятельного выполнения.

При подготовке пособия использованы учебники, учебные пособия и практикумы российских и зарубежных авторов, приведенные в списке использованной литературы статистические данные органов государственной статистики, а также лекционный материал и методические разработки автора.

1. ПРЕДМЕТ, МЕТОД И ЗАДАЧИ СТАТИСТИКИ

1.1. Развитие статистической науки

Слово «статистика» происходит от латинского STATUS – состояние, положение вещей. Первоначально оно употреблялось в значении «политическое состояние». В научный обиход слово «статистика» ввел в XVIII веке немецкий ученый Г. Ахенваль. К тому же времени относится начало преподавания статистики в университетах Германии.

История развития человечества показала, что без статистических данных невозможно управление государством, развитие отдельных отраслей и секторов экономики, обеспечение оптимальных пропорций между ними. Необходимость сбора и обобщения множества данных о населении страны, предприятиях, банках, фермерских хозяйствах и т.д. приводит к существованию специальных статистических служб-учреждений государственной статистики.

Уже в древний период истории человечества хозяйственные и военные нужды требовали наличия данных о населении, его составе, имущественном положении. С целью налогообложения организовывались переписи населения, производился учет земель.

Со временем собирание данных о массовых общественных явлениях приобрело регулярный характер.

С середины XIX в. благодаря усилиям великого бельгийского ученого-математика, астронома и статистика Адольфа Кетле (1796-1874 гг.) были выработаны правила переписей населения и установлена регулярность их проведения в развитых странах. Для координации развития статистики по инициативе А. Кетле проводились международные статистические конгрессы, а в 1885 г. был основан Международный статистический институт, существующий и сейчас.

Становление государственной статистики в России можно отнести к концу XII — началу XIII в., хотя первые переписи земель и населения с постоянно усложнявшейся программой проводились еще в Киевской Руси (IX — XII вв.). Реформы Петра I (1672 – 1725), которыми были охвачены все основные направления общественной жизни: экономика страны, административное управление, армия, культура и быт населения, а также войны вызывали потребность в полном и точном учете материальных ресурсов и населения. В этот период высший правительственный орган — Сенат — через систему коллегий не только руководил экономикой страны, но и являлся центром по проведе-

нию важнейших статистических работ, там собирались полученные материалы обследований, отчеты подведомственных коллегиям производств и заведений, а также местной администрации.

Петровская реформа налоговой системы связана с появлением новой единицы, ею стала «душа» мужского пола, что потребовало подушной переписи населения — ревизии. Первая ревизия была объявлена 26 ноября 1718 г., ревизию проводила армия.

В начале XIII в. В России зарождался и текущий учет населения. Так, в 1702 г. был издан указ о подаче в Патриарший Духовный приказ приходскими священниками недельных ведомостей о родившихся и умерших. В первой половине XIII в. Проводились уже переписи рабочих фабрик и мануфактур.

Первая половина XIX в. Связана с новым этапом в развитии отечественной статистики. В сентябре 1802 г. в соответствии с Высочайшим манифестом императора Александра I вводится письменная отчетность министерств. Так началось операционно-структурное оформление государственной статистики в России. Этот год принято считать годом рождения российской государственной статистики.

В 1811 г. впервые был создан официальный центр правительственной статистики — Статистическое отделение при Министерстве внутренних дел; сюда поступала отчетность губерний. Первым руководителем Статистического отделения был К.Ф. Герман.

Российские ученые внесли большой вклад в развитие статистической науки. Большое значение, например, имеет работа Д.П. Журавского «Об источниках и употреблении статистических сведений», изданная в 1846 году. Определив статистику как «счет по категориям», Журавский отмечал, что статистика необходима для «изучения всего, относящегося к человеку». Журавский определил важнейшие разделы социальной статистики:

- статистика народонаселения — необходимость его исчисления по классам и занятиям;
- изучение народного быта, жилища, питания;
- статистика театров, клубов, дворянских собраний, народных увеселений;
- статистика учреждений, охраняющих права собственности;
- статистика нищеты, бедности, сиротства;
- статистика самоубийств с указанием средств, причин, званий, возраста и прочих характеристик лиц, лишивших себя жизни.

Во всех предложениях Д.П. Журавский проводил идею как можно более точного и полного выявления *дифференциации* людей по условиям их жизни, по состоятельности.

Особое место в истории российской статистики принадлежит земской статистике. При земствах, органах местного самоуправления, с середины 70-х годов XIX века были созданы специальные статистические бюро. Земские статистики собирали и разрабатывали огромный статистический материал, который использовался для глубоких экономических и социальных исследований пореформенной России. Работа земской статистики характеризуется не только сбором и разработкой статистических данных, но и развитием статистической методологии.

Видными земскими статистиками были В.И. Орлов, П.П. Червинский, Ф.А. Щербина, А.П. Шликевич.

В 90-х годах были созданы фабрично-заводские инспекции, которые вели текущую статистику, разрабатывали данные по статистике труда, в том числе о составе рабочей силы, несчастных случаях, стачках и др.

Стала развиваться промышленная статистика. Под руководством В.Е. Варзара в 1900, 1908 и 1912 гг. были проведены первые переписи промышленности.

Начальный этап советской статистики (1917 – 1930 гг.) отличается исключительной интенсивностью: проводится большое число специально организованных, статистических переписей и обследований, плодотворно работают различные научные коллективы, строится первый баланс народного хозяйства.

Последующее развитие советской статистики тормозилось созданием в 30-е годы административно-бюрократической системой, массовыми репрессиями, в том числе и лучших экономистов и статистиков (Н.Д. Кондратьева, А.В. Чаянова, В.Г. Громана, О.А. Квитнина и многих других).

В это время формируются отраслевые статистики, складывается система объемных показателей, скрывающая негативные тенденции в развитии народного хозяйства. Активно разрабатываются и качественные статистические показатели (индексы производительности труда, себестоимости и др.). Статистика подчиняется решению оперативных задач, оценке выполнения плана в ущерб ее аналитическим функциям.

В годы Великой Отечественной войны, перед советской статистикой стояли задачи по оперативному учету трудовых, материальных ресурсов, перемещение производственных сил страны в восточные районы.

После войны роль и значение статистики возросли: развернулись балансовые работы, углубилась теория индексного метода и расшири-

лась практика его применения, получили распространение экономико-математические модели и методы, развитие прикладной статистики.

1.2. Предмет статистики

В настоящее время термин «статистика» употребляется в трех значениях:

- **отрасль практической деятельности** (статистический учет) по сбору, обработке, анализу и публикации массовых цифровых данных о самых различных явлениях и процессах общественной жизни. Эту деятельность на профессиональном уровне осуществляет государственная статистика;

- **совокупность цифровых сведений, статистические данные**, представляемые в отчетности предприятий, организаций, отраслей экономики, а также публикуемые в сборниках, справочниках, которые являются результатом статистической работы;

- **отрасль общественных наук, специальная научная дисциплина**, изучаемая в высших и средних специальных учебных заведениях. Статистика как наука представляет собой целостную систему научных дисциплин: теория статистики, экономическая статистика и ее отрасли, социальная статистика и ее отрасли, и т.д.

Теория статистики является наукой о наиболее общих принципах и методах статистического исследования социально-экономических явлений. Она разрабатывает понятийный аппарат и систему категорий статистической науки, рассматривает методы сбора, сводки, обобщения и анализа статистических данных, т.е. общую методологию статистического исследования массовых общественных процессов. Таким образом, **теория статистики – методологическая основа всех отраслевых статистик.**

Экономическая статистика разрабатывает и анализирует статистические показатели такие как национальный доход, валовой национальный продукт и др., отражающие состояние национальной экономики; структуру, пропорции, взаимосвязь отраслей; рассматривает особенности размещения производственных сил, состав и использование материальных, трудовых и финансовых ресурсов; наконец, осуществляет построение и анализ общей макростатистической модели рыночной экономики в виде национальных счетов.

Предметом любой науки являются свойства и отношения того объекта, природного и социального, на который направлена ее познавательная деятельность.

Предмет статистики — количественная сторона массовых общественных явлений в непрерывной связи с их качественной стороной, с их социально-экономическим содержанием.

Свой предмет статистика изучает методом обобщающих показателей. Она анализирует также природные ресурсы и природные условия, поскольку они влияют на жизнь общества. Статистика изучает:

- массовые общественные явления при помощи статистических показателей (численность населения, количество произведенной в стране конкретной промышленной, сельскохозяйственной, строительной и другой продукции за определенный период), их динамику (изменения уровня жизни населения и др.);

- количественную сторону массовых общественных явлений и дает количественное, цифровое освещение общественных явлений;

- количественную сторону массовых общественных явлений в неразрывной связи с их качественным содержанием; наблюдает в обществе процесс перехода количественных изменений в качественные;

- количественную сторону массовых общественных явлений в конкретных условиях места и времени (динамику численности населения и занятости его по секторам экономики, объема производства, распределения доходов, потребления и т.д.); характеризует явления общественной жизни в конкретных пространственных и временных границах;

- количественные связи между общественными явлениями с помощью специальной методологии; использует математические методы при исчислении ряда статистических показателей (ошибок выборки, тесноты связи и т.д.), в свою очередь гуманитарные и естественные науки широко применяют в своих исследованиях методы статистики для сбора, обработки и анализа данных.

Итак, статистика – отрасль общественной науки, изучающая методом обобщающих показателей количественную сторону определенных массовых социально-экономических явлений и закономерностей их развития в конкретных условиях места и времени.

Знание статистики необходимо специалисту для принятия решений в условиях стохастики (когда анализируемые явления подвержены влиянию случайностей), для анализа элементов рыночной экономики, в сборе информации, в связи с увеличением числа хозяйственных единиц и их типов, аудите, финансовом менеджменте, прогнозировании.

1.3. Отрасли статистики

Экономическая статистика разрабатывает и анализирует статистические показатели такие как национальный доход, валовой национальный продукт и др., отражающие состояние национальной экономики; структуру, пропорции, взаимосвязь отраслей; рассматривает особенности размещения производственных сил, состав и использование материальных, трудовых и финансовых ресурсов; наконец, осуществляет построение и анализ общей макростатистической модели рыночной экономики в виде национальных счетов.

Отрасли экономической статистики – статистика промышленности, сельского хозяйства, строительства, транспорта, связи, труда, природных ресурсов, охраны окружающей среды и т.д. – разрабатывают и изучают статистические показатели развития соответствующих отраслей.

Социальная статистика представляет собой одно из важнейших *приложений* статистического метода. Она дает количественную характеристику структуры общества, жизни и деятельности людей, их взаимоотношений с государством и правом, позволяет выявить и измерить закономерности в поведении людей, распределении благ между ними.

Следует подчеркнуть, что общество *неоднородно*, поэтому важнейшим направлением социальной статистики является изучение *дифференциации*, особенностей поведения отдельных групп населения.

Социальная статистика отличается от других отраслей статистики не только своим предметом и объектом исследования. Ее своеобразие состоит:

- в особых каналах получения исходной информации (например, социологические исследования);
- в применении специальных приемов обработки и обобщения этой информации (например, методы анализа нечисловой информации, непараметрические тесты и др.);
- в особых путях практического использования результатов.

При анализе социальных проблем необходимо учитывать конкретно-исторические условия того или иного общества. Невозможно правильно оценить и интерпретировать статистические показатели условий жизни, если абстрагироваться от таких факторов как, например:

- национально-культурные традиции;
- степень гуманизации общества;
- степень развития потребностей;
- территориальная мобильность населения и др.

Для социальной статистики характерна **множественность** объектов исследования. Их можно подразделить на два типа.

Первый и основной тип составляют потребители услуг, материальных и духовных ценностей, информации. Они представлены индивидуальными и групповыми объектами. Индивидуальный объект – человек (население как совокупность индивидов). Это также все население и отдельные его категории в зависимости от исследуемого социального процесса. Коллективный объект – группа лиц, совместно осуществляющая потребление, совместно участвующая в социальном процессе. Такими объектами являются семья, трудовой коллектив, садоводческое товарищество, гаражный кооператив и др.

Второй тип объектов охватывает лиц, организации, структуры, предоставляющие населению услуги, организующие тот или иной социальный процесс. Их деятельность определяет объем и качество предоставленных услуг и ценностей. Производство и потребление услуг, ценностей, информации составляют две взаимосвязанные стороны процесса. Этим предопределяется целесообразность их параллельного исследования. Так, жилищная проблема может быть раскрыта, если информация получена по разным видам объектов: семьям, где система показателей характеризует жилищные условия и их динамику, и организациям, формирующим рынок жилья. К ним относятся строительные организации, различные жилищные отделы и комиссии в составе местных органов управления, разнообразные посреднические конторы и фирмы по обмену, купле, продаже, найму жилья.

В отдельных случаях оба типа объектов представлены в единстве – когда, например, семьи сами своими силами осуществляют строительство жилого дома для себя.

Имеются и другие не менее важные особенности, присущие главным образом социальной статистике и сравнительно слабо выраженные, например, при изучении чисто экономических процессов.

В сфере производства в качестве единиц совокупности представлены предприятия, отличающиеся территориальной закреплённостью, не подверженные быстрому, частому и радикальным изменениям. Население же отличается большой территориальной подвижностью, поэтому затрудняется сбор информации. Рождаемость и смертность непрерывно меняют состав населения каждой территории. Каждый человек, каждая семья достаточно часто меняют свои демографические и соци-

ально-экономические показатели. В результате оказывается затруднительным регулярно следить за всеми изменениями. Лишь переписи населения один раз в десять лет (в пять лет) позволяют получить наиболее важные сведения о населении. Однако и они не в полной мере способны удовлетворить потребности социальной статистики в информации о структуре и качественных характеристиках объекта исследования.

Такая сложная ситуация нередко приводит к тому, что приходится оперировать преимущественно общими показателями потребления на уровне страны и отдельных регионов. Отсутствуют в большинстве своем показатели качества потребления по различным категориям семей, по различным социально-экономическим, демографическим и этническим группам населения.

Так, обеспеченность населения транспортными услугами в городах определяется только такими сводными показателями, как наличный подвижной состав по видам транспорта, общий объем пассажиро-перевозок. При этом трудно определить уровень удовлетворенности населения в транспортных услугах.

Лишь отчасти дефицит информации компенсируется тем, что учет некоторых данных осуществляется в выборочном порядке. Наиболее ценными в этом отношении являются материалы бюджетной статистики. Проводятся некоторые единовременные обследования по ряду других проблем социальной статистики. Как правило, в подобных работах содержатся обстоятельные характеристики потребителей, их дифференциация по условиям и уровням потребления. Слабой стороной этого источника информации служит то обстоятельство, что не все проблемы могут быть изучены по материалам, не всегда обеспечивается достаточная регулярность проведения таких работ, не все территории охвачены обследованиями. Региональные органы управления и отраслевые ведомства по своей инициативе и своими силами за счет своих средств часто предпринимают социальные исследования (обычно в форме договора на их проведение с научными учреждениями) по наиболее актуальным прикладным вопросам.

Чтобы не ошибиться при принятии решений при распределении ресурсов и обеспечить должную адресность различных практических мер, требуется оперативная и конкретная информация, которую заинтересованные субъекты получают, проводя сбор необходимых статистических данных по инициативе отдельных ведомств.

Социальная статистика решает задачи, которые решаются любой отраслевой статистикой применительно к *своему* объекту исследования:

- систематический анализ ситуаций в социальной сфере;
- изучение уровня и условий жизни населения в целом и по группам;
- анализ важнейших тенденций и закономерностей развития отраслей социальной инфраструктуры;
- анализ динамики социальных показателей;
- прогнозирование;
- исследование взаимодействия социальных процессов с другими составляющими общественного развития;
- оценка степени соответствия фактических параметров их нормативным значениям.

Кроме того, существуют особые задачи, присущие именно социальной статистике. Их специфика зависит, прежде всего, от трудностей, возникающих *в практике* изучения социальных процессов. К ним относятся следующие:

1) *преодоление автономности отдельных направлений* социальной статистики и обусловленной этим несопоставимости многих статистических показателей, действительное формирование единой взаимосвязанной системы социальной статистики;

2) *достижение соответствия* ряда статистических показателей оценке сущности социальных явлений и процессов, так как показатели не дают их качественных характеристик. Учитываются лишь отдельные формальные количественные параметры. Например, состояние системы здравоохранения трудно реально оценить, опираясь лишь на данные о числе приходящихся на 1000 человек населения врачей и больничных коек;

3) постепенное *интегрирование* исследований на макро- и микроуровнях, что позволит более глубоко и полно вскрыть первопричины и механизмы изучаемых процессов. Пока еще социальная статистика ориентирована преимущественно на исследования явлений и процессов на макроуровне, где обнаруживаются конечные результаты процесса. Решение этой задачи – одно из перспективных направлений, которое позволит радикально изменить уровень познания;

4) разработка показателей, построение моделей, оценка гипотез для наиболее характерных социально-культурных, социально-этнических, социально-демографических групп населения. Используемые при этом схемы группировок населения следует корректировать, по мере того как происходят сдвиги в составе населения. Действующая система показателей социальной статистики практически нивелирует реально существующую дифференциацию условий жизнедеятельности различных групп населения;

5) преодоление существующей *несопоставимости* показателей социальной статистики и показателей, представленных в других отраслевых статистиках;

6) *моделирование* социально-экономических связей с целью обнаружения механизмов взаимодействия в общественной системе. На макроуровне представлен ряд объективно существующих ограничительных факторов, предопределяющих пределы возможных колебаний социальных показателей в данных конкретных условиях (без разрушения системы). Это важно учитывать при выработке социальных программ;

7) *расширение круга показателей* статистики мнений. Актуальность этой задачи заключается в том, что важнейшей составляющей социальных процессов выступает психологический фактор. Субъективные личностные оценки факторов и событий предопределяют реакцию населения на них, поведение населения в различных сферах жизнедеятельности;

8) *проведение специальных мер*, компенсирующих по возможности такие слабые стороны многих показателей, как элементы субъективизма; неточности данных анамнеза (информация о событиях и фактах прошлых лет, получаемая при опросах населения); неполнота учета фактов, о которых люди неохотно дают информацию; отсутствие объективных однозначных критериев и шкал для различного рода оценочных суждений и т.д. Это одно из важных условий полноценной системы показателей социальной статистики, повышающее ее достоверность и информационную емкость.

Тест для самоконтроля

1. В научный обиход термин «статистика» был введен:
 - а) в 1896 году
 - б) в XVIII веке
 - в) в начале XX века
2. В древний период истории проводились такие статистические работы как:
 - а) переписи населения
 - б) учет земель
 - в) регистрация станочного парка предприятий
3. Первым руководителем Статистического отделения был.:
 - а) А.Кетле
 - б) К.Ф. Герман
 - в) В.И. Орлов

4. Первые переписи промышленности в России были проведены.
- а) в начале XIX века
 - б) в конце XIX века
 - в) в начале в начале XX века

Вопросы и упражнения

1. Охарактеризуйте роль реформ Петра I в развитии статистики.
2. Назовите выдающихся представителей земской статистики.
3. Перечислите статистические методы, получившие развитие во второй половине XX века
4. Приведите примеры использования статистики для решения задач управления
5. Объясните толкование термина «статистика» в значении «статистический учет».

2. ОСНОВНЫЕ КАТЕГОРИИ СТАТИСТИКИ

Статистика оперирует определенными категориями, т.е. понятиями, отражающими явления действительности.

Статистика изучает закономерности массовых явлений. Объект конкретного статистического исследования называют статистической совокупностью.

Статистическая совокупность – это множество единиц (объектов, явлений), объединенных единой закономерностью и варьирующих в пределах общего качества. Такова, например, совокупность предприятий, производящих однотипную продукцию, но различающихся между собой объемами производства, трудовыми и финансовыми ресурсами; совокупность домохозяйств; совокупность студентов и т.п.

Специфическим свойством статистической совокупности является **массовость** единиц, поскольку явление характеризуется массовым процессом и всем многообразием определяющих его причин и форм.

Под **единицами совокупности** понимаются ее неделимые первичные элементы, выражающие ее качественную однородность, т.е. являющиеся носителями признаков. Например, единицами совокупности могут выступать акционерные общества, фирмы, фермерские хозяйства, человек, семья, станок, изделие и т.д.

Под качественной **однородностью** единиц совокупности понимается сходство единиц (объектов, явлений) по каким-либо существенным признакам, но различающихся по каким-либо другим признакам. Например, множество промышленных предприятий наряду с качественной однородностью обладает различиями по размеру основных фондов, объему производства, численности работающих и т.д.

Единицы статистической совокупности характеризуются общими свойствами, именуемыми в статистике признаками.

Признак – показатель, характеризующий некоторое свойство объекта совокупности, рассматриваемый как случайная величина. Например, единица статистической совокупности – «предприятие» – имеет следующие признаки: объемы произведенной и реализованной продукции, соотношение собственных и заемных средств, издержки производства, численность работников и т.д.

Вариация – различия под совокупным влиянием разнообразных факторов, которые по-разному сочетаются в каждом конкретном случае. Например, успеваемость отдельного студента зависит от затрат времени на подготовку к занятиям, способностей к обучению, возраста и т.п.

Варьирующие признаки могут быть *количественными*, если их варианты выражаются числовыми значениями (возраст, стаж работы, оплата труда и пр.) и *неколичественными* (*атрибутивными*), не имеющими числового выражения и представляющими собой смысловые понятия (профессия, социальная принадлежность и т.д.)

Количественные признаки могут быть *дискретными* и *непрерывными*.

В случае, когда варианты признака могут принимать одно из двух противоположных значений, говорят об *альтернативном признаке*. Например, продукция может быть годной или бракованной.

Признаки подразделяются на *существенные* или главные, выражающие содержательную сторону явлений, и *несущественные* или второстепенные. Статистическому изучению подлежат существенные признаки.

Признаки, характеризующие статистическую совокупность, взаимосвязаны между собой, поэтому следует различать факторные и результативные признаки.

Факторные признаки – это независимые признаки, оказывающие влияние на другие, связанные с ними признаки.

Результативные признаки – это зависимые признаки, которые изменяются под влиянием факторных признаков. Так, квалификация, стаж работы – факторные признаки; производительность труда – результативный.

Статистический показатель – это количественно-качественная обобщающая характеристика какого-то свойства группы единиц или совокупности в целом. Этим он отличается от индивидуальных значений, которые называются признаками. Например, средний размер сберегательного вклада граждан страны – статистический показатель, размер вклада конкретного человека – признак.

Статистические показатели можно условно подразделить на *первичные* (объемные, количественные, экстенсивные) и *вторичные* (производные, качественные, интенсивные).

Первичные – характеризуют либо общее число единиц совокупности, либо сумму значений какого-либо признака (общая численность студентов вузов, объем выпускаемой продукции за год и т.д.).

Вторичные производные показатели обычно выражаются средними и относительными величинами и, взятые в динамике, характеризуют путь интенсивного развития (повышение эффективности использования ресурсов, рост (снижение) производительности труда, материалоемкости и трудоемкости единицы продукции и ее себестоимости).

Поскольку отдельные свойства совокупности не изолированы, а связаны между собой, то и статистические показатели, характеризующие эти свойства, не являются разрозненными, а образуют *систему показателей*.

Система статистических показателей – это совокупность взаимосвязанных показателей, объективно отражающая существующие между явлениями взаимосвязи, она охватывает все стороны жизни общества как на макро-уровне (страна, регион), так и микро-уровне (отдельное предприятие, фирма, объединение, домохозяйство, семья).

Случайное событие – событие, которое при заданной совокупности условий может произойти, а может и не произойти, но для которого определена вероятность его осуществления. **Случайность является формой проявления необходимости**. Влияние случайности затрудняет исследование присущих изучаемому явлению закономерностей. При соединении же большого числа явлений действия элементов случайностей взаимопогашаются, хотя они могут проявляться в признаках индивидуальных единиц статистической совокупности (вероятность того, что человек будет жив через год, значительно выше для юноши, чем для человека преклонного возраста, однако с юношей может произойти несчастный случай, и только наблюдая массу людей самого разного возраста, можно выявить закономерные возрастные различия уровня смертности).

Важнейшей категорией статистики является статистическая закономерность. Под закономерностью вообще принято понимать повторяемость, последовательность и порядок изменений в явлениях.

Статистическая закономерность – количественная закономерность изменения в пространстве и во времени массовых явлений и процессов общественной жизни, состоящих из множества элементов (единиц совокупности) Она проявляется не в индивидуальном явлении, а в массе однородных явлений, при обобщении данных статистической совокупности, т.е. в среднем. Следовательно, это средняя закономерность массовых явлений и процессов.

Статистическая закономерность отражает относящиеся к определенному пространству и времени причинно-следственные связи, выражающиеся в последовательности, регулярности, повторяемости событий с достаточно высокой степенью вероятности.

Статистическая закономерность устанавливается на основе анализа **массовых** данных, что обуславливает ее взаимосвязь с законом больших чисел.

Тесты для самоконтроля

1. Статистика изучает:
 - а) количественную сторону массовых явлений в непрерывной связи с их качественной стороной, с их социально-экономическим содержанием
 - б) количественную сторону единичных явлений, отдельные факты
 - в) наиболее общие законы общественного развития
2. К отраслям социальной статистики относят:
 - а) статистику транспорта
 - б) статистику занятости и безработицы
 - в) статистику сельского хозяйства
3. Примером качественного (атрибутивного) признака является
 - а) денежный доход
 - б) число детей
 - в) национальность
4. Под единицей совокупности понимается:
 - а) объект, который изучается статистическими методами
 - б) неделимый первичный элемент совокупности, выражающий ее качество
 - в) такого понятия в статистике не существует
5. «Средний размер стипендии по вузу» - это
 - а) статистический показатель
 - б) единица наблюдения
 - в) количественный признак
6. Примером дискретного признака является
 - а) средний размер заработной платы
 - б) национальность
 - в) число детей

Вопросы и упражнения

1. Что такое «статистическая совокупность» и в чем ее специфика?

2. Может ли в качестве единицы совокупности выступать семья? В каком случае?
3. В чем отличие понятий «статистический признак» и «статистический показатель»?
4. Приведите примеры альтернативных признаков.
5. Из двух признаков «уровень жизни» и «размер пенсии» какой Вы приняли бы в качестве факторного? Объясните свой выбор.
6. Какова взаимосвязь статистической закономерности с законом больших чисел?

3. СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ

3.1. Этапы статистического исследования

Для выявления и характеристики размеров и количественных соотношений социальных явлений и их измерения применяют специальные методы статистики, которые находят выражение в трех основных стадиях статистического исследования.

Первая стадия – массовое статистическое наблюдение. На этой стадии в результате регистрации фактов по определенной разработанной программе получают объективные данные об изучаемых общественных явлениях.

Вторая стадия – свodka и обработка, классификация и систематизация материалов, собранных в результате массового статистического наблюдения. На этой стадии применяется метод группировок, который дает возможность выявить различные социальные типы явлений путем выделения характерных групп и подгрупп по существенным признакам и подсчитать итоги по каждой из них.

