Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Кузбасский государственный технический университет имени Т. Ф. Горбачева»

Кафедра управленческого учета и анализа

Тамара Алексеевна Понкратова
Юлия Валерьевна Козлова

СТАТИСТИКА
ЧАСТЬ 1
ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Электронное учебное пособие

Кемерово 2017

Вперед

УДК 657(075.8)(086.76)

Рецензент(ы) Кучерова Е. В. – заведующий кафедрой управленческого учета и анализа, кандидат экономических наук, доцент ФГБОУ ВО «Кузбасский государственный технический университет имени Т. Ф. Горбачева»

Лубкова Э. М. – заведующий кафедрой финансов и кредита, кандидат экономических наук, доцент, председатель учебно-методической комиссии направления подготовки 38.05.01 «Экономическая безопасность» ФГБОУ ВО «Кузбасский государственный технический университет имени Т. Ф. Горбачева»

Понкратова Т. А. **Статистика. Часть 1. Общая теория статистики** [Электронный ресурс]: электронное учебное пособие / Т. А. Понкратова, Ю. В. Козлова; КузГТУ. – Кемерово, 2017. – 1 электрон. опт. диск (1,63 Мб)

Данное учебное пособие разработано с учетом последних изменений нормативной правовой базы, раскрывает теоретические и практические основы статистики. Предназначено для изучения дисциплины «Статистика» студентами направления подготовки 38.05.01 «Экономическая безопасность».

Может быть полезно специалистам, преподавателям, работающим в области экономических наук.

Текстовое (символьное) электронное издание

Минимальные системные требования: Частота процессора не менее 1,0 ГГц; ОЗУ 512 Мб; 20 Гб HDD; операционная система Windows XP; CD-ROM 4-скоростной; ПО для чтения файлов PDF-формата; SVGA-совместимая видеокарта; мышь.

Вперед

Сведения о программном обеспечении, которое использовано для создания электронного издания	MS Word
Сведения о технической подготовке материалов для электронного издания	Редактор З. М. Савина
Объем издания в единицах измерения объема носителя, занятого цифровой информацией (байт, Кб, Мб)	1,63 Мб
Наименование и контактные данные юридического лица, осуществившего запись на материальный носитель	Федеральное государственное бюджетное образовательное учреждение высшего образования «Кузбасский государственный технический университет имени Т. Ф. Горбачева» 650000, Кемерово, ул. Весенняя, 28 Тел./факс: 8(3842) 68-25-84

Вперед

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ 6
1 ПРЕДМЕТ И МЕТОД КУРСА «СТАТИСТИКА» 7
1.1 Предмет, содержание курса 7
1.2 Понятия и категории статистики 12
1.3 Стадии статистического исследования 13
2 СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ 14
2.1 Понятие, объект и единица статистического наблюдения 14
2.2 Формы и виды статистического наблюдения 16
2.3 Способы статистического наблюдения 18
2.4 Точность статистического наблюдения 21
3 СВОДКА И ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ 22
3.1 Понятие сводки и группировки 22
3.2 Виды группировок 24
3.3 Статистические распределения 26
3.3.1 Вариация признака в совокупности 26
3.3.2 Графическое изображение вариационного ряда 28
3.3.3 Показатели центра распределения 29
3.3.4 Показатели формы распределения 31
3.4 Статистические таблицы 33
3.5 Графическое изображение статистических показателей 36
3.5.1 Составные элементы графиков, их классификация 36
3.5.2 Виды диаграмм 39
3.5.3 Статистические карты, их виды 40
4 АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ 42
4.1 Формы выражения статистических величин 42
4.2 Абсолютные статистические величины 43
5 СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ 48
5.1 Средняя величина, ее сущность 48
5.3 Структурные средние величины 58
5.4 Вариация и ее показатели 61
6 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНИМИКИ СОЦИАЛЬНО-ЭКОМИЧЕСКИХ ЯВЛЕНИЙ 67
6.1 Определение и виды рядов динамики 67
6.2 Правила построения динамических рядов 69
6.3 Статистические характеристики (показатели)ряда динамики 69
6.4 Средние показатели ряда динамики 73
6.5 Механические методы выявления основной тенденции развития 76
6.6 Аналитическое выравнивание ряда 77

6.7 Интерполяция и экстраполяция 80
7 ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ 83
7.1 Понятие индексов и их классификация 83
7.2 Виды сложных индексов...84
7.3 Индексы с различной базой сравнения, постоянными и переменными весами 87
7.4 Индексы переменного состава, фиксированного состава и структурных сдвигов 89
7.5 Важнейшие экономические индексы и их взаимосвязь..........93
8 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ...95
8.1 Понятие выборочного наблюдения 95
8.2 Классификация методов отбора...96
8.3 Характеристики генеральной и выборочной совокупностей при выборочном наблюдении 96
8.4 Виды и способы отбора 101
9 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗЕЙ 104
9.1 Основные понятия корреляционного и регрессионного анализа 104
9.2 Парная корреляция и парная линейная регрессия 106
9.3 Множественная линейная регрессия 110
9.4 Нелинейная регрессия. Коэффициенты эластичности 112
9.5 Множественная корреляция 114
9.6 Оценка значимости параметров взаимосвязи 115
9.7 Непараметрические методы оценки связи.............................116
ВАРИАНТЫ ТЕСТОВЫХ ЗАДАНИЙ ПО ОСНОВНЫМ ТЕМАМ КУРСА 117
Темы: Предмет и метод курса «Статистика». Статистическое наблюдение. Сводка и группировка статистических данных. Статистические таблицы. 117
Тема: Графическое изображение статистических показателей..119
Тема: Абсолютные и относительные величины 122
Тема: Средние величины и показатели вариации........................125
Тема: Статистические индексы 127
Тема: Ряды динамики 130
КОНТРОЛЬНЫЕ ВОПРОСЫ 132
СПИСОК ЛИТЕРАТУРЫ 133

ПРЕДИСЛОВИЕ

Целью дисциплины «Статистика» (часть I. Общая теория статистики) является овладение студентами методикой и методами количественного исследования массовых процессов социально-экономических явлений с помощью статистических показателей, выработка практических навыков по сбору и обработке информации на всех этапах статистического исследования. Для достижения этих целей при изучении дисциплины ставятся следующие задачи:

– изучение общих свойств массовых явлений и методов их анализа;

– раскрытие содержания и конкретных методов построения системы экономико-статистических показателей.

Данное учебное пособие в полной мере и в доступной форме отражает вопросы, относящиеся к общей теории статистики. В приложениях с целью закрепления изученного материала даны тестовые задания и контрольные вопросы по основным темам курса.

Учебное пособие составлено в соответствии с требованиями Государственного образовательного стандарта по дисциплине «Статистика» для экономических специальностей и предназначено для студентов всех форм обучения.

Вернуться в начало.

1 ПРЕДМЕТ И МЕТОД КУРСА «СТАТИСТИКА»

1.1 Предмет, содержание курса

Статистика – это отрасль знаний и одновременно отрасль практической деятельности.

Например, по данным территориального органа Федеральной службы государственной статистики, по Кемеровской области наблюдается следующая динамика численности (рис. 1).

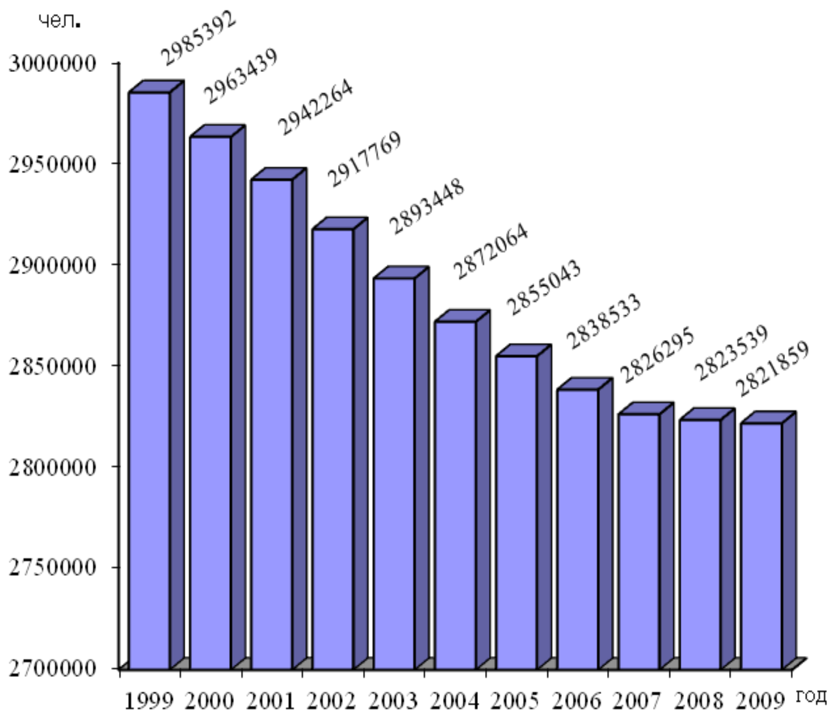

Рис. 1. Численность постоянного населения
в Кемеровской области на 1 января, чел.

Как видно на диаграмме, численность населения нашей области постоянно уменьшается. За последние 10 лет численность населения снизилась на 163 533 человека.

В табл. 1 представлены основные демографические показатели населения.

Таблица 1

Основные демографические показатели

Показатели	2004 г.	2005 г.	2006 г.	2007 г.	2008 г.
Численность постоянного населения, тыс. чел., в том числе:	2855,0	2838,5	2826,3	2823,5	2821,8
городское	2425,8	2411,7	2401,2	2398,6	2396,9
сельское	429,2	426,8	425,1	424,9	424,9
Из общей численности населения – население в возрасте, тыс. чел.: моложе трудоспособного	481,9	469,0	460,5	458,5	–
трудоспособном	1815,7	1813,8	1806,2	1797,9	–
старше трудоспособного	557,4	555,7	559,6	567,1	–
Ожидаемая продолжительность жизни при рождении, число лет: все население	62,2	61,6	63,0	64,0	–
мужчины	55,8	55,1	56,5	57,5	–
женщины	69,6	69,0	70,4	71,2	–
На 1000 человек населения: родившихся	10,7	10,8	11,3	12,1	13,0
умерших (всего),	17,9	18,7	17,3	16,6	16,3
в том числе детей в возрасте до 1 года	11,8	11,7	10,3	9,4	8,7
Естественный прирост, убыль (–) населения	–7,2	–7,9	–6,0	–4,5	–3,3
Число браков	7,4	8,2	8,5	8,9	8,9
Число разводов	4,5	4,9	5,3	5,5	5,5
Миграционный прирост, убыль (–) населения	1,3	2,0	1,6	3,5	2,8

Численность городского населения на 01.01.2008 составила 2821,8 тыс. человек, что на 33,2 тыс. человек меньше, чем в 2004 г., сельское население за тот же период снизилось на 4,3 тыс. человек и составило 424,9 тыс. человек.

Как найдены данные цифры, как определен рост и прирост, средняя продолжительность – что это? Какие методы использованы при определении данных показателей, вы узнаете из курса, который носит общее название «Статистика».

Слово «статистика» происходит от латинского «статус» – состояние, положение вещей. Оно вошло в употребление в XVII в., и вначале под статистикой понималось «государствоведение». Родоначальником статистики можно считать английского ученого Вильяма Петти. В XIX в. его последователи предприняли попытку представить статистику как науку о закономерностях развития общественных явлений.

Таким образом, статистика – наука, изучающая количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной, дающая количественное выражение закономерностей общественного развития.

Предметом статистики являются массовые общественные явления и складывающиеся в них количественные закономерности.

Статистика относится к общественным наукам. Ее особенности как общественной науки следующие:

1. Любому явлению природы и общества присущи две стороны: качественная и количественная. Статистика изучает количественную сторону общественных явлений, но в неразрывной связи с качественной стороной.

2. Как правило, статистика имеет дело не с единицами, а с совокупностями единиц, так как только в совокупности обнаруживаются закономерности.

3. Она изучает явления в условиях конкретного места и времени (где, когда).

4. Изучаемые общественные явления находятся в постоянном изменении, развитии.

5. Количественная сторона явлений и их взаимосвязи выражаются в цифрах через статистические показатели.

Статистика представляет собой систему научных дисциплин. Основными ее разделами являются:

– общая теория статистики, где рассматриваются наиболее общие категории, принципы и методы статистической науки;

– экономическая и социальная статистика, изучающие явления и процессы, происходящие в экономике, и социальные явления и процессы.

Структура статистической науки представлена на рис. 2.

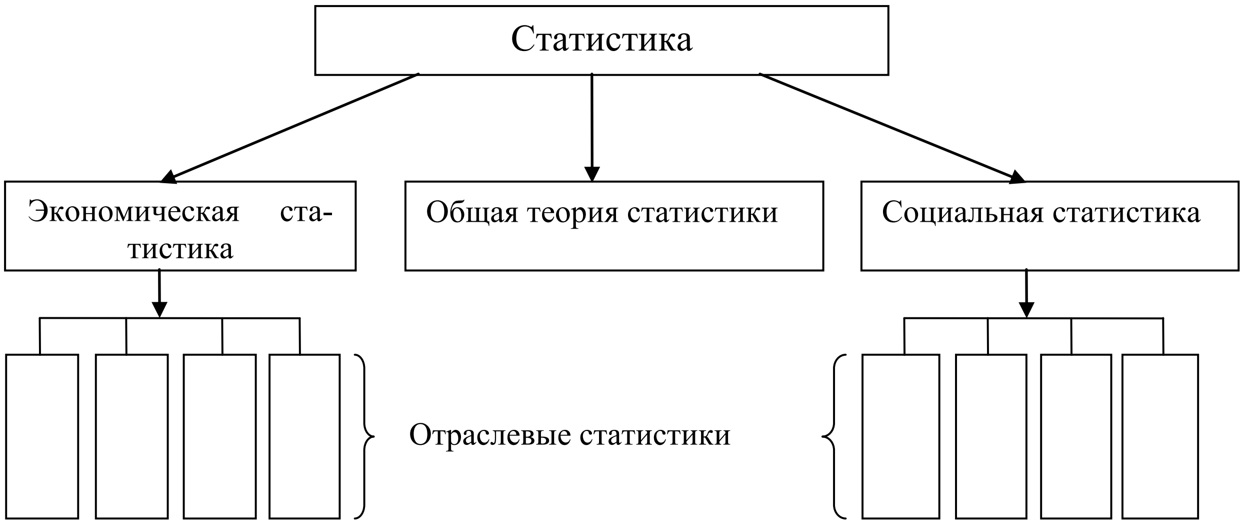

Рис. 2. Структура статистической науки

Задача общей теории статистики – разработка общих принципов и методов статистического исследования.

Задача экономической статистики – разработка и анализ показателей, отражающих состояние экономики в целом, взаимосвязь отраслей, наличие и особенности размещения материальных, трудовых и финансовых ресурсов, достигнутый уровень их использования.

Задача отраслевой экономической статистики (статистики промышленности, сельского хозяйства, строительства, транспорта, связи, труда, природных ресурсов и др.) – разработка и анализ статистических показателей развития соответствующей отрасли экономики.

Задача социальной статистики – формирование системы показателей для характеристики образа жизни населения и различных аспектов социальных отношений, проблем народонаселения, политики, культуры, здравоохранения, права и т.д.

Задача отраслевой социальной статистики (статистики уровня жизни и потребления, материальных благ и услуг, жилищно-коммунального хозяйства и бытового обслуживания населения, образования, культуры и искусства, здравоохранения и др.) – разработка и анализ статистических показателей развития соответствующей сферы социальной жизни общества.

Отраслевые статистики формируются на базе показателей экономической и социальной статистики, которые основываются на категориях и методах анализа, разработанных общей теорией статистики.

Статистическая методология представляет собой совокупность общих правил и специальных приемов и методов статистического исследования.

Общие правила статистического исследования исходят из положений социально-экономической теории и принципа диалектического метода познания.

Статистика применяет следующие специфические методы цифрового освещения явления.

1. Массовое научно организованное наблюдение, с помощью которого получают первичную информацию об отдельных единицах изучаемого явления. Получение сведений о достаточно большом числе единиц дает возможность освободиться от влияния случайных причин и установить характерные черты изучаемого объекта.

2. Группировка и сводка материала, представляющие собой расчленение всей массы случаев (единиц) на однородные группы и подгруппы, подсчет итогов по каждой группе и подгруппе и оформление полученных результатов в виде статистических таблиц.

Группировки дают возможность выделить из состава всех случаев единицы разного качества, показать особенности явлений, развивающихся в различных условиях. после проведения группировки приступают к обобщению данных наблюдения по выделенным частям и целому, то есть к получению статистических показателей в форме абсолютных величин, при помощи которых измеряют объемы (размеры) явлений. Эта ступень работы носит название сводки.

3. Обработка статистических показателей, полученных при сводке, и анализ результатов для получения обоснованных выводов о состоянии изучаемого явления и закономерностях его развития.

При обработке статистических данных исчисляют аналитические показатели, отражающие особенности отельных однородных групп (подгрупп), соотношения и взаимосвязи между ними. Для этого этапа исследования характерно применение всего арсенала статистических методов; применение специальных методов предопределяется поставленными задачами и особенностями первичной информации.

Таким образом, специфический метод статистики основан на соединении анализа и синтеза. Сначала выделяются в составе изучаемого явления и раздельно изучаются части (группы и подгруппы), оценивается существенность или несущественность наблюдаемых различий в величине признака, выявляются причины различий, а затем дается характеристика явления в целом, во всей совокупности его сторон, тенденций и форм развития.

Вернуться в начало.

1.2 Понятия и категории статистики

При изучении явлений статистика использует ряд особых понятий, составляющих ее специфический язык, а именно:

1. Признак.
2. Показатель.
3. Система показателей.
4. Совокупность.
5. Вариация.

Признаком принято называть свойство, характерную черту, особенности объектов (явлений), которые можно наблюдать или измерить (вид выпускаемой продукции, численность, величина основных производственных фондов).

Признаки делятся:

– на качественные и количественные;

– основные (существенные) и второстепенные;

– первичные и вторичные.

Под качественными понимают такие признаки, определенные значения которых отличаются друг от друга существенными моментами (так называемые атрибутивные свойства), например, профессия – токарь, слесарь и т.д. или пол – мужчины, женщины.

Количественные – признаки, отдельные значения которых отличаются друг от друга по величине (возраст, заработная плата), они выражаются числом.

Основные признаки определяют главное содержание процессов, второстепенные – дают добавочные сведения о свойствах и не связаны непосредственно с внутренним содержанием явлений.

Также признаки могут быть разделены на первичные – получаемые при сборе статистических данных и вторичные – получаемые при обработке этих данных.

Показатель – обобщающая количественная характеристика явлений (процессов) в их качественной определенности в условиях конкретного места и времени (уровень производительности труда по годам, себестоимость, заработная плата).

Систему показателей образуют несколько показателей, всесторонне отражающих развитие явления.

Совокупность – множество объектов, явлений, которые изучаются статистикой и которые имеют один или несколько общих признаков и различаются по другим (студенческая группа, поток, трудовой коллектив, количество цехов на предприятии).

Вариация – колеблемость, многообразие, изменяемость величин признака у отдельных единиц совокупности.

Пределы, в которых возможны изменения величины количественного признака, называются границами вариации: нижняя – минимальная величина, верхняя – максимальная величина (например, массовая доля жира в молоке). Отдельное значение признака называется вариантом.

Вернуться в начало

1.3 Стадии статистического исследования

Статистическое исследование проходит в несколько этапов. Начинается оно со сбора статистических данных, в результате получают первичную статистическую информацию. Такая планомерная регистрация существенных признаков элементов статистической совокупности называется статистическим наблюдением – это первый этап статистического исследования.

На втором этапе статистического исследования собранные данные подвергаются систематизации и группировке. Эта стадия называется сводкой статистических данных. Важнейшим методом, применяемым на этом этапе, является метод группировок, который состоит в выявлении типов, структуры, то есть в разделении совокупности на группы и подгруппы по определяемым признакам. На этом этапе переходят от описания отдельных единиц к описанию групп. При этом подсчитываются итоги, вычисляются обобщенные показатели в виде относительных и средних величин.

Третий этап – анализ и обобщение статистических фактов и обнаружение закономерностей в изучаемых явлениях. На этой стадии получают выводы о состоянии изучаемого явления, закономерностях его развития. Выводы и сам анализ излагается, как правило, в виде текста и сопровождается графическим и табличным материалом.

Таким образом, статистическому исследованию присущи несколько этапов:

- статистическое наблюдение;
- сводка статистических данных;
- анализ и обобщение статистических фактов, и обнаружение закономерностей в изучаемых явлениях.

Вернуться в начало

2 СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

2.1 Понятие, объект и единица статистического наблюдения

Статистическое наблюдение представляет собой планомерный, научно-организованный, как правило, систематический сбор данных о явлениях и процессах общественной жизни путем регистрации заранее намеченных существенных признаков.

В процессе статистического наблюдения формируется первичный материал – статистические данные, которые затем подвергаются обработке (сводке и группировке), анализу и обобщению. При проведении статистического наблюдения необходимо точно определить, что подлежит наблюдению, то есть объект (например, численность населения, в том числе городского и сельского; производство продукции, в том числе по отраслям и/или цехам, заводам, наименованию продукции).

Объектом статистического наблюдения называется совокупность, о которой должны быть собраны необходимые данные. Любая совокупность (объект статистического наблюдения) состоит из отдельных элементов. Характеристика объекта может быть получена лишь посредством характеристики отдельных элементов. При этом единицей наблюдения называют составной элемент объекта наблюдения, который является носителем признака, подлежащего регистрации.

Единицы наблюдения, как и объект в целом, обладают множеством различных свойств, качеств, называемых признаками. Учесть их все невозможно. Поэтому при организации статистического наблюдения необходимо определить, какие признаки следует регистрировать.

Перечень признаков, регистрируемых в процессе наблюдения, называется программой статистического наблюдения. Программа наблюдения оформляется в виде бланка (анкеты, формуляра), в который заносятся первичные сведения. Необходимым дополнением к бланку является инструкция (или указания на самих формулярах), разъясняющая смысл вопроса. Содержание программы определяется сущностью, свойствами объекта наблюдения, зависит также от цели наблюдения, потребности в определенных статистических данных. Объем программы зависит:

– от средств, которыми располагают статистические органы, проводящие наблюдение;

– от срочности, с которой нужно получить необходимые данные.

Программа должна содержать существенные признаки, непосредственно характеризующие изучаемые явления.

В программу не следует включать второстепенные вопросы, которые затрудняют работу по сбору информации, ее обработке и анализу.

В программу следует включать вопросы контрольного характера, служащие целям проверки и уточнения информации.

Для записи ответов на вопросы программы наблюдения разрабатывается формуляр наблюдения.

Формуляр наблюдения — это особым способом сформированный бланк, в котором содержатся перечень вопросов программы и куда заносятся собираемые сведения. Статистический формуляр должен быть удобен для чтения, записи и обработки. В формуляре должно быть предусмотрено место для ответа.

Вернуться в начало

2.2 Формы и виды статистического наблюдения

Все организационные формы статистического наблюдения могут быть сведены к двум основным типам, таким как:

а) отчетность;

б) специально организованные (переписи, единовременные учеты).

Отчетностью называют такую организационную форму статистического наблюдения, при которой сведения поступают в статистические органы в виде обязательных отчетов о деятельности предприятий, организаций, фирм и т.п. (например, форма П-1-предприятие «Основные сведения о деятельности предприятия», форма № 11 «Сведения о наличии и движении основных фондов (средств) и других нефинансовых активов» и др.). Она представляется по заранее установленной программе и в строго установленные сроки. Отчетность является одним из основных источников статистических сведений.

Специально организованное статистическое наблюдение представляет собой наблюдение, которое организуется с какой-либо специальной целью на определенную, как правило, дату для получения данных, которые в силу тех или иных причин невозможно собрать посредством отчетности, или для проверки, уточнения данных отчетности (перепись населения).

Кроме того, в зависимости от признака, положенного в основу деления на виды статистические наблюдения можно разделить:

1. В зависимости от периодичности, систематичности проведения, статистическое наблюдение может быть:

– текущим (непрерывным);

– прерывным.

Причем прерывное, в свою очередь, подразделяется:

– на периодическое;

– единовременное.

Текущим называется такое наблюдение, которое ведется систематически, непрерывно. Регистрация фактов проводится по мере их свершения (например, учет выработки продукции, учет явки рабочих).

Периодическим называется такое наблюдение, которое повторяется через определенные, равные промежутки времени (например, ежегодная перепись скота).

Единовременным называется такое наблюдение, которое проводится по мере надобности, время от времени, без соблюдения строгой периодичности или вообще проводится однажды и затем снова никогда не повторяется (например, учет товарных остатков и денежной наличности, проведенный в СССР в 1947 г.).

2. В зависимости от того, как проводится наблюдение, различают:

– непосредственное наблюдение;

– документальное;

– опрос.

Непосредственным называют такое наблюдение, при котором сами регистраторы путем непосредственного замера, взвешивания, подсчета устанавливают факт и на этом основании производят записи в формуляре наблюдения (например, инвентаризация товара в магазине, снятие остатков денежных средств в кассе и др.). Документальное – наблюдение, при котором запись ответов на вопросы формуляра наблюдения производится на основании соответствующих документов (например, рекламации на проданную продукцию).

Опрос – это наблюдение, при котором ответы на вопросы формуляра наблюдения записываются со слов опрашиваемого (например, перепись населения).

3. В зависимости от степени, полноты охвата наблюдением изучаемого объекта, статистические наблюдения подразделяют на сплошное и несплошное.

Сплошным называют такое наблюдение, при котором обследованию подвергаются все без исключения единицы изучаемой совокупности.

Несплошным называют такое наблюдение, при котором обследованию подвергаются не все, а только часть единиц изучаемой совокупности. Несплошное наблюдение проводится с целью получения характеристики объекта в целом. Причем заранее при проектировании наблюдения устанавливается:

а) сам факт, что наблюдение будет несплошным;

б) какая именно часть совокупности должна подвергнуться наблюдению;

в) каким образом следует отобрать эти единицы.

В практике статистики применяется несколько видов несплошного наблюдения:

1) выборочное;

2) монографическое обследование;

3) метод основного массива.

Выборочным называется наблюдение, основанное на принципе беспристрастного (случайного) отбора тех единиц изучаемой совокупности, которые должны быть подвергнуты наблюдению. При правильной организации оно дает достаточно достоверные данные для характеристики изучаемой совокупности в целом.

Монографическое обследование представляет собой детальное, глубокое изучение и описание отдельных, характерных в каком-либо отношении единиц интересующей исследователя совокупности. Оно проводится с целью выявления имеющихся или намечающихся тенденций в развитии, для изучения и распространения опыта и т. п.

Метод основного массива заключается в том, что обследованию подвергаются наиболее существенные, обычно наиболее часто встречающиеся единицы изучаемой совокупности. Взятые вместе, они имеют преобладающий удельный вес в совокупности по основному для данного исследования признаку.

Вернуться в начало

2.3 Способы статистического наблюдения

В статистике применяются следующие основные способы наблюдения:

1) отчетный;

2) экспедиционный;

3) саморегистрация (самоисчисление);

4) анкетный;

5) корреспондентский.

Отчетный способ заключается в том, что представляются отчеты о деятельности в строго обязательном порядке. Заполнение форм статистической отчетности и их представление осуществляют подразделения или специальные работники из штата предприятия (фирмы).

Экспедиционный способ заключается в том, что специально привлеченные и обученные работники, которых обычно называют счетчиками (или регистраторами), посещают каждую единицу наблюдения и сами заполняют формуляр наблюдения. Применяется при специально организованном наблюдении (перепись населения). Это громоздкий и дорогостоящий способ наблюдения.

Суть способа самоисчисления заключается в том, что задача счетчиков – раздать формуляры наблюдения, объяснить, как их заполнить и проверить правильность заполнения. Формуляры наблюдения заполняют сами опрашиваемые. Это позволяет значительно экономить время и снижает затраты на проведение наблюдения по сравнению с экспедиционным способом.

Анкетный способ – это сбор статистических данных с помощью специальных вопросников (анкет), распространяемых среди определенного круга лиц или публикуемых в периодической печати. Он основан на принципах добровольности и обычно анонимности заполнения анкет. Проводится в тех случаях, когда не требуется получения точных статистических данных, а нужны лишь приближенные характеристики, ориентировочные данные.

Корреспондентский способ заключается в том, что статистические органы договариваются с определенными лицами, которые берут на себя обязательство вести наблюдение за какими-либо явлениями, процессами и в установленные сроки предоставлять результаты наблюдения статистическим органам. Корреспонденты не состоят в штатах статистических органов, работают добровольно и дают субъективную оценку явлений.

Общая классификация статистических наблюдений по пяти признакам представлена на рис. 3.

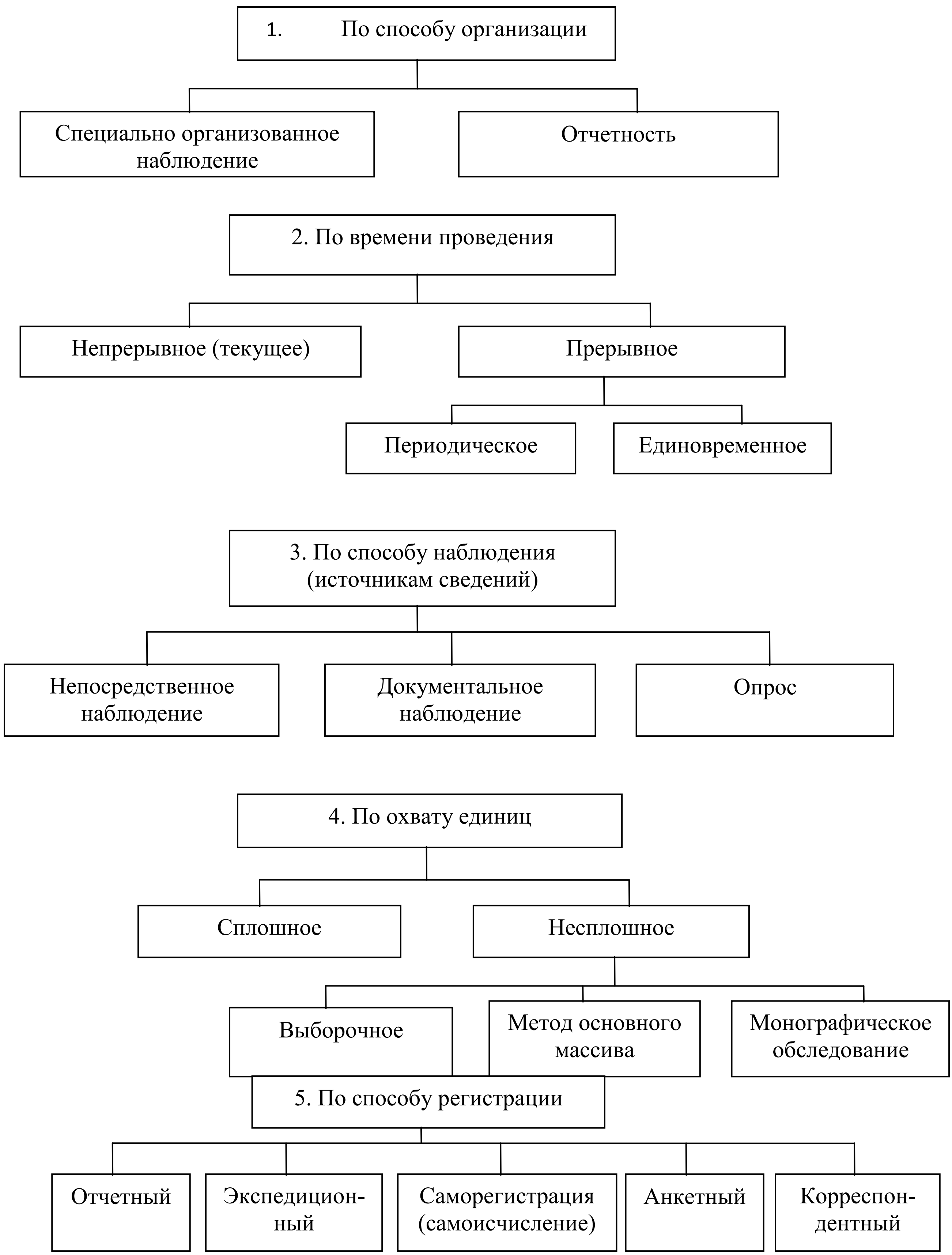

Рис. 3. Виды статистического наблюдения
по признакам классификации

[Вернуться в начало]

2.4 Точность статистического наблюдения

Точностью статистического наблюдения называется степень соответствия значения какого-либо признака, найденного путем статистического наблюдения, действительному его значению. Расхождения между установленным и действительным значениями называются ошибками наблюдения.

Различают два основных типа ошибок наблюдения:

1) ошибки регистрации;

2) ошибки репрезентативности (представительности).

Каждый из этих типов ошибок подразделяется, в свою очередь:

а) на случайные;

б) систематические.

1. Ошибки регистрации образуются вследствие неправильного установления фактов в процессе наблюдения, либо ошибочной записи, либо и того и другого вместе. При этом:

– случайными называются ошибки, которые возникают вследствие различных случайных причин (оговорился, записал в обратном порядке и т. п.). При достаточно большом числе наблюдений они, благодаря действию закона больших чисел, взаимоуничтожаются (так как действуют в различных направлениях);

– систематические возникают под действием определенных причин. В каждом случае они действуют в одном и том же направлении и приводят к серьезным искажениям общих результатов наблюдения (например, округление).

Ошибки регистрации могут иметь место при сплошном и несплошном наблюдении. В большинстве случаев точно величину ошибок регистрации определить не представляется возможным, можно получить только приближенную оценку.

2. Ошибки репрезентативности – отклонение величины изучаемого признака в отобранной и обследованной части совокупности от величины его во всей совокупности. Они свойственны только несплошному наблюдению.

Ошибки репрезентативности могут быть случайными и систематическими.

Случайные ошибки возникают в силу несплошного характера наблюдения, так как совокупность отобранных единиц на основе принципа беспристрастного случайного отбора единиц наблюдения неполно воспроизводит совокупность в целом. Величина случайной ошибки может быть оценена с помощью предельной ошибки выборки (Δ) и средней ошибки выборочного наблюдения (μ).

Систематические – возникают вследствие нарушения принципов беспристрастного отбора тех единиц изучаемой совокупности, которые должны быть подвергнуты наблюдению. Размеры систематической ошибки репрезентативности обычно не поддаются количественной оценке (отбирают лучших, худших и т. п.).

Вернуться в начало

3 СВОДКА И ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ

3.1 Понятие сводки и группировки

В результате статистического наблюдения получают данные о каждой единице совокупности, которые характеризуют ее со многих сторон. Эти сведения служат средством характеристики совокупности в целом. Но такую характеристику можно получить лишь после того, как проведена сводка материалов статистического наблюдения.

Сводка – второй после статистического наблюдения этап статистических исследований – представляет собой научную обработку первичных материалов статистического наблюдения для характеристики совокупности обобщающими показателями.

Статистическая сводка ведется по заранее составленной программе. В программе, прежде всего, определяют подлежащее и сказуемое.

Подлежащее сводки составляют группы или части, на которые разбивается совокупность. **Сказуемое** составляют показатели, характеризующие каждую группу и совокупность в целом. При проведении сводки используется метод группировок.

Простые итоговые данные дают слишком общее представление о совокупности, в то время как требуется не только характеристика всего наблюдаемого объекта, но и знание отдельных частей, групп. Чтобы иметь сведения о группах, совокупность необходимо разбить.

Разбиение множества единиц (совокупности) объекта наблюдения на однородные группы по определенным, существенным для них признакам называется **группировкой**.

При проведении метода группировки могут решаться следующие основные задачи:

– выделение социально-экономических типов;

– изучение структуры явления и структурных сдвигов, происходящих в нем;

– выявление связи и зависимости между явлениями.

Прежде чем проводить группировку, необходимо определить группировочный признак, то есть основание группировки. **Основанием** является признак, по которому совокупность разбивают на группы.

Выбор основания зависит от цели данной группировки. При этом основанием может служить либо атрибутивный (то есть выражающий свойства данного явления их наименованием) качественный признак, либо количественный – цифровое выражение. Отнесение их к соответствующим группам зависит от размера признака, взятого в основание группировки.

Если в основание группировки положен непрерывный количественный признак, то возникает вопрос о численных границах групп, то есть об интервалах.

Интервалы могут быть равными и неравными. **Равные** применяются в тех случаях, когда изменения количественного признака внутри совокупности происходят постепенно, равномерно. **Неравные** – изменения, прогрессивно увеличивающиеся или убывающие.

При образовании интервалов необходимо точно обозначать количественные границы, избегать такие группы, в которых отдельные значения можно отнести в две составные группы. Расчет величины интервала при равных интервалах осуществляется следующим образом

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

где h – шаг, величина интервала;

k – число групп.

Для построения интервального ряда из дискретного используется формула Стерджесса, с помощью которой определяется оптимальное количество интервалов (k):

$$k = 1 + 3{,}322 \lg N,$$

где N – число величин в дискретном ряде.

Вернуться в начало

3.2 Виды группировок

Различают следующие виды группировок.

1. **Типологические группировки**. При проведении типологической группировки исследуемая совокупность (общественное явление) разделяется на классы, социально-экономические типы.

При количественном группировочном признаке в типологической группировке необходимо правильно установить интервал группировки, количественно отделить один класс или тип от другого. Вопрос об интервалах типологической группировки решается на основании определения таких количественных границ, которые выделяют новое качество. Число групп в типологической группировке зависит от числа действительно имеющихся социально-экономических типов (уровень образования; разделение предприятий на крупные, мелкие).

2. **Структурные группировки**. Структурными группировками называется разбиение однородной в качественном отношении совокупности на группы, характеризующие строение совокупности, ее структуру.

Метод структурных группировок позволяет рассмотреть состав совокупности по экономическим и административным районам, по отраслям народного хозяйства, по географическим зонам и другим признакам. Если в структурной группировке сопоставить данные в динамике, то получится представление о структурных сдвигах в изучаемом явлении (например, все население, в том числе городское и сельское).

3. **Аналитические группировки**. Аналитические группировки дают возможность установить связь между отдельными признаками изучаемого социально-экономического явления (например, зависимость основных показателей работы предприятий от их размера).

4. **Вторичные группировки**. При статистических исследованиях иногда приходится производить вторичную группировку, то есть перегруппировывать статистический материал, уже сведенный в группы. К ней прибегают, если начальная группировка не удовлетворяет исследователя.

Ее можно производить путем сведения в новые группы по тому же признаку, как и первичную группировку. В этом случае интервалы либо укрупняются, либо, наоборот, уменьшаются.

5. **Комбинированные группировки**. Предыдущие четыре вида группировок являются группировками по одному признаку, то есть в основании выделенных групп лежит лишь один признак. Когда для расчленения совокупности на группы применяется не один, а два или более группировочных признаков – это комбинированная группировка.

Комбинация группировочных признаков приводит к резкому увеличению числа групп, что, в свою очередь, может привести к недостаточной численности единиц в каждой группе. Поэтому создается опасность на основании небольшого числа наблюдений сделать малообоснованные случайные выводы. К комбинированным группировкам прибегают при достаточно большом числе наблюдений, если задача заключается не в простом изучении структуры, а в исследовании зависимости результативного фактора от двух и более признаков.

Вернуться в начало

3.3 Статистические распределения

3.3.1 Вариация признака в совокупности

Составной частью сводной обработки данных статистического наблюдения является построение рядов распределения. **Рядом распределения** называется группировка, в которой для характеристики групп (упорядоченно расположенных по значению признака) применяется один показатель – численность группы. Цель построения ряда – выявление основных свойств и закономерностей исследуемой статистической совокупности.

В зависимости от того, является ли признак, взятый за основу группировки, качественным или количественным, различают соответственно два типа рядов распределения – **атрибутивные** и **вариационные**.

По характеру вариации различают дискретные и непрерывные признаки. **Дискретные признаки** отличаются друг от друга на некоторую конечную величину; **непрерывные** могут отличаться один от другого на сколь угодно малую величину и в определенных границах принимать любые значения.

Первый шаг в упорядочении первичного ряда – его *ранжирование*, то есть расположение всех вариантов ряда в возрастающем или убывающем порядке.

Число повторений отдельных вариантов значений признаков называют **частотой повторения**. Частота повторения обозначается f_i; сумма частот, равная объему изучаемой совокупности

$$\sum_{i=1}^{k} f_i .$$

Частоты, представленные в относительном выражении, называются частостями

$$w_i = f_i \Big/ \sum_{i=1}^{k} f_i .$$

Частости могут быть выражены в долях единицы или в процентах.

Например, распределение рабочих участка по квалификации и их характеристика представлены в табл. 2.

Таблица 2

Квалификационные характеристики рабочих участка «С»

Тарифный разряд рабочего x_i	Число рабочих, имеющих этот разряд, f_i	Частость w_i	Накопленная частота S_i
2	1	0,05	1
3	5	0,025	6
4	8	0,40	14
5	4	0,20	18
6	2	0,10	20
Итого	20	1,00	

В тех случаях, когда число вариантов дискретного признака достаточно велико, а также при анализе вариации дискретного признака, когда значение признака у отдельных единиц может вообще не повторяться, строятся **интервальные ряды распределения.**

Для определения величины интервала h для построения вариационного ряда с равными интервалами:

1) вычисляется разность между максимальным и минимальным значениями признака первичного ряда (размах вариации – R)

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

2) размах вариации делится на число групп k, то есть $h = R/k$. Число групп приближенно определяется по формуле Стэрджесса

$$k \approx 1 + 3{,}322 \cdot \lg N,$$

где N – общее число изучаемых единиц совокупности. Полученная величина округляется до целого числа.

Рассмотрим пример построения ряда распределения по первичным данным о размере прибыли 20 коммерческих банков за год (млн руб.):

$$x_i - 3{,}7; 4{,}3; 6{,}7; 5{,}6; 5{,}1; 8{,}1; 4{,}6; 5{,}7; 6{,}4; 5{,}9; 5{,}2; 6{,}2; 6{,}3; 7{,}2; 7{,}9; 5{,}8; 4{,}9; 7{,}6; 7{,}0; 6{,}9$$

Количество групп равно $k \approx 1 + 3{,}322 \cdot \lg 20 = 5{,}32$.

Округляя, получаем число групп, равное 5.

Величина интервала (h) определяется

$$(8{,}1-3{,}7)/5=0{,}9 \text{ млн руб.}$$

В результате группировки получаем ряд распределения (табл. 3).

Таблица 3

Распределение банков по величине прибыли

Размер прибыли, млн руб.	Число банков	Накопленная частота
3,7–4,6	2	2
4,6–5,5	4	6
5,5–6,4	6	12
6,4–7,3	5	17
7,3–8,1	3	20
Итого	20	

Знак «–» в первой строке соответствует принципу «исключительно» и означает, что значения признака, совпадающие с верхней границей интервала, в этот интервал не включаются, а попадают в следующий интервал. Если ставится знак «+», это соответствует принципу «включительно» и означает, что значения признака, совпадающие с верхней границей интервала, включаются в эту группу.

Вернуться в начало

3.3.2 Графическое изображение вариационного ряда

Для графического изображения дискретного ряда применяют **полигон распределения**. Для построения по оси абсцисс откладывают значения признака, по оси ординат – частоты или частости. Для замыкания полигона крайние вершины соединяются с точками на оси абсцисс, отстоящими на одно деление от x_{max} и x_{min}. Моде для дискретного ряда распределения соответствует значение, которому соответствует максимальная частота.

Для графического изображения интервальных вариационных рядов применяется **гистограмма.** Для построения по оси абсцисс откладываются равные отрезки, соответствующие величине интервалов, на которых строят прямоугольники с высотой, равной частотам или частостям интервала.

В ряде случаев для изображения вариационных рядов используется **кумулятивная кривая** (кумулята) и **огива**. Для построения кумуляты и огивы по оси абсцисс откладывают значения признака, по оси ординат – накопленные частоты (для кумуляты) или частости (для огивы).

Для графического определения медианы высоту наибольшей ординаты, которая соответствует общей численности совокупности, делят пополам. Через полученную точку проводят прямую, параллельную оси абсцисс, до пересечения ее с кумулятой. Абсцисса точки пересечения является медианой.

Вернуться в начало

3.3.3 Показатели центра распределения

Для характеристики среднего значения признака в вариационном ряду используются средняя арифметическая величина, медиана и мода. Рассмотрим расчет показателей центра распределения для вариационных рядов.

Средняя арифметическая:

– для дискретного ряда распределения

$$\bar{x} = \sum_{i=1}^{k} x_i f_i \Big/ \sum_{i=1}^{k} f_i ,$$

где x_i – вариант значений;

f_i – частота повторения данного варианта;

– для интервального ряда распределения

$$\bar{x} = \sum_{i=1}^{k} x'_i f_i \Big/ \sum_{i=1}^{k} f_i ,$$

где x'_i – средняя соответствующего интервала.

Для табл. 2 средняя арифметическая равна

$$\bar{x} = \frac{2 \cdot 1 + 3 \cdot 5 + 4 \cdot 8 + 5 \cdot 4 + 6 \cdot 2}{20} = 4{,}05 .$$

Для табл. 3 средняя арифметическая равна

$$\bar{x} = \frac{4{,}15 \cdot 2 + 5{,}05 \cdot 4 + 5{,}95 \cdot 6 + 6{,}85 \cdot 5 + 7{,}75 \cdot 3}{20} = 6{,}085 \text{ млрд руб.}$$

Медиана:

– для дискретного ряда распределения положение медианы определяется ее номером $N_{Me} = (n+1)/2$, где n – число единиц совокупности.

Для примера в табл. 2 $N_{Me} = (20+1)/2 = 10{,}5$, то есть медиана равна средней арифметической 10-го и 11-го значений признака. Me = 4;

– для интервального ряда распределения сразу можно определить интервал, в котором находится медиана. Затем определяем медиану по формуле

$$Me = x_{Me} + h\frac{(n+1)/2 - S_{Me-1}}{f_{Me}},$$

где x_{Me} – нижняя граница медианного интервала;

h – величина интервала;

S_{Me-1} – накопленная частота интервала, предшествующая медианному;

f_{Me} – частота медианного интервала.

Для примера, приведенного в табл. 3

$$Me = 5{,}5 + 0{,}9\frac{(20+1)/2 - 6}{6} = 6{,}175 \text{ млрд руб.}$$

Мода:

– для дискретного ряда распределения – наиболее часто встречающееся значение. Для табл. 2 мода равна 4 (максимальная частота 8);

– для интервального ряда распределения

$$Mo = x_{Mo} + h\frac{f_{Mo} - f_{Mo-1}}{[f_{Mo} - f_{Mo+1}] + [f_{Mo} - f_{Mo-1}]},$$

где x_{Mo} – нижняя граница модального интервала;

$_{Mo}$ – частота, соответствующая модальному интервалу;

f_{Mo-1} – предмодальная частота;

f_{Mo+1} – послемодальная частота.

Для примера, приведенного в табл. 3

$$Mo = 5{,}5 + 0{,}9\frac{6-4}{[6-5]+[6-4]} = 6{,}10 \text{ млрд руб.}$$

Вернуться в начало

3.3.4 Показатели формы распределения

Для получения приблизительного представления о форме распределения строят графики распределения (полигон и гистограмму). **Эмпирическое распределение** строится на основе выборки из исследуемой генеральной совокупности. Поэтому эмпирические данные в определенной степени связаны со случайными ошибками наблюдения. Кривая **теоретического распределения** показывает распределение, которое получилось бы при полном погашении всех случайных причин, затемняющих основную закономерность.

Выяснение общего характера распределения предполагает оценку степени его однородности, а также вычисление показателей асимметрии и эксцесса.

Симметричным является распределение, в котором частоты любых двух вариантов, равностоящих в обе стороны от центра распределения, равны между собой. Для сравнительного анализа **степени асимметрии** рассчитывают относительный показатель As

$$As = \frac{\bar{x} - Mo}{\sigma}.$$

Другой показатель асимметрии, предложенный Линдбергом

$$As = П - 50,$$

где П – процент тех значений признака, которые превосходят по величине среднюю арифметическую.

Наиболее точным и распространенным является показатель, рассчитываемый следующим образом

$$As = \frac{\mu_3}{\sigma^3}, \text{ где } \mu_3 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3 f_i}{\sum_{i=1}^{n} f_i}.$$

Положительная величина указывает на наличие правосторонней асимметрии (в этом случае $Mo < Me < \bar{x}$). Отрицательный знак свидетельствует о наличии левосторонней асимметрии ($Mo > Me > \bar{x}$).

Для симметричных распределений рассчитывается показатель **эксцесса** (островерши

нности). Эксцесс представляет собой выпад вершины эмпирического распределения вверх или вниз от вершины кривой нормального распределения. Показатель эксцесса, предложенный Линдбергом

$$Ex = П - 38{,}29,$$

где П – доля (%) количества вариантов, лежащих в интервале, равном половине среднего квадратического отклонения в ту и другую сторону от средней арифметической.

Наиболее точным является показатель, рассчитываемый по формуле

$$Ex = \frac{\mu_4}{\sigma^4} - 3, \text{ где } \mu_4 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^4 f_i}{\sum_{i=1}^{n} f_i}.$$

Значения эксцесса могут быть положительными (для островершинных распределений) и отрицательными (для плосковершинных).

Вернуться в начало

3.4 Статистические таблицы

Результаты сводки и группировки обычно излагаются в виде таблиц, в которых наглядно проявляется связь между признаками.

Основная особенность табличного изложения состоит в том, что показатели можно объединить под общим заголовком. При этом графленая сетка – это скелет таблицы. Вертикальные столбцы – графы, горизонтальные полосы – строки. Если записать заголовки граф и строк, то получим макет таблицы. Полная таблица – если внесены данные.

Перед сводкой необходимо составить макеты таблиц, которые являются конкретным выражением плана сводки. Макет статистической таблицы представлен в табл. 4.

Таблица 4

Макет таблицы

Сказуемое Подлежащее	Заголовки граф		
Перечень (группы) единиц совокупности			

Статистическая таблица имеет подлежащее и сказуемое.

Подлежащее – это то, о чем идет речь в таблице. В нем дается перечень отдельных элементов или групп характеризуемого явления. **Сказуемое** показывает, какими признаками характеризуется подлежащее. В нем отражаются численные характеристики элементов или групп данного явления, указанных в подлежащем.

Подлежащее составляет содержание строк, сказуемое – граф (записывается сверху).

Обязательной составной частью таблицы является общий заголовок, который кратко сообщает, о чем идет речь в таблице, к какому месту и времени она относится.

В зависимости от разработки подлежащего или от группировки единиц в подлежащем, выделяют три вида таблиц: простые, групповые и комбинационные.

При такой классификации разница между простой и групповой таблицами заключается в разном приеме объединения числовых данных. Различие между групповой и комбинационной таблицами состоит в разном числе признаков, по которым происходит разбиение совокупности единиц. Кроме того, простые таблицы ставят задачу дать лишь перечень изучаемых объектов, групповые и комбинационные – изучить взаимную зависимость.

1. **Простые** – таблицы, в подлежащем которых нет группировок (табл. 5).

Простые таблицы бывают:

– перечневые (подлежащее – перечень единиц, составляющих объект изучения);

– территориальные (дается перечень территорий, стран, областей и т. д.);

– хронологические (в подлежащем приводятся периоды или даты).

Таблица 5

Численность населения Кемеровской области

Показатель	Год						
	2003	2004	2005	2006	2007	2008	2009
Численность населения, тыс. чел.	2893,5	2872,1	2855,0	838,5	2846,3	2823,5	2821,9

2. **Групповые** – таблицы, в подлежащем которых изучаемый объект разделен на группы по какому-либо признаку (табл. 6).