Третья стадия – анализ показателей, полученных в результате сводки и обработки статистических материалов. Цель анализа – выявить тенденции, закономерности и связи в изучаемых социальных процессах, дать характеристику типичных их черт. Для этого применяются методы обобщающих показателей (суммарных, относительных, средних величин, индексов), оценки вариаций признаков и динамических рядов, дисперсионного, корреляционного, регрессионного анализа.

3.2. Подготовка статистического наблюдения

Подготовка статистического наблюдения — процесс, включающий разные виды работ. Сначала необходимо решить методологические вопросы, важнейшими из которых являются определение цели и объекта наблюдения, состава признаков, подлежащих регистрации; разработка документов для сбора данных; выбор отчетной единицы и единицы, относительно которой будет проводиться наблюдение, а также методов и средств получения данных.

Кроме методологических необходимо решить проблемы организационного характера, например, определить состав служб, проводящих наблюдение; подобрать и подготовить кадры для проведения наблюдения; составить календарный план работ по подготовке, прове-

дению и обработке материалов наблюдения; провести тиражирование документов для сбора данных.

Проведение массового сбора данных включает работы, связанные непосредственно с заполнением статистических формуляров. Он начинается с рассылки переписных листов, анкет, бланков, форм статистической отчетности и заканчивается их сдачей после заполнения в органы, проводящие наблюдение.

Собранные данные на этапе их подготовки к автоматизированной обработке подвергаются арифметическому и логическому контролю. Оба эти контроля основываются на знании взаимосвязей между показателями и качественными признаками.

На заключительном этапе проведения наблюдения анализируются причины, которые привели к неверному заполнению статистических бланков, и разрабатываются предложения по совершенствованию наблюдения. Это очень важно для организации будущих наблюдений.

Получение сведений в ходе статистического наблюдения требует немалых затрат финансовых и трудовых ресурсов, а также времени.

Цель наблюдения. Статистическое наблюдение чаще всего преследует практическую цель — получение достоверной информации для выявления закономерностей развития явлений и процессов. Например, целью микропереписи населения России в 1994 г. было получение данных о численности, составе населения, условиях его проживания.

Задача наблюдения предполагает его программу и формы организации. Неясно поставленная цель может привести к тому, что в процессе наблюдения будут собраны ненужные данные или, наоборот, не будут получены сведения, необходимые для анализа.

Объект и единица наблюдения. Отчетная единица. При подготовке наблюдения кроме цели следует точно определить, что именно подлежит обследованию, т.е. установить объект наблюдения.

Под **объектом наблюдения** понимается некоторая статистическая совокупность, в которой протекают исследуемые социально-экономические явления и процессы. Объектом наблюдения может быть совокупность физических лиц (население отдельного региона, страны; лица, занятые на предприятиях отрасли), физические единицы (станки, машины, жилые дома), юридические лица (предприятия, фермерские хозяйства, коммерческие банки, учебные заведения).

Всякий объект статистического наблюдения состоит из отдельных элементов — единиц наблюдения.

В статистике **единицей наблюдения** (в зарубежной литературе используется термин «элементарная единица») называют составной

элемент объекта, являющийся носителем признаков, подлежащих регистрации. Например, при демографических обследованиях единицей наблюдения может быть человек, но может быть и семья; при бюджетных обследованиях — семья или домашнее хозяйство.

Единицу измерения следует отличать от отчетной единицы. *Отчетной единицей* выступает субъект, от которого поступают данные о единице наблюдения. Так, при организации статистического наблюдения в капитальном строительстве информация может быть получена от проектных или подрядных организаций или от предприятий-застройщиков.

Единица наблюдения и отчетная единица могут совпадать. Например, если надо определить объем освоенных за год капитальных вложений, то предприятие-застройщик будет одновременно и единицей наблюдения, и отчитывающейся организацией. Однако, при изучении процесса концентрации капитальных вложений отчетной единицей будет по-прежнему застройщик, а единицей наблюдения — стройки и объекты, строительство которых ведет данный застройщик.

Программа статистического наблюдения. Всякое явление обладает множеством различных признаков. Собирать информацию по всем признакам нецелесообразно, а часто и невозможно. Поэтому необходимо отобрать те признаки, которые являются существенными, основными для характеристики объекта исходя из целей исследования. Для определения состава регистрируемых признаков разрабатывают программу наблюдения.

Программа наблюдения — это перечень признаков (или вопросов), подлежащих регистрации в процессе наблюдения. От того, насколько хорошо разработана программа статистического наблюдения, во многом зависит качество собранной информации.

Чтобы составить правильно программу наблюдения, исследователь должен ясно представлять задачи обследования конкретного явления или процесса, определить состав используемых в анализе методов, необходимые группировки и уже на основе этого выявить те признаки, которые нужно определить при проведении работы. Обычно программа выражается в форме вопросов переписного (прописного) листа.

Для обеспечения единообразия получаемых сведений от каждой отчетной единицы (это важно при последующей обработке информации) программа оформляется в виде документа, называемого статистическим формуляром.

Статистический формуляр — это документ единого образца, содержит программу и результаты наблюдения.

Обязательными элементами статистического формуляра является титульная и адресная части. Первая содержит наименование статистического наблюдения и органа, проводившего наблюдение, информацию о том, кто и когда утвердил этот формуляр, иногда его номер. Вторая содержит адрес отчетной единицы, ее подчиненность.

Формуляр может иметь разные названия: отчет, карточка, переписной лист, опросный бланк, анкета и т.д.

Кроме формуляра разрабатывается *инструкция*, определяющая порядок проведения наблюдения и заполнения формы отчетности, переписного листа, анкеты. В зависимости от сложности программы наблюдения инструкция публикуется в виде отдельной брошюры или помещается на обратной стороне формуляра. Формуляр и инструкция по его заполнению составляют инструментарий статистического наблюдения.

Место и время наблюдения. Выбор места проведения обследования зависит главным образом от цели наблюдения. Если необходимо получить данные для изучения состава населения по стране, то в этом случае наблюдение охватит территорию всей страны.

Выбор времени наблюдения заключается в решении двух вопросов:

- установление критического момента (даты) или интервала времени;
- определение срока (периода) наблюдения.

Под *критическим моментом* (датой) понимаются конкретный день года, час дня, по состоянию на который должна быть проведена регистрация признаков по каждой единице исследуемой совокупности.

Выбор критического момента или интервала времени определяется прежде всего целью исследования.

Срок (период) наблюдения — это время. В течение которого происходит заполнение статистических формуляров, т.е. время, необходимое для проведения массового сбора данных. Этот срок определяется исходя их объема работы.

3.3. Формы, виды и способы статистического наблюдения

Первым этапом статистического исследования является *статистическое наблюдение*, представляющее собой планомерный, организованный сбор первичных данных. Существуют различные органи-

зационные формы наблюдения, можно выделить различные способы и виды статистических наблюдений (рис. 3.1).

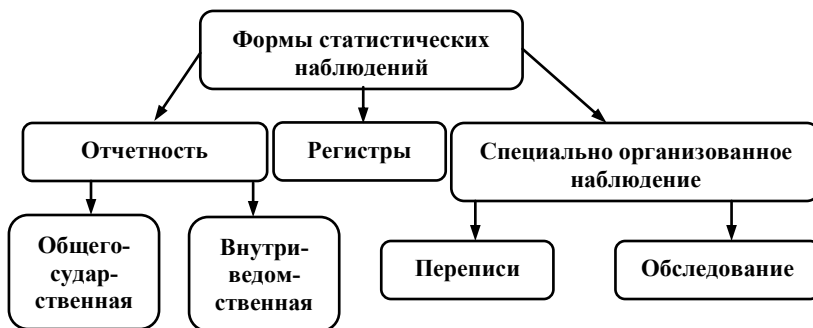


Рис. 3.1 Формы статистических наблюдений

Регистровая форма наблюдения. Регистровое наблюдение — это форма непрерывного статистического наблюдения за долговременными процессами, имеющими фиксированное начало, стадию развития и фиксированный конец. Оно основано на ведении статистического регистра. Регистр представляет собой систему, постоянно следящую за состоянием единицы наблюдения и оценивающую силу воздействия различных факторов на изучаемые показатели. В регистре каждая единица наблюдения характеризуется совокупностью показателей. Одни из них остаются неизменным в течение всего времени наблюдения и регистрируются один раз; другие показатели, периодичность изменения которых неизвестна, обновляются по мере изменения; третьи — представляют собой динамические ряды показателей с заранее известным периодом обновления. Все показатели хранятся до полного завершения наблюдения за единицей обследуемой совокупности.

В практике статистики различают регистры населения и регистры предприятий.

Регистр населения — поименованный и регулярно актуализируемый перечень жителей страны. Программа наблюдения ограничена общими признаками, такими, как пол, дата и место рождения, дата вступления в брак (эти данные остаются неизменными в течение всего периода наблюдения) и брачное состояние (переменный признак). Как правило, регистры хранят информацию только по тем переменным признакам, изменение значения которых документально оформлено.

Информация в регистр заносится на каждого родившегося и прибывшего из-за границы. Если человек умер или выехал на постоянное ме-

сто жительства из страны, то сведения о нем изымаются из регистра. Регистры населения ведутся по отдельным регионам страны. При перемещении места жительства сведения по единице наблюдения передаются в регистр соответствующей территории. В связи с тем, что правила регистрации довольно сложны и ведение регистра требует больших затрат, эта форма наблюдения практикуется в странах с небольшой численностью и высокой культурой населения (в основном это европейские страны).

Регистр предприятий включает в себя все виды экономической деятельности и содержит значения основных признаков по каждой единице наблюдаемого объекта за определенный период или момент времени. Регистры предприятий содержат данные о времени создания (регистрации) предприятия, его название и адрес, телефон, об организационно-правовой форме, структуре, виде экономической деятельности, количестве занятых (этот показатель отражает размер предприятия) и др.

В нашей стране были разработаны три регистра: промышленных предприятий,строек и подрядных организаций. Внедрение их в статистическую практику существенно повысило информационный и аналитический уровень статистики, позволило решить ряд экономико-статистических задач, для которых непригодны другие формы статистического наблюдения.

В настоящее время завершены работы по созданию единого регистра для всех хозяйственных единиц. Ему отводится важное значение во внедрении системы национальных счетов в статистическую практику.

Единый государственный регистр предприятий и организаций всех форм собственности (ЕГРПО) дает возможность организовать сплошное наблюдение по ограниченному кругу статистических показателей предприятий, зарегистрированных на территории России, позволяет получить непрерывные ряды показателей в случае изменения территориальной, отраслевой и других структур совокупности.

В регистр заносятся данные по всем предприятиям, организациям, учреждениям и объединениям независимо от их форм собственности, включая предприятия с иностранными инвестициями, банковские учреждения, общественные объединения и другие юридические лица.

Информационный фонд регистра содержит: во-первых, регистрационный код субъекта; во-вторых, сведения об отраслевой, территориальной принадлежности субъекта, его подчиненности, виде собственности, организационной форме; в-третьих, справочные сведения (фамилии руководителей, адреса, номера телефонов, факсов и т.д., сведения об учредителях) и, наконец, в-четвертых, экономические показатели. Значения последних будут заноситься в регистр на основе бухгалтерской и статистической отчетности, представляемых в региональные органы статистики.

Регистр содержит данные о следующих показателях: среднесписочная численность работников; средства, направляемые на потребление; остаточная стоимость основных средств; балансовая прибыль (убыток); уставной фонд. Так как регистр ведется по отдельным территориям, региональные статистические службы могут расширить состав экономических показателей в случае необходимости.

ЕГРПО позволяет производить отбор и группировку любой совокупности единиц по одному или нескольким признакам.

Сбор данных об единицах наблюдения осуществляется в процессе их государственной регистрации и последующего учета.

При закрытии предприятия ликвидационная комиссия в десятидневный срок уведомляет об этом службу ведения регистра.

Пользователями регистра могут быть любые юридические или физические лица, заинтересованные в получении информации.

Статистическая отчетность – форма наблюдения, при которой статистические органы в установленные сроки получают от предприятий и учреждений необходимые данные в виде установленных в законном порядке отчетных документов, скрепленных подписями лиц, ответственных за их представление и достоверность сообщаемых сведений.

Статистическая отчетность, представляемая организациями и учреждениями, отражает результаты их деятельности и носит периодический характер (ежегодная, квартальная, месячная). Отчетность по формам, не утвержденным в установленном порядке, является незаконной.

Совершенствование государственной статистики (в том числе и социальной) проводится в направлении упрощения форм отчетности, изменения системы показателей, автоматизации сбора, хранения и обработки учетно-статистической информации.

При изучении явлений общественной жизни, а в социальной статистике в особенности, возникает необходимость прибегать к **специально организованным статистическим наблюдениям**. Специально

организованные наблюдения (переписи и обследования) проводятся органами государственной статистики в тех случаях, когда возникает необходимость иметь более обширную и содержательную информацию о тех сторонах социальных процессов и явлений, которые недостаточно или вообще не получают отражения в статистической отчетности.

Например, в соответствии с Федеральной программой статистических работ со 2 по 22 ноября 1998г. Областной комитет государственной статистики Томской области проводил опрос населения по проблемам занятости и безработицы. В Томске было обследовано 800 человек в возрасте от 15 до 72 лет по специальной анкете. Такие обследования проводятся регулярно.

Особой формой специально организованного статистического наблюдения являются ***переписи*** (не только населения, но и производственных единиц, жилья и т.п.).

Основные принципы проведения переписей:

- полный охват изучаемого явления;
- установление критического момента переписи и правильный его выбор;
- единство программы и приемов переписи;
- централизованное руководство всей переписью;
- обеспечение специального контроля за полнотой учета;
- периодичность переписи (проведение ее через равные интервалы времени);
- преемственность программы (сохранение по возможности характера вопросов – для обеспечения сравнимости).

Статистическое наблюдение делится на виды по частоте регистрации и по полноте охвата единиц наблюдения (рис. 3.2).

Непрерывное (текущее) наблюдение ведется постоянно, систематически по мере возникновения явлений. Например, в загсе регистрируются рождения, смерти, браки, разводы.

Периодическое наблюдение – регистрация проводится через определенные промежутки времени (ежегодно, ежемесячно и т.п.).

Единовременное наблюдение проводится один раз для решения конкретной задачи или повторяется через неопределенные промежутки времени по мере необходимости.

При **сплошном наблюдении** регистрации подлежат все без исключения единицы совокупности.

Несплошное наблюдение предполагает отбор единиц совокупности для регистрации по способам основного массива, выборочному или монографическому.

При **способе основного массива** обследованию подвергается та часть единиц совокупности, которая вносит наибольший вклад в изу-



Рис. 3.2 Виды статистических наблюдений

чаемое явление. Например, при анализе структуры продовольственного рынка можно наблюдать только крупные оптовые фирмы – поставщики сахара, а не все мелкие.

Часто применение метода основного массива требует установления **ценза** – значения признака, которое ограничивает объект наблюдения. Например, возраст человека.

Заметим, что в США и Англии цензами называют переписи населения, промышленности и т.д.

При монографическом наблюдении подробно описываются отдельные единицы совокупности в целях их углубленного изучения.

Главное внимание обращается на качественные стороны явления, его поведение и т.д. Например, этнографические обследования.

Наиболее распространенным видом наблюдения в статистике (и социологии при проведении конкретных социологических исследований) является выборочное наблюдение.

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

Совокупность, из которой производится отбор, называется генеральной, и все обобщающие показатели – генеральными.

Совокупность отобранных единиц именуют выборочной совокупностью, и все ее обобщающие показатели – выборочными.

При любых статистических исследованиях возникают ошибки двух видов: регистрации и репрезентативности.

Ошибки регистрации могут иметь **случайный** и **систематический** характер. **Случайные ошибки** обычно уравнивают друг друга поскольку не имеют преимущественного направления в сторону преувеличения или преуменьшения значения изучаемого показателя. **Систематические ошибки** направлены в одну сторону вследствие преднамеренного нарушения правил отбора.

Ошибки репрезентативности присущи только выборочному наблюдению и возникают в силу того, что выборочная совокупность не полностью воспроизводит генеральную. Они представляют собой расхождение между значениями показателей, полученных при выборке и значениями показателей этих же величин, которые были бы получены при проведенном с одинаковой степенью точности сплошном наблюдении, т.е. между величинами выборочных и соответствующих генеральных показателей.

При любом обследовании источником получения первичных данных могут быть непосредственное наблюдение, документы и опрос, которые представляют собой **способы статистического наблюдения** (рис. 3.3).

Непосредственное наблюдение осуществляется путем регистрации изучаемых единиц и их признаков на основе непосредственного осмотра, подсчета, взвешивания и т.д. Например, регистрация цен на рынках.

Документальный способ – использование в качестве источника различных документов первичного учета предприятий, учреждений и организаций. Например, данные бухгалтерского учета.

При **опросе** источником данных являются сведения, которые дают опрашиваемые лица. Отметим, что при этом способе наблюдения точность обычно ниже, чем при непосредственном или документальном способах.

Экспедиционный способ заключается в том, что специально подготовленные регистраторы (счетчики, интервьюеры – в социологических исследованиях) заполняют переписные формуляры (вопросники, анкеты), одновременно контролируя (по возможности) правильность получаемых ответов. Это достаточно дорогостоящий способ. Используется в переписях.



Рис. 3.3 Виды статистических наблюдений

Способ саморегистрации отличается от экспедиционного тем, что формуляры (анкеты) заполняются самим опрашиваемым лицом, но процесс раздачи, заполнения и сбора анкет контролируется интервьюером. Этот способ используется, например, в государственной статистике при бюджетных обследованиях семей.

Корреспондентский способ заключается в том, что статистические организации рассылают специально разработанные бланки и ин-

струкции к их заполнению организациям или лицам, давшим согласие периодически заполнять бланки и присылать их в статистические органы в установленные сроки. Например, по такому способу функционирует потребительская панель.

Явочный способ предусматривает представление сведений в явочном порядке, например, при регистрации браков, разводов.

Тест для самоконтроля

1. Критический момент переписи - это
 - а) время, в течение которого происходит заполнение переписного листа
 - б) момент окончания переписи
 - в) момент, к которому приурочены регистрируемые события
2. Перепись населения - это
 - а) единовременное, специально организованное, сплошное наблюдение
 - б) периодическое, специально организованное, сплошное наблюдение
 - в) политическая акция по учету населения
3. Единицей наблюдения в статистике является:
 - а) перечень вопросов формуляра
 - б) социально-экономическое явление, подлежащие изучению
 - в) первичный элемент совокупности, являющийся носителем признаков, подлежащих регистрации
4. Единовременное наблюдение проводится
 - а) только один раз
 - б) периодически, через равные промежутки времени
 - в) один раз или повторяется через неопределенные промежутки времени
5. Перепись населения проводилась в период с 15 по 22 января. Критическим моментом было 12 часов ночи с 14 на 15 января. Счетчик пришел в семью 20 января. В этой семье 14 января родился ребенок. Счетчик должен
 - а) внести данные ребенка в переписной лист с отметкой о дате рождения

- б) не заполнять переписного листа на ребенка
 - в) внести сведения о ребенке в переписной лист родителей
6. С точки зрения статистического исследования ЕГРПО - это:
- а) внутриведомственная отчетность
 - б) единовременное специально организованное наблюдение
 - в) сплошное непрерывное наблюдение
7. Субъект, от которого поступают данные в ходе статистического наблюдения, называется
- а) отчетная единица
 - б) единица статистической совокупности
 - в) единица наблюдения

Вопросы и упражнения

1. В чем отличие статистического наблюдения от статистического исследования?
 2. Дайте характеристику статистической отчетности.
 3. Сформулируйте и поясните основные принципы проведения переписей.
 4. Проведите классификацию видов статистического наблюдения.
 5. Охарактеризуйте способы статистического наблюдения.
 6. Определите форму, вид и способ статистических наблюдений:
 - а) годовой баланс предприятия;
 - б) перепись населения;
 - в) мониторинг цен на рынке.
- Дайте пояснения.
7. Что такое «репрезентативность»?

4. АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ СТАТИСТИЧЕСКИЕ ВЕЛИЧИНЫ

4.1. Абсолютные статистические величины

Абсолютными в статистике называются суммарные обобщающие показатели, характеризующие размеры (уровни, объемы) общественных явлений в конкретных условиях места и времени. Например, численность населения Российской Федерации на январь 2009 г. составила 141,9 млн. человек; ввод в действие общей площади жилых домов за 2008 год составил 64,1 млн. м². ВВП России составил в текущих ценах 62356,9 трлн. руб.

Различают два вида абсолютных статистических величин: индивидуальные и суммарные.

Индивидуальными называют абсолютные статистические величины, характеризующие размеры признака у отдельных единиц совокупности (например, размер заработной платы отдельного работника, вклада гражданина в определенном банке и т.д.) Они получают непосредственно в процессе статистического наблюдения и фиксируются в первичных учетных документах.

В отличие от индивидуальных *суммарные* абсолютные статистические величины характеризуют итоговое значение признака по определенной совокупности объектов, охваченных статистическим наблюдением. Они являются суммой количества единиц совокупности (численность совокупности) или суммой значений варьирующего признака всех единиц совокупности (объем варьирующего признака).

Абсолютные статистические величины представляют собой *именованные числа*, т.е. имеют какую-либо единицу измерения.

В зависимости от сущности исследуемого социально-экономического явления абсолютные статистические величины выражаются в натуральных (тонны, штуки и т.п.), стоимостных (рубли) и трудовых (человеко-дни) *единицах измерения*. Абсолютные статистические величины могут быть положительными (доходы) и отрицательными (убытки, потери).

4.2. Относительные статистические величины

Относительная величина в статистике – это обобщающий показатель, который представляет собой частное от деления одного

абсолютного показателя на другой и дает числовую меру соотношений между ними.

Основные условия правильного расчета относительной величины – сопоставимость сравниваемых показателей и наличие реальных связей между изучаемыми явлениями.

Величина, с которой производится сравнение (знаменатель дроби), обычно называется *базой сравнения* или *основанием*.

Относительная величина динамики (i) рассчитывается как отношение уровней признаков в определенные периоды или моменты времени, т.е. она характеризует изменение уровня какого-либо явления во времени. Выбор базы сравнения при исчислении относительных показателей динамики определяется целью исследования. Относительные величины динамики называются *темпами роста*.

Относительная величина планового задания ($i_{пл.з}$) рассчитывается как отношение уровня, запланированного на предстоящий период, к уровню, фактически сложившемуся в базовом периоде.

Относительная величина выполнения планового задания ($i_{вып.пл}$) представляет собой отношение фактически достигнутого в данном периоде уровня к запланированному.

Относительные величины динамики, планового задания и выполнения планового задания связаны соотношением:

$$i = i_{пл.з} \cdot i_{вып.пл}.$$

Относительными величинами структуры называются показатели, характеризующие долю отдельных частей изучаемой совокупности во всем ее объеме. Они рассчитываются делением числа единиц в отдельных частях совокупности на общее число единиц совокупности. Выражаются они простым кратным отношением или в процентах.

Относительными величинами интенсивности называют показатели, характеризующие степень распространения или уровень развития того или иного явления в определенной среде. Они вычисляются путем сравнения разноименных величин, находящихся в определенной связи между собой. Например, показатель «обеспеченность населения холодильниками в расчете на 100 семей».

Относительными величинами координации называют показатели, характеризующие соотношение отдельных частей целого между собой. Вычисление этого вида показателей производится путем деления одной части целого на другую часть целого. В зависимости от поставленной задачи тот или иной признак может быть принят за базу. Поэтому для одной и той же совокупности можно исчислить несколько относительных показателей координации.

Относительными величинами сравнения называют показатели, представляющие собой частное от деления одноименных абсолютных статистических величин, характеризующих разные объекты (предприятия, фирмы, районы, области, страны и т.д.) и относящихся к одному и тому же периоду времени. Например, соотношение между уровнями себестоимости определенного вида продукции, выпущенной на двух предприятиях, между уровнями производительности труда в разных странах (при одинаковой методике расчетов).

Тест для самоконтроля

1. По данным газеты Frankfurter Allgemeine европейские государства продолжают наращивать долги - Германия должна 2,15 трлн евро, что соответствует 81,2% ее ВВП. В данном примере величина 2,15 трлн евро - это

- а) абсолютная статистическая величина
- б) относительная величина динамики
- в) относительная величина структуры
- г) относительная величина наглядности (сравнения)

2. Зная значения относительных величин динамики и планового задания, находят относительную величину

- а) структуры
- б) динамики
- в) координации
- г) выполнения плана

3. Относительная величина динамики показателя «объем производства» равна 1,03. Это означает

- а) снижение объема производства на 3 %
- б) рост объема производства на 3 %
- в) рост производства на 30 %
- г) рост объема производства на 103%

4. По предприятию за I квартал имеются следующие данные.

Показатель	План	Отчет
Выпуск валовой продукции, тыс. руб.	1800	2016

Плановое задание

- а) перевыполнено на 112%
- б) перевыполнено на 1,12 раза
- в) перевыполнено на 12%

5. Имеются данные о численности и составе студенческих групп

Пол	Число студентов	
	Группа 1	Группа 2
Юноши	15	27
Девушки	10	18

В какой группе доля юношей в общей численности группы больше

- а) в группе 1
- б) доля равны
- в) в группе 2

6. Выпуск продукции по плану должен был увеличиться по сравнению с предыдущим периодом на 10%, план перевыполнен на 10%. Фактическое увеличение выпуска продукции по сравнению с предыдущим периодом составило

- а) 21%
- б) 20%
- в) 100%
- г) в 2 раза

Вопросы и упражнения

1. Дайте определение абсолютной статистической величины. Приведите примеры индивидуальных и суммарных абсолютных величин.

2. Перечислите все возможные атрибуты показателя «фонд заработной платы предприятия».

3. Производство станков в 2001 году вырастет в два раза по сравнению с показателем 2000 года. Какая относительная величина применена в этом тезисе?

4. Покажите, что

$$i = i_{нл.з} \cdot i_{вып.нл.}$$

5. Укажите тип относительных величин в случаях:

- а) «треть студентов группы составляют юноши»;
- б) «в студенческой группе девушек вдвое больше юношей».

6. Из 44 студентов специальности 0610 сдали сессию на «отлично» 12 человек, а из 142 студентов специальности 2202 — 32 человека. Вычислить все возможные относительные величины.

7. Запишите определения относительных величин в виде формул.

5. СРЕДНИЕ ЗНАЧЕНИЯ И ПОКАЗАТЕЛИ ВАРИАЦИИ

5.1. Вариационные ряды

Вариацией значений какого-либо признака в совокупности называется различие его значений у разных единиц данной совокупности в один и тот же период или момент времени.

Отдельное конкретное значение варьирующего признака называют вариантом (или вариантой).

В отличие от вариации различия значений признака у одного и того же объекта, у одной и той же единицы совокупности в разные моменты или периоды времени следует называть *изменениями во времени* или *колебаниями*. Методы их измерения и анализа отличаются от методов измерения вариации.