Таблица 6

Численность населения Кемеровской области по полу

Контингент	Год	
	2007	2008
Оба пола, тыс. чел. в том числе:	2846,3	2823,5
мужчины	1267,1	1277,5
женщины	1579,2	1546

3. **Комбинационные** – таблицы, в которых совокупность разделяется на группы не по одному, а по нескольким признакам (табл. 7).

По построению сказуемого таблицы также можно разделить на простые и сложные. Простая разработка сказуемого предусматривает параллельное расположение показателей, а сложная разработка – комбинированное.

Таблица 7

Численность населения Кемеровской области
по полу и возрасту, тыс. чел.

Контингент	Год					
	базисный			отчетный		
	Всего	В том числе		Всего	В том числе	
		мужчины	женщины		мужчины	женщины

Всего, в том числе:	2846,3	1579,2	1267,1	2823,5	1546	1277,5
моложе трудоспособного возраста	458,5	364,9	93,6	448,3	358,6	89,7
трудоспособного возраста	1797,9	1123,5	674,4	1689,4	1069,7	619,7
старше трудоспособного возраста	589,9	90,8	499,1	685,8	117,7	568,1

Практикой выработаны определенные требования к составлению и оформлению таблиц:

1. Таблица по возможности должна быть краткой. Не следует загружать ее излишними подробностями, затрудняющими анализ явлений.

2. Каждая таблица должна иметь название, из которого становится известно:

– какой круг вопросов излагает и иллюстрирует таблица;

– каковы географические границы статистической совокупности, представленной таблицей;

– каковы единицы измерения (если они одинаковы для всех табличных клеток). Если единицы измерения неодинаковы, то в заголовках обязательно следует указывать, в каких единицах приводятся статистические данные.

Заглавия строк подлежащего и сказуемого должны быть сформулированы точно, кратко и ясно.

3. Если таблица разбивается на части, то делается нумерация граф. Это облегчает пользование таблицей, дает возможность лучше ориентироваться, показывает способ расчета цифр в графах.

4. Приводимые в подлежащем и сказуемом признаки должны быть расположены в логическом порядке с учетом необходимости рассматривать их совместно.

5. Таблица может сопровождаться примечаниями, в которых указываются источники данных и даются необходимые пояснения.

6. При оформлении таблиц обычно применяют следующие условные обозначения:

– знак тире «–» – когда явление отсутствует;

– «х» – если явление не имеет осмысленного содержания;

– многоточие «…» – когда отсутствуют сведения о его размере; или делается запись «нет сведений».

7. Округленные числа приводятся в таблице с одинаковой степенью точности.

Вернуться в начало

3.5 Графическое изображение статистических показателей

3.5.1 Составные элементы графиков, их классификация

Полученный в результате обработки статистический материал часто нуждается в наглядном изображении, для чего используются графики. В статистике **графиком** называют наглядное изображение статистических величин при помощи геометрических линий и фигур или географических схем.

Статистические графики – это в основном плоскостные геометрические знаки, отражающие размеры явлений. При этом в графическом изображении можно выделить следующие обязательные элементы:

1. Поле графика – пространство для размещения знаков, которое имеет определенные размеры и пропорции сторон (чистая бумага, географические или контурные карты и т. п.).
2. Геометрические знаки – символы понятий, отражаемых на графике (точки, отрезки прямых, круги, геометрические фигуры, силуэты).
3. Пространственные ориентиры – ориентиры, определяющие размещение знаков в поле графика. Эти ориентиры зависят от принятой системы координат (прямоугольная, косоугольная).
4. Масштабные ориентиры – эталоны знака, отражающие величину геометрических знаков. Они изображаются в виде кругов, прямоугольников и т. п., обычно выносятся с поля графика. Величину явления можно определить, сравнивая геометрический знак с эталоном.
5. Экспликация – словесное объяснение содержания графика и значения каждого его геометрического знака.

В основу классификации графиков положены различные признаки (рис. 4). Так, по различию полей графиков все графики делят на две группы. Полем графика может быть либо бумага (доска, ватман), либо географическая или контурная карта. С этой точки зрения различают:

1) диаграммы;

2) статистические карты (картосхемы).

В свою очередь, в зависимости от того, для какой цели применяется график, каждую группу делят на виды. Виды диаграмм бывают следующие:

1) сравнения;

2) структурные;

3) динамические.

Виды статистических карт:

1) картограммы;

2) картодиаграммы;

3) центрограммы.

Кроме того, в зависимости от применяемого геометрического знака, в каждом виде могут выделяться подвиды:

1) круговые;

2) столбиковые;

3) полосовые;

4) изобразительные;

5) линейные;

6) точечные и т. д.

Построение графиков требует соблюдения ряда правил:

1) заголовок графика должен кратко выражать всю сущность отображаемого явления;

2) обязательно должно быть указано место и время отображаемого явления (в заголовке или на самом графике);

3) масштабную шкалу нужно начинать от нуля, а не от числа, близкого к минимальному значению в изображаемом ряду. В случае необходимости следует делать разрыв на осях;

4) величина геометрического знака должна соответствовать в определенном масштабе величине отображаемого явления;

5) в диаграммах сравнения геометрические знаки должны располагаться в ранжированном порядке.

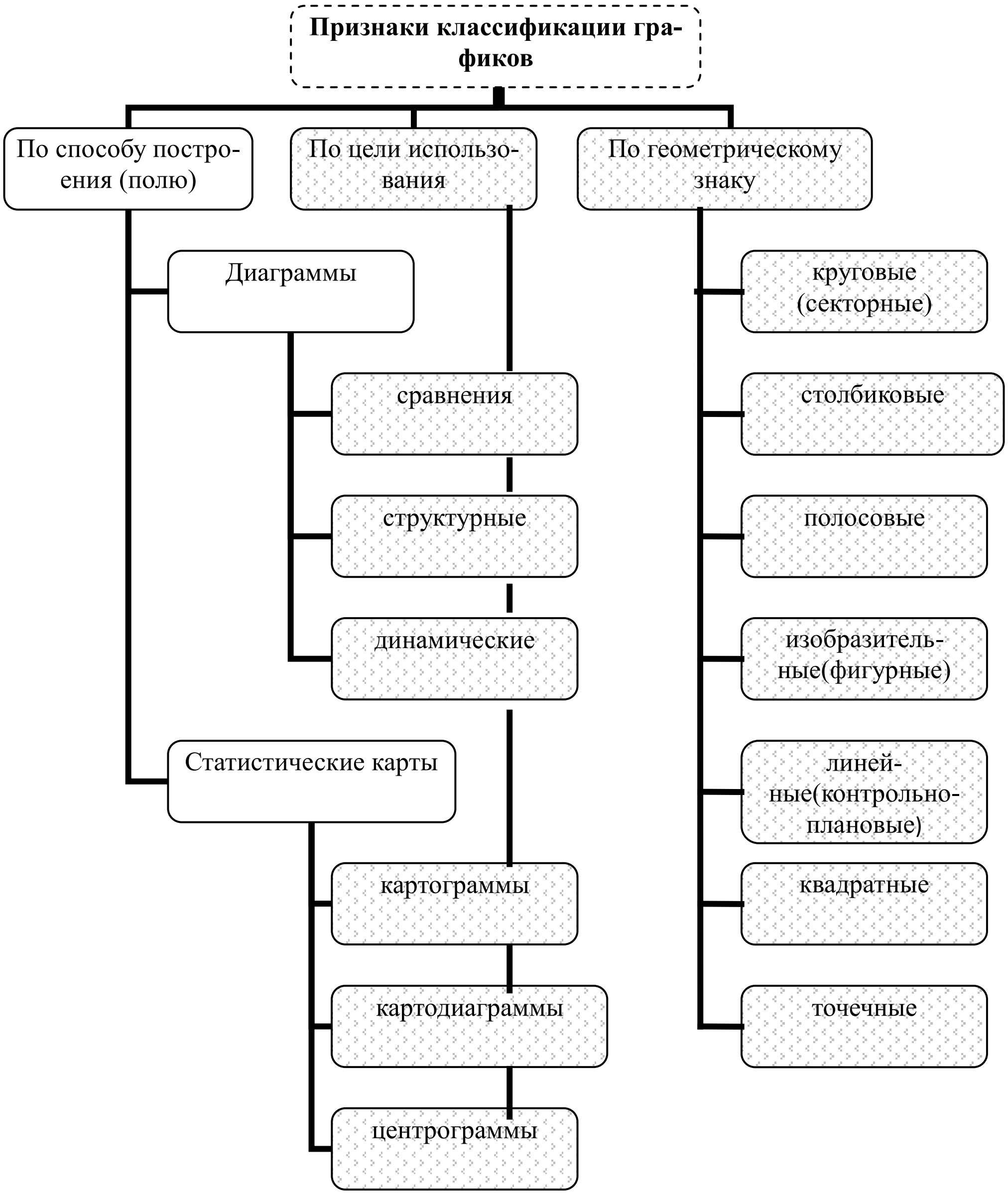

Рис. 4. Классификация графиков

Вернуться в начало

3.5.2 Виды диаграмм

Диаграмма – график, для которого полем является чистый лист бумаги (доска, ватман и т. п.).

В зависимости от цели различают следующие виды диаграмм:

1) диаграммы сравнения;
2) структурные диаграммы;
3) динамические диаграммы.

В каждой из диаграмм выделяют подвиды по геометрическому знаку.

1. **Диаграммы сравнения** используются для изображения различных статистических совокупностей по какому-либо изменяющемуся в пространстве признаку, то есть сравниваются различные объекты по одному признаку (например, численность населения различных стран).

Наиболее распространенными подвидами при этом являются:

а) столбиковые;

б) полосовые.

Столбики и полосы могут изображать явление в нескольких разрезах или по группам.

2. **Структурные диаграммы** позволяют сопоставить статистическую совокупность по составу, это прежде всего диаграммы удельных весов, характеризующих отношение отдельных частей совокупности к ее общему объему.

Среди них наиболее распространены следующие разновидности:

а) столбиковые;

б) секторные;

в) знаки Варзара.

Особая разновидность структурных диаграмм – знаки Варзара. Они позволяют отобразить на графике структуру по трем признакам, из которых произведение двух имеет определенный экономический смысл.

3. **Динамические диаграммы.** Их назначение состоит в показе изменения явления во времени. Чаще всего используются следующие подвиды:

а) столбиковые;

б) полосовые – каждый столбик или полоса отражают величину явления за определенный промежуток времени или на определенную дату;

в) круговые;

г) квадратные – в этих графиках величину явления отображают круги или квадраты, значения радиусов или сторон которых пропорциональны абсолютной величине признака;

д) изобразительные – на данных графиках статистические величины отображаются легко запоминаемыми символами, воспроизводящими внешний образ явления. При графическом изображении лучше использовать различное число одинаковых по размеру знаков-символов, причем каждому стандартному знаку придается определенное числовое значение;

е) контрольно-плановые (линейные) – в них геометрическим знаком является отрезок линии.

Вернуться в начало

3.5.3 Статистические карты, их виды

Статистические карты дают статистико-географический разрез данных, то есть показывают размещение явления по территории. В зависимости от условий применения различают:

1) картограммы;

2) картодиаграммы;

3) центрограммы.

1. **Картограммы** иллюстрируют статистические таблицы по одному признаку, то есть размещение одного показателя (например, численности) по территории. На них даются обычно лишь контуры географических или административных районов и необходимые пространственные ориентиры (города, районы и т. п.). Среди картограмм наибольшее распространение получили:

а) фоновые – отражают изменение одного статистического показателя по территории посредством различного цвета или интенсивности окраски и штриховки. Из них наиболее известны картограммы плотности населения. Они отображают распределение районов по среднему числу жителей на 1 км^2;

б) точечные – в них графическим знаком статистических данных являются точки строго определенного размера, размещенные в заданных границах. Каждая точка несет определенную числовую нагрузку. Они показывают степень концентрации этих объектов в различных районах и могут использоваться во многих отраслях статистики.

2. **Картодиаграммы.** Основная задача картодиаграмм состоит в показе географического распределения отображаемого статистического явления. Картодиаграммы иллюстрируют размещение по территории какого-либо явления, характеризуемого

по двум или более показателям, или территориальный разрез динамики одного показателя.

В самом простом виде картодиаграмму можно представить в виде столбиков, кругов разной величины, нанесенных на карту, где величина геометрического знака зависит от размера данного явления в отображаемом районе. Можно использовать и секторные диаграммы (при изображении только структуры (%) круги данных имеют одинаковый радиус).

К картодиаграммам относятся и схемы транспортных потоков. При построении этих схем на картах транспортных маршрутов отображаются объемы, а иногда и структура транспортных грузов.

3. **Центрограммы** позволяют составить целые статистико-географические описания. На них можно нанести ряды динамики различных статистических показателей для отдельных территорий, что позволяет наглядно представить отдельные стороны протекания изучаемого процесса в числовом виде. Центрограммы – это такие картосхемы, на которых размещаются целые таблицы, то есть для статистико-географического описания можно разместить соответствующие данные динамических рядов не в таблице, а на контурных, географических картах. Центрограммы позволяют отобразить динамику удельного веса отдельных районов, тенденцию размещения центра тяжести в расположении отдельных явлений. Они нашли широкое распространение при изучении миграции населения, изучении перемещения центров производства различных промышленных продуктов.

Вернуться в начало

4 АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

4.1 Формы выражения статистических величин

Статистическая величина – это количественная характеристика размеров явлений, их соотношения, степени изменения, взаимосвязи.

Различают абсолютные, относительные и средние величины.

Абсолютные величины выражают размеры явлений в единицах меры массы, протяженности, объема и т. п., то есть величину явления, взятую саму по себе, безотносительно к размерам других явлений (например, объем произведенной продукции, себестоимость продукции).

Относительные величины – это соотношение величины данного явления с величиной какого-нибудь другого явления, но взятой за другое время или по другой местности (например, процент выполнения плана по производству продукции, процент (доля) бракованной продукции в общем объеме производства).

Средние величины – это обобщающая характеристика, выражают величину признака (показателя) в расчете на единицу совокупности (например, средний процент выполнения плана про трем цехам, средний возраст рабочих на предприятии).

Кроме этого, статистические величины можно разделить:

1) на моментные и интервальные;

2) дискретные и непрерывные;

3) прямые и обратные величины.

Моментные величины отображают размер по состоянию на момент времени, на определенную дату (например, численность работающих на 1 января).

Интервальные величины отображают размер явления за промежуток времени (например, объем производства за сутки, год, пятилетку).

Дискретные (прерывные) – изменяющиеся прерывно, скачками, принимающие, как правило, целые значения (например, численность тракторов, поголовье скота).

Непрерывные – такие, которые в определенных пределах могут принимать любые значения (например, доход, заработная плата).

Прямые и обратные – это тесно взаимосвязанные показатели, один из которых является в математическом смысле обратной величиной другого (производительность и трудоемкость) и может быть рассчитан как отношение единицы к первому.

Вернуться в начало

4.2 Абсолютные статистические величины

Абсолютная величина отражает количественную сторону сущности, свойства явления (например, численность населения

Кемеровской области в 2009 г. была 2821 тыс. человек). **Абсолютные величины** – это числа именованные, имеющие определенную размерность, единицы измерения. Характеризуют наличие материальных, денежных, трудовых ресурсов, условно обозначаются: «а»

Различают три вида абсолютных величин: индивидуальные, групповые и общие.

Индивидуальные величины – это такие абсолютные величины, которые выражают размеры количественных признаков у отдельных единиц (например, возраст Петрова, стаж работы Иванова и т. д.). Они устанавливаются непосредственно в процессе статистического наблюдения.

Групповые и общие величины выражают величину того или иного признака у отдельной группы или всех единиц данной совокупности. Получают их в результате суммирования или других вычислений над индивидуальными абсолютными величинами.

Абсолютные величины, в зависимости от выражения их в определенных единицах измерения, бывают:

1. **Натуральные**. Натуральными принято называть такие единицы измерения, которые выражают величину предметов в физических мерах (веса, объема, длины, площади) в соответствии с физическими свойствами (например, объем производства – в т, $м^3$; электроэнергия – в кВт/час, численность – чел.).

2. **Стоимостные**. Стоимостные величины используются для характеристики величины в денежном выражении (доход населения – руб., объем производства – млн руб.).

3. **Трудовые**. Трудовые величины используются для измерения затрат труда на производство продукции, выполняемые работы и т. п. (в час и чел.-час, в днях и чел.-днях и т. п.).

Вернуться в начало

4.3 Относительные величины, их виды

Относительные величины – это величины, выражающие количественные соотношения между явлениями, их признаками. Получают их обычно путем деления одной величины на другую. Чаще всего – соотношением двух абсолютных величин.

Величина, с которой производится сравнение (знаменатель), называется **основанием** (базой, или базисной величиной), а та, которая сравнивается– **текущей** (сравниваемой, отчетной).

$$t = \frac{a_1}{a_0},$$

где t – относительная величина;

a_0, a_1 – абсолютная величина, соответственно базисная и текущая (0 – базисный период; 1 – текущий, отчетный, данный период).

Для конкретных показателей

$$t = z_1/z_0, \text{ или } t = q_1/q_0, \text{ или } t = w_1/w_0,$$

где z_0, z_1 – себестоимость единицы продукции в базисном и отчетном периодах;

q_0, q_1 – объем производства в базисном и отчетном периодах;

w_0, w_1 – выработка в базисном и отчетном периодах.

Относительная величина показывает, во сколько раз сравниваемая величина больше (меньше) базисной, или какую долю (часть) составляет, или сколько единиц приходится на сто, тысячу, сто тысяч и т. д. другой величины.

Важнейшее свойство относительных величин: они позволяют сравнивать такие явления, абсолютные величины, которые несравнимы, несопоставимы.

Относительные величины могут получаться в результате сопоставления одноименных и разноименных абсолютных величин.

Если сопоставляют одноименные величины, то получается относительная величина, не имеющая наименования (возраст, лет / возраст, лет; объем производства, тыс. руб. / объем производства, тыс. руб.). Они могут выражаться в виде кратного отношения (во сколько раз больше или меньше: например, численность РФ в 23 раза больше численности Азербайджана) или в процентном отношении, при котором базисная величина принимается за 100.

$$t = \frac{a_1}{a_0} \cdot 100\ \%$$

Если базисная величина принимается за 1000, то относительная величина выражается в промилле – ‰ (например, прирост населения по стране – 8 ‰, детская смертность – 25 ‰).

Если сопоставляются разноименные величины, то получаются именованные относительные величины, наименование которых образуется сочетанием наименований сравниваемой и базисной величин (например, плотность населения равна численности населения, деленной на площадь, чел./км^2).

В зависимости от содержания выделяют следующие группы и виды относительных величин.

Первая группа включает взаимосвязанные относительные величины:

– относительная величина планового задания ($t_{пл.з}$) – отношение величины показателя, устанавливаемой на планируемый период ($z_{пл.}$, $q_{пл.}$), к его величине, достигнутой в базисном периоде ($z_о$; $q_о$).

В общем виде

$$t_{пл.з.} = a_{пл}/a_о,$$

для конкретных показателей

$$t_{пл.з.} = z_{пл}/z_о;\ q_{пл}/q_о;\ w_{пл}/w_о;\ p_{пл}/p_о,$$

где $z_{пл}$ – себестоимость единицы продукции, планируемая на данный период;

$q_{пл}$ – объем продукции, планируемой на данный период;

$w_{пл}$ – выработка, планируемая на данный период;

$p_{пл}$ – цена единицы продукции, планируемая на данный период;

– относительная величина выполнения плана ($t_{вып.пл.}$) – величина, выражающая соотношение между фактическим и плановым уровнем показателя за данный период

$$t_{вып.пл.} = a_1\ a_{пл},$$

$$t_{вып.пл.} = z_1/z_{пл};\ q_1/q_{пл};\ w_1/w_{пл};\ p_1/p_{пл};$$

– относительная величина динамики ($t_{дин.}$) – величина, выражающая степень изменения явления во времени (характеризует скорость, темп развития), получается соотношением фактической величины показателя за данный период (z_1, q_1, w_1, p_1) к показателю за предыдущий период (z_0, q_0, w_0, p_0)

$$t_{дин.} = a_1/a_0,$$

$$t_{дин.} = z_1/z_0;\ q_1/q_0;\ w_1/w_0;\ p_1/p_0.$$

Относительную динамику можно получить также путем умножения двух предыдущих величин

$$t_{план.\ зад.} \times t_{вып.\ пл.} = t_{дин.},$$

$$q_{пл}/q_o \times q_1/q_{пл} = q_1/q_o.$$

Таким образом, зная любые две относительные величины, можно найти третью относительную.

Пример:

$q_{пл} = 200$ тыс. т $\quad t_{пл.з.} = 200/180 \times 100\ \% = 111\ \%$;

$q_1 = 220$ тыс. т $\quad t_{вып.пл.} = 220/200 \times 100\ \% = 110\ \%$;

$q_0 = 180$ тыс. т $\quad t_{дин.} = 220/180 \times 100\ \% = 122\ \%$ или

$(1{,}11 \times 1{,}1) \times 100\ \% = 122\ \%$.

Так как эти относительные величины взаимосвязаны, то они и объединены в одну группу.

Вторая группа включает относительные величины, которые определяют, если есть целое и в нем составные части:

– относительная величина структуры – соотношение размеров частей и самого целого. Она характеризует структуру, состав совокупности, то есть удельный вес, долю части в общем целом. Обычно выражается в процентах или долях единиц

где $\sum_{i=1}^{n} a_i$ – целое;

– относительная величина координации – соотношение частей целого между собой. Одну из составных частей принимают за базу сравнения и находят отношение к ней всех других частей данного целого

$$t_{коор} = a_i/a_1$$

С ее помощью определяют, сколько единиц данной части целого приходится на 1, 100, 1000 и т. д. другой, принятой за базу сравнения (см. табл. 6).

Таблица 8

Пример относительных величин структуры и координации

Численность	млн чел.	$t_{стр}$	$t_{коор}$
Мужчин	125,0	46,7 %	–
Женщин	142,7	53,3 %	$114\,\frac{\text{чел.}}{\text{100 чел.}}$
Всего	267,7	100 %	–

Третья группа включает относительные величины, характеризующие степень развития какого-либо явления:

– относительная величина интенсивности (степени) – величина, характеризующая степень распространения, развития какого-либо явления в определенной среде.

Получают ее соотношением разноименных величин:

в числителе величина явления, степень распространенную которого изучают, в знаменателе – объем среды, в которой происходит развитие этого явления (плотность населения – чел./км^2). Может быть в виде отвлеченных чисел (коэффициент рождаемости, процент текучести) или в виде именных величин (чел./км^2)

$$t_{инт} = a/q_{среды};$$

– относительная величина уровня экономического развития – показатель, характеризующий размеры производства различных видов продукции на душу населения (можно рассматривать как наиболее распространенную разновидность относительной величины интенсивности). Для ее определения в числителе берут объем производства данного вида продукции, в знаменателе – среднегодовую численность населения за тот же период. В ряде случаев показатель рассчитывают на 1 тыс., 1 млн человек (например, производство сахара составляет 44 кг/чел., картофеля – 133 кг/чел., молока – 332 кг/чел. (при норме 405 кг/чел.), овощей – 103 кг/чел. (при норме 130 кг/чел.).

Четвертая группа – это величины, характеризующие сравнение различных объектов:

– относительная величина сравнения (наглядности) – соотношение одноименных абсолютных величин, относящихся к разным объектам (например, доход Попова в 5 раз больше дохода Петрова)

$$t_{сравн} = a_А/a_Б$$

Пример относительной величины сравнения представлен в табл. 9.

Таблица 9

Относительная величина сравнения

Город	Численность, млн чел.	$t_{сравн}$
Москва	8203	–
Санкт-Петербург	4676	1,8
Киев	2248	3,6

Вернуться в начало

5 СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ

5.1 Средняя величина, ее сущность

Средняя величина – это обобщающая количественная характеристика признака в статистической совокупности, выражающая характерную, типичную величину признака в расчете на единицу совокупности. Величина, для которой исчисляется средняя (так называемый осредненный признак), обозначается Xi. Отдельные варианты этой величины – X_1, X_2, … Xn.

Средняя обозначается – $\overline{x}$.

Средняя величина обладает таким свойством, что в ней погашаются случайные отклонения индивидуальных величин от основного типа. И она выступает как характеристика общих черт явлений, типичных свойств.

Так как средняя величина является обобщающей характеристикой, она не может и не должна сходиться со всеми фактическими индивидуальными значениями, но ее величина лежит в пределах:

$$x_{min} \leq \overline{x} \leq x_{max}.$$

Основным условием правильного применения средней величины является однородность совокупности (в которой составные элементы сходны между собой по существенным для данного исследования признакам, относятся к одному типу). Средняя величина, вычисленная для неоднородной совокупности, то есть такой, в которой объединены качественно различные явления, не имеет смысла. Большое значение имеет и выбор формулы средней, по которой правильно можно ее вычислить. Для правильного выбора формулы средней величины лучше всего использовать среднее исходное соотношение (СИС), то есть логическую формулу средней.

Например, чтобы определить среднюю урожайность (ср. ур.), используют формулу

$$\text{Среднее исходное соотношение} = \frac{\sum_{i=1}^{n} \text{валовой сбор}}{\sum_{i=1}^{n} \text{посевная площадь}}.$$

а) Если в исходной формуле известны и числитель, и знаменатель, то в этом случае используется средняя агрегатная, то есть

$$\underset{(\overline{x})}{\text{ср.ур.}} = \frac{\sum_{i=1}^{n} \text{вал.сб.}}{\sum_{i=1}^{n} \text{пос.пл.}} = \overline{x}_{\text{агр.}} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} f_i},$$

где $\overline{x}_{\text{агр.}}$ – средняя агрегатная.

б) Если в исходной формуле неизвестен числитель (валовой сбор), то его выражают на основе известных значений

$$\underset{(\overline{x})}{\overline{\text{ср.ур.}}} = \frac{\sum_{i=1}^{n} \text{вал.сб.}}{\sum_{i=1}^{n} \text{пос.пл.}} = \frac{\sum_{i=1}^{n} \text{ур.сед.} \times \text{пос.пл.}}{\sum_{i=1}^{n} \text{пос.пл.}} = \overline{x}_{\text{арм.}} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i},$$

Валовой сбор = ур. с ед. × пос. пл.,

где $\overline{x_{\text{арм.}}}$ – средняя арифметическая.

в) Если в исходной формуле неизвестен знаменатель (посевная площадь), то его выражают на основе известных значений

$$\underset{(\bar{x})}{\text{ср.ур.}} = \frac{\sum_{i=1}^{n} \text{вал.сб.}}{\sum_{i=1}^{n} \text{пос.пл.}} = \frac{\sum_{i=1}^{n} \text{вал.сб.}}{\sum_{i=1}^{n} \text{вал.сб./ур.сед.}} = \overline{x}_{\text{гарм.}} = \frac{\sum_{i=1}^{n} W_i}{\sum_{i=1}^{n} \frac{W_i}{x_i}},$$

$$\text{Посевная площадь} = \text{вал. сб./ур. с ед.},$$

где $\overline{x_{\text{гарм.}}}$ – средняя гармоническая.

Вернуться в начало

5.2 Виды средних величин

Из всего многообразия средних величин наиболее часто в экономической статистике применяются: средняя агрегатная, средняя арифметическая, средняя гармоническая, средняя геометрическая, средняя хронологическая. Применение той или иной формы зависит от содержания осредняемого признака и конкретных данных, по которым ее необходимо рассчитать.

1. **Средняя агрегатная** вычисляется по формуле

$$\overline{x}_{\text{агр.}} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} f_i},$$

где w_i – объемный показатель;

f_i – вес признака, частота, численность.

Формула агрегатной средней используется, если известны значения числителя и знаменателя в логической формуле (СИС). Если известны фонд оплаты труда (ФОТ) и численность в отдельных цехах (участках), то средняя заработная плата по предприятию определяется по формуле (табл. 10)

$$\overline{\text{Ср.зп}} = \frac{\sum_{i=1}^{n} \text{ФОТ}}{\sum_{i=1}^{n} \text{численность}},$$

Таблица 10

Фонд оплаты труда ООО «Вымпел»

№ цеха	ФОТ, тыс. руб.	Численность, чел.
1	5300	550
2	4700	450

$$\overline{\text{Ср.зп}} = \frac{\sum_{i=1}^{n} \text{ФОТ}}{\sum_{i=1}^{n} \text{численность}} = \overline{x}_{\text{агр.}} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} f_i} = \frac{5300 + 4700}{550 + 450} = 10\,\frac{\text{тыс.руб.}}{\text{чел.}}$$

2. **Средняя арифметическая и ее свойства.** Средняя арифметическая – одна из наиболее распространенных форм средней величины. Средняя арифметическая используется, если даны отдельные значения признака или в логической формуле расчета показателя неизвестен числитель. Она применяется в виде простой и взвешенной средней арифметической.

Формула ***простой***

$$\overline{x}_{\text{арм.пр.}} = \frac{\sum_{i=1}^{n} x_i}{n},$$

где x_i – отдельные значения признака;

n – число единиц совокупности, число значений признака.

Часто в совокупности отдельные варианты могут принимать одинаковые значения, которые можно объединить в группы, подсчитав их численность, поэтому в этом случае осуществляется переход к средней взвешенной. Ее можно определить как частное от деления суммы произведения вариантов и их численностей (частот) – $\Sigma x \times f$ на сумму численностей (частот) – Σf.

$$\bar{x}_{\text{арм.взв.}} = \frac{\sum_{i=1}^{n} x_i \times f_i}{\sum_{i=1}^{n} f_i},$$

где x_i – отдельные значения признака, значения вариантов (показателей);

f_i – численность (частота, вес) каждого варианта (группы).

Основой для вычисления простой арифметической служат первичные записи результатов наблюдения, а для взвешенной – обработанный материал, сгруппированные данные по количественному признаку.

Простая средняя вычисляется в тех случаях, когда веса отсутствуют, или их очень трудно определить, или если численность отдельных групп (вариантов) не слишком отличается. В других случаях ее применение приводит к очень грубым ошибкам. Простая средняя соответствует простой совокупности объектов, в которой нет групп.

Средняя взвешенная отражает сложное строение совокупности, в ней учитывается удельный вес отдельных групп в совокупности.

Средняя арифметическая имеет ряд свойств, которые находят практическое применение:

1-е свойство. От увеличения (уменьшения) всех вариантов осредняемой величины в К раз их средняя величина соответственно увеличивается (уменьшается) в К раз.

$$Z_i = K \times X_i$$

$$\bar{Z} = \frac{\sum_{i=1}^{n} Z_i \times f_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} K \times (x_i \times f_i)}{\sum_{i=1}^{n} f_i} = \frac{K \times \sum_{i=1}^{n} (x_i \times f_i)}{\sum_{i=1}^{n} f_i} = K \times \bar{x}$$

2-е свойство. От уменьшения (увеличения) веса каждого варианта в К раз средняя не меняется.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i \times f_i / K}{\sum_{i=1}^{n} f_i / K} = \frac{1/K \sum_{i=1}^{n} x_i \times f_i}{1/K \sum_{i=1}^{n} f_i} = \bar{x}$$

3-е свойство. Величина средней зависит не от абсолютных значений весов отдельных вариантов, а от пропорций между ними.

Отношения отдельных частот $f_1, f_2, \dots f_n$ к Σf_i представляют долю отдельных вариантов в совокупности

$$d_i = \frac{f_i}{\sum_{i=1}^{n} f_i},$$

где d_i – удельный вес, часть, доля.

Поэтому вместо абсолютного значения f_i можно брать веса вариантов, выраженные в долях или процентах, тогда

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i \times x_i}{\sum_{i=1}^{n} f_i} = x_1 \times d_1 + \dots + x_i \times d_i = \sum_{i=1}^{n} x_i \times d_i .$$

4-е свойство. Если уменьшить (увеличить) все варианты осредненного признака на постоянное число (А), то средняя уменьшается (увеличивается) на то же число.

$$Z_i = x_i - A.$$

$$\bar{Z} = \frac{\sum_{i=1}^{n} Z_i \times f_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} (x_i - A) \times f_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} x_i \times f_i}{\sum_{i=1}^{n} f_i} - \frac{\sum_{i=1}^{n} A \times f_i}{\sum_{i=1}^{n} f_i} = \bar{x} - A$$

5-е свойство. Средняя, умноженная на численность всей совокупности, равна сумме произведения каждого варианта на ее численность.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i \times f_i}{\sum_{i=1}^{n} f_i}$$

$$\bar{x} \times \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} x_i \times f_i$$

6-е свойство. Сумма отклонений индивидуальных значений от их средней арифметической равна нулю.

$$\sum_{i=1}^{n} (x_i - \bar{x}) \times f_i = 0,$$

$$\sum_{i=1}^{n} x_i \times f_i - \bar{x} \times \sum_{i=1}^{n} f_i = 0,$$

так как $\sum_{i=1}^{n} x_i \times f_i = \bar{x} \times \sum_{i=1}^{n} f_i$ (свойство 5).

То есть если взять отклонения каждого варианта от средней величины и взвесить по численности, а затем сложить, то получим ноль.

7-е свойство. Сумма квадратов отклонений индивидуальных значений от их средней арифметической меньше, чем сумма квадратов отклонений индивидуальных значений от любой другой величины:

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 < \sum_{i=1}^{n} (x_i - \mathrm{A})^2.$$

Использование свойств средней арифметической позволяет значительно упростить ее вычисления. Упрощенный способ расчета средней арифметической, называемый **способом моментов** (первого порядка), состоит в следующем:

– уменьшим все значения вариантов на величину А, в качестве которой обычно принимается наиболее часто встречающееся значение признака: x_i – А;

– все полученные отклонения разделим на какое-нибудь общее кратное (обычно величину интервала) число, то же и для весов, то есть

$$Z_i = \frac{x_i - A}{K}; \qquad f_i' = \frac{f_i}{m};$$

– рассчитаем среднюю арифметическую условных значений (Zi)

$$\bar{Z} = \frac{\sum_{i=1}^{n} z_i \times f_i'}{\sum f_i'} = \frac{\frac{\sum_{i=1}^{n}(x_i - A)}{K} \times \frac{f_i}{m}}{\frac{\sum_{i=1}^{n} f_i}{m}} = \frac{1/K \times 1/m \sum_{i=1}^{n}(x_i - A) \times f_i}{\frac{1 \sum_{i=1}^{n} f_i}{m}} =$$

$$= \frac{\sum_{i=1}^{n} x_i \times f_i - A \sum_{i=1}^{n} f_i}{K \sum_{i=1}^{n} f_i} = \frac{\bar{x} - A}{K},$$

$$f_i' = f_i/m;$$

– на основании свойств средней арифметической, для того чтобы ее общее значение не изменялось, нужно условную среднюю Z увеличить в K раз и на A, то есть

$$\bar{x} = \bar{Z} \times K + A.$$

3. **Средняя гармоническая** используется в тех случаях, когда известны обратные значения показателя либо в логической формуле расчета показателя неизвестен знаменатель. Средняя гармоническая вычисляется из обратных значений признака и может также быть:

а) простой

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

где $1/x_i$ – обратные значения вариантов признака;

n – число вариантов;

б) взвешенной

$$\overline{x}_{\text{гарм.взв.}} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} \frac{w_i}{x_i}},$$

где $w_i = x_i \times f_i$ – объемный показатель.

Применяется средняя гармоническая в тех случаях, когда непосредственные данные о весах отсутствуют, а известны варианты осредняемого признака (x_i) и произведения значений вариантов на количество единиц, обладающих данным его значением ($w_i = x_i \times f_i$).

Пример. Рассчитаем среднюю гармоническую величину по данным о валовом сборе зерна, представленным в табл. 11.

Таблица 11

Валовой сбор зерна

Показатель	Валовой сбор (ВС)	Урожайность (ц/га) – х
1-й участок	110	22
2-й участок	650	26
3-й участок	600	30

$$\overline{\text{урож.}\ (\text{СИС})} = \frac{\sum_{i=1}^{n} BC}{\sum_{i=1}^{n} \text{Пл.}} = \frac{\sum_{i=1}^{n} BC}{\sum_{i=1}^{n} BC/\text{урож.}} = \overline{x}_{\text{гарм.}},$$

где площадь определяется как

$$\text{Площадь} = \frac{\text{Вал. сбор (ВС)}}{\text{Урожайность}}.$$

Подставляем в исходную формулу

$$\overline{x}_{\text{гарм.}} = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} \frac{w_i}{x_i}} = \frac{110 + 650 + 600}{\frac{110}{22} + \frac{650}{26} + \frac{600}{30}} = 27{,}2 \text{ ц/га}.$$

4. **Средняя геометрическая** используется в статистике в основном для вычисления темпов роста показателей динамиче-

ского ряда. В зависимости от имеющихся исходных данных может использоваться формула двух видов:

1. Если расчет ведется исходя из коэффициентов (темпов) роста, найденных по отношению к предыдущему периоду (цепных), то

$$\bar{x} = \sqrt[n]{Тр_{1цеп} \times Тр_{2цеп} \times ... Тр_{iцеп}} = \sqrt[n]{\Pi(Тр_{iцеп})},$$

где $Тр_i$ – цепные темпы роста;

n – число значений признака.

2. Если в распоряжении имеются абсолютные уровни ряда или базисный темп роста, то есть за весь период, то

$$\bar{x} = \sqrt[n]{\frac{Y_n}{Y_o}} = \sqrt[n]{Тр_{баз}},$$

где Y_o, Y_n – начальные и конечные абсолютные значения уровней ряда;

$Тр_{баз}$ – базисный темп роста за данный период.

Например, в 2009 г. валовой продукт (ВП) составил 120 млрд руб.; в 2004 г. ВП – 100 млрд руб.

$$Тр = \sqrt[5]{\frac{120}{100}} = \sqrt[5]{1,2}$$

5. **Средняя хронологическая** используется для вычисления средней величины из уровней моментного ряда динамики и может быть:

а) простой

$$\bar{x}_{хр.прост.} = \frac{1/2x_i + x_2 + ... + x_{n-1} + 1/2x_n}{n-1},$$

где x_i – абсолютные значения уровней для моментного ряда (на определенную дату, момент);

(n – 1) – продолжительность периода (например, если за год, то n – 1 = 12 месяцев);

б) взвешенной

$$\bar{x}_{хр.взв.} = \frac{(x_1+x_2)t_1+(x_2+x_3)t_2+...+(x_{n-1}+x_n)t_{n-1}}{2(t_i+t_2+t_3+...+t_{n-1})},$$

где t_i – период времени, отделяющий один уровень ряда от другого.

Средняя взвешенная используется, если интервалы времени между уровнями неравны и известно время, в течение которого сохранялось каждое значение уровня ряда.

Вернуться в начало

5.3 Структурные средние величины

В экономических расчетах кроме алгебраических средних используются еще особые разновидности средних величин, которые вытекают из характеристики статистических рядов, их условно можно назвать структурными средними: Мо, Ме.

Под **модой** (Мо) понимается вариант, который чаще всего встречается в данном статистическом ряду.

Под **медианой** (Ме) понимается значение варианта, расположенного в его середине, то есть такое, которое делит ряд по численности на две равные части.

Способ определения Мо и Ме зависит от вида ряда:

1. Если ряд представляет отдельные, дискретные значения, то структурные средние определяются исходя из понятия, то есть Мо = Xf_{max} (то есть значение варианта, имеющего наибольшую частоту).

Прежде чем найти значение Ме, необходимо найти ее номер, причем:

а) если всем единицам ряда придать порядковый номер, то номер медианы, в ряду с нечетным числом вариантов, определяется как

$$№_{Me} = (n + 1) / 2,$$

(например, n = 51, то $№_{Me}$ = (51 + 1) / 2 = 26, то есть 26-й вариант в ряду). Тогда Ме = $X_{№Me26}$, то есть Ме = x_{26} (Ме – это вариант, стоящий под данным номером);

б) если вариант – четное число, то медиану определяют как среднюю из двух центральных вариантов, порядковые номера которых $\frac{n}{2}$ и $\frac{n}{2}+1$.

Например, если n = 50, то:

$№_{Me1}$ = 50/2 = 25;

$№_{Me2}$ = 50/2 +1 = 26,

а $Me = \frac{X_{№_{Me1}} + X_{№_{Me2}}}{2}$, то есть $Me = \frac{X_{25} + X_{26}}{2}$.

Пример. Определим моду и медиану по данным, представленным в табл. 12.

Таблица 12

Численность и заработная плата работников ООО «Весна»

Показатели	Значения						Всего
Среднегодовая заработная плата, тыс. руб.	203	214	232	255	264	276	
Численность, чел.	1	1	5	3	2	2	14

Мо = 232 тыс. руб. (так как fmax = 5), Σfi = 14 человек.

$№_{Me} = \frac{14}{2}$ и $\frac{14}{2}+1$, то есть 7 и 8 человек.

X_7 = 232 тыс. руб. (заработная плата 7-го человека);

X_8 = 255 тыс. руб. (заработная плата 8-го человека),

$$M_e = \frac{x_7 + x_8}{2} = \frac{232 + 255}{2} = 243{,}5 \text{ тыс. руб.}$$

2. Если динамический ряд представлен в виде интервалов (то есть не дискретный, а интервальный), то для вычисления M_o и M_e прибегают к следующим формулам

$$M_o = x_o + h\frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})},$$

$$M_o = x_o + h\frac{\frac{\sum f_i}{2} - S_{m-1}}{f_m},$$

где x_o – нижняя граница модального, медианного интервалов соответственно;

h – шаг, величина интервала;

f_{m-1} и f_{m+1} – частота предшествующего и последующего за модальным интервалов;

f_m – частота модального, медианного интервала (соответственно);

S_{m-1} – сумма накопленных (куммулятивных) частот в интервалах, предшествующих медианному.

Прежде чем рассчитывать M_o и M_e, определяют модальный и медианный интервал.

Модальный интервал – это тот, где наибольшая частота; **медианный интервал** – это тот, где накопленная частота превышает половину общей численности совокупности.

M_o и M_e в отличие от алгебраических средних, являющихся в значительной мере абстрактными характеристиками, выступают как конкретные величины, совпадающие с вполне определенными вариантами этого ряда. Поэтому они имеют большое практическое применение (например, чтобы определить наиболее ходовой размер обуви (одежды), средняя арифметическая, дающая абстрактную величину, не подходит и используется M_o).

Вернуться в начало

5.4 Вариация и ее показатели

Вариацией признака называется изменение его у единиц совокупности. Элементы совокупности характеризуются различными количественными значениями признака, их изменение порождается разнообразием условий, окружающих факторов, воздействующих на элементы (например, вариация оценок на экзамене порождается различными способностями студентов, затратами на подготовку, социально-бытовыми условиями и т. д.).

Измерение вариации позволяет определить степень воздействия на данный признак других признаков. Вариация может быть в пространстве и во времени (например, изменяется урожайность по районам или в одном районе по годам).

Показатели вариации относят к числу обобщающих показателей, они измеряют вариацию в совокупности явлений.

Значение показателей вариации состоит в следующем:

– они дополняют среднюю величину, за которой скрываются индивидуальные значения;

– характеризуют степень однородности статистической совокупности по данному признаку;

– характеризуют границы вариации признака;

– соотношение показателей вариации характеризует взаимосвязь между признаками.

В статистике чаще всего используются следующие показатели вариации:

1. Размах вариации – R. __.
2. Среднее абсолютное (линейное) отклонение – d.
3. Среднее квадратичное отклонение – σ.
4. Дисперсия – σ^2.
5. Коэффициенты вариации – V.

Размах вариации (R) – это разность между max и min значениями признака, он характеризует предел изменения признака (имеет ту же размерность, что и сам признак).

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

Среднее абсолютное (линейное) отклонение ($\bar{d}$) – это средняя арифметическая из абсолютных отклоненных значений признака всех единиц совокупности (так как сумма индивидуальных отклонений в силу свойств средней равна нулю, то берут абсолютную величину):

– простая

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

взвешенная

$$\bar{d} = \frac{\sum_{i=1}^{n} |d_i| \times f_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}| \times f_i}{\sum_{i=1}^{n} f_i},$$

где f_i – частота, вес отдельных вариантов.

Среднее абсолютное отклонение, так же как и размах, – число именованное, размерность его соответствует размерности признака.

Среднеквадратическое отклонение (σ) является характеристикой рассеивания, имеет ту же размерность, что и признак, и представляет собой корень квадратный из среднего квадрата отклонений значения признака от их средней величины.

Простая

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n}}, \quad d_i = x_i - \bar{x}.$$

Взвешенная

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 \times f_i}{\sum_{i=1}^{n} f_i}},$$

или

$$\sigma = \sqrt{\overline{x^2} - \bar{x}^2},$$

где $\overline{x^2}$ – средняя величина квадрата значений признака (то есть средняя из квадратов);

$\bar{x}^2$ – квадрат средней величины признака.

При его определении принимаются в расчет все отклонения значений признака (так как $\sum_{i=1}^{n}(x_i - \bar{x}) = 0,$ поэтому возводят в квадрат).

Между средним абсолютным и средним квадратическим отклонением существует следующее примерное соотношение: $\sigma \approx 1{,}25d$ (если фактическое распределение близко к нормальному). Чем меньше величина среднего квадратического отклонения, тем однороднее совокупность.

Дисперсия (σ^2) вычисляется для всей статистической совокупности в целом как средний квадрат отклонений значения признака от общей средней, измеряет степень колеблемости признака, его вариацию, порождаемую всей совокупностью действующих на него факторов, и определяется по формулам

– простая

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n} d_i^2}{n};$$

– взвешенная

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \times f_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} d_i^2 \times f_i}{\sum_{i=1}^{n} f_i};$$

$$\sigma = \overline{x^2} - \bar{x}^2.$$

Дисперсия имеет ряд свойств, которые находят практическое применение:

1-е свойство. Уменьшение (увеличение) всех значений признака на одну и ту же величину (А) не меняет величины σ^2 (так как разность между «новым» значением признака и «новой» средней остается без изменения).

$$x_i - A = Z_i, \text{ тогда } \bar{Z} = \overline{(x_i - A)} = \bar{x} - A.$$

Отсюда

$$\sigma_z^2 = \frac{\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2}{n} =$$

$$= \frac{\sum_{i=1}^{n}\left[(x_i - A) - (\bar{x} - A)\right]^2}{n} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \sigma^2 x,$$

то есть

$$\sigma^2{}_z = \sigma^2{}_x.$$

2-е свойство. Уменьшение (увеличение) всех значений признака в К раз уменьшает (увеличивает) дисперсию в $К^2$ раз.

$$\frac{x_i}{K} = Z_i,$$

тогда

$$\sigma_Z{}^2 = \frac{\left[\sum_{i=1}^{n} \frac{1}{K} \times x_i - \frac{1}{K} \times \bar{x}\right]^2}{n}.$$

Согласно свойству средней

$$\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{K^2 \times n} = \sigma^2{}_z,$$

то есть

$$\sigma^2{}_x = K^2 \times \sigma^2{}_z.$$

На основании данных свойств разрабатывается упрощенный метод вычисления дисперсии с помощью способа моментов (2-го порядка):

а) исходные значения вариант признака Xi заменяют условными

$$Z_i = \frac{x_i - A}{K},$$

где А – обычно значение признака, которое чаще всего встречается в совокупности,

К – величина интервала, кратное число;

б) определяется дисперсия условной величины (Z)

$$\sigma_z^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i - A}{K} - \frac{\bar{x} - A}{K} \right)^2 = \underbrace{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{K} \left[(x - A) - (\bar{x} - A) \right] \right)}_{\text{исходя из свойств дисперсии}} = \frac{\sigma^2 x}{K^2};$$

в) определяется дисперсия исходного признака

$$\sigma_x{}^2 = K^2 \times \sigma_z{}^2.$$

Правило сложений дисперсий.

Общая дисперсия измеряет вариацию результативного признака по всей совокупности под влиянием всех факторов, обусловивших эту вариации.

Межгрупповая дисперсия (σ^2) характеризует систематическую вариацию под воздействием признака – фактора, положенного в основание группировки. Она равны среднему квадрату отклонений групповых (частных) средних от общей средней для всей совокупности:

$$\delta^2 = \frac{\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2 * f_i}{\sum_{i=1}^{n} f_i}$$

Внутригрупповая дисперсия применяется к отдельной группе и обозначается как и общая дисперсия (σ^2), но с индексом i, который подчеркивает, что расчет выполняется для отдельной i-ой группы.

Внутри групповая дисперсия отражает случайную вариацию, то есть ту ее часть, которая обусловлена влиянием прочих (неучтенных) факторов, отличных от основания группировки.

Существует формула, связывающая общую дисперсию, межгрупповую дисперсию и среднюю по внутригрупповым дисперсиям

$$\sigma^2 = \delta^2 + \sigma_t^2.$$

Это означает, что общая дисперсия равна сумме межгрупповой дисперсии и средней по внутригрупповым дисперсиям.

Правило сложения дисперсий показывает, что чем больше доля межгрупповой дисперсии в общей дисперсии, тем сильнее влияние группировочного признака на изучаемый результативный признак.

Относительные показатели вариации - являются относительной мерой вариации и представляют собой отношение именованного показателя вариации (R, d, σ) и средней величины ($\bar{x}$, Mo или Me). Таким образом, в принципе возможен расчет девяти коэффициентов вариации. Они дают представление о степени однородности совокупности. Чем меньше их величины, тем меньше варианты признака отличаются один от другого по величине, тем, следовательно, однороднее совокупность. Будучи относительной величиной, абстрагируют различия абсолютных величин вариации различных признаков и дают возможность сравнения. То есть с их помощью можно сравнивать размеры вариации одного признака в нескольких совокупностях. Чаще на практике используются следующие:

а) коэффициент осцилляции:

$$K_R = \frac{R}{\bar{x}} \times 100\%,$$

б) относительное линейное отклонение

$$K_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \times 100\%,$$

в) коэффициент вариации

$$V = \frac{\sigma}{\bar{x}} \times 100\%$$

При этом из них чаще всего используется коэффициент вариации (V).

Коэффициент вариации используют не только для сравнительной оценки вариации, но для характеристики однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33 % (для распределений близких к нормальному).

Вернуться в начало

6 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНИМИКИ СОЦИАЛЬНО-ЭКОМИЧЕСКИХ ЯВЛЕНИЙ

6.1 Определение и виды рядов динамики

Любое явление развивается во времени. Для характеристики изменения (развития) явления строятся хронологические ряды, существенное значение в которых имеет последовательность показателя (то есть временная его хронология). Ряды, где показатели располагаются во временной хронологии, называют **рядами динамики, или временными рядами**. Таким образом, рядами динамики называется временная последовательность значений статистических показателей. То есть ряды динамики представляют собой ряд численных значений определенного показателя в последовательные моменты или периоды времени.

Любой ряд динамики состоит из двух элементов (табл. 13):

– моментов (дат) или периодов времени, к которым относятся статистические данные;

– самих данных, то есть числовых значений показателя, составляющих динамический ряд и называемых уровнями (yi).

Таблица 13

Численность работающих на предприятии, чел.

Момент (дата)	На 1 января	На 1 февраля	На 1 марта
Уровень	100	110	108

Оба элемента – время и уровень – называются **членами динамического ряда.**

Уровни ряда обладают следующими особенностями:

– уровень последующего времени зависит от уровня, достигнутого в предыдущий период;

– чем больше интервал времени между событиями, тем больше отличаются их количественные и качественные состояния.

В зависимости от группировки элементов по различным признакам ряды динамики делят на виды. Классификацию обычно осуществляют по времени, по полноте охвата и по способу выражения уровней ряда.