Причиной вариации являются разные условия существования разных единиц совокупности, влияние самых разнообразных факторов. Вариация присуща всем явлениям природы и общества, кроме законодательно закрепленных нормативных значений отдельных социальных признаков: не варьирует признак «число президентов республики» – все республики имеют значение этого показателя равное единице, что закреплено в конституциях. Предметом изучения статистики является вариация, ее измерение и анализ.

Первый этап статистического изучения вариации состоит в построении вариационного ряда. *Вариационный ряд* – совокупность значений варьирующего признака и соответствующих им численностей единиц совокупности. Он может быть дискретным (прерывным) или интервальным (непрерывным). Если варианты расположены в возрастающем или убывающем порядке вариационный ряд называется ранжированным или упорядоченным.

Число элементов (групп) в дискретном вариационном ряду определяется числом реально существующих значений варьирующего признака. Если же признак принимает хотя и дискретные значения, но их число очень велико (например, численность населения городов), тогда строится интервальный вариационный ряд. При построении интервального вариационного ряда необходимо выбрать оптимальное число групп (интервалов признака) и установить длину интервала.

Интервал – значения варьирующего признака, лежащие в определенных границах. Различают интервалы закрытые (замкнутые) с конкретными численными значениями обеих границ интервала

(например, заработная плата работников в границах «от 100 до 110 руб.») и открытые, в которых нижняя или верхняя граница не фиксирована (например, «до 100 руб.» или «свыше 200 руб.»).

Величина интервала (интервальная разность) есть разность между верхней $x_{i \max}$ и нижней $x_{i \min}$ границами интервала

$$l_i = x_{i \max} - x_{i \min}.$$

Различают интервальные вариационные ряды с равными интервалами, в которых все интервалы имеют одну и ту же величину, и интервальные вариационные ряды с неравными интервалами. Неравные интервалы позволяют выделить типы явлений.

Для расчета оптимальной величины интервала рекомендуется приближенная формула Стерджесса (Sturges):

$$z = \frac{x_{\max} - x_{\min}}{1 + 3,32 \lg n},$$

где n – число единиц совокупности;

$x_{\max}$ и $x_{\min}$ – наибольший и наименьший варианты;

$1 + 3,32 \lg n$ – рекомендуемое Стерджессом число интервалов.

Следует учитывать также, что для построения ряда и анализа удобнее иметь по возможности округленные значения границ интервалов.

Центр интервала (среднее значение) – это полусумма граничных значений для каждого интервала вариационного ряда.

$$x_{i \text{ центр}} = \frac{x_{i \max} + x_{i \min}}{2}.$$

Вариационные ряды обычно представляют в виде статистических таблиц. Для наглядного представления вариационного ряда используют графические изображения в виде полигона, гистограммы, кумуляты, кривой концентрации (Лоренца), огивы и т.д.

5.2. Виды средних величин

Индивидуальные значения одного и того же признака у различных единиц совокупности неодинаковы. В процессе обработки и обобщения статистических данных возникает необходимость в определении средних величин.

Средняя величина — обобщающий показатель, характеризующий типический уровень признака в расчете на единицу совокупности в конкретных условиях места и времени.

Средняя величина обобщает количественную вариацию признака, т.е. в средних величинах погашаются индивидуальные различия единиц совокупности, обусловленные случайными обстоятельствами. Средние величины применяются для оценки достигнутого уровня изучаемого показателя, при расчете нормативов, при анализе и планировании. Средняя величина всегда именованная, имеет ту же размерность, что и признак у отдельных единиц совокупности.

Средняя величина должна вычисляться по однородной совокупности. Если совокупность неоднородна, то **общие средние**, рассчитанные по всей совокупности, должны подкрепляться **групповыми средними**, дающими характеристику размера явления, складывающуюся в конкретных условиях данной группы. Средние величины делятся на две большие категории:

- степенные средние (к ним относятся средняя арифметическая, средняя гармоническая, средняя геометрическая, средняя квадратическая);
- структурные средние (мода и медиана).

5.3. Степенные средние

Степенные средние в зависимости от представления исходных данных могут быть простыми и взвешенными. Простая средняя рассчитывается по несгруппированным данным и имеет следующий вид:

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

где x_i — значение осредняемого признака у i -й единицы совокупности;

m — показатель степени средней;

n — число единиц совокупности.

Если данные сгруппированы, то рассчитывается взвешенная средняя по общей формуле

$$\bar{X} = \sqrt[m]{\frac{\sum x_i^m f_i}{\sum f_i}},$$

где x_i — значение осредняемого признака или срединное значение интервала, в котором измеряется значение признака;

m — показатель степени средней;

f_i — частота, показывающая сколько раз встречается i -е значение осредняемого признака.

В зависимости от значения показателя степени m различают следующие виды степенных средних:

- средняя гармоническая, если $m = -1$;
- средняя геометрическая, если $m = 0$;
- средняя арифметическая, если $m = 1$;
- средняя квадратическая, если $m = 2$;
- средняя кубическая, если $m = 3$.

С увеличением значения показателя степени m увеличивается величина степенной средней, вычисленная для одних исходных данных.

Рассмотрим средние величины, наиболее часто встречаемые в статистической практике.

Средней арифметической величиной называется такое среднее значение признака, при вычислении которого общий объем признака в совокупности сохраняется неизменным. При ее вычислении общий объем признака мысленно распределяется поровну между всеми единицами совокупности.

Рассмотрим пример. Известны размеры месячной заработной платы рабочих бригады за май 1999 года:

№ п/п	Месячная зарплата, руб.
1.	650
2.	650
3.	650
4.	800
5.	800
6.	800
7.	800
8.	950
9.	1000
10.	1000

Среднюю заработную плату считаем по формуле простой средней

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_{10}}{10} = \frac{8100}{10} = 810 \text{ руб.}$$

Сгруппируем исходные данные. Получим следующий ряд распределения – дискретный вариационный ряд:

Месячная зарплата, руб.	1000	950	800	650	Всего
Число рабочих	2	1	4	3	10

В результате группировки получаем новый показатель — частоту f_i , указывающий число рабочих с зарплатой x_i рублей. Следовательно, средняя заработная плата должна рассчитываться по формуле взвешенной средней

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{1000 \cdot 2 + 950 \cdot 1 + 800 \cdot 4 + 650 \cdot 3}{10} = \frac{8100}{10} = 810 \text{ руб.}$$

Средняя арифметическая величина обладает следующими свойствами:

- сумма отклонений отдельных значений признака от средней арифметической равна нулю;
- если к каждому значению признака прибавить постоянное число, то средняя увеличится на это же число;
- если каждое значение признака умножить на постоянное число, то средняя увеличится во столько же раз;
- если все частоты разделить на какое-либо число, то средняя величина не изменится. Это свойство дает возможность заменять абсолютные значения частоты удельными весами.

Если по условиям задачи необходимо, чтобы неизменной оставалась при осреднении сумма величин, обратных индивидуальным значениям признака, то средняя величина является *гармонической средней*.

Формула гармонической средней величины такова:

$$\bar{x}_{\text{гарм}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

Для сгруппированных данных применяется средняя гармоническая взвешенная

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

Средняя гармоническая применяется, когда необходимые веса в исходных данных не заданы непосредственно, а входят множителем в один из имеющихся показателей.

Рассмотрим на примере порядок расчета и выбор формы средней величины.

На основании следующих данных по двум сельскохозяйственным предприятиям необходимо определить, в каком из них и насколько выше средняя урожайность зерновых культур:

Культура	Предприятие 1		Предприятие 2	
	Валовый сбор, ц	Урожайность, ц/га	Посевная площадь, га	Урожайность, ц/га
Пшеница	32500	25	1540	20
Рожь	1620	18	120	19
Ячмень	13640	22	460	18
Просо	1650	15	80	13
ИТОГО	49410	—	2200	—

Показатель урожайности является вторичным признаком, так как он задан на единицу первичного признака (посевной площади, выраженной абсолютной величиной) и может быть представлен в виде **логической формулы** как отношение первичных признаков:

$$\text{урожайность} = \frac{\text{валовый сбор}}{\text{посевная площадь}} \quad \text{или} \quad y = \frac{\text{вс}}{\text{пп}}.$$

Для предприятия № 1 среднюю урожайность определяем как среднюю гармоническую взвешенную

$$\begin{aligned} \bar{y}_1 &= \frac{\sum_{i=1}^4 \text{вс}_i}{\sum_{i=1}^4 \frac{\text{вс}_i}{y_i}} = \frac{32500 + 1620 + 13640 + 1650}{\frac{32500}{25} + \frac{1620}{18} + \frac{13640}{22} + \frac{1650}{15}} = \\ &= \frac{32500 + 1620 + 13640 + 1650}{1300 + 90 + 620 + 110} = \frac{49410}{2120} = 23,31 \text{ ц/га} \end{aligned}$$

Здесь слагаемые знаменателя — посевные площади, занятые под различные культуры.

Для предприятия 2 средняя урожайность определяется как средняя арифметическая взвешенная, причем в качестве весов берутся посевные площади:

$$\bar{Y}_2 = \frac{\sum_{i=1}^4 y_i \cdot \text{пп}_i}{\sum_{i=1}^4 \text{пп}_i} = \frac{20 \cdot 1540 + 19 \cdot 120 + 18 \cdot 460 + 13 \cdot 80}{1540 + 120 + 460 + 80} = \frac{42400}{2200} = 19,24 \text{ ц/га}$$

Таким образом, средняя урожайность зерновых культур на предприятии 1 по сравнению с предприятием 2 выше на 4,04 ц/га (или на 21 %).

Если при замене индивидуальных величин признака на среднюю величину необходимо сохранить неизменным произведение индивидуальных величин, то следует применить *геометрическую среднюю величину*. Ее формула такова:

$$\bar{x}_{\text{геом}} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \quad \text{или} \quad \bar{X} = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}.$$

Основное применение геометрической средняя находит при определении средних темпов роста. Пусть, например, в результате инфляции за первый год цена товара возросла в 2 раза к предыдущему году, а за второй год еще в 3 раза к уровню предыдущего года. Ясно, что за два года цена выросла в 6 раз. Каков средний темп роста цены за год? Арифметическая средняя здесь непригодна, ибо если за год цены возросли бы в $\frac{2+3}{2} = 2,5$ раза, то за два года цена возросла бы в $2,5 \times 2,5 = 6,25$ раза, а не в 6 раз. Геометрическая средняя дает правильный ответ: $\sqrt{6} = 2,45$ раза.

5.4. Структурные средние

Структурные средние — *медиана* и *мода* — характеризуют величину значения признака, занимающего определенное значение в ранжированном вариационном ряду.

Медиана (*Me*) — значение признака, которое делит упорядоченную последовательность его значений на две равные по численности

части. В итоге у одной половины единиц совокупности значение признака не превышает медианного уровня, а у другой — не меньше его.

Положение медианы определяется ее номером $N_{Me} = (n + 1) / 2$, где n — число единиц в совокупности.

В интервальном вариационном ряду для нахождения медианы применяется формула

$$Me = X_0 + h \cdot \frac{\left(\sum_{i=1}^k f_i \right) / 2 - f'_{Me-1}}{f_{Me}},$$

где Me — медиана;

X_0 — нижняя граница интервала, в котором находится медиана;

h — величина интервала;

k — число интервалов;

f'_{Me-1} — накопленная частота в интервале, предшествующем медианному;

f_{Me} — частота в медианном интервале.

Аналогично медиане вычисляются значения признака, делящие совокупность на четыре равные части — **квартили** Q . Второй квартиль совпадает с медианой, а первый и третий рассчитываются по формулам:

$$Q_1 = X_0 + h \cdot \frac{\left(\sum_{i=1}^k f_i \right) / 4 - f'_{Q_1-1}}{f_{Q_1}};$$

$$Q_3 = X_0 + h \cdot \frac{\left(3 \sum_{i=1}^k f_i \right) / 4 - f'_{Q_3-1}}{f_{Q_3}}.$$

Значения признака, делящие ряд на пять равных частей называются квинтилями, на десять частей — децилями, на сто частей — перцентилями.

Мода (Mo) — наиболее часто встречающееся значение признака в совокупности. В дискретном ряду мода определяется без вычислений как значение признака с наибольшей частотой. В интервальном ряду модальным является интервал с наибольшей частотой. Внутри него определяется **точечная мода**

$$M_0 = X_0 + h \cdot \frac{f_{M_0} - f_{M_0-1}}{(f_{M_0} - f_{M_0-1}) + (f_{M_0} - f_{M_0+1})},$$

где X_0 — нижняя граница модального интервала;

f_{M_0} — частота в модальном интервале;

f_{M_0-1} — частота в предыдущем интервале;

f_{M_0+1} — частота в следующем за модальным интервале;

h — величина интервала.

Вычислим структурные средние величины для данных из следующей таблицы:

Таблица Распределение хозяйств по урожайности зерновых культур

Группы хозяйств по урожайности x_i	Число хозяйств f_i	Середина интервала, ц/га x'_i	$x'_i f_i$	Накопленная частота f'_i
10–15	6	12,5	75,0	6
15–20	9	17,5	157,5	15
20–25	20	22,5	450,0	35
25–30	41	27,5	1127,5	76
30–35	26	32,5	845,0	102
35–40	21	37,5	787,5	123
40–45	14	42,5	595,0	137
45–50	5	47,5	237,5	142
50–55	1	52,5	52,5	143
ИТОГО	143		4327,5	

Медианным является среднее из 143 значений, т.е. семьдесят второе от начала ряда значение урожайности, оно находится в четвертом интервале.

$$Me = 25 + 5 \cdot \frac{72 - 35}{41} = 29,5 \text{ ц/га};$$

$$Q_1 = 25 + 5 \cdot \frac{35,75 - 35}{41} = 25,09 \text{ ц/га};$$

$$Q_3 = 35 + 5 \cdot \frac{107,25 - 102}{21} = 36,25 \text{ ц/га}$$

$$M_0 = 25 + \frac{41 - 20}{(41 - 20) + (41 - 26)} = 27,9 \text{ ц/га}$$

Средняя арифметическая взвешенная для этой таблицы вычисляется с учетом середины интервала

$$\bar{X} = \frac{\sum_{i=1}^k x'_i f_i}{\sum_{i=1}^k f_i} = \frac{4327,5}{143} = 30,3 \text{ ц/га}$$

где x'_i — середина интервала.

5.5. Показатели вариации

Для измерения вариации признака применяются различные абсолютные и относительные показатели. Простейшим из них является **размах** или **амплитуда вариации** — абсолютная разность между максимальным и минимальным значениями признака изучаемой совокупности

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

Точнее характеризуют вариацию показатели, учитывающие колеблемость всех значений признака.

$$\text{Среднее линейное отклонение } d = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \quad i = 1, 2, \dots, n.$$

$$\text{Если данные сгруппированы } d_{\text{гр}} = \frac{\sum_{j=1}^k |x_j - \bar{x}| f_j}{\sum_{j=1}^k f_j} \quad j = 1, 2, \dots, k,$$

где k — число групп;

f_j — число единиц совокупности j -й группы.

Например, для данных из таблицы среднее линейное отклонение $d = 6,85$ ц/га, т.е. в среднем урожайность в изучаемой совокупности хозяйств отклонялась от средней урожайности на 6,85 ц/га.

В математической статистике важную роль играют показатели **дисперсии** и **стандартного отклонения**.

$$\text{Дисперсия } D_x = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad \text{и} \quad D_{x_{\text{гр}}} = \frac{\sum_{j=1}^k (x_j - \bar{x})^2 f_j}{\sum f_j}.$$

Стандартное отклонение равно квадратному корню из дисперсии, т.е.

$$\sigma = \sqrt{D_x}.$$

Для данных из таблицы стандартное отклонение равно

$$\sigma = \sqrt{\frac{10 \cdot 185}{143}} = 8,44 \text{ ц/га.}$$

Для оценки интенсивности вариации и для сравнения ее в различных совокупностях и для разных признаков **необходимы относительные показатели вариации**. Они вычисляются как отношения абсолютных показателей силы вариации к средней арифметической величине признака:

- относительный размах вариации ρ : $\rho = R : \bar{x}$;
- относительное отклонение по модулю m : $m = d : \bar{x}$;
- коэффициент вариации v : $v = \sigma : \bar{x}$.

Для данных из таблицы:

$$\rho = 42,4 : 30,3 = 1,4 \text{ или } 140 \%;$$

$$m = 6,85 : 30,3 = 0,226 \text{ или } 22,6 \%;$$

$$v = 8,44 : 30,3 = 0,279 \text{ или } 27,9 \%.$$

Принято считать, что если коэффициент вариации более 30%, то совокупность является неоднородной и ее следует делить на группы.

Тест для самоконтроля

1. Число интервалов (групп) в вариационном ряду можно определить по формуле Стерджесса

- а) $k = 1 + \lg n$, где n – постоянное число
- б) $k = 1 + 3,32 \lg n$, где n – численность совокупности
- в) $k = 1 + 3,32 \ln m$, где m – численность совокупности

2. Величина равного интервала определяется по формуле

- а) $h_i = R/k$, где k – число интервалов, $R = x_{\max} - x_{\min}$
- б) $h_{i+1} = h_i + a$, где a – постоянная величина
- в) $h_{i+1} = h_i - a$, где a – постоянная величина

3. Наибольшее значение признака в интервале называется
- а) нижней границей интервала
 - б) центром интервала
 - в) верхней границей интервала
4. При уменьшении всех весов в 10 раз
- а) значение средней арифметической увеличится на 10
 - б) значение средней арифметической не изменится
 - в) значение средней арифметической уменьшится в 10 раз
5. Если все значения признака уменьшить в 4 раза, то средняя арифметическая
- а) не изменится
 - б) уменьшится в 4 раза
 - в) уменьшится в 16 раз
 - г) уменьшится в 2 раза

Вопросы и упражнения

1. Дайте определение вариационного ряда. Приведите пример.
2. В выборочном бюджетном обследовании собрана информация о доходах 1000 семей. Сколько интервалов, на Ваш взгляд, должен содержать вариационный ряд? Ответ обоснуйте.
3. Что такое средняя величина? Где применяются средние величины?
4. Каковы виды средних величин и методы их расчета?
5. В каких случаях используется средняя гармоническая?
 - а) когда неизвестен числитель исходного соотношения;
 - б) когда неизвестен знаменатель исходного соотношения.
6. Изменится ли средняя величина, если все ее веса изменить на некоторую постоянную величину?
 - а) изменится;
 - б) не изменится.
7. Дайте определение моды и медианы в статистике. Каковы особенности определения моды и медианы в дискретном и интервальном ряде распределения?

6. СТАТИСТИЧЕСКИЕ ГРУППИРОВКИ

6.1. Значение и сущность группировки

Русский статистик Д.П. Журавский (1810-1856 гг.) определил статистику как «счет по категориям». Действительно, среди бесконечного разнообразия явлений мы, как правило, улавливаем наличие некоторого конечного числа групп или типов.

Группировка – это распределение единиц по группам в соответствии со следующим принципом: различия между единицами, относенными к одной группе, должны быть меньше, чем между единицами, отнесенными к разным группам.

Группировка лежит в основе всей дальнейшей работы с собранной информацией. На основе группировки рассчитываются сводные показатели по группам, появляется возможность их сравнения, анализа причин различий между группами, изучения взаимосвязей между признаками. Если рассчитать сводные показатели только в целом по совокупности, то мы не сможем уловить ее структуры, роли отдельных групп, их специфики.

Однородность (гомогенность) данных является исходным условием их статистического описания и анализа – вычисления и интерпретации обобщающих показателей, построения уравнения регрессии, измерения корреляции, статистического вывода.

Таким образом, значение группировки состоит в том, что этот метод обеспечивает обобщение данных, представление их в компактном, обозримом виде. Кроме того, группировка создает основу для последующей сводки и анализа данных.

Сводные показатели для отдельных групп являются типичными и устойчивыми, если, во-первых, группировка проведена правильно, во-вторых, группы имеют достаточную численность. Первое условие связано с тем, что деление на группы далеко не всегда очевидно. Выполнение второго условия необходимо, так как только при достаточно большом числе единиц (не менее 5 единиц в группе) в сводных показателях взаимопогашаются случайные характеристики и проявляются закономерные, типичные.

Для решения задачи группировки нужно установить правила отнесения каждой единицы к той или иной группе. В эти правила входят определение тех характеристик (признаков), по которым будет производится группировка (так называемые группировочные признаки), и

их значений, отделяющих одну группу от другой (интервалы группировки).

При определении числа групп следует обращать внимание на то, чтобы в одну группу не попало бы свыше половины всех единиц совокупности и в средних группах было больше единиц, чем в крайних.

Часто число групп определяется опытным путем на основе перебора вариантов группировки, выявления такого варианта, который наилучшим образом позволяет увидеть различия между группами.

Группировка называется **простой (монотетической)**, если для ее построения используется один группировочный признак. Если группировка проводится по нескольким признакам, она называется **сложной (политетической)**. Обычно такая группировка проводится как комбинационная, т.е. группы, выделенные по одному признаку, подразделяются на подгруппы по другому признаку. Недостатком комбинационной группировки является дробление совокупности на большое число категорий, включающих малое число единиц, что, во-первых, снижает надежность групповых показателей и, во-вторых, делает данные труднообозримыми. Альтернативой является проведение многомерных группировок или многомерных классификаций.

Очевидно, что метод группировок тесно связан с представлением данных в виде групповых и комбинационных таблиц, а также с графическим представлением структуры совокупности, ее частей и соотношений между ними.

6.2. Виды группировок

Группировка проводится с целью установления статистических связей и закономерностей, построения описания объекта, выявления структуры изучаемой совокупности. Различия в целевом назначении группировки выражаются в существующей в отечественной статистике классификации группировок: типологические, структурные, аналитические.

Типологическая группировка служит для выделения социально-экономических типов. Этот вид группировок в значительной степени определяется представлениями экспертов о том, какие типы могут встретиться в изучаемой совокупности. Последовательность ее построения следующая:

- 1) называются те типы явлений, которые могут быть выделены;
- 2) выбираются группировочные признаки, формирующие описание типов;
- 3) устанавливаются границы интервалов;

4) группировка оформляется в таблицу, выделенные группы (на основе комбинации группировочных признаков) объединяются в намеченные типы, и определяется численность каждого из них.

Типологический метод группировки субъективен, зависит от квалификации экспертов. Однако, если объект исследования достаточно изучен, то этот метод дает хорошо интерпретируемые результаты.

Правильность проведения типологической группировки проверяется. Для этого рассчитываются сводные показатели по группам (средние, относительные величины). Если различие между группами статистически незначимо (по t-критерию Стьюдента, χ^2 -критерию и т.п.), то схема группировки должна быть пересмотрена.

Структурная группировка характеризует структуру совокупности по какому-либо одному признаку.

В таблице 6.1 приведены сведения о распределении населения по душевому доходу.

Таблица 6.1 Распределение населения РФ по среднему душевому денежному доходу (в процентах)

Среднедушевой денежный доход в мес., тыс. руб.	1995	1996
400 и менее	47,4	25,3
400,1 - 600,0	22,3	22,4
600,1 - 800,0	12,6	16,9
800,1 - 1000,0	7,1	11,5
1000,0 - 1200,0	4,0	7,6
1200,1 - 1600,0	3,9	8,4
1600,1 - 2000,0	1,5	3,9
Свыше 2000,0	1,2	4,0
Всего	100	100

Источник: Российский стат. ежегодник. 1997. – М.: Госкомстат России, 1997. – С. 139

В структурной группировке часто используются равные интервалы. Структурная группировка позволяет изучать интенсивность вариации группировочного признака. Можно изучать динамику структуры совокупности.

Пусть w_{i0} и w_{i1} - доли i -ой группы в период «0» и «1». Показатели среднего абсолютного изменения структуры:

$$d_{w_1-w_0} = \frac{\sum_{j=1}^k |w_{j1} - w_{j0}|}{k},$$

где k — число групп.

Другой сводный показатель структурных сдвигов строится на основе формулы стандартного отклонения:

$$S_{w_1-w_0} = \sqrt{\frac{\sum_{j=1}^k (w_{j1} - w_{j0})^2}{k}}.$$

Для таблицы $d_{w_1-w_0} = 5,5$, $S_{w_1-w_0} = 8,5$.

Если показатели структуры выразить в процентах, то оба показателя покажут, на сколько процентов в среднем различаются удельные веса отдельных оцениваемых групп сравниваемых структур. При отсутствии структурных сдвигов эти показатели равны нулю; их величина тем больше, чем значительнее абсолютные изменения удельных весов групп. При сравнении предполагается, что число групп в одном и другом периодах остается неизменным.

Деление группировок на типологические и структурные достаточно условно. Если задать, например, границы среднедушевого дохода, соответствующие определенным типам благосостояния, то группировку структурную можно с полным основанием назвать типологической.

Аналитическая группировка характеризует взаимосвязь между двумя и более признаками, один из которых рассматривается как результат, другой (другие) — как фактор (факторы).

Рассмотрим пример однофакторной аналитической группировки.

Таблица 6.2 Характеристика зависимости прибыли малых предприятий от оборачиваемости оборотных средств

Продолжительность оборота средств, в днях	Число малых предприятий	Середина интервала, дни	Средняя прибыль, млн. руб.	Изменение средней прибыли, млн. руб
x_i	n_i	x'_i	$\bar{y}_i$	$\bar{y}_i - \bar{y}_{i-1}$
20 – 30	6	25	14.57	-
31 – 50	8	40	12.95	-1.62
51 - 80	6	65	7.40	-5.55
Итого	20	43	11.77	

В данном примере:

x — оборачиваемость в днях (фактор);

y — прибыль (результат).

Данные группируются по признаку-фактору. Затем по каждой группе рассчитывается среднее значение. Задача состоит в том, чтобы

увидеть, есть связь между признаками или нет; прямая связь или обратная; линейная или нелинейная.

Сопоставим изменения средних значений результата с изменениями фактора. Чтобы эти изменения были сравнимыми надо делать группировку с равными интервалами или рассчитывать изменения результата на единицу изменения фактора.

В примере средняя прибыль изменяется от группы к группе, следовательно, связь между оборачиваемостью и прибылью есть. Причем, обратная: чем медленнее оборачиваются средства, тем меньше прибыль.

Рассчитаем, на сколько снижается прибыль при замедлении оборачиваемости:

$$1) b_{xy} = \frac{\bar{y}_3 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} = \frac{-1.62}{15} = -0.108 \text{ млн. руб./день};$$

$$2) b_{xy} = \frac{\bar{y}_3 - \bar{y}_2}{\bar{x}_3 - \bar{x}_2} = \frac{-5.55}{25} = -0.222 \text{ млн. руб./день}.$$

Полученные значения показывают величину снижения прибыли при замедлении оборачиваемости на 1 день. Здесь: $b_{xy} \neq \text{const}$, следовательно, связь нелинейная.

b_{xy} - показатели силы связи.