1. По времени ряды динамики делят на моментные и интервальные.

Моментным называется ряд, уровни которого характеризуют величину явления по состоянию на определенный момент времени, определенную дату (например, численность либо стоимость ОПФ на первое число каждого квартала или месяца).

Интервальным называется такой ряд, уровни которого характеризуют величину изучаемого показателя, полученную в итоге за определенный период времени (например, объем производства за год, месяц, квартал).

В отличие от моментного ряда, суммирование уровней которого не имеет смысла, в интервальном можно суммировать уровни следующих друг за другом периодов и сумму можно рассматривать как итог (уровень) за более длительный период времени (например, объем производства за пять лет). Кроме того, можно дробить каждый из уровней (объем производства за каждый месяц вместо квартала).

2. По полноте охвата времени, отражаемого в рядах динамики, их можно разделить на полные и неполные.

В полных рядах даты или периоды времени следуют друг за другом с равными интервалами. В неполных – равный интервал не соблюдается.

3. По способу выражения уровней ряда могут быть рядами абсолютных, средних и относительных величин, например:

– производство электроэнергии в РФ по годам;

– динамика средней месячной заработной платы рабочего;

– темп роста национального дохода по годам.

Вернуться в начало

6.2 Правила построения динамических рядов

При формировании рядов динамики нужно соблюдать важнейшее требование сопоставимости всех уровней ряда, которое выражается следующим образом:

– сопоставимость территории, к которой относятся уровни ряда – изменение границ области, района, страны приводит к различию, несравнимости статистических показателей;

– сопоставимость уровней рядов динамики по кругу охватываемых объектов – несопоставимость может возникнуть при переходе объекта из одного подчинения в другое;

– сопоставимость по критическому моменту регистрации для явлений с сезонным характером уровней (например, численность скота летом больше, чем зимой);

– несопоставимость из-за различия единиц измерения – при возможности измерения в различных единицах уровни ряда нужно выражать в одних;

– сопоставимость по методике учета и расчета показателей (например, производительность рабочих и работающих);

– сопоставимость в понимании единиц совокупности, характеризуемой рядом динамики, так как определить единицы можно по-разному.

Вернуться в начало

6.3 Статистические характеристики (показатели) ряда динамики

Для характеристики изменения явления во времени находят статистические показатели.

Большинство статистических показателей основано на абсолютном и относительном сравнении уровней ряда. К таким показателям относятся:

1) абсолютный прирост (Δy);
2) темп роста (Тр);
3) темп прироста (Тпр);
4) абсолютное содержание 1 % прироста (α).

Все характеристики могут определяться двумя методами: цепным и базисным. При цепном методе каждый данный уровень сравнивается с предыдущим. При базисном методе каждый данный уровень сравнивается с одним и тем же принятым за базу сравнения.

1. **Абсолютный прирост** показывает, на сколько абсолютных единиц изменяется данный уровень по сравнению:

а) с предыдущим уровнем при цепном методе

$$\Delta y_{i\ \text{цеп}} = y_i - y_{i-1},$$

то есть

$$y_1 - y_0;\ y_2 - y_1;\ y_3 - y_2;\ \dots\ ;\ y_n - y_{n-1};$$

б) с начальным уровнем при базисном методе

$$\Delta y_{i\ \text{баз}} = y_i - y_o,$$

то есть

$$y_1 - y_0;\ y_2 - y_0;\ y_3 - y_0;\ \dots\ ;\ y_n - y_0,$$

где Δy_i – абсолютный прирост;

y_i – текущий уровень;

y_{i-1} – предыдущий уровень;

y_o – начальный уровень.

За весь период, описываемый рядом, абсолютный прирост выразится как разность между последними и начальными уровнями:

$$\Delta у_{\text{за весь период}} = \Delta у_{\text{баз}} = \sum_{i=1}^{n} (y_i - y_{i\text{-}1}) = (y_1 - y_0) +$$
$$+ (y_2 - y_1) + \ldots + (y_n - y_{n\text{-}1}) = y_n - y_{\text{o}}.$$

Таким образом, сумма цепных абсолютных приростов дает соответствующий базисный абсолютный прирост.

Абсолютный прирост может иметь знак «+» или «–» и показывает, насколько уровень текущего периода выше или ниже предыдущего (базисного), имеет те же единицы измерения, что и уровни ряда (табл. 14).

Таблица 14

Потребление мяса на душу населения

Показатели	Год				
	Базисный ($у_0$)	Первый ($у_1$)	Второй ($у_2$)	Третий ($у_3$)	Четвертый ($у_4$)
Годовое потребление мяса, кг/чел.	48	62	62	64	65
Абсолютный прирост (цепной) по годам, кг/чел.	–	+14	0	+2	+1
за 3 года	ΔY = +3 кг/чел.				
Абсолютный прирост (базисный) по годам, кг/чел.	–	+14	+14	+16	+17
за 4 года	ΔY = +17 кг/чел.				

2. **Темп роста** – это отношение данного уровня к предыдущему или какому-либо другому, принятому за базу сравнения, показывает, во сколько раз данный уровень изменяется по сравнению с предыдущим (базисным).

При этом сравниваемый уровень называется **текущим**, а уровень, с которым производится сравнение, – **базисным**.

Если производится сравнение каждого уровня с предыдущим, то получают ***цепные*** темпы (коэффициенты) роста ($Тр_{\text{цеп.}}$)

$$\frac{у_1}{у_0};\ \frac{у_2}{у_1};\ \frac{у_3}{у_2};\ \ldots\ \frac{у_n}{у_{n-1}}.$$

Если каждый уровень сравнивается с начальным уровнем или каким-либо другим, принятым за базу, то получают ***базисные*** (коэффициенты) темпы роста ($Тр_{баз}$)

$$\frac{y_1}{y_0};\ \frac{y_2}{y_0};\ \dots\ \frac{y_n}{y_0}.$$

Таким образом, темпы роста могут вычисляться двумя способами

$$Тр_{цепной} = \frac{y_i}{y_{i-1}};\ \frac{62}{48};\ \frac{62}{62};\ \frac{64}{62};\ \frac{65}{64};$$

$$Тр_{базисный} = \frac{y_i}{y_0};\ \frac{62}{48};\ \frac{62}{48};\ \frac{64}{48};\ \frac{65}{48}.$$

Произведение цепных темпов роста дает соответствующий базисный темп роста. Таким образом, темп роста за весь период времени, представленный рядом динамики, будет равен конечному базисному темпу роста, то есть произведению соответствующих цепных рядов

$$\mathbf{Т_р = \frac{У_n}{У_0} = \frac{У_1}{У_0} \times \frac{У_2}{У_1} \times \frac{У_3}{У_2} \times \frac{У_n}{У_{n-1}}}.$$

Темп роста можно выражать в коэффициентах или в процентах

$$Тр\ (\%) = Тр \times 100\ \%$$

Если Тр > 1 (100 %), то, следовательно, идет увеличение уровней ряда, если Тр < 1 (100 %) – снижение.

3. **Темп прироста** – это отношение абсолютного прироста к предыдущему или базисному уровню

$$Тпр_{цеп} = \frac{\Delta y_i}{y_{i-1}} = \frac{y_i - y_{i-1}}{y_{i-1}} = Тр_{цеп} - 1;$$

$$Тпр_{баз} = \frac{\Delta y_i}{y_0} = \frac{y_i - y_0}{y_0} = Тр_{баз} - 1.$$

Темп прироста показывает, на сколько процентов увеличивается (+) или уменьшается (–) текущий уровень по сравнению с базисным (предыдущим), принятым за 100 %.

Так как абсолютный прирост за весь период $y = y_n - y_o$, то темп прироста за весь период

$$Тпр_{баз} = \frac{y_n - y_0}{y_0} = \frac{y_n}{y_0} - 1 = \underset{\text{за весь период}}{Тр - 1}$$

или

$$Тпр\ (\%) = Тр\ (\%) - 100\ \%.$$

4. **Абсолютное содержание (значение) 1 % прироста** показывает, какая абсолютная величина скрывается за относительным показателем 1 % прироста, и представляет собой отношение абсолютного прироста к темпу прироста в %

$$\alpha_{цеп} = \frac{\Delta y_{i\,цеп.}}{Т_{пр.\ цеп.\ (\%)}} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \times 100\%} = \frac{y_{i-1}}{100\%},$$

$$\alpha_{баз} = \frac{\Delta y_{i\,баз.}}{Т_{пр.\ баз.\ (\%)}} = \frac{y_i - y_0}{\frac{y_i - y_0}{y_0} \times 100\%} = \frac{y_0}{100\%}.$$

Следовательно, абсолютное значение 1 % прироста – это 0,01 от величины предыдущего или базисного уровня.

Вернуться в начало

6.4 Средние показатели ряда динамики

Статистические показатели, рассчитанные по уровням ряда динамики, изменяются во времени, варьируют по годам. Это требует их обобщения и расчета средних показателей, которые характеризуют общее развитие явления за данный период. К таким обобщающим характеристикам динамического ряда относят:

1) средний уровень ряда ($\bar{y}$);

2) средний абсолютный прирост ($\Delta\overline{y}$);

3) средний темп роста ($\overline{T_p}$);

4) средний темп прироста ($\overline{T_{пр}}$).

1. **Средний уровень ряда** – временная или хронологическая средняя – рассчитывается как средняя величина из уровней ряда. Вычисляется по-разному для интервальных и моментных рядов.

Чтобы найти средний уровень (среднюю величину показателя) интервального ряда, достаточно сумму уровней ряда разделить на число периодов, к которым они относятся

$$\overline{y_{арм}} = \frac{y_0 + y_1 + y_2 + \cdots + y_n}{n+1} = \frac{\sum_{i=0}^{n} y_i}{n+1},$$

где n+1 – число уровней ряда.

Используя данные табл. 16, получим

$$\overline{y_{арм}} = \frac{48+62+62+64+65}{5} = 60{,}2 \text{ кг / год.}$$

Средний уровень моментного ряда рассчитывается на основе средней хронологической

$$\overline{y_{хр}} = \frac{\frac{1}{2}y_0 + y_1 + y_2 + \cdots + y_{n-1} + \frac{1}{2}y_n}{n},$$

Например, n показывает период, за который определяется вид средней величины: так, средняя за квартал n = 3 месяца, за год n = 12 месяцев. Следовательно, для расчета среднего уровня за квартал нужно иметь 4 значения, за год – 13.

Для неполных моментных рядов применяется взвешивание суммы каждой смежной пары уровней по продолжительности периода между ними, то есть средняя хронологическая взвешенная:

$$\overline{y_{хр}} = \frac{(y_1 + y_2) \times t_1 + (y_2 + y_3) \times t_2 + \ldots (y_{n-1} + y_n) \times t_{n-1}}{2 \times (t_1 + t_2 + \ldots + t_{n-1})},$$

где t_i – время между моментом регистрации y_1 и y_2, y_2 и y_3 и т. п.

2. **Средний абсолютный прирост** есть средняя из абсолютных приростов за промежутки времени данного периода.

$$\bar{\Delta y} = \frac{\Delta y_1 + \Delta y_2 + ... + \Delta y_{n-1}}{n} = \frac{\sum_{i=1}^{n} \Delta y_{i\ цеп}}{n} =$$

$$= \frac{(y_1 - y_0) + (y_2 - y_1) + ... + (y_n - y_{n-1})}{n} = \frac{y_n - y_0}{n} = \frac{\Delta y_{баз}\ (\text{за весь период})}{n},$$

где n – число абсолютных приростов (тогда уровней соответственно n + 1), то есть

$$\bar{\Delta y} = \frac{y_n - y_0}{n}.$$

Средний абсолютный прирост показывает, на сколько абсолютных единиц в среднем за период изменяются уровни ряда. Может быть со знаком «+» – прирост или «–» – снижение.

Используя данные табл. 16 (за период второй – четвертый год), получим

$$\Delta\bar{y} = \frac{65\ кг - 62\ кг}{3\ года} = 1\ кг/год,$$

(четыре уровня и три абсолютных прироста).

$$\Delta\bar{y} = \frac{65\ кг - 48\ кг}{4\ года} = 4{,}25\ кг/год$$

(пять уровней и четыре абсолютных прироста).

3. **Средний темп роста** ($\overline{Тр}$) является обобщающим показателем темпов роста уровней ряда динамики и показывает, как в среднем изменялись уровни ряда динамики на протяжении исследуемого периода. Для его расчета всегда используется формула средней геометрической:

– если имеются данные об абсолютных уровнях ряда или базисные темпы роста, то

$$\overline{Тр} = \sqrt[n]{\frac{y_n}{y_0}} = \sqrt[n]{Тр_{баз}};$$

где y_n, y_0 – конечный и начальный уровни ряда;

n – число приростов;

$Тр_{баз}$ – темп роста базисный за весь период.

Для данных из табл. 14

$$\overline{Тр} = \sqrt[4]{\frac{65}{48}}$$ (за четыре года)

– если имеются данные о цепных коэффициентах роста, то

$$\overline{Тр} = \sqrt[n]{T_1 \times T_2 \times ... \times Tn},$$

где $T_{1, 2, ..., n}$ – цепные коэффициенты роста;

n – число цепных коэффициентов роста.

4. **Средний темп прироста ($\overline{Тпр}$)** показывает, на сколько процентов в среднем изменяются уровни ряда за данный период, вычисляется только исходя из средних темпов роста, для чего пользуются соотношением

$$\overline{Тпр} = \overline{Тр} - 1$$

или

$$\overline{Тпр}\ (\%) = \overline{Тр}\ \% - 100\ \%.$$

Следовательно, для вычисления среднего темпа прироста вначале нужно обязательно определить средний темп роста, а затем уже средний темп прироста.

Применение перечисленных показателей динамики является первым этапом анализа ряда динамики, позволяющим выявить скорость, интенсивность развития явления, представленного рядом.

Вернуться в начало

6.5 Механические методы выявления основной тенденции развития

Если рассматривать уровни экономических показателей на коротких промежутках времени (например, день, месяц), то в силу влияния различных факторов, действующих в разных направлениях, их значения будут значительно колебаться. Поэтому очень трудно бывает выявить основную тенденцию развития изучаемого явления за какой-то период. Для того чтобы выявить основную тенденцию развития явления за период, можно использовать различные методы. К простейшим методам относятся следующие:

1. **Метод укрупнения интервалов.** Он применяется, когда данный период времени заменяется на более крупный (например, месяц на год; год на пятилетку). В таком новом ряду за уровень ряда применяется либо общий размер уровня за год, пятилетку, получаемые как сумма уровней, входящих в данный период, либо среднее значение за укрупненный интервал.

2. **Метод скользящей средней.** При выявлении основной тенденции развития с помощью данного метода, по-особому укрупняются интервалы времени: вместо каждого данного уровня берутся средние из рядом стоящих. Полученная средняя охватывает группу из некоторого числа уровней (3, 5, 7 и т. п., в середине которых находится взятый). Она будет скользящей, поскольку период осреднения меняется: из него убирается один уровень (первый) и добавляется следующий (например, $\frac{y_1+y_2+y_3}{3}$; $\frac{y_2+y_3+y_4}{3}$). В такой средней сглаживаются случайные отклонения:

Месяцы 1 2 3 4 5 6 7 8 9 10 11 12

Вернуться в начало

6.6 Аналитическое выравнивание ряда

Определение тренда (аналитического выражения) является наиболее эффективным способом выявления основной тенденции развития явления.

При этом уровни ряда динамики выражаются в виде функции времени

$$y_t = f(t).$$

Аналитическое выравнивание состоит в подборе для данного ряда динамики теоретической кривой, выражающей основные черты фактической динамики, то есть в подборе теоретической кривой, наилучшим образом описывающей эмпирические (фактические) данные.

Аналитическое выравнивание может проводиться с использованием различных трендов.

1. Наиболее простым является **выравнивание по прямой**:

$$y_t = a_0 + a_1 \times t,$$

где t – условное обозначение времени;

a_0, a_1 – параметры искомой прямой.

Параметры прямой, удовлетворяющие методу наименьших квадратов, находятся из решения системы уравнений

$$\begin{cases} na_0 + a_1 \sum_{i=1}^{n} t = \sum_{i=1}^{n} y_i \\ a_0 \sum_{i=1}^{n} t + a_1 \sum_{i=1}^{n} t^2 = \sum_{i=1}^{n} y_i \times t \end{cases},$$

где y – фактические значения уровней;

n – число уровней ряда.

Систему уравнений можно упростить, если t подобрать так, чтобы сумма была равна 0, то есть начало отсчета времени перенести в середину рассматриваемого периода, тогда

$$a_0 = \frac{\sum_{i=1}^{n} y}{n}; \qquad a_1 = \frac{\sum_{i=1}^{n} y_i \times t}{\sum_{i=1}^{n} t^2}.$$

При этом подбор t осуществляется:

а) если число уровней ряда четное, то условное обозначение времени t строится следующим образом: …–7–5–3–1+1+3+5+7… (то есть два серединных момента принимаются –1, +1. Все остальные, соответственно, обозначаются через 2 интервала);

б) при нечетном числе отсчет ведется от середины, принятой за ноль, через единицу: …–3–2–1 0+1+2+3…

Значения можно находить, пользуясь следующими формулами:

$$\sum_{i=1}^{n} = \frac{n(n^2-1)}{3},$$

(n – четное)

$$\sum_{i=1}^{n} t^2 = \frac{n(n^2-1)}{12}. \text{(n – нечетное)}$$

2. **Выравнивание по параболе** (2-го порядка). Если выравнивание производить по многочлену более высокой степени

(например, 2-го порядка): $yt = a_0 + a_1t + a_2t^2$, то система нормальных уравнений, получаемых методом наименьших квадратов, для определения параметров параболы имеет вид:

$$\begin{cases} na_0 + a_1\sum_{i=1}^{n} t + a_2\sum_{i=1}^{n} t^2 = \sum_{i=1}^{n} y_i; \\ a_0\sum_{i=1}^{n} t + a_1\sum_{i=1}^{n} t^2 + a_2\sum_{i=1}^{n} t^3 = \sum_{i=1}^{n} y_i t; \\ a_0\sum_{i=1}^{n} t^2 + a\sum_{i=1}^{n} t^3 + a\sum_{i=1}^{n} t^4 = \sum_{i=1}^{n} y_i t^2. \end{cases}$$

Для упрощения системы t подбираем так, чтобы $\sum_{i=1}^{n} t = 0$ и $\sum_{i=1}^{n} t^3 = 0$ (как показано ранее), тогда система упростится

$$\begin{cases} na_0 + a_2\sum_{i=1}^{n} t^2 = \sum_{i=1}^{n} y_i; \\ a_1\sum_{i=1}^{n} t^2 = \sum_{i=1}^{n} y_i t; \end{cases}$$

$$a_0 t^2 + a_2\sum_{i=1}^{n} t^4 = \sum_{i=1}^{n} y_i t^2,$$

следовательно,

$$a_1 = \frac{\sum_{i=1}^{n} y_i t}{t^2},$$

где a_0, a_2 определяются решением системы из 2 уравнений;

a_0 – величина, выражающая средние условия образования уровней ряда;

a_1 – скорость развития;

a_2 – ускорение этого развития.

Аналитическое выравнивание можно проводить и по многочленам более высоких степеней, например, параболе 3-го порядка

$$yt = a_0 + a_1t + a_2t^2 + a_3t^3.$$

3. **Показательная функция** применительно к выравниванию имеет следующий вид:

$$yt = a_0 \times a_1^t,$$

где a_0 – начальный уровень ряда;

a_1 – среднегодовой темп роста.

Для определения параметров уравнения методом наименьших квадратов, предварительно логарифмируют уровни, тогда логарифмы уровней отражаются линейной функцией

$$\lg yt = \lg a_0 + t \times \lg a_1;$$

если $\sum_{i=1}^{n} t = 0$, то

$$lga_0 = \frac{1}{n};\ \lg a_1 = \sum_{i=1}^{n} t \times \lg y \,/ \sum_{i=1}^{n} t^2.$$

При этом, чем выше порядок параболы, тем более точно она воспроизводит фактические данные.

Однако основной целью построения аналитического уравнения является не просто воспроизведение фактических данных, а определение тенденции развития данного явления во времени. Основанием для выбора формы кривой для выравнивания служит анализ сущности явления.

Вернуться в начало

6.7 Интерполяция и экстраполяция

Выравнивание рядов динамики используют не только для выявления тенденций, но и для того, чтобы найти недостающее значение уровня ряда. Такой способ нахождения недостающего значения внутри рассматриваемого периода, основанный на выравнивании рядов динамики, называется **интерполяцией**. Другой прием заключается в том, что, продолжая данные математических кривых, как бы предсказывается дальнейшее развитие явления, то есть на основе выявления особенностей изменения явлений за данный период можно предугадать поведение явления в будущем (прошлом). Он называется **экстраполяцией**.

Экстраполяцию и интерполяцию можно осуществить различными способами. Но они обязательно основываются на предположении о том, что закономерность (тенденция) изменения изучаемого явления, выявленная для определенного периода времени, сохранится на ограниченном отрезке времени как в будущем (прошлом), так и внутри данного периода.

Так как в действительности тенденция может измениться, то полученные таким путем данные надо рассматривать как своего рода оценку. Рассмотрим некоторые простейшие приемы, помогающие прогнозировать те или иные показатели за определенный отрезок времени.

1. **Прогнозирование на основе среднего абсолютного прироста.** Если при анализе ряда динамики обнаруживается, что абсолютные приросты уровней примерно постоянны, то в этом случае рассчитывают средний абсолютный прирост и последовательно прибавляют (вычитают) его к последнему известному уровню ряда столько раз, на сколько периодов экстраполируют (интерполируют) ряд.

Разберем это на основе данных, представленных в табл. 15.

Таблица 15

Производство обуви ОАО «Вест»

Показатель	2005 г.	2006 г.	2007 г.	2008 г.
Производство обуви, тыс. пар	788	801	809	820
$\Delta y_i = y_i - y_{i-1}$	–	+13	+8	+11

$$\Delta y = \frac{\sum_{i=1}^{n} \Delta y_i}{n} = \frac{y_n - y_0}{n} = \frac{13+8+11}{3} = \frac{820-788}{3} = 10,7 \text{ млн. пар}$$

Определим объем производства 2010 г.

$$y_{26} = 820 + 10,7 \times 2 = 841,4 \text{ тыс. пар.}$$

Аналогично можно и интерполировать: на основе среднего абсолютного прироста и последнего перед недостающим уровнем значения ряда

$$y_{2007} = y_{2008} - \overline{\Delta y} \times n = 820 - 10,7 = 809,3 \text{ тыс. пар,}$$

то есть оценку, а не точные значения дает прогнозирование.

2. **Прогнозирование на основе среднегодового темпа роста.** Если за исследуемый ряд лет годовые коэффициенты роста относительно постоянны, то в этом случае можно рассчитать

средний коэффициент роста и последний известный уровень ряда умножить (разделить) на средний коэффициент роста в степени, соответствующей периоду экстраполяции.

Например, динамика численности населения региона имеет следующую тенденцию (табл. 16).

Таблица 16

Динамика численности населения региона

Показатель	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.
Численность населения на 31 декабря, чел.	52	53	54,5	55,8	57,5	59,1
К роста (цеп.)	–	1,02	1,03	1,025	1,03	1,028

$$\overline{\text{Тр}} = \sqrt[5]{\frac{59,1}{52}} = \sqrt[5]{1,02 \times 1,03 \times 1,025 \times 1,03 \times 1,028} = 1,026.$$

Если исходить из предположения, что данный темп роста сохранится на определенные промежутки и в дальнейшем или прошедшем, то можно определить численность, например, на 2022 г.

$$у_{2022} = 59,1 \times 1,026^{17} = 91,43 \text{ млн чел.}$$

или на 1999 г.

$$у_{1999} = \frac{52}{1,026} = 50,7 \text{ млн. чел.}$$

Аналогично производят и интерполяцию.

3. **Прогнозирование на основе какой-либо аналитической формулы.** Зная уравнение для исчисления теоретических уровней и подставляя в него значения t за пределами (внутри) исследуемого ряда, можно рассчитать для данных t вероятные значения уровней (yt).

Например, тенденция производства стали за 1985–1989 г.г. характеризуется следующим образом

$$yt = 120,8 + 5,21 \times t,$$

	1985	1986	1987	1988	1989	1990
t	–2	–1	0	1	2	3

то для 2007 г.

$$y_{2007} = 120{,}8 + 5{,}21 \times 17 = 209{,}37 \text{ млн т},$$

$$y_{1986} = 120{,}8 + 5{,}21 \times (-1) = -115{,}59 \text{ млн т}.$$

Вернуться в начало

7 ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ

7.1 Понятие индексов и их классификация

Индексами в статистике называются относительные величины, характеризующие соотношение показателей во времени, в пространстве или фактических с плановыми.

При всем их разнообразии, индексы можно разделить на два класса:

1. Индивидуальные, элементарные, простые (i).
2. Сводные, сложные, общие (I).

Под **элементарными индексами** понимаются относительные величины, характеризующие изменение во времени показателей, относящихся к одному объекту, или сравнивающие размеры показателей для одновременно существующих однородных объектов.

$$\mathrm{i} = \frac{a_1}{a_0} \text{ или } i = \frac{a_1}{a_{пл}}, \quad \mathrm{i} = \frac{a_В}{a_С},$$

то есть соотносятся величины простого явления (объекта), характеризуемые одним показателем (например, цена или себестоимость по одному конкретному виду продукции),

$$i_p = \frac{P_1}{P_0}; \qquad \mathrm{i}_z = \frac{Z_1}{Z_0}.$$

При этом обязательным условием для его вычисления является максимальная однородность объекта, для которого он вычисляется.

Однако в экономике существуют в основном сложные явления, составные части которых нельзя непосредственно сложить. Относительные показатели, характеризующие изменение сложного явления в целом, характеризуемого двумя и более показателями, называются **сводными, общими индексами**.

$$I_p = \frac{\sum_{i=1}^{n} P_1 \times q_1}{\sum_{i=1}^{n} P_0 \times q_1}; \quad I_{pq} = \frac{\sum_{i=1}^{n} P_1 \times q_1}{\sum_{i=1}^{n} P_0 \times q_0},$$

где P_1, P_0 – цена изделия соответственно в отчетном и базисном периодах;

q_1, q_0 – объем произведенной продукции соответственно в отчетном и базисном периодах.

Исходными величинами для их построения служат индивидуальные индексы, размеры явлений, специальные расчетные показатели. Выражаются индексы в коэффициентах или процентах.

Вернуться в начало

7.2 Виды сложных индексов

Сложные индексы, в свою очередь, можно разделить на основные два вида:

1) агрегатные;

2) средние.

Агрегатным называется сложный **индекс**, полученный путем сопоставления итогов, выражающих величину сложного явления в отчетном (базисном) периодах при помощи соизмерителей, а способ их получения – агрегатным. Отличительной особенностью любого агрегатного индекса является то, что в числителе и знаменателе этого индекса фигурирует сумма произведений двух и более показателей, один может меняться, то есть индексируется, другой же выступает в роли соизмерителя, то есть остается неизменным.

Например, на основе агрегатного способа построим общий индекс товарооборота (Т)

$$I_{pq}=\frac{\sum_{i=1}^{n}T_1}{\sum_{i=1}^{n}T_0}=\frac{\sum_{i=1}^{n}P_1\times q_1}{\sum_{i=1}^{n}P_0\times q_0}.$$

Данный индекс зависит от двух величин, то есть изменение величины товарооборота зависит от изменения цен (P) на продаваемые товары и объема продаж (q) этих товаров, поэтому можно построить два индекса, каждый из которых характеризует влияние лишь одного фактора:

1. Индекс цен (I_P) характеризует изменение товарооборота за счет изменения цен (то есть во сколько раз изменится товарооборот за счет изменения цен)

$$I_p=\frac{\sum_{i=1}^{n}P_1\times q_1}{\sum_{i=1}^{n}P_0\times q_1};$$

$$Э_p=\sum_{i=1}^{n}(P_1-P_0)\times q_1.$$

Разность числителя и знаменателя данного индекса характеризует абсолютную величину такого изменения (в руб.).

В индексе цен цена изменяется, так как ее влияние определяется, объем фиксируется на уровне отчетного периода, так как это количественный фактор и его влияние устраняется.

2. Индекс физического объема (I_q) характеризует изменение товарооборота за счет изменения объема (то есть во сколько раз изменится товарооборот за счет изменения объема)

$$I_q=\frac{\sum_{i=1}^{n}P_0\times q_1}{\sum_{i=1}^{n}P_0\times q_0};$$

$$Э_q=\sum_{i=1}^{n}(q_1-q_0)\times P_0.$$

Разность числителя и знаменателя данного индекса характеризует абсолютную величину такого изменения (в руб.).

В индексе физического объема объем изменяется, так как его влияние устанавливается, цена фиксируется на уровне базисного периода, так как это качественный фактор и его влияние устраняется.

При построении этих индексов используется правило фиксации Г. Пааше: в индексе, характеризующем влияние качественного показателя (p, z, w), данный показатель индексируется, (то есть изменяется), тогда как другой количественный показатель, влияние которого устраняется, фиксируется (то есть остается неизменным), причем на уровне отчетного периода (q_1, T_1).

Если же характеризуем влияние количественного фактора, то он индексируется, другой же – качественный – фиксируется на уровне базисного периода (Zo, Po, Wo) (то есть на уровне базисного периода «0» фиксируется качественный показатель, а на уровне отчетного «1» – количественный).

Средние индексы. Агрегатный способ исчисления общих индексов является основным, но не единственным. Другой способ состоит в том, что по отдельным видам показателей рассчитываются индивидуальные индексы, а затем из них рассчитывается средний. При построении среднего индекса возникает вопрос о форме средней величины, используемой для его вычисления, и о весах. В практике статистики средний индекс рассчитывается как средние арифметические или гармонические величины, обязательно взвешенные

$$I_{арм} = \frac{\sum_{i=1}^{n} i \times f}{\sum_{i=1}^{n} f}; \quad I_{гарм.} = \frac{\sum_{i=1}^{n} M}{\sum_{i=1}^{n} \frac{M}{i}},$$

где i – индивидуальный индекс;

f, M – веса.

Для того чтобы правильно выбрать веса и форму среднего индекса, следует руководствоваться тем, что средний индекс должен быть тождественен агрегатной форме, которая является основной.

Исходя из этого все индексы можно свести в две группы:

1) агрегатные индексы с базисными весами, им соответствуют средние арифметические с базисными весами.

$$I_q = \frac{\sum_{i=1}^{n} q_1 \times P_0}{\sum_{i=1}^{n} q_0 \times P_0} = \frac{\sum_{i=1}^{n} i_q \times q_0 \times P_0}{\sum_{i=1}^{n} q_0 \times P_0} = I_{арм.} ,$$

Для того чтобы перейти к среднему индексу, показатель q_1 (в числителе индекса) выражают через индивидуальный индекс

$$i_q = \frac{q_1}{q_0}; \quad q_1 = i_q \times q_0.$$

Данное выражение подставляем в агрегатную форму (в числитель) и получаем среднеарифметический индекс;

2) агрегатные индексы с текущими весами, им соответствуют средние гармонические с текущими весами.

$$I_p = \frac{\sum_{i=1}^{n} q_1 \times P_1}{\sum_{i=1}^{n} q_1 \times P_0} = \frac{\sum_{i=1}^{n} q_1 \times P_1}{\sum_{i=1}^{n} \frac{q_1 \times P_1}{i_p}} = I_{гарм.} \quad .$$

Для того чтобы перейти к среднему индексу, показатель Ро (в знаменателе индекса) выражают через индивидуальный индекс

$$i_p = \frac{P_1}{P_0}; \quad P_0 = \frac{P_1}{i_p}.$$

Данное выражение подставляем в агрегатную форму (в знаменатель индекса) и получаем среднегармонический индекс.

В этом смысле общий индекс изучаемого явления рассматривается как результат изменения уровня данного явления у отдельных единиц совокупности. В процессе осреднения индивидуальных индексов веса подбираются так, чтобы был возможен переход от общего индекса в форме средней величины к общему индексу в агрегатной форме.

Вернуться в начало

7.3 Индексы с различной базой сравнения, постоянными и переменными весами

Когда возникает необходимость изучить развитие явления за определенный период времени, то используется система индексов, которая последовательно характеризует изменения, происходящие в течение выбранного интервала времени. Система содержит (n–1) индексов, где n – это число абсолютных уровней (y_i) в данном ряду динамики.

Возможны два варианта построения системы индексов:

– показатели периода сравниваются с одним, принятым за базу:

$$\frac{y_2}{y_1}, \frac{y_3}{y_1}, \frac{y_4}{y_1}, \ldots \frac{y_n}{y_1},$$

то есть система базисных индексов;

– показатели сравниваются между собой последовательно (последующий с предыдущим):

$$\frac{y_2}{y_1}, \frac{y_3}{y_2}, \frac{y_4}{y_3}, \ldots \frac{y_n}{y_{n-1}},$$

то есть система цепных индексов.

Базисная система дает представление об общем изменении изучаемого явления (как, во сколько раз за весь период);

система цепных индексов – последовательное изменение уровней (как, во сколько раз, например, ежегодно).

Произведение цепных индексов дает соответствующий базисный индекс

$$\frac{y_2}{y_1} \times \frac{y_3}{y_2} \times \frac{y_4}{y_3} \times \frac{y_n}{y_{n-1}} = \frac{y_n}{y_1}.$$

Однако эта взаимосвязь безусловна только для индивидуальных индексов. Для общих индексов эта зависимость будет сохраняться, если система общих индексов рассчитана с одними и теми же весами, то есть для так называемой системы индексов с постоянными весами.

Так, при исчислении цепных индексов физического объема, продукцию можно оценить в одних и тех же ценах

$$I_{p_1} = \frac{\sum_{i=1}^{n} q_2 \times P_1}{\sum_{i=1}^{n} q_1 \times P_1}; \frac{\sum_{i=1}^{n} q_3 \times P_1}{\sum_{i=1}^{n} q_2 \times P_1}; \frac{\sum_{i=1}^{n} q_4 \times P_1}{\sum_{i=1}^{n} q_3 \times P_1}; \ldots \frac{\sum_{i=1}^{n} q_n \times P_1}{\sum_{i=1}^{n} q_{n-1} \times P_1}.$$

То есть все индексы имеют одни и те же веса и поэтому представляют систему цепных индексов с постоянными весами, и, следовательно, для них сохраняется взаимосвязь между цепными и базисными индексами.

$$\frac{\sum_{i=1}^{n} q_2 \times P_1}{\sum_{i=1}^{n} q_1 \times P_1} \times \frac{\sum_{i=1}^{n} q_3 \times P_1}{\sum_{i=1}^{n} q_2 \times P_1} \times \frac{\sum_{i=1}^{n} q_4 \times P_1}{\sum_{i=1}^{n} q_3 \times P_1} \times \frac{\sum_{i=1}^{n} q_n \times P_1}{\sum_{i=1}^{n} q_{n-1} \times P_1} = \frac{\sum_{i=1}^{n} q_n \times P_1}{\sum_{i=1}^{n} q_1 \times P_1}$$

Однако при построении системы цепных индексов можно использовать разные соизмерители

$$I_q = \frac{\sum_{i=1}^{n} q_2 \times P_1}{\sum_{i=1}^{n} q_1 \times P_1}; \frac{\sum_{i=1}^{n} q_3 \times P_2}{\sum_{i=1}^{n} q_2 \times P_2}; \frac{\sum_{i=1}^{n} q_4 \times P_3}{\sum_{i=1}^{n} q_3 \times P_3} \ldots \frac{\sum_{i=1}^{n} q_n \times P_n}{\sum_{i=1}^{n} q_{n-1} \times P_n}.$$

Это система цепных индексов с переменными весами, так как вес меняется при переходе к другому индексу. В этом случае переход от цепных к базисным индексам невозможен.

Аналогично строится система индексов цен:

– система цепных индексов с постоянным весом

$$I_p = \frac{\sum_{i=1}^{n} P_2 \times q_1}{\sum_{i=1}^{n} P_1 \times q_1}; \frac{\sum_{i=1}^{n} P_3 \times q_1}{\sum_{i=1}^{n} P_2 \times q_1}; \frac{\sum_{i=1}^{n} P_4 \times q_1}{\sum_{i=1}^{n} P_3 \times q_1}; \ldots \frac{\sum_{i=1}^{n} P_n \times q_1}{\sum_{i=1}^{n} P_{n-1} \times q_1};$$

с переменными весами

$$I_p = \frac{\sum_{i=1}^{n} P_2 \times q_2}{\sum_{i=1}^{n} P_1 \times q_2}; \frac{\sum_{i=1}^{n} P_3 \times q_3}{\sum_{i=1}^{n} P_2 \times q_3}; \frac{\sum_{i=1}^{n} P_4 \times q_4}{\sum_{i=1}^{n} P_3 \times q_4}; \ldots \frac{\sum_{i=1}^{n} P_n \times q_n}{\sum_{i=1}^{n} P_{n-1} \times q_n}.$$

Вернуться в начало

7.4 Индексы переменного состава, фиксированного состава и структурных сдвигов

При изучении динамики качественных показателей (p, z, w) часто приходится определять изменение средней величины индексируемого показателя для однородной совокупности.

В общем виде динамику таких средних показателей можно выразить в виде отношения:

$$I = \bar{X}_1 : \bar{X}_0 = \frac{\sum_{i=0}^{n} X_1 \times f_1}{\sum f_1} : \frac{\sum_{i=0}^{n} X_0 \times f_0}{\sum f_0} = \frac{\sum_{i=0}^{n} X_1 \times d_1}{\sum_{i=0}^{n} X_0 \times d_0},$$

где

$$d_i = \frac{f_i}{\sum_{i=0}^{n} f_i}.$$

Такую относительную величину, характеризующую динамику средних показателей для однородной совокупности, называют **индексом переменного состава.**

Для различных качественных показателей (в однородной совокупности) индексы переменного состава легко записать в виде следующих отношений (например, для средней себестоимости):

$$I_{п.с.} = \bar{Z}_1 \div \bar{Z}_0 = \frac{\sum_{i=1}^{n} Z_1 \times q_1}{\sum_{i=1}^{n} q_1} \div \frac{\sum_{i=1}^{n} Z_0 \times q_0}{\sum_{i=1}^{n} q_0} = \frac{\sum_{i=1}^{n} Z_1 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_0};$$

где

$$\mathbf{d_i} = \frac{\mathbf{q_i}}{\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{q_i}}.$$

То есть d_i – это доля, удельный вес каждого предприятия (подразделения) в общем объеме производства.

Средние величины, динамику которых эти индексы отражают, могут меняться не только за счет изменения самого показателя (z, p) у отдельных объектов, но и за счет изменения удельного веса (доли) этих частей в общей совокупности ($qi / \Sigma qi = di$).

Таким образом, изменение средней величины, то есть индекс переменного состава зависит от изменения двух показателей (факторов): изменения данного показателя у отдельных объектов (z, p), а также изменения удельного веса этих частей в общей совокупности (di).

Поэтому индекс переменного состава можно разложить на два индекса-сомножителя, причем первый показывает изменение среднего показателя под влиянием изменения данного показателя у отдельных объектов, второй – под влиянием изменения удельного веса (доли) частей в общей совокупности.

Чтобы определить влияние на общее изменение средней величины только одного фактора, например данного показателя у отдельных объектов, влияние другого фактора нужно устранить, то есть зафиксировать, оставить неизменным.

Так, чтобы исключить влияние изменения удельного веса, то есть структуры совокупности на динамику средних величин, нужно для двух периодов рассчитать средний показатель по одной и той же структуре, то есть удельный вес (долю) зафиксировать, как правило, доля фиксируется на уровне отчетного периода (так как это количественный фактор).

1. Индекс, показывающий динамику средних величин при одной и той же фиксированной структуре, носит название **индекса фиксированного состава**

$$I_{ф.с.} = \frac{\sum_{i=1}^{n} Z_1 \times q_1}{\sum_{i=1}^{n} q_1} \div \frac{\sum_{i=1}^{n} Z_0 \times q_1}{\sum_{i=1}^{n} q_1}$$

или в агрегатной форме

$$I_{ф.с.} = \frac{\sum_{i=1}^{n} Z_1 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_1}, \text{ где } \quad \mathbf{d_i = \frac{q_i}{\sum_{i=1}^{n} q_i}}.$$

Индекс фиксированного состава показывает, как изменяется средняя себестоимость в результате изменения себестоимости данного вида продукции на отдельных предприятиях (цехах).

В этом индексе влияние структурного фактора устранено. Величина его не может выходить за пределы значения частных индексов, так как он является средним из них.

2. Для того чтобы определить влияние на общее изменение средней величины изменения структурного фактора (то есть доли, удельного веса), необходимо построить индекс, в котором устранено влияние осредняемого показателя у отдельных объектов, то есть его необходимо зафиксировать, так как это качественный показатель, то его обычно фиксируют на уровне базисного периода. Такой индекс, характеризующий влияние изменения структуры на общее изменение средней величины в однородной совокупности, называют **индексом структурных сдвигов**

$$I_{стр.} = \frac{\sum_{i=1}^{n} Z_0 \times q_1}{\sum_{i=1}^{n} q_1} \div \frac{\sum_{i=1}^{n} Z_0 \times q_0}{\sum_{i=1}^{n} q_0}$$

или в агрегатной форме

$$I_{стр.} = \frac{\sum_{i=1}^{n} Z_0 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_0}.$$

Данный индекс характеризует изменения средней себестоимости данного вида продукции в результате изменения доли, удельного веса данного предприятия (цеха) в общем объеме производства.

Разность числителя и знаменателя соответствующих индексов характеризует абсолютную величину такого изменения. Так как индекс переменного состава отражает на себе влияние двух факторов, а индекс фиксированного состава только влияние изменения осредняемого показателя без учета изменения структуры совокупности, то индекс структурных сдвигов можно найти и путем деления индекса переменного состава на индекс фиксированного состава

$$\mathbf{I}_{стр.} = \frac{\mathbf{I}_{п.с.}}{\mathbf{I}_{ф.с.}}.$$

Так как индекс переменного состава характеризует общее изменение средней величины под влиянием двух факторов, то

$$I_{п.с.} = I_{ф.с.} \times I_{стр.};$$

$$\frac{\sum_{i=1}^{n} Z_1 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_0} = \frac{\sum_{i=1}^{n} Z_1 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_1} \times \frac{\sum_{i=1}^{n} Z_0 \times d_1}{\sum_{i=1}^{n} Z_0 \times d_0}.$$

Данная группа индексов находит практическое применение в планировании и статистико-экономическом анализе деятельности предприятий и отраслей для характеристики изменения средней величины качественных показателей (цен, себестоимости, производительности труда и т. д.)

Вернуться в начало

7.5 Важнейшие экономические индексы и их взаимосвязь

1. **Индекс затрат** характеризует изменение общих затрат на производство продукции и зависит от изменения затрат на производство единицы продукции по отдельным ее видам, а также объема данного вида продукции, так как

Затраты = Себестоимость единицы продукции ×
× Объем производства,

то индекс

$$I_{затрат} = I_z \times I_q,$$

где Iz – индекс, который характеризует влияние изменения себестоимости единицы продукции (то есть Z);

Iq – индекс, характеризующий влияние изменения объема производства отдельных видов продукции.

2. **Индекс товарооборота** характеризует изменение объема товарооборота в стоимостном выражении (то есть объема продукции в определенных ценах) и зависит от изменения цен на отдельные виды продукции, а также объема продаж продукции, так как

Товарооборот = Цена × Объем реализации,

то

$$I_{т.о.} = I_p \times I_q,$$

где Ip – индекс, который характеризует влияние изменения цен на продукцию по ее отдельным видам;

Iq – индекс, характеризующий влияние изменения величины физического объема продукции.

3. **Индекс производительности труда** (w) характеризует изменение производительности труда и зависит от изменения объема производства продукции и изменения численности (затрат времени). Поскольку

$$\text{Производительность труда} = \frac{\text{Объем производства}}{\text{Численность}},$$

то

$$I_{пр.тр.} = \frac{I_q}{I_т},$$

где Iq – индекс, характеризующий влияние изменения объема производства;

Iт – индекс, характеризующий влияние изменения затрат времени при производстве продукции или изменения численности.

4. **Индекс заработной платы** характеризует изменения заработной платы и зависит от изменения фонда оплаты труда и численности.

Заработная плата = Фонд оплаты труда / Численность,

тогда

$$I_{з/п} = \frac{I_{ФОТ}}{I_т},$$

где $I_{ФОТ}$ – индекс, который характеризует влияние изменения фонда оплаты труда;

Iт – индекс, характеризующий изменение влияния численности работающих.

5. **Индекс удельного расхода сырья** характеризует изменение удельного расхода сырья на единицу продукции (m) и зависит от изменения общего объема затраченного сырья (М) и объема произведенной продукции (q)

$$m = \frac{M}{q},$$

где М – общий расход сырья на весь объем производимой продукции;

q – объем производства продукции.

$$I_m = \frac{I_M}{I_q},$$

Таким образом, общее для всех индексов: как взаимосвязаны между собой показатели, так и взаимосвязаны между собой индексы, характеризующие их изменение.

Вернуться в начало

8 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

8.1 Понятие выборочного наблюдения

Среди несплошного наблюдения наиболее распространенным видом является выборочное.

Выборочным наблюдением называется такое несплошное наблюдение, при котором обследованию подвергаются не все единицы изучаемой совокупности, а лишь отобранные в определенном порядке.

Цель выборочного наблюдения – по характеристикам отобранной части единиц судить о всей совокупности. Выборочное наблюдение включает следующие этапы:

1. Постановка цели наблюдения.
2. Составление программы наблюдения.
3. Решение организационных вопросов.
4. Определение процента отбора и способа (повторный или бесповторный).
5. Проведение отбора.
6. Регистрация признаков.
7. Обобщение данных наблюдения и расчет характеристик.

8. Расчет ошибок выборки.

9. Пересчет выборочных характеристик на генеральную совокупность.

К выборочным наблюдениям прибегают, когда:

– необходимо сэкономить силы и средства при проведении исследования;

– невозможно провести сплошное наблюдение (например, качество изделий определяется путем их разрушения).

Вернуться в начало

8.2 Классификация методов отбора

Отбор единиц из исследуемой совокупности можно производить по-разному, в зависимости от целого ряда условий. При этом систему организации отбора единиц из совокупности называют **методом отбора**.

Различают повторный и бесповторный методы отбора.

Повторный – такой метод, при котором отобранная однажды единица возвращается обратно в (генеральную) совокупность и снова участвует в отборе. При таком методе сохраняется постоянной вероятность попасть в выборку для всех единиц совокупности.

Бесповторный – такой метод отбора, при котором отобранная единица в совокупность единиц, из которых производится отбор, обратно не возвращается. При таком отборе для каждой новой единицы вероятность попасть в выборку увеличивается.

Так как бесповторный отбор охватывает все новые и новые совокупности, а повторный одну и ту же совокупность на всем его протяжении, следовательно, бесповторный отбор дает более точные результаты, чем повторный (так как вероятность попасть в выборку возрастает и отбираются каждый раз новые единицы из совокупности).

Вернуться в начало

8.3 Характеристики генеральной и выборочной совокупностей при выборочном наблюдении

В статистике совокупность, из которой производится отбор единиц, называют **генеральной**, а совокупность отобранных единиц – **выборочной совокупностью**. Эти совокупности характеризуются рядом показателей:

1) долей (p и w);

2) средним размером признака ($\bar{x}$, $\tilde{x}$);

3) ошибками выборки (μ, Δ).

1. Доля единиц, обладающих тем или иным признаком в генеральной совокупности, называется **генеральной долей** – p, а в выборочной совокупности – **выборочной долей, или частостью** – w.

2. Средняя величина показателя в генеральной совокупности называется **генеральной средней** – $\bar{x}$, в выборочной – **выборочной средней** – $\tilde{x}$.

3. Расхождение между выборочными и генеральными характеристиками называется ошибками выборки, причем отклонение генеральной доли от выборочной доли называется ошибкой при определении доли – μ_p, Δ_p, а генеральной средней от выборочной средней называется ошибкой при определении средней – μ_x, Δ_x.

Появление ошибок обусловлено тем, что:

– выборочному наблюдению, как и сплошному, свойственны ошибки регистрации как случайные, так и систематические;

– распространение результатов выборки на всю генеральную совокупность связано с возникновением ошибки репрезентативности. При достаточно большом числе наблюдений, должной организации выборки, величина таких ошибок выборки может быть доведена до сколь угодно малых размеров.

При соблюдении принципа случайного отбора ошибка выборки, прежде всего, зависит:

– от численности – чем больше численность выборки при прочих равных условиях, тем меньше величина ошибки;

– однородности совокупности – при одинаковой численности выборки ошибка будет меньше в той совокупности, в которой изучаемый признак варьирует в меньшей степени, то есть совокупность более однородна;

– метода отбора – при бесповторном методе, когда вероятность попасть в выборку возрастает с каждой отобранной единицей, ошибка выборки меньше, чем при повторном отборе.

Такая зависимость ошибки выборки от перечисленных факторов находит выражение в формулах средней ошибки выборки, которая характеризует среднюю величину отклонения характеристик генеральной и выборочной совокупностей.

При проведении выборочного наблюдения могут ставиться две задачи:

1. Измерить среднее значение самого варьирующего признака в генеральной совокупности на основе выборочной средней.
2. Определить генеральную долю по выборочным данным.

В связи с этими задачами определяются средние ошибки выборки, соответственно:

– для самого варьирующего признака μ_x;

– для генеральной доли μ_p.

Величина ошибки выборки зависит и от метода отбора, поэтому для ее определения используются различные формы.

Следовательно, когда выборочное обследование ставит задачей измерить среднее значение варьирующего признака в генеральной совокупности, тогда средняя ошибка выборки при определении средней рассчитывается в зависимости от метода отбора:

– *при повторном отборе*

$$\mu_x = \sqrt{\frac{\sigma^2}{n}},$$

где σ^2 – дисперсия варьирующего признака;

n – численность выборочной совокупности;

– *при бесповторном отборе*

$$\mu_x = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)},$$

где N – численность генеральной совокупности.

Когда же задачей ставится измерить долю признака в генеральной совокупности, тогда средняя ошибка выборки при определении доли рассчитывается в зависимости от метода отбора, соответственно

– *при повторном отборе*

$$\mu_x = \sqrt{\frac{w(1-w)}{n}},$$

где w – доля единиц, обладающих данным признаком в выборочной совокупности;

– при бесповторном отборе

$$\mu_x = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)},$$

Средняя ошибка выборки показывает, какие возможны отклонения характеристик выборочной совокупности от соответствующих характеристик генеральной. Однако о величине этой ошибки можно судить только с определенной вероятностью, то есть то, что генеральные характеристики не выйдут за определенные пределы, можно утверждать не с абсолютной достоверностью, а лишь с определенной степенью вероятности.

Чтобы определить изменение генеральных характеристик и утверждать, что эти характеристики не выйдут за данные пределы с определенной вероятностью, на основе средней ошибки выборки определяют предельную ошибку выборки (Δ), которая зависит от коэффициента доверия (t).

Доверительное число (t) указывает, что расхождение не превысит кратную ему среднюю ошибку выборки (μ)

$$\Delta = t \times \mu,$$

где t – коэффициент кратности ошибки, зависящий от вероятности, с которой можно гарантировать, что предельная ошибка не превышает t-кратной средней ошибки.

Каждому значению коэффициента доверия соответствует значение вероятности, табл. 17.

Например, если t = 3, то с вероятностью 0,997 можно утверждать, что расхождение не превысит трехкратную среднюю ошибку выборки.