По аналитической группировке измеряют связь при помощи эмпирического корреляционного отношения. Он основан **на правиле сложения (разложения) дисперсии**: общая дисперсия D_y равна сумме средней из внутригрупповой и межгрупповой дисперсий:

Правило сложения дисперсий

$$D_y^2 = \overline{D_{yx}^2} + D_{\bar{y}x}^2$$

$$\text{или } \sum_j \sum_i (y_{ij} - \bar{y})^2 + \sum_j \sum_i (y_{ij} - \bar{y}_j)^2 + \sum_j (\bar{y}_j - \bar{y})^2 n_j.$$

Дисперсия внутри группы при относительном постоянстве признака-фактора возникает за счет других факторов (не связанных с изучением). Эта дисперсия называется остаточной

$$D_{yx}^2 = \frac{\sum_i^{n_j} (y_{ij} - \bar{y}_j)^2}{n_j}, \quad \begin{matrix} j = 1, 2, \dots, m \\ i = 1, 2, \dots, n, \end{matrix}$$

y_{ij} — значение признака для i -ой единицы в j -ой группе;

$\bar{y}_j$ — среднее значение признака в j -ой группе;

n_j — число единиц в j -ой группе;

m — число групп.

Внутригрупповые дисперсии объединяются в средней величине внутригрупповых дисперсий:

$$\overline{D_{yx}^2} = \frac{\sum_j^m D_{yx}^2 \cdot n_j}{\sum_j^m n_j}.$$

Межгрупповая дисперсия относится на счет изучаемого фактора, она называется факторной

$$D_{yx}^2 = \frac{\sum_j^m (\bar{y}_j - \bar{y})^2 \cdot n_j}{\sum_j^m n_j}.$$

Эмпирическое корреляционное отношение измеряет, какую часть общей колеблемости результативного признака вызывает изучаемый фактор. Отношение межгрупповой (факторной) дисперсии к общей дисперсии:

$$R^2 = \frac{D_{yx}^2}{D_y^2} - \text{коэффициент детерминации};$$

$$r = \sqrt{\frac{D_{yx}^2}{D_y^2}} - \text{эмпирическое корреляционное отношение}.$$

Коэффициент детерминации изменяется от 0 до 1.

Рассмотрим пример.

Пример.

С целью установления зависимости между урожайностью и сортом винограда в одном из хозяйств на основе выборки определили урожай на 10 кустах винограда.

Наименование сорта винограда	Число проверенных кустов	Урожай винограда с каждого куста, кг				
		куст №1	куст №2	куст №3	куст №4	куст №5
Сорт «А»	3	6	5	7	—	—
Сорт «Б»	5	7	6	8	5	9
Сорт «В»	2	9	7	—	—	—

Исчислите общую, межгрупповую и среднюю из групповых (частных) дисперсий. Определите связь между сортом и его урожайностью.

Решение. Если совокупность разбита на группы, то дисперсия признака D может быть определена как сумма межгрупповой дисперсии $D_{м.гр.}$ и средней из групповых дисперсий $\overline{D_i}$:

$$D = \overline{D_i} + D_{м.гр.}$$

$$\overline{D_i} = \frac{\sum (D_i f_i)}{\sum f_i},$$

где $D_i = \frac{\sum (x - \overline{x_i})^2}{f_i}$ — дисперсия признака в группе i (групповая дисперсия);

x — индивидуальное значение признака;

$\overline{x_i}$ — среднее значение признака в группе i ;

f_i — число наблюдений в группе i .

$$D_{м.гр.} = \frac{\sum (\overline{x_i} - \overline{x})^2 f_i}{\sum f_i},$$

где $\overline{x}$ — среднее значение признака в совокупности.

Отношение межгрупповой дисперсии к общей дает возможность измерить вариацию результативного признака за счет факторного, то есть признака, положенного в основание группировки, и тем самым судить о связи между изучаемыми признаками:

$$R^2 = \frac{D_{м.гр.}}{D},$$

где R^2 — коэффициент детерминации.

При $R^2 = 0$ связи нет, при $R^2 = 1$ — связь полная.

Применим правило сложения дисперсии к исходным данным.

1. Групповые средние, т.е. средняя урожайность по каждому сорту винограда равны:

$$\bar{x}_A = \frac{\sum x}{n} = \frac{6+5+7}{3} = 6 \text{ кг};$$

$$\bar{x}_B = \frac{7+6+8+5+9}{5} = 7 \text{ кг};$$

$$\bar{x}_B = \frac{9+7}{2} = 8 \text{ кг}.$$

2. Определим среднюю урожайность винограда по хозяйству:

$$\bar{x} = \frac{\sum (x_i \cdot f_i)}{\sum f_i} = \frac{6 \cdot 3 + 7 \cdot 5 + 8 \cdot 2}{10} = 6,9 \text{ кг}.$$

3. Определим групповую (частную) дисперсию урожайности для каждого сорта отдельно:

$$D_A = \frac{(6-6)^2 + (5-6)^2 + (7-6)^2}{3} = 0,667;$$

$$D_B = \frac{(7-7)^2 + (6-7)^2 + (8-7)^2 + (5-7)^2 + (9-7)^2}{5} = 2,000$$

$$D_B = \frac{(8-9)^2 + (8-7)^2}{2} = 1,000.$$

4. Средняя из частных дисперсий:

$$\overline{D_i} = \frac{0,667 \cdot 3 + 2,000 \cdot 5 + 1,000 \cdot 2}{3+5+2} = \frac{14,001}{10} = 1,4.$$

5. Межгрупповая дисперсия:

$$D_{\text{м.гр}} = \frac{(6-6,9)^2 \cdot 3 + (7-6,9)^2 \cdot 5 + (8-6,9)^2 \cdot 2}{3+5+2} = \frac{4,9}{10} = 0,49$$

6. Определяем общую дисперсию урожайности по всей совокупности, используя правило сложения дисперсий:

$$D = 1,40 + 0,49 = 1,89.$$

Проверим этот вывод путем расчета общей дисперсии обычным способом:

$$D = \frac{\sum (x - \bar{x})^2}{n} = \frac{(6 - 6,9)^2 + (5 - 6,9)^2 + (7 - 6,9)^2 + (7 - 6,9)^2 + (6 - 6,9)^2 + (8 - 6,9)^2 + (5 - 6,9)^2 + (9 - 6,9)^2 + (9 - 6,9)^2 + (7 - 6,9)^2}{10} = 1,89.$$

7. Определим коэффициент детерминации — R^2 :

$$R^2 = \frac{0,49}{1,89} = 0,26 \text{ или } 26\%.$$

Таким образом, только на 26% вариация урожайности обусловлена различиями между сортами, а на 74% — другими факторами (характером почвы, удобренностью участков, поливом и т.п.).

8. Определяем эмпирическое корреляционное отношение:

$$r = \sqrt{\frac{0,49}{1,89}} = \sqrt{0,26} \approx 0,5.$$

Пользуясь шкалой Чэдока, можно утверждать, что связь умеренная.

Шкала Чэдока

Значение r	0,1 – 0,3	0,3 – 0,5	0,5 – 0,7	0,7 – 0,9	0,9 – 0,99
Характеристика силы связи	Слабая	Умеренная	Заметная	Тесная	Весьма тесная

Тест для самоконтроля

- Группировочным признаком может быть
 - качественный признак
 - количественный признак
 - как качественный, так и количественный признак
- При построении простой группировки используется
 - три и более признаков
 - один признак
 - два признака

3. Задача аналитической группировки:

- а) показать структуру изучаемой совокупности
- б) выявить зависимость между признаками
- в) выделить новые группы на основе уже проведенной группировки

4. В таблице представлены данные об инвестициях в России в основном капитал (млрд.руб.)

Показатели	1998г.	1999г.
Инвестиции в основной капитал		
в том числе в отрасли:	402,4	565,6
производящие товары		
оказывающие рыночные и	153,8	269,4
нерыночные услуги	238,6	296,2

Показатель среднего абсолютного изменения структуры равен

- а) 1,4
- б) 28,8
- в) 6,9

5. Межгрупповая дисперсия при отсутствии различия между вариантами внутри групп равна

- а) единице
- б) колеблется от нуля до единицы
- в) общей дисперсии

6. Коэффициент детерминации представляет собой:

- а) долю межгрупповой дисперсии в общей
- б) долю межгрупповой дисперсии в остаточной
- в) сумму межгрупповой и внутригрупповой дисперсий

5. Если межгрупповая дисперсия равна 40, а среднее значение внутригрупповых равно 10, коэффициент детерминации равен

- а) 0,8
- б) 0,2
- в) 4

7. Если среднее значение внутригрупповых дисперсий равно 10, коэффициент детерминации равен 0,8, то полная дисперсия равна

- а) 40
- б) 80
- в) 50

8. Величина, вычисляемая по формуле

$$S_{yx}^2 = \frac{\sum_j^m (\bar{y}_j - \bar{y})^2 \cdot n_j}{\sum_j^m n_j}.$$

является

- а) внутригрупповой дисперсией
- б) межгрупповой дисперсией
- в) остаточной дисперсией

Вопросы и упражнения

1. Сформулируйте принцип построения группировок.
2. Приведите примеры простой и сложной группировок.
3. Для чего строится типологическая группировка?
4. Какая группировка позволяет сделать выводы о наличии связи между признаками?
5. Если средняя внутригрупповых дисперсий равна полной дисперсии признака в совокупности, чему равна межгрупповая дисперсия?
6. Если коэффициент детерминации близок к 1, какой вывод можно сделать относительно зависимости между признаками?
7. Перечислите задачи исследования связи между признаками на основе аналитической группировки.
8. В таблице приведены данные о результатах обследования работников предприятия по размеру заработной платы.

Группы работников по возрасту, лет	Число работников	Дисперсия заработной платы
До 20	100	300
20-30	120	400
30 и старше	150	500

Общая дисперсия признака «зарплата» составила 450. Оказывает ли возраст существенное влияние на заработную плату?

7. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

7.1. Понятие о выборочном наблюдении

Наиболее совершенным и научно обоснованным способом несплошного наблюдения является выборочное наблюдение, получившее в настоящее время широкое применение в работе органов государственной статистики, научно-исследовательских лабораторий, институтов, предприятий. Его использование позволяет лучше организовать наблюдение, обеспечивает быстроту проведения, экономию труда и средств на получение и обработку информации.

Выборочное наблюдение при строгом соблюдении условий случайности и достаточно большой численности отобранных единиц репрезентативно (представительно); по результатам изучения определенной части единиц с достаточной для практики степенью точности можно судить о всей совокупности. Однако вычисленные по материалам выборочного наблюдения статистические показатели не будут точно совпадать с соответствующими характеристиками для всей совокупности (генеральной совокупности). Величина этих отклонений называется **ошибкой наблюдения**, которая складывается из ошибок двоякого рода: ошибки регистрации (точности) и ошибки репрезентативности.

Ошибки регистрации свойственны любому наблюдению (сплошному и несплошному). Они вызываются несовершенством измерительных приборов, недостаточной квалификацией наблюдателя, неточностью подсчетов и т. п. Однако при выборочном наблюдении они значительно меньше, так как в этом случае используются более квалифицированные и подготовленные кадры.

Ошибки репрезентативности свойственны только несплошным наблюдениям. Они характеризуют размер расхождений между величинами показателя, полученного в выборочной и генеральной совокупности в условиях одинаковой точности единичных наблюдений. Ошибки репрезентативности могут быть систематическими и случайными. Систематические ошибки возникают при нарушении установленных правил отбора единиц. Случайные ошибки репрезентативности обязаны своим возникновением недостаточно равномерным представлением в выборочной совокупности различных категорий единиц генеральной совокупности.

Величина случайной ошибки определяет надежность данных выборочного наблюдения, их пригодность для суждения о генеральной

совокупности. При помощи формул теории вероятностей можно рассчитать возможную максимальную случайную ошибку — вероятный (стохастический) предел ошибки.

Максимально возможная ошибка — это такая величина отклонения выборочной средней (доли) от генеральной, вероятность превышения которой вследствие случайных причин в условиях данной выборки очень мала.

Величина случайной ошибки репрезентативности зависит от:

- степени колеблемости изучаемого признака в генеральной совокупности;
- способа формирования выборочной совокупности;
- объема выборки.

По **степени охвата единиц** исследуемой совокупности различают большие и малые выборки.

По **способу формирования** выборочной совокупности различают следующие виды выборочного наблюдения: простая случайная (собственно случайная) выборка, расслоенная (типическая или районированная), серийная, механическая, комбинированная, ступенчатая, многофазная.

Совокупность единиц, из которых производится отбор, принято называть **генеральной совокупностью**. Совокупность отобранных единиц из генеральной совокупности называется **выборочной совокупностью**.

N — объем генеральной совокупности (число входящих в нее единиц);

n — объем выборочной совокупности (число единиц, попавших в выборку);

$\bar{X}$ — генеральная средняя (среднее значение признака в генеральной совокупности);

$\bar{x}$ — выборочная средняя (среднее значение признака в выборочной совокупности);

P — генеральная доля (доля единиц, обладающих данным признаком в генеральной совокупности);

p — выборочная доля (доля единиц, обладающих данным признаком в выборочной совокупности);

σ^2 — генеральная дисперсия (дисперсия признака в генеральной совокупности);

S^2 — выборочная дисперсия (дисперсия признака в выборочной совокупности);

σ — среднее квадратическое отклонение признака в генеральной совокупности;

S — среднее квадратическое отклонение признака в выборочной совокупности.

7.2. Простая случайная выборка

При простой случайной выборке отбор единиц в выборочную совокупность производится непосредственно из всей массы единиц генеральной совокупности в форме случайного отбора, при котором каждой единице генеральной совокупности обеспечивается одинаковая вероятность (возможность) быть выбранной. Единица отбора совпадает с единицей наблюдения. Случайный отбор осуществляется путем применения жеребьевки (лотереи) или путем использования таблиц случайных чисел.

Случайный отбор может быть проведен в двух формах: в форме возвратной (повторной) выборки и в форме безвозвратной (бесповторной) выборки. При повторном отборе вероятность попадания каждой единицы генеральной совокупности остается постоянной, так как после отбора какой-то единицы она снова возвращается в генеральную совокупность и может быть выбранной. При бесповторном отборе выбранная единица не возвращается в генеральную совокупность и вероятность попадания отдельных единиц в выборку все время изменяется (для оставшихся единиц она возрастает).

Применение простой случайной повторной выборки на практике весьма ограничено; обычно используется бесповторная выборка.

Теорема П.Л. Чебышева утверждает принципиальную возможность определения генеральной средней по данным случайной повторной выборки. Теорема Чебышева дополняется **теоремой А.М. Ляпунова**, которая позволяет рассчитать максимальную ошибку выборочной средней при данном достаточно большом числе независимых наблюдений. Согласно этой теореме при достаточно большом числе независимых наблюдений в генеральной совокупности с конечной средней и ограниченной дисперсией вероятность того, что расхождение между выборочной и генеральной средней ($\bar{x} - \bar{X}$) не превзойдет по абсолютной величине некоторую величину $t\mu$ равна **интегралу Лапласа**. Это можно записать так:

$$P(|\bar{x} - \bar{X}| \leq t\mu) = \Phi(t);$$

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt.$$

где $\Phi(t)$ — интеграл Лапласа (нормированная функция Лапласа).

Величина $t\mu$ обозначаемая Δ , называется **предельной ошибкой выборки**. Следовательно,

$$\Delta_{\bar{x}} = t\mu_{\bar{x}}; \Delta_p = t\mu_p,$$

где $\Delta_{\bar{x}}$ — предельная (максимально возможная) ошибка средней;

Δ_p — предельная (максимально возможная) ошибка доли;

μ — величина средней квадратической стандартной ошибки;

t — коэффициент кратности средней ошибки выборки, зависящий от вероятности, с которой гарантируется величина предельной ошибки.

В зависимости от принятой вероятности P определяется значение коэффициента кратности (t) по удвоенной нормированной функции Лапласа.

Величина средней ошибки в условиях большой выборки ($n > 30$) рассчитывается по известным из теории вероятностей формулам:

а) при случайной повторной выборке:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}; \quad \mu_p = \sqrt{\frac{p \cdot (1-p)}{n}};$$

б) при случайной бесповторной выборке:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)};$$

$$\mu_p = \sqrt{\frac{p \cdot (1-p)}{n} \cdot \left(1 - \frac{n}{N}\right)}.$$

При расчете ошибок возникает существенное затруднение: величины σ и p по генеральной совокупности неизвестны. Эти величины в условиях большой выборки заменяют величинами S (выборочная дисперсия) и w (выборочная доля), рассчитанными по выборочным данным. В табл. 7.1 приведены формулы расчета ошибок простой случайной выборки.

Таблица 7.1 Формулы ошибок простой случайной выборки

	Способ отбора единиц	
	повторный	бесповторный
Средняя ошибка μ : для средней	$\mu_x = \sqrt{\frac{S^2}{n}}$	$\mu_x = \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right)}$
для доли	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Предельная ошибка Δ : для средней	$\Delta_x = t \sqrt{\frac{S^2}{n}}$	$\Delta_x = t \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right)}$
для доли	$\Delta_p = t \sqrt{\frac{w(1-w)}{n}}$	$\Delta_p = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

Формулы предельной ошибки позволяют решать задачи трех видов:

1. Определение пределов генеральных характеристик с заданной степенью надежности (доверительной вероятностью) на основе показателей, полученных по данным выборки.

Доверительные интервалы для генеральной средней —

$$x = \tilde{x} \pm \Delta_x;$$

$$\tilde{x} - \Delta_x \leq x \leq \tilde{x} + \Delta_x.$$

Доверительные интервалы для генеральной доли —

$$p - w \pm \Delta_p;$$

$$w - \Delta_p \leq p \leq w + \Delta_p.$$

2. Определение доверительной вероятности того, что генеральная характеристика может отличаться от выборочной не более чем на определенную заданную величину.

Доверительная вероятность является функцией от t , определяемой по формуле

$$t = \frac{\Delta_x}{\mu_x}.$$

По величине t определяется доверительная вероятность.

3. Определение необходимого объема выборки, который с практической вероятностью обеспечивает заданную точность выборки.

Для расчета объема выборки необходимо иметь следующие данные:

- а) размер доверительной вероятности (P);
- б) коэффициент t , зависящий от принятой вероятности;

в) величину σ^2 (или pq) в генеральной совокупности; они заменяются величинами, полученными в предшествующих обследованиях или при пробных выборках.

Таблица 7.2 Формулы для определения численности простой случайной выборки

	Способ отбора единиц	
	повторный	бесповторный
Численность выборки (n): для средней	$n = \frac{t^2 S^2}{\Delta_x^2}$	$n = \frac{t^2 N S^2}{\Delta_x^2 N + t^2 S^2}$
для доли	$n = \frac{t^2 w(1-w)}{\Delta_p^2}$	$n = \frac{t^2 N w(1-w)}{\Delta_p^2 N + t^2 w(1-w)}$
В случаях, когда частость w даже приблизительно неизвестна, в расчет вводят максимальную величину дисперсии доли, равную 0,25 (если $w=0,5$, то $w(1-w)=0,25$).		

г) величину максимально допустимой ошибки (Δ_x или Δ_p);

д) объем генеральной совокупности (N).

Необходимый объем выборки определяется на основе допустимой величины ошибки: $\Delta_x = t\mu_x$ или $\Delta_p = t\mu_p$.

В табл. 7.2 приведены формулы для расчета численности простой случайной выборки.

7.3. Расслоенная (типическая или районированная) выборка

В составе генеральной совокупности с различным уровнем изучаемого признака желательно обеспечить более равномерное представительство в выборочной совокупности различных типов. Эта цель достигается при применении расслоенной (типической или стратифицированной) выборки. Эту выборку применяют также в целях более рав-

номерного представления в выборке различных районов, и в этом случае ее называют районированной выборкой.

При типической выборке неоднородная генеральная совокупность подразделяется на более однородные в отношении изучаемых признаков группы (типы, районы). По каждой группе определяются ее объем (N_i) и число подлежащих наблюдению единиц (n_i). Отбор обследуемых единиц производится в каждой группе при помощи одного из способов случайного отбора — повторного или бесповторного.

Общее число единиц выборочной совокупности распределяется между группами пропорционально численности групп в составе генеральной совокупности. Такой отбор называется **пропорциональным**.

N — общая численность единиц в генеральной совокупности

$$N = N_1 + N_2 + \dots + N_k,$$

где $N_1, N_2, \dots, N_k$ — численность отдельных групп генеральной совокупности;

n — общий объем выборочной совокупности.

Объем выборки для каждой группы —

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

где $\frac{N_i}{N}$ удельный вес данной (i -й) группы в генеральной совокупности;

$$n = n_1 + n_2 + \dots + n_k.$$

Кроме пропорционального размещения по группам численности единиц выборочной совокупности применяется так называемое оптимальное размещение, при котором число наблюдений в группе определяется по формуле

$$n_i = \frac{N_i \sigma_i}{\sum_1^k N_i \sigma_i} \cdot n.$$

Формулы для расчета ошибок типической выборки приведены в табл. 7.3.

В табл. 7.3 приняты следующие условные обозначения:

S^2 — средняя групповая выборочная дисперсия средней;

$$S^2 = \frac{\sum S_i^2 n_i}{\sum n_i};$$

S^2 — внутригрупповая дисперсия данной (i -й) группы в выборочной совокупности;

$\overline{w(1-w)}$ — средняя групповая выборочная дисперсия доли.

Как видно из приведенных формул, величина стандартной ошибки типической выборки зависит только от точности определения групповых средних, т. е. от величины внутригрупповых дисперсий. Согласно правилу сложения дисперсий общая дисперсия складывается из межгрупповой дисперсии и средней из внутригрупповых дисперсий. Отсюда следует, что ошибка типической случайной выборки меньше, чем ошибка простой случайной выборки.

Предельная (максимально возможная) ошибка типической выборки:

$$\Delta_x = t\mu_x; \Delta_p = t\mu_p.$$

Таблица 7.3 Формулы ошибок типической выборки

	Способ отбора единиц	
	повторный	бесповторный
Средняя ошибка (μ): для средней: а) при пропорциональном размещении единиц	$\mu_{\bar{x}} = \sqrt{\frac{\overline{S^2}}{n}}$	$\mu_{\bar{x}} = \sqrt{\frac{\overline{S^2}}{n} \left(1 - \frac{n}{N}\right)}$
б) при оптимальном размещении единиц	$\mu_{\bar{x}} = \frac{1}{N} \sqrt{\frac{S_i^2 N_i^2}{n_i}}$	$\mu_{\bar{x}} = \frac{1}{N} \sqrt{\sum \frac{S_i^2 N_i^2}{n_i} \left(1 - \frac{n_i}{N_i}\right)}$
для доли: в) при пропорциональном размещении единиц	$\mu_p = \sqrt{\frac{\overline{w(1-w)}}{n}}$	$\mu_p = \sqrt{\frac{\overline{w(1-w)}}{n} \left(1 - \frac{n}{N}\right)}$
г) при оптимальном размещении единиц	$\mu_p = \frac{1}{N} \sqrt{\frac{w_i (1-w_i) N_i^2}{n_i}}$	$\mu_p = \frac{1}{N} \sqrt{\sum \frac{w_i (1-w_i) N_i^2}{n_i} \left(1 - \frac{n_i}{N_i}\right)}$

Необходимый объем выборки определяется на основе формулы и величины допустимой ошибки.

7.4. Серийная выборка

Сущность серийной выборки заключается в том, что вместо случайного отбора единиц совокупности осуществляется отбор групп (серий, гнезд). Внутри отобранных серий производится сплошное наблюдение. Серии (гнезда) состоят из единиц, связанных между собой или территориально, или организационно, или, наконец, во времени. Отбор серий может производиться в порядке повторного и бесповторного отбора. Серии могут быть равновеликими и неравновеликими. На практике чаще применяется серийный отбор с равными сериями.

Стандартная ошибка при равновеликих сериях определяется по формулам, представленным в табл. 7.4.

Таблица 7.4 Стандартная ошибка при равновеликих сериях

	Способ отбора серий	
	повторный	бесповторный
Средняя ошибка (μ): для средней	$\mu_{\bar{x}} = \sqrt{\frac{\delta_{\bar{x}}^2}{m}}$	$\mu_{\bar{x}} = \sqrt{\frac{\delta_{\bar{x}}^2}{m} \left(1 - \frac{m}{M}\right)}$
для доли	$\mu_p = \sqrt{\frac{\delta_w^2}{m}}$	$\mu_p = \sqrt{\frac{\delta_w^2}{m} \left(1 - \frac{m}{M}\right)}$

В табл. 7.4 приняты следующие условные обозначения:

$\delta_{\bar{x}}^2$ — межгрупповая выборочная дисперсия средней;

$$\delta_{\bar{x}}^2 = \frac{\sum (\tilde{x}_i - \tilde{x}_0)^2}{m},$$

где $\tilde{x}_i$ — средний уровень признака в серии;

$\tilde{x}_0$ — средний уровень признака для всей выборочной совокупности;

m — число равных серий в выборочной совокупности;

M — число равных серий в генеральной совокупности;

δ_w^2 — межгрупповая выборочная дисперсия доли;

$$\delta_w^2 = \frac{\sum (w_i - w)^2}{m},$$

где w_i — доля единиц, обладающих данным признаком в серии;

w — доля единиц, обладающих данным признаком во всей выборочной совокупности.

Ошибка серийной выборки больше, чем при любом другом способе отбора. Тем не менее, серийный отбор широко применяется на практике, что объясняется его организационными преимуществами.

7.5. Механическая выборка

Механическая выборка заключается в отборе единиц из генеральной совокупности через равные промежутки из определенного расположения их в генеральной совокупности (по алфавиту, в пространстве, последовательности появления во времени).

При организации механического отбора возникают две задачи:

- определение «шага отчета» (расстояния между отбираемыми единицами);

- выбор единицы, с которой надо начинать отчет.

«Шаг отчета» определяется путем деления численности генеральной совокупности на численность выборочной совокупности: $\frac{N}{n}$.

Выбор начала отчета рекомендуется производить путем случайного отбора из единиц первого интервала — первого «шага отчета». Механический отбор может осуществляться в самом процессе наблюдения, и его удобно применять тогда, когда выборочно наблюдается масса постепенно возникающих перед наблюдателем единиц (например, производят проверку каждой 10-й, 20-й и т. д. детали, обработанной на станке).

Если в генеральной совокупности единицы располагаются случайным образом по отношению к изучаемому признаку, то механический отбор можно рассматривать как разновидность случайного бесповторного отбора; поэтому для оценки ошибки механической выборки применяются формулы случайной бесповторной выборки:

$$\mu_{\bar{x}} = \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right)}, \quad \Delta_{\bar{x}} = t \mu_{\bar{x}};$$

$$\mu_p = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}, \quad \Delta_p = t\mu_p.$$

7.6. Другие виды

Комбинированная выборка предполагает использование нескольких способов выборки. Можно комбинировать, например, серийную выборку и случайную. В этом случае, разбив генеральную совокупность на серии (группы) и отобрав нужное число серий, производят случайную выборку единиц в серии. Такая комбинированная выборка может быть повторной (для групп и единиц) и бесповторной.

Средняя ошибка комбинированной выборки определяется по формулам:

$$\text{при повторном отборе} \quad \mu_{\bar{x}} = \sqrt{\frac{S^2}{n} + \frac{\delta^2}{m}};$$

$$\text{при бесповторном отборе} \quad \mu_{\bar{x}} = \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right) + \frac{\delta^2}{m} \left(1 - \frac{m}{M}\right)}.$$

Многоступенчатая выборка предполагает извлечение из генеральной совокупности сначала укрупненных групп единиц, затем групп, меньших по объему, и так до тех пор, пока не будут отобраны те группы (серии) или отдельные единицы, которые будут подвергнуты наблюдению. Выборка может быть двухступенчатой, когда генеральная совокупность разбивается на группы и производится отбор групп, а затем внутри групп — отбор единиц наблюдения. На обеих ступенях отбор может вестись в случайном порядке.

В отличие от типического отбора, где отбор производится из всех без исключения групп, при многоступенчатом отборе производится отбор самих групп, и, следовательно, не все они попадают в выборку.

Число ступеней отбора может быть и более трех. Если число ступеней отбора больше двух, то средняя ошибка выборки определяется по формуле

$$\mu = \sqrt{\mu_1^2 + \frac{\mu_2^2}{n_1} + \frac{\mu_3^2}{n_1 n_2} + \dots},$$

где μ_1, μ_2, μ_3 — средние ошибки выборки на отдельных ступенях отбора;

n_1, n_2 численность выборок на соответствующих ступенях..

При **многофазной выборке** выборочные совокупности образуются так, что одни сведения собираются от всех единиц отбора, затем отбираются еще некоторые единицы, которые и обследуются по более широкой программе. Расчет ошибки многофазной выборки производится для каждой фазы в отдельности.

7.7. Малые выборки

Выборки, при которых наблюдением охватывается небольшое число единиц ($n < 30$), принято называть малыми выборками. Они обычно применяются в том случае, когда невозможно или нецелесообразно использовать большую выборку (исследование качества продукции, если это связано с ее разрушением, в частности на прочность, на продолжительность срока службы и т. д.).

Предельная ошибка малой выборки определяется по формуле

$$\Delta_{\text{М.В}} = t \cdot \mu_{\text{М.В}}.$$

Средняя ошибка малой выборки:

$$\mu_{\text{М.В}} = \sqrt{\frac{S^2}{n}},$$

где S^2 — дисперсия малой выборки.

$$S^2 = \frac{\sum (x - \tilde{x})^2}{n - 1},$$

где $\tilde{x}$ — среднее значение признака по выборке;

$n - 1$ — число степеней свободы ($n - 1 = k$);

t — коэффициент доверия малой выборки, зависящий не только от заданной доверительной вероятности, но и от численности единиц выборки.

Вероятность того, что генеральная средняя находится в определенных границах, определяется по формуле

$$P[\tilde{x} - t\mu_{\text{М.В}} \leq \bar{x} \leq \tilde{x} + t\mu_{\text{М.В}}] = 2S_{(t)} - 1,$$

где $S_{(t)}$ — значение функции Стьюдента.

Для расчета коэффициента доверия t определяют значение функции $S_{(t)}$ по формуле

$$S_{(t)} = (P + 1) : 2.$$

Затем по таблице распределения Стьюдента в зависимости от значения функции $S_{(t)}$ и числа степеней $k(n - 1)$ определяют значение t .

Функция $S_{(t)}$ используется также для определения вероятностей того, что фактическое нормированное отклонение ($t_{\phi} = \frac{\tilde{x} - \bar{x}}{\mu_{\text{м.в}}}$) не превзойдет или превзойдет табличное значение.

Вероятность того, что фактическое отношение (t_{ϕ}) не превзойдет по абсолютной величине табличное значение (t), определяется по формуле

$$P[t_{\phi} < |t|] = 2S_{(t)} - 1.$$

Вероятность того, что фактическое отношение превзойдет по абсолютной величине табличное значение

$$P[t_{\phi} > |t|] = 2[1 - S_{(t)}].$$

7.8. Решение типовых задач

Задача 1. Из партии электроламп взята 20%-ная случайная бесповторная выборка для определения среднего веса спирали.

Результаты выборки следующие (табл. 7.5).

Таблица 7.5

Вес, мг	38 – 40	40 – 42	42 – 44	44 – 46
Число спиралей	15	30	45	10

Определить с вероятностью 0,95 доверительные пределы, в которых лежит средний вес спирали, для всей партии электроламп.

Решение.

Доверительные интервалы для генеральной средней с вероятностью P :

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

где $\tilde{x}$ — средний уровень признака по выборке:

$$\tilde{x} = \frac{\sum x' f}{\sum f} = \frac{39 \cdot 15 + 41 \cdot 30 + 43 \cdot 45 + 45 \cdot 10}{15 + 30 + 45 + 10} = \frac{4200}{100} = 42.0 \text{ мг}$$

$$\Delta_{\bar{x}} = t\mu_{\bar{x}} = t\sqrt{\frac{S^2}{n} \cdot \left(1 - \frac{n}{N}\right)}; \quad N = \frac{100}{0,2} = 500.$$

При вероятности $P = 0,95$ $t = 1,96$.

$$S^2 = \frac{\sum (x' - \tilde{x})^2 \cdot f}{\sum f} =$$

$$= \frac{(39 - 42)^2 \cdot 15 + (41 - 42)^2 \cdot 30 + (43 - 42)^2 \cdot 45 + (45 - 42)^2 \cdot 10}{100} =$$

$$= \frac{300}{100} = 3,0$$

$$\Delta_{\bar{x}} = 1,96 \sqrt{\frac{3,0}{100} \cdot \left(1 - \frac{100}{500}\right)} = 0,3 \text{ мг.}$$

Доверительные интервалы для генеральной средней с вероятностью $P = 0,95$:

$$42,0 - 0,3 \leq \bar{x} \leq 42,0 + 0,3;$$

$$41,7 \text{ мг} \leq \bar{x} \leq 42,3 \text{ мг.}$$

Задача 2. По городской телефонной сети в порядке случайной выборки (механический) отбор произвели 100 наблюдений и установили среднюю продолжительность одного телефонного разговора 5 мин при среднем квадратическом отклонении 2 мин.

Какова вероятность того, что ошибка репрезентативности при определении средней продолжительности разговора не превысит 18 с?

Решение.

По условию задачи известны:

объем выборки — $n = 100$;

выборочная средняя — $\tilde{x} = 5$ мин;

выборочное среднее квадратическое отклонение — $S = 2$ мин;

предельная ошибка выборки — $\Delta_{\bar{x}} = 18 \text{ сек} = 0,3 \text{ мин.}$

$$\Delta_{\bar{x}} = t\mu_{\bar{x}}; \quad \mu_{\bar{x}} = \sqrt{\frac{S^2}{n}} = \sqrt{\frac{2^2}{100}} = 0,2 \text{ мин.}$$

$$t = \frac{\Delta_{\bar{x}}}{\mu_{\bar{x}}} = \frac{0,3}{0,2} = 1,5.$$

Затем по таблице на основе значения t определяется вероятность того, что ошибка не превысит заданной величины.

При $t = 1,5$ вероятность $P = 0,866$.

Задача 3.

На основе выборочного обследования в отделении связи города предполагается определить долю писем частных лиц в общем объеме отправляемой корреспонденции. Никаких предварительных данных об удельном весе этих писем в общей массе отправляемой корреспонденции не имеется.

Определить численность выборки, если результаты выборки дать с точностью до 1% и гарантировать это с вероятностью 0,95.

По условию задачи известны:

размер допустимой (предельной) ошибки — $\Delta_p = 1\%$ или 0,01;

принятая вероятность — $P = 0,95$;

при $P = 0,95$ $t = 1,96$.

Необходимая численность выборки:

$$n = \frac{t^2 \cdot w(1-w)}{\Delta_p^2}.$$

Так как значение w не дано, то следует ориентироваться на наибольшую дисперсию, которой соответствует значение $w = 0,5$.

$$n = \frac{1,96^2 \cdot 0,5 \cdot 0,5}{0,01^2} = 9604.$$

Таким образом, чтобы с заданной точностью определить долю частных писем в общем объеме отправляемой корреспонденции, необходимо в порядке случайной выборки отобрать 9604 письма.

Тест для самоконтроля

1. Выборочное наблюдение является

- а) несплошным
- б) сплошным
- в) основного массива

2. Расхождение выборочных и генеральных характеристик, обусловленное нарушением правил случайности отбора, характеризуют

- а) ошибки регистрации
- б) случайные ошибки репрезентативности
- в) систематические ошибки репрезентативности

3. Совокупность единиц, из которой производится отбор, называется

- а) статистическая совокупность
- б) генеральная совокупность
- в) выборочная совокупность

4. Совокупность единиц, отобранная из некоторой совокупности, называется

- а) статистическая совокупность
- б) генеральная совокупность
- в) выборочная совокупность

5. Проведен выборочный опрос населения города S с целью изучения намерения голосовать за кандидата N. При механическом способе отбора из списка избирателей в 20 тысяч человек было опрошено 800 человек, из которых 640 человек высказались в поддержку кандидата N. Определите с вероятностью 0,997 возможный процент поддерживающих кандидата N среди всех избирателей

- а) от 76% до 84%
- б) от 72% до 80%
- в) от 76% до 80%

6. Формула

$$\mu_{\bar{x}} = \frac{1}{N} \sqrt{\frac{S_i^2 N_i^2}{n_i}}$$

применяется для расчета ошибки

- а) при случайной повторной выборке
- б) при случайной бесповторной выборке
- в) при типической выборке

7. В результате повторной случайной выборки в городе предполагается определить долю многодетных семей. На основе предыдущих исследований известно, что дисперсия признака равна 0,27. Чтобы с вероят-

ностью 0,954 ошибка выборки не превышала 2%, необходимо опросить

- а) 1048 семей
- б) 2700 семей
- в) 954 семьи

Вопросы и упражнения

1. Укажите преимущества выборочного наблюдения перед сплошным наблюдением.
2. Согласны ли Вы, что правильно организованное выборочное наблюдение может повысить достоверность результатов по сравнению со сплошным наблюдением? Ответ обоснуйте.
3. Что такое «ошибки репрезентативности»?
4. Определите понятия «генеральная совокупность», «выборочная совокупность».
5. Сформулируйте основной принцип построения простой случайной выборки.
6. Какова максимальная величина дисперсии доли?
7. Что такое «сфотографированная» выборка?
8. При каких условиях механическую выборку можно рассматривать как случайную?
9. В чем отличие «большой» и «малой» выборок?
10. Определите с вероятностью 0,997 пределы, в которых будет находиться процент брака для партии электроламп 16 000 штук, если из проверенных 1600 штук (случайный бесповторный отбор) оказались бракованными 40 штук. (Ответ: $1,4\% \leq p \leq 3,6\%$).

8. РЯДЫ ДИНАМИКИ

8.1. Виды рядов динамики

Ряд расположенных в хронологической последовательности значений статистических показателей образует **временной (динамический) ряд**. Каждый временной ряд включает два элемента: момент или период времени и конкретное значение показателя (уровень ряда).

Ряды динамики классифицируются по времени (интервальные и моментные), по форме представления уровней (ряды абсолютных, относительных и средних величин), по расстоянию между датами или интервалами времени (полные и неполные хронологические ряды), по числу показателей (изолированные и комплексные ряды динамики).

Интервальный ряд динамики — последовательность, в которой уровень явления относится к результату, накопленному или произведенному за определенный интервал времени. Например, ряды показателей объема выпуска продукции по месяцам. Если же уровень ряда характеризует фактическое наличие изучаемого явления в конкретный момент времени, то совокупность уровней образует **моментный ряд** динамики. Примером моментного ряда может служить последовательность показателей численности населения на начало года.

В результате суммирования уровней интервального динамического ряда получаются **накопленные итоги**. Например, общий объем продукции, выпущенной за год. Сумма уровней моментного ряда динамики не имеет реального содержания, и накопленные итоги для таких рядов не подсчитываются.

Если даты регистрации или окончания периодов следуют друг за другом, то ряд динамики называется **полным**. **Неполные** ряды — когда принцип равных интервалов не соблюдается.

Если ведется временной анализ одного показателя, то имеем **изолированный ряд** динамики, если анализируется система показателей, то — **комплексный** или многомерный динамический ряд.

8.2. Статистические показатели динамики

Для количественной оценки динамики социально-экономических явлений применяются статистические показатели: абсолютные приросты, темпы роста и прироста, темпы 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 способом показатели динамики называются *цепными*. Например. Для ряда динамики розничного товарооборота магазина в 1994 — 1998 гг. за постоянную базу сравнения принят уровень 1994г. — года начала функционирования данного торгового предприятия.

Способы расчета показателей динамики рассмотрим на данных о товарообороте магазина в 1994 — 1998 гг.

Важнейшим статистическим показателем динамики является *абсолютный прирост*, который определяется в разностном сопоставлении двух уровней ряда динамики в единицах измерения исходной информации.

Базисный абсолютный прирост $\Delta y_{\text{б}}$ исчисляется как разность между сравниваемым уровнем y_i и уровнем, принятым за постоянную базу сравнения y_{0_i} :

$$\Delta y_{\text{б}_i} = y_i - y_{0_i} . \quad (8.1)$$

Цепной абсолютный прирост $\Delta y_{\text{ц}}$ — разность между сравниваемым уровнем y_i и уровнем, который ему предшествует y_{i-1} :

$$\Delta y_{\text{ц}_i} = y_i - y_{i-1} . \quad (8.2)$$

Расчет абсолютных приростов по формулам (8.1) и (8.2) приведен в табл. 8.1.

Из табл. 8.1 видно, что по сравнению с 1994 г. в каждом последующем году происходило систематическое увеличение абсолютных приростов товарооборота (тыс. руб.): $46,9 < 94,4 < 143,0 < 202,7$. Цепные абсолютные приросты показывают, что нарастание объемов товарооборота происходило из года в год:

$$46,9 < 47,5 < 48,6 < 59,7.$$

Абсолютный прирост может иметь и отрицательный знак, показывающий, насколько уровень изучаемого периода ниже базисного.

Таблица 8.1 Основные показатели динамики товарооборота магазина в 1994 — 1998 гг.

Показатель	1994	1995	1996	1997	1998
Товарооборот, тыс. руб., Y	885,7	932,6	980,1	1028,7	1088,4
Абсолютный прирост, тыс. руб.					
базисный ($\Delta y_{\text{б}_i} = y_i - y_{0_i}$)	—	$932,6 - 885,7 = 46,9$	$980,1 - 85,7 = 94,4143,0$	202,7	
цепной ($\Delta y_{\text{ц}_i} = y_i - y_{i-1}$)	—	$932,6 - 885,7 = 46,9$	$980,1 - 932,6 = 47,5$	48,6	59,7
Темп роста, %					
базисный ($\text{Тр}_{\text{б}_i} = (y_i : y_{0_i}) \cdot 100$)	—	$932,6 : 885,7 \cdot 100 = 105,3$	$980,1 : 885,7 \cdot 100 = 110,6$	116,1	122,9
цепной ($\text{Тр}_{\text{ц}_i} = (y_i : y_{i-1}) \cdot 100$)	—	$932,6 : 885,7 \cdot 100 = 105,3$	$980,1 : 932,6 \cdot 100 = 105,1$	104,9	105,8
Темп прироста, %					
базисный ($\text{Тп}_{\text{б}_i} = (\Delta y_{\text{б}_i} : y_{0_i}) \cdot 100$)	—	$46,9 : 885,7 \cdot 100 = 5,3$	$94,4 : 885,7 \cdot 100 = 10,6$	16,1	22,9
цепной ($\text{Тп}_{\text{ц}_i} = (\Delta y_{\text{ц}_i} : y_{i-1}) \cdot 100$)	—	$46,9 : 885,7 \cdot 100 = 5,3$	$47,5 : 932,6 \cdot 100 = 5,1$	4,9	5,8
Темп наращивания, %					
($\text{Тн}_{\text{ц}_i} = (\Delta y_{\text{ц}_i} : y_{0_{\text{н}}}) \cdot 100$)	—	$46,9 : 885,7 \cdot 100 = 5,3$	$47,5 : 885,7 \cdot 100 = 5,4$	5,5	6,8

Между базисными и цепными абсолютными приростами имеется связь: сумма цепных абсолютных приростов $\sum \Delta y_{\text{ц}}$ равна базисному абсолютному приросту последнего периода ряда динамики $\Delta y_{\text{б}_n}$:

$$\Delta y_{\text{б}_n} = \sum \Delta y_{\text{ц}}. \quad (8.3)$$

Так, применяя формулу (8.3), можно по вычисленным в табл. 8.1 цепным абсолютным приростам определить базисный абсолютный прирост:

$$\Delta y_{\text{б}_n} = 46,9 + 47,5 + 48,6 + 59,7 = 202,7 \text{ тыс. руб.}$$

Распространенным статистическим показателем динамики является **темп роста**. Он характеризует отношение двух уровней ряда и может выражаться в виде коэффициента или в процентах.

Базисные темпы роста $\text{Тр}_{\text{б}}$ исчисляются делением сравниваемого уровня (y_i) на уровень, принятый за постоянную базу сравнения y_{0_i} :

$$\text{Тр}_{\text{б}_i} = y_i : y_{0_i}. \quad (8.4)$$

Цепные темпы роста $\text{Тр}_{\text{ц}}$ исчисляются делением сравниваемого уровня y_i на предыдущий уровень y_{i-1} :

$$\text{Тр}_{\text{ц}_i} = y_i : y_{i-1}. \quad (8.5)$$

Если темп роста больше единицы (или 100%), то это показывает на увеличение изучаемого уровня по сравнению с базисным. Темп роста, равный единице (или 100%), показывает, что уровень изучаемого периода по сравнению с базисным не изменился. Темп роста меньше единицы (или 100%) показывает на уменьшение уровня изучаемого периода по сравнению с базисным. Темп роста всегда имеет положительный знак.

Расчет темпов роста по формулам (8.4) и (8.5) приведен в табл.8.1.

Показатели базисных темпов роста табл. 8.1 свидетельствуют, что с 1994 г. происходило систематическое увеличение товарооборота магазина, достигшего в 1998 г. 122,9% базисного уровня. Цепные темпы роста показывают, что в развитии товарооборота имело место замедление годовых темпов (%):

$$105,3 > 105,1 > 104,9 < 105,8.$$

Между базисными и цепными темпами роста имеется взаимосвязь: произведение последовательных цепных темпов роста равно базисному темпу роста, а частное от деления последующего базисного темпа роста на предыдущий равно соответствующему цепному темпу роста:

$$\frac{Y_{1995}}{Y_{1994}} \cdot \frac{Y_{1996}}{Y_{1995}} \cdot \frac{Y_{1997}}{Y_{1996}} \cdot \frac{Y_{1998}}{Y_{1997}} = \frac{Y_{1998}}{Y_{1994}}. \quad (8.6)$$

Так, подставляя в левую часть формулы (8.6) вычисленные в табл. 1.3 цепные темпы роста (в коэффициентах):

$$1,053 \cdot 1,051 \cdot 1,049 \cdot 1,058,$$

получаем базисный темп роста в 1998 г. = 1,229.

Темпы прироста характеризуют абсолютный прирост в относительных величинах. Исчисленный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень с уровнем, принятым за базу сравнения.

Базисный темп прироста T_{Π_b} вычисляется делением сравниваемого базисного абсолютного прироста Δy_{b_i} на уровень, принятый за постоянную базу сравнения y_{0_i} :

$$T_{\Pi_b} = \Delta y_{b_i} : y_{0_i}. \quad (8.7)$$

Цепной темп прироста $T_{\Pi_{ц}}$ — это отношение сравниваемого цепного абсолютного прироста $\Delta y_{ц_i}$ к предыдущему уровню y_{i-1} :

$$T_{\Pi_{ц}} = \Delta y_{ц_i} : y_{i-1}. \quad (8.8)$$

Расчет темпов прироста по формулам (8.7) и (8.8) приведен в табл. 8.1.

Между показателями темпа прироста и темпа роста имеется взаимосвязь:

$$T_{\Pi_i}(\%) = Tr_i(\%) - 100 \quad (8.9)$$

(при выражении темпа роста в процентах).

$$T_{\Pi_i} = Tr_i - 1 \quad (8.10)$$

(при выражении темпа роста в коэффициентах).

Формулы (8.9) и (8.10) используются для определения темпов прироста по темпам роста. Например, на основе вычисленного для 1998 г. базисного темпа роста товарооборота 112,9% по формуле (8.9) можно определить темп прироста:

$$Тп_{б_{1998}} = 112,9 - 100 = 22,9\%.$$

Если уровни ряда динамики сокращаются, то соответствующие показатели темпа прироста будут со знаком минус, так как они характеризуют относительное уменьшение прироста уровня ряда динамики. Например, если в I квартале 1998 г. объем реализации товаров А составил 92% уровня продажи этого товара в IV квартале 1997 г., то, применяя формулу (1.5), получим такой темп прироста:

$$Тп_{Iкв.} = 92 - 100 = -8\%,$$

т.е. произошло сокращение продажи на 8%.

Важным статистическим показателем динамики социально-экономических процессов является **темп наращивания**, который в условиях интенсификации экономики измеряет наращивание во времени экономического потенциала.

Вычисляются темпы наращивания $ТН$ делением цепных абсолютных приростов $\Delta y_{ц_i}$ на уровень, принятый за постоянную базу сравнения, y_{0_i} :

$$ТН_i = \Delta y_{ц_i} : y_{0_i}. \quad (8.11)$$

Применение формулы (1.11) проиллюстрировано в табл. 8.1.

В отличие от вычисленных в табл. 8.1 цепных темпов прироста $Тп_{ц}$, показывающих затухание, развитие товарооборота в 1995 — 1997 гг: $5,3 > 5,1 > 4,9$, темп наращивания $ТН$ свидетельствует о наращивании из года в год объема товарооборота:

$$5,3 < 5,4 < 5,5 < 6,8.$$

Противоречивость вычисленных в табл. 8.1 показателей темпов прироста и наращивания объясняется особенностями принятого метода счета. Так, в ряду чисел 1,2,3,4,... при одинаковых цепных абсолютных приростах (равных 1) отношения чисел дают следующие результаты (%):

$$\frac{1}{2} 100 = 200, \frac{3}{2} 100 = 150, \frac{4}{3} 100 = 133 \text{ и т.д.,}$$

т.е. налицо затухание темпов роста. Аналогичные действия производятся и при вычислении в табл. 8.1 темпов роста и прироста.

Из преобразования в формуле (8.11) следует, что темпы наращивания можно непосредственно определять по базисным темпам роста:

$$ТН_i = \frac{\Delta y_{ц_i}}{y_{0_i}} = \frac{y_i - y_{i-1}}{y_{0_i}} = Тр_{б_i} - Тр_{б_{i-1}}. \quad (8.12)$$

Формула (8.12) удобна для практики, так как статистическая информация о динамике социально-экономических явлений публикуется чаще всего в виде базисных рядов динамики.

Применение формулы (8.12) проиллюстрируем на базисных темпах роста товарооборота магазина (см. табл. 8.1). Для определения темпа наращивания объема товарооборота в 1998 г. найдем разность между базисными темпами роста 1998 и 1997 гг.:

$$T_{H_{1998}} = 122,9 - 116,1 = 6,8\%.$$

8.3. Средние показатели

Для получения обобщающих показателей динамики социально-экономических явлений определяются средние величины: средний уровень, средний абсолютный прирост, средний темп роста и прироста и др.

Средний уровень ряда ($\bar{y}$) динамики характеризует типическую величину абсолютных уровней.

В интервальных рядах динамики средний уровень $\bar{y}$ определяется делением суммы $\sum y_i$ на число уровней n :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{y_1 + y_2 + \dots + y_n}{n}. \quad (8.13)$$

Применение формулы (8.13) проиллюстрируем на данных *интервального ряда динамики* товарооборота магазина в 1994 — 1998 гг.

$$\bar{y} = \frac{885,7 + 932,6 + 980,1 + 1028,7 + 1088,4}{5} = 983,1 \text{ тыс. руб.}$$

В *моментном ряду динамики* с равностоящими датами времени средний уровень определяется по формуле

$$\bar{y} = \frac{\frac{1}{2} y_1 + y_2 + \dots + \frac{1}{2} y_n}{n - 1}. \quad (8.14)$$

Пример. Имеются данные о валютном курсе ММВБ (руб./долл.):

Дата	13.12.93	14.12.93	15.12.93	16.12.93	17.12.93
Курс	1231	1237	1247	1247	1250

$$\bar{y} = \frac{\frac{1231}{2} + 1237 + 1247 + 1247 + \frac{1250}{2}}{5 - 1} = \frac{4971,5}{4} = 1242,9 \text{ руб./долл.}$$

В моментном ряду динамики с неравноотстоящими датами средний уровень определяется по формуле:

$$\bar{y} = \frac{\sum t_i y_i}{\sum t_i} = \frac{t_1 y_1 + t_2 y_2 + \dots + t_n y_n}{t_1 + t_2 + \dots + t_n}, \quad (8.15)$$

где y_i — уровни ряда динамики, сохранившиеся без изменения в течение промежутка времени t_i .

Применение формулы (8.15) проиллюстрируем на данных о состоянии численности работников магазина в апреле 1998г. С 1 по 20 апреля в списочном составе работников магазина значилось 190 человек, с 21 апреля и до конца месяца числилось 196 человек. Тогда в соответствии с формулой (1.15) среднедневная (списочная) численность работников магазина в апреле составила:

$$\bar{y} = \frac{20 \cdot 190 + 10 \cdot 195}{30} = 192 \text{ человека.}$$

Средний абсолютный прирост ($\Delta \bar{y}$) представляет собой обобщенную характеристику индивидуальных абсолютных приростов ряда динамики. Для определения среднего абсолютного прироста $\Delta \bar{y}$ сумма цепных абсолютных приростов $\sum \Delta y_{ц_i}$ делится на число приростов n :

$$\Delta \bar{y} = \sum \Delta y_{ц_i} : n. \quad (8.16)$$

Применение формулы (8.16) проиллюстрируем на данных табл. 1.2 и цепных абсолютных приростах товарооборота магазина:

$$\Delta \bar{y} = \frac{46,9 + 47,5 + 48,6 + 59,7}{4} = 50,7 \text{ тыс. руб.}$$

Средний абсолютный прирост может определяться по абсолютным уровням ряда динамики. Для этого определяется разность между конечным y_n и базисным y_0 уровнями изучаемого периода, которая делится на $m - 1$ субпериодов:

$$\Delta \bar{y} = \frac{y_n - y_0}{m - 1}. \quad (8.17)$$

Применение формулы (8.17) проиллюстрируем на данных табл.8.1.

$$\Delta \bar{y} = \frac{1088,4 - 885,7}{5 - 1} = 50,7 \text{ тыс. руб.}$$

Основываясь на взаимосвязи между цепными и базисными абсолютными приростами (8.3), показатель среднего абсолютного прироста можно определить по формуле

$$\Delta \bar{y} = \frac{\Delta y_{\delta_n}}{m-1}. \quad (8.18)$$

Применение формулы (8.18) проиллюстрируем на данных о базисном приросте товарооборота магазина в 1998 г. (см. табл.8.1):

$$\Delta \bar{y} = \frac{202,7}{5-1} = 50,7 \text{ тыс. руб.}$$

Средний темп роста ($\bar{T}_p$) — обобщающая характеристика индивидуальных темпов роста ряда динамики. Для определения среднего темпа роста $\bar{T}_p$ применяется формула

$$\bar{T}_p = \sqrt[n]{T_{p_1} \cdot T_{p_2} \cdot \dots \cdot T_{p_n}}. \quad (8.19)$$

где $T_{p_1}, T_{p_2}, \dots, T_{p_n}$ — индивидуальные (цепные) темпы роста (в коэффициентах);

n — число индивидуальных темпов роста.