Таблица 17

Соответствие коэффициента доверия и вероятности

Значение коэффициента доверия (t)	P
1	2
1	0,683
1	2
1,5	0,866
2	0,954
2,5	0,988
3	0,997
3,5	0,999

Зная выборочную среднюю величину ($\tilde{x}$), долю признака (w) и предельные ошибки выборки (Δ), можно определить границы, в которых заключены генеральные средняя и доля:

– для средней

$$\Delta x = t \times \mu x,$$

то есть

$$\bar{x} - \tilde{x} = \pm \Delta x$$

или

$$\tilde{x} - \Delta x \leq \bar{x} \leq \tilde{x} + \Delta x;$$

– для доли

$$\Delta p = t \times \mu p;$$

$$p - w = \pm \Delta p,$$

$$w - \Delta p \leq p \leq w + \Delta p,$$

то есть с определенной вероятностью, соответствующей установленному коэффициенту доверия (t), можно утверждать, что изменение генеральных характеристик будет лежать в указанных пределах.

Данные формулы дают возможность не только определить ошибки, но и рассчитать предварительно, какую необходимо взять численность выборки, чтобы ошибка не превышала определенные заданные размеры.

Так, численность выборки при собственно случайном и механическом отборе определяется следующим образом, табл. 18.

Таблица 18

Численность выборки при собственно случайном и механическом отборе

Метод отбора	Формулы объема выборки	
	для средней величины	для доли
повторный	$n = \dfrac{t^2\sigma^2}{\Delta_x^2}$	$n = \dfrac{t^2 w(1-w)}{\Delta_p^2}$
бесповторный	$n = \dfrac{t^2\sigma^2 N}{N\Delta_x^2 + t^2\sigma^2}$	$n = \dfrac{t^2 w(1-w)}{N\Delta_p^2 + t^2 w(1-w)}$

Вернуться в начало

8.4 Виды и способы отбора

В практической деятельности в сочетании с повторным и бесповторным методами отбора применяются три вида отбора:

1) индивидуальный – отбор единиц из совокупности;

2) групповой – отбор групп единиц из совокупности;

3) комбинированный – это комбинация первого и второго вида.

Различные виды отбора могут осуществляться различными способами:

– случайной выборкой;

– механической;

– типической;

– серийной;

– комбинированной.

Случайная выборка. При случайном способе выборки включение единиц в выборочную совокупность осуществляется наудачу. При этом выборка может осуществляться путем повторного и бесповторного отбора.

При случайной повторной выборке соблюдается независимость отбора единиц и сохраняется равная возможность для всех единиц совокупности оказаться включенными в состав выборки. Случайная выборка может вестись при помощи жеребьевки или с использованием таблиц случайных чисел, в которых дан набор чисел.

Средняя ошибка выборки определяется по формуле

$$\mu_x = \frac{\sigma}{\sqrt{n}}.$$

Бесповторная выборка производится также наудачу, но попавшая однажды в совокупность единица не возвращается и поэтому в другой раз в выборку попасть не может.

При этом средняя ошибка выборки определяется по формуле

$$\mu_x = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}.$$

Механическая выборка заключается в отборе единиц из генеральной совокупности, производимом в каком-либо механическом порядке, например, в отборе каждой пятой, десятой и т. д. единицы, при определенном расположении единиц в генеральной совокупности. При этом промежуток, через который попадают единицы в выборку, зависит от принятой пропорции отбора, которая устанавливается делением численности совокупности на объем выборки (N/n).

Чаще всего механический отбор применяется там, где имеется объективная последовательность в расположении единиц. Для определения средней ошибки механической выборки и ее численности следует использовать формулу

$$\mu_x = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}.$$

Типический отбор – это такая выборка, когда перед ее производством генеральная совокупность делится на группы по какому-либо типическому признаку (на типические группы), а затем внутри каждой группы производится случайная выборка.

Из всех типических групп можно отбирать число единиц, пропорциональное и непропорциональное их численности. В зависимости от этого различают пропорциональный (численности групп) и непропорциональный (например, 5-й из каждой группы) типический отборы. Кроме того, он может быть повторный и бесповторный.

Средняя ошибка при пропорциональной типической выборке определяется следующим образом:

при повторном отборе

$$\mu_x = \sqrt{\frac{\overline{\sigma^2}}{n}},$$

при бесповторном отборе

$$\mu_x = \sqrt{\frac{\overline{\sigma^2}}{n}\left(1 - \frac{n}{N}\right)},$$

где $\overline{\sigma^2}$ – средняя из дисперсий групп.

Для других разновидностей μх определить очень сложно.

Серийная выборка заключается в том, что вместо случайного отбора единиц совокупности осуществляется отбор групп, серий. Внутри отобранных серий производится сплошное наблюдение. Точность серийной выборки зависит не от величины общей дисперсии, а от дисперсии групповых средних величин.

Серийная выборка может производиться в порядке повторного и бесповторного отбора. Кроме того, серии могут быть равновеликими и неравновеликими. Средняя ошибка при отборе равновеликими сериями определяется по формулам:

– при повторном

$$\mu_x = \sqrt{\frac{\overline{\sigma^2}}{r}},$$

при бесповторном

$$\mu_x = \sqrt{\frac{\overline{\sigma^2}}{r}\left(1 - \frac{r}{R}\right)},$$

где

$$\overline{\sigma^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{r},$$

где r – число отобранных серий;

x_i – средняя ошибка в отобранных сериях;

$\bar{x}$ – общая средняя ошибка для всей совокупности;

R – общее число серий в генеральной совокупности.

Комбинированная выборка предполагает использование нескольких способов выборки, например, серийной и случайной с индивидуальным отбором единиц. В этом случае, разбив генеральную совокупность на серии (группы) и отобрав нужное число серий, производят случайную выборку единиц в сериях.

Средняя ошибка при разных комбинациях ее способов исчисляется по-разному, в зависимости от ступенчатости отбора.

Отбор называется **одноступенчатым**, если отобранные единицы подвергаются наблюдению и по ним делают обобщение. **Многоступенчатый** предполагает извлечение из генеральной совокупности сначала укрупненных групп единиц, затем групп, меньших по объему, и так до тех пор, пока не будут отобраны те, которые подвергаются наблюдению.

Вернуться в начало

9 СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗЕЙ

9.1 Основные понятия корреляционного и регрессионного анализа

Исследуя природу, общество, экономику, необходимо считаться со взаимосвязью наблюдаемых процессов и явлений.

Формы проявления взаимосвязей весьма разнообразны. В качестве двух самых общих их видов выделяют **функциональную (полную) и корреляционную (неполную)** связи. В первом случае величине факторного признака строго соответствует одно или несколько значений функции. Достаточно часто функциональная связь проявляется в физике, химии. В экономике примером может служить прямо пропорциональная зависимость между производительностью труда и увеличением объема производства продукции.

Корреляционная связь (которую также называют неполной, или статистической) проявляется в среднем, для массовых наблюдений, когда заданным значениям зависимой переменной соответствует некоторый ряд вероятных значений независимой переменной. Объяснение тому сложность взаимосвязей между анализируемыми факторами, на взаимодействие которых влияют неучтенные случайные величины. Поэтому связь между признаками проявляется лишь в среднем, в массе случаев. При корреляционной связи каждому значению аргумента соответствуют случайно распределенные в некотором интервале значения функции.

По направлению связи бывают **прямыми**, когда зависимая переменная растет с увеличением факторного признака, и **обратными**, при которых рост последнего сопровождается уменьшением зависимой переменной. Такие связи также можно назвать соответственно **положительными и отрицательными**.

Относительно своей аналитической формы связи бывают **линейными** и **нелинейными**. В первом случае между признаками в среднем проявляются линейные соотношения. Нелинейная взаимосвязь выражается нелинейной функцией, а переменные связаны между собой в среднем нелинейно.

Существует еще одна достаточно важная характеристика связей с точки зрения взаимодействующих факторов. Если характеризуется связь двух признаков, то ее принято называть **парной**. Если изучаются более чем две переменные – **множественной.**

Указанные выше классификационные признаки наиболее часто встречаются в статистическом анализе. Но, кроме перечисленных, различают также **непосредственные, косвенные** и **ложные** связи. Суть каждой из них очевидна. В первом случае факторы взаимодействуют между собой непосредственно, для косвенной связи характерно участие какой-то третьей переменной, которая опосредует связь между изучаемыми признаками. Ложная связь – это связь, установленная формально и, как правило, подтвержденная только количественными оценками. Она не имеет под собой качественной основы или же бессмысленна.

По силе различаются **слабые** и **сильные** связи. Эта формальная характеристика выражается конкретными величинами и интерпретируется в соответствии с общепринятыми критериями силы связи для конкретных показателей.

Задачи собственно **корреляционного анализа** сводятся к измерению тесноты связи между варьирующими признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на результативный признак.

Задачи **регрессионного анализа** лежат в сфере установления формы зависимости, определения функции регрессии, использования уравнения для оценки неизвестных значений зависимой переменной.

Решение названных задач опирается на соответствующие приемы, алгоритмы, показатели, применение которых дает основание говорить о статистическом изучении взаимосвязей.

Методы оценки тесноты связи подразделяются на корреляционные (параметрические) и непараметрические. Параметрические методы основаны на использовании, как правило, оценок нормального распределения и применяются в случаях, когда изучаемая совокупность состоит из величин, которые подчиняются закону нормального распределения. На практике это положение чаще всего принимается априори. Собственно, эти методы – параметрические – и принято называть корреляционными.

Непараметрические методы не накладывают ограничений на закон распределения изучаемых величин. Их преимуществом является простота вычислений.

Вернуться в начало

9.2 Парная корреляция и парная линейная регрессия

Простейшим приемом выявления связи между двумя признаками является построение **корреляционной таблицы** (табл. 19).

В основу группировки положены два изучаемых во взаимосвязи признака – X и Y. Частоты f_{ij} показывают количество соответствующих сочетаний X и Y. Если f_{ij} расположены в таблице беспорядочно, можно говорить об отсутствии связи между переменными. В случае образования какого-либо характерного сочетания f_{ij} допустимо утверждать о связи между X и Y. При этом, если f_{ij} концентрируются около одной из двух диагоналей, имеет место прямая или обратная линейная связь.

Наглядным изображением корреляционной таблицы служит **корреляционное поле**. Оно представляет собой график, где на оси абсцисс откладывается значение X, на оси ординат – Y, а точками показывается сочетание X и Y. По расположению точек, их концентрации в определенном направлении можно судить о наличии связи. В итогах корреляционной таблицы по строкам и столбцам приводятся два распределения – одно по X, другое по Y.

Таблица 19

Пример корреляционной таблицы

X \ Y	Y_1	Y_2	…	Y_z	Итого	$\overline{Y_i}$
X_1	f_{11}	12	…	f_{1z}	$\sum_{I}^{z} f_{1j}$	$\overline{Y_1}$
X_2	f_{21}	22	…	f_{2z}	$\sum_{I}^{z} f_{2j}$	$\overline{Y_2}$
…	…	…	…	…	…	…
X_k	f_{k1}	k2	…	f_{kz}	$\sum_{I}^{z} f_{kj}$	$\overline{Y_k}$
Итого	$\sum_{i=1}^{k} f_{i1}$	$\sum_{I}^{k} f_{i2}$	…	$\sum_{I}^{k} f_{iz}$	n	$\overline{Y}$
$\overline{X_j}$	$\overline{X_1}$	$\overline{X_2}$	…	$\overline{X_z}$	$\overline{X}$	–

Рассчитываем для каждого X_i среднее значение Y, то есть $\overline{Y_i}$, как

$$\overline{Y_i} = \sum_{j=1}^{z} Y_j f_{ij} / \sum_{j=1}^{z} f_{ij}, \quad i = 1,\ 2,\ ...,\ k.$$

Последовательность точек (X_i, $\overline{Y_i}$) дает график, который иллюстрирует зависимость среднего значения результативного признака Y от факторного X – **эмпирическую линию регрессии**, наглядно показывающую, как изменяется Y по мере X.

По существу, и корреляционная таблица, и корреляционное поле, и эмпирическая линия регрессии предварительно уже характеризуют взаимосвязь, когда выбраны факторный и результативный признаки и требуется сформировать предложения о форме и направленности связи. В то же время количественная оценка тесноты связи требует дополнительных расчетов.

Практически для количественной оценки тесноты связи широко используют **линейный коэффициент корреляции**. Иногда его называют просто коэффициентом корреляции. Если заданы значения переменных X и Y, то он вычисляется по формуле

$$r_{yx} = r_{xy} = \frac{\overline{XY} - \overline{X} \cdot \overline{Y}}{\sigma_x \cdot \sigma_y}.$$

Можно использовать и другие формулы, но результат должен быть одинаковым для всех вариантов расчета.

Коэффициент корреляции принимает значение в интервале от –1 до +1. Принято считать, что если $|r| < 0{,}30$, то связь слабая; при $|r| = (0{,}3/0{,}7)$ – средняя; при $|r| = > 0{,}70$ – сильная, или тесная. Когда $|r| = 1$ – связь функциональная. Если же $r \approx 0$, то это дает основание говорить об отсутствии линейной связи между Y и X. Однако в этом случае возможно нелинейное взаимодействие, что требует дополнительной проверки и других измерителей, рассматриваемых ниже.

Для характеристики влияния изменений X на вариацию Y служат методы регрессионного анализа. В случае парной линейной зависимости строится регрессионная модель

$$Y_i = a_0 + a_1 \cdot X_i + \varepsilon_i, \quad i = 1, \dots, n,$$

где n – число наблюдений;

a_0, a_1 – неизвестный параметр уравнения;

ε_i – ошибка случайной переменной Y.

Уравнение регрессии записывается как

$$Y_{iТЕОР} = a_0 + a_1X_i,$$

где $Y_{iТЕОР}$ – рассчитанное значение результативного признака после подстановки в уравнение X.

Параметры a_0 и a_1 оцениваются с помощью процедур, наибольшее распространение из которых **получил метод наименьших квадратов**. Его суть заключается в том, что наилучшие оценки a_0 и a_1 получают, когда

То есть сумма квадратов отклонений эмпирических значений зависимой переменной от вычисленных по уравнению регрессии должна быть минимальной. Сумма квадратов отклонений является функцией параметров a_0 и a_1. Ее минимизация осуществляется решением системы уравнений

$$\begin{cases} na_0 + a_1\sum_{i=1}^{n} X = \sum_{i=1}^{n} Y; \\ a_0\sum_{i=1}^{n} X + a_1\sum_{i=1}^{n} X^2 = \sum_{i=1}^{n} XY. \end{cases}$$

Можно воспользоваться и другими формулами, вытекающими из метода наименьших квадратов, например

$$a_1 = \frac{\sum_{i=1}^{n}\left(\overline{X}_i - \overline{X}\right)\left(\overline{Y}_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - X\right)^2} \text{ или } a_1 = r_{xy}\frac{\sigma_y}{\sigma_x},$$

$$a_0 = \overline{Y} - a_1\overline{X}.$$

Аппарат линейной регрессии достаточно хорошо разработан и, как правило, имеется в наборе стандартных программ оценки взаимосвязи для ЭВМ. Важен смысл параметров: a_1 – это коэффициент регрессии, характеризующий влияние, которое оказывает изменение X на Y. Он показывает, на сколько единиц в среднем изменится Y при X на одну единицу. Если a_1 больше 0, то наблюдается положительная связь. Если a_1 имеет отрицательное значение, то увеличение X на единицу влечет за собой уменьшение Y в среднем на a_1. Параметр a_1 обладает размерностью отношения Y к X.

Параметр a_0 – это постоянная величина в уравнении регрессии. В ряде случаев его интерпретируют как начальное значение Y.

Например, по данным о стоимости оборудования X и производительности труда Y методом наименьших квадратов получено уравнение

$$Y = -12{,}14 + 2{,}08X.$$

Коэффициент «a_1» означает, что увеличение стоимости оборудования на 1 млн руб. ведет в среднем к росту производительности труда на 2,08 тыс. руб.

Значение функции $Y = a_0 + a_1X$ называется расчетным значением и на графике образует **теоретическую линию регрессии.**

Смысл теоретической регрессии состоит в том, что это оценка среднего значения переменной Y для заданного значения X.

Вернуться в начало

9.3 Множественная линейная регрессия

Парная корреляция или парная регрессия могут рассматриваться как частный случай отражения связи некоторой зависимости переменной, с одной стороны, и одной из множества независимых переменных – с другой. Когда же требуется охарактеризовать связь всего указанного множества независимых переменных с результативным признаком, говорят о **множественной корреляции** или **множественной регрессии**.

Рассмотрим вопрос о регрессии. В ряде случаев именно от его решения – оценки уравнения регрессии – зависят оценки тесноты связи, а они, в свою очередь, дополняют результаты регрессивного анализа. Прежде всего следует определить перечень независимых переменных X, включаемых в уравнение. Это должно делаться на основе теоретических положений. Список X может быть достаточно широк и ограничен только исходной информацией. На практике теоретические положения о сути взаимосвязи подкрепляются парными коэффициентами корреляции между зависимой и независимыми переменными. Отбор наиболее значимых из них можно провести с помощью ЭВМ, выбирая в соответствии с коэффициентами корреляции и другими критериями факторы, наиболее тесно связанные с Y. Параллельно решается вопрос о форме уравнения. Современные средства вычислительной техники позволяют за относительно короткое время рассчитать достаточно много вариантов уравнений. В ЭВМ вводятся значения зависимой переменной Y и матрица независимых переменных X, принимается форма уравнения, например линейная. Ставится задача включить в уравнение k наиболее значимых X. В результате получим уравнение регрессии с k наиболее значимыми факторами. Аналогично можно выбрать наилучшую форму связи. Этот традиционный прием, называемый пошаговой регрессией, если он не противоречит качественным посылкам, достигает приемлемых результатов. Первоначально обычно берется линейная модель множественной регрессии

$$Y_i = a_0 + a_1X_{i1} + a_2X_{i2} + \ldots + a_kX_{ik} + \varepsilon_i$$

или в форме уравнения регрессии

$$Y_{ТЕОР} = a_0 + a_1X_1 + a_2X_2 + \ldots + a_kX_k,$$

где $Y_{теор}$ – расчетное значение регрессии, которое представляет собой оценку ожидаемого значения Y при фиксированных значениях $X_1, X_2, \ldots, X_k$;

$a_1, a_2, \ldots, a_k$ – коэффициенты регрессии, каждый из которых показывает, на сколько единиц изменится Y с изменением соответствующего признака X на единицу при условии, что остальные признаки останутся на прежнем уровне.

Параметры уравнения множественной регрессии, как правило, находятся методом наименьших квадратов. В матричной записи система уравнений имеет вид

$$\left(X^T X\right) \cdot A = X^T Y,$$

где

$$X = \begin{pmatrix} 1X_{11} \ldots X_{1k} \\ 1X_{21} \ldots X_{2k} \\ \ldots\ \ldots\ \ldots \\ 1X_{n1} \ldots X_{nk} \end{pmatrix}; \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \ldots \\ Y_n \end{pmatrix}; \quad A = \begin{pmatrix} a_0 \\ a_1 \\ \ldots \\ a_k \end{pmatrix}.$$

Оценка параметров множественной регрессии вручную затруднительна, приводит к потерям точности и может лишь удовлетворить любопытство. Получение же оценок параметров на ЭВМ в настоящее время не представляет большой проблемы. Гораздо важнее, насколько линейная форма связи соответствует реально существующей зависимости между Y, с одной стороны, и множеством X, с другой.

Вернуться в начало

9.4 Нелинейная регрессия. Коэффициенты эластичности

Представление связи через линейную функцию там, где на самом деле существуют нелинейные соотношения, вызовет ошибки аппроксимации и в конечном итоге упрощенные или даже ложные положения и выводы на основе аналитического уровня.

Вопрос о нелинейности формы уравнения следует решать на стадии теоретического анализа. Как правило, анализ должен опираться на суть взаимодействия изучаемых явлений и процессов и формально подкрепляться различного рода статистическими критериями. Но на практике допускается и другое решение: нелинейность формулируется как гипотеза и очерчивается лишь круг возможных уравнений, а затем форма и вид уравнения уточняются на ЭВМ. Существуют разные формы нелинейных уравнений регрессии, но в общем виде можно выделить два их класса.

К первому отнесем регрессии нелинейные относительно включенных в исследование переменных, но линейные по параметрам. Это, например, полиномы. В случае парной регрессии имеем уравнение

$$Y = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + \ldots .$$

Множественная регрессия $Y = f\left(X_1, \mathrm{X}_2\right)$ по аналогии выглядит следующим образом

$$Y = a_0 + a_1 X_1 + a_2 X_1^2 + a_3 X_1^3 + \ldots + b_1 X_2 + b_2 X_2^2 + b_3 X_2^3 + \ldots +$$
$$+ \mathrm{c}_1 \mathrm{X}_1 \mathrm{X}_2 + \mathrm{c}_2 \mathrm{X}_1 \mathrm{X}_2^2 + \mathrm{c}_3 \mathrm{X}_1^2 \mathrm{X}_2 + \ldots .$$

Возможно применение гиперболы, других функций. При желании с помощью стандартных программ ЭВМ может быть образовано любое нелинейное сочетание переменных, линейных относительно коэффициентов уравнения. Последние оцениваются с помощью метода наименьших квадратов.

Второй класс нелинейных функций отличается нелинейностью по оцениваемым параметрам. Таких уравнений также существует множество. Наиболее распространена степенная функция вида

$$Y = a_0 X^{a1} \text{ (парная регрессия)},$$

либо

$$Y = a_0 X_1^{a1} X_2^{a2} X_3^{a3} \cdots \text{ (множественная регрессия)}.$$

Даже по приведенным примерам можно составить мнение о широком спектре возможных аналитических представлений нелинейной формы связи. Ограничивает их использование сложность процедур оценивания параметров уравнений. Это подчас требует специальных приемов, алгоритмов, программ для ЭВМ.

Относительно просто решается такая задача для функций, преобразуемых к линейному виду. Например, степенную функцию можно прологарифмировать, получив линейную зависимость Y и X в логарифмах, и применить для оценки параметров уже упоминавшийся метод наименьших квадратов. Однако надо иметь в виду, что при этом оценивается не сама нелинейная функция, но ее линейное преобразование, а это может вызвать смещение оценок параметров.

Интерпретация коэффициента регрессии как углового коэффициента в линейном уравнении для нелинейной зависимости не годится. Определить изменение Y при изменении X на единицу можно с помощью производной (простой или частной), взятой по соответствующему фактору X. Так, для степенного уравнения $Y = a_0 X^{a1}$ производная по X равна

$$f'(X) = \frac{dY}{dX} = a_0 a_1 X^{a_1 - 1}.$$

Видно, что она является величиной переменной, а это усложняет экономическую интерпретацию результатов.

Чаще всего для характеристики влияния изменения X на Y используют так называемый **коэффициент эластичности** (Э), который показывает, на сколько процентов изменится Y при изменении X на один процент, то есть

$$Э = \frac{dY}{dX} \cdot \frac{X}{Y} = f'(X) \frac{X}{Y}.$$

Например, для линейного уравнения коэффициент эластичности фактора X выглядит как

$$Э = \frac{a_1 X}{Y} = \frac{a_1 X}{a_0 + a_1 X}.$$

Для парной степенной функции $Y = a_0 X^{al}$ коэффициент эластичности X равен a_1.

Коэффициенты эластичности – это, собственно, относительные величины. Их использование расширяет возможности сопоставления, экономической интерпретации результатов в дополнение к абсолютным величинам – коэффициентам регрессии.

Вернуться в начало

9.5 Множественная корреляция

Оценки тесноты связи (корреляции) могут играть двоякую роль. Это самостоятельные характеристики, дающие представление и о взаимодействии изучаемых факторов, и об аппроксимации фактических данных аналитической функцией. Поэтому расчет показателей множественной корреляции предполагает оценку уравнений регрессии.

При оценке линейной множественной связи рассчитывают **коэффициент множественной корреляции.** По смыслу он отражает тесноту связи между вариацией зависимой переменной и вариациями всех включенных в анализ независимых переменных. Обычно сначала строится линейная множественная регрессия, а затем оценивается сам коэффициент.

Наиболее общие формулы для его определения имеют следующий вид

$$R = \sqrt{1 - \frac{\sigma^2_{ост.}}{\sigma^2}}; \quad R^2 = \frac{\sum_{i=1}^{n}\left(Y_{iтеор.} - \overline{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2},$$

где σ^2 – общая дисперсия фактических данных результативного признака (дисперсия Y);

$\sigma^2_{ост.}$ – остаточная дисперсия, характеризующая вариацию Y за счет факторов, не включенных в уравнение регрессии.

Коэффициент множественной корреляции изменяется от 0 до 1, чем ближе R к 1, тем более сильная связь между Y и множеством X. Эта же оценка R используется и как мера точности аппроксимации фактических данных в уравнении. Если R незначительно по величине (как правило, $R \leq 0{,}3$), то можно утверждать, что либо не все важнейшие факторы взаимосвязи учтены, либо выбрана неподходящая форма уравнения. В этом случае следует пересмотреть список переменных модели, а возможно, и сам ее вид.

Для нелинейной множественной связи рассчитывают индекс корреляции. Форма и процедура его вычисления аналогичны указанным выше, только взаимодействие факторов аппроксимируется нелинейной функцией. Он также изменяется в пределах от 0 до 1. На практике, как правило, используется одно название – коэффициент множественной корреляции.

Квадрат R равен так называемому **коэффициенту детерминации** (D, или R^2). Он показывает, какая часть вариации зависимого признака объясняется включенными в модель факторами.

Вернуться в начало

9.6 Оценка значимости параметров взаимосвязи

Получив оценки корреляции и регрессии, необходимо проверить их на соответствие истинным параметрам взаимосвязи.