Применение формулы (8.19) проиллюстрируем на данных табл.8.1 о цепных темпах роста розничного товарооборота магазина, заменив процентное их выражение коэффициентами: 1995 — 1,053; 1996 г. — 1,051 и т.д.

По формуле (8.19) имеем:

$$\bar{T}_p = \sqrt[4]{1,053 \cdot 1,051 \cdot 1,049 \cdot 1,058} = 1,053 \text{ или } 105,3\%.$$

Средний темп роста можно определить и по абсолютным уровням ряда динамики по формуле

$$\bar{T}_p = \sqrt[m-1]{y_n : y_0}. \quad (8.20)$$

Применение формулы (1.20) проиллюстрируем на данных об объемах розничного товарооборота магазина в 1994 — 1998 гг.

$$\bar{T}_p = \sqrt[5-1]{1088,4 : 885,7} = 1,053 \text{ или } 105,3\%.$$

На основе взаимосвязи между цепными и базисными темпами роста (1.6) средний темп роста можно определить по формуле

$$\bar{T}_p = \sqrt[m-1]{T_{p_{\delta_i}}}. \quad (8.21)$$

Применяя формулу (8.21), рассчитаем среднегодовой темп роста товарооборота магазина за 1994 — 1998 гг.:

$$\bar{T}_p = \sqrt[5-1]{1,229} = 1,053 \text{ или } 105,3\%.$$

Средний темп прироста ($\bar{T}_П$) можно определить на основе взаимосвязи между темпами роста и прироста. При наличии данных о средних темпах роста $\bar{T}_р$ для получения средних темпов прироста $\bar{T}_П$ используется зависимость:

$$\bar{T}_П = \bar{T}_р - 1 \quad (8.22)$$

(при выражении среднего темпа роста в коэффициентах).

Применяя формулу (8.22) можно вычислить средний темп прироста объема розничного товарооборота магазина за 1994 — 1998 гг. (см. табл. 8.1) на основе среднего темпа роста по (8.19) и (8.20) как 1,05:

$$\bar{T}_П = 1,053 - 1 = 0,053 \text{ или } 5,3\%.$$

8.4. Изучение основной тенденции развития

Важным направлением в исследовании закономерностей динамики социально-экономических процессов является изучение общей тенденции развития (тренда). Это можно осуществить, применяя специальные методы анализа рядов динамики. Конкретное их использование зависит от характера исходной информации и предопределяется задачами анализа.

Изменения уровней рядов динамики обуславливается влиянием на изучаемое явление ряда факторов, которые, как правило, неоднородны по силе, направлению и времени их действия. Постоянно действующие факторы оказывают на изучаемые явления определяющее влияние и формируют в рядах динамики основную тенденцию развития (тренд). Воздействие других факторов проявляется периодически. Это вызывает повторяемые во времени колебания уровней рядов динамики. Действие разовых (спорадических) факторов отображается случайными (кратковременными) изменениями уровней рядов динамики.

Различные результаты действия постоянных, периодических и разовых причин и факторов на уровни развития социально-экономических явлений во времени обуславливают необходимость изучения основных компонентов рядов динамики: тренда, периодических колебаний, случайных отклонений.

При изучении в рядах динамики основной тенденции развития (тренда) решаются две взаимосвязанные задачи: выявление в изучаемом явлении наличия тренда с описанием его качественных особенностей; измерение выявленного тренда, т.е. получение обобщающей количественной оценки основной тенденции развития.

На практике наиболее распространенными методами статистического изучения тренда являются: укрупнение интервалов, сглаживание скользящей средней, аналитическое выравнивание.

Метод укрупнения интервалов применяется для выявления тренда в рядах динамики колеблющихся уровней, затушевывающих основную тенденцию развития. Главное в этом методе заключается в преобразовании первоначального ряда динамики в ряды более продолжительных периодов (месячные в квартальные, квартальные в годовые и т.д.).

Рассмотрим применение метода укрупнения интервалов на данных о реализации радиоприемников в магазинах города (шт):

Январь	3662	Июль	3803
Февраль	3096	Август	3812
Март	2956	Сентябрь	3921
Апрель	3805	Октябрь	4442
Май	3364	Ноябрь	3824
Июнь	2946	Декабрь	3976

Различные направления изменений по отдельным месяцам уровней данного ряда динамики затрудняют выводы об основной тенденции продажи радиоприемников. Решение этой задачи упростится, если соответствующие месячные уровни объединять в квартальные: I — 9714 шт., II — 10115 шт., III — 11536 шт., IV — 12242 шт. После укрупнения интервалов основная тенденция роста продажи радиоприемников стала очевидной (тыс. шт.):

$$9,7 < 10,1 < 11,5 < 12,2.$$

Для статистического изучения тренда применяется так называемое **сглаживание методом скользящей средней**. 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.

Применение этого метода рассмотрим на данных о реализации продуктов сельскохозяйственного производства магазинами города (таб. 8.2, гр. 1, 2).

Особенностью данных табл. 8.2 являются периодическая колеблемость квартальных уровней, увеличение уровня продажи во II и III кварталах и некоторое снижение в IV квартале. Основная тенденция развития непосредственно не просматривается.

Для выявления основной тенденции развития методом скользящей средней прежде всего устанавливаются ее звенья. Звенья сколь-

зующей средней должны составляться из числа уровней, отвечающих длительности внутригодовых циклов в изучаемом явлении.

Таблица 8.2 Реализация продуктов сельскохозяйственного производства (тыс.руб)

Год, квартал	Исходные уровни y_i	Скользящие средние $\bar{y}_i$	Сглаженные уровни с центрированием $\bar{y}_{\hat{C}}$
1	2	3	4
1-й год			
I кв.	175	—	—
II кв.	263	1061:4=265,25	—
III кв.	326	1133:4=283,25	274,25
IV кв.	297	1168:4=292,00	287,60
2-й год			
I кв.	247	1208:4=302,00	297,00
II кв.	298	1252:4=313,00	307,50
III кв.	366	1425:4=356,25	334,60
IV кв.	341	1568:4=392,00	374,10
3-й год			
I кв.	420	—	402,90
II кв.	441	1655:4=412,75	421,00
III кв.	453	1713:4=428,25	429,00
IV кв.	399	1719:4=429,25	430,75
4-й год			
I кв.	426	1727:4=431,75	435,37
II кв.	449	1756:4=439,00	446,62
III кв.	482	1817:4=454,25	—
IV кв.	460	—	—

Для ряда динамики, отображающего развитие товарооборота по кварталам, скользящие средние обычно составляются из четырехчленных звеньев. Их расчет состоит в определении средних величин из четырех уровней ряда с отбрасыванием при вычислении каждой новой скользящей средней одного уровня слева и присоединением одного уровня справа:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3 + y_4}{4}; \bar{y}_2 = \frac{y_2 + y_3 + y_4 + y_5}{4} \text{ и т.д.}$$

В нашем примере исчисляются 13 скользящих средних (табл. 8.2, гр. 3).

Для четного числа уровней каждое значение скользящей средней приходится на промежуток между двумя смежными кварталами. Так, первая скользящая средняя (265,25) записывается между II и III квар-

талами, вторая (283,25) — между III и IV кварталами и т.д. Для определения сглаженных уровней производится центрирование ($\hat{C}$). Для III квартала определяется срединное значение между первой и второй скользящими средними: $(265,25 + 283,25) : 2 = 274,25$ тыс. руб., для IV квартала центрируются вторая и третья скользящие средние:

$$(283,25 + 292) : 2 = 287,6 \text{ тыс. руб. и т.д.}$$

Полученные значения сглаженных уровней помещены в гр. 4 табл. 8.2. Из их графического изображения отчетливо видна основная тенденция развития торговли (рис. 8.1).

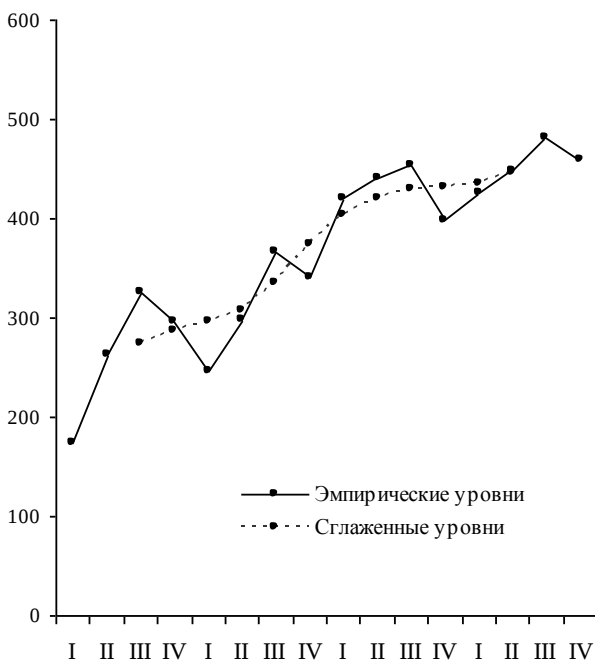


Рис. 8.1. Реализация сельскохозяйственной продукции по кварталам четырехлетия

Если при сглаживании рядов динамики звенья скользящей средней составляют из нечетного числа уровней, то необходимость в центрировании отпадает.

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 более высокого

порядка задачи — измерением тренда — достигается *методом аналитического выравнивания*.

Основным содержанием метода аналитического выравнивания в рядах динамики является то, что основная тенденция развития y_i рассчитывается как функция времени

$$y_{t_i} = f(t_i). \quad (8.23)$$

Определение теоретических (расчетных) уровней y_{t_i} производится на основе так называемой *адекватной математической функции*, которая наилучшим образом отображает основную тенденцию ряда динамики.

Подбор адекватной функции осуществляется методом наименьших квадратов — минимальностью отклонений суммы квадратов между теоретическими y_{t_i} и эмпирическими y_i уровнями:

$$\sum (y_{t_i} - y_i)^2 = \min. \quad (8.24)$$

Значение уравнения (8.24) состоит в том, что при изучении тренда оно принимается в качестве критерия оценки соответствия расчетных (теоретических) уровней с фактическими (эмпирическими) уровнями ряда динамики.

Важнейшей проблемой, требующей своего решения при применении метода аналитического выравнивания, является подбор математической функции, по которой рассчитываются теоретические уровни тренда.

В практике статистического изучения тренда различают следующие эталонные типы развития социально-экономических явлений во времени:

1) *равномерное развитие*. Для этого типа динамики присущи постоянные абсолютные приросты:

$$\Delta y_n = \text{const}. \quad (8.25)$$

Основная тенденция развития в рядах динамики со стабильными приростами отображается уравнением линейной функции:

$$\bar{y}_t = a_0 + a_1 t, \quad (8.26)$$

где a_0 и a_1 — параметры уравнения;

t — обозначение времени.

Параметр a_1 является коэффициентом регрессии, определяющим направление развития. Если $a_1 > 0$, то уровни ряда динамики равно-

мерно возрастают, а при $a_1 < 0$ происходит их равномерное снижение;

2) *равноускоренное (равнозамедленное) развитие*. Этому типу динамики свойственно постоянное во времени увеличение (замедление) развития. Уровни таких рядов динамики изменяются с постоянными темпами прироста:

$$T_{n_{\text{ц}}} = \text{const}. \quad (8.27)$$

Основная тенденция развития в рядах динамики со стабильными темпами прироста отображается функцией параболы второго порядка:

$$\bar{y} = a_0 + a_1 t + a_2 t^2. \quad (8.28)$$

В формуле (1.28) значения параметров a_0 и a_1 идентичны параметрам, используемым в формуле (1.26). Параметр a_2 характеризует *постоянное изменение интенсивности развития (в единицу времени)*. При $a_2 > 0$ происходит ускорение развития, при $a_2 < 0$ идет процесс замедления роста. Параметр a_1 может быть как со знаком плюс, так и со знаком минус;

3) *развитие с переменным ускорением (замедление)*. Для этого типа динамики основная тенденция развития выражается функция параболы третьего порядка:

$$\bar{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3. \quad (8.29)$$

В уравнении (1.29) параметр a_3 отображает изменение ускорения. При $a_3 > 0$ ускорение возрастает, а при $a_3 < 0$ ускорение замедляется;

4) *развитие по экспоненте*. Этот тип динамики характеризуют стабильные темпы роста

$$T p_{\text{ц}} = \text{const}. \quad (8.30)$$

Основная тенденция в рядах динамики с постоянными темпами роста отображается показательной функцией:

$$\bar{y}_t = a_0 a_1^t, \quad (8.31)$$

где a_1 — темп роста (снижения) изучаемого явления в единицу времени, т.е. интенсивность развития;

5) *развитие в замедлением в конце периода*. У этого типа динамики показание цепного абсолютного прироста сокращается в **конечных уровнях ряда динамики**:

$$\Delta y_{цп} \rightarrow 0. \quad (8.32)$$

Основная тенденция развития в таких рядах динамики выражается полулогарифмической функцией:

$$\bar{y} = a_0 + a_1 \lg t. \quad (8.33)$$

При аналитическом выравнивании в рядах динамики можно применить и другие математические функции. Так, при изучении основной тенденции **неудовлетворенного и реализованного спроса населения** применяются:

$$\text{степенная функция} — \bar{y} = a_0 t^{a_1}, \quad (8.34)$$

$$\text{функция гиперболы} — \bar{y} = a_0 + a_1 \frac{1}{t}. \quad (8.35)$$

Применение метода аналитического выравнивания при статистическом изучении тренда проиллюстрируем на примере.

Пример. По данным о розничном товарообороте региона в 1992 — 1997 гг. (табл. 8.3) нужно произвести анализ основной тенденции развития товарооборота:

Таблица 8.3

Год	Объем розничного товарооборота, млрд. руб.	Темп роста по годам, %	Абсолютный прирост по годам, млрд. руб.
1	2	3	4
1992	11,18	—	—
1993	12,23	109,4	1,05
1994	13,28	108,6	1,05
1995	14,31	107,7	1,03
1996	15,36	107,3	1,05
1997	16,40	106,8	1,04
В среднем	14,32	107,9	1,04

Из табл. 8.3 видно, что развитие товарооборота происходило с **затухающими темпами роста** (гр. 3) и **относительно стабильными абсолютными приростами** (гр.4).

Для установления в данном ряду динамики типа развития определяющим признаком является характер изменения абсолютных приростов. Поскольку при среднем абсолютном приросте, равном 1,04 млрд. руб., величина их изменений незначительная ($\pm 0,01$ млрд. руб.), то анализируемый ряд динамики можно считать с равномерным развитием (см. формулу (8.25)). Поэтому для аналитического выравнивания применяется функция (8.26): $\bar{y}_t = a_0 + a_1 t$.

Для вычисления параметров функции (8.26) на основе требований метода наименьших квадратов (см. (8.24)) составляется система нормальных уравнений:

$$\begin{aligned} na_0 + a_1 \sum t &= \sum y; \\ a_0 \sum t + a_1 \sum t^2 &= \sum t \cdot y. \end{aligned} \quad (8.36)$$

Для решения системы уравнений (8.36) обычно применяют **способ определителей**, позволяющий получать более точные результаты за счет сведения к минимуму ошибки из-за округлений в расчетах параметров:

$$a_0 = \frac{\sum y \sum t^2 - \sum ty \sum t}{n \sum t^2 - \sum t \sum t}; \quad (8.37)$$

$$a_1 = \frac{n \sum ty - \sum t \sum y}{n \sum t^2 - \sum t \sum t}. \quad (8.38)$$

Применительно к анализируемым данным для определения алгоритмов (8.37) и (8.38) составляется матрица расчетных показателей (табл. 8.4).

Таблица 8.4

Год	Объем розничного товарооборота, млрд. руб.	t_i	t_i^2	$t_i y_i$	y_{t_i}
1992	11,18	1	1	11.18	11,183
1993	12,25	2	4	24.46	12,226
1994	13,28	3	9	39.84	13,269
1995	14,31	4	16	57.24	14,312
1996	15,36	5	25	76.80	15,355
1997	16,40	6	36	98.40	16,398
Σ	82.76	21	91	307.92	82.743

По итоговым данным табл. 8.4 определяем по формуле (8.37):

$$a_0 = \frac{82,76 \cdot 91 - 307,92 \cdot 21}{6 \cdot 91 - 21 \cdot 21} = 10,14.$$

по формуле (8.28):

$$a_1 = \frac{6 \cdot 307,92 - 21 \cdot 82,76}{6 \cdot 91 - 21 \cdot 21} = 1,043.$$

По вычисленным данным производим **синтезирование трендовой модели функции** (см. формулу 8.26)):

$$\bar{y}_t = 10,14 + 1,043 \cdot t. \quad (8.39)$$

На основе модели (8.39) определяются теоретические уровни тренда (y_{t_i}) для каждого года анализируемого ряда динамики:

$$\bar{y}_{t_{1992}} = 10,14 + 1,043 \cdot 1 = 11,183 \text{ млрд. руб.};$$

$$\bar{y}_{t_{1993}} = 10,14 + 1,043 \cdot 2 = 12,226 \text{ млрд. руб.};$$

$$\dots\dots\dots$$

$$\bar{y}_{t_{1997}} = 10,14 + 1,043 \cdot 6 = 16,398 \text{ млрд. руб.}$$

Вычисленные теоретические уровни помещены в гр. 6 табл. 8.4. Правильность расчетов проверяется по равенству:

$$\sum y_i = \sum y_{t_i}. \quad (8.40)$$

Несовпадение в равенстве (8.40) на 0,017 млрд. руб. объясняется округлениями в расчетах.

Параметр a_1 трендовой модели (8.39) показывает, что объем розничного товарооборота региона возрастал в среднем на 1,042 млрд. руб. в год.

Для подтверждения гипотезы о возможном типе развития можно использовать графический метода. Наглядное изображение анализируемого ряда динамики позволяет получать образное представление о размещении на поле графика эмпирических уровней. Это способствует лучшему осмыслению специфики изменений в ряду динамики. Но дать обобщенную статистическую оценку выявленного тренда графический метод не может.

Из характера размещения уровней анализируемого ряда динамики на поле графика (рис. 8.2) можно сделать предположение о возможном применении тренда при аналитическом изучении ряда математических функций. Это может быть и уравнение прямолинейной функции (8.26) и уравнение показательной кривой (8.31). Для выбора наиболее адекватной из них следует осуществить сравнительный анализ тренда исходных данных способом перебора решений по намеченным математическим функциям.

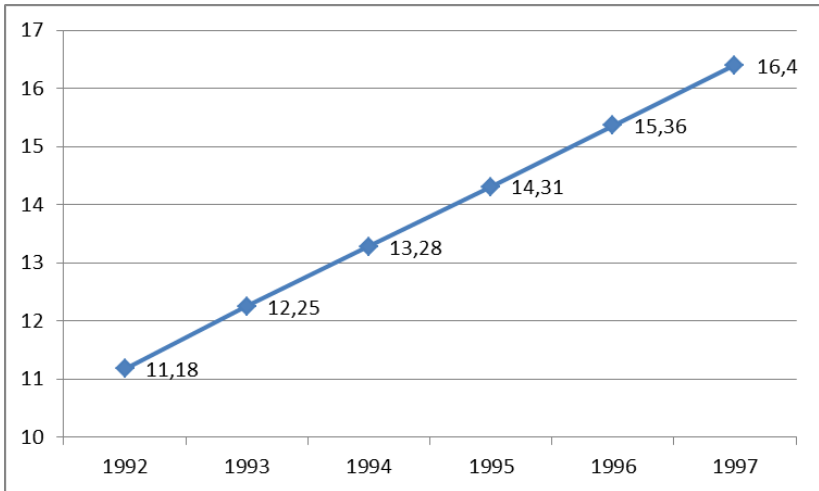


Рис. 8.2. Розничный товарооборот региона в 1992 — 1997 гг

Одним из применяемых в практике статистического изучения тренда показателей *адекватности математической функции* является *стандартизованная ошибка аппроксимации* σ_{y_t} :

$$\sigma_{y_t} = \sqrt{\frac{\sum (y_{t_i} - \hat{y}_t)^2}{n}}. \quad (8.41)$$

Применение в изучении тренда формулы (1.48) основано на том, что за наиболее адекватную принимается функция, у которой стандартизованная ошибка аппроксимации минимальная.

Построив уравнение регрессии, проводят оценку его надежности. Это делается посредством критерия Фишера (F). Фактический уровень ($F_{\text{факт}}$) сравнивается с теоретическим (табличным) значением:

$$F_{\text{факт}} = \frac{\frac{1}{k-1} \sigma_{\text{факт}}^2}{\frac{1}{n-k} \sigma_{\text{ост}}^2}, \quad F_{\text{факт}} = \frac{\sigma_{\text{факт}}^2 (n-k)}{\sigma_{\text{ост}}^2 (k-1)},$$

где k — число параметров функции, описывающей тенденцию;
 n — число уровней ряда.

$$\sigma_{\text{ост}}^2 = \frac{\sum (Y - f(t))^2}{n} - \text{остаточная дисперсия;}$$

$$\sigma_{\text{факт}}^2 = \sigma_y^2 - \sigma_{\text{ост}}^2 = \frac{\sum (f(t) - y)^2}{n} - \text{факторная дисперсия;}$$

$$\sigma_y^2 = \frac{\sum (y - \bar{y})^2}{n} = \sigma_{\text{факт}}^2 + \sigma_{\text{ост}}^2 - \text{полная дисперсия.}$$

$F_{\text{факт}}$ сравнивается с табличным значением $F_{\text{теор}}$ при $\nu_1 = (k-1)$, $\nu_2 = (n-k)$ степенях свободы и уровне значимости α (обычно $\alpha = 0,05$). Если $F_{\text{факт}} > F_{\text{теор}}$, то уравнение регрессии значимо, т.е. основная модель адекватна фактической временной тенденции.

Для оценки точности модели вычисляют коэффициент детерминации:

$$R^2 = \frac{\sigma_{\text{факт}}^2}{\sigma_y^2}, \quad 0 \leq R^2 \leq 1.$$

Если R^2 близко к 1, то модель близка к реальному процессу.

8.5. Анализ сезонных колебаний

Если в анализируемой временной последовательности наблюдаются устойчивые отклонения от тенденции (как в большую, так и в меньшую сторону), то можно предположить наличие в ряду динамики некоторых (одного или нескольких) колебательных процессов. Это особенно заметно, когда изучаемые явления имеют сезонный характер, — возрастание или убывание уровней повторяется регулярно с интервалом в один год (например, производство молока и мяса по месяцам года, потребление топлива и электроэнергии для бытовых нужд, сезонная продажа товаров и т.д.).

Чтобы избавиться от влияний особенностей некоторых лет, часто внутригодовые и среднегодовые данные вычисляются как средние за несколько лет.

Уровень сезонности оценивается с помощью:

- 1) индексов сезонности;
- 2) гармонического анализа.

Индексы сезонности показывают, во сколько раз фактический уровень ряда в момент или интервал времени t больше среднего уровня либо уровня, вычисляемого по уравнению функции тенденции $f(t)$. При анализе сезонности уровни временного ряда показывают развитие явления по месяцам (кварталам) одного или нескольких лет. Для каждого месяца (квартала) получают обобщенный индекс сезонности как среднюю арифметическую из одноименных индексов каждого года. Индексы сезонности — это, по существу, относительные величины координации, когда за базу сравнения принят либо средний уровень ряда, либо уровень тенденции. Способы определения индексов сезонности зависят от наличия или отсутствия основной тенденции.

Если тренда нет или он незначителен, то для каждого месяца (квартала) индекс сезонности i_t равен

$$i_t = \frac{Y_t}{\bar{Y}},$$

где Y_t — средний уровень показателя за месяц (квартал) t , вычисленный по одноименным месяцам нескольких лет;

$\bar{Y}$ — общий средний уровень показателя.

Пример. Имеются данные об объеме продаж акций на 15 крупнейших биржах России за 2003 г. (млн. руб.):

Месяц	Уровень показателя (Y_t)	$i_{t,сез}$
Январь	12,78	0,027
Февраль	122,08	0,254
Март	709,98	1,477
Апрель	1602,61	3,334
Май	651,83	1,356
Июнь	220,80	0,459
Июль	327,68	0,682
Август	277,12	0,576
Сентябрь	418,31	0,870
Октябрь	521,18	1,084
Ноябрь	396,20	0,824
Декабрь	508,34	1,057

$$\bar{Y} = \frac{5768,91}{12} = 480,425 \text{ млн. руб. в месяц.}$$

Как видно, в 2003 г. были зарегистрированы три пика объемов продаж акций: в апреле, октябре и декабре.

При наличии тренда индекс сезонности определяется на основе методов, исключающих влияние тенденции. Порядок расчета следующий:

1) рассчитывают отношения $i_t = Y_t / f(t)$;

2) при необходимости находят среднее из этих отношений для одноименных месяцев (кварталов) $\bar{i}_t = \frac{\sum i_t}{T}$, где T — число лет.

Для анализа внутригодовой динамики социально-экономических явлений могут применяться *гармоники ряда Фурье*.

При аналитическом выражении изменений уровней ряда динамики используется формула

$$y_t = a_0 + \sum (a_k \cos kt + b_k \sin kt). \quad (8.42)$$

В формуле (1.49) k определяет номер гармоники, который используется с различной степенью точности (обычно от 1 до 4).

При решении уравнения (8.42) параметры определяются на основе положений метода наименьших квадратов (8.22). Определяя для функции (8.42) частные производные и приравнивая их к нулю, получают систему нормальных уравнений, параметры которых вычисляются по формулам:

$$a_0 = \frac{\sum y_i}{n}; \quad (8.43)$$

$$a_k = \frac{2}{n} \sum y_i \cos kt_i; \quad (8.44)$$

$$b_k = \frac{2}{n} \sum y_i \sin kt_i. \quad (8.45)$$

При анализе ряда внутригодовой динамики по месяцам значение k принимается за 12. Подставляя месячные периоды как части окружности, ряд внутригодовой динамики можно записать в таком виде:

Периоды (t_i)	0	$\frac{1}{6}\pi$	$\frac{1}{3}\pi$	$\frac{1}{2}\pi$	$\frac{2}{3}\pi$	$\frac{5}{6}\pi$	π	$\frac{7}{6}\pi$	$\frac{4}{3}\pi$	$\frac{3}{2}\pi$	$\frac{5}{3}\pi$	$\frac{11}{6}\pi$
Уровни (y_i)	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}

Проиллюстрируем построение модели внутригодовой динамики по первой гармонике ряда Фурье на данных о розничном товарообороте по месяцам 1997 г. (табл. 8.5).