Существующие программы для ЭВМ включают, как правило, несколько наиболее распространенных критериев. Для оценки значимости коэффициента парной корреляции рассчитывают стандартную ошибку коэффициента корреляции

$$\sigma_{r_{xy}} = \sqrt{\frac{1 - r_{xy}^2}{n - 2}}.$$

В первом приближении нужно, чтобы $\sigma_{r_{xy}} < r_{xy}$. Значимость r_{xy} проверяется его сопоставлением с $\sigma_{r_{xy}}$, при этом получают

$$t_{расч.} = r_{xy} \cdot \sqrt{\frac{n - 2}{1 - r_{xy}^2}},$$

где $t_{расч.}$ – так называемое расчетное значение *t*-критерия.

Если $t_{расч.}$ больше теоретического (табличного) значения критерия Стьюдента ($t_{табл.}$) для заданного уровня вероятности и (n – 2) степеней свободы, то можно утверждать, что r_{xy} значимо.

Подобным же образом на основе соответствующих формул рассчитывают стандартные ошибки параметров уравнения регрессии, а затем и t-критерии для каждого параметра. Важно опять-таки проверить, чтобы соблюдалось условие $t_{расч.} > t_{табл.}$. В противном случае доверять полученной оценке параметра нет оснований.

Вывод о правильности выбора вида взаимосвязи и характеристику значимости всего уравнения регрессии получают с помощью F-критерия, вычисляя его расчетное значение

$$F_{расч.} = \frac{R^2 (n - m)}{\left(1 - R^2\right)(m - 1)},$$

где *n* – число наблюдений;

m – число параметров уравнения регрессии.

$F_{расч.}$ также должно быть больше $F_{теор.}$ при $V_1 = (m - 1)$ и $V_2 = (n - m)$ степенях свободы. В противном случае следует пересмотреть форму уравнения, перечень переменных и т. д.

Вернуться в начало

9.7 Непараметрические методы оценки связи

Методы корреляционного и дисперсионного анализа не универсальны: их можно применять, если все изучаемые признаки являются количественными. При использовании этих методов нельзя обойтись без вычисления основных параметров распределения (средних величин, дисперсий), поэтому они получили название **параметрических методов**.

Между тем в статистической практике приходится сталкиваться с задачами измерения связи между качественными признаками, к которым параметрические методы анализа в их обычном виде неприменимы. Статистической наукой разработаны методы, с помощью которых можно измерить связь между явлениями, не используя при этом количественные значения признака, а значит, и параметры распределения. Такие методы получили название **непараметрических**.

Если изучается взаимосвязь двух качественных признаков, то используют комбинационное распределение единиц совокупности в форме так называемых **таблиц взаимной сопряженности.**

Вернуться в начало

ВАРИАНТЫ ТЕСТОВЫХ ЗАДАНИЙ ПО ОСНОВНЫМ ТЕМАМ КУРСА

Темы: Предмет и метод курса «Статистика». Статистическое наблюдение. Сводка и группировка статистических данных. Статистические таблицы.

1. Предмет изучения статистики:

1) массовые общественные явления;

2) единицы совокупности;

3) совокупность;

4) нет правильного ответа;

5) показатели.

2. Наибольшее значение признака в интервале называется:

1) нижней границей;

2) величиной интервала;

3) разницей между нижней границей и верхней границей интервала;

4) верхней границей интервала;

5) разницей между нижней границей и величиной интервала.

3. Статистическое наблюдение – это:

1) сбор статистической информации;

2) научно-организованная работа по сбору статистической ин-формации;

3) обработка информации;

4) источник получения информации;

5) нет правильного ответа.

4. Виды группировок:

1) количественные и качественные;
2) типологические, структурные, аналитические;
3) комбинационные;
4) вторичные;
5) нет правильного ответа.

5. Статистическая отчетность – это:
1) вид статистического наблюдения;
2) способ статистического наблюдения;
3) форма статистического наблюдения;
4) критерий статистического наблюдения;
5) тип статистического наблюдения.

6. По характеру разработки сказуемого различают статистические таблицы:
1) простые, сложные;
2) перечневые;
3) комбинационные, групповые, простые;
4) монографические;
5) сводные.

7. Способы статистического наблюдения:
1) отчетный, экспедиционный, самоисчисление, анкетный, кор-респондентский;
2) текущий, прерывный;
3) отчетность, специально организованное;
4) выборочное, монографическое, основного массива;
5) сплошное, несплошное.

8. Основные типы ошибок наблюдения:
1) регистрации и репрезентативности;
2) случайные и систематические;
3) случайные регистрации;
4) систематические, репрезентативности;
5) систематические регистрации.

9. Методы отбора:
1) сплошной, несплошной;

2) повторный, бесповторный;
3) текущий, прерывный;
4) случайный, систематический;
5) непрерывное, единовременное, периодическое.

10. С определенной вероятностью можно утверждать, что генеральная средняя будет лежать в пределах:

1) $x - \Delta_x \le \bar{x} \le x + \Delta_x$;
2) $\bar{x} - x = \pm\Delta_x$;
3) $W - \Delta_p \le P \le W + \Delta_p$;
4) $P - W = \pm\Delta_p$;
5) $W - \Delta_p \le \bar{x} \le W + \Delta_p$.

11. Группировки бывают:

1) типологические, структурные, аналитические;
2) вторичные;
3) комбинированные;
4) по одному признаку;
5) по двум и более признакам.

Вернуться в начало

Тема: Графическое изображение статистических показателей

1. Экспликация – это:

1) словесное объяснение содержания графика и геометрических знаков;
2) символы понятий;
3) эталоны геометрического знака;
4) элементы классификации;
5) элемент графика.

2. Классификация графиков проводится:

1) по различию полей графика;
2) по геометрическим знакам;
3) по различию полей, геометрических знаков, по цели, для чего применяется график и т. п.;
4) по пространственным ориентирам;

5) по экспликации.

3. Диаграммы делятся на виды:

1) сравнения, динамические, структурные;

2) картограммы, картодиаграммы, центрограммы;

3) столбиковые, полосовые, секторные, квадратные, точечные, линейные, изобразительные;

4) столбиковые, полосовые;

5) по цели использования и по геометрическому знаку.

4. Статистические карты бывают:

1) картограммы, картодиаграммы, центрограммы;

2) фоновые, точечные;

3) сравнения, структурные, динамические;

4) по цели использования;

5) столбиковые и полосовые.

5. По способу построения (полю) графика различают:

1) структурные, динамические, сравнения;

2) диаграммы и картосхемы;

3) столбиковые, полосовые, секторные, квадратные, точечные, линейные, изобразительные и т. д.;

4) картограммы, картодиаграммы, центрограммы;

5) фоновые, точечные.

6. По цели использования:

1) структурные, динамические, сравнения;

2) диаграммы и картосхемы;

3) столбиковые, полосовые, секторные, квадратные, точечные, линейные, изобразительные и т. д.;

4) картограммы, картодиаграммы, центрограммы;

5) фоновые, точечные.

7. По геометрическому знаку:

1) структурные, динамические, сравнения;

2) диаграммы и картосхемы;

3) столбиковые, полосовые, секторные, квадратные, точечные, линейные, изобразительные и т. д.;

4) картограммы, картодиаграммы, центрограммы;

5) фоновые, точечные.

8. В зависимости от условий применения различают следующие статистические карты:

1) структурные, динамические, сравнения;

2) диаграммы и картосхемы;

3) столбиковые, полосовые, секторные, квадратные, точечные, линейные, изобразительные и т. д.;

4) картограммы, картодиаграммы, центрограммы;

5) фоновые, точечные.

9. В диаграммах сравнения отображаются:

1) один и тот же показатель по разным объектам в данный период времени;

2) один и тот же показатель в динамике;

3) разные показатели по одному объекту;

4) состав и структура совокупности;

5) несколько показателей в динамике.

10. В структурных диаграммах отображаются:

1) один и тот же показатель по разным объектам в данный период времени;

2) один и тот же показатель в динамике;

3) разные показатели по одному объекту;

4) состав и структура совокупности;

5) несколько показателей в динамике.

11. В динамических диаграммах отображаются:

1) один и тот же показатель по разным объектам в данный период времени;

2) один и тот же показатель в динамике;

3) разные показатели по одному объекту;

4) состав и структура совокупности;

5) несколько показателей в динамике.

12. В столбиковых диаграммах сравнения столбики располагаются:

1) в хронологическом порядке;

2) в ранжированном порядке;
3) в порядке возрастания;
4) в порядке убывания;
5) не имеет значения.

13. В столбиковых динамических диаграммах столбики располагаются:
1) в хронологическом порядке;
2) в ранжированном порядке;
3) в порядке возрастания;
4) в порядке убывания;
5) не имеет значения.

14. В структурных диаграммах дается:
1) доля каждой составляющей в общем объеме;
2) изменение показателя;
3) изменение доли каждой части во времени;
4) изменение состава;
5) изменение удельного веса.
Вернуться в начало

Тема: Абсолютные и относительные величины

1. Относительная координация – это:
1) соотношение частей целого между собой;
2) соотношение части целого к самому целому;
3) соотношение показателя по одному объекту к такому же показателю по другому объекту;
4) соотношение показателя в данном периоде к такому же показателю в предшествующем периоде;
5) нет правильного ответа.

2. Планом предусматривалось снизить себестоимость на 5 %. Относительная величина выполнения планового задания составила 102 %. Определите, как изменилась себестоимость по сравнению с предыдущим годом:

1) снизилась на 3,1 %; 2) снизилась на 6,8 %;
3) 96,9 %; 4) выросла на 7,1 %;
5) снизилась на 5 %.

3. План перевыполнен на 1 %. Показатель динамики – 110 %. Определите относительную величину планового задания:

1) 108,9 %; 2) 110 %;
3) 8,9 %; 4) 109 %.

4. Планом предусматривалось увеличить выпуск на 5 %, фактически он вырос на 2 % и составил 20,4 млн руб. Определите абсолютную величину планового задания и процент выполнения плана:

1) 21 млн руб.; 97,1 %; 2) 20 млн руб.; 105 %;
3) 20,4 млн руб.; 102 %; 4) 20,4 млн руб.; 105 %;
5) 20 млн руб.; 102 %.

5. Относительный показатель выполнения плана производства продукции составил 103 %, при этом объем производства по сравнению с предшествующим периодом вырос на 2 %. Что предусматривалось планом?

1) снижение объема производства;
2) уровень производства последующего периода;
3) рост объема производства;
4) уровень производства предыдущего периода;
5) прежний уровень производства.

6. Планом предусмотрено увеличение объема продукции предприятия против прошлого года на 2,1 %. Фактически прирост продукции против прошлого года составил 4,8. Определите процент выполнения плана по выпуску продукции:

1) 102,0 %; 2) 102,64 %;
3) 102,67 %; 4) 107,0 %;
5) 97,42 %.

7. Предприятие перевыполнило план реализации продукции в отчетном году на 3,8 %. Увеличение реализации продукции *в отчетном году по сравнению с прошлым составило 5,6 %. Определите, каково было плановое задание по росту объема реализации продукции*:

1) увеличилось на 1,73 %;
2) увеличилось на 1,75 %;
3) увеличилось на 2,01 %;
4) увеличилось на 9,61 %;
5) уменьшилось на 1,71 %.

8. Объем сельскохозяйственной продукции в отчетном году составил 220 млн руб., в том числе животноводство – 126 млн руб. Какие относительные можно определить?

1) структуры и координации;
2) сравнения, интенсивности;
3) наглядности;
4) координации;
5) структуры, сравнения.

9. Какие это относительные: а) удельный вес городского населения – 65 %; б) объем за пять лет вырос в 3 раза; в) на 100 мужчин приходится 114 женщин?

1) структуры, динамики, координации;
2) координации, динамики, структуры;
3) интенсивности, сравнения, структуры;
4) интенсивности, динамики, координации;
5) динамики, координации, структуры.

10. Что это за относительные величины: а) производство сахара – 33 кг/чел.; б) удельный вес городского населения – 66 %; в) плотность населения – 10 чел./км^2?

1) интенсивности, координации, интенсивности;
2) уровня экономического развития, структуры, степени;
3) степени, координации, уровня экономического развития;
4) динамики, сравнения, наглядности;
5) степени, структуры, наглядности.

Вернуться в начало

Тема: Средние величины и показатели вариации

1. Определите среднюю производительность труда по трем бригадам в целом, если в первой бригаде производительность составила 10 т/чел., во второй – 12 т/чел., в третьей – 16 т/чел. При этом в первой бригаде работает 20 %, во второй – 30 % всех рабочих.

1) 13,6 т/чел; 2) 12,66 т/чел;
3) 15,3 т/чел; 4) 10 т/чел;
5) 12 т/чел.

2. Распределение семей по числу детей характеризуется следующим образом

Число детей	0	1	2	3	4	5	6
Количество семей	6	28	22	19	13	5	7

Определите медиану:

1) 28; 2) 2; 3) 22; 4) 1; 5) 3.

3. Средняя гармоническая применяется:

1) для расчета среднего темпа роста;

2) когда в логической формуле расчета показателя неизвестен числитель;

3) когда в логической формуле неизвестен знаменатель;

4) в моментных рядах;

5) нет правильного ответа.

4. Вариация – это:

1) изменение массовых явлений во времени;

2) изменение структуры статистической совокупности в пространстве;

3) изменение массовых явлений в пространстве;

4) изменение значений признака во времени и в пространстве;

5) изменение состава совокупности.

5. Взвешенные средние применяются:

1) для сгруппированных данных;

2) для характеристики структуры совокупности;

3) для характеристики однородности совокупности;

4) для характеристики координации совокупности;

5) для однородной совокупности.

6. Цехом произведены бракованные детали в трех партиях: в первой партии – 90 шт., что составило 3,0 % от общего числа деталей; во второй партии – 140 шт., или 2,8 %; в третьей партии – 160 шт., или 2 %.

Определите средний процент бракованных деталей на основе соответствующей средней:

1) $x_{гарм}$ = 2,44 %;

2) $x_{агр}$ = 2 %;

3) $\bar{x}_{хр}$ = 3 %;

4) $x_{геом}$ = 2,55 %;

5) $\bar{x}_{арм}$ = 2,6 %.

7. Возраст брокеров универсальной биржи колеблется в пределах от 20 до 26 лет

Возраст, лет	20	21	22	23	24	25	26	Итого
Количество брокеров	15	27	29	30	38	35	26	200

Определите медиану:

1) 23; 2) 30;

3) 24; 4) 38;

5) 29.

8. Остаток средств на расчетном счете предприятия был на 1 января 180 тыс. руб. 15 января поступило на расчетный счет 900 тыс. руб.

22 января списано с расчетного счета 530 тыс. руб. 27 января поступило на расчетный счет 380 тыс. руб.

С 28 января до конца месяца остаток средств на расчетном счете не изменился.

Определите среднесуточный остаток средств на расчетном счете предприятия в январе:

1) $\bar{y}$ = 563,17 млн руб.; 2) $\bar{y}$ = 563,87 млн руб.;

3) $\bar{y}$ = 563,86 млн руб.; 4) $\bar{y}$ = 563 млн руб.;

5) $\bar{y}$ = 600 млн руб.

9. Определите Мо в данном ряду

Стаж работы, лет	до 2	2–4	4–	6–8	9 и >
Численность, чел.	3	20	7	11	10

1) 3, 13 года; 2) 4, 25 года;
3) 5, 13 лет; 4) 4, 71 года;
5) 3 года.

10. Определите Ме по данным предыдущего задания

1) 7 чел.; 2) 4, 25 года;
3) 5, 13 лет; 4) 4,71 года;
5) 5 лет.

Вернуться в начало

Тема: Статистические индексы

1. Что характеризует индекс фиксированного состава (себестоимости):

1) изменение затрат за счет изменения себестоимости;

2) изменение средней себестоимости за счет себестоимости на отдельных предприятиях;

3) изменение себестоимости;

4) изменение средней себестоимости за счет изменения доли каждого предприятия в общем объеме производства;

5) изменение затрат.

2. Если себестоимость увеличилась на 14 %, а количество продукции снизилось на 6 %, то индекс затрат на производство будет равен

1) 107,16 %; 2) 120,16 %;
3) 121,16 %; 4) 108,16 %;
5) 110,16 %.

3. Для определения изменения стоимости продукции за счет изменения ее объема (товарной массы) необходимо:

1) из числителя индекса стоимости продукции вычесть его знаменатель;

2) из знаменателя индекса стоимости продукции вычесть его числитель;

3) из числителя индекса объема продукции вычесть его знаменатель;

4) из знаменателя индекса стоимости продукции вычесть его числитель;

5) числитель индекса объема продукции умножить на его знаменатель.

4. Цены на платные услуги в текущем периоде по сравнению с базисным выросли в 2,1 раза, а количество предоставленных услуг сократилось на 30 %.

Определите индекс стоимости предоставленных услуг

1) 3,0; 2) 1,47;
3) 1,64 %; 4) 0,70;
5) 1.

5. Индекс фиксированного состава рентабельности производства составлял 58 %, а индекс структурных сдвигов – 105 %. Определите индекс переменного состава рентабельности

1) 1,659; 2) 0,609;
3) 1,81; 4) 0,552;
5) 0,7.

6. Изменение показателей однородной совокупности характеризуют:

1) общие индексы;
2) индивидуальные индексы;
3) средние индексы;
4) агрегатные индексы;
5) индексы структурных сдвигов.

7. Как изменилась производительность труда, если объем производства вырос на 5 %, а численность снизилась на 2 %?

1) выросла на 10 %; 2) выросла на 3,26 %;
3) выросла на 7 %; 4) выросла на 2,5 %;
5) выросла на 7,14 %.

8. Определите абсолютный прирост товарооборота, вызванный изменением цен, если цена единицы изделия составляла ρ_0 *= 8 руб.,* ρ_1 *= 9 руб., реализовано продукции* q_0 *= 150 шт.,* q_1 *= 200 шт.*

1) вырос на 600 руб.; 2) выросла на 3,26 %;
3) вырос на 400 руб.; 4) выросла на 2,5 %;
5) снизился на 200 руб.

9. Определите абсолютную величину изменения затрат на производство продукции за счет изменения объема производства, если z_0 *= 5 руб./т,* z_1 *= 5,5 руб./т,* q_0 *= 150 тыс. т,* q_1 *= 130 тыс. т*:

1) –100 тыс. руб.; 2) +65 тыс. руб.;
3) –110 тыс. руб.; 4) –20 тыс. т;
5) снизился на 200 руб.

10. Определите индекс структурных сдвигов, если себестоимость продукции в базисном периоде в первом цехе составила 14 руб./т при объеме производства 500 тыс. т, во втором цехе соответственно 13 руб./т и 500 тыс. т, в отчетном периоде в первом цехе себестоимость составила 13 руб./т при объеме 400 тыс. т, во втором цехе 12 руб./т и 600 тыс. т соответственно.

1) 91,85 %; 2) 99,26 %;
3) 92,54 %; 4) 104,42 %;
5) 98,12 %.

11. Сумму экономии от изменения цен, полученную при покупке товаров, можно определить

1) $\sum_{i=1}^{n}(p_1 - p_0)q_1$; 2) $\sum_{i=1}^{n}(p_1 - p_0)d_1$;

3) $\dfrac{\sum_{i=1}^{n} p_1 - q_1}{\sum_{i=1}^{n} p_0 - q_1}$; 4) $\dfrac{\sum_{i=1}^{n} p_1 - d_1}{\sum_{i=1}^{n} p_0 - d_1}$;

5) $\sum_{i=1}^{n} p_1 q_1 - \sum_{i=1}^{n} p_1 q_0$.

Вернуться в начало

Тема: Ряды динамики

1. *Абсолютный прирост, найденный цепным способом – это:*

1) $y_i - y_{i-1}$; 2) $y_i - y_o$;
3) y_i / y_{i-1}; 4) y_i / y_o;
5) $(y_i - y_{i-1}) / y_{i-1}$.

2. Темп прироста цепной показывает:

1) на сколько процентов изменяется данный уровень по сравнению с предыдущим;
2) на сколько процентов изменяется по сравнению с базисным;
3) во сколько раз изменяется по сравнению с базисным;
4) во сколько раз изменяется по сравнению с предыдущим;
5) нет правильного ответа.

3. Средний уровень моментного ряда находится по:

1) средней арифметической;
2) средней хронологической;
3) средней геометрической;
4) средней гармонической;
5) средней агрегатной.

4. Ряд динамики характеризует:

1) структуру совокупности по какому-либо признаку;
2) изменение характеристики совокупности в пространстве;
3) изменение характеристики совокупности во времени;
4) структуру совокупности по какому-либо признаку во времени;
5) изменение характеристики совокупности во времени и в пространстве.

5. Интерполяция уровней динамического ряда – это:

1) выравнивание динамического ряда;
2) выявление тренда;
3) нахождение уровней будущего динамического ряда;

4) нахождение промежуточных неизвестных уровней динамического ряда;

5) нахождение суммы уровней динамического ряда.

6. Темп роста исчисляется как:

1) отношение уровней;

2) разность уровней ряда;

3) произведение уровней ряда;

4) сумма уровней ряда;

5) нет правильного ответа.

7. За период 2005–2010 гг. объем производства вырос в 1,4 раза, в том числе за период 2005–2008 гг. в 2 раза. Определите, на сколько процентов изменился объем за период 2008–2010 гг.:

1) снизился на 30 %; 2) 280 %;
3) вырос на 180 %; 4) снизился на 60 %;
5) вырос на 42,9 %.

8. По годам снижение себестоимости составило 2 и 1 %. Как изменилась себестоимость за два года?

1) снизилась на 3 %; 2) снизилась на 2,98 %;
3) 97,02 %; 4) 1 %;
5) выросла на 3,02 %.

9. *Тенденция роста производства картофеля за 1992–1997 гг. была описана уравнением тренда* $y_t = 27{,}32 + 1{,}62t$. *Определите возможный уровень производства картофеля в 2005 г.*:

1) 61,34 млн т; 2) 45,14 млн т;
3) 59,72 млн т; 4) 38,56 млн т;
5) 38,46 млн т.

10. Как изменилась себестоимость продукции за три года, если по годам снижение составляло: 3 %; 2 %; 5 %?

1) снизилась на 10 %; 2) снизилась на 9,69 %;
3) снизилась на 17,06 %; 4) снизилась на 5 %;
5) снизилась на 90,31 %.

Вернуться в начало

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что является предметом изучения статистики?
2. Как называется наибольшее значение признака в интервале?
3. Что понимается под статистическим наблюдением?
4. Способы статистического наблюдения?
5. Сводка – это …?
6. Какие виды группировок существуют?
7. Что такое статистическая отчетность?
8. Виды статистических таблиц.
9. Основные типы ошибок наблюдения.
10. Каковы основные методы отбора?
11. Способы формирования выборочной совокупности?
12. От чего зависит величина ошибки выборки?
13. Чем средняя ошибка выборки отличается от предельной?
14. В чем разница между абсолютными и относительными величинами?
15. Какие виды относительных величин существуют?
16. Что характеризуют средние величины?
17. Виды средних величин?
18. Свойства средней арифметической?
19. Условия применения средних величин?
20. Показатели вариации?
21. В чем преимущество относительных показателей вариации?
22. Показатели структурных средних величин.
23. Что характеризует мода? Каким образом рассчитывается?
24. Как рассчитывается медиана?
25. Понятие индексов.
26. Классификация индексов.
27. Простые и сложные индексы.
28. Виды сложных индексов.
29. Важнейшие экономические индексы и их взаимосвязь.
30. Виды рядов динамики.
31. Показатели динамических рядов.
32. Что характеризует ряд динамики?
33. Что такое интерполяция?

34. Что такое экстраполяция?
35. Понятия и элементы графического изображения?
32. Виды диаграмм?
33. Сущность корреляционного анализа?
34. Задачи регрессионного анализа?
35. Классификация корреляционного анализа?
36. Критерии оценки значимости параметров взаимосвязи?
37. Что такое ряд распределения?
38. Что такое полигон распределения?
39. Когда применяется гистограмма?
40. Что такое кумулята и огива?
41. Показатели центра распределения – это …?

[Вернуться в начало]

СПИСОК ЛИТЕРАТУРЫ

1. Статистический справочник / Адамов В. Е. [и др.]. – Москва: Статистика, 1999. – 263 с.
2. Ефимова, М. Р. Общая теория статистики: учеб. для вузов. – Москва: ИНФРА-М, 1998.
3. Назаров, М. Г. Курс социально-экономической статистики: учеб. для вузов. – Москва: Финстатинформ, ЮНИТИ-ДАНА, 2000. – 432 с.
4. Пасхавер, И. С. Общая теория статистики: учебник / И. С. Пасхавер, А. Л. Яблочник. – Москва: Финансы и статистика, 1993. – 602 с.
5. Салин, В. Н. Социально-экономическая статистика: учеб. для вузов / В. Н. Салин, Е. П. Шпаковская. – Москва: Юристъ, 2001. – 461 с.
6. Минашкин, В. Г. Статистика: учеб. для вузов / В. Г. Минашкин [и др.]. – Москва: Проспект, 2006. – 266 с.
7. Статистика: учеб. пособие / А. В. Багат, М. М. Конкина, В. М. Симчера [и др.]. – Москва: Финансы и статистика, 2006. – 367 с.
8. Экономика и статистика фирм: учебник / В. Е. Адамов, С. Д. Ильенкова, Т. П. Сиротина, С. А. Смирнов. – Москва: Финансы и статистика, 1997. – 240 с.

9. Понкратова, Т. А. Общая теория статистики: конспект лекций для студентов всех специальностей и форм обучения. – Кемерово, 1997. – 63 с.

10. Сборник задач по общей теории статистики / Т. А. Понкратова [и др.]. – Кемерово, 1996. – 36 с.

11. Понкратова, Т. А. Общая теория статистики: учеб. модуль. – Кемерово, 2002. – 72 с.

12. Понкратова, Т. А. Комплексное учеб. пособие по экономическим дисциплинам для студентов специальности «Бухгалтерский учет и аудит». – Кемерово, 1999. – 68 с.

13. Понкратова, Т. А. Прикладная статистика: конспект лекций / Т. А. Понкратова, О. С. Кузнецова; КемТИПП. – Кемерово, 2008. – 120 с.

Вернуться в начало