Применяя первую гармонику ряда Фурье, определяются параметры уравнения (8.42):

по формуле (8.43)
$$a_0 = \frac{366,4}{12} = 30,5 ;$$

по формуле (8.44)
$$a_1 = \frac{2 \cdot (-2,4)}{12} = -0,4 ;$$

по формуле (8.45)
$$b_1 = \frac{2 \cdot (-7,8)}{12} = -1,3 .$$

По полученным параметрам синтезируется математическая модель:

$$\bar{y}_t = 30,5 - 0,4 \cos t - 1,3 \sin t . \quad (8.46)$$

Таблица 8.5 Розничный товарооборот по месяцам 2007 г.

Месяц	t_i	Объем розничного товарооборота, млрд. руб.	$\cos t_i$	$\sin t_i$	$y_i \cos t_i$	$y_i \sin t_i$	y_{t_i}
1	2	3	4	5	6	7	8
Январь	0	27,3	1,0	0,0	27,3	0,0	30,1
Февраль	(1:6) π	28,0	0,866	0,5	24,2	14,0	29,6
Март	(1:3) π	31,2	0,5	0,866	15,6	27,0	29,2
Апрель	(1:2) π	30,1	0,0	1,0	0,0	30,1	29,2
Май	(2:3) π	29,2	-0,5	0,866	-14,6	25,3	29,6
Июнь	(5:6) π	30,0	-	0,5	-26,6	15,0	30,2
			0,866				
Июль	π	30,1	-1,0	0,0	-30,1	0,0	30,9
Август	(7:6) π	32,0	-	-0,5	-27,7	-16,0	31,5
			0,866				
Сентябрь	(4:3) π	31,4	-0,5	-0,866	-15,7	-27,2	31,8
Октябрь	(3:2) π	32,3	0,0	-1,0	0,0	-32,3	31,8
Ноябрь	(5:3) π	31,2	0,5	-0,866	15,6	-27,0	31,4
Декабрь	(11:6) π	33,5	0,866	-0,5	29,0	-16,7	30,9
	$\times$	366,4	$\times$	$\times$	-2,4	-7,8	366,2

На основе модели (8.46) определяются для каждого месяца расчетные уровни y_{t_i} :

$$y_{t_{\text{я}}} = 30,5 - 0,4 \cdot 1,0 - 1,3 \cdot 0 = 30,1 \text{ млрд. руб.};$$

$$y_{t_{\phi}} = 30,5 - 0,4 \cdot 0,866 - 1,3 \cdot 0,5 = 29,5 \text{ млрд. руб.};$$

$$y_{t_d} = 30,5 - 0,4 \cdot 0,866 - 1,3 \cdot (-0,5) = 30,9 \text{ млрд. руб.}$$

Вычисленные для каждого месяца 1997 г. теоретические уровни y_{t_i} записаны в гр. 8 табл. 8.5. Итоговые данные этой графы свидетельствуют о достаточно точном распределении выровненных данных. Отклонение $\sum y_{t_i}$ от $\sum y_i$ объясняется неизбежными округлениями в расчетах.

Тест для самоконтроля

1. Если уровень ряда текущего года меньше уровня предыдущего, то абсолютный цепной прирост
 - а) положительный;
 - б) отрицательный;
 - в) равен нулю.
2. Тренд — это
 - а) сезонные колебания;
 - б) основная тенденция;
 - в) случайные изменения.
3. Ряд динамики характеризует
 - а) структуру совокупности по какому-либо признаку
 - б) колебания признака во времени
 - в) вариацию признака в совокупности
4. Выпуск продукции фирмой за первый год увеличился в 3 раза, а за следующий год поднялся еще на 100%. Среднегодовой темп роста выпуска продукции составил
 - а) 250 % б) 300 % в) 167 % г) 245 %
5. Численную оценку точности модели тренда можно получить при помощи
 - а) критерия Фишера
 - б) индекса сезонности
 - в) коэффициента детерминации
 - г) гармонического анализа

6. Формула $\bar{y}_i = \frac{y_{i-1} + y_i + y_{i+1}}{3}$ используется для выравнивания вре-

менного ряда методом

- а) укрупнения интервалов.
- б) аналитическим (линейная модель).
- в) моделирования
- г) скользящей средней.

Вопросы и упражнения

1. Назовите элементы временного ряда.
2. Каковы цели статистического изучения рядов динамики социально-экономических явлений?
3. В чем отличие базисных и цепных показателей рядов динамики?
4. Может ли темп роста быть отрицательной величиной?
5. Какая средняя величина применяется для расчета среднегодовой численности населения?
6. В каком случае коэффициент детерминации может быть равен единице?
7. Приведите пример социально-экономического явления с сезонными колебаниями.
8. Что такое «индекс сезонности»?
9. Назовите виды математических функций, которые применяются для аналитического выравнивания рядов динамики.
10. Какие виды статистических графиков применяются для изображения рядов динамики?

9. ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ

9.1. Основные понятия и определения

Индексы относятся к важнейшим обобщающим показателям. Слово «индекс» имеет несколько значений: показатель, указатель, опись, реестр. Оно используется как понятие в математике, экономике, в метеорологии и других науках.

В статистике под **индексом** понимается относительный показатель, который выражает соотношение величин какого-либо явления во времени, в пространстве или сравнение фактических данных с любым эталоном (план, прогноз, норматив и т.д.).

В международной практике индексы принято обозначать символами i и I (начальная буква латинского слова index). Буквой « i » обозначаются индивидуальные (частные) индексы, буквой « I » — общие индексы. Знак внизу справа означает период: 0 — базисные, 1 — отчетный. Помимо этого используются определенные символы для обозначения индексируемых показателей:

q — количество (объем) какого-либо товара в натуральном выражении;

p — цена единицы товара;

z — стоимость единицы товара;

t — затраты времени на производство единицы продукции;

w — выработка продукции в стоимостном выражении на одного рабочего или в единицу времени;

v — выработка продукции в натуральном выражении на одного рабочего или в единицу времени;

T — общие затраты времени (tq) или численность рабочих;

pq — стоимость продукции или товарооборот;

zq — издержки производства.

Все экономические индексы можно классифицировать по следующим признакам:

- степень охвата явления;
- база сравнения;
- вид весов (соизмерителя);
- характер объекта исследования;
- объект исследования;
- состав явления;

- форма построения;
- период исчисления.

По степени охвата явления индексы бывают индивидуальные и сводные. *Индивидуальные* индексы служат для характеристики изменения отдельных элементов сложного явления. Их примером могут быть измерения объема производства отдельных видов продукции (телевизоров, электроэнергии и т.д.), а также цен на акции какого-либо предприятия. Для измерения динамики сложного явления, составные части которого непосредственно несоизмеримы (изменение физического объема продукции, включающей разноименные товары, индекса цен акций предприятия, региона и т.д.) рассчитывают *сводные* или *общие* индексы.

Если индексы охватывают не все элементы сложного явления, а только часть их, то такие индексы называются *групповыми* или *субиндексами*, например, индексы физического объема продукции по отдельным отраслям промышленности, индексы цен по группам продовольственных и непродовольственных товаров. Групповые индексы отражают закономерности в развитии отдельных частей изучаемых явлений. В таких индексах проявляется их связь с методом группировок.

По базе сравнения индексы можно разделить на две группы: *динамические* и *территориальные*. Первая группа индексов отражает изменение явления во времени. Например, индекс цен на продукцию в 2006 г. по сравнению с предыдущим годом; индекс стоимости потребительской корзины в августе по сравнению с июлем 2007 г.

При исчислении динамических индексов происходит сравнение значения показателя в отчетный период со значением этого показателя за предыдущий период, который называют *базисным*. Однако в качестве последнего могут быть использованы и прогнозные, и плановые показатели.

Динамические индексы бывают *базисные* и *цепные*.

Вторая группа индексов (*территориальные*) применяется для межрегиональных сравнений. Большое значение эти индексы имеют в международной статистике при сопоставлении показателей социально-экономического развития различных стран. Например, индекс цен на фототовары в Италии по сравнению с Германией, стоимости потребительской корзины в Москве по сравнению с Санкт-Петербургом.

По виду весов индексы бывают с *постоянными* и *переменными* весами.

В зависимости от формы построения различаются индексы *агрегатные* и *средние*. Последние делятся на арифметические и гармо-

нические. Агрегатная форма общих индексов является основной формой экономических индексов. Средние индексы — производные, они получаются в результате преобразования агрегатных индексов.

По характеру объема исследования общие индексы подразделяются на индексы *количественных (объемных)* и *качественных* показателей. В основе такого деления индексов лежит вид индексируемой величины. К первой группе индексов относятся, например, индексы объема продаж долларов США на Московской межбанковской валютной бирже, а ко второй — индекс курса немецкой марки.

По объекту исследования индексы бывают: производительности труда, себестоимости, физического объема продукции, стоимости продукции и т.д.

По составу явления можно выделить две группы индексов: постоянного (фиксированного) состава и переменного состава. Деление индексов на эти две группы используется для анализа динамики средних показателей.

По периоду исчисления индексы подразделяются на годовые, квартальные, месячные, недельные.

С помощью экономических индексов решаются следующие задачи:

- измерение динамики социально-экономического явления за два и более периодов времени;
- измерение динамики среднего экономического показателя;
- определение соотношения показателей по разным регионам;
- определение степени влияния изменений значений одних показателей на динамику других;
- пересчет значения макроэкономических показателей из фактических цен в сопоставимые.

Каждая из этих задач решается с помощью различных индексов.

9.2. Индивидуальные и общие индексы

Индивидуальные индексы получают в результате сравнения однотоварных явлений. Например, индекс цен на подсолнечное масло определяется как отношение цены на этот товар в текущем периоде к цене базисного периода.

Индивидуальные индексы представляют собой относительные величины динамики, выполнения плана, сравнения, и их расчет не требует специальных правил.

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

Индекс физического объема продукции i_q рассчитывается по формуле

$$i_q = \frac{q_1}{q_0}. \quad (9.1)$$

Этот индекс показывает, во сколько раз возрос (уменьшился) выпуск какого-либо товара в отчетном периоде по сравнению с базисным, или сколько процентов составляет рост (снижение) выпуска товара. Если из значения индекса, выраженного в процентах, вычесть 100%, то полученная величина покажет, на сколько процентов возрос (уменьшился) выпуск продукции. В знаменателе может быть не только количество продукции, произведенной за какой-то предыдущий период, но и плановое значение ($q_{пл}$), нормативное (q_n) или эталонное значение, принятое за базу сравнения (q_3). Тогда формула (2.1) примет соответствующий вид:

$$i_q = \frac{q_1}{q_{пл}}; \quad (9.2)$$

$$i_q = \frac{q_1}{q_n}; \quad (9.3)$$

$$i_q = \frac{q_1}{q_3}. \quad (9.4)$$

Индексы других показателей строятся аналогично.

Индивидуальный индекс цен

$$i_p = \frac{p_1}{p_0} \quad (9.5)$$

характеризует изменение цены одного определенного товара в текущем периоде по сравнению с базисным.

Индивидуальный индекс себестоимости единицы продукции

$$i_z = \frac{z_1}{z_0} \quad (9.6)$$

показывает изменение себестоимости единицы продукции в текущем периоде по сравнению с базисным.

Производительность труда может быть измерена количеством продукции, производимой в единицу времени (v), или затратами рабочего времени на производство единицы продукции (t). Поэтому можно построить:

- индекс количества продукции, произведенной в единицу времени:

$$i_v = \frac{v_1}{v_0} = \frac{q_1}{T} \cdot \frac{q_0}{T_0}; \quad (9.7)$$

- индекс производительности труда по трудовым затратам:

$$i_t = \frac{t_0}{t_1}. \quad (9.8)$$

Так как между количеством продукции, произведенной в единицу времени, и затратами рабочего времени на производство единицы продукции существует обратно пропорциональная зависимость, т.е.

$$t = \frac{I}{v}, \quad (9.9)$$

то индекс (9.8) получается в результате деления величины показателя в базисном периоде на величину в текущем периоде.

Для характеристики производительности труда часто используется индивидуальный индекс выработки продукции в стоимостном выражении на одного рабочего:

$$i_w = \frac{w_1}{w_0} = \frac{q_1 p}{T_1} : \frac{q_0 p}{T_0}, \quad (9.10)$$

где p — сопоставимые цены.

Индивидуальные индексы (9.7 и (9.10) показывают, во сколько раз производительность труда в базисном периоде выше (ниже), чем в отчетном.

Индекс, исчисленный по формуле (9.8), показывает, во сколько раз производительность труда в базисном периоде выше (ниже), чем в отчетном.

Индивидуальный индекс стоимости продукции отражает, во сколько раз изменилась стоимость какого-либо товара в текущем периоде по сравнению с базисным, или сколько процентов составляет рост (снижение) стоимости товара, и определяется по формуле

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0}. \quad (9.11)$$

Индивидуальный индекс численности рабочих можно рассчитать следующим образом:

$$i_T = \frac{T_1}{T_0} = \frac{t_1 q_1}{t_0 q_0}. \quad (9.12)$$

Он показывает, во сколько раз изменилась численность рабочих в текущем периоде по сравнению с базисным, или сколько процентов составляет рост (снижение) численности рабочих.

По данным о цене, количестве и стоимости проданных товаров (табл. 9.1) рассчитаем индивидуальные индексы. Результаты расчетов (представлены в гр. 7 – 9 этой таблицы) показывают, что больше всего возросли цены на кофе — почти на 6%, а индекс стоимости товаров и физического объема продукции самый высокий по чаю — соответственно 5,2 и 5 раз.

В экономических расчетах чаще всего используются общие индексы, которые характеризуют изменение совокупности в целом. Построение этих индексов и является содержанием индексной методологии. В индексной теории сложились две концепции: *синтетическая и аналитическая*. Они по-разному интерпретируют общие индексы.

Согласно **синтетической** концепции особенность общих индексов состоит в том, что они выражают относительное изменение сложных (разнотоварных) явлений, отдельные части или элементы которых непосредственно несоизмеримы, и поэтому индексы — показатели синтетические. Например, промышленные предприятия производят несколько видов продукции, имеющей различное назначение. Следовательно, путем суммирования количества произведенных товаров различных видов нельзя получить показатель физического объема продукции. Методология построения общих индексов предусматривает прежде всего приведение разнотоварных явлений к соизмеримому виду.

В **аналитической** теории индексы трактуются как показатели, необходимые для измерения влияния изменения составных частей, компонентов, факторов сложного явления на изменение уровня этого явления. Например, изменение общей величины товарооборота в текущем периоде по сравнению с базисным связано как с изменением физического объема продаж товаров, так и с изменением цен по каждому виду товаров. Поэтому индексная методология предусматривает определение влияния каждого из факторов путем элиминирования влияния других факторов на уровень изучаемого явления.

Таким образом, общие индексы являются синтетическими и аналитическими показателями.

Таблица 9.1 Цена и количество проданных товаров

То-вар	Еди-ница изме-рения	Цена, руб.		Количество проданных товаров		Стоимость проданной продукции, тыс. руб.		Индивидуальный индекс, %			Стои-мость продукции, проданной в мае, в ценах ап-реля, тыс. руб. p_0q_1	$i_q \cdot p_0q_0$	$\frac{p_1q_1}{i_p}$	Стои-мость про-дукции, про-данной в апреле, в ценах мая, тыс. руб. p_1q_0
		ап-рель, p_0	май, p_1	ап-рель, q_0	май, q_1	ап-рель, p_0q_0	май, p_1q_1	цен $i_p = \frac{p_1}{p_0}$	физичес-кого объ-ема про-дукции $i_q = \frac{q_1}{q_0}$	стои-мости $i = \frac{p_1q_1}{p_0q_0}$				
A	Б	1	2	3	4	5	6	7=2:1	8=4:3	9=6:5	10=1·4	11=8·5	12=6:7	13=2·3
A	Т	1638	1704	1000	5000	1638	8520	104,03	500	520	8190	8190	8190	1704
B	Шт	6925	7340	2000	2500	13850	18350	105,99	125	132	17312,5	17312,5	17312,5	14680
C	Кг	5040	5240	400	500	2013	2620	103,97	125	130	2520	2520	2520	2096
Всего		—	—	—	—	17504	29490	—	—	—	28022,5	28022,5	28022,5	18480

Общие индексы строят для количественных (объемных) и качественных показателей. В зависимости от цели исследования и наличия исходных данных используют различные формы построения общих индексов: агрегатную или средневзвешенную.

9.3. Агрегатный индекс как исходная форма индекса

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

Латинское слово «агрегат» (aggregatus) означает «складываемый, суммируемый». Особенность этой формы индекса состоит в том, что в агрегатной форме непосредственно сравниваются две суммы одноименных показателей. В настоящее время это наиболее распространенная форма индексов, используемая в практической статистике многих стран мира.

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

Индексируемой величиной называется признак, изменение которого изучается (цена товаров, курс акций, затраты рабочего времени на производство продукции, количество проданных товаров и т.д.). **Вес индекса** — это величина, служащая для целей соизмерения индексируемых величин.

За каждым экономическим индексом стоят определенные экономические категории. Экономическое содержание индекса предопределяет методику его расчета.

Методика построения агрегатного индекса предусматривает решение трех вопросов:

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

При выборе веса индекса принято руководствоваться следующим правилом: если строится индекс количественного показателя, то веса берутся за базисный период, при построении индекса качественного показателя используются веса отчетного периода.

Построим три индекса: стоимости продукции, физического объема продукции и цен.

Стоимость продукции — это произведение количества продукции в натуральном выражении (q) на ее цену (p).

Индекс стоимости продукции, или товарооборота (I_{pq}), представляет собой отношение стоимости продукции текущего периода ($\sum p_1 q_1$) к стоимости продукции в базисном периоде ($\sum p_0 q_0$) и определяется по формуле

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}. \quad (9.13)$$

Такой индекс показывает, во сколько раз возросла (уменьшилась) стоимость продукции (товарооборота) отчетного периода по сравнению с базисным, или сколько процентов составляет рост (снижение) стоимости продукции. Если из значения индекса стоимости (9.13) вычесть 100% ($I_{pq} - 100$), то разность покажет, на сколько процентов возросла (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным. Разность числителя и знаменателя ($\sum p_1 q_1 - \sum p_0 q_0$) показывает, на сколько рублей увеличилась (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным. Аналогично строятся индексы для показателей, которые являются произведением двух сомножителей: издержек производства (произведение себестоимости единицы продукции на количество продукции); затрат времени на производство всей продукции (произведение затрат времени на производство единицы продукции на количество выработанной продукции).

Рассчитаем индекс стоимости (товарооборота) продукции по данным табл. 9.1.

$$I_{pq} = \frac{29490}{17504} = 1,685, \text{ или } 168,5\%.$$

Следовательно, стоимость продукции (товарооборота) в мае по сравнению с апрелем возросла почти в 1,7 раза (рост составил 168,5%). Стоимость продукции увеличилась на $168,5\% - 100\% = 68,5\%$, или на 11980 тыс. руб. ($29940 - 17504$). Значение индекса стоимости продукции (товарооборота) зависит от двух факторов: изменения количества продукции и цен, что обуславливает возможность и необходимость построения еще двух индексов: объема продукции и цен.

Индекс физического объема продукции — это индекс количественного показателя. В этом индексе индексируемой величиной будет количество продукции в натуральном выражении, а весом — цена.

Только умножив несоизмеримые между собой количества разнородной продукции на их цены, можно перейти к стоимостям продукции, которые будут уже величинами соизмеримыми. Так как индекс физического объема — индекс количественного показателя, то весами будут цены базисного периода. Тогда формула индекса примет следующий вид:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}, \quad (9.14)$$

где в числителе дроби — условная стоимость произведенных в текущем периоде товаров в ценах базисного периода, а в знаменателе — фактическая стоимость товаров, произведенных в базисном периоде. Если объектом исследования является отдельное предприятие, то индекс определяется по совокупности произведенных товаров; когда объект исследования — отрасль промышленности, индекс рассчитывается по совокупности всех товаров, произведенных в отрасли, или по отдельным их группам в зависимости от цели анализа. Если же объектом исследования является какой-либо регион, то индекс рассчитывается по товарам, произведенным предприятиями региона.

Индекс физического объема продукции (9.14) показывает, во сколько раз возросла (уменьшилась) стоимость продукции из-за роста (снижения) объема ее производства или сколько процентов составляет рост (снижение) стоимости продукции в результате изменения физического объема ее производства. Если из значения индекса физического объема продукции (2.14) вычесть 100% ($I_q - 100$), то разность покажет, на сколько процентов возросла (уменьшилась) стоимость продукции в текущем периоде по сравнению с базисным из-за роста (снижения) объема ее производства. Разность числителя и знаменателя ($\sum p_1 q_0 - \sum p_0 q_0$) показывает, на сколько рублей изменилась стоимость продукции в результате роста (уменьшения) ее объема. Изменение цен на продукцию в текущем периоде по сравнению с базисным не влияет на величину индекса.

Для расчета индекса воспользуемся данными табл. 9.1.

$$I_q = \frac{28022,5}{17504} = 1,6009, \text{ или } 160,09\%.$$

Следовательно, стоимость продукции в мае по сравнению с апрелем увеличилась в 1,6 раза (или рост стоимости составил 160%). Вычитая из числителя индекса знаменатель, получим:

$$28022,5 - 17504 = 10518,5 \text{ тыс. руб.},$$

т.е. за счет увеличения объема продукции на 60% (160% – 100%) ее стоимость в абсолютном выражении увеличилась на 10518,5 тыс. руб. При построении агрегатного индекса цен, который в условиях рыночной экономики является наиболее широко распространенным показателем инфляции, исходят из тех же предпосылок, что и при построении индекса физического объема продукции.

Индекс цен — это индекс качественного показателя. Индексируемой величиной будет цена товара, так как этот индекс характеризует изменение цен. Весом будет выступать количество произведенных товаров. Умножив цену товара на его количество, получаем величину, которую можно суммировать и которая представляет собой показатель, соизмеримый с другими подобными ему величинами.

Индекс цен определяется по следующей формуле:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}, \quad (9.15)$$

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

Индекс показывает, во сколько раз возросла (уменьшилась) стоимость продукции из-за изменения цен, или сколько процентов составляет рост (снижение) стоимости продукции в результате изменения цен. Если из значения индекса (9.15) вычесть 100% ($I_p - 100$), то разность покажет, на сколько процентов возросла (уменьшилась) стоимость продукции из-за изменения цен, а разность числителя и знаменателя ($\sum p_1 q_1 - \sum p_0 q_1$) — на сколько рублей изменилась стоимость продукции в результате роста (снижения) цен. Изменение количества произведенной продукции в текущем периоде по сравнению с базисным не влияет на величину индекса.

Определим индекс по данным табл.9.1.

$$I_p = \frac{29490}{28022,5} = 1,0523 \text{ или } 105,23\%.$$

Таким образом, в среднем по трем товарам цены выросли в 1,0523 раза (или рост цен составил 105,23%). В результате увеличения цен на 5,23% (105,23% – 100%) покупатели заплатили на 1467,5 тыс. руб. больше в мае, чем в апреле (29490 – 28022,5 = 1467,5).

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

$$I_{pq} = I_p \cdot I_q \quad (9.16)$$

или

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot \frac{\sum p_0 q_1}{\sum p_0 q_0}. \quad (9.17)$$

Разность числителя и знаменателя каждого индекса-сомножителя выражает размер изменения общей абсолютной величины под влиянием изменения одного фактора. Алгебраическая сумма этих разностей равна разности числителя и знаменателя индекса стоимости продукции:

$$(\sum p_0 q_1 - \sum p_0 q_0) + (\sum p_1 q_1 - \sum p_0 q_1) = \sum p_1 q_1 - \sum p_0 q_0 \quad (9.18)$$

Равенства (9.16 – 9.18) выполняются в том случае, если при исчислении индекса объемного показателя веса были зафиксированы на уровне базисного периода, а при расчете индекса качественного показателя — на уровне отчетного периода.

Для нашего примера (табл. 9.1) индекс стоимости, исчисленный по формуле (9.16), равен:

$$1,6009 \cdot 1,0523 = 1,685, \text{ или } 168,5\%,$$

а алгебраическая сумма этих разностей (формула (9.18) равна:

$$10518,5 + 1467,5 = 11986 \text{ тыс. руб.}$$

9.4. Средние индексы

Помимо агрегатных индексов в статистике применяется другая их форма — средневзвешенные индексы. К их исчислению прибегают тогда, когда имеющаяся в распоряжении информация не позволяет рассчитать общий агрегатный индекс. Так, если отсутствуют данные о ценах, но имеется информация о стоимости продукции в текущем периоде и известны индивидуальные индексы цен по каждому товару, то нельзя определить общий индекс цен как агрегатный, но возможно исчислить его как средний из индивидуальных. Точно так же, если неизвестны количества произведенных отдельных видов продукции, но известны индивидуальные индексы и стоимость продукции базисного периода, можно определить общий индекс физического объема продукции как средневзвешенную величину.

Средний индекс — это индекс, вычисленный как средняя величина из индивидуальных индексов. Агрегатный индекс является основной формой общего индекса, поэтому средний индекс должен быть

тождественен агрегатному индексу. При исчислении средних индексов используются две формы средних: арифметическая и гармоническая.

Средний арифметический индекс тождественен агрегатному индексу, если весами индивидуальных индексов будут слагаемые знаменателя агрегатного индекса. Только в этом случае величина индекса, рассчитанного по формуле средней арифметической, будет равна агрегатному индексу.

Средний арифметический индекс физического объема продукции вычисляется по формуле:

$$I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}. \quad (9.19)$$

Так как $i_q \cdot q_0 = q_1$, то формула этого индекса легко преобразуется в формулу (9.14). Весами в формуле (9.19) является цена продукции базисного периода.

Средний арифметический индекс производительности труда определяется следующим образом:

$$I_t = \frac{\sum i_t t_1 q_1}{\sum t_1 q_1} = \frac{\sum i_t T_1}{\sum T_1}. \quad (9.20)$$

Так как $i_t \cdot t_1 = t_0$, то формула этого индекса может быть преобразована в агрегатный индекс трудоемкости продукции. Весами являются общие затраты времени на производство продукции в текущем периоде.

В статистике широко известен и другой средний арифметический индекс, который используется при анализе производительности труда. Он носит название **индекса Струмилина** и определяется следующим образом:

$$I_v = \frac{\sum \left(\frac{q_1}{T_1} : \frac{q_0}{T_0} \right) \cdot T_1}{\sum T_1} \quad (9.21)$$

Индекс показывает, во сколько раз возросла (уменьшилась) производительность труда, или сколько процентов составил рост (снижение) производительности труда в среднем по всем единицам исследуемой совокупности.

Средние арифметические индексы чаще всего применяются на практике для расчета сводных индексов количественных показателей,

а из качественных показателей — для исчисления двух приведенных выше индексов (9.20 – 9.21).

Индексы двух качественных показателей (цен, себестоимости и т.д.) определяются по формуле средней гармонической взвешенной величины.

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

$$I_z = \frac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}, \quad (9.22)$$

а индекс цен:

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}. \quad (9.23)$$

таким образом, весами при определении среднего гармонического индекса себестоимости являются издержки производства текущего периода, а индекса цен — стоимость продукции этого периода.

Рассчитаем средние индексы цен и физического объема продукции по данным табл. 9.1. (гр. 11 – 12):

$$I_q = \frac{28022,5}{17504} = 1,6009 \text{ или } 160,09\%;$$

$$I_p = \frac{29490}{28022,5} = 1,0523 \text{ или } 105,23\%.$$

Этот же результат получился при расчете агрегатных индексов. Средние индексы широко используются для анализа рынка ценных бумаг. Наиболее известными являются индексы Доу-Джонса, Стэрнарда и Пура.

Индекс Доу-Джонса (Dow Jones Industrial Average Index) определяется как средний арифметический индекс значений курсов акций, котирующихся на Нью-Йоркской фондовой бирже. Один сводный и три групповых индекса рассчитываются каждые полчаса, и ежедневно публикуется их значение на момент закрытие биржи. Групповые индексы определяются по ценам акций 30 промышленных, 20 транспортных и 15 компаний сферы услуг. Общий индекс рассчитывается по всем 65 компаниям. Их перечень был составлен в 1928 г. Первоначальная методика исчисления индекса была разработана основателем и

редактором крупнейшей в США газеты «Уолл-стрит джорнел» Чарлзом Доу.

Индекс Стэндарда и Пура (Standard and Poor's 500 Stock Index) — индекс, рассчитываемый по курсам акций 500 крупнейших компаний Нью-Йоркской фондовой биржи как средний взвешенный показатель, учитывающий общее число выпущенных компанией акций. В число компаний, акции которых включены в индекс, входят 400 промышленных корпораций, 40 — финансовых, 20 — транспортных и 40 — сферы услуг.

9.5. Выбор базы и весов индексов

Выбор базы сравнения и весов индексов — это два важнейших методологических вопроса построения систем индексов. Система используется при изучении динамики социально-экономических явлений за некоторый интервал времени, включающий более двух периодов времени.

Системой индексов называется ряд последовательно построенных индексов. Такие системы характеризуют изменения, происходящие в изучаемом явлении в течение исследуемого периода времени.

В зависимости от базы сравнения системы индексов бывают базисными и цепными.

Система базисных индексов — это ряд последовательно вычисленных индексов одного и того же явления с постоянной базой сравнения, т.е. в знаменателе всех индексов находится индексируемая величина базисного периода.

Система цепных индексов — это ряд индексов одного и того же явления, вычисленных с меняющейся от индекса к индексу базой сравнения.

В экономико-статистических исследованиях выбор системы индексов (базисные или цепные) проводится в зависимости от цели анализа. Базисные индексы дают более наглядную характеристику общей тенденции развития исследуемого явления, а цепные — четче отражают последовательность изменения уровней во времени.

Системы цепных и базисных индексов могут быть построены для индивидуальных и общих индексов. Системы индивидуальных индексов стоимости продукции, физического объема продукции и цен (табл. 9.2) просты по построению. Аналогично им строятся системы индивидуальных индексов и для других показателей.

Между цепными и базисными индексами существуют различные виды связи.

Таблица 9.2 Системы индивидуальных индексов

Название индивидуального индекса	Система индексов	
	базисных	цепных
Индекс стоимости	$\frac{p_1 q_1}{p_0 q_0}; \frac{p_2 q_2}{p_0 q_0}; \dots; \frac{p_n q_n}{p_0 q_0}$	$\frac{p_1 q_1}{p_0 q_0}; \frac{p_2 q_2}{p_1 q_1}; \dots; \frac{p_n q_n}{p_{n-1} q_{n-1}}$
Индекс физического объема	$\frac{q_1}{q_0}; \frac{q_2}{q_0}; \dots; \frac{q_n}{q_0}$	$\frac{q_1}{q_0}; \frac{q_2}{q_1}; \dots; \frac{q_n}{q_{n-1}}$
Индекс цен	$\frac{p_1}{p_0}; \frac{p_2}{p_0}; \dots; \frac{p_n}{p_0}$	$\frac{p_1}{p_0}; \frac{p_2}{p_1}; \dots; \frac{p_n}{p_{n-1}}$

Если известны цепные индексы, то путем их последовательного перемножения можно получить базисные индексы. Например,

$$\frac{p_1}{p_0} \cdot \frac{p_2}{p_1} = \frac{p_2}{p_0} \quad (9.24)$$

или

$$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} = \frac{q_3}{q_0} \quad (9.25)$$

Зная последовательные значения базисных индексов, легко рассчитать на их основе цепные индексы:

$$\frac{p_2}{p_1} = \frac{p_2}{p_0} \cdot \frac{p_1}{p_0} \quad (9.26)$$

или

$$\frac{q_3}{q_2} = \frac{q_3}{q_0} \cdot \frac{q_2}{q_0} \quad (9.27)$$

Системы базисных и цепных индексов могут быть построены **для агрегатных индексов**.

Система индексов стоимости имеет следующий вид:

- цепные индексы:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0}; \frac{\sum p_2 q_2}{\sum p_1 q_1}; \dots; \frac{\sum p_n q_n}{\sum p_{n-1} q_{n-1}};$$

- базисные индексы:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0}, \frac{\sum p_2 q_2}{\sum p_0 q_0}, \dots, \frac{\sum p_n q_n}{\sum p_0 q_0}.$$

Формирование системы индексов, например, цен или физического объема отличается от уже рассмотренных в этом параграфе систем индексов. Это связано с тем, что при построении систем этих индексов можно использовать *постоянные и переменные веса*.

Системой весов с постоянными весами называется система сводных индексов одного и того же явления, вычисленных с весами, *не меняющимися* при переходе от одного индекса к другому. Постоянные веса позволяют исключить влияние изменения структуры на величину индекса.

Например, *система базисных индексов физического объема* продукции с постоянными весами (p_0) имеет следующий вид:

$$\frac{\sum p_0 q_1}{\sum p_0 q_0}, \frac{\sum p_0 q_2}{\sum p_0 q_0}, \dots, \frac{\sum p_0 q_n}{\sum p_0 q_0},$$

а систему цепных индексов с теми же постоянными весами можно представить так:

$$\frac{\sum p_0 q_1}{\sum p_0 q_0}, \frac{\sum p_0 q_2}{\sum p_0 q_1}, \dots, \frac{\sum p_0 q_n}{\sum p_0 q_{n-1}}.$$

Система индексов с переменными весами представляет собой систему сводных индексов одного и того же явления, вычисленных с весами, *последовательно меняющимися* от одного индекса к другому. Переменные веса — это веса *отчетного* периода.

Например, *система базисных индексов цен с переменными весами* следующая:

$$\frac{\sum p_1 q_1}{\sum p_0 q_1}, \frac{\sum p_2 q_2}{\sum p_0 q_2}, \dots, \frac{\sum p_n q_n}{\sum p_0 q_n}.$$

Элементами этой системы являются **индексы-дефляторы**, которые необходимы для пересчета стоимостных показателей системы национальных счетов в сопоставимые цены.

Система цепных индексов цен с переменными весами выглядит так:

$$\frac{\sum p_1 q_1}{\sum p_0 q_1}, \frac{\sum p_2 q_2}{\sum p_1 q_2}, \dots, \frac{\sum p_n q_n}{\sum p_{n-1} q_n}.$$

Отдельные индексы этой системы используются для пересчета стоимостных показателей отчетного периода в цены предыдущего периода.

Системы общих индексов других показателей строятся аналогично.

Системы агрегатных индексов обладают теми же свойствами, что и системы индивидуальных индексов, т.е., зная базисные индексы, можно рассчитать цепные; при наличии цепных индексов легко получить соответствующие им базисные. Например,

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} \cdot \frac{\sum p_2 q_2}{\sum p_1 q_1} = \frac{\sum p_2 q_2}{\sum p_0 q_0} \quad (9.28)$$

9.6. Индексы структурных сдвигов

При изучении динамики качественных показателей 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. Так, средняя заработная плата на предприятии может вырасти в результате оплаты труда работников или увеличением доли высокооплачиваемых сотрудников. Снижение трудоемкости производства единицы продукции по совокупности предприятий отрасли может быть обусловлено повышением производительности труда на предприятиях или концентрацией производства продукции на заводах с низкой трудоемкостью. Так как на изменение среднего значения показателя оказывают воздействие два фактора, возникает задача определить степень влияния каждого из факторов на общую динамику средней.

Эта задача решается с помощью индексного метода, т.е. путем построения системы взаимосвязанных индексов, в которую включают три индекса: переменного состава, постоянного состава и структурных сдвигов.

Индексом переменного состава называют индекс, выражающий соотношение средних уровней изучаемого явления, относящихся к разным периодам времени. Например, индекс переменного состава

себестоимости продукции одного и того же вида рассчитывается по формуле

$$I_{\text{пс}} = \frac{\bar{z}_1}{z_0} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0}, \quad (9.29)$$

где $I_{\text{пс}}$ — индекс переменного состава;

z_0 — себестоимость продукции в базисный период;

z_1 — себестоимость продукции в отчетный (текущий) период;

q_0, q_1 — физический объем продукции;

Напомним, что $\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$ — среднее арифметическое взвешенное.

Индекс переменного состава отражает изменение не только индексируемой величины (в данном случае себестоимости), но и структуры совокупности (весов).

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. Так, индекс фиксированного состава себестоимости продукции рассчитывают по формуле

$$I_{\text{фс}} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1}, \quad (9.30)$$

где $I_{\text{фс}}$ — индекс фиксированного состава.

Под **индексом структурных сдвигов** понимают индекс, характеризующий влияние изменения структуры изучаемого явления на динамику среднего уровня этого явления. Индекс определяется по формуле (при изучении изменения среднего уровня себестоимости)

$$I_{\text{сс}} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{\sum z_0 q_1}{\sum z_0 q_0} : \frac{\sum q_1}{\sum q_0}, \quad (9.31)$$

где $I_{\text{сс}}$ — индекс структурных сдвигов.

Система взаимосвязанных индексов при анализе динамики средней себестоимости имеет следующий вид:

$$I_{\text{пс}} = I_{\text{фс}} \times I_{\text{сс}}, \quad (9.32)$$

То есть индекс переменного состава равен произведению индексов фиксированного состава и структурных сдвигов.

Рассмотрим применение такой системы на конкретном примере. Пусть имеются данные о себестоимости единицы продукции на трех предприятиях в текущем и базисном периодах (табл. 9.3).

Таблица 9.3 Количество произведенной продукции и себестоимость единицы продукции одного вида по трем предприятиям отрасли

Но- мер пред- прия- тия	Произведено продукции				Себестои- мость едини- цы продук- ции, тыс. руб.		Индивиду- альные индексы себестои- мости, % $i_z = \frac{z_1}{z_0}$	Издержки произ- водства, млн. руб.		
	Всего единиц		% к итогу							
	Базис- ный пери- од, q_0	Теку- щий пери- од, q_1	Базис- ный пери- од, z_0	Теку- щий пери- од, z_1	Базис- ный пери- од, $z_0 q_0$	Теку- щий пери- од, $z_1 q_1$		$z_0 q_1$		
A	1	2	3	4	5	6	7=6:5	8=5:1	9=6:2	10=5:2
1	1680	1500	70	50	20	20,3	101,5	33,6	30,45	30
2	480	600	20	20	18	18,4	102,2	8,64	11,04	10,8
3	240	900	10	30	15	15,5	103,3	3,6	13,95	13,5
Всего	2400	3000	100	100	19,10	18,48	—	45,84	55,44	54,3

В текущем периоде по сравнению с базисным себестоимость производства продукции возросла на каждом предприятии (гр. 5 – 6); изменилась структура производства: уменьшилась доля первого предприятия в общем выпуске продукции, возросла доля третьего, а доля второго не изменилась (гр. 3 – 4).

Рассчитаем индекс переменного состава. Для этого сначала определим среднюю себестоимость единицы продукции в текущем и базисном периодах:

$$\bar{z}_0 = \frac{45840}{2400} = 19,10 \text{ тыс. руб.};$$

$$\bar{z} = \frac{55440}{3000} = 18,48 \text{ тыс. руб.}$$

Тогда

$$I_{\text{пс}} = 18,48 : 19,10 = 0,9675 \text{ или } 96,75\%.$$

Следовательно, средняя себестоимость по трем предприятиям снизилась в текущем периоде по сравнению с базисным на 3,25%, хотя на каждом из них в отдельности она возросла. Это — результат того,

что исчисленный индекс учитывает влияние еще и структурного фактора.

Определим индекс себестоимости фиксированного состава.

$$I_{\text{фс}} = \frac{55,44}{54,3} = 1,021 \text{ или } 102,1\%.$$

Таким образом, себестоимость в текущем периоде по сравнению с базисным возросла в среднем на 2,1%.

Вычислим влияние изменения структуры на динамику средней себестоимости:

$$I_{\text{сс}} = \frac{54,3}{45,84} : \frac{3000}{2400} = 0,9476 \text{ или } 94,76\%.$$

Изменение доли предприятий в общем объеме произведенной продукции привело к снижению себестоимости на 5,24%.

9.7. Индексы пространственно-территориального сопоставления

В статистической практике часто возникает потребность в сопоставлении уровней экономического явления в пространстве: по странам, экономическим районам, областям, т.е. в исчислении территориальных индексов. При построении территориальных индексов приходится решать вопрос, какие веса использовались при их исчислении. Например, если стоит задача сравнения цены двух регионов (А и Б), то можно построить два индекса:

$$I_{\text{А/Б}} = \frac{\sum p_{\text{А}} q_{\text{А}}}{\sum p_{\text{Б}} q_{\text{А}}} \quad (9.33)$$

и

$$I_{\text{Б/А}} = \frac{\sum p_{\text{Б}} q_{\text{Б}}}{\sum p_{\text{А}} q_{\text{Б}}}, \quad (9.34)$$

где $I_{\text{А/Б}}$ — индекс, в котором в качестве базы сравнения применяются данные по региону А;

$I_{\text{Б/А}}$ — индекс, используемый в качестве базы сравнения данных по региону Б.

Эти формулы могут дать совершенно различное представление о соотношении уровней явления. Например, при расчете по формуле (9.33) значение признака будет ниже в регионе А, а по формуле (9.34) — в регионе Б.

Рассмотрим пример (табл. 9.4).

Рассчитав индексы по формулам получаем:

$$I_{A/B} = \frac{\sum p_A q_A}{\sum p_B q_A} = \frac{28900}{23100} = 1,2510 \text{ или } 125,1\%$$

и

$$I_{B/A} = \frac{\sum p_B q_B}{\sum p_A q_B} = \frac{14230}{14200} = 1,0035 \text{ или } 100,35\%.$$

Индексы показывают, что при сравнении региона А с регионом Б цены выше в регионе А на 25,1%, а при сравнении региона Б с регионом А оказывается, что цены несколько выше в регионе Б. Таким образом, расчет индексов не позволяет определить в каком регионе выше цены. Причина заключается в резком различии структуры продаж в отдельных регионах.

В теории и практике статистики предлагаются различные методы построения территориальных индексов, в том числе *метод стандартных весов*. Этот метод заключается в том, что значения индексированной величины взвешиваются не по весам какого-либо одного региона, а по весам суммарным (или средним)

$$I_p = \frac{\sum p_A (q_A + q_B)}{\sum p_B (q_A + q_B)} = \frac{43100}{33350} = 1,154 \text{ или } 115,4\%.$$

Цены в регионе А выше, чем в регионе Б на 15,4%.

9.8. Индексы – дефляторы

Пересчет важнейших стоимостных показателей системы национальных счетов (национальный доход, валовой национальный продукт и т.д.) из фактических цен в сопоставимые осуществляется с помощью индекса дефлятора. **Дефлятор** — это коэффициент, переводящий значение стоимостного показателя за отчетный период в стоимостные измерители базисного. Например, индекс-дефлятор валового внутреннего продукта (ВВП) представляет собой индекс цен, применяемый для корректировки номинального объема ВВП с учетом инфляции и получения на этой основе реального его объема.

Согласно экономической теории, с помощью номинального ВВП измеряется объем продукции текущего года (q_1) в текущих ценах (p_1). С помощью реального ВВП измеряется объем продукции текущего года (q_1) в ценах, которые сложились в базисном году (p_0).

Таблица 9.4 Цены на продукты питания и количество проданной продукции по двум регионам

Наименование продуктов	Единица изме- рения	Регион А		Регион Б		Товaroоборот				Коли- чество продук- ции, про- данной в регионах А и Б $q_A + q_B$		
		цена, руб. p_A	коли- чество продан- ной про- дукции q_A	цена, руб. p_B	коли- чество продан- ной про- дукции q_B	регион А $p_A q_A$	регион Б $p_B q_B$					
А	Б	1	2	3	4	5=1·2	6=3·4	7=3·2	8=1·4	9=2+4	10=1·9	11=3·9
Масло	Кг	22	1200	16,5	300	26400	4950	19800	6600	1500	33000	24750
Хлеб	Кг	2	500	3	300	1000	900	1500	600	800	1600	2400
Яйца	десяток	5	300	6	1400	1500	8400	1800	7000	1700	8500	10200
Всего	—	—	—	—	—	28900	14250	23100	14200	—	43100	37350

Система базисных индексов с переменными весами:

$$\frac{\sum p_1 q_1}{\sum p_0 q_1}, \frac{\sum p_2 q_2}{\sum p_0 q_2}, \dots, \frac{\sum p_n q_n}{\sum p_0 q_n}.$$

Индекс-дефлятор рассчитывается как отношение фактической стоимости продукции отчетного периода к стоимости объема продукции, структура которого аналогична структуре отчетного года, но определенного в ценах базисного года. В основе расчета индекса-дефлятора лежит формула Паше — агрегатная формула индекса с текущими весами. Индекс-дефлятор для ВВП в 2002 г. определяется по формуле:

$$I_d = \frac{\sum p_{2002} \cdot q_{2002}}{\sum p_0 \cdot q_{2002}}, \quad (9.35)$$

где I_d — индекс-дефлятор;

q_{2002} — объем продукции в 2002 г.;

p_{2002}, p_0 — цены фактически действующие в 2002 г. и базисном году соответственно.

Реальный ВВП за 2002 г. определяется по формуле:

$$R_{2002} = Q_{2002} : I_d, \quad (9.36)$$

где Q_{2002} — номинальный ВВП.

Формулу (9.35) можно представить и в следующем виде:

$$I_d = \frac{Q_{2002}}{R_{2002}}. \quad (9.37)$$

Индекс-дефлятор для 2003 г. может быть исчислен так:

$$I_d = \frac{\sum p_{2003} \cdot q_{2003}}{\sum p_0 \cdot q_{2003}}, \quad (9.38)$$

где q_{2003} — объем продукции в 2003 г.;

p_{2003} — цены, фактически действующие в 2003 г.

Сравним формулы (9.35) и (9.38). Легко заметить, что в них используются различные веса (q_{2002} и q_{2003}). Поэтому важной особенностью индекса-дефлятора является то, что он не может быть использован для сравнительной оценки динамики цен за два периода, в данном случае за 2002 и 2003 гг. Индексы-дефляторы дают представление

только об отношении стоимости продукции в текущем периоде к ее стоимости в базисном периоде.

При этом не учитывается отличие состава и структуры продукции в базисный период по сравнению с отчетным.

Таким образом, индекс-дефлятор — это самостоятельный показатель.

В статистической практике индексы-дефляторы определяются не только в целом по народному хозяйству; они исчисляются по отдельным регионам, различным товарным группам, каналам реализации потребительских благ, отраслям экономики и т.д.

Тест для самоконтроля

1. Производство индивидуальных индексов цены и физического объема — это

- а) сумма выручки от реализации продукции
- б) объем товарооборота
- в) индивидуальный индекс товарооборота
- г) общий индекс товарооборота

2. Индивидуальный индекс цен определяется по формуле

а) $\frac{\sum p_1 q_1}{\sum p_0 q_1}$ б) $\frac{p_1}{p_0}$ в) $\frac{p_0}{p_1}$ г) $\frac{\sum p_1}{\sum p_0}$

3. Имеются данные о ценах и количестве проданного товара

Вид товара	Единица измерения	Цена, руб.		Реализовано, тыс.ед.	
		2007 г.	2008 г.	2007 г.	2008 г.
Хлеб	булка	20	24	600	500
Сахар	кг	50	50	100	150

Общий индекс стоимости продукции равен

- а) 114,7 %
- б) 101,1 %
- в) 102,9 %

4. За год объем товарооборота увеличился на 4000 рублей. Какова была цена товара в базисном году, если известно, что индекс товарооборота равен 300%, а продано товара в базисном году была 100 единиц.

- а) 10 руб
- б) 20 руб
- в) 30 руб
- г) 100 руб

5. Имеются данные о себестоимости и произведенной продукции на двух предприятиях

Вид продукции	Предприятие А		Предприятие Б	
	цена, руб.	произведено, шт.	цена, руб.	произведено, шт.
1	370	1000	380	600
2	120	900	120	200
3	400	300	450	1500

индекс цен продукции предприятия А по сравнению с предприятием Б- равен

- а) 1200 рублей
- б) 48,4%
- в) 0,932

6. Цена товара увеличилась в 1,3 раза, товарооборот снизился на 1%. Индекс физического объема равен

- а) 131,3%
- б) 100%
- в) 84,6%
- г) 76,2%

Вопросы и упражнения

1. Дайте определение экономического индекса.
2. Каковы цели индексного метода изучения социально-экономических явлений?
3. В чем отличие базисных и цепных индексов?
4. В чем отличие индивидуального и агрегатного индексов стоимости продукции?
5. Какая величина применяется в качестве веса при вычислении территориального индекса цен?
6. Каково соотношение между общими индексом товарооборота, индексом цен Пааше и индексом физического объема Ласпейреса?
7. Как определяется индекс структурных сдвигов, если известны индексы переменного и постоянного состава.
8. Дайте определение показателя «индекс-дефлятор»?

ЛИТЕРАТУРА

Основная литература

1. Харченко Н.М. : Статистика : Учебник / Н. М. Харченко. - М. : Дашков и К°, 2007. - 366 с. : ил., табл. - Библиогр.: с. 361-363. - ISBN 5-91131-126-7. (Допущен УМО в качестве учебника для студентов вузов). Имеется в библиотеке ТУСУРа: Экземпляры всего: 40.
2. Гмурман, В.Е. Теория вероятностей и математическая статистика [Текст]: учебное пособие / В. Е. Гмурман. - 12-е изд., перераб. - М. : Юрайт, 2010. - 480 с. : табл. - (Основы наук). - - ISBN 978-5-9916-0616-5. - ISBN 978-5-9692-0874-2 (Рекомендовано УМО в качестве учебного пособия для студентов вузов) Имеется в библиотеке ТУСУРа: Экземпляры всего: 1.
3. Гмурман, В.Е. Теория вероятностей и математическая статистика : учебное пособие / В. Е. Гмурман. - 12-е изд., перераб. - М. : Высшее образование, 2006. - 478[2] с. - ISBN 978-5-9692-0192-7. (Рекомендовано УМО в качестве учебного пособия для студентов вузов) Имеется в библиотеке ТУСУРа: Экземпляры всего: 21.

Дополнительная литература

1. Лепихина З.П. Статистика: Учебное пособие/ З. П. Лепихина; Федеральное агентство по образованию, Томский государственный университет систем управления и радиоэлектроники. - Томск: ТУСУР, 2005. - 284 с.: ил.. - Библиогр.: с. 277-278. - ISBN 5-86889-273-9. (Рекомендовано СибРУМС в качестве учебно-методического пособия для студентов вузов) Имеется в библиотеке ТУСУРа: Экземпляры всего: 20.
2. Статистика: Учебное пособие/ ред. : М. Р. Ефимова. - М.: Инфра-М, 2006. - 335[1] с.: табл., ил.. - (Вопрос - ответ: Основана в 2000 г.). - Библиогр.: с. 330. - ISBN 5-16-000308-8. Имеется в библиотеке ТУСУРа: Экземпляры всего: 90.
3. Статистика: Учебное пособие для вузов/ Виктор Максимович Гусаров. - М.: ЮНИТИ-ДАНА, 2001. - 464 с.: ил. - Библиогр.: с. 456-457. - ISBN 5-238-00206-8). Имеется в библиотеке ТУСУРа: Экземпляры всего: 18.
4. Есипова Э.Ю. Система национальных счетов: краткий курс. – М.: Финансы и статистика, 2009г.- 176с. ISBN 978-5-279-03225-9 [Электронный ресурс]. — URL: http://e-lanbook.com/books/element.php?pl1_cid=25&pl1_id=1042

5. Лялин В.С. Статистика: теория и практика в Excel. – М.: Финансы и статистика, 2010 г.- 448с. ISBN 978-5-279-03381-2 [Электронный ресурс]. — URL: http://elanbook.com/books/element.php?pl1_cid=25&pl1_id=1048.

6. Демография и статистика населения : Учебник для вузов / И. И. Елисеева [и др.] ; ред. : И. И. Елисеева. - М. : Финансы и статистика, 2006. - 687[1] с. : ил., табл. - Библиогр.: с. 685-687. - ISBN 5-297-02743-X. Имеется в библиотеке ТУСУРа: Экземпляры всего: 20.

Методические указания

1. Методические указания к курсовой работе по дисциплине «Статистика» для студентов специальности 080504.65 «Государственное и муниципальное управление» [Электронный ресурс]. — Томск: ТУСУР, каф.АОИ, 2012. — 12с. — URL: http://aoi.tusur.ru/upload/methodical_materials/STAT_Kurs_MU_2012_file__264_1796.pdf

2. Методические указания к лабораторным работам дисциплине «Статистика» для студентов специальности 080504.65 Государственное и муниципальное управление» [Электронный ресурс]. — Томск: ТУСУР, каф.АОИ, 2012. - 34с. — URL: http://aoi.tusur.ru/upload/methodical_materials/STAT_Lab_MU_2012_file__265_8996.pdf

3. Методические указания к практическим занятиям по дисциплине «Статистика» для студентов специальности 080504.65 «Государственное и муниципальное управление» [Электронный ресурс]. — Томск: ТУСУР, каф.АОИ, 2012. – 46с. — URL: http://aoi.tusur.ru/upload/methodical_materials/STAT_Prakt_MU_2012_file__266_8290.pdf

4. Методические указания для выполнения самостоятельной работы по дисциплине «Статистика» для студентов специальности 080504.65 «Государственное и муниципальное управление» [Электронный ресурс]. — Томск: ТУСУР, каф.АОИ, 2012. – 26с. — URL: http://aoi.tusur.ru/upload/methodical_materials/STAT_Sam_MU_2012_file__267_1834.pdf